BIOS 522: Project 1

Falcons Group

2020-09-15

Contents

1 Introduction

From 1974 to 1984, a randomized placebo-controlled Mayo Clinic trial was conducted to investigate the effect of the drug D-penicillamine on the mortality of 424 patients diagnosed with primary biliary cirrhosis (PBC) of the liver. Right censorship due to surviving or receiving a liver transplant is present in the data. More details can be found in the original text by Fleming and Harrington that includes this dataset (Counting Processes and Survival Analysis, Wiley, 1991). In this report we develop a regression model to describe the time to death as a function of treatment, accounting for age and serum bilirubin levels. We first implement a suggestion from a fictional collaborator, Dr. Blum, to use linear and logistic regression models to analyze the survival data. At the end, we conduct a log-linear Weibull survival model to address the shortcomings of the linear and logistic approaches.

2 Methods

Our goal is to develop a model to quantify the relationship between time to death and certain predictors in our data. The effect of the drug D-penicillamine on time to PBC death is of particular interest. This model can then be used for statistical inference of the model parameters, as well as prediction of time to death for future observations. We take four different approaches to analyze the survival data: a linear model that excludes all censored observations from the analysis, a linear model that treats all censored times as death times, a logistic regression, where the binary outcome is either death or censorship, and a log-linear Weibull regression that addresses the censoring nature of the survival data in the likelihood.

Each regression was run multiple times on the same sets of predictors. The sets consist of three univariate analyses on the predictors administered drug type (D-penicillamine vs. Placebo), age in years, and level of serum bilirubin in mg/dL as a categorical variable, and one multivariate analysis including all three predictors. Contrasts for categorical variables for all regressions were set so that the intercept represents the treatment group with <1.1 serum bilirubin levels. The model with the lowest Akaike information criterion (AIC), a measure of in-sample fit that adjusts for model complexity, is selected as the best performing model for describing the dataset and future predictive accuracy.

3 Results

3.1 Linear Regression

Linear models were run under two scenarios: complete exclusion of censored data and treatment of the time of censorship as time of death. Our data shows 187 censored patients and 125 dead patients at the end of the study ($\delta(0 = \text{alive}, 1 = \text{dead})$). For serum bilirubin levels, 116 patients had less than 1.1 mg/dL, 113 patients had between 1.1 to 3.3 mg/dL of serum bilirubin, and 83 patients had greater than 3.3 mg/dL.

Our linear models are formulated as

$$Y = \beta_0 + \beta^\top \mathbf{X} + \epsilon, \quad \text{where } \epsilon \sim N(0, \sigma^2)$$

The coefficients in linear regression represent how the mean survival time changes with a unit increase of a given covariate value, keeping all other covariate values constant. When the covariate is categorical, the coefficient more directly represents the mean difference in survival time between a default group (coded as 0), and the other groups coded by that covariate. Specific values and interpretations of regression coefficients are presented in Section 7: Supplementary Tables and Figures.

3.1.1 Excluding censored observations

3.1.1.1 Univariate regressions on treatment (drug), age (in years), and serum bilirubin (mg/dL) Univariate linear regression models were fit to model survival time using treatment (drug), age (in years), and categorical serum bilirubin (mg/dL) as predictors. For these models, censored observations were excluded from the data set leaving 125 observations in the data. Supplementary Tables 1-3 organize the results of these regressions, along with interpretations of each of the coefficients.

We determine that the linear regression coefficients for treatment and age are not statistically significant from 0, since their p-values are larger than the prespecified confidence level of $\alpha = 0.05$ (treatment p-value: 0.647; age p-value: 0.060) though there seems to be stronger evidence for the significance of the age predictor compared to the drug. The conclusion of a poor model fit is supported by the adjusted R^2 values for each regression being close to 0 (treatment adj. $R^2 = -0.00641$; age adj. $R^2 = 0.02069$).

The linear regression coefficient of serum bilirubin is statistically significant with a p-value of ($3.712 \times 10^{-6} < 0.05$). The R^2 value is also low (Adjusted $R^2 = 0.172$) but is higher than the other two univariate models.

3.1.1.2 Regression using all three covariates Results of the multivariate regression are in Supplementary Table 4. Our R^2 value remains low (Adjusted $R^2 = 0.2099$) even with all covariates combined. One interesting point was that age by itself was not statistically significant at the 0.05 level but taken with the other covariates it is significant so this would require more investigation. Though, the univariate p-value was pretty close to 0.05 already.

3.1.2 Treating censored times as death times

3.1.2.1 Univariate regressions on treatment (drug), age (in years), and serum bilirubin (mg/dL) For these models, censored observations were treated like death times and we had 312 observations. Supplementary Tables 5-7 organize the results of these regressions.

Even after we treat censored times as death times, we do not have enough evidence to conclude that the covariate of treatment (drug) effect is a statistically significant predictor of survival due to our p-value being larger than our alpha value ($0.883 > 0.05$) when we control for age, and serum bilirubin levels. Furthermore, our R^2 value is close to 0 (Adjusted $R^2 = 0.0147$) which suggests that the model still has very small amount of explanation towards our response variable.

There is enough evidence to conclude that both covariates of age and serum bilirubin levels are statistically significant from 0 due to their p-value being smaller than our prespecified confidence level of $\alpha = 0.05$ (age p-value: 0.0182; serum bilirubin p-value: 2.2×10^{-16}) when we control for treatment, and age. However, both adjusted R^2 values are low (age adj. $R^2 = 0.0147$; serum bilirubin adj $R^2 = 0.2176$) which suggests that both models have little explanatory power toward survival time.

3.1.2.2 Regression on all three covariates Results of the multivariate linear regression are in Supplementary Table 8. Our R^2 value is low (Adjusted $R^2 = 0.2283$) which suggests that the model has very low explanatory power towards our response variable. This is the best performing linear regression model in terms of explaining the most variance of the response, by the metric of R-squared. It would be recommended to use a different modeling approach to describe the dataset.

3.2 Logistic Regression

As in Section 4.1, we will analyze the results of the logistic regression models on the survival data. Our logistic regression models use the general format,

$$Y_i \sim \text{Bernoulli}(p_i) \qquad \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ki}.$$

In logistic regression, a coefficient value represents the change in the log odds of death when the covariate is increased by 1. Specific values of coefficients and their interpretations for each model are discussed in Section 7: Supplementary Tables and Figures.

3.2.1 Univariate regressions on treatment (drug), age (in years), and serum bilirubin (mg/dL)

Univariate logistic regression models were separately fit to model survival time using treatment (drug), age (in years), and serum bilirubin (mg/dL) as predictors. For these models, censored observations were treated like death times and we had 312 observations. Supplementary Tables 9-11 organize the results of these regressions.

In other words, we do not have enough evidence to conclude that treatment effect is statistical significant from 0 due to our p-value being larger than our prespecified confidence level of $\alpha = 0.05$ (treatment p-value: 0.695) when we control for age and serum bilirubin levels. The model produces the largest AIC = 423.97 out of the models that we fit for logistic regression.

We determine that the logistic regression coefficients for age and serum bilirubin levels do represent a statistically significant change in log odds when all other factors are held constant. The effects of age and serum bilirubin levels, when modeled with logistic regression, produce p-values smaller than the prespecified confidence level (age p-value: 1.81×10^{-5}; serum bilirubin p-value: 6.84×10^{-6}) when we control for other covariates. The univariate model of serum bilirubin had the lowest AIC when fit (age AIC: 404.3; serum bilirubin AIC: 354.73) compared to the other predictors.

3.2.2 Regression on all three covariates

Results of the regression are in Supplementary Table 12. Our R^2 value is low (Adjusted $R^2 = 0.2283$) which suggests that the model has very small amount of explanation towards our response variable. This model produces the lowest Akaike's information criterion (AIC = 335.33) of the logistic regression models and so within that class of models provides the best fit.

4 Discussion

4.1 Critique

Dr. Blum's suggestion to conduct both linear and logistic regressions are good places to start but remain inappropriate methods to model the data from the trial.

The main reason is the models do not appropriately address the right censorship present in the data. A patient is right censored if the event is not directly observed during the study and all that is known is the time to event is bounded below by the censorship time. In this dataset, patients are censored due to still being alive by the end of the ten-year study period, or due to receiving a liver transplant. The standard linear and logistic regression model objects in R do not address this issue (though it is still possible to incorporate censoring into estimation); therefore, estimation of the model parameters is based on a data likelihood that is not reflective of the actual information present in the sample. In the linear regression, we made some modifications to the response variable to get around this issue. However, excluding censored observations is not an efficient use of the entire dataset, and treating the censored time as the death times is fundamentally incorrect because deaths were never observed for those patients. An ideal model would encode the true state of information present in the dataset into the likelihood.

The fit statistics, particularly the R^2 in the linear models which intuitively can be interpreted as the percentage of total variation in the data that is explained by the model, is very low at 0.23 even for the multivariate linear model. The linear model does not do a good job of describing changes in time to event between patients that are present in our data.

There are other technical shortcomings of the linear and logistic approaches. In the linear model, each patient's time to death is assumed to follow a normal distribution, so the range of values is the entire real line. However, time to death should be nonnegative and the linear model does not restrict the response to the proper space. The logistic regression is also unsatisfactory because we are interested in a different response variable than what the logistic regression models. Our dataset provides us with rich information on time to death which we can use to estimate a survival curve, but a logistic regression models binary outcome which can be considered a summary of the time to death, and a survival function cannot be derived. Therefore, the basic model formulation of a logistic regression does not fully address the question that we are looking to answer: How does the treatment affect the time to death among PBC patients?

4.2 Parametric Survival Analysis

We conduct a parametric Weibull survival analysis from the `flexsurvreg` R package on the data to address the shortcomings from the earlier analysis using the same combination of predictors as before. Results can be found in Section 7.3. Our choice of a Weibull regression was motivated by visualizing log cumulative hazard against log of survival time (Section 7.3.5), broken up by serum bilirubin level and treatment group. The cumulative hazard was estimated by the Nelson-Aalen method. The scatterplots we observed were very linear, suggesting the data follows a Weibull distribution with a scale parameter written as a function of our covariates. This model assumes that each patient's time to death follows a nonnegative Weibull distribution, and are independent of each other. Specifically, the log of the scale parameter is assumed to be a linear function of our covariates, and the shape parameter is assumed to be constant. Right censoring is now also incorporated when conducting maximum likelihood estimation (MLE), contrary to the earlier models. More details on estimation of Weibull survival models can be found in the technical appendix, including steps to derive the mean and standard error of each of the estimated parameters.

The best performing model based on lowest AIC is the multivariate model with all three covariates. We focus on this model for interpretation of the coefficient estimates. The first two parameters in the regression output are the shape and scale. If a Weibull survival function is defined as

$$S(x) = \exp\left(-\left(\frac{x}{\lambda}\right)^{\alpha}\right), \quad x \geq 0$$

Then the shape corresponds to α and the scale corresponds to λ. This is exactly the notation used for the `dweibull` function in R, which is different from that used in class. The estimate corresponding to shape from the regression output is the estimate of the constant shape parameter α of our time to event Weibull distribution. However, the scale is assumed to vary by patient and is reparametrized so that its log is a linear function of the covariates.

$$\log(\lambda_i) = \mu + \beta^\top Z_i$$

With covariate values Z_i for patient i and model parameters μ and β. Therefore, the scale estimate in the regression output can be thought of as an intercept when all the covariate values Z are set to 0, corresponding to a scale parameter value of $\lambda_0 = \exp(\mu)$. In other words patients in the treatment group that are age 0 and have serum bilirubin levels below 1.1 mg/dL have an estimated Weibull survival curve with a scale parameter of $\lambda = \exp(\mu)$.

The estimates for the rest of the parameters are the linear effect each of the three covariates have on $\log(\lambda)$. The estimate for the drug predictor is –0.167, which means the mean survival time of the placebo group is $\exp(-0.167) = 0.846$ of the treatment group, when all other covariate values are fixed. However, the confidence interval of the estimate overlaps 0, which is not compelling evidence that the drug has a significant effect on survival time. The estimate for the age in years is –0.029, meaning an increase in age of one year is associated with a decrease in mean survival time by a multiplicative factor of $\exp(-0.029) = 0.97$, a 3% decrease, so not a huge change between ages. Our data does span a large age range from 26 to 78 years, so our model may not accurately capture differences in the effect between younger ages and older ages.

The serum 1.1-3.3 estimate tells us that the mean survival time of the serum 1.1-3.3 group is lower than the serum <1.1 group by a multiplicative factor of $\exp(-0.96) = 0.383$ or a 61.7% decrease. For the serum >3.3 estimate, the difference between the serum <1.1 group and the serum >3.3 group is a multiplicative factor of $\exp(-1.90) = 0.15$, or a 85% decrease.

The direction of each of the coefficients in this survival model is reasonable. The estimates tell us that mean survival time decreases for those in the placebo group, older patients, and patients with high serum bilirubin levels. The age and serum effects match with our preconceived notions of PBC mortality, and suggests there is weak evidence that the drug improves the prognosis within each age and serum level group.

One notable difference between the univariate analyses and the multivariate analysis is the direction of the effect of the treatment. The effects of age and serum bilirubin remain relatively unchanged between the univariate and multivariate models. However, In the univariate drug survival model, the placebo group is estimated to have an increased multiplicative effect compared to the treatment group on mean survival time of $\exp(0.04) = 1.043$, or a 4.3% increase. While this estimate is close to 1 and therefore does not suggest a strong effect, it is contrary to the multivariate model's estimate of 0.846 / 15.4% decrease in mean survival time, where the placebo group has a shorter mean time to event. This result suggests that accounting for age and serum bilirubin levels helps alleviate potential confounding of the drug and time to event relationship.

5 Conclusions

In this analysis, we experimented with several regression approaches to model the time to PBC death as a function of treatment, age, and serum bilirubin levels. The linear and logistic models were unsatisfactory because the assumed distributions are not correct descriptions of time to event data, which is a continuous and nonnegative quantity. In addition, the linear and logistic models as used did not account for right censorship in the maximum likelihood estimation.

The Weibull survival model we run in the last section addresses these critical issues, and the model including all three covariates did the best job of capturing the variation in the data. This is supported by the multivariate model having the lowest AIC out of all survival models tested. We recommend that Dr. Blum use this for further research of the data in this trial.

6 Appendix

Parameters are estimated by maximum likelihood using the algorithms available in the standard R *optim* function.

A Weibull regression model assumes

$$\log X = \mu + \gamma^T Z + \sigma W, W \sim S_W(x) = \exp\{-\exp(x)\} \tag{1}$$

$X|Z$ follows Weibull distribution, with probability density function

$$f(x) = \alpha \lambda_Z x^{\alpha-1} \exp(-\lambda_Z x^\alpha) \tag{2}$$

and survival function

$$S_X(x|Z) = \exp(-\lambda_Z x^\alpha) \tag{3}$$

where

$$\lambda_Z = \exp\left[-\frac{\mu + \gamma^T Z}{\sigma}\right], \alpha = \frac{1}{\sigma}$$

For right censored sample with n observations, the likelihood function should be changed to

$$L(x_1, ..., x_n; \alpha, \lambda_Z) = \prod_{i:c_i=1} \alpha \lambda_{Z_i} x_i^{\alpha-1} \exp(-\lambda_{Z_i} x_i^\alpha) \prod_{i:c_i=0} \exp(-\lambda_{Z_i} x_i^\alpha) \tag{4}$$

The log-likelihood function is

$$\log L = \sum_{i:c_i=1} \left\{\log[\alpha \lambda_{Z_i} x_i^{\alpha-1}] - \lambda_{Z_i} x_i^\alpha\right\} - \sum_{i:c_i=0} \lambda_{Z_i} x_i^\alpha \tag{5}$$

6.1 Estimates

Take derivatives of the log-likelihood with respect to μ, σ and γ. Set equations to zero, and solve them through iterative procedures. The BFGS optimization algorithm is the default in `flexsurvreg`. Method "BFGS" is a quasi-Newton method (also known as a variable metric algorithm), specifically that published simultaneously in 1970 by Broyden, Fletcher, Goldfarb and Shanno. This uses function values and gradients to build up a picture of the surface to be optimized. It uses the analytic derivatives of the likelihood with respect to the model parameters, if these are available, to improve the speed of convergence to the maximum.

In `flexsurvreg`, the reported *shape* is $\alpha = \frac{1}{\sigma}$, *scale* is $\lambda_Z^{-\frac{1}{\alpha}} = exp(\mu + \gamma^T Z)$ and covariate effects γ have the same "accelerated failure time" interpretation, as linear effects on $logX$.

6.2 Standard errors

Standard errors are calculated via the observed Fisher information. Let $\theta = [\mu, \sigma, \gamma]$. The Hessian matrix of (5) is

$$\mathbf{H} = \nabla^2\ l(x_1, \ldots x_n; \theta) = \nabla^2\ l(\theta) \tag{6}$$

where,

$$\mathbf{H}_{i,j} = \frac{\partial^2 l(\theta)}{\partial \theta_i \partial \theta_j}$$

The observed Fisher information is

$$\mathbf{I}\left(\hat{\theta}\right) = -\mathbf{H} \tag{7}$$

A standard asymptotic approximation to the distribution of the MLE for large N is a normal distribution with mean θ and variance $\left[\mathbf{I}\left(\hat{\theta}\right)\right]^{-1}$.
Therefore, the standard errors of $\hat{\theta}$ are

$$\hat{se}\left(\hat{\theta}\right) = \left[\mathbf{I}\left(\hat{\theta}\right)\right]^{-\frac{1}{2}} \tag{8}$$

The corresponding approximate 95% confidence intervals for θ is

$$\hat{\theta} \pm 1.96\left[\mathbf{I}\left(\hat{\theta}\right)\right]^{-\frac{1}{2}} \tag{9}$$

According to the delta-method and the relationship between *shape*, *scale*, μ, σ and γ, we can get the standard errors and confidence intervals of *shape* and *scale*.

7 Supplementary Tables and Figures

7.1 Linear Regression Models

This part of the supplementary section goes over the results of the linear regression analysis conducted in section 3.1. The survival data was examined under two scenarios:

- i) complete exclusion of censored data
- ii) treating the time of censorship as time of death.

7.1.1 Excluding censored observations

In this section, the data used for the linear models excludes any censored data. As a result we have 125 observations.

7.1.1.1 Regression on treatment (drug)

Table 1: Linear regression results using treatment as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1608	305.3	5.268	5.975e-07
drug	-89.63	195.5	-0.4585	0.6474

Table 1 organizes the results of the linear regression of time $\sim$ treatment. This model had a multiple $R^2 = 0.001707$ and adjusted $R^2 = -0.00641$. The F-statistic was $F_{1,123} = 0.2103$ with a p-value of $p = 0.6474$.

In this model, $\beta_{\text{drug}} = -89.63$. This means being in the placebo group will decrease the survival time by 89.63 days on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.647$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant.

7.1.1.2 Regression on age

Table 2: Linear regression results using age as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2445	518.3	4.716	6.403e-06
ageinyear	-18.2	9.566	-1.902	0.05945

Table 2 organizes the results of the linear regression of time $\sim$ age. This model had a multiple $R^2 = 0.02858$ and adjusted $R^2 = 0.2069$. The F-statistic was $F_{1,123} = 3.619$ with a p-value of $p = 0.05945$.

In this model, $\beta_{\text{age}} = -18.199$. This means a year increase in age will decrease the survival time by 18.199 days. However, because the p-value for this coefficient ($p = 0.0594$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant.

7.1.1.3 Regression on serum bilirubin

Table 3: Linear regression results using serum bilirubin as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2109	240.2	8.779	1.222e-14
bili_cat1.1-3.3	-259.5	279.5	-0.9284	0.355
bili_cat>3.3	-1111	272.1	-4.083	7.97e-05

Table 3 organizes the results of the linear regression of time $\sim$ serum bilirubin. This model had a multiple $R^2 = 0.1853$ and adjusted $R^2 = 0.172$. The F-statistic was $F_{2,122} = 13.88$ with a p-value of $p = 3.712 \times 10^{-6}$.

In this model, $\beta_{\text{SB}\in[1.1,3.3]} = -259.5$. This means if a person's serum bilirubin level is within 1.1 to 3.3 mg/dL it will decrease the survival time by 259.5 days on average compared to people whose serum bilirubin levels are below 1.1 mg/dL. However, because the p-value for this coefficient ($p = 0.355$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant. Furthermore, $\beta_{\text{SB}>3.3} = -1111.1$. This means if a person's serum bilirubin level is greater than 3.3 mg/dL it will decrease the survival time by 1111.1 days on average compared to people whose serum bilirubin levels are below 1.1 mg/dL.

7.1.1.4 Regression on all three covariates

Table 4: Linear regression results using treatment, age, and serum bilirubin as predictors

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3675	627.6	5.856	4.237e-08
drug	-144.3	174.4	-0.8271	0.4098
ageinyear	-23.98	8.71	-2.753	0.006818
bili_cat1.1-3.3	-336.1	274.5	-1.224	0.2232
bili_cat>3.3	-1208	268	-4.508	1.533e-05

Table 4 organizes the results of the linear regression of time $\sim$ treatment + age + serum bilirubin. This model had a multiple $R^2 = 0.2354$ and adjusted $R^2 = 0.2099$. The F-statistic was $F_{4,120} = 13.88$ with a p-value of $p = 1.533 \times 10^{-6}$.

In this model, $\beta_{\text{drug}} = -144.27$. This means being in the placebo group will decrease the survival time by 144.27 days on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.40982$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant. For age, $\beta\text{age} = -23.98$. This means a year increase in age will decrease the survival time by 23.98 days. For serum bilirubin, $\beta_{\text{SB}\in[1.1,3.3]} = -336.07$ and $\beta_{\text{SB}>3.3} = -1208.34$. This means if a person's serum bilirubin level is between 1.1 to 3.3 mg/dL then their survival time decreases by 336.07 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. If a person's serum bilirubin level is greater than 3.3 mg/dL then their survival time decreases by 1208.34 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. However, the p-value for $\beta_{\text{SB}\in[1.1,3.3]}$ is greater than the prespecified confidence level $\alpha = 0.05$, thus this term is not statistically significant. Interpretations for each coefficient are are only relevent given that the other coefficients stay constant.

7.1.2 Treating censored times as death times

In this section, the data used for the linear models treats censored data as equal to the death time. As a result we have 312 observations.

7.1.2.1 Regression on treatment (drug)

Table 5: Linear regression results using treatment as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2034	200.7	10.14	4.749e-21
drug	-18.76	127.4	-0.1472	0.883

Table 5 organizes the results of the linear regression of time $\sim$ treatment. This model had a multiple $R^2 = 6.992 \times 10^{-5}$ and adjusted $R^2 = -0.003156$. The F-statistic was $F_{1,310} = 0.02168$ with a p-value of $p = 0.883$.

In this model, $\beta_{\text{drug}} = -18.76$. This means being in the placebo group will decrease the survival time by 18.76 days on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.883$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant.

7.1.2.2 Regression on age

Table 6: Linear regression results using age as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2716	305.5	8.892	5.059e-17
ageinyear	-14.19	5.975	-2.375	0.01816

Table 6 organizes the results of the linear regression of time $\sim$ age. This model had a multiple $R^2 = 0.01787$ and adjusted $R^2 = 0.0147$. The F-statistic was $F_{1,310} = 5.641$ with a p-value of $p = 0.01816$.

In this model, $\beta_{\text{age}} = -14.191$. This means a year increase in age will decrease the survival time by 14.191 days.

7.1.2.3 Regression on serum bilirubin

Table 7: Linear regression results using serum bilirubin as the predictor

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2504	92.25	27.14	8.397e-84
bili_cat1.1-3.3	-397.7	131.3	-3.028	0.002665
bili_cat>3.3	-1330	142.8	-9.31	2.463e-18

Table 7 organizes the results of the linear regression of time $\sim$ serum bilirubin This model had a multiple $R^2 = 0.2226$ and adjusted $R^2 = 0.2176$. The F-statistic was $F_{2,309} = 44.24$ with a p-value of $p < \epsilon_{\text{Machine}} = 2.22 \times 10^{-16}$.

In this model, $\beta_{\text{SB}\in[1.1,3.3]} = -397.73$. This means if a person's serum bilirubin level is within 1.1 to 3.3

mg/dL it will decrease the survival time by 397.73 days on average compared to people whose serum bilirubin levels are below 1.1 mg/dL. Furthermore, $\beta_{\text{SB}>3.3} = -1329.90$. This means if a person's serum bilirubin level is greater than 3.3 mg/dL it will decrease the survival time by 1329.90 days on average compared to people whose serum bilirubin levels are below 1.1 mg/dL.

7.1.2.4 Regression on all three covariates

Table 8: Linear regression results using treatment, age, and serum bilirubin as predictors

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3233	348.6	9.274	3.288e-18
drug	-40.12	113.2	-0.3543	0.7234
ageinyear	-13.39	5.339	-2.509	0.01262
bili_cat1.1-3.3	-400.3	130.8	-3.06	0.00241
bili_cat>3.3	-1322	141.9	-9.319	2.365e-18

Table 8 organizes the results of the linear regression of time $\sim$ treatment + age + serum bilirubin. This model had a multiple $R^2 = 0.2382$ and adjusted $R^2 = 0.2283$. The F-statistic was $F_{4,307} = 24$ with a p-value of $p < \epsilon_{\text{Machine}} = 2.22 \times 10^{-16}$.

In this model, $\beta_{\text{drug}} = -40.120$. This means being in the placebo group will decrease the survival time by 40.120 days on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.72337$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant. For age, $\beta\text{age} = -13.394$. This means a year increase in age will decrease the survival time by 13.394 days. For serum bilirubin, $\beta_{\text{SB}\in[1.1,3.3]} = -400.348$ and $\beta_{\text{SB}>3.3} = -1322.395$. This means if a person's serum bilirubin level is between 1.1 to 3.3 mg/dL then their survival time decreases by 400.348 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. If a person's serum bilirubin level is greater than 3.3 mg/dL then their survival time decreases by 1322.395 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. Interpretations for each coefficient are are only relevent given that the other coefficients stay constant.

7.2 Logistic Regression Models

7.2.1 Regression on treatment (drug)

Table 9: Logistic regression results using treatment as the predictor

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.2675	0.3631	-0.7366	0.4614
drug	-0.09074	0.2312	-0.3925	0.6947

Table 9 organizes the results of the logistic regression of dead $\sim$ treatment. This model had an AIC of 423.97.

In this model, $\beta_{\text{drug}} = -0.09074$. This means being in the placebo group will decrease the log odds of death by 0.09074 on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.695$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant.

7.2.2 Regression on age

Table 10: Logistic regression results using age as the predictor

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.951	0.6114	-4.828	1.383e-06
ageinyear	0.05045	0.01177	4.287	1.808e-05

Table 10 organizes the results of the logistic regression of dead $\sim$ age. This model had an AIC of 404.3.

In this model, $\beta_{\text{age}} = 0.01177$. This means a year increase in age will increase the log odds of death by 0.01177 on average.

7.2.3 Regression on serum bilirubin

Table 11: Logistic regression results using serum bilirubin as the predictor

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.762	0.2625	-6.711	1.931e-11
bili_cat1.1-3.3	1.459	0.3243	4.499	6.838e-06
bili_cat>3.3	2.721	0.3593	7.573	3.642e-14

Table 11 organizes the results of the logistic regression of dead $\sim$ serum bilirubin This model had an AIC of 354.73.

In this model, $\beta_{\text{drug}} = -0.09074$. This means being in the placebo group will decrease the log odds of death by 0.09074 on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.695$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant.

In this model, $\beta_{\text{SB}\in[1.1,3.3]} = 1.4587$. This means if a person's serum bilirubin level is within 1.1 to 3.3 mg/dL it will increase the log odds of death by 1.4587 on average compared to people whose serum bilirubin levels are below 1.1 mg/dL. Furthermore, $\beta_{\text{SB}>3.3} = 2.7208$. This means if a person's serum bilirubin level is greater than 3.3 mg/dL it will increase the log odds of death by 2.7208 on average compared to people whose serum bilirubin levels are below 1.1 mg/dL.

7.2.4 Regression on all three covariates

Table 12: Logistic regression results using treatment, age, and serum bilirubin as predictors

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-5.15	0.9439	-5.456	4.865e-08
drug	0.07332	0.2761	0.2656	0.7906
ageinyear	0.06316	0.01389	4.547	5.44e-06
bili_cat1.1-3.3	1.562	0.3389	4.607	4.077e-06
bili_cat>3.3	2.914	0.3808	7.653	1.966e-14

Table 12 organizes the results of the logistic regression of dead $\sim$ treatment + age + serum bilirubin. This model had an AIC of 335.33.

In this model, $\beta_{\text{drug}} = 0.07332$. This means being in the placebo group will increase the log odds of death by 0.07332 on average compared to being given D-penacillamine. However, because the p-value for this coefficient ($p = 0.791$) is greater than the prespecified confidence level $\alpha = 0.05$, this term is not statistically significant. For age, $\beta\text{age} = 0.06316$. This means a year increase in age will increase log odds of death by 0.06316. For serum bilirubin, $\beta_{\text{SB}\in[1.1,3.3]} = 1.56164$ and $\beta_{\text{SB}>3.3} = 2.91444$. This means if a person's serum bilirubin level is between 1.1 to 3.3 mg/dL then their log odds of death increase by 1.56164 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. If a person's serum bilirubin level is greater than 3.3 mg/dL then their log odds of death increases by 2.91444 days on average compared to people whose serum bilirubin level is below 1.1 mg/dL. Interpretations for each coefficient are are only relevent given that the other coefficients stay constant.

7.3 Parametric Survival Analysis

7.3.1 Parametric survival analysis on drug

Table 13: Weibull survival analysis results using treatment as the predictor

	Estimate	Std. Error	L95%	U95%
shape	1.127	0.08867	0.9659	1.315
scale	4335	1095	2642	7111
drug	0.04176	0.1589	-0.2696	0.3531

Table 13 organizes the results of the Weibull regression of time $\sim$ treatment. The log-likelihood was -1178.718 with 3 degrees of freedom. This model had an AIC of 2383.436.

In this model, $\beta_{\text{drug}} = 0.0418$. This means the mean survival time of the placebo group is longer and 1.0427 times that of the D-penacillamine group. However because the 95% confidence interval contains 1 ($\lambda_{\text{drug}} \in [0.7025, 1.309]$), we cannot say that this coefficient is statistically significant at the predetermined confidence level of $\alpha = 0.05$.

7.3.2 Parametric survival analysis on age

Table 14: Weibull survival analysis results using age as the predictor

	Estimate	Std. Error	L95%	U95%
shape	1.144	0.08947	0.9815	1.334
scale	26540	11824	11083	63553
ageinyear	-0.03436	0.007959	-0.04996	-0.01877

Table 14 organizes the results of the Weibull regression of time $\sim$ age. The log-likelihood of -1178.773 with 3 degrees of freedom. This model had an AIC of 2363.467.

In this model, $\beta_{\text{age}} = -0.0344$. This means a year increase in age decreases the mean survival time by a multiplicative factor of 0.9662 times.

7.3.3 Parametric survival analysis on serum bilirubin

Table 15: Weibull survival analysis results using serum bilirubin as the predictor

	Estimate	Std. Error	L95%	U95%
shape	1.337	0.1007	1.153	1.549
scale	10785	2242	7176	16209
bili_cat1.1-3.3	-0.9427	0.2215	-1.377	-0.5085
bili_cat>3.3	-1.902	0.23	-2.353	-1.452

Table 15 organizes the results of the Weibull regression of time $\sim$ serum bilirubin. The log-likelihood of -1136.177 with 4 degrees of freedom. This model had an AIC of 2280.354.

In this model, $\beta_{\text{SB}\in[1.1,3.3]}$ = -0.943. This means a person's serum bilirubin level being between 1.1 and 3.3 mg/dL decreases the mean survival time by 0.3895 times compared to a person whose serum bilirubin level is less than 1.1 mg/dL with all other covariates constant. Furthermore, $\beta_{\text{SB}>3.3}$ = -1.9. This means a person's serum bilirubin level being greater than 3.3 mg/dL decreases the mean survival time by 0.1496 times compared to a person whose serum bilirubin level is less than 1.1 mg/dL.

7.3.4 Parametric survival analysis on all three covariates

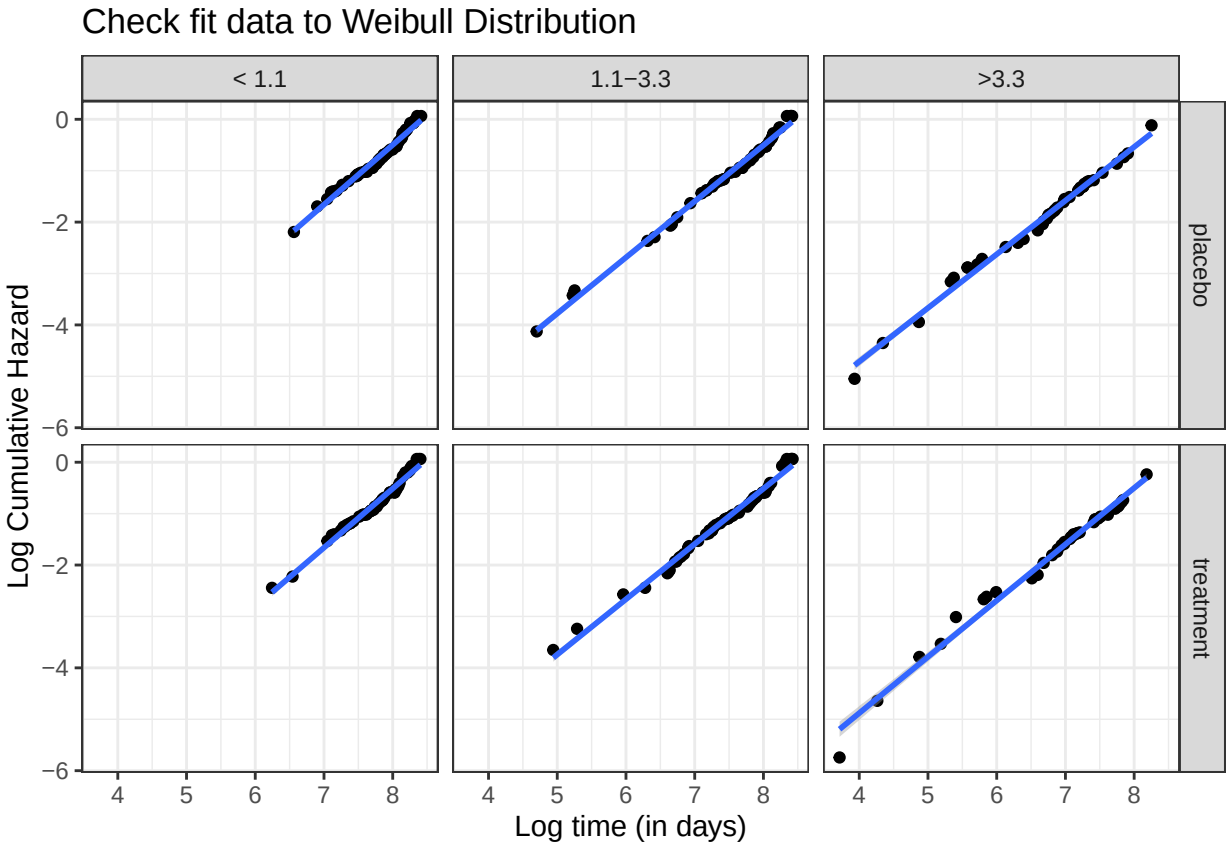
Table 16: Weibull survival analysis results using treatment, age, and serum bilirubin as predictors

	Estimate	Std. Error	L95%	U95%
shape	1.365	0.1016	1.18	1.58
scale	61014	30192	23133	160928
drug	-0.1669	0.1349	-0.4313	0.09752
ageinyear	-0.02925	0.006207	-0.04141	-0.01708
bili_cat1.1-3.3	-0.9603	0.2174	-1.386	-0.5342
bili_cat>3.3	-1.9	0.2251	-2.341	-1.459

Table 16 organizes the results of the Weibull regression of time $\sim$ treatment + age + serum bilirubin. The log-likelihood of -1124.358 with 6 degrees of freedom. This model had an AIC of 2260.716.

In this model, β_{drug} = -0.167. This means being in the placebo group decreases the mean survival time by 0.8462 times compared to being given D-penacillamine. However because the 95% confidence interval contains 1 ($\lambda_{\text{drug}} \in [0.907, 1.539]$), we cannot say that this coefficient is statistically significant at the predetermined confidence level of $\alpha = 0.05$. For age, βage = -0.0292. This means a year increase in age decreases mean survival time by 0.9712 times. For serum bilirubin, $\beta_{\text{SB}\in[1.1,3.3]}$ = -0.96 and $\beta_{\text{SB}>3.3}$ = -1.9. This means a person's serum bilirubin level being between 1.1 and 3.3 mg/dL decreases mean survival time by 0.3829 times compared to a person whose serum bilirubin level is less than 1.1 mg/dL. A person's serum bilirubin level being greater than 3.3 mg/dL decreases mean survival time by 0.1496 times compared to a person whose serum bilirubin level is less than 1.1 mg/dL. Interpretations for each coefficient are only relevent given that the other coefficients stay constant.

7.3.5 Checking Weibull distribution is a good fit with log-log plot



The plot of log time against the log cumulative hazard (estimated by the Nelson-Aalen estimator) is mostly linear in each of the treatment / serum bilirubin groups and deviates mostly at the earlier time points. Due to the linearity within each of the groups, we conclude the Weibull distribution is a good fit for this dataset. It is harder to visualize along with continuous age (our third covariate), but based on this diagnostic it seems Weibull is a good fit regardless.