Homework 4 Score: 82.5/105

1. Descriptive Statistics

Methods:

Descriptive statistics are presented in groups defined by the primary outcome of survival. Continuous variables (age, bilirubin) are presented according to the mean, standard deviation, minimum and maximum. Categorical and binary variables (gender, treatment with d-penicillamine) are presented according to absolute number of participants and percentages. Because bilirubin levels are right skewed with many outliers, bilirubin levels are log-transformed with base 2. Kapalan –Meier estimate of the censoring distribution curve is included. Data that are missing are subsequently excluded from analysis. For this study, we assumed that censoring was non-informative.

Results:

Of the 418 participants in the study, 161 (38.5%) had died by the end of the study and 257 had survived or had been censored. The minimum survival time was 41 days, with 1 year survival of approximately 93%, 2 year survival of 88% and 5 year survival of 70%. Median survival was approximately 9.3 years. Please see the Kaplan Meier Survival Curve included for overall censorship distribution of study population.

As primary biliary cirrhosis (PBC) is a disease that effects women more than men, the majority of study participants were female (88.5%), middle aged and with moderate- severe liver disease. Compared to the participants who did not die during the study, the participants who survived were more likely to be male (17.4% vs 7.5%), have higher mean bilirubin (5.54mg/dL vs 1.77 mg/dL), more severe liver disease (mean stage 3.4 vs 2.8) and shorter follow up time. Interestingly, only 310 participants had treatment status recorded, with about half of participants receiving active drug, d-penicillamine. 106 participants are also missing information on gender and stage of disease.

Although other variables, such as measures of liver synthetic function, transaminitis and markers of portal hypertension are likely associated with survival, our analysis is limited to the variables examined below. Based on a priori assumptions of development of PBC and survival, age and sex should be evaluated as potential confounders.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Censored (n= 257) | Died (n= 161) | All (n= 418) |
| Age (years)1 | 48.7 (10.36, 26-78) | 53.9 (9.8, 30.86-76.71) | 50.7 (10.4, 26.3-78.4) |
| Female (%) | 173 (92.5) | 103 (82.4) | 276 (88.5) |
| Bilirubin (mg/dL)1 | 1.77 (2.2, 0.3-1) | 5.54 (5.8,0.3-28) | 3.22 (4.4, 0.3-28) |
| Stage | 2.8 (0.89, 1-4) | 3.4 (0.74, 1-4) | 3.0 (0.88, 1-4) |
| 1 (%) | 15 (8) | 1 (0.8) | 16 (5.3) |
| 2 (%) | 51 (27.3) | 16 (12.8) | 67 (21.5) |
| 3 (%) | 77(41.2) | 43 (34.4) | 120 (38.5) |
| 4 (%) | 44 (23.5) | 65 (52) | 109 (34.9) |
| Treated with d-penicillamine (%) | 92 (49.7) | 65 (52) | 157 (51) |
| Follow up time (days)1 | 2256.6 (1000.2, 533-4795) | 1376.9 (1049.2, 41-4191) | 1917.8 (1104.7, 41-4795) |
| 1 Descriptive statistics presented as mean (standard deviation, minimum- maximum) | | | |







Score: 5/10

2. In homework 2, we were able to use logistic regression to analyze our data by dichotomizing survival by 4year vital statistics. This was possible, because participants were followed for a long time with no censored individuals dropping out of the study before 4 years. In this data set, there is huge variability in follow up time – of the 257 participants who survived/ were censored, one participant was only followed by 533 days and only about ½ were followed for more than 5 years. To use logistic regression, we would have to limit our data to the time our first participant was censored (i.e. 533 days) when only about 35 people had died. By arbitrarily picking this time point to dichotomize our data, we would lose a lot of data and statistical power. This would result in a loss of precision in determining an association between bilirubin and death. In addition, we may introduce confounding as the patients who died earlier may have been sicker.

Score : 10/10; Perfect answer mentioning both confounding and precision consideration;

3.

a. Methods: Proportional hazard ratio was used to model the distribution of time to death from any cause across continuous levels of serum bilirubin. The quantification of the association between bilirubin level and all cause mortality was summarized by a hazard ratio computed from the regression model. The Huber-White sandwich estimator was used to determine standard error and Wald statistics used to generate subsequent confidence intervals and two-sided p-value.

Results: Data was available on 418 participants with mean serum bilirubin of 1.77mg/dL. During an average follow up of 5.3 years, 161 participants were observed to die. From our proportional hazard regression, we estimate that for each 1mg/dL increase in serum bilirubin levels, there is a 15.2% rise in instantaneous risk of death. From 95% CI, this estimate would not be surprising if the true instantaneous risk of death was between 12.1% to 18.5% higher for each 1mg/dL increase in serum bilirubin. A p-value of < 0.001, means that we can reject with great confidence the null hypothesis that risk of death is not associated with serum bilirubin levels.

b. Please see question 6 for modeled answer of bilirubin -1

Score: 9.5/10

4. a. A priori, we could predict that log transformation of serum bilirubin might be a good analysis. For technical reasons, log transformation is useful in minimizing the outliers of our data: serum bilirubin levels are right skewed (see the descriptive statistics above), with most bilirubin levels < 5 mg/dL. In addition, log-transforming bilirubin may make sense from a biologic standpoint because it is a biochemical product that has effects on a multiplicative scale – i.e. a 2-3mg/dL change in serum bilirubin has much different effects/ implications than a 32 – 33mg/dL change. Finally, because we are analyzing our data based on instantaneous risk of death, it makes more sense to think of our predictor on a multiplicative scale – i.e. a 2 fold increase in serum bilirubin levels causes a “n percent” increased risk of death.

Score: 5/5

b. Methods: Serum bilirubin levels were transformed using log base 2. Proportional hazard ratio was used to model the distribution of time to death from any cause across continuous levels of log transformed serum bilirubin levels. Inference on the slope of the regression model was used to generate a hazard ratio to describe the association between log transformed bilirubin level and all cause mortality. The Wald statistics was used to compute 95% confidence intervals and two-sided p-value based on the Huber Sandwich estimator. No subjects were excluded from analysis

Results: 418 participants were available for analysis with mean serum bilirubin of 1.77mg/dL. During an average follow up of 5.3 years, 161 participants were observed to die. From our proportional hazard regression, we estimate that for each doubling (or 2 fold) increase in serum bilirubin levels, there is a 98.4% increase in instantaneous risk of death. From 95% CI, this estimate would not be surprising if the true instantaneous risk of death was between 1.78 to 2.21 times higher for each doubling in serum bilirubin. With a p-value of < 0.001, we can reject with great confidence the null hypothesis that risk of death is not associated with serum bilirubin levels.

Score: 10/10

c. Please see question 6 for modeled answer of log bilirubin

5. a. Methods: The distribution of time to death from any cause was compared across groups defined by log transformed bilirubin and untransformed, continuous linear bilirubin using a proportional hazard ratio. Inference on the slope of the regression model was used to generate a hazard ratio to describe the association between mean bilirubin, log transformed bilirubin and all cause mortality. The Wald statistics was used to compute 95% confidence intervals and two-sided p-value based on the Huber Sandwich estimator. No subjects were excluded from analysis

Results: 418 participants were available for analysis with mean serum bilirubin of 1.77mg/dL. During an average follow up of 5.3 years, 161 participants were observed to die. From our proportional hazard regression modeling linear and log transformed bilirubin, we see a statistically significant association only between log transformed bilirubin and instantaneous risk of death (P< 0.001). This suggests that the true association between instantaneous risk of death and serum bilirubin, is best described by log transforming bilirubin.

b. Please see question 6 for fitted model of log transformed + linear bilirubin

Score: 7/10

6.



Plot of fitted hazard ratio (compared to group having a bilirubin of 1mg/dL) from the proportional hazard regression models

The graph above of all three of our models shows a positive association between bilirubin levels and risk of death: i.e. as bilirubin increases, so does the instantaneous risk of death. The models that include a linear fit for bilirubin (problems 3 and 5) produce a curvilinear association between bilirubin levels and risk of death. The linear fit of bilirubin in 3 produces a curvilinear association with risk of death that asymptotically approaches infinity on the y-axis. The model produced in question 5 produces a curvilinear fit ,which plateaus at higher levels of bilirubin. The log-transformed bilirubin creates a linear association between our predictor of interest and our fitted hazard ratio. Thus, as stated in problem 5, a logarithmic transformation for our data holds the greatest appeal.

Score: 10/10

7. a. To be considered a confounder, the covariates must be related to both our predictor of interest (bilirubin) and our outcome of survival. Looking at our descriptive statistics, it is clear that both age and gender are associated with survival. Although overall, women are more likely to develop PBC, men with the disease seem more likely to die. Older participants were also more likely to die. It is less intuitive to engineer a reason why age or gender would be associated with our predictor of interest, serum bilirubin levels. There fore, age or gender would be better considered as precision variables or effect modifiers than as confounding variables.

Score: 5/10; it seems that you didn't check the association between age/sex with bilirubin in the data.

b. To be considered a precision variable, age and gender should be associated with our outcome of interest (i.e. survival) but not our predictor. By categorizing survival according to these variables, we can reduce the variability within our model and improve the fit.

Score: 3/10; I agree with your answer. But the key is using comparisons of adjusted and unadjusted regression to see whether we gain precision by adjusting sex or age.

c. Methods: A proportional hazard regression was used to model the relationship between death and log transformed bilirubin, using age and gender as covariates. Inference on the slope of the regression model was used to generate a hazard ratio to describe the association between log-transformed bilirubin and all cause mortality. The Wald statistics was used to compute 95% confidence intervals and two-sided p-value based on the Huber Sandwich estimator. Of note, the 106 participants without gender or age recorded were excluded from all analysis.

Results: 312 participants were available for analysis. The instantaneous risk of death increased by 2.11 for participants that had a doubling of serum bilirubin levels. By 95% confidence intervals, this increase would not be unusual if the true increased risk of death was between 1.84 to 2.44 for patients with a doubling of serum bilirubin levels. A p-value of < 0.001 indicates that we can reject with great confidence the null hypothesis that the risk of death is not associated with the bilirubin level.

Score: 8/10;

8. We ignored treatment with pencillamine in our above analyses. Based on our descriptive statistics, it does not seem that treatment significantly impacted survival and about 50% of survivors and 50% of censored participants received treatment. Treatment with the drug started after enrollment in the study, so the drug could not be considered a confounder. Hypothetically, if the treatment was effective, it may have altered survival and thus acted as an effect modifier.

Score: 10/10