Homework #4

10 February 2015

**Question 1.**

Methods: Descriptive statistics are presented for the variables of interest in the following analyses. Because bilirubin is a blood marker, it is likely to act on a multiplicative scale. Thus, bilirubin levels were divided into four categories for ease of presentation based on an approximately doubling: 0-1 (not including 1), 1-2 (not including 2), 2-4 (not including 4), and >=4. Kaplan-Meier survival estimates for all-cause mortality are shown by bilirubin category, sex, age categories, and treatment arm. Table 1 shows mean, standard deviation, minimum, maximum, and sample size of age in each bilirubin category. Table 1 also shows the proportion of females and proportion in the intervention arm by bilirubin category, as well as the number of non-missing observations in each category.

Results: There are a total of 418 subjects in this dataset, all of whom have a bilirubin measurement. The mean bilirubin level is 3.22 mg/dL (standard deviation of 4.41 and range of 0.300 to 28.0 mg/dL). Sex is missing for 106 subjects: 38 with bilirubin <1, 30 with bilirubin >=1 and <2, 15 with bilirubin >=2 and <4, and 23 with bilirubin >=4. Treatment arm (drug versus placebo) is also missing for a large number of subjects (108 total missing): 39 with bilirubin <1, 31 with bilirubin >=1 and <2, 15 with bilirubin >=2 and <4, and 23 with bilirubin >=4. The survival curves by bilirubin group illustrate a distinct difference in survival, with the highest bilirubin group (>=4) having the lowest survival. As shown in table 1, the mean age is very similar across all four bilirubin groups (about 51 years) and thus is unlikely a confounder. The proportion of females is also similar across the bilirubin groups. The proportion of subjects in each treatment arm (drug or placebo) is also similar across groups. Kaplan-Meier survival curves are shown by sex and by treatment arm for those subjects with non-missing data for sex (n=312) and treatment arm (n=310). Males appear to have worse survival as illustrated in the Kaplan-Meier survival curves by sex below. Survival was relatively similar between each treatment arm, shown in the Kaplan-Meier curve below.

**Table 1**. Descriptive statistics by serum bilirubin level.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Serum Bilirubin Level (mg/dL)** | | | | |
|  | **<1**  **(n=142)** | **>=1 & <2 (n=107)** | **>=2 & <4 (n=78)** | **>=4**  **(n=91)** | **Overall (n=418)** |
| **Age (years)a** | 51.2 (9.80; 30.6-75.0; n=142) | 50.6 (10.8; 26.3-76.7; n=107) | 50.2 (11.4; 29.6-71.9; n=78) | 50.7 (10.4; 30.9-78.4; n=91) | 50.7 (10.4; 26.3-78.4; n=418) |
| **Female**b **(%)** | 93.3% (n=104) | 88.3% (n=77) | 82.5% (n=63) | 86.8% (n=68) | 88.5% (n=312) |
| **Treatment armc** | 48.5% (n=103) | 55.3% (n=76) | 54.0% (n=63) | 45.6% (n=68) | 50.6% (n=310) |

aDescriptive statistics shown for continuous variable age are mean (standard deviation; minimum-maximum; sample size). bDescriptive statistics for binary variable female are proportion of the group that are female and total sample size (male and female). cDescriptive statistics for binary variable treatment arm are proportion in treatment arm (receiving d-penicillamine).



**Question 2.**

We are not able to use logistic regression to investigate associations between mortality and various covariates because we have censored data. Not all subjects were observed to die. Due to this censoring, time to death or censoring could confound the potential association between bilirubin level and mortality if it is associated with both of them. Time could be a confounder if it causes both increased bilirubin and increased likelihood of death. If we ignore the group with censored data and those individuals went on to live many years, we would underestimate the probability of survival.

**Question 3a.**

Methods: Proportional hazards regression with robust standard errors was used to evaluate the instantaneous risk (hazard) of death by any cause over the entire observation period across groups defined by serum bilirubin level. The two-sided p-value and 95% confidence interval are Wald-based estimates. All subjects were included in the analysis.

Results: 418 participants were included in this analysis and a total of 161 deaths were observed. For each difference of 1 in bilirubin level, we estimate the risk of death is 1.152 times higher (15.2% higher) in the group with the higher bilirubin. Based on the 95% confidence interval, this hazard ratio would not be considered unusual if the true risk of death were between 1.121 and 1.185 times higher among those with a 1 higher bilirubin. A highly statistically significant p-value (P<0.0005) suggests that we can reject the null hypothesis of no increased risk of death across groups increasing in bilirubin for the alternative hypothesis of a higher instantaneous risk of death in those with higher bilirubin levels.

**Question 3b.**

See graph in problem 6.

**Question 4a.** We might have preferred a log-transformed bilirubin as a predictor a priori on a scientific basis, because as a biologic marker we might expect it to be relevant on a multiplicative scale.

**Question 4b.**

Methods: Proportional hazards regression with robust standard errors was used to evaluate the instantaneous risk (hazard) of death by any cause over the entire observation period across groups defined by serum bilirubin level, using base 2 log-transformed serum bilirubin as the predictor. The two-sided p-value and 95% confidence interval are Wald-based estimates. Subjects missing bilirubin level were excluded from analysis.

Results: 418 participants were included in this analysis and a total of 161 deaths were observed. For each 2-fold increase in bilirubin level, we estimate the risk of death is 1.984 times higher (98.4% higher) in the group with the higher bilirubin. Based on the 95% confidence interval, this hazard ratio would not be considered unusual if the true risk of death were between 1.781 and 2.212 times higher in a group with a 2-fold higher bilirubin than another group. A highly statistically significant p-value (P<0.0005) suggests that we can reject the null hypothesis of no increased risk of death across groups increasing in bilirubin for the alternative hypothesis of a higher instantaneous risk of death in those with higher bilirubin levels.

**Question 4c.**

See graph in problem 6.

**Question 5a.**

Methods: Proportional hazards regression was used to compare the distributions of time to death across groups defined by bilirubin level. Bilirubin was included in this model as both an untransformed variable and as a base 2 logarithmically-transformed variable. All subjects had information on baseline bilirubin levels and were included. Standard error estimates were based on the Huber-White sandwich estimator and the two-sided p-value and 95% confidence intervals were based on the Wald statistic for both tests. The null hypothesis that β1=β2=0 was tested. The test for nonlinearity is of primary interest with the null hypothesis that the coefficient of the logarithmically-transformed bilirubin is equal to zero.

Results: 418 subjects were included in this analysis. There were 161 observed deaths during an average of 5.25 years of observation. The mean bilirubin level at baseline in the entire group was 3.22 mg/dL (standard deviation of 4.41 and range of 0.300 to 28.0 mg/dL). We find a statistically significant (p<0.0005) association between bilirubin and instantaneous risk of death and thus reject the null hypothesis that both β1 and β2 are equal to zero in favor of the alternative hypothesis that either β1 or β2 is not equal to zero. The test for nonlinearity is statistically significant (p<0.0005) suggesting that the linear model of bilirubin in modeling the log hazard function with mortality as the outcome is not sufficient.

**Question 5b.**

See graph in problem 6.

**Question 6.** The three different models produce fitted values that are similar in that they indicate an upward trend – increasing relative hazard as bilirubin increases. All three models are shown below with the relative hazard compared to 1 mg/dL bilirubin. There is very little difference between the model with log-transformed bilirubin as the predictor (labeled as log fit below) and the model with both bilirubin and the log-transformed bilirubin in the same model (labeled as log-linear fit below) when bilirubin levels are less than about 15 mg/dL. When bilirubin is less than about 20 mg/dL the model with untransformed bilirubin (labeled as linear fit below) underestimates the relative hazard compared to the other two models, but with increasing bilirubin levels greater than 20 mg/dL (where the data is much more sparse) this model overestimates the relative hazard compared to the other two models. While the log linear and log fit models are both similar over lower bilirubin levels, we would select the log-fit model for ease of interpretability and communication.



**Question 7a.** A confounder would be associated with the predictor of interest in the sample and causally associated with outcome of interest but not in the causal pathway of interest. The association of interest in each stratum in the confounder will be similar to each other but different from the association in the combined sample. Both sex and age do not appear to be associated with the predictor of interest, bilirubin (see Table 1 in Question 1). There are a similar proportion of females in each group defined by bilirubin level and the mean age is very similar across all four bilirubin groups. Because neither of these variables (age or sex) are associated with the predictor of interest in the sample, we do not think they are confounders.

**Question 7b.** Adjustment for a precision variable will impact the standard errors leading to a narrower confidence interval. As established above in part a, neither age nor sex are associated with bilirubin level (predictor of interest in this sample). However, they are causally associated with mortality (women tend to live longer than men as shown in the Kaplan-Meier survival curves in Question 1 and older age is associated with mortality). Thus both age and sex are precision variables. Adjusting for these precision variables moves the estimate of the hazard ratio slightly further from the null.

**Question 7c.**

Methods: Proportional hazards regression with robust standard errors was used to estimate the association between bilirubin (predictor of interest) and all-cause mortality (response), adjusted for sex and age. Bilirubin was log-transformed using base 2. The two-sided p-value and 95% confidence interval are Wald-based estimates with standard errors computed using the Huber-White sandwich estimator. All subjects had data on age. Subjects missing data on sex were excluded from this analysis.

Results: 312 subjects were included in this analysis. Comparing two groups with a 2-fold difference in bilirubin level but of the same sex and age, the group with the higher bilirubin is estimated to have a 2.108 times higher instantaneous risk of death. Based on the 95% confidence interval, this estimate would not be judged unusual if the true instantaneous risk of death were between 1.840 and 2.416 times higher in the group with higher bilirubin. A highly statistically significant p-value (p<0.0005) suggests that we can reject the null hypothesis of no difference in risk of death in favor of the null hypothesis of a higher instantaneous risk of death among those with a higher bilirubin of the same age and sex as those with a lower bilirubin.

**Question 8.** The treatment arm (drug versus placebo) is missing for 108 subjects. The treatment arm is not associated with the baseline bilirubin level (similar proportions were designated to the intervention arm within each bilirubin strata, see Table 1 in Question 1) and thus would not be considered a confounder. We might expect that treatment arm would be an effect modifier and that the association between baseline bilirubin and mortality could be different among those taking the drug versus those taking the placebo. However, we see a similar hazard ratio within each group (2.062 among those on the placebo and 2.099 among those on the active drug). We might also consider treatment arm as a precision variable as it is not associated with the exposure (bilirubin). However, the Kaplan-Meier survival curve in Question 1 does not provide evidence that the treatment is associated with mortality and thus treatment would not be considered a precision variable.