**BIOST 518/515: Applied Biostatistics II/Biostatistics II**

Emerson, Winter 2015

**Homework 04**

February 4, 2015

This homework investigates associations between death from any cause and age, sex, and serum bilirubin in a population of patients with primary biliary cirrhosis who were enrolled in a randomized clinical trial (RCT) of D-penicillamine.

1. Provide suitable descriptive statistics pertinent to the scientific questions addressed in this homework.

**Method:**

As stated, the homework investigates associations between death from any cause and age, sex, and serum bilirubin level. Descriptive statistics were provided in different sets. Vital status were present for groups defined by serum bilirubin level, groups defined by age and groups defined by sex, separately. Descriptive statistics include total number of subjects and number of deaths in each groups, as well as other summaries based on Kaplan-Meier estimates. Additional descriptive statistics for the relationship between serum bilirubin level, age and sex were also provided, including total number of observations with non-missing data, percentage (sex), mean, standard deviations, minimum and maximum (age).

**Answer:**

The total number of observations in the dataset is 418. None of the observations was missing data on serum bilirubin level, sex, observed time and censoring status. 106 observations were missing data on sex, among which 36 died and 70 were censored by the end of the trial.

Table 1a presents estimates of the survival estimates within groups defined by serum bilirubin level and the entire sample. Those in the groups with higher serum bilirubin level tended to have lower survival probabilities, especially those with serum level of 4 mg/dl and above had the lowest survival probabilities. On an average of the first 6 years following study enrollment, those with serum level of less than 1 mg/dl had a mean survival of 5.77 years while those with serum level of 4 mg/dl and above only had a mean survival of 3.15 years. Figure 1 presents the Kaplan-Meier estimates graphically for the same 4 groups defined by serum bilirubin level. Group with serum level lower than 1 mg/dl had a remarkably higher survival curve than other groups (no overlaps with any other groups).

Table 1b presents the relationship between age, sex and serum bilirubin level. The group with the lowest serum bilirubin level had the highest mean age, although the difference in the means of age were only very slightly. The other three groups defined by bilirubin level had very similar average age. Overall 88.5% of the sample excluding missing data on sex were female. The percentage of female were different from each other, with the group with lowest bilirubin level having the highest percentage of female.

Table 1c presents estimates of the survival estimates within groups defined by age and the entire sample. Those in the youngest group tended to have much higher survival probabilities. However, the difference in mean survival within the first 6 years following study enrollment were not drastically different as in the case of table 1a. Figure 2 presents the Kaplan-Meier estimates graphically for the groups defined by age. The youngest group (less than 40 years old) had the highest survival curve, but over time this curve started overlapping with other curves.

Table 1d presents estimates of the survival estimates within groups defined by sex and the entire sample. Female subjects tended to have much higher survival probabilities. That being said, the vast majority of the subjects were female. Figure 3 presents the Kaplan-Meier estimates graphically for groups defined by sex. The female had the higher survival curve and over time did not overlap with the male survival curve.

**Table 1a. Descriptive Statistics of Vital Status by Serum Bilirubin Level**

|  |  |
| --- | --- |
|  | **Serum Bilirubin Level (mg/dl)** |
| **Level < 1** | **1 ≤ Level < 2** | **2 ≤ Level < 4** | **Level ≥ 4** | **All levels** |
| **N subjects[1]** | 142 | 107 | 78 | 91 | 418 |
| **N deaths** | 20 | 30 | 44 | 67 | 161 |
| **Survival at 1 year** | 99.3% | 96.3% | 91.0% | 79.12% | 92.8% |
| **Survival at 2 years** | 96.5% | 95.3% | 84.6% | 69.18% | 88.8% |
| **Survival at 5 years** | 93.1% | 87.1% | 52.5% | 24.77% | 69.9% |
| **Survival at 10 years** | 76.3% | 53.9% | 9.5% | 6.55% | 44.2% |
| **25th pct of survival[2]** | 10.55 yrs | 7.11 yrs | 2.74 yrs | 1.34 yrs | 4.00 yrs |
| **Median survival[2]** | NA[4] | 11.47 yrs | 5.70 yrs | 2.95 yrs | 9.30 yrs |
| **60th pct of survival[2]** | NA[4] | NA[4] | 8.82 yrs | 3.55 yrs | 10.55 yrs |
| **6-year mean survival[3]** | 5.77 yrs | 5.48 yrs | 4.45 yrs | 3.15 yrs | 4.90 yrs |

*[1] None of the observations was missing data on age*

*[2] Estimates were approximated and based on the Kaplan-Meier estimates computed within strata by serum bilirubin level and in the entire sample. Estimates were obtained in years then divided by 365.25*

*[3] Average number of years alive during the first 6 years following study enrollment, as computed by the area under Kaplan-Meier survival curves within strata and sample restricted to 6 years*

*[4] Corresponding percentile cannot be estimated from the available data*

**Figure 1. Kaplan-Meier Estimates by Serum Bilirubin Level**



**Table 1b. Descriptive Statistics of Age and Sex by Serum Bilirubin Level**

|  |  |
| --- | --- |
|  | **Serum Bilirubin Level (mg/dl)** |
| **Level < 1****(n = 142)** | **1 ≤ Level < 2****(n = 107)** | **2 ≤ Level < 4****(n = 78)** | **Level ≥ 4****(n = 91)** | **All levels****(n = 418)** |
| **Age** |  |  |  |  |  |
| **N** | 142 | 106 | 78 | 91 | 418 |
| **Mean (sd)** | 51.2 (9.8) | 50.6 (10.8) | 50.2 (11.4) | 50.7 (10.4) | 50.7 (10.4) |
| **Min - Max** | 30.6 – 75.0 | 26.3 – 76.7 | 29.6 – 71.9 | 30.9 – 78.4 | 26.3 – 78.4 |
| **Sex** |  |  |  |  |  |
| **N** | 104 | 77 | 63 | 68 | 312 |
| **Female - %** | 93.27% | 88.31% | 82.54% | 86.76% | 88.46% |

**Table 1c. Descriptive Statistics of Vital Status by Age**

|  |  |
| --- | --- |
|  | **Age (years)** |
| **Age < 40** | **40 ≤ Age < 50** | **50 ≤ Age < 60** | **Age ≥ 60** | **All age groups** |
| **N subjects[1]** | 69 | 126 | 135 | 88 | 418 |
| **N deaths** | 13 | 42 | 59 | 47 | 161 |
| **Survival at 1 year** | 98.6% | 93.7% | 91.1% | 89.8% | 92.8% |
| **Survival at 2 years** | 98.6% | 88.1% | 86.7% | 81.8% | 88.8% |
| **Survival at 5 years** | 83.8% | 73.8% | 67.2% | 56.4% | 69.9% |
| **Survival at 10 years** | 57.7% | 60.3% | 38.0% | 14.8% | 44.2% |
| **25th pct of survival[2]** | 7.36 yrs | 4.63 yrs | 3.33 yrs | 2.48 yrs | 4.00 yrs |
| **Median survival[2]** | NA[4] | 11.47 yrs | 8.68 yrs | 6.85 yrs | 9.30 yrs |
| **60th pct of survival[2]** | NA[4] | NA[4] | 9.43 yrs | 8.82 yrs | 10.55 yrs |
| **6-year mean survival[3]** | 5.55 yrs | 5.03 yrs | 4.77 yrs | 4.39 yrs | 4.90 yrs |

*[1] None of the observations was missing data on age*

*[2] Estimates were approximated and based on the Kaplan-Meier estimates computed within strata by age group and in the entire sample. Estimates were obtained in years then divided by 365.25*

*[3] Average number of years alive during the first 6 years following study enrollment, as computed by the area under Kaplan-Meier survival curves within strata and sample restricted to 6 years*

*[4] Corresponding percentile cannot be estimated from the available data*

 **Figure 2. Kaplan-Meier Estimates by Age**



**Table 1d. Descriptive Statistics of Vital Status by Sex**

|  |  |
| --- | --- |
|  | **Sex** |
| **Male** | **Female** | **All** |
| **N subjects[1]** | 36 | 276 | 312 |
| **N deaths** | 22 | 103 | 125 |
| **Survival at 1 year** | 94.4% | 92.8% | 92.9% |
| **Survival at 2 years** | 88.7% | 89.5% | 89.4% |
| **Survival at 5 years** | 52.2% | 73.0% | 70.6% |
| **Survival at 10 years** | 34.3% | 44.9% | 43.9% |
| **25th pct of survival[2]** | 2.77 yrs | 4.63 yrs | 4.07 yrs |
| **Median survival[2]** | 6.53 yrs | 9.39 yrs | 9.30 yrs |
| **60th pct of survival[2]** | 7.66 yrs | 10.55 yrs | 10.55 yrs |
| **6-year mean survival[3]** | 4.41 yrs | 5.00 yrs | 4.93 yrs |

*[1] There were 106 observations with missing data on sex, among which 36 died and 70 observations were censored on observed time to death.*

*[2] Estimates were approximated and based on the Kaplan-Meier estimates computed within strata by sex and in the entire sample. Estimates were obtained in years then divided by 365.25*

*[3] Average number of years alive during the first 6 years following study enrollment, as computed by the area under Kaplan-Meier survival curves within strata and sample restricted to 6 years*

**Figure 3. Kaplan-Meier Estimates by Sex**



2. In prior homeworks using the Cardiovascular Health Study datasets, we were able to use logistic regression to investigate associations between mortality and various covariates. Why might such an approach not seem advisable with these data? (Consider the extent to which such analyses might be confounded and/or lack precision.)

**Answer**:

In prior homework using the Cardiovascular Health Study datasets, we were able to use logistic regression to investigate associations between mortality and various covariates. It was because we were able to dichotomize the time to death at 4 years and get decent sample size.

Such approach might not seem advisable with these data. The first reason is lack of precision. It is always of interest not to dichotomize over continuous variable, especially those with censored data. In addition, the first censored time appears at approximately 1.5 years. Dichotomizing at 1.5 years will give such a small sample size of those who died within 1.5 years. We might not have enough information about other covariates within such small sample size.

The second reason is the potential of confounding. If the probability of dying was systematically associated with the length of observed time in the sample, we would unknowingly create a confounder for our analysis by dichotomizing time to death. This could happen if for example, by chance prevalent cases (both old and newly diagnosed cases) were enrolled in the study at the very beginning of the trial, while as the trial was going on, later in the study more incident cases were recruited instead. In such case, by defining a time point to dichotomize observed time, a potential confounder was created out of the process.

3. Perform a statistical regression analysis evaluating an association between serum bilirubin and all-cause mortality by comparing the instantaneous risk (hazard) of death over the entire period of observation across groups defined by serum bilirubin modeled as a continuous variable.

a. Include a full report of your inference about the association.

**Method:**

The instantaneous risks of death over the entire period of observation across groups defined by serum bilirubin were compared using proportional hazards regression modeling serum bilirubin as a continuous variable. The hazard ratios were computed from the outputs of the regression. The 95% confidence interval and two-sided p-values were constructed using Wald statistics based on the Huber-White sandwich estimator. There were no subjects with missing data on serum bilirubin

**Result:**

There were 418 observations with no missing data on serum bilirubin level at baseline. From a proportional hazards regression analysis, the instantaneous risk of death is a relative 15.24% higher (hazard ratio of 1.1524) for each 1 mg/dl higher in serum bilirubin level at baseline. With 95% confidence, this is consistent with the true instantaneous risk of death being anywhere between 12.1% and 18.5% higher for each 1 mg/dl higher in serum bilirubin level at baseline. This result is statistically significant at the 5% critical threshold with a two-sided p-value of less than 0.0001. As a result, we reject the null hypothesis, in favor of an alternative hypothesis that the higher risk of death from any cause was associated with higher serum bilirubin levels.

b. For each population defined by serum bilirubin value, compute the hazard ratio relative to a group having serum bilirubin of 1 mg/dL. (This will be used in problem 6). - See plot in problem 6

4. Perform a statistical regression analysis evaluating an association between serum bilirubin and all-cause mortality by comparing the instantaneous risk (hazard) of death over the entire period of observation across groups defined by serum bilirubin modeled as a continuous logarithmically transformed variable.

a. Why might this analysis be preferred *a priori?*

**Answer:**

A priori this analysis might be preferred because the effect of serum bilirubin on all-cause mortality could potentially be on a multiplicative scale instead of an additive scale, which is typical in analyzing biological data. In addition, the use of log transformation could potentially reduce heteroscedasticity, reduce the influence of larger measurements at the tails and thus improve statistical precision.

b. Include a full report of your inference about the association.

**Method:**

The instantaneous risks of death over the entire period of observation across groups defined by serum bilirubin were compared using proportional hazards regression modeling serum bilirubin as a continuous logarithmically transformed variable. The hazard ratios were computed from the outputs of the regression. The 95% confidence interval and two-sided p-values were constructed using Wald statistics based on the Huber-White sandwich estimator. There were no subjects with missing data on serum bilirubin

**Result:**

There were 418 observations with no missing data on serum bilirubin level at baseline. From a proportional hazards regression analysis, the instantaneous risk of death is a relative 9.89% higher (hazard ratio of 1.0989) for a 10% increase (1.1-fold) in serum bilirubin level at baseline. With 95% confidence, this is consistent with the true instantaneous risk of death being anywhere between 8.26% and 11.53% higher for a 10% increase in serum bilirubin level at baseline. This result is statistically significant at the 5% critical threshold with a two-sided p-value of less than 0.0001. As a result, we reject the null hypothesis, in favor of an alternative hypothesis that the higher risk of death from any cause was associated with higher serum bilirubin levels.

c. For each population defined by serum bilirubin value, compute the hazard ratio relative to a group having serum LDL of 1 mg/dL. - See plot in problem 6

5. One approach to testing to see whether an association between the response and the predictor of interest is adequately modeled by an untransformed continuous variable is to add some other transformation to the model and see if that added covariate provides statistically significant improved “fit” of the data. In this case, we could test for “linearity” of the bilirubin association with the log hazard ratio by including both the untransformed and log transformed bilirubin. (Other alternatives might have been bilirubin and bilirubin squared, but in this case our *a priori* interest in the log bilirubin might drive us to the specified analysis.)

a. Provide full inference related to the question of whether the association is linear.

**Method:**

The instantaneous risks of death over the entire period of observation across groups defined by serum bilirubin were compared using proportional hazards regression including both the continuously untransformed and continuously logarithm transformed serum bilirubin level. The 95% confidence intervals for the coefficients and two-sided p-values were constructed using Wald statistics based on the Huber-White sandwich estimator. There were no subjects with missing data on serum bilirubin.

**Result:**

A test for “linearity” of the bilirubin association with the log hazard ratio was performed by including both the untransformed and the log transformed bilirubin. At the 5% critical threshold, the test on log-transformed term was statistically significant (p-value of less than 0.0001) while the test on the untransformed term was not significant (p-value of 0.148). This result suggests that we have strong evidence that the true association between death from any causes and serum bilirubin was adequately described by a log hazard function that is linear in the log-transformed of serum bilirubin, but not in the untransformed value of serum bilirubin.

b. Again, save the fitted values from this model by obtaining the estimated HRs relative to a group with bilirubin of 1 mg/dl. (This will be used in problem 6.) - See plot in problem 6

6. Display a graph with the fitted hazard ratios from problems 3 - 5. Comment on any similarities or differences of the fitted values from the three models.

**Answer:**

Figure 4 presents the fitted hazard ratios from problem 3 to 5. Figure 5 presents the log-transformed of the fitted values. In figure 5, the fitted curve does not differ very much from the proportional hazards regression fits based on log-transformed predictor and the one based on the combination of the log-transformed and untransformed predictors. In figure 4, the fitted curves for those two models are also closely related for bilirubin level of less than 15 mg/dl, which were also where the vast majority of our data falls in.

Because the model of the combination of untransformed and log-transformed predictor were more flexible and it agreed with the model of log-transformed predictor the most, it suggested that the model that using log-transformed predictor should be preferred in this case over the model with untransformed predictor.

**Figure 4. Relative Hazard Ratio to Group with Serum Bilirubin of 1 mg/dl**



**Figure 5. Log-Transformed Relative Hazard Ratio to Group with Serum Bilirubin of 1 mg/dl**



7. We are interested in considering analyses of the association between all cause mortality and serum bilirubin after adjustment for age and sex.

a. What evidence is present in the data that would make you think that either sex or age might have confounded the association between death and bilirubin? (In real life, we would ideally decide whether to adjust for potential confounding in our pre-specified statistical analysis plan (SAP)).

**Answer**:

By common sense and prior knowledge, one could safely assume that sex could be causally associated with all-cause mortality and not in the causal pathway of interest. Indeed, all others being equal, female individuals tend to live longer than their male counterparts.

In table 2b, we can see that the proportion of female subjects by groups defined by serum bilirubin were different from each other. 93.27% of patients with bilirubin of less than 1 mg/dl were female. These proportions for the remaining groups (between 1 and 2 mg/dl, between 2 and 4 mg/dl, and greater than 4 mg/dl) were 88.31%, 82.54%, and 86.76%, respectively. This suggested that the proportion of female might be associated with serum bilirubin in the sample. As a result, sex might have confounded the association between death and bilirubin.

b. What evidence is present in the data that would make you think that either sex or age might have added precision to the analysis of the association between death and bilirubin? (In real life, we would ideally decide whether to adjust in our pre-specified SAP).

**Answer:**

By common sense and prior knowledge, one could safely assume that age could be causally associated with all-cause mortality and not in the causal pathway of interest. Indeed, all others being equal, older people tend to die earlier than younger people.

In table 2c, we can see that the mean age by groups defined by serum bilirubin were not much different from each other. The mean age for patients with bilirubin of less than 1 mg/dl was 51.2 years. Meanwhile, the mean age for the remaining groups (between 1 and 2 mg/dl, between 2 and 4 mg/dl, and greater than 4 mg/dl) were 51.2, 50.6, and 50.2 years, respectively. This suggested that there was little to no association between age and serum bilirubin in the sample. As a result, including age might have added precision to the analysis of the association between death and bilirubin.

c. Provide full inference regarding an association between death and bilirubin after adjustment for sex and age.

**Method:**

The instantaneous risks of death over the entire period of observation across groups defined by serum bilirubin were compared using proportional hazards regression modeling serum bilirubin as a continuous logarithmically transformed variable. The model also included age as a continuous variable for precision, and sex as a binary variable for adjusting of potential confounder. The hazard ratios were computed from the outputs of the regression. The 95% confidence interval and two-sided p-values were constructed using Wald statistics based on the Huber-White sandwich estimator. 106 observations with missing data on sex at baseline were omitted from this analysis.

**Result:**

Out of 418 observations, there were 312 patients with complete data on serum bilirubin level, age and sex at baseline. From a proportional hazards regression and analysis, when controlling for age and sex (i.e. among groups that have the same age and sex), the instantaneous risk of death is a relative 10.80% higher (hazard ratio of 1.1080) for a 10% increase (1.1-fold) in serum bilirubin level at baseline. With 95% confidence, this is consistent with the true instantaneous risk of death being anywhere between 8.75% and 12.90% higher for a 10% increase in serum bilirubin level at baseline, controlling for age and sex. This result is statistically significant at the 5% critical threshold with a two-sided p-value of less than 0.0001. As a result, we reject the null hypothesis, in favor of an alternative hypothesis that the higher risk of death from any cause was associated with higher serum bilirubin levels when controlling for age and sex.

8. Note that in the above analyses, we completely ignored the intervention in the RCT? What impact could this have had on our results?

**Answer:**

By the design of RCT, patients were randomized to receive either placebo or intervention. Assuming the randomization was done correctly, this design prevent the potential association between serum bilirubin level and the treatment arm in the sample (people with 2 mg/dl serum bilirubin has the same probability to receive placebo as the probability to receive intervention). On the other hand, treatment arm could potentially be associated with all-cause mortality which would not be in the pathway between serum bilirubin and all-cause mortality. Therefore, by completely ignoring the intervention in the RCT as in the above analysis, our results could have lost some precision. Including treatment arms in the RCT in the model could have improve our precision in investigating the association between serum bilirubin and all-cause mortality.