**Overall grade: 53.5 / 105**

**Question 1**

Grading: 2/10. -5 for not presenting descriptive statistics stratified by bilirubin levels. – 2 for not using Kaplan-Meier methods and presenting descriptives (boxplots etc) that do not use KM. -1 for not noting that there was missingness in sex data.

*Methods*: Since the analysis is interested in the association between death and age, sex and serum bilirubin, the data was first summarized by looking at all of the data. However, since we are specifically interested in death and serum bilirubin, summary data was conducted more thoroughly looking at serum bilirubin levels.

*Inference*: The dataset had 418 observations, with 161 observed deaths (Table 1). The average age was 50.74 years old and mean serum bilirubin of 3.22 mg/dl. The sample had only 11.5% males. The mean observation time was 1918 days.

***Table 1. Descriptive Statistics for All Observations***

|  |  |
| --- | --- |
|  | **All** |
| *N Subjects* | 418 |
| *N Deaths* | 161 |
| *Age (years) 1* | 50.74 (10.45; 26.28-78.44) |
| *Serum Bilirubin (mg/dl) 1* | 3.22 (4.41; 0.30-28.00) |
| *Male* | 11.5% |
| *Observation Time (days) 1* | 1918 (1104.67; 41-4795) |

Note: 1 Mean (SD; min-max)

The distribution of serum bilirubin is shown below in the figure, where there are the majority of the observations below 5 mg/dl.



***Figure 1. Histogram of bilirubin***

The spread in serum bilirubin and time to death by gender is provided in the boxplots below. Females had lower median bilirubin levels, but higher mean levels because there were a lot of high level outliers. Females had on average a longer time until death as compared to males.



***Figure 2. Observations by Gender***

Serum bilirubin by time to death was plotted to identify any trends between these two variables. There was larger variation for lower bilirubin, but lower survival (time to death) for higher levels of bilirubin. This is shown below.



***Figure 3. Time to death by bilirubin***

**Question 2**

Grading: 6/10. -2 for not mentioning that not including follow-up time could confound our results. -2 for not mentioning the distribution of censoring with regard to the time of study (e.g., in past CHS data we were able to dichotomize and use logistic regression because all participants had 100% follow-up to that point).

In the prior homework using the Cardiovascular Health Study datasets, we used logistic regression to investigate associations between mortality and various covariates. That method is not advisable with this dataset because this dataset includes observation time (time until death or censoring). This provides an added level of precision; not only do we know the binary information of death or survival, but we know how long they survived. An increase in survival time can be a good indicator of a successful drug. Therefore it is important to capture survival time and a logistic regression cannot encompass this.

**Question 3a**

Grading: 8/10. -1 for not noting what kind of confidence intervals (90%? 95%? 99%?). -1 for misinterpreting the 15.2% increase (should be per 1 mg/dl bilirubin).

*Methods*: A survival analysis evaluating the association between serum bilirubin and all-cause mortality was performed. This was done by comparing the instantaneous risk of death over the observation time by serum bilirubin modeled as a continuous variable. Since serum bilirubin was included as a continuous variable, a cox proportional hazards regression model was used. The association between all cause mortality and bilirubin levels was computed from a regression model with confidence intervals and two-sided p values computed using Wald statistics based on the Huber-White sandwich estimator.

*Inference*: There were a total of 418 subjects’ data used in the model, with 161 deaths. These subjects had a mean bilirubin level of 3.22 mg/dl. From the proportional hazards regression analysis, the instantaneous risk of death is relative 15.2% higher (hazard ratio 1.152) for each 10% higher serum bilirubin level. Based on a 95% confidence interval, this observed hazard ratio suggesting higher death rates for groups of patients with higher serum bilirubin would not be unusual if the true instantaneous risk of death were anywhere from 12.7% and 17.9% higher in a group having a baseline bilirubin 10 mg/dl higher than that in another group (95% CI for hazard ratio 1.127 – 1.179). The two sided Wald Statistic has a p-value < 0.0001, therefore we can reject the null hypothesis that serum bilirubin levels are not associated with death, favoring a tendency for higher mortality with higher bilirubin. The following plot shows the Kaplan-Meier survival estimates controlling for serum bilirubin as a continuous predictor.



***Figure 3. Kaplan-Meier survival estimates***

**Question 3b**

*Methods*: The hazard ratio was computed for each serum bilirubin value by using the coefficient for serum bilirubin from the regression model above. This hazard ratio was computed by taking e to the coefficient times the serum bilirubin level.

*Inference*: The following table shows the hazard ratio for each serum bilirubin level relative to a group having a serum bilirubin level of 1 mg/dl.

***Table 3. Hazard Ratios***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Serum Bilirubin* *(mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* |
| 0.3 | 1.0435 | 3.6 | 1.6664 | 8.6 | 3.3869 |
| 0.4 | 1.0584 | 3.7 | 1.6902 | 8.7 | 3.4353 |
| 0.5 | 1.0735 | 3.8 | 1.7143 | 8.9 | 3.5341 |
| 0.6 | 1.0888 | 3.9 | 1.7388 | 9 | 3.5846 |
| 0.7 | 1.1044 | 4 | 1.7637 | 9.5 | 3.8481 |
| 0.8 | 1.1202 | 4.2 | 1.8144 | 10.8 | 4.6273 |
| 0.9 | 1.1362 | 4.4 | 1.8666 | 11 | 4.7605 |
| 1 | 1.1524 | 4.5 | 1.8933 | 11.1 | 4.8285 |
| 1.1 | 1.1689 | 4.6 | 1.9204 | 11.4 | 5.0384 |
| 1.2 | 1.1856 | 4.7 | 1.9478 | 12.2 | 5.6439 |
| 1.3 | 1.2025 | 5 | 2.0325 | 12.6 | 5.9734 |
| 1.4 | 1.2197 | 5.1 | 2.0615 | 13 | 6.3221 |
| 1.5 | 1.2371 | 5.2 | 2.0910 | 13.6 | 6.8837 |
| 1.6 | 1.2548 | 5.4 | 2.1511 | 13.8 | 7.0818 |
| 1.7 | 1.2727 | 5.5 | 2.1819 | 14 | 7.2856 |
| 1.8 | 1.2909 | 5.6 | 2.2130 | 14.1 | 7.3897 |
| 1.9 | 1.3093 | 5.7 | 2.2446 | 14.4 | 7.7109 |
| 2 | 1.3280 | 5.9 | 2.3092 | 14.5 | 7.8211 |
| 2.1 | 1.3470 | 6 | 2.3422 | 15 | 8.3960 |
| 2.2 | 1.3663 | 6.1 | 2.3757 | 16 | 9.6755 |
| 2.3 | 1.3858 | 6.3 | 2.4440 | 16.2 | 9.9540 |
| 2.4 | 1.4056 | 6.4 | 2.4790 | 17.1 | 11.3094 |
| 2.5 | 1.4256 | 6.5 | 2.5144 | 17.2 | 11.4710 |
| 2.6 | 1.4460 | 6.6 | 2.5503 | 17.4 | 11.8011 |
| 2.7 | 1.4667 | 6.7 | 2.5867 | 17.9 | 12.6685 |
| 2.8 | 1.4876 | 6.8 | 2.6237 | 18 | 12.8494 |
| 2.9 | 1.5089 | 7.1 | 2.7377 | 20 | 17.0645 |
| 3 | 1.5304 | 7.2 | 2.7769 | 21.6 | 21.4122 |
| 3.1 | 1.5523 | 7.3 | 2.8165 | 22.5 | 24.3279 |
| 3.2 | 1.5745 | 8 | 3.1106 | 24.5 | 32.3083 |
| 3.3 | 1.5970 | 8.1 | 3.1550 | 25.5 | 37.2322 |
| 3.4 | 1.6198 | 8.4 | 3.2921 | 28 | 53.0800 |
| 3.5 | 1.6429 | 8.5 | 3.3392 |  |  |

**Question 4a**

Grading: 2.5/5. -2.5 for not mentioning multiplicative scale (though you did mention the distribution).

*Methods*: The log transformation of serum bilirubin was considered for use in the regression analysis similar to above.

*Inference*: Using log transformed serum bilirubin may be preferred a priori because the bilirubin is skewed towards zero. By transforming the variable, this can be addressed. The histogram below shows this skewed distribution of serum bilirubin.



***Figure 4. Distribution of Untransformed Serum Bilirubin***

**Question 4b**

*Grading: 5/10. -4 for report of HR and 95% CI. -1 for not mentioning width of CI in methods.*

*Methods*: Similar to Question 3, a survival analysis evaluating the association between serum bilirubin and all-cause mortality was performed. The only difference from question 3 above is that serum bilirubin was log transformed. The association between all cause mortality and bilirubin levels was computed from a regression model with confidence intervals and two-sided p values computed using Wald statistics based on the Huber-White sandwich estimator. In order to compare by a 10% increase in bilirubin, the log of bilirubin was also divided by the log base 1.1

*Inference*: There were a total of 418 subjects’ data used in the model, with 161 deaths. These subjects had a mean bilirubin level of 3.22 mg/dl. From the proportional hazards regression analysis, the instantaneous risk of death is relative 9.9% higher (hazard ratio 1.098) for each 10% higher serum bilirubin level. Based on a 95% confidence interval, this observed hazard ratio suggesting higher death rates for groups of patients with higher serum bilirubin would not be unusual if the true instantaneous risk of death were anywhere from 8.3% and 11.5% higher in a group having a baseline bilirubin 10 mg/dl higher than that in another group (95% CI for hazard ratio 1.083 – 1.115). The two sided Wald Statistic has a p-value < 0.0001, therefore we can reject the null hypothesis that serum bilirubin levels are not associated with death, favoring a tendency for higher mortality with higher bilirubin. The following plot shows the Kaplan-Meier survival estimates controlling for serum bilirubin as a continuous predictor.



***Figure 5. Kaplan-Meier survival estimates for log transformation of bilirubin***

**Question 4c**

*Methods*: The hazard ratio was computed for each serum bilirubin value by using the coefficient for serum bilirubin from the regression model above. This hazard ratio was computed by taking e to the coefficient times the log serum bilirubin level.

*Inference*: The following table shows the hazard ratio for each serum bilirubin level relative to a group having a serum bilirubin level of 1 mg/dl.

***Table 4. Hazard Ratios for log transformed bilirubin***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Serum Bilirubin* *(mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* |
| 0.3 | 0.5962 | 3.6 | 1.7336 | 8.6 | 2.5201 |
| 0.4 | 0.6746 | 3.7 | 1.7542 | 8.7 | 2.5326 |
| 0.5 | 0.7425 | 3.8 | 1.7744 | 8.9 | 2.5575 |
| 0.6 | 0.8030 | 3.9 | 1.7943 | 9 | 2.5698 |
| 0.7 | 0.8579 | 4 | 1.8139 | 9.5 | 2.6302 |
| 0.8 | 0.9086 | 4.2 | 1.8523 | 10.8 | 2.7791 |
| 0.9 | 0.9558 | 4.4 | 1.8897 | 11 | 2.8011 |
| 1 | 1.0000 | 4.5 | 1.9080 | 11.1 | 2.8120 |
| 1.1 | 1.0418 | 4.6 | 1.9261 | 11.4 | 2.8444 |
| 1.2 | 1.0815 | 4.7 | 1.9440 | 12.2 | 2.9285 |
| 1.3 | 1.1193 | 5 | 1.9964 | 12.6 | 2.9694 |
| 1.4 | 1.1555 | 5.1 | 2.0134 | 13 | 3.0095 |
| 1.5 | 1.1903 | 5.2 | 2.0303 | 13.6 | 3.0684 |
| 1.6 | 1.2237 | 5.4 | 2.0635 | 13.8 | 3.0877 |
| 1.7 | 1.2560 | 5.5 | 2.0798 | 14 | 3.1069 |
| 1.8 | 1.2872 | 5.6 | 2.0960 | 14.1 | 3.1164 |
| 1.9 | 1.3175 | 5.7 | 2.1120 | 14.4 | 3.1447 |
| 2 | 1.3468 | 5.9 | 2.1435 | 14.5 | 3.1540 |
| 2.1 | 1.3753 | 6 | 2.1590 | 15 | 3.2003 |
| 2.2 | 1.4031 | 6.1 | 2.1744 | 16 | 3.2903 |
| 2.3 | 1.4301 | 6.3 | 2.2047 | 16.2 | 3.3079 |
| 2.4 | 1.4565 | 6.4 | 2.2197 | 17.1 | 3.3856 |
| 2.5 | 1.4823 | 6.5 | 2.2345 | 17.2 | 3.3941 |
| 2.6 | 1.5075 | 6.6 | 2.2492 | 17.4 | 3.4110 |
| 2.7 | 1.5321 | 6.7 | 2.2638 | 17.9 | 3.4528 |
| 2.8 | 1.5562 | 6.8 | 2.2783 | 18 | 3.4610 |
| 2.9 | 1.5799 | 7.1 | 2.3209 | 20 | 3.6213 |
| 3 | 1.6031 | 7.2 | 2.3349 | 21.6 | 3.7430 |
| 3.1 | 1.6258 | 7.3 | 2.3488 | 22.5 | 3.8092 |
| 3.2 | 1.6481 | 8 | 2.4430 | 24.5 | 3.9511 |
| 3.3 | 1.6700 | 8.1 | 2.4561 | 25.5 | 4.0196 |
| 3.4 | 1.6916 | 8.4 | 2.4947 | 28 | 4.1844 |
| 3.5 | 1.7128 | 8.5 | 2.5075 |  |  |

**Question 5a**

Grading: 8/10. -1 for not noting the width of the CI in the methods. -1 for incorrect p-value report in Results (see key, p=0.148)

*Methods*: A model was created including both covariates bilirubin and log transformed bilirubin to test for a linear association. The analysis compared model fit to determine whether the association was linear. The regression model used confidence intervals and two-sided p values computed using Wald statistics based on the Huber-White sandwich estimator. Similar to question 4, the log transformed bilirubin was divided by log 1.1 to compare relative to the 10% increase in bilirubin.

*Inference*: There were a total of 418 observations used in the model, with 161 deaths. The model shows that the log transformation of bilirubin is significant (p < 0.001), but the non-transformed bilirubin is not significant (p = 0.191). Therefore the relationship is not linear between death and bilirubin. The following table further shows the results from this regression model.

***Table 5. Regression Output for Inclusion of Both Regressors***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Variable* | *Coefficient* | *HR* | *SE* | *z* | *p-value* | *HR CI* |
| **Bilirubin** | -0.0395 | 0.9614 | 0.0301 | -1.308 | 0.191 (ns) | (0.906, 1.020) |
| **Log(Bilirubin)** | 0.1128 | 1.1194 | 0.0160 | 7.071 | < 0.0001 | (1.085, 1.155) |
| R2 = 0.31 Wald Test: p – value < 0.0001 |

**Question 5b**

*Methods*: Similar to before, the hazard ratio was computed for each serum bilirubin value by using the coefficient for serum bilirubin from the regression model above. This hazard ratio was computed by taking e to the corresponding coefficient times the log serum bilirubin level plus e to the corresponding coefficient times the serum bilirubin level.

*Inference*: The following table shows the hazard ratio for each serum bilirubin level relative to a group having a serum bilirubin level of 1 mg/dl.

***Table 6. Hazard Ratios for log transformed bilirubin + untransformed bilirubin***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Serum Bilirubin* *(mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* | *Serum Bilirubin (mg/dl)* | *Hazard Ratio* |
| 0.3 | 1.5269 | 3.6 | 2.7995 | 8.6 | 3.7347 |
| 0.4 | 1.6088 | 3.7 | 2.8234 | 8.7 | 3.7500 |
| 0.5 | 1.6808 | 3.8 | 2.8471 | 8.9 | 3.7801 |
| 0.6 | 1.7458 | 3.9 | 2.8704 | 9 | 3.7950 |
| 0.7 | 1.8053 | 4 | 2.8934 | 9.5 | 3.8685 |
| 0.8 | 1.8607 | 4.2 | 2.9385 | 10.8 | 4.0509 |
| 0.9 | 1.9125 | 4.4 | 2.9824 | 11 | 4.0780 |
| 1 | 1.9614 | 4.5 | 3.0040 | 11.1 | 4.0914 |
| 1.1 | 2.0078 | 4.6 | 3.0253 | 11.4 | 4.1314 |
| 1.2 | 2.0521 | 4.7 | 3.0464 | 12.2 | 4.2355 |
| 1.3 | 2.0945 | 5 | 3.1082 | 12.6 | 4.2863 |
| 1.4 | 2.1352 | 5.1 | 3.1283 | 13 | 4.3363 |
| 1.5 | 2.1744 | 5.2 | 3.1483 | 13.6 | 4.4100 |
| 1.6 | 2.2122 | 5.4 | 3.1876 | 13.8 | 4.4342 |
| 1.7 | 2.2488 | 5.5 | 3.2070 | 14 | 4.4583 |
| 1.8 | 2.2843 | 5.6 | 3.2262 | 14.1 | 4.4702 |
| 1.9 | 2.3188 | 5.7 | 3.2452 | 14.4 | 4.5059 |
| 2 | 2.3523 | 5.9 | 3.2826 | 14.5 | 4.5177 |
| 2.1 | 2.3849 | 6 | 3.3011 | 15 | 4.5761 |
| 2.2 | 2.4167 | 6.1 | 3.3194 | 16 | 4.6904 |
| 2.3 | 2.4478 | 6.3 | 3.3556 | 16.2 | 4.7128 |
| 2.4 | 2.4781 | 6.4 | 3.3735 | 17.1 | 4.8123 |
| 2.5 | 2.5078 | 6.5 | 3.3912 | 17.2 | 4.8232 |
| 2.6 | 2.5368 | 6.6 | 3.4088 | 17.4 | 4.8450 |
| 2.7 | 2.5653 | 6.7 | 3.4262 | 17.9 | 4.8988 |
| 2.8 | 2.5932 | 6.8 | 3.4435 | 18 | 4.9095 |
| 2.9 | 2.6205 | 7.1 | 3.4946 | 20 | 5.1180 |
| 3 | 2.6474 | 7.2 | 3.5114 | 21.6 | 5.2783 |
| 3.1 | 2.6738 | 7.3 | 3.5281 | 22.5 | 5.3663 |
| 3.2 | 2.6998 | 8 | 3.6415 | 24.5 | 5.5567 |
| 3.3 | 2.7253 | 8.1 | 3.6573 | 25.5 | 5.6495 |
| 3.4 | 2.7504 | 8.4 | 3.7041 | 28 | 5.8753 |
| 3.5 | 2.7751 | 8.5 | 3.7194 |  |  |

**Question 6**

Grading: 2/10. +2 correct that log transformed and model with both versions of bilirubin are similar. Otherwise fitted values very far off from those presented in the key.

*Methods*: The fitted values for the hazard ratios were plotted for the three scenarios from questions 3, 4, and 5 for comparison of the three methods.

*Inference*: The following plot shows the fitted hazard ratios based on serum bilirubin levels and the coefficients from questions 3, 4, and 5. All three models are similar for low levels of bilirubin, however they part around 5 mg/dl. The transformed model and the model with both versions of bilirubin are very similar, but the one including both is shifted right.

 

***Figure 6. Model Comparisons of Bilirubin Transformations***

**Question 7a**

Grading: 6/10. -2 for not using KM methods to report statistics with censored data. -2 for incorrect descriptive statistic (men should tend to have higher bilirubin than women; see key).

The data suggest that sex may have a confounding association between death and bilirubin. Males had a mean bilirubin level of 2.872 mg/dl, while females had a mean level of 3.306 mg/dl. Furthermore, the proportion of deaths observed in females was lower, 103 deaths from 276 observations (37.3%); whereas males had 22 deaths from 36 observations (61.1%).

The data suggest that age may have a confounding association between death and bilirubin, however this is not as strong as the effect of sex. The following figure shows the scatter of serum bilirubin by age, where there is not an apparent trend between serum bilirubin and age. However, the middle age group appears to have more variance in bilirubin. Literature suggests that older patients will be more vulnerable to death as compared to the younger patients.



***Figure 7. Serum bilirubin levels across ages***

**Question 7b**

Grading: 0/10. See key. Needed cox proportional hazards regression model including these variables.

Since sex is not necessarily a strong indicator of death, this may actually act as added precision to the analysis. Mean bilirubin levels were different between the groups, where females had on average higher levels. Therefore sex could be a precision variable.

There was no strong association between age and serum bilirubin, as shown above. However, age is generally a good indicator of death. Therefore, sex would make a better precision variable.

**Question 7c**

Grading: 4/10. -1 for not noting the type of regression model (cox PH); -1 for not noting the width of the confidence interval. -1 for not noting how age/sex were modeled in the regression (age continuous? In 10 year intervals?); sex – male or female as the reference group?); -3 for not reporting correct HR/CIs (see key).

*Methods*: Similar to the methods of analysis above, a survival analysis evaluating the association between serum bilirubin and all-cause mortality was performed. This time, age and sex were included in the model. The association between all cause mortality and bilirubin levels was computed from a regression model with confidence intervals and two-sided p values computed using Wald statistics based on the Huber-White sandwich estimator.

*Inference*: There were a total of 312 subjects’ data used in the model, with 125 deaths. There were 106 observations removed for missing data. These subjects had a mean bilirubin level of 3.22 mg/dl. From the proportional hazards regression analysis, the instantaneous risk of death is relative 16.2% higher (hazard ratio 1.162) for each 10% higher serum bilirubin level. Based on a 95% confidence interval, this observed hazard ratio suggesting higher death rates for groups of patients with higher serum bilirubin would not be unusual if the true instantaneous risk of death were anywhere from 13.2% and 19.2% higher in a group having a baseline bilirubin 10 mg/dl higher than that in another group (95% CI for hazard ratio 1.1322 – 1.192).

The instantaneous risk of death is relative 3.8% higher (hazard ratio 1.038) for each 10 year increase in age. Based on a 95% confidence interval, this observed hazard ratio suggesting higher death rates for older patients would not be unusual if the true instantaneous risk of death were anywhere from 2.0% and 5.7% higher in a group having a baseline age of 10 years higher than that in another group (95% CI for hazard ratio 1.020 – 1.057).

The instantaneous risk of death is relative 33.9% lower (hazard ratio 0.660) for females as compared to males. However, this is not significant, as the 95% confidence interval suggest that this observed hazard ratio suggesting higher death rates for females would not be unusual if the true instantaneous risk of death were anywhere from 58.95% lower and 6.2% higher in females as compared to males (95% CI for hazard ratio 0.410 – 1.062).

From the p –value, we can reject the null hypothesis that serum bilirubin level and age are not associated with death, favoring a tendency for higher mortality with higher bilirubin and males. However, we cannot reject the null hypothesis that gender and death are not associated. The following table and plots further show the output of this analysis.

***Table 6. Regression Output Adjusting for Sex and Age***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Variable* | *Coefficient* | *HR* | *SE* | *z* | *p-value* | *HR CI* |
| **Bilirubin** | 0.150 | 1.162 | 0.0130 | 11.452 | < 0.0001 | (1.132, 1.192) |
| **Age** | 0.038 | 1.038 | 0.0090 | 4.1900 | < 0.0001 | (1.020, 1.057) |
| **Female** | -0.415 | 0.660 | 0.2425 | -1.711 | 0.087 | (0.410, 1.062) |
| R2 = 0.291 Wald Test: p – value < 0.0001 |



***Figure 8. Kaplan-Meier survival estimates after adjusting for age and sex***

**Question 8**

Grading: 10/10.

The above analyses ignored the intervention of the RCT. The models do not account for the treatment/control groups from the drug study. We could have stratified our analysis by treatment. However, this would have had a very small impact on our results. Since the treatment/control was random, the association between death and bilirubin should have been the similar across these two groups. The random assignment of treatment would have resulted in similar bilirubin distributions within the two treatment groups. Therefore the results would be very similar.