Biostatistics 518: Winter 2015

Homework #1

1665

1. If we are interested in death from any cause by 4 years after enrollment, this is appropriate because the minimum follow-up time in patients not dying during the study period is 1,480 days, or 4.05 years. There is no right-censoring for the first 4 years of the study, therefore we can use this outcome without having to deal with the issue of right-censoring.

Methods: The predictor of interest, C-Reactive Protein (CRP) was divided into 3 levels (< 1, 1-3, and > 3) based on the low, medium and high risk of cardiovascular disease associated with these 3 levels. An outcome variable was created for death from any cause within 4 years of study enrollment. Descriptive statistics were presented within the 3 CRP levels. Binary variables (sex, smoking status, history of CVD and death within 4 years) were presented as number and percentage and continuous variables (age, BMI, cholesterol) were expressed as means and standard deviations.

Results: CRP levels were not available for 67 subjects, therefore data for those subjects was not used in the analysis. For the other variables, if data was missing, the mean was calculated using only the available data. There was no variable with a large (> 5%) percentage of missing data. Of the remaining 4,933 subjects, 428 had CRP < 1, 3,330 had CRP between 1 and 3, and 1175 had CRP > 3. The table below presents descriptive statistics between these groups and for the group as a whole. Of note, the proportion of women, BMI, proportion of smokers, and proportion with history of cardiovascular disease appear to increase with increasing CRP level and is most marked in the group with CRP > 3. There does not appear to be any trend in age or serum cholesterol across CRP levels. There appears to be an increased incidence of mortality as CRP increases, with a cumulative mortality of 4.9% in those with CRP < 1 and 15.6% in those with CRP of > 3. Associations between CRP and obesity, smoking, and pre-existing cardiovascular disease likely confound the observed association between CRP and 4 year mortality and these interactions should be investigated further.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Serum C-Reactive Protein (CRP)** | | | |
|  |  | **Below 1 mg/L**  **n=428** | **1-3 mg/L**  **n=3330** | **Above 3 mg/L**  **n=1175** | **Any level**  **n = 4,933** |
| **Male** | n(%) | 195 (45.6%) | 1,442 (43.3%) | 435 (37%) | 2,072 (42%) |
| **Age (years)** | Mean (SD) | 73.451 (5.804) | 72.737 (5.524) | 72.735 (5.581) | 72.799 (5.565) |
| **BMI n=4920** | Mean (SD) | 23.816 (3.639) | 26.389 (4.306) | 28.455 (5.463) | 26.658 (4.721) |
| **Current smoker n=4927** | n(%) | 50 (9.6%) | 366 (11%) | 193 (16.4%) | 600 (12.2%) |
| **Serum cholesterol (mg/dL) n=4930** | Mean (SD) | 205.995 (40.527) | 212.828 (38.573) | 210.502 (40.387) | 211.683 (39.226) |
| **History of CVD** | n(%) | 78 (18.2%) | 715 (21.5%) | 338 (28.8%) | 1,131 (22.9%) |
| **Death within 4 years** | n(%) | 21 (4.9%) | 280 (8.4%) | 183 (15.6%) | 484 (9.8%) |

1. Methods: Mean CRP level was compared between the group of subjects that died within 4 years of study enrollment and those who did not. A t-test that did not presume equal variances was performed and 95% confidence interval calculated for the difference in means.

Results: The mean CRP level among subjects who died within 4 years of enrollment was 1.954 mg/L higher than those still alive at 4 years (5.376 mg/L vs. 3.422 mg/L). With 95% confidence, we can say that this difference in means would not be unusual if the true population difference in mean CRP was between 1.21 and 2.70 mg/L higher in patients who die within 4 years. This difference is statistically significant with a p-value < 0.0001 and thus we can state with high confidence that the mean CRP level is higher in patients who have a higher likelihood of death within a 4 year period.

1. Methods: Geometric mean serum CRP was compared between subjects who died within 4 years of study enrollment and those who did not. CRP was log-transformed to account for its presence on a multiplicative scale in inflammatory disorders. In order to accomplish this, all CRP values reported as 0 were changed to 0.5 (50% of the lower limit of detection for the assay). A t-test that did not presume equal variances was performed to compare differences in mean log-transformed CRP levels. A p-value and 95% confidence interval were calculated and the results were then exponentiated to provide the results below.

Results: Geometric mean CRP in the 484 subjects who died within 4 years of study enrollment was 2.03 and for those still alive at 4 years was 1.115, 46.4% higher in those who died than those who did not. Using a 95% confidence interval, we can say that this result would not be unusual if the true population ratio of geometric means showed a difference between 33.2% and 60.9% higher CRP in people who died within 4 years (p<0.0001).

1. Methods: A dichotomous predictor of interest variable was created by dividing CRP levels into less than or equal to 3, which I will call “low” and greater than three, which I will call “high”. The outcome remained the same. A chi-square test was used to compare the probabilities of death within 4 years among the high and low CRP groups. A 95% confidence interval and p-value were calculated for the difference in probabilities.

Results: Among those with CRP > 3, 183 or 15.6% died within 4 years, compared to 301 or 8% of subjects with CRP ≤ 3, for a difference of 7.6% higher probability to death in the high CRP group. Using a 95% confidence interval, this result would not be unusual if the true population difference in probability of death within 4 years between the 2 CRP groups was 5.3% to 9.8% higher in patients with CRP >3 (p<0.0001). With high confidence, we can therefore reject the null hypothesis that there is no difference in probability of death within 4 years between the high and low CRP groups.

1. Methods: A dichotomous predictor of interest variable was created by dividing CRP levels into less than or equal to 3, which I will call “low” and greater than three, which I will call “high”. The outcome remained the same. A chi-square test was used to compare the probabilities of death within 4 years among the high and low CRP groups. An odds-ratio and 95% confidence interval for the ratio was also calculated.

Results: Among those with CRP > 3, 183 or 15.6% died within 4 years, compared to 301 or 8% of subjects with CRP ≤ 3. Within this sample, the odds of dying within 4 years was 2.12 times higher in the high CRP group when compared to the low CRP group. Using a 95% confidence interval, this result would not be unusual if the true population odds of death within 4 years was between 1.74 and 2.58 times higher in patients with CRP >3 (p<0.0001). With high confidence, we can therefore reject the null hypothesis that there is no difference in odds of death within 4 years between the high and low CRP groups.

1. Methods: The probability of survival over time was estimated using a Kaplan-Meier curve (pictured below) for subjects with high (> 3 mg/L) and low (≤ 3 mg/L) CRP levels. The logrank test was used to determine the difference in survival between the two groups. Cox proportional hazards regression was then used to calculate the hazard ratio (difference in risk of instantaneous death) and 95% confidence interval for the hazard ratio.

Results: As seen below in the Kaplan-Meier curves for survival in 1,175 subjects with high and 3,758 subjects with low CRP levels, the probability of survival is lower in the high CRP group at all time points. The difference in survival increases steadily until year 5 and then remains relatively steady around 10% for the remainder of the observation period. A Hazard ratio of 1.687 indicates a 68.7% higher risk of instantaneous death in the high CRP group. Based on a 95% confidence interval, this result would not be unusual if the true population hazard ratio were between 1.485 and 1.917. The logrank test gives a p-value of <0.0001, indicating that we can with high confidence reject the null hypothesis that there is no difference in survival between the two groups.



1. I think the best choice in this case would be to treat both death and CRP as continuous variables, and likely to log-transform CRP in that analysis to account for its presence on a multiplicative scale in inflammatory disorders. Unfortunately I don’t know how to do than analysis at this time. Therefore, I think best and most straightforward analysis would be to look at mean log-transformed CRP levels among subjects who die during pre-specified intervals. 4 year mortality could be used given that all survivors were followed to that point, but one could also do 1-, 2-, 3-year and 4-year mortality to get a better idea of the time-course of how CRP levels might be associated with mortality. Another option that I might consider would be to present several Kaplain-Meier curves stratified by several levels of log-transformed CRP and estimate survival probabilities and hazard ratios between those CRP levels. This would be advantageous in that we could use data from the entire observation period but could be cumbersome to perform and potentially difficult to present clearly.