(31 points given in total)

Biostat 518 HW 01

1. I don’t know how to verify this with the dataset, because there was no recorded time of enrollment, but I would assume that 4 years is the time-frame for which we have information on all participants. The maximum follow-up time was 8 years, so there are some people for whom we know more, but in order to make use of more data points we’ve chosen to look only at time of death up to 4 years. （ 0 points for question 1, because to get any credit, the answer must include the minimum time of follow-up time for a censored observation)
2. Table 1: There were 67 missing observations for CRP. Descriptive statistics are categorized by risk of heart disease. (5 points given)(no title for the table, the last column should be >3mg/L instead of <3mg/L, the subjects numbers are wrong for the first two groups, should be 428 and 3330 respectively. And there’s no column about the situation for total population. The percentages of Male in three columns are wrong for the first two groups, so are the answers for the BMI, smoker%, prevdis% of the first two groups, and there’s no information about the percentage of Death w/in 4 years. ( I give 3 points to the general table layout , 2 points to the choice of descriptive statistics, no points to the discussion of the finding)

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk of heart disease** | **Low risk (<1mg/L CRP)** | **Avg. risk (1-3 mg/L CRP)** | **High risk (<3 mg/L CRP)** |
| Subjects(n) | 1969 | 1789 | 1175 |
| Male(%) | 44.8 | 42.2 | 37.0 |
| Age (yrs)\*  | 73, 5.71, 65-98  | 72.5, 5.37, 65-100 | 72.7, 5.58, 65-93 |
| BMI\* | 25.14, 3.84, 14.7-43.2 | 27.15, 4.56, 15.1-53.2 | 28.45, 5.46, 15.3-58.8 |
| Smoker(%) | 9.3 | 12.6 | 16.4 |
| Chol. (mg/dL)\* | 210.31, 38.03, 73-407 | 213.96, 39.67, 78-354 | 210.5, 40.39, 97-430 |
| Prev. CVD (%) | 18.9 | 23.5 | 28.8 |
| Death w/in 8 yrs(%) | 18.3 | 22.3 | 29.7 |

\*avg, std dev, range

1. Mean Blood C reactive protein levels (mg/L) differed significantly between participants who died within four years of enrolling in our study vs. those who survived more than four years. We estimate a mean difference between groups of 1.95 mg/L, with a 95% confidence interval of 1.21mg/L to 2.70mg/L, p<.0001.(4 points given) (no information about what statistic analysis were performed(t test equal or unequal variance?), no point estimates for the two groups, and the point estimate for the difference should be 1.95 mg/dL, no interpretation of the results)
2. Mean Blood C reactive protein levels (mg/L) differed significantly between participants who died within four years of enrolling in our study vs. those who survived more than four years. We estimate a geometric mean difference between groups of 6.89 mg/L, with a 95% confidence interval of 3.28mg/L to 14.47, p<.0001. Note: values of 0 were recoded as .5. UGGGHHHH HOW TO INTERPRET? (5 points given) (no interpretation of the results and no mention about the specific analysis performed, and the answers were wrong)
3. There was a significant difference between the proportion of people with low CRP who died within the 4-year study time, and deaths within the high CRP group. (8.01% vs. 15.57% respectively). Our chi-squared value was 57.89, with p<.001.(5 points were given) (no definition of low and high crp group, which should be specified as <=3mg/L Vs >3mg/L. No details about which analysis were employed to compare the proportion(there are many analyses could do this), and no interpretation about the results)
4. The odds of death were significantly higher (p<.001) for the group with high CRP with an estimated odds ratio of 2.12 (95% confidence interval of 1.74-2.57).(6 points given) (no interpretation of the results and no mentioning about the specific analysis performed, no point estimates for the odds for each group,)
5. The instantaneous risk of death is 2.048 times higher for people with high CRP levels than those with low (p<.0001, 95%CI 1.70-2.46).(3 points given) (the answers are wrong, and no description about the analyses performed, no interpretation about the results)
6. I would have chosen to report the odds ratio, as I think it is the easiest to interpret and make sense of. The vagueness of a unit of time makes me dislike the hazard ratio, and the difference in probability is also somewhat abstract, but hearing that the odds of death being twice as high with high CRP levels would make me want to do something to lower my CRP levels. ( 3 points given) (Simpler comparisons of means and proportions are actually preferred here, no mentioning of other points in the keys,)