1. It is valid to dichotomize the data by 5-year all-cause mortality because all survivors (those without an all-cause mortality) have an observation time greater than 5 years and therefore no censoring is necessary when setting a 5-year cutoff. The range of observation time for survivors is 1827 to 2159 days. By dividing each by 365 days/year, we obtain a range of follow-up time for survivors of 5.01 – 5.92 years. Thus, all survivors have been noted to be alive at a visit greater than the 5-year cutoff for this dichotomization variable.

2. For the table below, I stratified by 5-year all-cause mortality status and presented the mean and 95% confidence interval. **The 95% confidence interval for questions 2-4 were calculated from a t-distribution with n-1 degrees of freedom.** P-values represent the difference in means for each variable between the 5-year all-cause mortality groups, as evaluated by a two-sample, two-sided t-test without the assumption of equal variance. As noted in the table below, the two groups are significantly different (t-test p-value < 0.05) for all variables except for weight.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **No 5-Year Mortality****(N=602)\*** | **5-Year Mortality****(N=133)\*** | **P-Value** |
| Age, years | 74.1 (73.7 – 74.5) | 76.7 (75.6 – 77.8) | 1.64 x 10-5 |
| Male gender, % | 46.5 (42.5 – 50.5) | 64.7 (56.4 – 72.9) | 1.22 x 10-4 |
| Weight, lb | 160.1 (157.7 – 162.5) | 159.2 (153.6 – 164.8) | 0.76 |
| Pack years, years | 17.9 (15.9 – 19.9) | 27.2 (22.0 – 32.4)d | 0.0046 |
| LDL-C, mg/dl | 127.4 (124.8 – 130.0)c | 118.6 (112.4 – 124.8)e | 0.011 |
| CHD History, %a | 17.8 (11.1 – 24.3) | 36.1 (27.8 – 44.3) | 6.25 x 10-5 |
| CHF History, % | 3.8 (0.5 – 7.1) | 13.5 (7.6 – 19.4) | 0.0019 |
| Stroke History, %b | 9.9 (4.8 – 15.1) | 29.3 (21.5 – 37.2) | 6.45 x 10-6 |

Abbreviations: CHD = coronary heart disease; CHF = congestive heart failure; LDL-C = low density lipoprotein concentration.

\*95 percent confidence intervals calculated from a t-distribution are presented in parentheses, with the lower and upper bound, respectively.

a CHD defined as the diagnosis of either angina or a myocardial infarction.

b Stroke defined as the diagnosis of either a TIA or a stroke.

c Missing data; n=594

d Missing data; n=132

e Missing data; n=131

3. Using a two-sample, two-sided t-test without the assumption of equal variance, I tested for a difference in mean LDL-C between the survivor and 5-year mortality groups. Those who survived to 5 years had a mean LDL of 127.4 with a 95% confidence interval (CI) of 124.8 to 130.0. In comparison, the 5-year mortality group had a mean LDL-C of 118.6 with a 95% CI of 112.4 to 124.8. The difference in mean LDL-C between the two groups was significant, with a p-value of 0.011. However, it should be noted that those who suffered a 5-year mortality had a lower mean LDL-C than the survivor group.

4. Using a natural log transformation of LDL-C values, I used a two-sample, two-sided t-test without the assumption of equal variance to test for a difference in the mean log-transformed LDL-C value between the survivor and 5-year mortality groups. **The estimates and confidence intervals were then transformed back through exponentiation for the purposes of reporting results.** Those who survived to 5 years had a geometric mean LDL-C of 122.9, with a 95% CI of 120.3 to 125.2. In comparison, the 5-year mortality group had a mean log10-transformed LDL-C of 112.2 with a 95% CI of 104.6 to 120.3. The difference in the geometric mean LDL-C between the two groups was significant, with a p-value of 0.0075, though it should be noted that those who suffered a 5-year mortality had a lower geometric mean LDL-C than the survivor group.

5. Both questions 5 and 6 will be answered by using the table below:

|  |  |  |
| --- | --- | --- |
|  | **Normal LDL-C** | **High LDL-C** |
| **Survivor** | 502 | 92 |
| **5-Year All-Cause Mortality** | 116 | 15 |

The probability of death given a high LDL-C can be calculated by the relative risk (RR). The RR is calculated given the numbers in this table by: [(15 / (15+92)) / (116 / (116+502))], which equals 0.75. The CI was calculated using the log(RR) transformation to obtain the standard error (SE), then the SE was transformed back to the original units and multiplied by 1.96 to obtain a 95% CI of 0.45 to 1.23 risk of death during the interval given a high LDL-C.

6. The odds of death given a high LDL-C can be calculated by the odds ratio (OR), given by: [(502\*15) / (92\*116)], which equals 0.706. The CI was calculated given the method described in the answer to #5, yielding a 95% CI of 0.39 to 1.26 for the OR of death during the interval given a high LDL-C.

7. The instantaneous hazard of death was calculated from a Kaplan-Meier survival function in R using the package “epiR” and the library function “epi.insthaz”. This instantaneous hazard of death was calculated separately for both the high LDL-C (>160 mg/dl) and the low LDL-C group and compared. The 95% CI was calculated using a z-score method with a critical value of 1.96. From these methods, an instantaneous hazard of all-cause mortality of 1.07 x 10-4 with a 95% CI of 3.52 x 10-5 to 0.0302 was obtained for the subjects with normal LDL-C. In comparison, the group with high LDL-C had an instantaneous hazard of 5.88 x 10-5 with a 95% CI of 1.04 x 10-5 to 0.063. Thus, the group with high LDL-C had a lower instantaneous risk of death when compared with the group with normal LDL-C levels.

8. *A priori*, I would have chosen to perform a two-sample two-sided t-test without the assumption of equal variance to test the alternative hypothesis that the mean LDL-C are different between those who died during the 5 year interval and those who did not. Continuous variables (not stratifying based on “high” or “low” HDL) offer greater statistical power to reject the null hypothesis. Note that if regression were an option to answer this question, I would have used regression methods to adjust for potential covariates that are also associated with the outcome of death during the 5-year interval (see the descriptive table).