Biomarkers of Inflammation and Mortality

Background

The Cardiovascular Health Study is a government sponsored cohort study of adults aged 65 years and older in four communities. The primary goals of the study were to observe the incidence of cardiovascular disease (especially heart attacks and congestive heart failure) and cerebrovascular disease (especially strokes) in the elderly over an 11 year period, and to relate the incidence of those diseases to various risk factors measured in the population on a regular basis. By such an observational study, greater insight into the natural history of chronic diseases in the elderly would be obtained. This is of particular importance, because there is increasing evidence that some of the associations observed between cardiovascular or cerebrovascular disease and various risk factors in middle aged adults are not observed in older adults. Possible mechanisms for such disparities include

- effects due to survivorship: those people particularly susceptible to getting diseases from specific risk factors have already died, and are thus not present in an older cohort, and
- effects due to increased risks from other diseases: the elderly might have increased risk of diseases that rarely affect middle aged adults and that participants are actually protected from those diseases by characteristics that are associated with increased risk of the diseases prevalent in middle aged people. An example of such a mechanism might be that the increased risk of infectious diseases in the elderly and the protective effect of greater energy reserves against infection combine to produce a tendency for greater weight to be associated with greater longevity in the elderly.

In this study, over 5,000 elderly, generally healthy (cancer was an exclusion criterion), adults were randomly selected from Medicare rolls in four communities. Agreement to participate was high, and thus the sample can be regarded as a fairly accurate representation of healthy older Americans. At the time of study enrollment, and on annual visits over the length of the study, the participants' data regarding various behavioral (e.g., smoking, alcohol consumption), functional (e.g., ability to perform routine tasks), and clinical (e.g., blood pressure, laboratory tests) measures are recorded. In addition, all serious medical events (e.g., hospitalizations, heart attacks) are investigated and categorized based on standardized, study-wide definitions.

For this project, we are interested in the role of inflammation in the pathogenesis of atherosclerotic disease. In particular, we are interested in two biochemical markers of inflammation, the C reactive protein and fibrinogen. We would like to assess whether patients with higher levels of these biomarkers are at increased risk of death in general and cardiovascular death in particular.

Questions of Interest

The data to be analyzed for this assignment is a subset of the thousands of variables on a subset of the thousands of participants in this study. The questions to be addressed are:

1. What associations exist between the inflammatory biomarkers and other known risk factors for cardiovascular disease?
2. Are the inflammatory biomarkers predictive of overall mortality and cardiovascular disease specific mortality either individually or in combination?

3. Does any association between the inflammatory biomarkers and mortality differ between shortterm (death within 3 years) and longterm (death after having survived at least three years)?
4. Does any predictive value of the biomarkers differ between sexes?
5. Do any associations found above persist after adjustment for known risk factors for cardiovascular disease?

Available Data

The data are stored in the file inflamm.txt in ASCII format. The first line of the file contains variable names, and each succeeding line corresponds to the observations on one of 5,000 participants. When data is missing for a particular variable, `NA' is recorded. The descriptions of the variables are as follows:

- *id*= Participant identification number
- *site*= Clinical site for participant (coded 1 – 4)
- *age*= Participant age at study enrollment (years)
- *maler*= Indicator of participant's sex (0= female,1=male)
- *bkrace*= Indicator that the participant's race is black (0 = no, 1= yes)
- *smoker*= Indicator that the participant smokes (0 = no, 1= yes)
- *estrogen*= Indicator of estrogen use (0 = no, 1= yes)
- *prevdis*= Indicator of prevalent atherosclerotic disease at study enrollment (0 = no, 1= previous angina, MI, TIA, stroke)
- *diab2*= Indicator of diagnosed diabetes at enrollment (0 = no, 1= yes)
- *bmi*= Body mass index (weight in kg divided by square of height in meters)
- *systBP*= Systolic blood pressure in mmHg. The systolic blood pressure is the maximum pressure generated during a contraction of the heart muscle. Persons with high blood pressure have been found to be at increased risk for heart disease, cerebrovascular disease, and kidney disease. The ``normal'' range for systolic blood pressure is 110 to 140 mmHg.
- *aai= The* ratio of systolic blood pressure measured in the participant's ankle at the time of MRI to the systolic blood pressure measured in the particpant's arm. Typically, we measure blood pressure in the arm. However, in patients with severe hardening of the arteries, the arteries in the legs may become so blocked as to restrict blood flow to the lower extremities. Thus, measuring the ankle blood pressure relative to the arm blood pressure is a marker of extent of arterial disease: A low ankle : arm index suggests more severe peripheral arterial disease. A person with no peripheral arterial disease might be expected to have aai=1.
- *cholest*= Serum cholesterol in mg/dl.
- *crp*= Blood C reactive protein in mg/l.
- *fib*= Blood fibrinogen in mg/dl.
- *ttodth*= The total time (in days) that the participant was observed on study between the date of study enrollment and death or data analysis, whichever came first.
- *death*= An indicator that the participant was observed to die while on study. If death=1, the number of days recorded in ttodth is the number of days between that participant's enrollment and his/her death. If death=0, the number of days recorded in ttodth is the number of days observed without a death.

- *cvddeath*= An indicator that the participant's observed death was due to cardiovascular disease.