Objective: Our goal is to identify confounding variables for the relationship between having a large medical expenditure and having a major smoking caused disease (MSCD)

Data: We will utilize data from the 1987 National Medical Expenditure Survey which contains the outcome, annual medical expenditures (binary, greater than $1000 vs. not) and exposure, presence of MSCD. In addition, the data has information on several potential confounding variables: age, marital status, poverty status, education, seat belt use, and smoking history.

Methods: Proportions were used to summarize the variables. A logistic regression model was used to estimate the unadjusted (marginal) association between having a large medical expenditure and MSCD. Adjusted (conditional) logistic regression models were constructed by adding each of the potential confounding variables one at a time. We identified confounding using the methods of Janes et al (Biostatistics, 2010). Specifically, we compared the adjusted odds ratios to the unadjusted odds ratios, as well as, the corresponding test statistics (estimated odds ratios divided standard error of the estimated odds ratios). In analyses where the adjusted odds ratio exceeded the unadjusted odds ratios but the test statistics were similar, the difference is attributed to the non-linearity behavior of logistic regression models and not to the presence of confounding. Therefore, variables were labeled as confounding when the unadjusted and adjusted odds ratios and the test statistics were different. NOTE: you may want to quantify "different".

Results:

Guidance: Start with descriptive statistics for the sample. This is important in that it allows the reader to get a snapshot of who is in the sample and the severity/degree of outcome and exposure. Next, following along with the methods section, present the findings from the unadjusted analysis. Follow-up with your results that support the findings for confounders.

Here are some suggested sentences from the 10:30am session:

<u>Descriptive analysis:</u>

11:30:03   From Scott Mu : In our data which included observations on __ individuals, major smoking caused disease occured in __ individuals and a large medical expenditure occurred in __ individuals. Of those with a major smoking caused disease, __ had a large medical expenditure.

11:30:05   From Elizabeth Colantuoni : The analysis included N persons, X% who had a large expenditure and X% with a MSCD. Further, the average age was X, ?% were married, ?% were considered to be living below the poverty status and X? were ever smokers.

11:30:51   From Annie Martin : The final analysis population was made up of xx individuals; y of whom were high school educated or higher, x of whom were married, z of whom reported consistent seat belt use, and c of whom where above the poverty line. The average medical expenditure was abc.

<u>Unadjusted analysis:</u>

11:30:14   From Stephanie Yan : Individuals who have a MSCD are X times more likely [95% CI] to have a large medical expenditure.

11:33:36   From Elizabeth Colantuoni : Individuals who have a MSCD have odds of a large expenditure that are 6.2 times greater than those without a MSCD (OR: , 95% CI: and Z statistic)

11:38:04   From Xiaobin Zhou : Among individual with same marital status, the odds of having big expenditure among people with MSCD is ?? (95% CI_____Chisq____) times the odds among people without MSCD, which is different from the odds ratio for all individual regardless of marital status.

Adjusted analysis

11:38:08   From Scott Mu : The unadjusted odds ratio was __(95% CI, Zstat) were compared to adjusted odds ratios. Educational status was identified as a positive confounder (OR: Zstat), as was __ and __.

11:39:00   From Elizabeth Colantuoni : Based on the adjusted analyses, we identified marital status as a confounding variable for the large expenditure and MSCD relationship (OR: , 95% CI: and Z statistic). There was no evidence in the data to suggest that the remaining variables were confounders (Poverty status: OR, z, seat belt use: OR, z, etc).

11:39:10   From David Johanson : Confounder analysis suggests that this variable may be a confounder in the relationship between smoking status and mscd status. The absolute difference in the odds ratio between mscd and smoker status in the adjusted and unadjusted models was x and the difference in z-statistics was z