Your Name: _____

Names of people you worked with: _____

Instructions: Work on this problem in class with your group. Do your best. This piece of paper will be collected during class.

Task: Consider the following study:

After World War II, evidence began mounting that there was a link between cigarette smoking and pulmonary carcinoma (lung cancer). In the 1950s, three now classic articles were published on the topic. One of these studies was conducted in the United States by Wynder and Graham (1950). They found records from a large number of patients with a specific type of lung cancer in hospitals in California, Colorado, Missouri, New Jersey, New York, Ohio, Pennsylvania, and Utah. Of those in the study, the researchers focused on 605 male patients with this form of lung cancer. Another 780 male hospital patients with similar age and economic distributions without this type of lung cancer were interviewed in St. Louis, Boston, Cleveland, and Hines, IL. Subjects (or family members) were interviewed to assess their smoking habits, occupation, education, etc. The table below classifies them as non-smoker or light smoker, or at least a moderate smoker.

Group	non/light smoker	moderate/heavy smoker	total
cancer	22	583	605
healthy	204	576	780
total	226	1159	1385

1. Assume that the explanatory variable is the smoking status (heavy smoking in numerator and light smoking in denominator) and the response variable is cancer status. Using the data above, calculate $\widehat{RR}, \widehat{OR}$. Let "success" be defined as having cancer.

2. Assume that the explanatory variable is the cancer status (cancer in numerator and healthy in denominator) and the response variable is smoking status. Using the data above, calculate $\widehat{RR}, \widehat{OR}$. Let "success" be defined as being a moderate/heavy smoker.

3. Which direction of variables do you want to conclude (circle one)?

   - explanatory variable is the smoking status and the response variable is cancer status?
   - explanatory variable is the cancer status and the response variable is smoking status?

4. Which direction of variables give the appropriate proportions / estimates (circle one)?

   - explanatory variable is the smoking status and the response variable is cancer status?
   - explanatory variable is the cancer status and the response variable is smoking status?

Solution:

1.

$$\widehat{RR} = \frac{583/1159}{22/226} = 5.17$$

$$\widehat{OR} = \frac{583/576}{22/204} = 9.39$$

2.

$$\widehat{RR} = \frac{583/605}{576/780} = 1.31$$

$$\widehat{OR} = \frac{583/22}{576/204} = 9.39$$

3. Which direction of variables do you want to conclude?

   - explanatory variable is the smoking status and the response variable is cancer status?

4. Which direction of variables give the appropriate proportions / estimates?

   - explanatory variable is the cancer status and the response variable is smoking status?