

# STA 6384, Report 2.15

Carson Slater *Baylor University*

**Problem:** Work Agresti's problem 2.12, p. 63.

A 20-year cohort study of British male physicians (R. Doll and R. Peto, *Br. Med. J.* 2: 1525–1536, 1976) noted that the proportion per year who died from lung cancer was 0.00140 for cigarette smokers and 0.00010 for nonsmokers. The proportion who died from coronary heart disease was 0.00669 for smokers and 0.00413 for nonsmokers.

(a)

**Describe the association of smoking with each of lung cancer and heart disease, using the difference of proportions, relative risk, and odds ratio. Interpret.**

Let  $P_S$  denote the proportion for smokers and  $P_{NS}$  denote the proportion for nonsmokers.

*Lung Cancer*

The given proportions are  $P_S = 0.00140$  and  $P_{NS} = 0.00010$ .

- **Difference of Proportions ( $\Delta P$ ):** This measures the absolute excess risk.

$$\Delta P = P_S - P_{NS} = 0.00140 - 0.00010 = 0.00130$$

**Interpretation:** The risk of dying from lung cancer is 0.00130 higher for smokers than for nonsmokers. This means that for every 10,000 smokers, there are 13 additional deaths from lung cancer per year compared to 10,000 nonsmokers.

- **Relative Risk (RR):** This measures how many times more likely the outcome is for smokers compared to nonsmokers.

$$RR = \frac{P_S}{P_{NS}} = \frac{0.00140}{0.00010} = 14.0$$

**Interpretation:** Smokers are **14 times** more likely to die from lung cancer than nonsmokers.

- **Odds Ratio (OR):** This compares the odds of the outcome for smokers versus nonsmokers. The odds are calculated as  $P/(1 - P)$ .

$$OR = \frac{P_S/(1 - P_S)}{P_{NS}/(1 - P_{NS})} = \frac{0.00140/0.99860}{0.00010/0.99990} \approx \frac{0.001402}{0.000100} \approx 14.02$$

**Interpretation:** The odds of a smoker dying from lung cancer are about **14.02 times** the odds for a nonsmoker. Since the disease is rare, the odds ratio provides a close approximation of the relative risk.

## Coronary Heart Disease

The given proportions are  $P_S = 0.00669$  and  $P_{NS} = 0.00413$ .

- **Difference of Proportions ( $\Delta P$ ):**

$$\Delta P = P_S - P_{NS} = 0.00669 - 0.00413 = 0.00256$$

**Interpretation:** The risk of dying from coronary heart disease is 0.00256 higher for smokers. This means that for every 10,000 smokers, there are approximately 26 additional deaths from heart disease per year compared to 10,000 nonsmokers.

- **Relative Risk (RR):**

$$RR = \frac{P_S}{P_{NS}} = \frac{0.00669}{0.00413} \approx 1.62$$

**Interpretation:** Smokers are about **1.62 times** more likely to die from coronary heart disease than nonsmokers.

- **Odds Ratio (OR):**

$$OR = \frac{P_S/(1 - P_S)}{P_{NS}/(1 - P_{NS})} = \frac{0.00669/0.99331}{0.00413/0.99587} \approx \frac{0.006735}{0.004147} \approx 1.624$$

**Interpretation:** The odds of a smoker dying from coronary heart disease are about **1.624 times** the odds for a nonsmoker.

(b)

**Which response is more strongly related to cigarette smoking, in terms of the reduction in number of deaths that would occur with elimination of cigarettes? Explain.**

In terms of the reduction in the number of deaths that would occur with the elimination of cigarettes, **coronary heart disease** is more strongly related to smoking.

The question asks for the impact on the *number of deaths*, which corresponds to the **absolute risk reduction**. The best measure for this from part (a) is the **difference of proportions** (also called the risk difference or attributable risk), as it quantifies the excess number of deaths attributable to smoking.

From the analysis in part (a):

- For **coronary heart disease**, the risk difference is:

$$\Delta P_{HD} = 0.00256$$

This indicates that eliminating smoking would prevent approximately **26 deaths per 10,000 smokers** each year.

- For **lung cancer**, the risk difference is:

$$\Delta P_{LC} = 0.00130$$

This indicates that eliminating smoking would prevent **13 deaths per 10,000 smokers** each year.

Since the excess risk due to smoking is greater for heart disease ( $0.00256 > 0.00130$ ), eliminating smoking would save more lives from coronary heart disease than from lung cancer in the population of smokers.

While smoking has a much higher *relative risk* for lung cancer ( $RR = 14.0$ ) than for heart disease ( $RR \approx 1.62$ ), the absolute public health burden is greater for heart disease. This is because the baseline risk of dying from heart disease is substantially higher than for lung cancer. Therefore, from a public health perspective aimed at reducing the total number of fatalities, the association with coronary heart disease is stronger.