Biostatistics

In-class exercise week 6: Solution
Topic: Study types

1) State the type of study described in each of the following situations.

a) To investigate the relationship between egg consumption and heart disease, a group of patients admitted to hospital with myocardial infarction were questioned about their egg consumption. A group patients admitted to a fracture clinic were also questioned about their egg consumption using an identical protocol.

This is a **case control study**, since the sampling is disease based (heart disease yes or no) and the information about the exposure (egg consumption) is retrospectively gathered.

b) To investigate the relationship between certain solvents and cancer, all employees at a factory were questioned about their exposure to an industrial solvent, and the amount and length of exposure measured. These subjects were regularly monitored, and after 10 years a copy of the death certificate for all those who had died was obtained and checked for cancer as cause of death.

This is a **cohort study**, since it is prospective and the sampling is exposure based (exposure to solvent) and then the subjects are followed over time and after 10 years information about diseases status (cancer or dead because of cancer) is gathered.

2) Relative Risk

Recall the definition of the risk ratio for disease (D) among individuals exposed to an risk factor or exposure (E) relative to unexposed individuals:

$$\text{risk ratio}: \quad RR = \frac{P(D \mid E)}{P(D \mid \bar{E})}$$

In a prospective cohort study, researchers reported that the risk of fatal coronary heart disease was increased for women with diagnosed diabetes compared with women without (risk ratio adjusted for age is 3.50 with a 95% confidence interval from 2.70 to 4.53).

a) Fill the missing number in the following cross-table (disease status vs exposure status) so that we get an estimated RR of 3.5.

	diseased D	not diseased $\bar{D}$
exposed E	**35**	**65**
not exposed $\bar{E}$	**10**	x =90

$$RR = \frac{P(D \mid E)}{P(D \mid \bar{E})} \overset{!}{=} 3.5$$

$$P(D \mid E) = 35/(35+65) = 0.35$$

$$P(D \mid \bar{E}) = 0.35/3.5 = 0.10$$

$$P(D \mid \bar{E}) = 10/(10+\mathrm{x}) = 0.10 \;\Leftrightarrow\; 10 = 0.1 \cdot 10 + 0.1 \cdot \mathrm{x}$$

$$x = P(\bar{D} \text{ and } E) = 90$$

b) Which of the following statements accurately describes the reported risk of fatal coronary heart disease?

a. ~~95% of women with diabetes have an increased risk of fatal coronary heart disease between 2.70 and 4.53 times that of women without diabetes~~
   ~~Because the confidence interval does not contain zero the result is statistically insignificant~~
- ✖ For women with diagnosed diabetes, the risk of having fatal coronary heart disease is 3.5 times that of women without diagnosed diabetes
- ✖ The reported risk of fatal coronary heart disease is statistically significant at the 5% level

d. ~~In women, diabetes causes fatal coronary heart disease~~