EMBLEM
High Quality, High Impact Research

Cancer Research Methods Workshop

Mayur M. Desai, PhD, MPH

Mwanza, Tanzania

September 18-19, 2012

Yale SCHOOL OF PUBLIC HEALTH

Measures of Association and Potential Impact

The "2 x 2" table

Disease

Exposure	+	-	
+	a	b	n_1
-	c	d	n_2
	m_1	m_2	N

Fixed in a cohort study

Fixed in a case-control study

Fixed in a cross-sectional study

Risk Ratio or Relative Risk in a cohort study

Disease

Exposure	+	-	
+	a	b	n_1
-	c	d	n_2

also called **Relative Risk** →

$$\text{Risk Ratio} = \frac{\text{incidence in the exposed}}{\text{incidence in the unexposed}} = \frac{a / n_1}{c / n_2}$$

Prevalence Ratio in a cross-sectional study

Disease

Exposure	+	-	
+	a	b	n_1
-	c	d	n_2

$$\text{Prevalence Ratio} = \frac{\text{prevalence in the exposed}}{\text{prevalence in the unexposed}} = \frac{a / n_1}{c / n_2}$$

Odds Ratio = ad/bc

Disease

Exposure	+	-	
+	a	b	n_1
-	c	d	n_2
	m_1	m_2	

Cohort (or cross-sectional) study

$$\text{Risk Odds Ratio} = \frac{(a/n_1) / (b/n_1)}{(c/n_2) / (d/n_2)} = \frac{(a / b)}{(c / d)} = \frac{ad}{bc}$$

(or Prevalence Odds Ratio)

Case-control study

$$\text{Exp Odds Ratio} = \frac{(a/m_1) / (c/m_1)}{(b/m_2) / (d/m_2)} = \frac{(a / c)}{(b / d)} = \frac{ad}{bc}$$

Example: association between sex & diabetes

$X^2 = 17.2$, DF = 1, $p < 0.001$

$$RR = \frac{648 / 7007 = 9.25}{907 / 8016 = 11.31} = 0.82$$

$$OR = \frac{648 * 7109}{6359 * 907} = 0.80$$

→ Lower risk or odds of diabetes in men compared with women

Table of sex by diabetes

sex	diabetes		
Frequency Percent Row Pct Col Pct	YES 1	NO 2	Total
MALE 1	648 4.31 9.25 41.67	6359 42.33 90.75 47.22	7007 46.64
FEMALE 2	907 6.04 11.31 58.33	7109 47.32 88.69 52.78	8016 53.36
Total	1555 10.35	13468 89.65	15023 100.00

Example 1

- Among 2,390 women aged 16 to 49 years who were free from bacteruria, 482 were OC users at the initial survey in 1973, while 1,908 were not. At a second survey in 1976, 27 of the OC users had developed bacteruria, as had 77 of the non-users. Identify the study design and calculate the measure of association and interpret it.

Example 1

		Bacteruria		
		Yes	No	Total
OC use	Yes	27	455	482
	No	77	1,831	1,908
	Total			2,390

Example 1

- Study design: ?
- Calculate RR:

$$RR = \frac{\text{Incidence in exp}}{\text{Incidence in unexp}} = \frac{27 / 482}{77 / 1{,}908} = 1.4$$

- **Interpretation:** Women who used oral contraceptives had 1.4 times (or 40%) higher risk of developing bacteruria over the follow-up period compared with non-users

Example 2

- Of 156 women with myocardial infarction (MI), 23 were current OC users at the time of their hospital admission. Of the 3,120 control women without MI, 304 were current OC users. Identify the study design and calculate the measure of association and interpret it.

Example 2

		MI		
		Yes	No	Total
Current OC use	Yes	23	304	
	No	133	2,816	
	Total	156	3,120	

Example 2

- Study design: ?
- Calculate OR:

$$OR = \frac{a * d}{b * c} = \frac{23 * 2{,}816}{304 * 133} = 1.6$$

- **Interpretation:** Odds of OC use were 1.6 times (or 60%) higher among MI cases than among controls – often phrased as 1.6 times higher risk associated with exposure

Consider the following two 2x2 tables…

Disease A

		+	-	
Exp	+	10	90	100
	-	1	99	100

Disease B

		+	-	
Exp	+	60	40	100
	-	30	70	100

$I_{exp} = 0.10$ $\quad$ $I_{exp} = 0.60$

$I_{unexp} = 0.01$ $\quad$ $I_{unexp} = 0.30$

RR = 10.0 $\quad$ RR = 2.0

The RRs indicate that the exposure is more strongly associated with Disease A

Attributable risk (in exposed individuals)

$$AR = I_{exp} - I_{unexp}$$

- Excess risk in exposed individuals that can be attributed to the exposure (i.e., # of cases among the exposed that could be eliminated if the exposure were eliminated)

Disease A

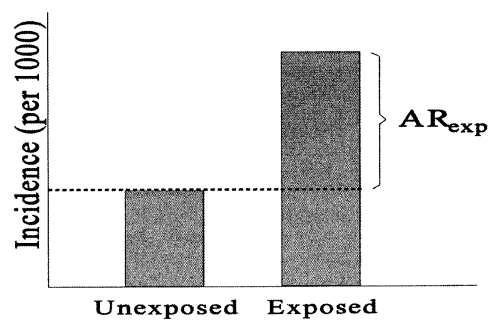
		+	-	
Exp	+	10	90	100
	-	1	99	100

Disease B

		+	-	
Exp	+	60	40	100
	-	30	70	100

$AR = 0.10 - 0.01 = 0.09$ $\quad$ $AR = 0.60 - 0.30 = 0.30$

Attributable risk in the exposed



Attributable risk percent (in exposed individuals)

$$AR\% = \left(\frac{I_{exp} - I_{unexp}}{I_{exp}}\right) \times 100$$

- Proportion of the disease among the exposed that is attributable to the exposure (i.e., could be prevented by eliminating the exposure)

Disease A

		+	-	
Exp	+	10	90	100
	-	1	99	100

Disease B

		+	-	
Exp	+	60	40	100
	-	30	70	100

$AR\% = \left(\frac{0.10 - 0.01}{0.10}\right) x\,100 = 90\%$ $\quad$ $AR\% = \left(\frac{0.60 - 0.30}{0.60}\right) x\,100 = 50\%$

Refer to Example 1

Calculate AR (risk difference):

$$AR = \frac{27}{482} - \frac{77}{1{,}908} = 0.0560 - 0.0404 = 0.0156$$

→ 156 excess cases per 10,000 OC users

Interpretation: The excess occurrence of bacteruria among OC users attributable to their OC use is 156 cases per 10,000 OC users

Refer to Example 1

Calculate AR%:

$$AR\% = \frac{(27 / 482) - (77 / 1{,}908)}{(27 / 482)} = \frac{0.0560 - 0.0404}{0.0560}$$

→ 0.2786 = 27.86%

Interpretation: About 28% of bacteruria cases among OC users is attributable to their OC use and could be eliminated if they did not use oral contraceptives

Relative risk vs. risk difference

- Relative risk
  - etiologic research, identifying risk factors (large RRs and ORs satisfy the "strength" criterion of causal association)
- Risk difference (absolute risk)
  - public health importance of exposures and diseases (assuming causal relationships)

Comparison of Relative Risk and Attributable Risk in Mortality from Lung Cancer and from Coronary Heart Disease for Heavy Smokers and Nonsmokers

	Annual Death Rates per 100,000 Persons	
Exposure Category	*Lung Cancer*	*Coronary: Heart Disease*
Heavy smokers	166	599
Nonsmokers	7	422
Measure of Excess Risk		
Relative risk:	166/7 = 23.7	599/422 = 1.4
Attributable risk:	166 – 7 = 159	599 – 422 = 177

*Doll, R., and Hill, A. B.: Lung cancer and other causes of death in relation to smoking. A second report on the mortality of British doctors. Br. Med. J., 2:1071, 1956.

AR% = 96% AR% = 30%

Exercise based on previous slide

- Which disease mortality is more strongly associated with cigarette smoking? Why?
- If the number of deaths attributable to smoking is used as an index of public health importance, which disease has more significance?

Population attributable risk

$$PAR = I_{pop} - I_{unexp} = AR * \text{prevalence of exp in the pop}$$

- Excess risk in the total study population that can be attributed to the exposure (i.e., # of cases in the pop that could be eliminated if the exposure were eliminated)

Population attributable risk

- Using data from the Doll and Hill study (earlier slide) and assuming the prevalence of smoking is 20% in the general population, calculate the PAR for lung cancer mortality associated with smoking:

PAR = AR * prevalence of exp in the general pop

= (159 per 100,000 per year) * 20%

= 31.8 per 100,000 per year

Interpretation of AR and PAR

- AR: In a population of 100,000 smokers, 159 deaths from lung cancer per year would have been avoided if they had not smoked
- PAR: In the general population of 100,000 persons, with a 20% prevalence of smoking, about 32 deaths from lung cancer per year would be prevented by eliminating smoking

Population attributable risk percent

$$\text{PAR\%} = \left(\frac{I_{pop} - I_{unexp}}{I_{pop}}\right) \times 100 = \left(\frac{p_{exp} \times (RR - 1)}{1 + p_{exp} \times (RR - 1)}\right) \times 100$$

- Proportion of the disease in the total study population that is attributable to the exposure (i.e., could be prevented by eliminating the exposure)