Measures of Disease Association

HSSP 100b

September 11 & September 18, 2015

What is an Association?

- In epidemiology, an association implies a relationship between an exposure and an outcome
  - Exposure can increase the likelihood of an outcome or decrease the likelihood of an outcome
- Exposure = any potential risk (or preventive) factor
- Outcome = disease occurrence (could also be death, some other outcome)

How Do We Measure an Association?

- We calculate a single measure of the magnitude (strength) of the association between the exposure and outcome
- To do so, we must first summarize the data (e.g., construct Row-by-Column [RxC] table)
- Then we calculate a summary measure of association appropriate for the study design
  - Relative measures of association (divide 2 measures of disease frequency)
  - Absolute measures of association (subtract 2 measures of disease frequency)

How Do We Measure an Association?

- Relative measures:
  - important measures of *the strength of an exposure-outcome association*
- Absolute measures:
  - tells us *how much of the disease/outcome can be attributed to a certain exposure (or the amount of disease we can hope to prevent if we are able to eliminate the exposure)*

Summarizing the Data: The 2x2 Table – The Basic Tool of The Epidemiologist

	Outcome		
Exposure status	**Disease**	**No Disease**	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

What does 'a' represent?
Those with exposure and disease
What does 'b' represent?
Those with exposure but no disease
What does 'c' represent?
Those with disease but no exposure
What does 'd' represent?
Those with neither exposure nor disease

What does 'a+c' represent?
all who have the outcome
What does 'b+d' represent?
all who do not have the outcome
What does 'a+b' represent?
all who have the exposure
What does 'c+d' represent?
all who do not have the exposure

Relative Measures of Disease Association

Relative Risk (RR)

- Risk Ratio: Cumulative Incidence Ratio (CIR)
  - CIe / CIo in cohort studies with count data
- Rate Ratio: Incidence Rate Ratio (IRR)
  - IRe / IRo in cohort studies with person-time data

Odds Ratio

- Odds Ratio* (OR):
  - OR = $\frac{ad}{bc}$ in case-control studies

*approximates risk when particular assumptions are met

…A Few Words on the Odds Ratio

- Compares the odds of exposure among those with the outcome to the odds of exposure among those without the outcome
- Is used when incidence cannot be calculated (e.g., case-control studies, cross-sectional studies, ecological studies)
- Generally, in case-control studies we can't calculate incidence since there is no rate of development of the disease.

Relative Risk / Odds Ratio

- Range of values: 0 to infinity
  - If RR, OR=1 then no association (null value)
  - If RR, OR>1 then exposure is associated with increased risk of disease
  - If RR, OR<1 then exposure is associated with decreased risk of disease

- Interpretation: Those who are exposed have X times the risk (odds) of developing the outcome as those who are non-exposed

Example 1: Cohort Study with Count Data (complete follow-up)

An investigator wants to know whether there is an association between texting while driving and car accidents. She enrolls a group of participants and asks them if they text while driving; 30 people report that they text while driving and 30 people report not texting while driving. She follows them for 1 year to determine whether they have been in a car accident. Among those who reported texting while driving, 10 were in a car accident. Among those who did not report texting while driving, 5 were in a car accident.

- What is the exposure?
- What is the outcome?
- What is the hypothesis to be tested?

  …remember the key elements one needs to include in a formal hypothesis:

  - What is the exposure?
  - What is the outcome?
  - Who is the comparison group?
  - What is the proposed direction of the relationship?

Risk Ratio (Cumulative Incidence Ratio)

	Outcome		
Exposure status	**Car Accident**	**No Car Accident**	Total
Texting while driving	10	20	30
No texting while driving	5	25	30
Total	15	45	60

$$\text{Risk Ratio} = \frac{\text{CIe}}{\text{CIo}} = \frac{10/30}{5/30} = 2$$

People who text while driving have 2 times the risk of being in a car accident compared to people who do not text while driving.

Example 2: Cohort Study with Person Time Data (incomplete follow-up)

An investigator wants to know whether women's age at the birth of their first child is related to risk of developing breast cancer. He enrolls women who have given birth into his study (who are free of breast cancer at baseline) and asks them how old they were when their first child was born and classifies them into two groups: 35+ vs. <35. He then follows them for 10 years to determine who develops breast cancer. The investigator observes 50 cases among women with their first birth at 35+ years of age in 26,423 person-years of follow-up, and 620 cases of breast cancer among women whose age at first birth <35 in 952,500 person-years of observation.

- What is the exposure?
- What is the outcome?
- What is the hypothesis to be tested?

Rate Ratio (Incidence Rate Ratio)

	Outcome		
Exposure status	**Breast Cancer**	**No Breast Cancer**	Total (person-yrs)
35+ at first birth	50	--	26,423
<35 at first birth	620	--	952,500
Total	670		

$$\text{Rate Ratio} = \frac{\text{IRe}}{\text{IRo}} = \frac{50/26{,}423 \text{ p-y}}{620/952{,}500 \text{ p-y}} = 2.91$$

Women who give birth to their first child at 35+ years of age have 2.91 times the risk of breast cancer compared to women who first give birth at <35 years of age.

Example 3: Case-Control Study

An investigator wants to know if there's an association between getting 8 or more hours of sleep and risk of obesity. He enrolls 60 obese people and 85 non-obese people into his study and asks them about their sleeping habits. Among the obese people, 10% usually get at least 8 hours of sleep each night. Among the non-obese people, 20% usually get at least 8 hours of sleep each night.

- What is the exposure?
- What is the outcome?
- What is the hypothesis to be tested?

Odds Ratio (OR)

	Outcome		
Exposure status	**Obese**	**Non-obese**	Total
8+ hours of sleep	6	17	23
<8 hours of sleep	54	68	122
Total	60	85	145

$$OR = \frac{\text{odds of exposure among diseased}}{\text{odds of exposure among non-diseased}} = \frac{a*d}{b*c}$$

$$OR = \frac{6*68}{17*54} = 0.44$$

Those who usually sleep 8+ hours a night have 0.44 times the odds of obesity compared to those who usually sleep <8 hours a night.

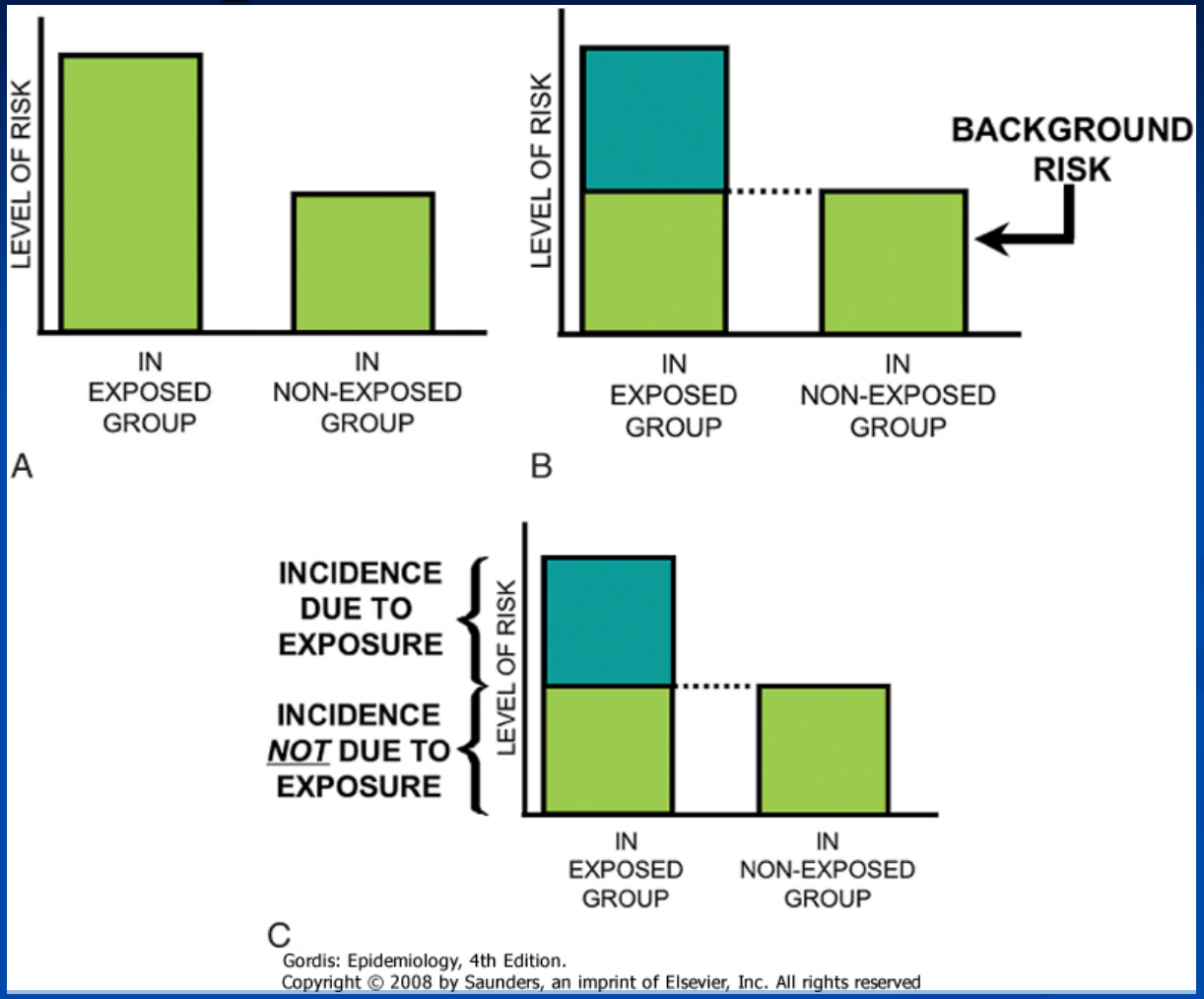
We can also say that the exposed have 56% lower odds compared to the unexposed since comparing to the null value of 1.

Absolute Measures of Disease Association

Measures of Absolute Risk: Attributable Risk for Exposed Group

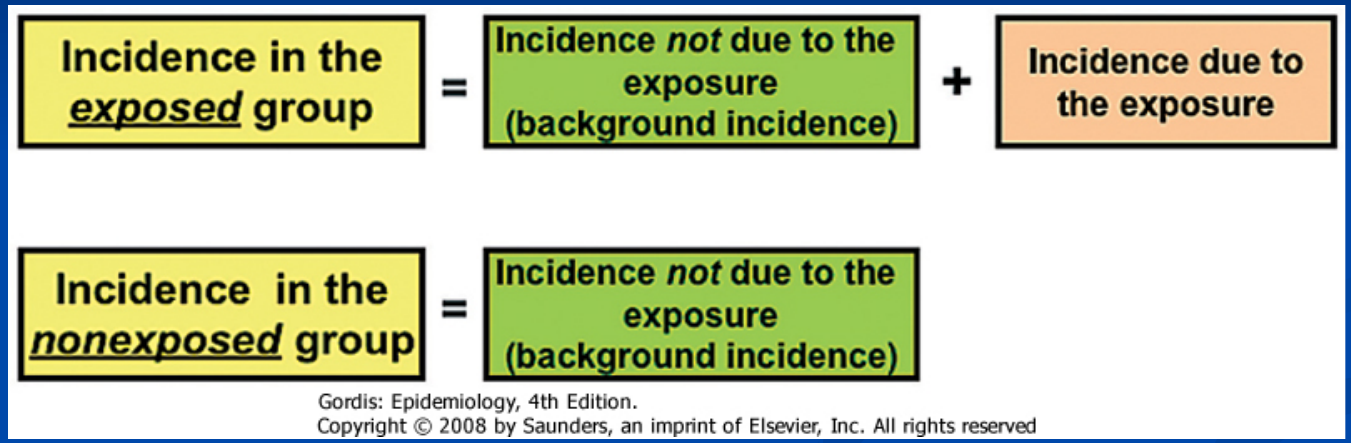
- Attributable risk among the exposed (ARe) tells us *how much of the disease/outcome can be attributed to a certain exposure (or the amount of disease we can hope to prevent if we are able to eliminate the exposure)*
  - Example: how much of lung cancer risk experienced by smokers can be attributed to smoking?
  - The calculation of ARe…
    - assumes causality has been established
    - depends on study design

Concept of Attributable Risk



A, Total risks in exposed and nonexposed groups. B, Background risk. C, Incidence attributable to exposure and incidence not attributable to exposure.

Concept of Attributable Risk (Cont.)



Gordis: Epidemiology, 4th Edition.

Absolute Risk Among the Exposed

- ARe
  - Risk difference: CIe - CIo in cohort studies with count data
    - Range of values: -1 to 1
  - Rate difference: IRe – IRo in cohort studies with person-time data
    - Range of values: $-\infty$ to $+\infty$
  - Null value = 0
  - Interpretation: Assuming exposure causes the outcome, X number of cases of the outcome among the exposed are due to the exposure, or could be avoided if the exposure was eliminated

Absolute Risk Among the Exposed

- ARe%
  - CIe – CIo / CIe in cohort studies with count data
  - IRe – IRo / IRe in cohort studies with person-time data
  - OR – 1 / OR in case-control studies

  - Interpretation: Assuming exposure causes the outcome, X% of the cases of the outcome among the exposed are due to the exposure, or could be avoided if the exposure was eliminated

Measures of Attributable Risk: Cohort Studies

- Risk difference (or attributable risk): studies with count data

$$CI_e - CI_o = \frac{a}{a+b} - \frac{c}{c+d}$$

- Rate difference (or attributable rate): studies with person-time data

$$IR_e - IR_o = \frac{a}{PT_e} - \frac{c}{PT_o}$$

Example 1: ARe in a Cohort Study

Environmental tobacco smoke (ETS) exposure	Myocardial infarction (MI) +	Myocardial infarction (MI) -	
+	120	19,880	20,000
-	100	22,400	22,500
	220	42,280	42,500

$AR_e = CI_e - CI_o$

$= (120/20{,}000)-(100/22{,}500)$

$= 15.6/100{,}000$

Assuming ETS exposure causes MI, 15.6 cases of MI among 100,000 people exposed to ETS are due to their exposure, or could be eliminated if ETS exposure were eliminated.

Measures of Attributable Risk: Cohort Studies

Attributable risk can also be expressed as a percent, rather than an absolute number. Addresses question: *"What proportion of the risk in exposed persons is due to the exposure?"*

Example:

$$AR_e\% = \frac{AR_e}{CI_e} = \frac{CI_e - CI_o}{CI_e}$$

Example 1: ARe% in a Cohort Study

Environmental tobacco smoke (ETS) exposure	Myocardial infarction (MI) +	Myocardial infarction (MI) -	
+	120	19,880	20,000
-	100	22,400	22,500
	220	42,280	42,500

$$AR_e\% = \frac{ARe}{CI_e} = \frac{CI_e - CI_o}{CI_e}$$

$$= \frac{(120/20{,}000)-(100/22{,}500)}{120/20{,}000}$$

$$= 25.9\%$$

If ETS exposure causes MI, 25.9% of MIs among those with ETS exposure is due to their exposure, or could be avoided if ETS exposure were eliminated

Example 2: ARe/ARe% in a Cohort Study

Age at first birth	Breast cancer: Yes	Breast cancer: No	
35+	50	—	26,423 person-yrs
<35	620	—	952,500 person-years
	670		

ARe = 50/26,423 py – 620/952,500 py

= 0.00189 – 0.00065

= 0.0012 or 12/10,000

$$\text{ARe\%} = \frac{0.0012}{0.00189} = 63.5\%$$

If older age at first birth causes breast cancer, for every 10,000 women whose first birth is over the age of 35 followed for one year, 12 cases of breast cancer are due to their age at first birth, or 63.5% of the breast cancer among these women could be avoided with an earlier age at first birth.

Measures of Attributable Risk: Case-Control Study

In a case-control study, ARe cannot generally be calculated since can't usually calculate incidence. But can calculate $AR_e\%$.

$$AR_e\% = \frac{OR-1}{OR}$$

Example: ARe% in a Case-Control Study

Cigarette smoking	Incident stroke +	Incident stroke -	
+	66	424	490
-	66	1,162	1,228
	132	1,586	1,718

Attributable risk % among the exposed (ARe %)

$$= \frac{OR-1}{OR}$$

$= 2.74-1/2.74 = 63.5\%$

If cigarette smoking use causes stroke, 63.5% of strokes among those who smoked cigarettes can be attributed to this exposure, or could be eliminated if they were to stop smoking.

Relative Risk vs. Absolute Risk

- Relative risk/Odds ratio:
  - Indicator of the strength of the association between exposure and disease
  - Used to assess importance of factor in the etiology of the disease

- Absolute risk:
  - Indicator of the absolute effect of an exposure on disease
  - Used when assessing public health impact of eliminating an exposure

- Give different but complementary information

Example: Relative Risk vs. Absolute Risk

Lung Cancer and CHD Mortality in Male British Physicians: Smokers vs. Nonsmokers					
	Age-adjusted mortality rates per 100,000				
	Smokers	Nonsmokers	Relative Risk	ARe	ARe%
Lung cancer	140	10	14.0	130	92.9
CHD	669	413	1.6	256	38.3
From Boll R, Peto R: Mortality in relation to smoking: twenty years' observation on male British doctors. Br Med J 2:1525-1536, 1976.					

Questions:

- Is smoking a stronger risk factor for lung cancer or CHD? What measure will tell us this?
- If smoking were eliminated, would we save more smokers from lung cancer or CHD? What measure will tell us this?

Example: Relative Risk vs. Absolute Risk (cont.)

Lessons:

- A modest RR may be important from a public health standpoint if the exposure and disease are common.
- A strong RR may be important from the standpoint of understanding a disease, but not compelling from a public health standpoint if the outcome is rare (especially weighing risks and benefits).

Sources

- As noted in slides
- Gordis text
- Oleckno, William A. *Epidemiology: Concepts and Methods*. Long Grove, IL: Waveland Press, Inc., 2008.
- Poole E. *HSSP 100b*, Brandeis, Fall 2010
- Buring J. *EPI200: Principles of Epidemiology Course Notes,* Harvard, Fall 2007