Measures of Effect

KP Fennie, PhD MPH

Objectives

- Compare and contrast absolute and relative effects
- Calculate and interpret risk difference, population risk difference, etiologic fraction and population etiologic fraction
- Compare and contrast type I and type II errors

Absolute vs. relative effects

- Absolute effects
  - Subtracted measures of disease frequency
  - Give us information about public health impact
  - Also called attributable risk
- Relative effects
  - Divided measures of disease frequency
  - Gives us information about strength of association between exposure and outcome

Rate difference

- Absolute effect
- Also called attributable risk
- Describes disease burden among exposed
- Difference between incidence rate of disease of exposed and unexposed groups
- RD = $I_e - I_{ne}$ where

  $$I = \frac{\text{\# new cases during a specified time period}}{\text{population at risk during same time period}} \times K$$

  (K = some multiplier)
- Helpful because risk of exposed is not just a function of exposure but also other factors:
  - Example: Heart disease among those with high cholesterol function of cholesterol but also genetics, tobacco use, etc. If you subtract risk of unexposed from risk of exposed, have a better understanding of cholesterol's contribution

Six Cities Study example of rate difference

- Difference between mortality incidence in most polluted city (Steubenville, Ohio) and least polluted city (Portage, Wisconsin)
- $I_{(\text{Steubenville})} - I_{(\text{Portage})}$
- 16.24 per 1000 py – 10.73 per 1000 py = 5.51 per 1000 py
- Interpretation: 5.51 excess deaths per 1000 person-years among Steubenville residents due to pollution

Risk difference

- Absolute effect
- Also called attributable risk
- Difference between cumulative incidence of exposed and unexposed
- Reminder cumulative incidence =

$$\frac{\text{\# new cases in pop. at risk for a specified time}}{\text{\# in population at risk at beginning of time period}} \times 100$$

Population risk difference

- Measure of benefit to population derived by modifying a risk factor
- How many cases in whole population can be attributed to a particular exposure
- $I_p = (I_e * P_e) + (I_{ne} * P_{ne})$
- Population risk difference = $I_p - I_{ne}$
- Also PRD = $RD \times P_e$
- Function of risk difference and proportion of population exposed

Population risk difference example

- Assume 10% of US pop exposed to pollution levels as high as in Steubenville
- Method 1: PRD = $RD \times P_e$
  - 5.51 per 1000 py * 0.1 = 0.55 per 1000 py
- Method 2: $I_p - I_{ne}$
  - $[(I_e * P_e) + (I_{ne} * P_{ne})] - I_{ne}$
  - [(16.24 per 1000 py * 0.1) + (10.73 per 1000 py* .9)] - 10.73 per 1000 py = 0.55 per 1000 py
- Interpretation: 0.55 deaths per 1000 py attributed to pollution in US.

Etiologic fraction

- Also called “attributable risk percent among exposed” and “attributable proportion among exposed” or “attributable fraction (exposed)”
- Relative effect
- Estimates percentage of cases among exposed due to exposure
- Problem with relative risk is that it is the risk between the exposed and non exposed, but non exposed also may have a high risk due to other factors
- Etiologic fraction = $(I_e - I_{ne})/I_e$ * 100

Etiologic fraction (cont.)

- Etiologic fraction = $(I_e - I_{ne})/I_e$
- Also because I_{ne} / I_e = 1/RR

  Etiologic fraction = (RR-1)/RR
- Six Cities example

(16.24 per 1000 py – 10.73 per 1000)/16.24 per 1000 py = 0.34 or 34%

34% of deaths among Steubenville participants may be attributed to pollution

Population etiologic fraction

- Proportion of rate of disease in population that is due to exposure
- Also called attributable proportion among total population or attributable fraction (population) or population attributable risk
- $(I_p - I_{ne})/I_p$
- Also $\frac{P_e\ (RR-1)}{\bar{P}_e\ (RR-1) + 1}$
- Six Cities example

Method 1: (11.28 per 1000 py – 10.73 per 1000)/11.28 per 1000 py = 4.9%

Method 2: [.1 (1.51-1)] / [[.1 (1.51-1)] + 1] = 4.9%

Interpretation: 4.9% of deaths among all US residents (assuming 10% exposed) may be attributed to pollution

Impact of disease on population

- Depends on
  - Strength of association between exposure and resulting disease
  - Overall incidence rate of disease in population
  - Prevalence of exposure

Absolute and relative measures of comparison

Type of measure	Formula	Interpretation
Rate or risk difference	$I_e - I_{ne}$	Excess rate of disease among exposed pop.
Population risk difference	$I_p - I_{ne}$	Excess rate of disease in total pop.
Etiologic fraction	$(I_e - I_{ne})/I_e$	Excess proportion of disease among exposed
Population etiologic fraction	$(I_p - I_{ne})/I_p$	Excess proportion of disease among total pop.
Relative risk	I_e / I_{ne}	Strength of association between exposure and disease

Adapted from Ashengrau Essentials of Epidemiology in Public Health, 2003.

Hypothesis testing

- Testing an assertion about a parameter in a population
  - Example: Cigarette smoking affects lung cancer
- Test to see whether data support (not prove) hypothesis
- Usually formulated as null hypothesis (Ho)
  - Example: There is no association between smoking and lung cancer

Hypothesis testing

- Hypothesis testing begins with assumption that null hypothesis is true
  - H_O = null hypothesis - no relationship
  - H_1 = research (alternative) hypothesis - there is a relationship
- Example: There is no association between smoking and lung cancer

p-value

- When inferring from sample phenomena to pop. phenomena, need to be fairly sure that what is observed in sample is not a function of mere chance
- It is possible to determine statistically - with considerable precision - if sample phenomena is attributable to chance
- p value = probability that the observed result is due to chance alone; level of significance
- p = .05: 5 chances in 100 that event occurred by chance

p = .10? p = .01? p = .001?

α vs p

- α = a predetermined choice for level of significance; most widely accepted standard for α is .05

  - *Ex*: α = .05 → "I am willing to risk a 5% chance of saying that there is a difference when - in truth - there is none"

- p = observed or attained level of significance; the actual probability of saying that there is a difference when - in truth - there is none

Statistical Significance

- If p value is < what α was set at (usually .05), then we say that the result is statistically significant
- If a result is statistically significant, it does not necessarily mean that it is clinically significant or clinically relevant
- With a large enough sample size, it is very likely that results will be statistically significant, but that does not mean that they will have practical value

Clinical vs. statistical significance

- Very small differences can be statistically significant if sample size large enough
- Example: 1 mm Hg difference in blood pressure (120 mm vs. 119 mm systolic) could be statistically significant if there was an extremely large sample size
- Prior to study
  - decide what a meaningful difference is
  - calculate sample size needed
  - design study so large enough sample can be recruited

- Statistical significance is a necessary precondition for clinical significance; if a difference is not statistically significant, it can’t be clinically significant, unless the sample size is too small & study lacks power

Hypothesis testing revisited

- Type I error: probability of rejecting null hypothesis when H_o is true (measured by *P*-value)
- Type II error: probability of failing to reject null hypothesis when it is false
- Power (1-β) probability of rejecting null hypothesis when it is false (our goal!)

Your decision

Truth

	H_o true	H_o false
Don't reject H_o (not stat. sig.)	Correct	β (type II error)
Reject H_o (stat. sig.)	α (type I error)	correct

Types of Error

- Type I Error: In reality there is no difference, but you conclude that there is.
  - alpha (α) = probability of making a Type I error; to avoid, lower the level of significance (e.g., .05 to .01)
  - rejecting a true H_O
- Type II Error: In reality there is a difference, but you conclude that there is not.
  - beta (β) = probability of making a Type II error; to avoid, raise the level of significance (e.g., .05 to .10)
  - accepting a false H_O

Relationship Between Type of Error

- As α is decreased, β is increased.
- As α is increased, β is decreased.

- Increase N to decrease probability of both Type I & Type II errors

Example: Error

- Test efficacy of new drug (Drug A) on pts w/ MI
- We hypothesize that mortality in patients taking Drug A will be lower than mortality in patients taking Drug B (old drug)

Consequences of each type of error

α

- We use Drug A because in the sample we found that mortality was improved, when in reality it doesn't make a difference
- Type I error
- *Consequence*: Pts won't benefit from the drug. Presuming that the drug is not harmful, we do not directly hurt pts, but since we think we have found a “cure”, we may no longer test other drugs.

Consequences of each type of error β

- We don't use Drug A because in the sample we found that it made no difference in mortality, when in reality it would have made a difference
- Type II error
- *Consequence*: By withholding the drug, pts may die who might otherwise have lived

Story

Once upon a time, there was a King who was very jealous of the Queen. He had 2 Knights, Alpha who was very handsome, & Beta, who was very ugly. It happened that the Queen was in love with Beta. The King, however, suspected the Queen was having an affair with Alpha & had him beheaded. Thus the King made both kinds of errors: he suspected a relationship (with Alpha) where there was none, & he failed to detect a relationship (with Beta) where there really was one. The Queen fled the kingdom with Beta & lived happily ever after, while the King suffered torments of guilt about his mistaken & fatal rejection of Alpha.

The End

Confidence interval

- Computed interval around a value

  e.g. OR 2.0 (95% CI 1.5-3.0)
- Indicates the amount of random error in estimate
- 95% confidence interval interpretation: CI contains “true” population estimate 95% of time
  - If study repeated 100 times, 95 times the CI will contain true estimate and 5 times it won’t
- Depends on data variability and sample size
- Wide CI indicates low precision, narrow indicates high precision
- If includes 1.0 not statistically significant

Confidence interval

- Depending on situation, confidence interval is calculated in various ways
- Based on standard deviation
  - 95% CI based on 1.96 units
  - 99% CI based on 2.58 units
- Related to p-value with respect the standard deviation
- Confidence Intervals are not always equidistant
  - RR 2.0 (95%CI 1.5, 2.5)
  - RR 2.0 (95%CI 1.2, 4.6)
  - RR 0.5 (95%CI 0.23, 0.62)
- Many in Epidemiology feel CI's should be used and not p-values

Interpretation of Confidence interval

- The width of the 95% CI indicates the range of variation for point estimates that may be expected by chance differences from one random sample to another.
- CI represents the range within which the true magnitude of effect lies, with a certain degree of assurance (e.g., 95%).
- You are 95% sure (*confident*) that the CI includes the population parameter; 95% of all 95% CIs do include the population parameter.

Assuming the same n, what happens to the width of the CI as the value of the confidence coefficient (AKA, confidence level) is increased (goes from 95% to 99%)?

A. The CI narrows

B. The CI widens

C. The CI stays the same

Odds ratios, *P* values, and 95% Confidence Intervals for a Case-Control Study with Three Different Sample Sizes

Parameter	n= 20	n= 50	n=500
Odds ratio	2.0	2.0	2.0
P	0.50	0.20	0.001
95% CIs	0.5, 7.7	0.9, 4.7	1.5, 2.6

HIV Infection and Associated Risks Among Young Men Who Have Sex with Men in a Florida Resort Community (Webster R, et al. JAIDS 2003;33:223-31)

- Objective: obtain population-based estimates of HIV prevalence and risk behaviors among young MSM in South Beach
- Methods:
  - Sampled based on residential site
  - Inclusion criteria, 18-29 yo unmarried males, had resided ≥ 30 days in South Beach, reported ever having sex with another man
  - Anonymous
  - Interview-administered and self-administered surveys
  - HIV (OraSure) test

Young MSM survey (continued)

- Results:
  - 2622 residential units screened between 1/20/96 and 12/19/96
  - 108 men met entry criteria
  - 92.6% consented

Unadjusted Odds Ratios for Selected Correlates of Oral HIV Antibody Test Results, Young Men Who Have Sex with Men, Miami Beach, 1996

Characteristic	HIV- (n=85) %	HIV+ (n=15) %	Unadjusted odds ratio
Currently has primary partner	44.7	20.0	0.31 (0.08-1.18)
< 2 years residency in South Beach	50.6	80.0	3.91 (1.03-14.84)**
# anal sex partners			
0-1	27.4	13.3	1.00
2-9	56.0	20.0	0.75 (0.12-4.70)
10 or more	16.7	66.7	8.21 (1.57-43.08)**
"Cruised" for sex monthly	41.2	80.0	5.71 (1.50-21.73)**
Belief that stopping to put on condom takes fun out	27.1	60.0	4.04 (1.30-12.62)**

Source: Webster, RD, et al. *JAIDS* 2003; 33:223-31. ***P*<0.005

Which measure would I calculate?

- What percentage of myocardial infarctions (heart attacks) would be prevented among people in the United States if there was no tobacco use?
- How many myocardial infarctions among tobacco smokers would be prevented if the smokers didn't smoke?
- How many myocardial infarctions in the United States would we prevent if there was no tobacco use?
- What percentage of myocardial infarctions among smokers are due to smoking?