

# Session Objectives

1. Students will be able to apply basic epidemiological concepts to specific disease areas as described in lecture when presented with a topic following that lecture
2. Students will be able to calculate basic epidemiological measures for a given circumstance as described in lecture following that lecture
3. Students will be able to interpret basic epidemiological results as described in lecture following that lecture
4. Students will be able to evaluate an epidemiological study for statistical significance as described in lecture following that lecture
5. Students will be able to describe an epidemiological study from beginning to end as described in lecture following that lecture

# Case

When you were a child, your mother constantly told you that you “should not go out in the winter without a hat as you will catch a cold.” Now, as a medical professional, you are ready to prove her wrong. You decide to dedicate your life to investigating this issue.

**NEW YORK INSTITUTE  
OF TECHNOLOGY**

College of Osteopathic  
Medicine



# Case-Control Studies

- In a case-control study with two groups, one group has the disease of interest (cases) and a comparable group is free from the disease (controls).
- The case-control study identifies possible causes of disease by finding out how the two groups differ with respect to exposure to some factor.

# Selecting Controls

## Advantages of general population controls

- Because of selection process, investigator is usually assured that they come from the same base population as the cases.

## Disadvantages of general population controls

- Time consuming, expensive, hard to contact and get cooperation; may remember exposures differently than cases

# Selecting Controls

What illnesses make good hospital controls?

- Those illnesses that have no relation to the risk factor(s) under study

Example: Should respiratory diseases be used as controls for a study of smoking and myocardial infarction? Do they represent the distribution of smoking in the entire population that gave rise to the cases of MI?

- Special control groups like friends, spouses, siblings, and deceased individuals.

# Matching in Epidemiology

## Types

- Individual matching
- Frequency matching

Data are analyzed in terms of case-control pairs rather than for individual subjects

# Matching in Epidemiology

- Can be time efficient, cost effective, and improve statistical power
- The more variables that are chosen as matching characteristics, the more difficult it is to find a suitable control to match to the case
- Once a variable is used for matching, no relationship can be discerned between this variable and the disease
- Don't match on anything you think might be a risk factor!

## Odds Ratio: Calculated in Case-Control Studies

---

Odds Ratio=

Odds of exposure in diseased

Odds of exposure in healthy

*or*

Odds of disease in exposed

Odds of disease in unexposed



# Analysis of case-control studies

---

Exposure Status	Disease Status		ODDS
	Yes	No	
Yes	A	B	A/B
No	C	D	C/D

Odds Ratio= $[A/B]/[C/D] = [A \times D]/[B \times C]$

Or  $[A/C]/[B/D] = [A \times D]/[B \times C]$

# Analysis of case-control studies

---

EXAMPLE: Case control study of Parkinson's Disease (PD) and coffee consumption

		Case (Has PD)	Control (healthy)
Coffee Drinker	Yes	42	247
No Coffee	No	107	825

$$\text{Odds ratio} = [(a/b) / (c/d)] = [(42/247) / (107/825)] = 1.3$$

“A PD patient is 1.3 times as likely as a healthy individual to have been a coffee drinker”

# Statistical Significance

- P is short for probability: the probability of getting something more extreme than your result, when there is no effect in the population.
- P-values below 0.05 are often considered significant
- *Usually report a 95% CI*
- You can be 95% confident you will get a number between your reported range if experiment is repeated
- ▶ IF your CI includes 1, then you automatically will NOT have statistically significant results (this will be supported by a p-value)

## Strengths of case-control studies

- Efficient for rare diseases and diseases with long induction and latent period.
- Can evaluate many risk factors for the same disease so good for diseases about which little is known

# Weaknesses of case-control studies

- Inefficient for rare exposures
- Vulnerable to bias because of retrospective nature of study
- May have poor information on exposure because retrospective
- Difficult to infer temporal relationship between exposure and disease

(Gordis, Epidemiology) 7th Edition by Leon Gordis MD MPH DrPH, Chapter 7

You decide that the  
case-control study is not  
the optimal way of  
conducting your study,  
and instead opt to  
conduct a cohort study.

# COHORT STUDIES

## NEW YORK INSTITUTE OF TECHNOLOGY

College of Osteopathic  
Medicine

- Definition: A study in which two or more groups of people that are free of disease and that differ according to the extent of exposure (e.g. exposed and unexposed) are compared with respect to disease incidence. Can be prospective or retrospective.
- Cohort studies are the observational equivalent of experimental studies but the researcher cannot allocate exposure -he must locate a natural experiment to observe the relationship between the exposure and disease

# Weaknesses of Cohort Studies

- Inefficient for rare outcomes
- If retrospective, poor information on exposure and other key variables, more vulnerable to bias
- If prospective, expensive and time consuming, inefficient for diseases with long induction and latent period
- Keep these strengths and weaknesses in mind for comparison with case-control studies



## Relative Risk: Calculated in a Cohort Study

Relative risk =

$$\frac{\text{Incidence rate (risk) in the exposed}}{\text{Incidence rate (risk) in the non-exposed}}$$

- Relative Risk = 1.0      No Risk

(meaning the rate of poor outcome among those with risk factor is the same as those without)

- Relative Risk > 1      Elevated Risk

(meaning the rate of poor outcome is more common among those with the risk factor present)

- Relative Risk < 1      Reduced Risk

(meaning the rate of poor outcome is less common among those with the risk factor present)

You decide that you would really like to conduct a randomized controlled trial, since this is the gold standard in epidemiology studies. You spend all of your time, money and effort on creating a controlled “winter chamber” and recruit 400 people to live under a controlled setting where you will randomize people to the “hat” or “no hat” groups.

Analysis is the same as a cohort study!

# Use of Placebo and Blinding

Placebos are used to make the groups as comparable as possible (recall laboratory experiment)

Blinding: subjects do not know if they are receiving treatment or placebo (single blind); neither subjects nor investigators know who is receiving treatment or placebo (double blind).

Purpose of blinding: To avoid bias in ascertainment of outcome

Placebo allows study to be blind

# Number Needed to Treat (NNT)

Number of persons who would have to receive an intervention for 1 to benefit.

$$\text{NNT} = 1 / (\text{rate in untreated group} - \text{rate in treated group}) = 1 / \text{Absolute risk reduction}$$

- To determine absolute risk reduction subtract incidence with treatment from incidence without treatment :  $0.0043 - 0.0028 = 0.0015 = \text{Absolute Risk Reduction}$
- The number needed to treat is the inverse of the Absolute risk reduction:  
 $1/0.0015=667$
- This means that if 667 individuals are exposed to the risk factor, 1 case will be prevented (NNT).
- Number Needed to Harm (NNH) is similar but is  $1/\text{attributable risk}$  when you are looking at the harmful side effects of a drug.

# Confounding

- An alternate explanation for observed association between an exposure and disease.
- What if you see that those who wear hats get fewer colds. BUT, then you find out the same people who wear hats wear gloves and scarves. What if it is really the scarf that protects against colds?
- This is why you collect data on variables other than the ones you are studying- namely, confounders.
- Common confounders are age, gender and race.
- Can address by stratification OR with more complicated regression analysis

# Stratified Analysis-Confounding

Composite OR=2.15			
Female		Male	
Smoking	Yes	Case 5	Control 8
	No	Case 45	Control 72
Stratum-specific OR = 1		Stratum-specific OR = 1	

Smoking	Yes	Case 25	Control 10
	No	Case 25	Control 10
Stratum-specific OR = 1		Stratum-specific OR = 1	

- In this case control study of smoking and risk of heart attack. Gender is a confounder.

## Interaction (Effect Modification)

- When an exposure behaves differently in different groups.
- For example, if you find smoking increases the risk for heart attacks much more in males than females (these are fictional numbers and not based on fact!)





How would you address  
interaction and/or confounding  
in an analysis?

**Stratification!**

Both issues, though very  
different from one another,  
have the same solution

# Bias

- Bias does not mean that the investigator is “prejudiced.”
- Bias can arise in all study types: experimental, cohort, case-control
- Bias occurs in the design and conduct of a study. It can be evaluated but not fixed in the analysis phase.

# Recall bias

People with disease remember or report exposures differently (more or less accurately) than those without disease.

Can result in over- or under-estimate of measure of association. (see following slide)

Solutions: Use controls who are themselves sick; use standardized questionnaires that obtain complete information, mask subjects to study hypothesis

This concludes this  
epidemiology session. Please  
make certain you understand  
all of the listed objectives.

# Lecture and Lab Feedback Form:

**NEW YORK INSTITUTE  
OF TECHNOLOGY**

College of Osteopathic  
Medicine

<https://comresearchdata.nyit.edu/redcap/surveys/?s=HRCY448FWYXREL4R>