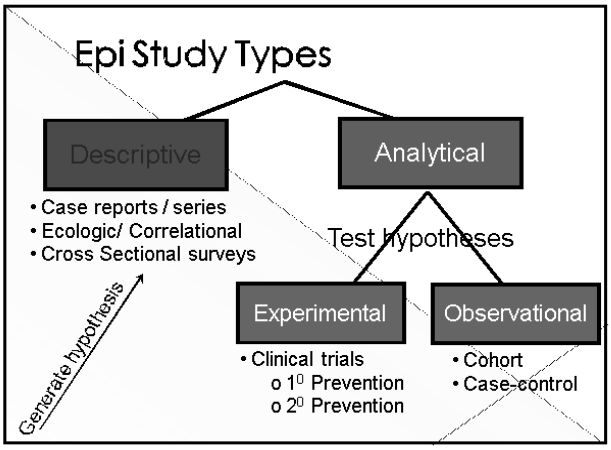
Study Design: Experimental

Introduction

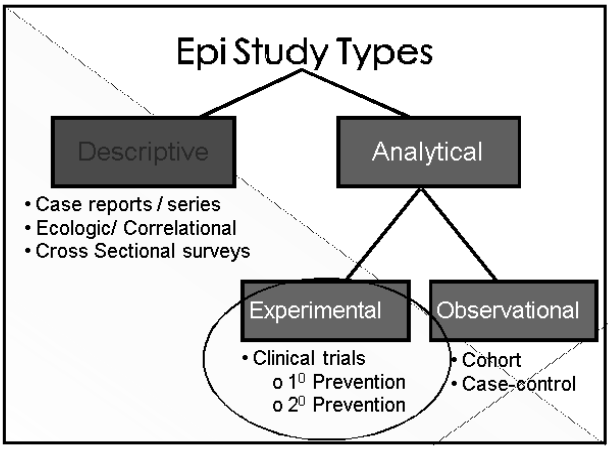
Learning Objectives

- Recognize characteristics of epi study designs
- Recognize when each design is appropriate
- Understand strengths & limitations of each study design
- Understand how data from each type contributes to assessing causality

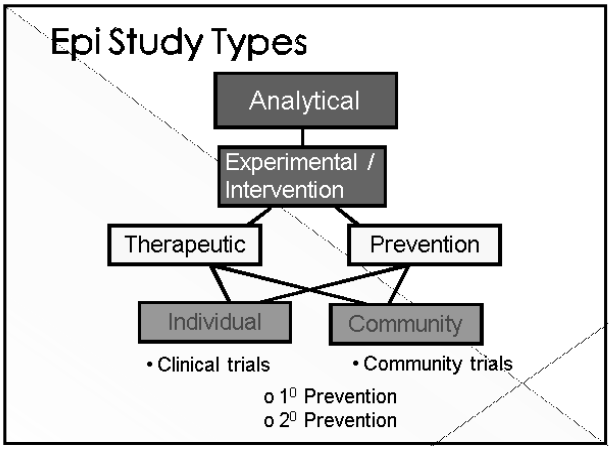
Epi Study Types



Epi Study Types



Epi Study Types



Therapy vs Prevention

- Therapy
  - Already diseased individual
  - Intervention to:
    - ameliorate symptoms
    - improve survival, quality of life
    - reduce risk of recurrence
- Prevention
  - Healthy individual – may be high risk
  - Intervention to prevent disease

Clinical Trials

- Most closely resemble laboratory experiments
- Investigator assigns exposure
- "Gold standard"
- May have greater validity

Study Design: Experimental Procedures

Sub topic learning objective:

Learn about the conduct of experimental studies

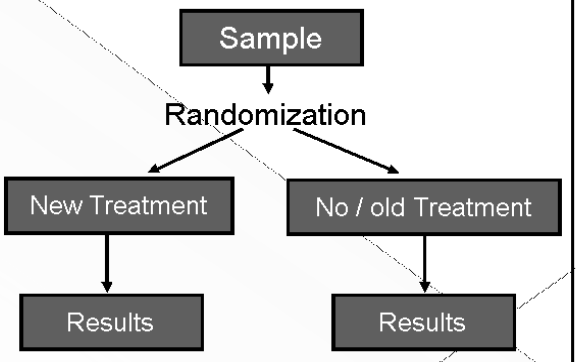
Clinical Trial Procedures

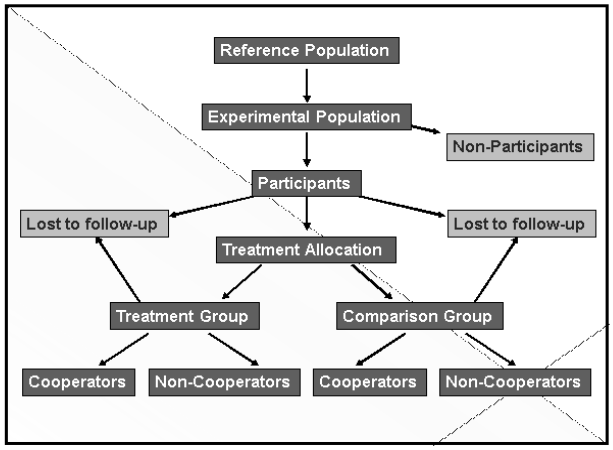
- State Hypothesis
  Treatment A will produce better outcomes than Treatment B
- Recruitment
  - Determine study population
  - Inclusion criteria
  - Exclusion criteria
  - Informed consent procedure
  - Apply for IRB approval
- Random allocation to interventions being compared

Clinical Trial Procedures

- Monitor for outcome under study
  - Disease occurrence
  - Disease reoccurrence
  - Symptom improvement / cure
  - Side effects
- Compare rates of outcomes in each group

Clinical Trial





Selection of Study Population

- Reference population
  - Group to whom results should apply
  - Examples:
    - Men ages 25 – 45
    - Hispanics aged 65+
    - Children < 16 years of age

Selection of Study Population

- Considerations
  - Accessibility of subjects
  - Likelihood of compliance with study protocol
  - Feasibility of follow-up
  - Is study population representative of the reference population?

Selection of Study Population

- Sample Size
  - How many endpoints are expected
  - Subgroup analysis
  - Length of follow up
  - If no difference detected:
    - Because no true difference OR
    - Too few patients to detect a difference

Selection of Study Population

- Eligibility criteria
  - Examples:
    - No prior treatment
    - Incident cases only
    - No other health conditions…

Selection of Study Population

- Eligibility criteria
  - Examples:
    - No prior treatment
    - Incident cases only
    - No other health conditions…

Placebos

- What is it?
  - Sham treatment
    - Pharmacologically inactive treatment
    - Indistinguishable from test treatment
- Goals
  - Used to make the groups as comparable as possible
- Allows study to be masked

Placebo - Example

- Acupuncture treatment of depression
  - Symptoms map to points for acupuncture
    - 1st practitioner evaluates symptoms & determines the correct needle placements
    - 2nd practitioner performs the acupuncture
    - Patients randomly assigned to receive acupuncture that corresponds to their symptoms or acupuncture in some set of points that does not correspond to symptoms
- Result:
  - Acupuncture (in the correct locations) performed about as well as drugs, but with fewer side effects

Study Design: Experimental
Bias & Analysis

Sub topic learning objective:

- Understand the basics of the sources of potential bias in experimental study designs and the analysis of these designs

Compliance

- Active participation / cooperation of participants
  - *Required*
- Deviations from the protocol will occur
  - Side effects
  - Illness
  - Level of interest
  - Length of follow-up

Non-Compliance

- Noncompliance makes the compared groups more alike
  - reduces the ability to detect a difference between the groups (diminishes study power)

Non-Compliance

- What can you do about it?
- Compliance enhancement strategies
  - Design phase
    - interested group
    - simple protocol
    - run-in periods
    - cross over
  - During study
    - frequent contact with subjects
    - incentives to continue, such as free check-ups
  - Measuring compliance

Outcome Determination

- Non-Uniform ascertainment of outcome is BIAS
- Follow-up: don't lose people
  - equally vigilant follow-up in all compared groups

Clinical Trials

- Analysis by intention to treat:

"Once randomized, always analyzed"

Clinical Trial – Measure of Association

- Relative Rate

$$RR = \frac{Rate_{treatment}}{Rate_{placebo}}$$

Clinical Trial Study Data

ID	Treatment	Outcome
1	A	Y
2	B	Y
3	A	N
4	B	N
etc.	…	…
n	A	Y

2 X 2 Tables

		Outcome: Yes	Outcome: No	TOTAL
Group	Treatment	*a*	*b*	*a + b*
	Placebo	*c*	*d*	*c + d*
	TOTAL	*a + c*	*b + d*	*a + b + c + d*

$$Rate_{treatment} =$$

2 X 2 Tables

		Outcome: Yes	Outcome: No	TOTAL
Group	Treatment	*a*	*b*	*a + b*
	Placebo	*c*	*d*	*c + d*
	TOTAL	*a + c*	*b + d*	*a + b + c + d*

$$Rate_{treatment} = \frac{a}{a+b}$$

2 X 2 Tables

		Outcome: Yes	Outcome: No	TOTAL
Group	Treatment	*a*	*b*	*a + b*
	Placebo	*c*	*d*	*c + d*
	TOTAL	*a + c*	*b + d*	*a + b + c + d*

$$Rate_{placebo} = \frac{c}{c+d}$$

Clinical Trials - Measure of Association

		Outcome: Yes	Outcome: No	TOTAL
Group	Treatment	a	b	$a+b$
	Placebo	c	d	$c+d$
	TOTAL	$a+c$	$b+d$	$a+b+c+d$

$$RR = \frac{Rate_{treatment}}{Rate_{placebo}} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{a(c+d)}{c(a+b)}$$

Study Design: Experimental
Ethics & Summary

Sub topic learning objective:

Understand the ethical considerations in the conduct of experimental studies

Clinical Trials - Ethics

- Equipoise
  - Must be genuine doubt about efficacy of treatment yet sufficient belief that it may work

Ethical Considerations

- Is the treatment safe?
- Can treatment ethically be withheld?
- Is a placebo group ethical?
- Which patients qualify?
  - Exclusions
  - Inclusions
  - "Sensitive" populations
- Is it ethically acceptable to "mask" the trial?

Clinical Trials - Ethics

- Stopping rules
  - At what point is there sufficient evidence of:
    - Benefits
    - Harm

  to end the trial early?
- How will you know if trial is masked?

Randomized Clinical Trials Advantages

- Randomization balances other variables between treatment groups
- Baseline data collected
- Temporal relationship
- Masking minimizes bias
- Statistical assumptions met
- Control of amount & timing of exposure

Randomized Clinical Trials Caveats

- Exclusions limit generalizability
- Large numbers of participants
- Expensive
- Ethics
- Compliance

Design Characteristics

1. # of observations
2. Data collection methods / source of data
3. Timing of data collection
4. Unit of observation
5. Availability of subjects
6. Method of defining "*study population*"
7. Measure of association

Clinical Trial Example

Puoti, M. et al. 2004 A randomized, controlled trial of triple antiviral therapy as initial treatment of chronic hepatitis C in HIV-infected patients. J Hepatol 41(2):312-8.

Selection of Subjects

- April 2000 – October 2001
- n=80
- Inclusion Criteria:
  1. 18 – 60 years old
  2. ALT levels above upper limit of normal @ 6 months prior to enrollment (↑ levels indicate liver damage)
  3. HIV seropositivity by ELISA confirmed by Western blot
  4. CD4 count persistently > 300/µl over prior 8 months
  5. Anti-retroviral treatment for ≥ 3 months
  6. Willingness to abstain from alcohol for the duration of study

Clinical Trial Example

- Selection of Subjects
  - Exclusions
    - Reactivity for Hepatitis B surface antigen
    - Neutropenia (< 1500 neutrophils / µl)
    - Anemia (< 12 g/dl Hb < 13 g/dl Hb)
    - ....
    - Pregnancy / lactation
    - Unwillingness to practice contraception during study
- Randomization (Treatment A or B)
  - Stratified on HCV genotype
- Masking - none
- Placebos - none

Clinical Trial Example

- Compliance
  - 25 / 80 (31.3%) stopped treatment early
    - 10 due to adverse effects
    - 15 other reasons
  - No difference between treatment groups

Clinical Trial Example

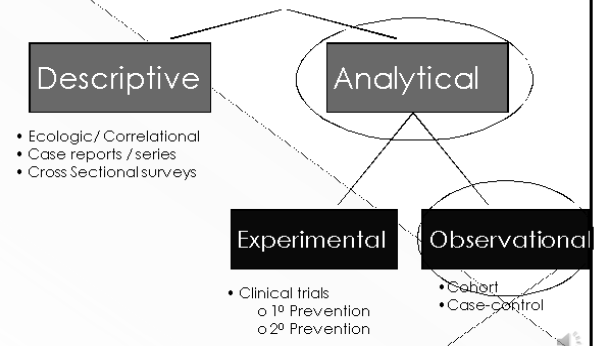
Table 1. Baseline Characteristics of Study Population

	Total	Group A	Group B
Number	80	41	39
Sex [% male (n)]	76.3 (61)	80.5 (33)	71.8 (28)
Age	37 ± 5	38 ± 5	37 ± 5
BMI	23.8 ± 3	23.6 ± 2.8	24.1 ± 3.2
Estimated HIV infection duration [yrs, range]	11 (10 – 13)	12 (9 – 14)	11 (9 – 13)
Estimated HCV infection duration [yrs, range]	5 (4 - 7)	5 (4 - 8)	5 (3 - 7)

Cohort Studies
Introduction

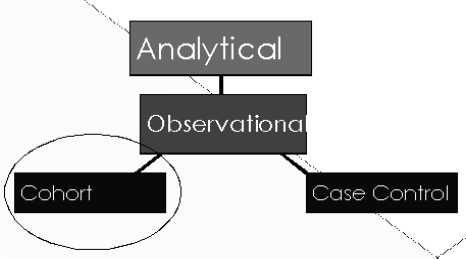
43

Epi Study Types



44

Epi Study Types



45

Cohort - Definition

1. **group of people:** a united group of people
2. **supporter:** a supporter, accomplice, or associate of a leader
3. statistics **group with statistical similarities:** a group of people sharing a common factor
4. history **unit of Roman army:** an ancient Roman military unit that formed one tenth of a legion and that consisted of 300 to 600 men
5. **soldiers:** a group of soldiers or warriors

[15th century. From, ultimately, the Latin stem *cohort-* , literally "enclosure," thus "people within an enclosure," hence "company of infantry," variant of *cort-* .]

http://encarta.msn.com/encnet/features/dictionary/DictionaryResults.aspx?search=cohort

Cohort - Definition

"A group of people that is defined in some way"

- S. Pettygrove

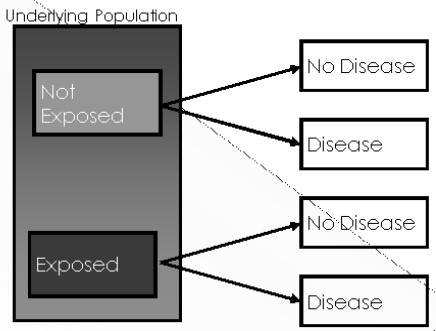
47

Cohort Study

- Two or more defined groups of people
  - free of disease at outset
  - differ in extent of exposure (e.g. exposed and unexposed)
- Followed to determine disease outcome
- Disease incidence is compared across exposure groups

48

Cohort Study



49

Cohort Studies

- Goal
  - complete, comparable, unbiased information on health experience of every study subject

50

Cohort Studies

- Cohorts
  - Occupational groups
  - Groups undergoing particular medical treatment
  - Groups with unusual dietary or life style factors
  - Professional groups (nurses, doctors)
  - Students or alumni of colleges
  - Geographically defined areas (e.g. Framingham)

51

Study Design: Cohort
Procedures

52

Learning Objectives

- Learn about the **conduct** of **cohort** studies

53

Selection of Cohorts

- Selection from
  - General populations
  - Special exposure populations
    - Occupational groups
    - Groups undergoing particular medical treatment
    - Groups with unusual dietary or life style factors
  - Higher risk than general population
  - Higher probability of obtaining complete exposure histories
    - Professional groups (nurses, doctors)
    - Students or alumni of colleges
    - Geographically defined areas (e.g. Framingham)

54

Cohort Studies

- Rare exposure
  - Special exposure populations
    - Occupational groups
    - Groups undergoing particular medical treatment
    - Groups with unusual dietary or life style factors
- Example: Rubber workers in Akron, Ohio
  - Exposure: industrial solvent
  - Outcomes: cancer

55

Cohort Studies

- Common exposure
  - cohort to facilitate accurate and complete ascertainment of data
    - Doctors
    - Nurses
    - well-defined communities

56

Cohort Study – Example 1

Framingham Study

- Exposures
  - Smoking
  - Hypertension
  - family history
- Outcomes
  - heart disease
  - Stroke
  - Gout
  - etc.

57

Cohort Example 2

- Ranch Hand Study
  - Exposed
    - 1,264 Air Force service personnel
    - sprayed agent orange during Vietnam War, 1962-1971
  - Unexposed
    - 1,264 Air Force service personnel
    - flew other missions during Vietnam War
  - Outcomes
    - cancer, post traumatic stress, adverse pregnancy outcomes etc

58

Cohort Example 2

- If Agent Orange is not associated with the outcomes,

$$Risk_{exposed} = Risk_{unexposed}$$

59

Comparison Group

- As similar as possible to the exposed group
  - except for exposure
- If exposure has no effect on disease, then
  - Rate of disease in the exposed = Rate of disease in the unexposed

60

Comparison Group

- Internal comparison
  - unexposed members of same cohort
    - Framingham study
    - Ranch Hand study

61

Comparison Group

- Comparison cohort
  - Unexposed cohort from a different, but similar population
  - Ex: Asbestos textile vs. cotton textile workers

62

Comparison Group

- General population data
  - Pre-existing data from the general population for comparison
  - Commonly used in occupational studies, but healthy worker effect
- Ex. A study of asbestos and lung cancer with U.S. male population as the comparison group

63

Healthy Worker Effect

- People who are unhealthy may not participate in the workforce
  - Therefore, the health outcomes of people employed may be expected to be better than those of the general population

64

Exposure Data Sources

- Pre-existing records
  - Occupational records
  - Medical records
  - Other records
  - Advantages:
    - Inexpensive
    - Data recorded before disease occurrence
  - Caveats:
    - Level of detail may be inadequate
    - Missing records
    - Information on confounders?

65

Exposure Data

- Questionnaires / Interviews
  - Advantage
    - Can obtain information not routinely recorded
  - Caveat
    - recall bias

66

Exposure Data

- Direct physical exams / tests
- Environmental monitoring
- Same for both:
  - Advantage
    - Can obtain information not routinely recorded
  - Caveat
    - Relation of today's findings to pre-disease condition

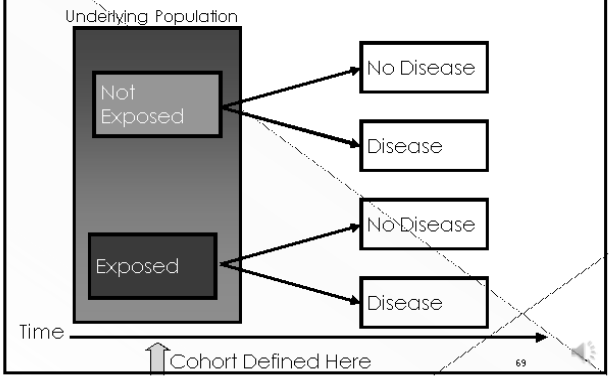
67

Outcome Data

- Death certificates
- Physician, hospital, health plan records
- Questionnaires (verify by records)
  - But, what if subject is deceased?
- Medical exams

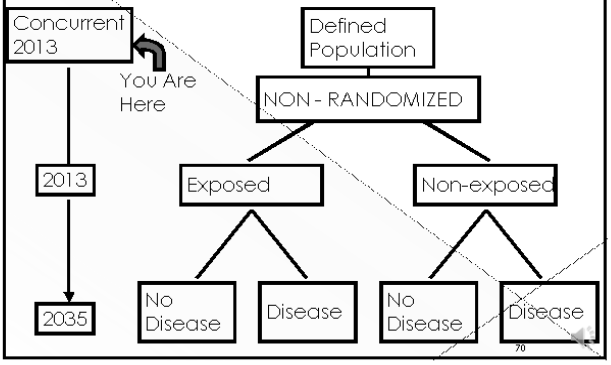
68

Cohort Study



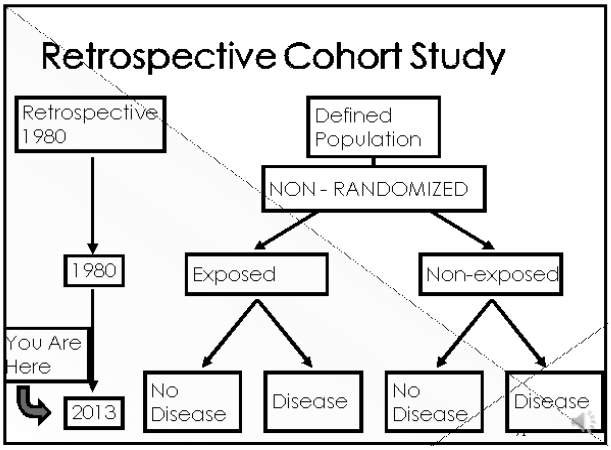
69

Concurrent Cohort Study

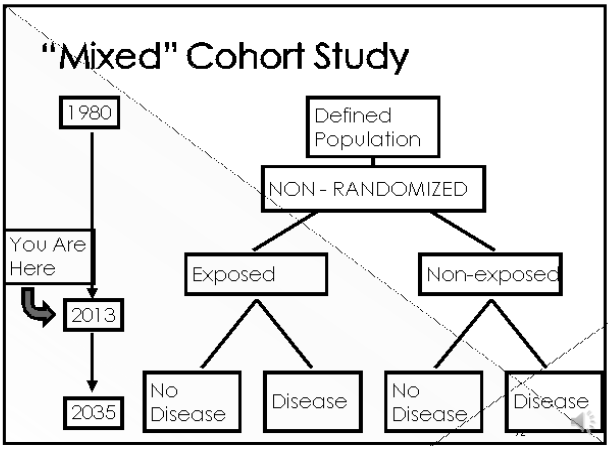


70

Retrospective Cohort Study



"Mixed" Cohort Study



Retro- vs Prospective Cohort

- Retrospective
  - Cheaper, faster
  - Efficient with diseases with long latent period
  - Exposure data may be inadequate
- Prospective
  - More expensive, time consuming
  - Not efficient for diseases with long latent periods
  - Better exposure and confounder data
  - Less vulnerable to bias (if good follow-up)
  - Clearer temporal relationship of exposure & outcome
- But availability of cohorts…

73

Feasibility Considerations

- Selection of population for study
  - Frequency of the outcome under study
  - Need to obtain complete and accurate exposure and follow-up

74

Study Design: Cohort
Bias, Analysis, Summary

75

Learning Objectives

- Understand
  - potential sources of bias in a cohort study
  - analysis of a cohort study

76

Bias Reduction

- Masking to avoid bias in outcome ascertainment
  - Mask exposure status to those conducting follow up and confirming outcomes
- Loss to follow up
  - Crucial to have high follow-up rates
  - Comparable ascertainment of outcomes in the exposed and comparison groups

77

Loss to Follow-up

- Follow all subjects from exposure to outcome or end of study
  - town lists
  - Polk directories
  - telephone books
  - Vital Records: birth, death, marriage records
  - driver's license lists
  - physician / hospital records
  - relatives, friends
- Expensive, time consuming

78

Issues in Interpretation

- Role of bias
  - Misclassification
    - Nondifferential or random
    - Differential
  - Loss to follow-up
  - Effect of nonparticipation
  - Generalizability vs. validity
- Effect of nonrandomization of exposure
- Measurement of confounders

79

Cohort Study Data

ID	Exposure T_0	Disease T_0	Disease T_1
1	Y	N	Y
2	N	N	Y
3	Y	N	N
4	N	N	N
Etc.	…	…	…
n	Y	N	Y

80

Cohort Study - Measure of Association

	Disease: YES	Disease: No	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$RR = \frac{Risk_{exposed}}{Risk_{unexposed}} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{a(c+d)}{c(a+b)}$$

81

Analysis Example

	Breast Cancer Cases	Woman-Years of follow-up
Exposed	41	28,001
Unexposed	15	19,025
Total	56	47,036

IR = Incidence Rate = events / person-time

IR_e = 41/28,011 = 1.5/1,000 woman-years

R_U = 15/19,025 = 0.8/1,000 woman-years

IRR = IR_1/IR_0 = 1.9

82

Cohort Studies Advantages

- Temporal sequence well established
- Rare exposures
- Allows for examination of multiple effects of a single exposure
- Minimizes selection bias

83

Cohort Studies Caveats

- Not good for rare diseases
- If prospective
  - Time consuming
  - Expensive
  - Not good for diseases with long induction or latent period
- If retrospective
  - Exposure assessment vulnerable to bias
- Maximized follow-up bias

84

Design Characteristics

1. # of observations
2. Data collection methods / source of data
3. Timing of data collection
4. Unit of observation
5. Availability of subjects
6. Method of defining "*study population*"
7. Measure of association

85

Cohort AND Clinical

- RR is measure of association
- Groups: exposed / not exposed

86

Cohort vs Clinical

- Clinical
  - Investigator controls exposure
- Cohort
  - Investigator does not control exposure

87

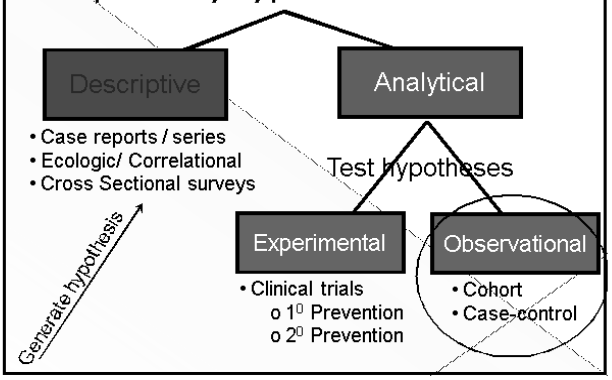
Study Design: Case Control

Introduction

Learning Objectives

- Understand the rationale for & set up of the **case control** study design
- Understand **procedures** involved in the conduct of a **case control** study

Epi Study Types



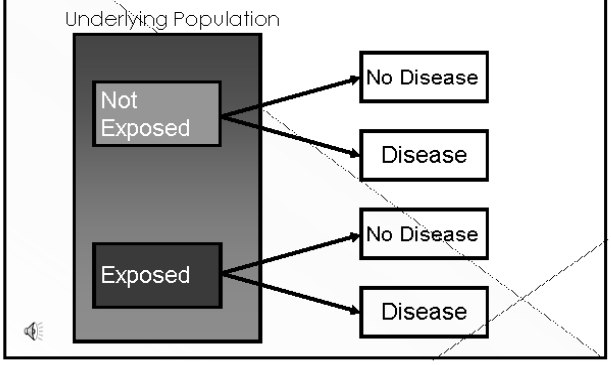
What's an Epidemiologist to do?

- Potentially harmful exposure
  - Clinical trial is unethical
- Rare disease
  - Cohort study would require a huge sample size
- Long induction or latent period
  - Prospective cohort study would take a long time
- Retrospective cohort
  - But what if no retrospective cohort is available?

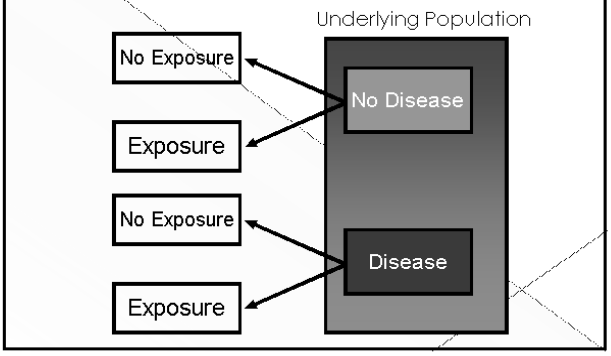
Case-Control Study

- Cases of disease identified & enrolled
- Population that produced the cases identified, a sample enrolled as a comparison (controls)
- Exposures determined for cases and controls

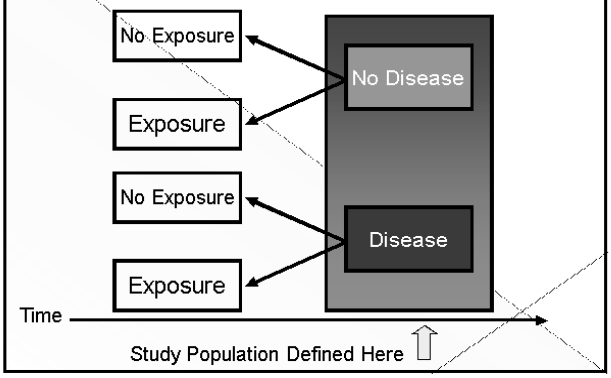
Cohort Study



Case-Control Study



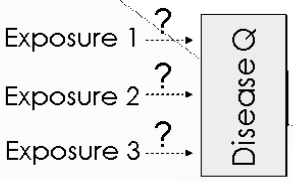
Case-Control Study



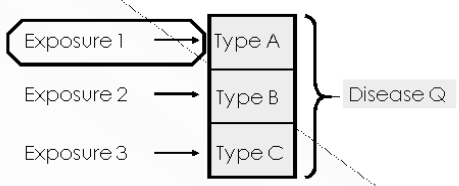
Case Selection

- Who are the cases?
  - Cases should be as homogeneous a disease entity as possible

Homogeneity of Cases



Homogeneity of Cases



Case Selection

- Case definition -> accurate classification of disease
- Efficient and accurate sources to identify cases: existing registries, hospitals
  - Hospital-based
  - Population based
    - Registries
    - Surveillance systems
- Incident vs. prevalent cases
  - Increased risk for disease OR
  - Increased duration of disease

But, *Compared to what?*

Controls

- Sample of source population that gave rise to the cases
- To estimate the exposure distribution in the source population that produced the cases

Control Selection

- Controls represent
  - NOT the entire nondiseased population
  - But population who meet:
- The "Would Criterion"
  - If the potential control actually had the outcome WOULD he/she end up as a case in your study?

Control Sources

- Where might we find controls?
  - Population controls
  - Hospital controls
  - Special control series

General Population

- Controls selected from defined geographic population
  - Sources
    - random digit dialing
    - residence lists
    - drivers' license records

General Population

- Advantages
  - Usually from same base population as the cases (might meet "would" criterion)
- Disadvantages
  - Time
  - Expensive
  - Low response rate
  - *may remember exposures differently than cases*

Hospital Controls

- Cases selected from hospital population
  - Example
    - Cigarette smoking and myocardial infarction in women
    - Cases: admissions to hospital coronary care units
    - Controls: from surgical, orthopedic, and medical unit of same hospital
      - musculoskeletal & abdominal disease, trauma, & other non-coronary conditions

Hospital Controls

- Advantages
  - Same selection factors leads cases & controls to hospital
  - Easily identifiable and accessible
    - so less expensive than population-based controls
  - Controls are also sick so accuracy of exposure recall comparable to cases
  - More willing to participate than population-based controls
- Caveats
  - Hospital based controls are ill
    - may not accurately represent the exposure experience in the population that produced the cases
  - Hospital catchment areas may be different for different diseases

Hospital Controls

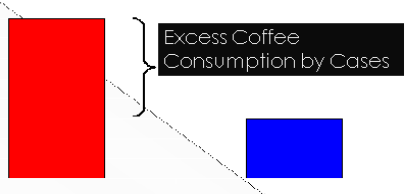
- What Illnesses?
  - those with no relation to the risk factor(s) under study

Example

- Would you use people with respiratory diseases as controls for a study of smoking and myocardial infarction?

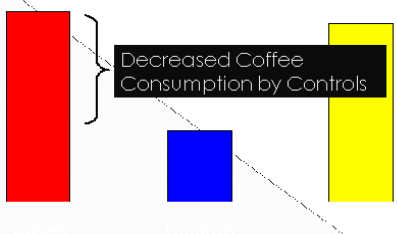
Do they represent the distribution of smoking in the entire population that gave rise to the cases of MI?

Coffee & Pancreatic Cancer



MacMahon B, et al. 1981. Coffee and cancer of the pancreas. N Engl J Med 304:630-3.

Coffee & Pancreatic Cancer



Control Groups

- Friends, spouses, siblings, & deceased individuals as controls
  - Rarely used
  - Some cases
    - few appropriate friends
    - widowed
    - only or adopted children
  - Dead controls are more likely than living controls to have been smokers and drinkers

Outcome / Exposure

- Outcome
  - Cases, already
  - Ensure controls are not cases
- Exposure Assessment
  - Questionnaires / interviews
  - Environmental monitoring
  - Biomarkers

Matching

- What is it?
  - Some characteristics may be associated with exposure and/or the outcome
  - Some cases have these characteristics
  - Select controls with the same characteristics
  - Example:
    - Cases all age > 65
    - Select controls matched to cases on age

Matching

- Individual matching
  - As each case is enrolled find a case that matches
- Frequency matching
  - Calculate the frequency of the variable to be matched in the case group and select controls to have the same frequency

Matched Analysis

- Specialized statistical techniques for analysis of matched case control data,
- BUT
  - "The results of the matched analysis were consistent with the unmatched analysis. Only the results of the unmatched analysis are presented here"

Bias Reduction

- ALL selection criteria must apply equally to cases and controls
- Information must be as similar as possible for cases and controls

Study Design: Case Control

Analysis I

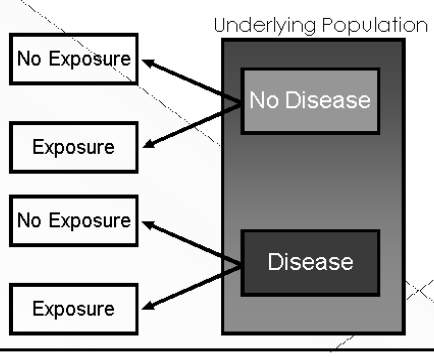
Analysis Learning Objective:

Understand the issues in the analysis of case control studies

Cross–Classified Data

ID	Disease Status	Exposure Status
1	Y	Y
2	N	N
3	Y	N
4	N	Y
Etc.	…	…
n	Y	Y

Case-Control Study



2 X 2 Table

Disease Status

Exposure Status	Yes	No	TOTAL
Yes	*a*	*b*	*a + b*
No	*c*	*d*	*c + d*
TOTAL	*a + c*	*b + d*	*a + b + c + d*

Cohort Study: Proportion with disease in total population can be determined from proportion in study population

2 X 2 Table Case-Control

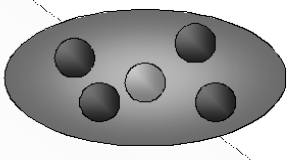
m = exposed controls
n = 2*(a + c) - *m*

Disease Status

Exposure Status	Yes	No	TOTAL
Yes	*a*	*m*	*a + m*
No	*c*	*n*	*c + n*
TOTAL	*a + c*	*2*(a + c)*	*a + c + 2*(a + c)*

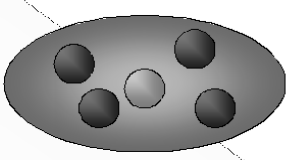
Case-Control Study: Proportion with disease in total population can NOT be determined: proportion in study set by investigator

Proportion



1 / 5

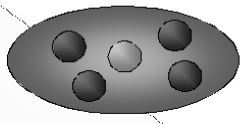
Odds



1:4

1/4

Odds vs Proportion



Odds

Yes / No

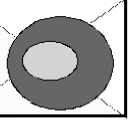
1:4 = ¼ = .25

Proportion

Yes / (Yes + No)

Yes / All

1/5 = .20



2 X 2 Tables

		Disease Status		
		Yes	No	TOTAL
Exposure Status	Yes	***a***	***b***	***a + b***
	No	***c***	***d***	***c + d***
	TOTAL	***a + c***	***b + d***	***a + b + c + d***

Odds of Exposure in Cases = ***a / c***

2 X 2 Tables

		Disease Status		
		Yes	**No**	TOTAL
Exposure Status	Yes	***a***	***b***	***a + b***
	No	***c***	***d***	***c + d***
	TOTAL	***a + c***	***b + d***	***a+b+c+d***

Odds of Exposure in Controls = ***b / d***

Case-Control - Analysis

Case-control		Disease Status		
		Yes	No	TOTAL
Exp. Stat.	Yes	*a*	*b*	*a+b*
	No	*c*	*d*	*c+d*
	TOTAL	*a+c*	*2*(a+c)*	*3*(a+c)*

Cases – exposed or not exposed

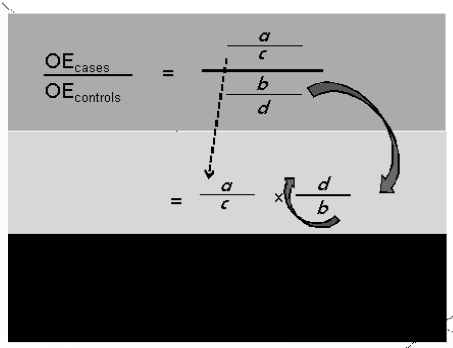
› Odds of exposure among cases = a/c

Controls - exposed or not exposed

› Odds of exposure among controls = b/d

Odds Ratio = OR = $\frac{\text{Odds of exposure in cases}}{\text{Odds of exposure in controls}} = \frac{a/c}{b/d}$

Odds Ratio



2 X 2 Tables

		Disease Status		
		Yes	No	TOTAL
Exposure Status	Yes	*a*	*b*	*a + b*
	No	*c*	*d*	*c + d*
	TOTAL	*a + c*	*b + d*	*a + b + c + d*

$$OR = \frac{ad}{bc}$$

"Large" Samples

OR ► RR

Case-Control Analysis

- Example
  - Case-control study
  - spontaneous abortion and prior induced abortion
    - OUTCOME = spontaneous abortion
    - EXPOSURE = prior induced abortion

		Case (Spon. Ab.)	Control (Live birth)
Prior Induced Abortion	Yes	42	247
	No	107	825

Case-control Analysis

		Case (Spon. Ab.)	Control (Live birth)
Prior Induced Abortion	Yes	42	247
	No	107	825

$$\text{Odds Ratio} = OR = \frac{\text{Odds of exposure in cases}}{\text{Odds of exposure in controls}}$$

$$= \frac{42/107}{247/825} = \frac{0.392523}{0.299394} = 1.31$$

Interpretation

- In this study, women who had a prior induced abortion were 1.31 times as likely to have a spontaneous abortion as women who had not had a prior induced abortion
- In this study, women who had a spontaneous abortion were 1.31 times as likely to have had a prior induced abortion as women who delivered a live baby

Nested Case-Control

- Cohort study ongoing
  - Some cohort members develop disease
- Sample members of cohort without the disease as controls

Nested Case-Control

- Hypothetical Example
  - Prospective cohort study
    - 89,949 women aged 34-59
  - Blood drawn on all 89,949 at baseline & frozen

Nested Case-control

- Does pesticide exposure increase breast cancer risk?
  - 1,439 breast cancer cases identified over 8 years of follow-up
- Exposure
  - Level of pesticides in blood
    - High
    - Low

Nested Case-control

- Quantifying blood pesticide levels
  - $$$
  - all 89,949 blood samples!!
- Analyze blood on all cases (N=1,439)
- Sample women who did not get breast cancer
  - say two times as many cases (N=2,878)

Nested Case-Control Study

Breast Cancer

DDE	Cases	Controls
High	360	432
Low	1,079	2,446
Total	1,439	2,878

$$\text{Odds Ratio} = OR = \frac{\text{Odds of exposure in cases}}{\text{Odds of exposure in controls}}$$

$$= \frac{360/1079}{432/2446} = \frac{0.33364}{0.17661} = 1.89$$

Nested Case-Control

Breast Cancer

Pesticide	Yes	No	Total
High	360	13,276	13,636
Low	1,079	75,234	76,313
Total	1,439	88,510	89,949

Relative Risk = RR = (360/13,636) / (1,079/76,313) = 1.88

Case-Control Advantages

- Good when exposure data expensive or difficult to obtain
  - Pesticide / breast cancer study
- Disease has long induction / latent period
  - Cancer, cardiovascular disease
- When the disease is rare
  - Studying risk factors for birth defects
- When little is known about the disease
  - Evaluate multiple exposures
  - Early studies of AIDS
- When underlying population is dynamic
  - Migrant workers

Case-Control Caveats

- Inefficient for rare exposures
- Exposures and diseases have already occurred at the time of the interview
  - Assessment of past exposure may lack precision
- Design susceptible to
  - Differential selection of cases or controls on basis of exposure
  - Differential reporting of exposure based on disease status
- Temporal relationship of exposure and outcome?
- Some skepticism about logic of such studies

Design Characteristics

1. # of observations
2. Data collection methods / source of data
3. Timing of data collection
4. Unit of observation
5. Availability of subjects
6. Method of defining *"study population"*
7. Measure of association

Study notes…….. You should be

- ❖ Be able to describe & discuss scenarios in which to use the various study design types
- ❖ Differentiate between Case – Control & Cohort Study
- ❖ How participants are selected
- ❖ Appropriate measure of association for study type
- ❖ Correct interpretation of findings

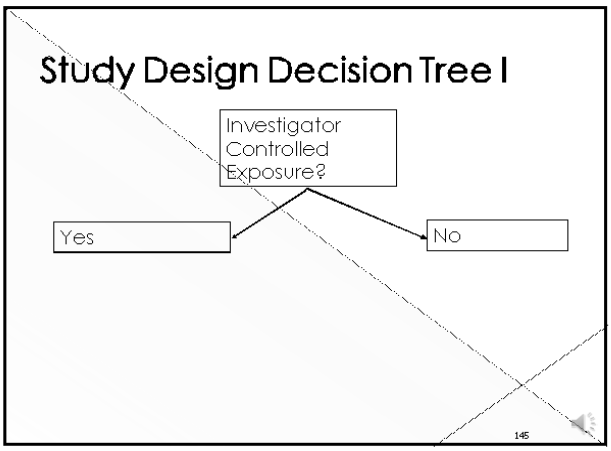
Study Design Summary

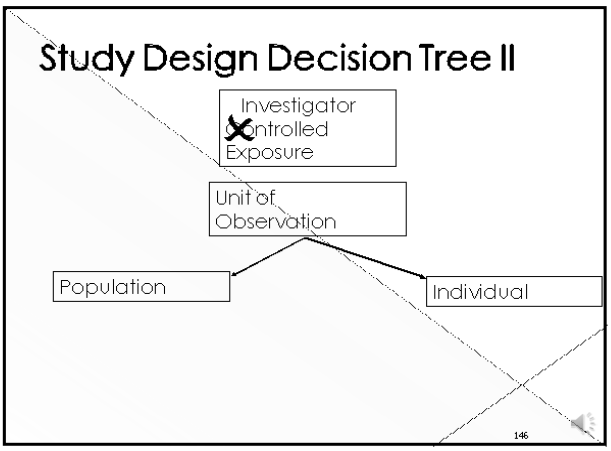
143

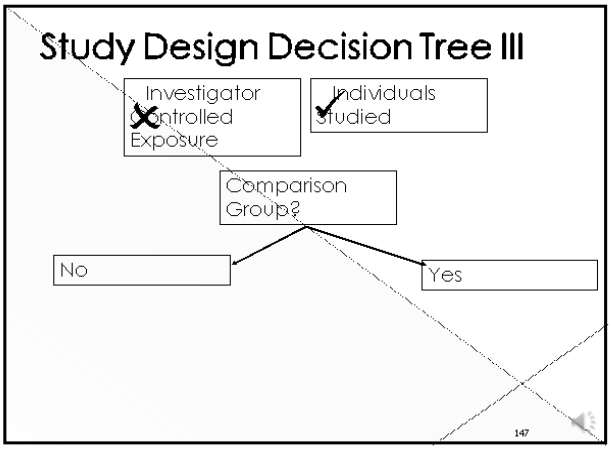
Learning Objectives

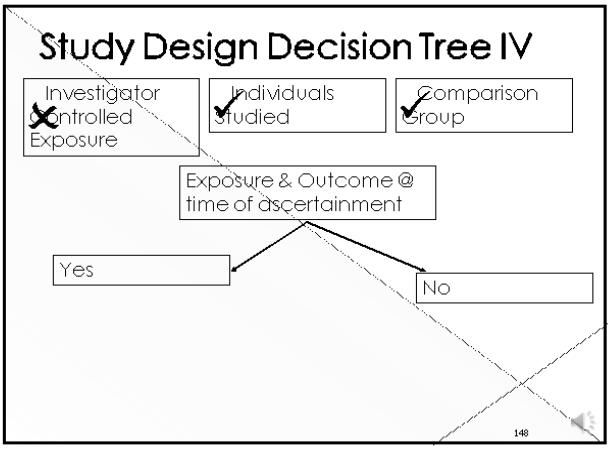
- Understand the **key elements** that will help you determine which **study design** is being described in a paper or test question

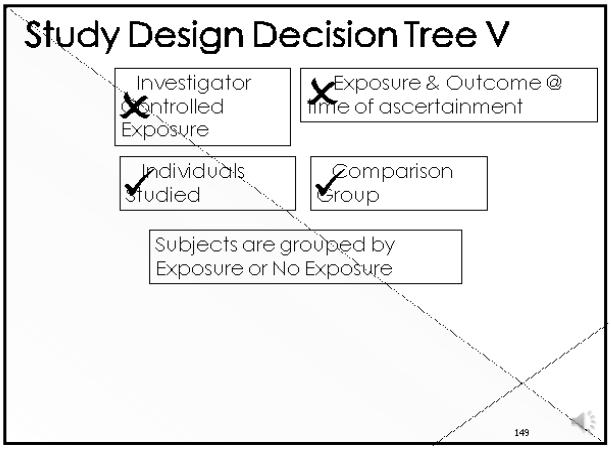
144

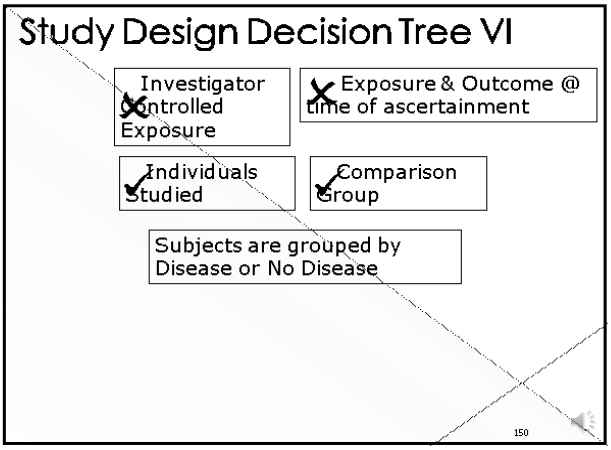

Study Design Decision Tree I
Investigator Controlled Exposure?
Yes
No
145


Study Design Decision Tree II
✗ Investigator Controlled Exposure
Unit of Observation
Population
Individual
146


Study Design Decision Tree III
✗ Investigator Controlled Exposure
✓ Individuals Studied
Comparison Group?
No
Yes
147


Study Design Decision Tree IV
✗ Investigator Controlled Exposure
✓ Individuals Studied
✓ Comparison Group
Exposure & Outcome @ time of ascertainment
Yes
No
148


Study Design Decision Tree V
✗ Investigator Controlled Exposure
✗ Exposure & Outcome @ time of ascertainment
✓ Individuals Studied
✓ Comparison Group
Subjects are grouped by Exposure or No Exposure
149


Study Design Decision Tree VI
✗ Investigator Controlled Exposure
✗ Exposure & Outcome @ time of ascertainment
✓ Individuals Studied
✓ Comparison Group
Subjects are grouped by Disease or No Disease
150

Additional Hints

- If all subjects have a disease, it is neither the exposure or the outcome
  - A study of the effect of vitamins on survival of cancer patients
- Personal characteristics can be considered "exposure"
  - Risk of accident for left-handers vs right handers

151

	Ecological	Cross Sectional	Case-Control	Cohort
Rare *disease*				
Rare *cause*				
Multiple exposures				
Multiple outcomes				
Temporal relationship				
Measurement of incidence				
Long latent periods				
Time Required				
Cost				

152

* = usually

	Descriptive Studies			Analytic: Experimental		Analytic: Observational	
	Case Rpt/Srs	Ecol/ Corr	X-Sxnl	Clin Trial	Comm Trial	Cohort	Case Control
# Obs	Small*	Var	Var	Var	Var	Large*	Var
Dat Source	Clin	2ndry	Qs	Clin/qs	2ndry	Clin/qs	Clin/qs
Time of Dat Collect	After	After	After	Both	Both	Before*	After
Unit of Obs	Individ	Comm	Individ	Individ	Comm	Individ	Individ
Avail of Subj	Var	Var	Var	Var	Var	Var	Var
Define Study Pop	Clin	2ndry	Sample	Exp or not	Exp or not	Exp or not	Disease or not
Meas Assn	None	Corr coef	Corr coef	Rr	Rr	Rr	Or

153

Interpreting Relative Risk

- In this study,
- the risk of *specify your disease*
- in *add specifics about your study population*
- who were exposed to *specify your exposure*
- was *specify the RR* times as great
- as the risk to *add specifics about your study population*
- who were not exposed.

154

Interpreting Relative Risk

- In this study,
- the risk of *asthma attacks*
- in *left handed, middle aged, epidemiologists*
- who were exposed to *scented lotion applied to coworkers hands*
- was *3.5* times as great
- as the risk of asthma attacks in *left handed, middle aged, epidemiologists*
- who were not exposed

155

Interpreting Odds Ratios

- In this study,
- *specify details about your case population*
- who had *specify your disease*
- were *specify the OR* times as likely to have been
- exposed to *specify your exposure*
- as *add specifics about your control population*

156

Interpreting Odds Ratios

- In this study,
- *left handed, middle aged, epidemiologists*
- *who had asthma attacks*
- were 3.5 times as likely to have been
- *exposed to scented lotion applied to coworkers hands*
- *as left handed, middle aged, epidemiologists* who did not *have asthma attacks*

157

Memorization

- DO know:
  - What is a proportion?
  - What is an odds?
  - How do you compare risk?
  - Which design allows you to calculate actual risk?
  - What is the epidemiologic measure of risk?

158

Memorization

- If you choose to memorize the a, b, c, d formulas,
- Be SURE to ALSO memorize the table layout, AND
- Be prepared to set up the table that way

159

2 X 2 Tables

Exposure Status	Disease Status: Yes	Disease Status: No	TOTAL
Yes	a	b	$a+b$
No	c	d	$c+d$
TOTAL	$a+c$	$b+d$	$a+b+c+d$

$$OR = \frac{ad}{bc}$$

160

2 X 2 Tables

Exposure Status	Disease Status: No	Disease Status: Yes	TOTAL
Yes	a	b	$a+b$
No	c	d	$c+d$
TOTAL	$a+c$	$b+d$	$a+b+c+d$

$$\frac{\text{Odds of Exposure in Cases}}{\text{Odds of Exposure in Controls}}$$

Memorized Formula

$$OR = \frac{ad}{bc}$$

161

2 X 2 Tables

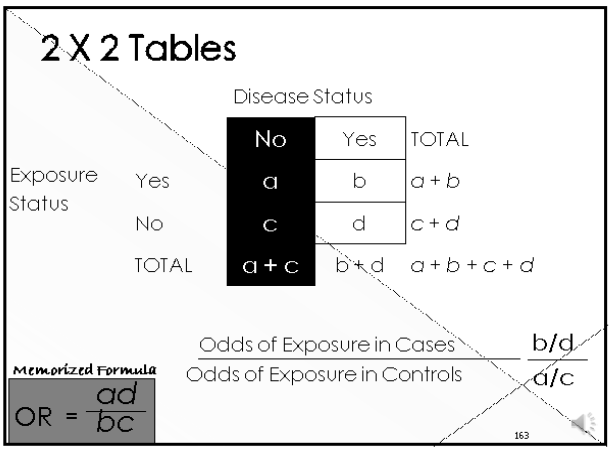
Exposure Status	Disease Status: No	Disease Status: Yes	TOTAL
Yes	a	b	$a+b$
No	c	d	$c+d$
TOTAL	$a+c$	$b+d$	$a+b+c+d$

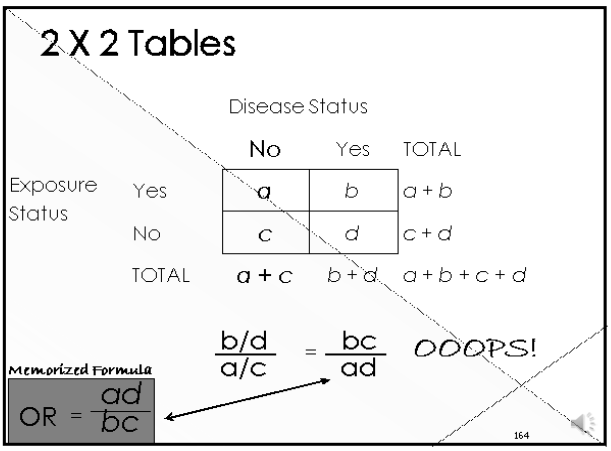
$$\frac{\text{Odds of Exposure in Cases}}{\text{Odds of Exposure in Controls}} \quad b/d$$

Memorized Formula

$$OR = \frac{ad}{bc}$$

162


2 X 2 Tables
Disease Status
No
Yes
TOTAL
Exposure Status
Yes
a
b
a + b
No
c
d
c + d
TOTAL
a + c
b + d
a + b + c + d
Odds of Exposure in Cases
Odds of Exposure in Controls
b/d
a/c
Memorized Formula
OR = ad/bc
163


2 X 2 Tables
Disease Status
No
Yes
TOTAL
Exposure Status
Yes
a
b
a + b
No
c
d
c + d
TOTAL
a + c
b + d
a + b + c + d
b/d / a/c = bc / ad
OOOPS!
Memorized Formula
OR = ad/bc
164