

Measures of Effect

Attributable Risk (AR) or Risk Difference (RD)

■ Definition:

Measure of effect that provides information about the absolute effect of the exposure or the excess risk of disease in those exposed compared with those nonexposed.

■ Group of Interest: Exposed

- The attributable risk is used to quantify the risk of disease in the "exposed" group that can be considered attributable to the exposure by removing the risk of disease that would have occurred anyway due to other causes.

How to Calculate?

- Attributable risk is defined as the difference between the incidence rates (or cumulative incidence) in the exposed and nonexposed.

- For example, in a cohort study

$$AR = I_e - I_o \text{ or } AR = CI_e - CI_o$$

where I_e is the incidence rate in the exposed and I_o is the incidence rate in the nonexposed

Assumptions

- Interpretation of the AR is dependent on the assumption that a cause-effect relationship exists between exposure and disease.
- If no association between the exposure and disease, $I_e - I_o = 0$ so $AR = 0$.

Interpretation

- If there is an association between exposure and disease and $AR > 0$, the AR is the number of cases of the disease among the exposed that could be eliminated if the exposure was eliminated.
- Thus, the AR is useful as a measure of public health impact of an exposure.

Example:

| OC Use | Bacteriuria | | Total |
|--------|-------------|------|-------|
| | Yes | No | |
| Yes | 27 | 455 | 482 |
| No | 77 | 1831 | 1908 |
| Total | 104 | 2286 | 2390 |

$$CI_e = 27/482 = .05601 \quad CI_o = 77/1908 = .04036$$

$$AR = CI_e - CI_o = .05601 - .04036 = .01566$$

Interpretation: The excess occurrence of bacteriuria among OC users attributable to their OC use is 156.6 per 10,000.

Attributable Risk Percent (AR%)

- The attributable risk is often expressed as a percentage, the AR%, which can be interpreted as the "proportion" of the disease in the exposed that could be prevented by eliminating the exposure.

How to Calculate?

- $AR\% = (I_e - I_o) / I_e$ or $AR\% = (RR - 1) / RR * 100$

where RR is the risk ratio or relative risk of the exposure and disease (cohort study)

- Note: In a case control study, the OR may be substituted for the RR.

- Recall OC use and bacteriuria example...

$$CI_e = 27/482 = .0560 \quad CI_o = 77/1908 = .0404$$

$$AR\% = (CI_e - CI_o) / CI_e = (.0560 - .0404) / .05601 = .2796 \text{ or } 27.96\%$$

Interpretation: If OC use does cause bacteriuria, about 28% of bacteriuria among women who use OCs can be attributed to their OC use.

- Example continued...

$$CI_e = .05601 \quad CI_o = .04036$$

$$RR = CI_e / CI_o = .05601 / .04036 = 1.388$$

$$AR\% = (RR - 1) / RR * 100 = (1.388 - 1) / 1.388 = 27.96$$

Note: Result is the same as previous.

Population Attributable Risk (PAR)

- Recall that the AR quantifies the excess risk in the "exposed" group.
- The PAR estimates the excess rate of disease in the "total" study population of exposed and nonexposed that is attributable to the exposure.

Uses of PAR

- The PAR helps determine which exposures have the most relevance to the public's health.
- Can help guide the allocation of resources aimed at interventions

How to Calculate?

$$PAR = I_T - I_0 \text{ or } PAR = CI_T - CI_0$$

where I_T is the incidence rate in the total population and I_0 is the incidence rate in the nonexposed

Alternatively, the PAR can be calculated as
 $PAR = (AR) * P_e$

where AR is the attributable risk and P_e is the proportion of exposed people in the population

Recall OC use and bacteriuria example...

| OC Use | Bacteriuria | | Total |
|--------------|-------------|-------------|-------------|
| | Yes | No | |
| Yes | 27 | 455 | 482 |
| No | 77 | 1831 | 1908 |
| Total | 104 | 2286 | 2390 |

$$CI_T = 104/2390 = .0435$$

$$PAR = CI_T - CI_0 = .0435 - .04036 = .00314$$

Interpretation: If OC use were eliminated in this study population, the excess annual incidence rate of bacteriuria that could be eliminated is 31.4 per 10,000.

PAR%

- Similar to the AR%, the PAR% is the "proportion" of disease in the study population that is attributable to the exposure and could be eliminated if the exposure was eliminated

How to Calculate?

- $PAR\% = (I_T - I_o) / I_T$
or
- $PAR\% = [P_e(RR - 1) / (P_e(RR - 1) + 1)] * 100$

where RR is the risk ratio or relative risk of the exposure and disease (cohort study) and P_e is the proportion of exposed people in the population

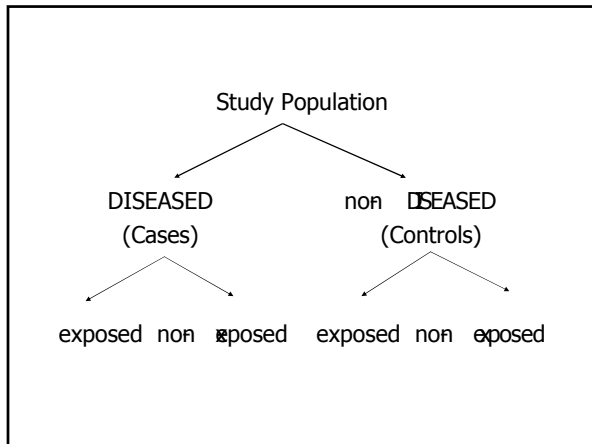
- Note: In a case control study, the OR may be substituted for the RR.

Case-Control Studies

Mount Sinai School of Medicine
Center for Multicultural &
Community Affairs
March 25, 2003

CASE-CONTROL STUDIES SOME KEY POINTS

- Most frequently used study design
- Participants selected on the basis of whether or not they are DISEASED
- Those who are diseased are called CASES.
- Those who are not diseased are called CONTROLS.



Because participants are selected on the basis of disease, exposures for ALL PARTICIPANTS are obtained RETROSPECTIVELY...

| | | |
|----------|-------------------|---------------------------------|
| PAST | | PRESENT |
| Exposure | ← <i>recall</i> → | Cases & Controls Selected |

Example: lung cancer cases and non-cancerous controls recall past exposure to cigarette smoke

SELECTION OF CASES

- FIRST decide on a specific case definition based on a medically diagnosed condition.
- Must consider what criteria will confirm the case definition:
 - lung cancer confirmed by biopsy
 - osteoporosis confirmed by bone density measurements
 - atherosclerosis confirmed by ultrasound of carotid arteries

SELECTION OF CASES

- SECOND will you use INCIDENT or PREVALENT cases?
- Incident...
 - must wait for new cases to occur (time consuming)
 - study can specifically measure exposure relating to development of disease (temporality)
- Prevalent...
 - don't have to wait while cases occur with time - more practical!
 - study will specifically measure exposure relating to survival with disease (may miss cases with severe disease that have already died)

X = deceased

Time →

Sample Selected

Your sample is biased because you have only selected survivors

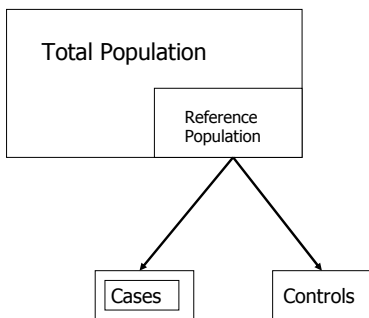
SELECTION OF CASES

- THIRD be aware of the unique qualities of certain groups:
 - hospital admissions
 - nursing homes
 - screening participants
 - day care facilities
- some groups may have better supporting medical records
- some groups may be more homogenous and present less confounding variables

SELECTION OF CONTROLS

■ THE BIG PICTURE...

- Controls should be representative of the referent population from which cases are selected (i.e. comparable)
- They don't have to be representative of the whole population
- Controls should have the potential to become cases



Controls should be comparable to cases.

SELECTION OF CONTROLS

■ Different Types of Controls...

- Population controls
 - randomly selected individuals from the population like RDD (random digit dialing)
- Neighborhood controls
 - individuals that live in the same neighborhoods as cases

SELECTION OF CONTROLS

- Different Types of Controls (con't)...
- Friend controls
 - best friends of cases
 - spouses or siblings of cases
- Hospital controls
 - individuals at the same hospital with cases

MEASURING ASSOCIATION

- Because study participants in Case Control studies are selected based on disease status...
- Case control studies are ideal for the study of rare diseases.
- Incidence can't be calculated.

MEASURING ASSOCIATION

- Because incidence can't be calculated, a relative risk can't be calculated.
- Instead of the RR, an ODDS RATIO is calculated in case control studies.

MEASURING ASSOCIATION

- Odds: NOT a proportion, but the ratio of the # of ways an event CAN occur relative to the # of ways an event CAN NOT occur

$$\text{Odds} = \frac{P(\text{event occurs})}{1 - P(\text{event occurs})} = \frac{p}{1 - p}$$

- Odds Ratio: $\frac{\text{Odds of case being exposed}}{\text{Odds of control being exposed}}$

Classic 2 x 2 Table for a Case-Control Study

| | <u>Disease</u> | <u>No Disease</u> |
|--------------------|----------------|-------------------|
| <u>Exposure</u> | a | b |
| <u>No Exposure</u> | c | d |

$$\text{Odds Ratio} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

Example: Is Use of Artificial Sweeteners associated with Bladder Cancer?

| | <u>Cases</u> | <u>Controls</u> |
|-------------------|--------------|-----------------|
| <u>Ever Used</u> | 1,293 | 2,455 |
| <u>Never Used</u> | 1,707 | 3,321 |
| <u>Total</u> | 3,000 | 5,776 |

$$\text{ODDS RATIO} = \frac{1,293 * 3,321}{2,455 * 1,707} = 1.025$$

Hoover and Strasser (1980) Lancet 1: 837-840

Interpretation of the Odds Ratio...

If:

OR = 1 then exposure NOT related to disease

OR >1 then exposure POSITIVELY related to disease

OR <1 then exposure NEGATIVELY related to disease

Hoover and Strasser concluded what from their study?

Interpretation:

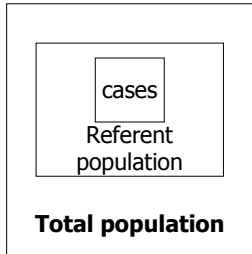
The odds that those with bladder cancer had used artificial sweeteners is 1.025 times greater than those who do not have bladder cancer.

A Special Case...

When the disease is RARE and the duration of the given disease is SHORT...

$$O.R. \cong R.R.$$

CASE-CONTROL STUDY SUMMARY



- cases and controls are representative of a referent population
- controls have the potential to become cases
- selection based on disease and exposure assessed retrospectively

Twists and Turns in Case-Control Studies

SELECTION OF CONTROLS

- The investigator can elect to use more than one TYPE of control for each case... when there is no ONE group similar enough to cases.

EXAMPLE: A particular leukemia case may have both a neighborhood control (similar to case in terms of environment) and a sibling control (similar to case in terms of genetic background).

SELECTION OF CONTROLS

■ POTENTIAL PROBLEMS of Control Selection...

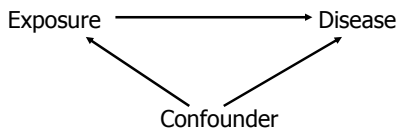
- You must assume that there is equal exposure in your case and control groups.
- You must assume a steady state population so that controls and cases have had similar lifetime exposures.

Example: If you're studying lung cancer in adults and a relevant exposure is asbestos in local schools you assume both cases and controls to have attended local schools.

SELECTION OF CONTROLS

■ To avoid potential problems of confounding some studies use MATCHING.

- MATCHING: The process of selecting controls so that they are similar to cases on certain specific characteristics.



Confounders are third variables that are associated with both the disease and the exposure.

SELECTION OF CONTROLS

■ CHARACTERISTICS THAT ARE OFTEN USED FOR MATCHING...

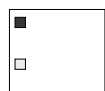
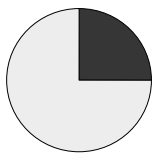
- age
- gender
- body mass index (weight / height²)
- smoking status
- marital status

SELECTION OF CONTROLS

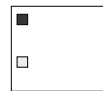
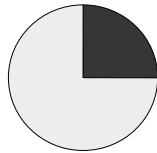
- There are two types of matching...
 - 1) GROUP MATCHING (frequency matching)
 - based on proportions
 - Idea is to select a control group with a certain characteristic identical to cases in the same proportion as it appeared in cases.
- Example: If 25% of cases in your study smoke you would select a control population that included 25% smokers.

GROUP MATCHING EXAMPLE

CASE POPULATION



CONTROL POPULATION



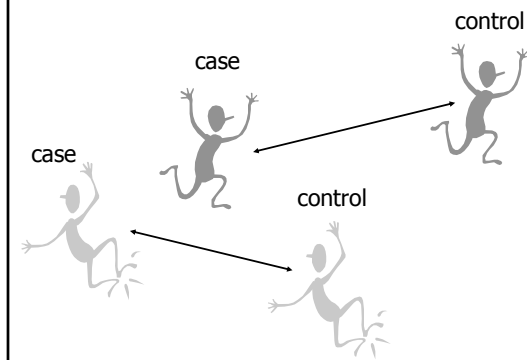
SELECTION OF CONTROLS

2) INDIVIDUAL MATCHING (matched pairs)

- For every individual case a control is selected who is identical to the case on certain characteristics.

Example: If your first case is a 25 year old woman who smokes then you would find a control who is 25, female and a smoker. So you are matching on age, gender, and smoking status.

MATCHED PAIRS EXAMPLE



POTENTIAL PROBLEMS WITH MATCHING

- It will be difficult to find controls if too many variables are selected for matching.
- Variables used for matching can not be studied as exposures or confounders.
- **OVERMATCHING**- when variables related to disease are inadvertently matched upon.

OVERMATCHING EXAMPLE

A study of breast cancer uses women matched on BMI and smoking status...

PROBLEM: BMI and smoking may be related to the disease of interest (breast cancer) but because they were used for matching they can not be studied as they will be the same in cases and controls BY DESIGN.

NESTED CASE-CONTROL STUDIES

- A new study design...
- Case control studies can be NESTED within a cohort study- a hybrid design...

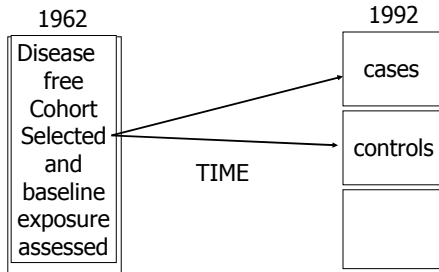
1. A cohort of study participants is assembled.
2. Over time those in the cohort who develop DISEASE become CASES.
3. At the end of the study a sample of the cohort who did NOT develop disease become CONTROLS.

NESTED CASE-CONTROL STUDIES

EXAMPLE: Lung cancer and Vitamin C in Tecumseh

1. A cohort collected in 1962 (information about dietary Vitamin C assessed and blood samples drawn)
2. With time those who developed lung cancer became CASES and their blood samples were then analyzed.
3. At the end of observation period in 1992 a group of non disease CONTROLS was chosen and their blood samples were then analyzed.

EXAMPLE: Lung cancer and Vitamin C in Tecumseh



This example illustrates the main advantages of a NESTED case control study...

- It is more *COST EFFECTIVE* than a cohort because only have to process selected samples (wait to see who are cases, etc.)
- *ELIMINATES RECALL BIAS* because instead of using previously collected data (that may not even have been intended for epidemiology!) all information is collected at baseline

BIAS IN CASE-CONTROL STUDIES

- **BIAS:** any systematic error (not random or by chance) in a study which leads to an incorrect estimate of the association between an exposure and the disease of interest
- **MAIN TYPES** of bias in Case- Control Studies...
 - selection bias
 - recall bias

BIAS IN CASE-CONTROL STUDIES

- **SELECTION BIAS:** systematic error due to differences in characteristics between those selected for a study and those not selected

EXAMPLE in CC Studies: When cases are selected from a hospitalized population with unique exposures, controls often are not representative of the population that gave rise to cases.

BIAS IN CASE-CONTROL STUDIES

- **RECALL BIAS:** Systematic error due to differences in accuracy or completeness of recall to memory of past events or experiences

EXAMPLE in CC Studies: Often cases faced with a serious illness will more closely scrutinize their past exposures and will be more accurate and complete in their recall than controls.

What do you think will happen to our estimate of the Odds Ratio if cases recall their exposure status better than controls?

| | <u>D</u> | <u>ND</u> |
|-------------|----------|-----------|
| Exposure | ↑ a | b |
| No exposure | c | d |

*Odds Ratio = $(\uparrow a) d / bc$

*Note that this will be biased away from the null (i.e. give a 'bigger' OR than what is 'true')

CASE-CONTROL STUDY SUMMARY

■ **ADVANTAGES...**

- good for studying rare diseases
- can use smaller sample sizes
- cost/time effective when using previously collected (RETROSPECTIVE) exposures

■ **DISADVANTAGES...**

- subject to bias (selection and recall)
- can't calculate incidence
- selecting appropriate controls can be challenging
