CPH EXAM REVIEW– EPIDEMIOLOGY

Paul Terry, PhD
Associate Professor
Department of Medicine
University of Tennessee Medical Center - Knoxville

January 13, 2017

Today's Objectives:

- Review of basic topics covered in the epidemiology section of the exam

- Materials covered cannot replace basic epidemiology course

- This review will be archived on the NBPHE website under Study Resources
  www.nbphe.org

Outline

- Definition and Terminology
- Measures of Disease Frequency
- Epidemiologic Study design
- Causation
- Screening for Disease

Epidemiology

is the **study** of **distribution** and **determinants** of **health-related states or events** in **specified populations** and **the application of this study to control health problems**

Disease Distribution

- **How common?**

- **Who is affected?**

- **When does it occur?**

- **Where does it occur?**

Disease Distribution

- **How common?**
- **Who is affected?**
- **When does it occur?**
- **Where does it occur?**

Endemic, Epidemic and Pandemic

- **Endemic: usual presence of a disease within a given population.**

- **Epidemic: occurrence of a disease clearly in excess of normal expectancy in a defined community or region**

- **Pandemic: worldwide epidemic**

Disease Distribution

- **How common?**
- **Who is affected?**
- **When does it occur?**
- **Where does it occur?**

DESCRIPTIVE EPIDEMIOLOGY - PERSON

Some personal characteristics that are examined with respect to disease occurrence are:

- age
- sex / gender
- race / ethnicity
- education
- income
- occupation
- marital status

Populations

- Membership can be permanent or transient

  - Population with permanent membership is referred to as "Fixed" or "Closed"

    - People present at Hiroshima
    - Passengers on an airplane

  - Population with transient membership is referred to as "Dynamic" or "Open"

    - Population of Omaha

Disease Distribution

- **How common?**
- **Who is affected?**
- **When does it occur?**
- **Where does it occur?**
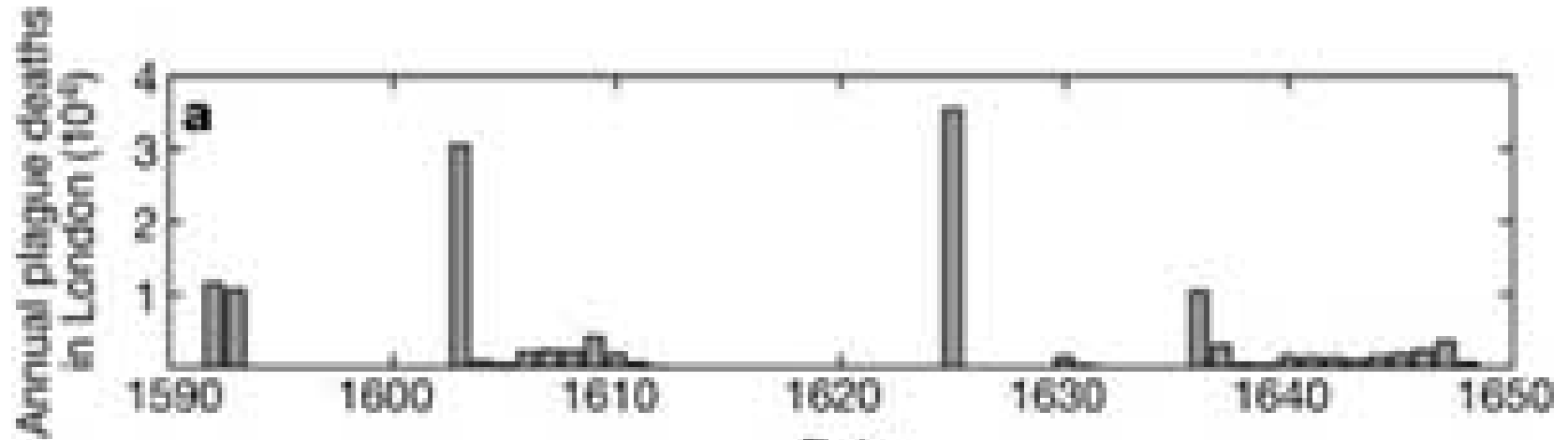
Annual Plague Deaths, London
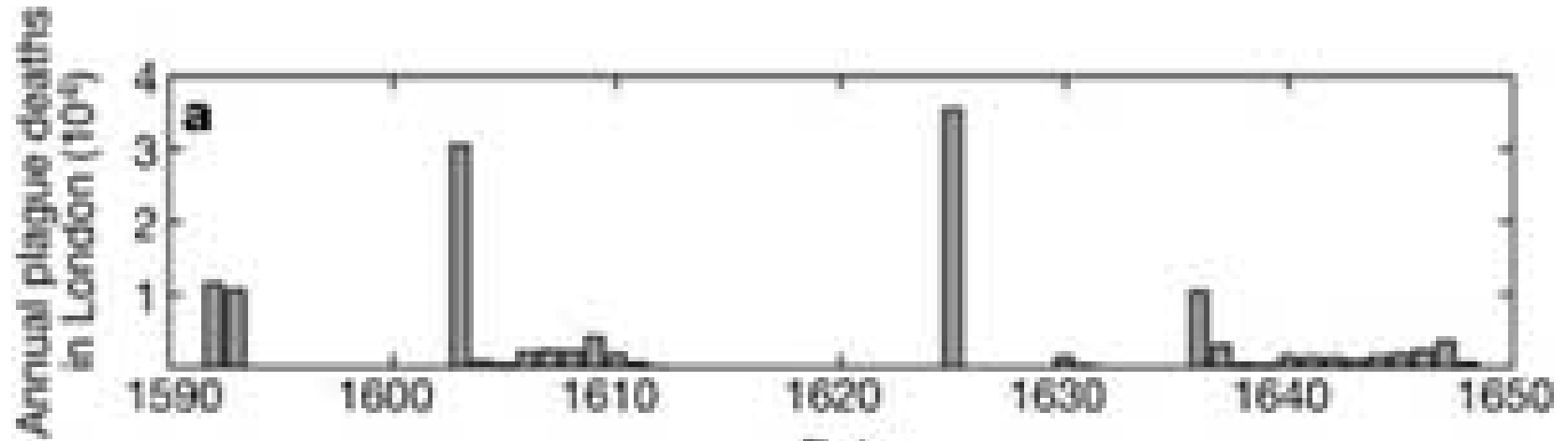
Source: Keeling, M.J. & Gilligan, C. A. (2000) *Nature* **407**, 903-906

Endemic, Epidemic and Pandemic

WHAT IS A RARE (SPORADIC) DISEASE?

Annual Plague Deaths, London

Source: Keeling, M.J. & Gilligan, C. A. (2000) *Nature* **407**, 903-906

Disease Distribution

- **How common?**
- **Who is affected?**
- **When does it occur?**
- **Where does it occur?**

Cholera cases in the Golden Square area of London, August-September 1854

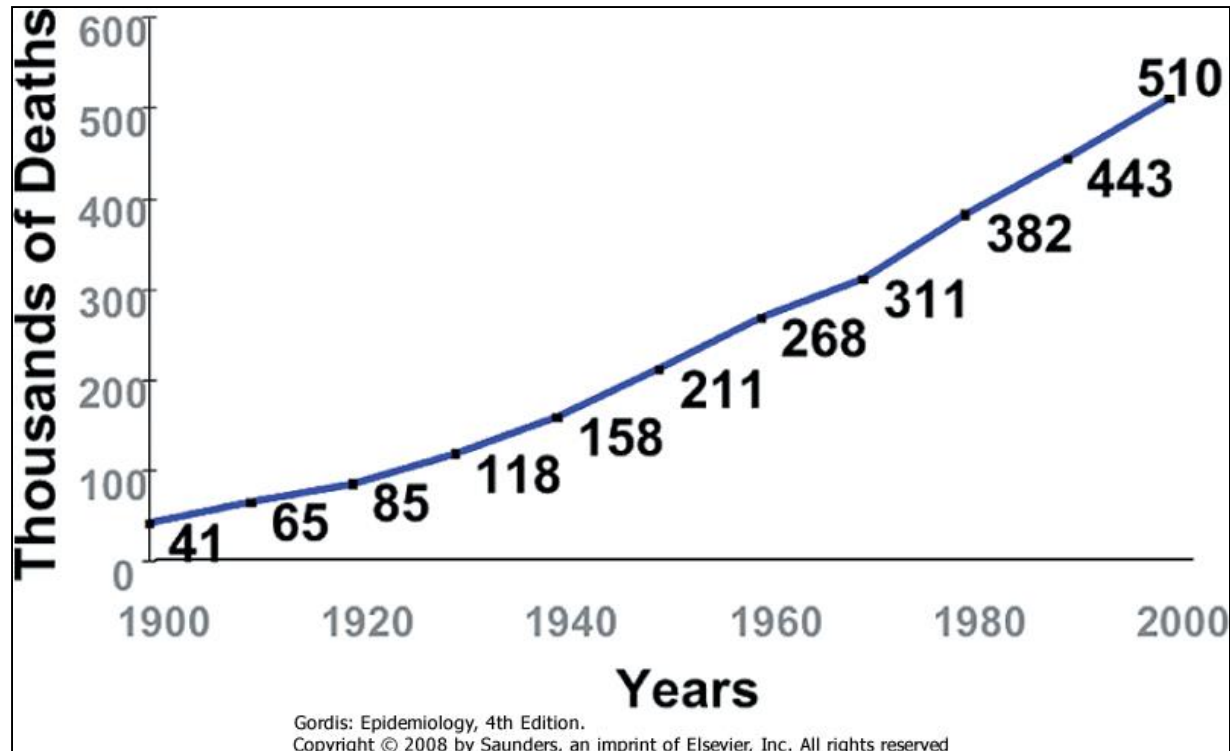# Measures of Morbidity

Four simple mathematical parameters

- Counts
- Ratios
- Proportions
- Rates

Measures of Frequency

"Count" - the most basic epidemiologic measure

- – Expressed as integers (1, 2, 3, ...)
- – Answers the question, "How many people have this disease?"
- – Important to distinguish between *incident* (new) and *prevalent* (existing) cases

Deaths in the U.S. 20th Century

Gordis: Epidemiology, 4th Edition.
Copyright © 2008 by Saunders, an imprint of Elsevier, Inc. All rights reserved

Ratio

- One number *(x)* divided by another *(y):* $\dfrac{x}{y}$

- Range: zero *(0)* to infinity *(∞)*

- *(x)* and *(y)* may be related or completely independent

- Sex of children
- Attending a clinic

$$\frac{females}{males}$$

Ratio

Obtained by dividing one quantity by another.

Example: Number of stillbirths per thousand live births.

$$\frac{50 \text{ stillbirths}}{2500 \text{ live births}} \quad * \text{ 1,000}$$

= 20 stillbirths for every 1,000 live births

Ratio

- One number *(x)* divided by another *(y):* $\dfrac{x}{y}$

- Range: zero *(0)* to infinity *(∞)*

- *(x)* and *(y)* may be related or completely independent

- Sex of children
- Attending a clinic

$$\dfrac{females}{males}$$

Ratio

- One number *(x)* divided by another *(y):* $\dfrac{x}{y}$

- Range: zero *(0)* to infinity *(∞)*

- *(x)* and *(y)* may be related or completely independent

Proportion

- Sex of children
- Attending a clinic

$$\dfrac{females}{males}$$

$$\dfrac{females}{all}$$

Proportion

- Ratio in which the numerator *(x)* is included in the denominator *(x+y):*

- Range: zero *(0)* to one *(1)*

- Often expressed as percentage (e.g., Among all children who attended a clinic, what proportion was female)?

$$\frac{females}{all}$$

Proportion

A ratio in which the numerator is included in the denominator (expressed as %)

Example: The number of fetal deaths out of the total number of births

$\underline{50\ \text{stillbirths}}$ $0.0196*100 = 1.96\%$

2550 total births

Proportion

A ratio in which the numerator is included in the denominator (expressed as %)

Example: The number of fetal deaths out of the total number of births

$$\underline{50 \text{ stillbirths}} \qquad 0.0196*100 = 1.96\%$$
2550 total births

RISK

RISK

$$Risk = \frac{a}{N}$$

Where:

a = number of **<u>new onset cases</u>** (events)
N = population-at-risk at beginning of follow-up

Cohort Follow-up

1 _____

2 _____ ✖ (6 months)

3 _____

4 _____

5 _____

Time (12 months)

RISK

$$Risk = \frac{a}{N}$$

1 / 5 = 0.2 = 20%

We follow 2000 newborns to measure development of respiratory infection in the first year

- Suppose 50 infants develop respiratory infection in first year of life

$$Risk = \frac{50}{2000} = 0.025 = 2.5\%$$

- The risk (probability) of developing a respiratory infection in the first year of life is ~ 2.5%

- 25 of 1000 infants in this population or 1 in 40 will develop infection in the first year of life.

We follow 2000 newborns to measure development of respiratory infection in the first year

- Suppose 50 infants develop respiratory infection in first year of life

$$Risk = \frac{50}{2000} = 0.025 = 2.5\%$$

Can we assume this risk applies to other populations?

RISK

- Must specify time period of observation because risk changes with time

- Must specify population because risk varies across populations

- Must specify region / place (for same reason)

RATE

What is a rate?

Rate

Generally speaking, a quantity per unit of time.

Example: The woman's heart rate was 60 beats per minute.

Example: The driver's rate of speed was 60 miles per hour.

Example: The man's pay rate was $60 per day.

Rate

- Can be expressed as *(a/T) where (a)* = cases and (*T*) involves a component of time

- Range: zero *(0)* to infinity *(∞)*

- Measures speed at which things happen

Incidence rate (IR)

$$IR = \frac{\# \text{ of new cases of disease (I)}}{\# \text{ of } \textit{person-time (PT)} \text{ units of observation}}$$

Incidence rate (IR)

$$IR = \frac{\# \text{ of new cases of disease (I)}}{\# \text{ of } \textit{person-time (PT)} \text{ units of observation}}$$

1 / 4.5 yrs. = 0.22 cases
per person per year

Rate Example

A measure of how fast something of interest happens in a population

Example: The number of new cases of Parkinson's disease that develops per 1,000 people

$$\frac{\#\text{ of new cases of Parkinson's disease}}{\text{Total \# years disease - free subjects observed}}\times 1000$$

Rate in Epidemiology

A measure of how fast something of interest happens in a population

Example: The number of new cases of Parkinson's disease that develops per 1,000 Knox residents over 30 years of age per year

$$\frac{\#\ of\ new\ cases\ of\ Parkinson's\ disease}{Total\ \#\ years\ disease\text{-}free\ subjects\ observed} \times 1000$$

Time, place and population should be specified

Rate in Epidemiology

A measure of how fast something of interest happens in a population

Example: The number of new cases of Parkinson's disease that develops per 1,000 Knox residents over 30 years of age per year

$$\frac{\# \text{ of new cases of Parkinson's disease}}{\text{Total } \# \text{ years disease - free subjects observed}} \times 1000$$

Time, place and population should be specified

Rate in Epidemiology

A measure of how fast something of interest happens in a population

Example: The number of new cases of Parkinson's disease that develops per 1,000 Knox residents over 30 years of age per year

$$\frac{\# \text{ of new cases of Parkinson's disease}}{\text{Total \# years disease - free subjects observed}} \times 1000$$

Time, place and population should be specified

Incidence Rate

- Measures how rapidly new cases develop during specified time period

- <u>NEW</u> cases per person-time

- Synonyms: incidence, incidence density, rate

Incidence Rate

$$IR = \frac{a}{T}$$

Where:

a = number of <u>new onset cases</u>
T = person-time at risk during study period
    (follow-up)

Person-time

- Accounts for all the time each person is in the population at risk

- The length of time for each person is called person-time

- Sum of person-times is called the total person-time at risk for the population
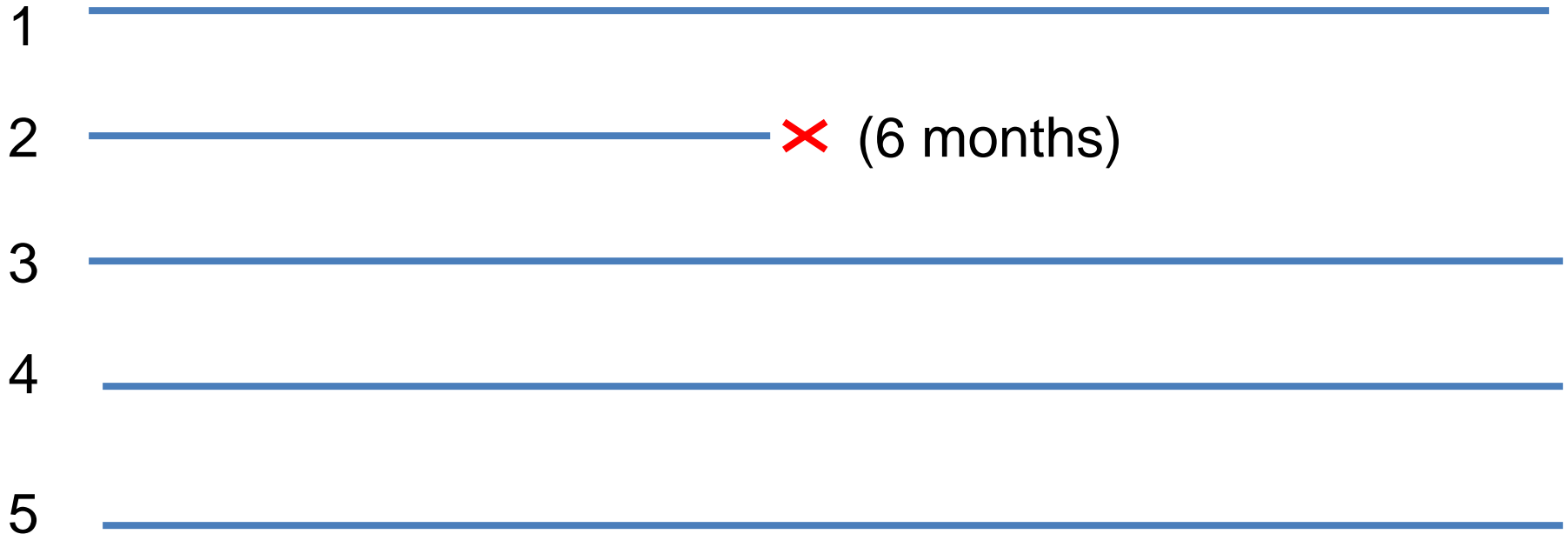
Person-time Assumption

- 100 persons followed 10 years = 1000 person years

- 1000 persons followed for 1 year = 1000 person years

Assumes rate is constant over different periods of time

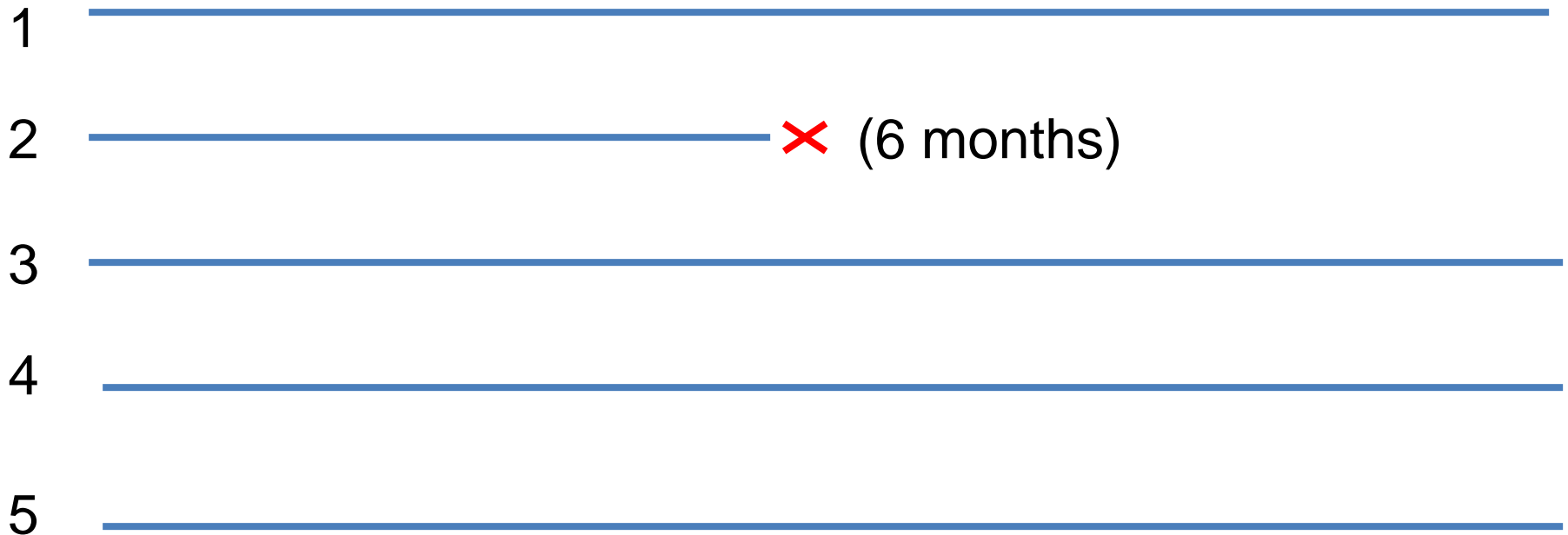# Example: Cohort Follow-up

1 ─────────────────────────────────────

2 ─────────────────────── ✖ (6 months)

3 ─────────────────────────────────────

4 ─────────────────────────────────────

5 ─────────────────────────────────────

→ Time (12 months) →

Cohort Follow-up

1 _____

2 _____ ✖ (6 months)

3 _____

4 _____

5 _____

1 new case / 4.5 person-years = 0.2222 cases per person per year (per "person-year")

Cohort Follow-up

1 ——————————————————————————

2 ———————————————— ✖ (6 months)

3 ——————————————————————————

4 ——————————————————————————

5 ——————————————————————————

1 new case / 54 person-months= 0.0185 cases per person per month (per person-month)

We follow 2000 newborns to measure development of respiratory infection in the first year

- Suppose 50 infants develop respiratory infection in first year of life

$$Risk = \frac{50}{2000} = 0.025 = 2.5\%$$

From this information, can we calculate a rate?

Example: Cohort Follow-up

1 ───

2 ─────────────────────── ✖ (6 months)

3 ───

4 ───

5 ───

⟶ Time (12 months) ⟶

PREVALENCE

What is prevalence?

Prevalence

- Proportion

- Not a rate – no time component in the calculation

- Measures proportion of <u>existing</u> disease in the population at a given time

- "Snapshot"

- Dimensionless, positive number (0 to 1)

Prevalence proportion

$$Prevalence = \frac{A}{N} = \frac{A}{A+B}$$

Where:

A = number of _existing_ cases
B = number of non-cases
N = total population

Example: Cohort Follow-up

1 ————————————————————————

2 ———————————————— ✖ (6 months)

3 ————————————————————————

4 ————————————————————————

5 ————————————————————————

——→ **Time (12 months)** ——→

Incidence vs. Prevalence

- New disease vs. existing disease

- Different implications in study design, analysis and interpretation

Incidence vs. Prevalence

- For each of the following, determine whether the statement requires measurement of **incidence** or **prevalence**

Incidence vs. Prevalence

- The Dept. of Education wants to organize an after-school program for children with learning disabilities and needs to know if there are sufficient children in need within the county to warrant such a service.
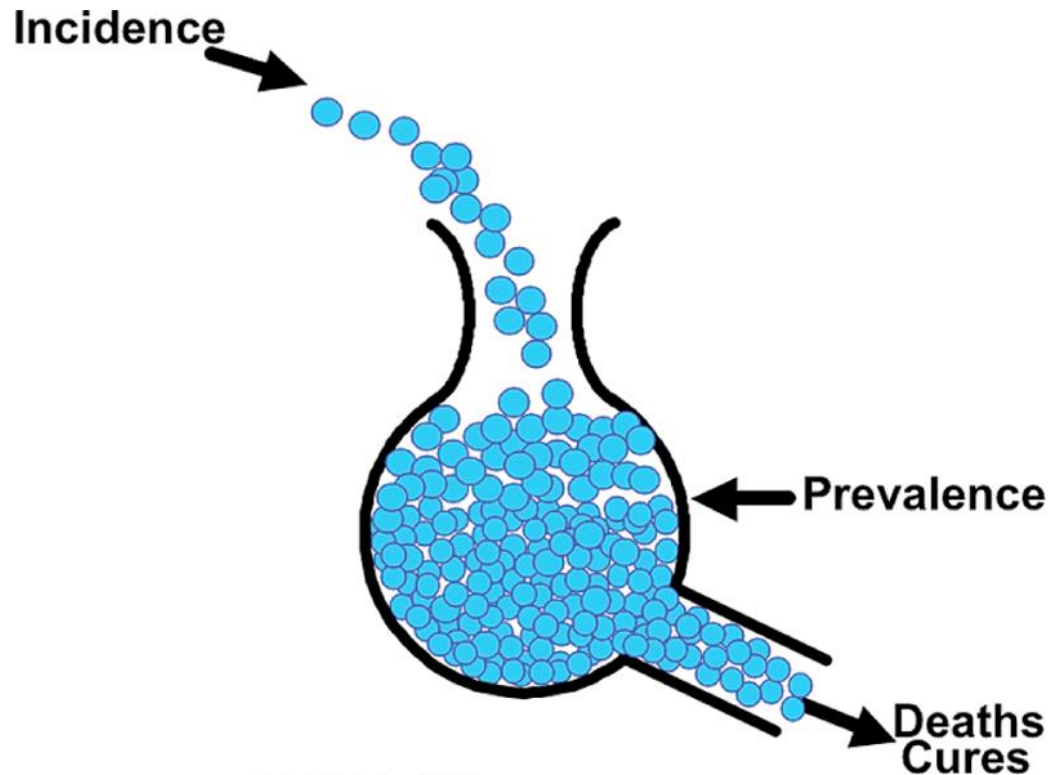
Incidence vs. Prevalence

- The university hospital epidemiologist wants to know the rate of tuberculin skin-test conversion (going from negative to positive indicates an infection with the causative agent of tuberculosis) occurring in third year medical students during 2001.

Incidence vs. Prevalence

- A medical school research team has developed a new drug which is purported to cure chronic schizophrenia and the team wants to study a large number of patients to determine the efficacy of the drug.
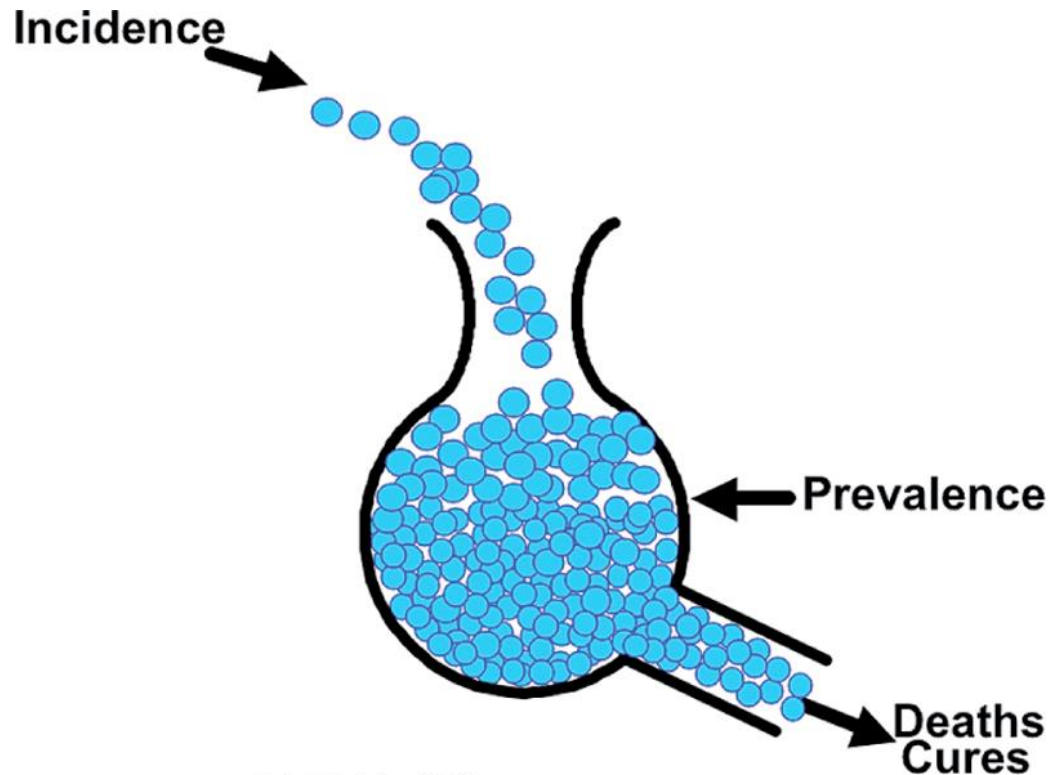
Incidence, Prevalence, Duration

Incidence

Prevalence

Deaths
Cures

- Prevalence increases as new cases added to the existing cases (i.e., incidence)

- Prevalence decreases as people are cured or die

- Prevalence = Incidence * Duration

HIV / AIDS

- If the prevalence of people with AIDS has risen over the past 10 years, is that necessarily a bad sign regarding public health progress?

Incidence, Prevalence, Duration

Incidence

Prevalence

Deaths
Cures

Gordis: Epidemiology, 4th Edition.
Copyright © 2008 by Saunders, an imprint of Elsevier, Inc. All rights reserved

- Prevalence increases as new cases added to the existing cases (i.e., incidence)

- Prevalence decreases as people are cured or die

- Prevalence = Incidence * Duration

Measures of Mortality

Mortality

- Measures the occurrence death
- Can be measured as a <u>proportion</u> or a <u>rate</u>
- Risk of death
- Rate of death

The statistical calculations for <u>risks</u> and <u>rates</u> for mortality are similar to those for disease morbidity

Case Fatality Rate

- This is <u>not a rate</u>, this is a proportion
- Proportion of deaths from a specific illness

$$\text{Case Fatality Rate} = \frac{a}{N}$$

Where:
a = Number of deaths from an illness
N = Number of people with that illness

What percentage of people diagnosed as having a disease die within a certain time after diagnosis?

Case-fatality rate

- Case-fatality – a measure of the severity of the disease

- Case-fatality – can be used to measure benefits of a new therapy
  - As therapy improves - the case-fatality rate would be expected to decline
  - e.g. AIDS deaths with the invention of new drugs

Proportionate Mortality

- Of all deaths, the proportion caused by a certain disease

- Can determine the leading causes of death

- Proportion of cause-specific death is dependent on all other causes of death

- This does not tell us the risk of dying from a disease

Proportionate Mortality

"One of every four deaths in the United States is due to cancer." -- CDC

25%

Other Mortality Rates

- Crude Mortality Rate
  - Includes all deaths, total population, in a time period

- Cause-Specific Mortality Rate
  - Includes deaths from a specific cause, total population, in a time period

- Age-Specific Mortality Rate
  - Includes all deaths in specific age group, population in the specific age group, in a time period

Age Adjustment

Should I Move?

Mortality rate in 1996
387 per 100,000/year

Mortality rate in 1996
1,026 per 100,000/year

Table 1. Vital Statistics for Alaska and Florida, 1996

Age (Years)	Alaska Deaths	Alaska Population	Florida Deaths	Florida Population
0-4	122	57,000	2,177	915,000
5-24	144	179,000	2,113	3,285,000
25-44	382	222,000	8,400	4,036,000
45-64	564	88,000	21,108	2,609,000
65-74	406	16,000	30,977	1,395,000
75+	582	7,000	71,483	1,038,000
Totals	2,200	569,000	136,258	13,278,000

$$\textit{Crude } \text{MR in AL} = \frac{2,200}{569,000} \times 10^5 = 387 \text{ per } 100,000$$

$$\textit{Crude } \text{MR in FL} = \frac{136,258}{13,278,000} \times 10^5 = 1,026 \text{ per } 100,000$$

Table 1. Vital Statistics for Alaska and Florida, 1996

Age (Years)	Alaska		Florida	
	Deaths	Population	Deaths	Population
0-4	122	57,000	2,177	915,000
5-24	144	179,000	2,113	3,285,000
25-44	382	222,000	8,400	4,036,000
45-64	564	88,000	21,108	2,609,000
65-74	406	16,000	30,977	1,395,000
75+	582	7,000	71,483	1,038,000
Totals	2,200	569,000	136,258	13,278,000

$$\textit{Crude } \text{MR in AL} = \frac{2,200}{569,000} \times 10^5 = 387 \text{ per } 100,000$$

$$\textit{Crude } \text{MR in FL} = \frac{136,258}{13,278,000} \times 10^5 = 1,026 \text{ per } 100,000$$

Table 1. Vital Statistics for Alaska and Florida, 1996

	Alaska		Florida	
Age (Years)	Deaths	Population	Deaths	Population
0-4	122	57,000	2,177	915,000
5-24	144	179,000	2,113	3,285,000
25-44	382	222,000	8,400	4,036,000
45-64	564	88,000	21,108	2,609,000
65-74	406	16,000	30,977	1,395,000
75+	582	7,000	71,483	1,038,000
Totals	2,200	569,000	136,258	13,278,000

$$\textit{Crude } \text{MR in AL} = \frac{2,200}{569,000} \times 10^5 = 387 \text{ per } 100,000$$

$$\textit{Crude } \text{MR in FL} = \frac{136,258}{13,278,000} \times 10^5 = 1,026 \text{ per } 100,000$$

Can we remove this confounding by age?

- Separate (stratify) the population into age groups and calculate rates for each age

  – Compare age-specific mortality rates

- If two different populations, adjust (standardize) the mortality rates of the two populations, taking into account the age structures

  – Results in comparable rates between populations or in the same population over time

Direct Standardization

- If the age composition of the populations were the same, would there be any differences in mortality rates?

- Direct age adjustment is used to remove the effects of age structure on mortality rates in <u>two different populations</u>

- Apply actual <u>age-specific rates</u> to a standard population (US population 2000)

Table 1. Vital Statistics for Alaska and Florida, 1996

Age (Years)	Alaska Deaths	Alaska Population	Florida Deaths	Florida Population
0-4	122	57,000	2,177	915,000
5-24	144	179,000	2,113	3,285,000
25-44	382	222,000	8,400	4,036,000
45-64	564	88,000	21,108	2,609,000
65-74	406	16,000	30,977	1,395,000
75+	582	7,000	71,483	1,038,000
Totals	2,200	569,000	136,258	13,278,000

Table 2. Age-specific mortality rates, Alaska and Florida, 1996

Stratum	Age (Years)	Age-specific mortality per 100,000 Alaska	Florida
1	0-4	214	238
2	5-24	80	64
3	25-44	172	208
4	45-64	640	808
5	65-74	2528	2221
6	75+	8314	6887

Table 1. Vital Statistics for Alaska and Florida, 1996

	Alaska		Florida	
Age (Years)	Deaths	Population	Deaths	Population
0-4	122	57,000	2,177	915,000
5-24	144	179,000	2,113	3,285,000
25-44	382	222,000	8,400	4,036,000
45-64	564	88,000	21,108	2,609,000
65-74	406	16,000	30,977	1,395,000
75+	582	7,000	71,483	1,038,000
Totals	2,200	569,000	136,258	13,278,000

Table 2. Age-specific mortality rates, Alaska and Florida, 1996

		Age-specific mortality per 100,000	
Stratum	Age (Years)	Alaska	Florida
1	0-4	214	238
2	5-24	80	64
3	25-44	172	208
4	45-64	640	808
5	65-74	2528	2221
6	75+	8314	6887

Hypothetical:

- If the deaths in the US were based on Alaska's age-specific mortality rates…

- If the deaths in the US were based on Florida's age-specific mortality rates…

$$\textbf{\textit{Crude}} \text{ MR in AL} = \frac{2,200}{569,000} \times 10^5 = 387 \text{ per } 100,000$$

$$\textbf{\textit{Crude}} \text{ MR in FL} = \frac{136,258}{13,278,000} \times 10^5 = 1,026 \text{ per } 100,000$$

$$\text{Age - adjusted MR in AK} = \frac{\sum_{i=1}^{k} N_i r_i}{\sum_{i=1}^{k} N_i} = \frac{232,421,256,000}{265,406,000} = 875 \text{ per } 100,000$$

$$\text{Age} - \text{adjusted MR in FL} = \frac{\sum_{i=1}^{k} N_i r_i}{\sum_{i=1}^{k} N_i} = \frac{216,324,617,000}{265,406,000} = 815 \text{ per } 100,000$$

Indirect Standardization

- When age-specific rates are not available – use age-specific mortality rates from the general population to calculate expected number of deaths

Standardized mortality ratios (SMR)

= observed deaths/ expected deaths

Study Design

Here is a good time for a break!!!

Study Design

- Experimental studies (Clinical Trial, Randomized Controlled Trial)
- Observational studies
  - Cohort
  - Case-control
  - Cross-sectional
  - Ecological

Experimental studies are characterized by:

- Manipulation of the exposure by the researcher

Randomized Controlled Trials

- A randomized controlled trial is a type of <u>experimental</u> research design for comparing different treatments, in which the assignment of treatments to patients is made by a random mechanism.

- Customary to present table of patient characteristics to show that the randomization resulted in a balance in patient characteristics.

Randomized Controlled Trials
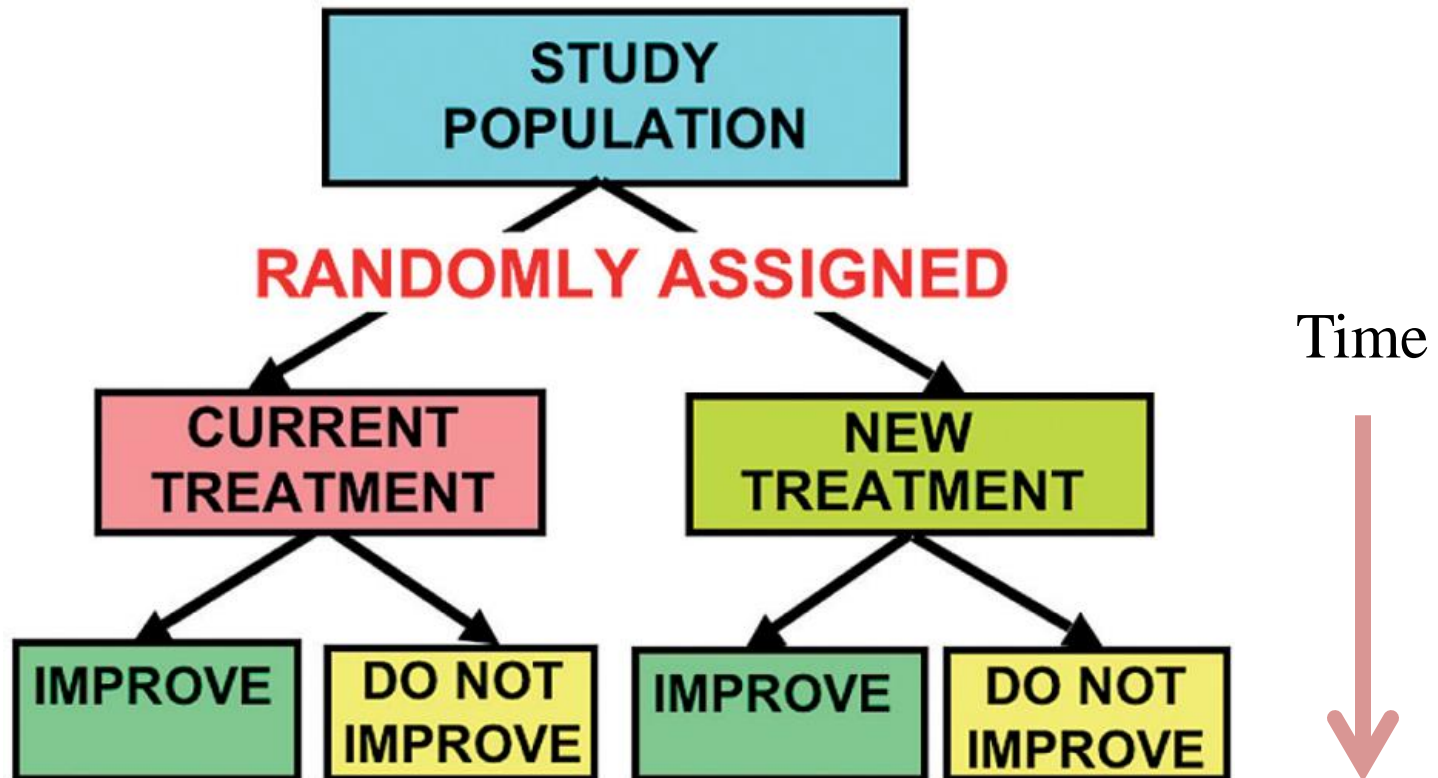
Gordis: Epidemiology, 4th Edition.
Copyright © 2008 by Saunders, an imprint of Elsevier, Inc. All rights reserved

Type of Study?

Methods: Fifteen patients were randomized to receive a preoperative beverage with high (125 mg/ml) or low (25 mg/ml) carbohydrate content. Postoperative cognitive ability was subsequently measured.

Type of Study?

Methods: Fifteen patients were randomized to receive a preoperative beverage with high (125 mg/ml) or low (25 mg/ml) carbohydrate content. Postoperative cognitive ability was subsequently measured.

CLINICAL TRIAL

Type of Study?

Methods: Ninety eight individuals 18-65 years of age were randomized to placebo or sertraline 25 mg/day for 2 days, followed by 50 mg from day 3 to 90, and buspirone 5 mg three times a day for 7 days, and 10 mg from day 8 to 90.

Type of Study?

Methods: Ninety eight individuals 18-65 years of age were randomized to placebo or sertraline 25 mg/day for 2 days, followed by 50 mg from day 3 to 90, and buspirone 5 mg three times a day for 7 days, and 10 mg from day 8 to 90.

CLINICAL TRIAL

Some Limitations of a Clinical Trial

1. Ethical considerations

2. Select population

3. Duration

4. Adherence / compliance

Use of "Blinding"

- Important when knowing treatment could influence the interpretation of results

- Placebo- ensure control and treatment group have same "experience"

Treat

- Blinding helps ensure that bias is avoided

  – Single-blind: <u>patient</u> does not know what treatment they are receiving

  – Double–blind: <u>patient and investigator</u> do not know what treatment (cannot be used for some treatments, e.g. surgery)

It All Comes Down to…

- Obtaining groups that are comparable for everything except the treatment…

- So that differences in outcome can fairly be ascribed only to the difference between the groups (i.e., to the treatment).

A Clinical Trial…

- Can be viewed as a type of <u>prospective cohort study</u>

- It involves active follow-up of a group of people and determines their outcomes (disease, cure, side-effects, etc.)

- However, a cohort study typically referred to is OBSERVATIONAL (no assigned treatments)

PROSPECTIVE COHORT STUDY

Example: Cohort Follow-up

1

2 ✕ (6 months)

3

4

5
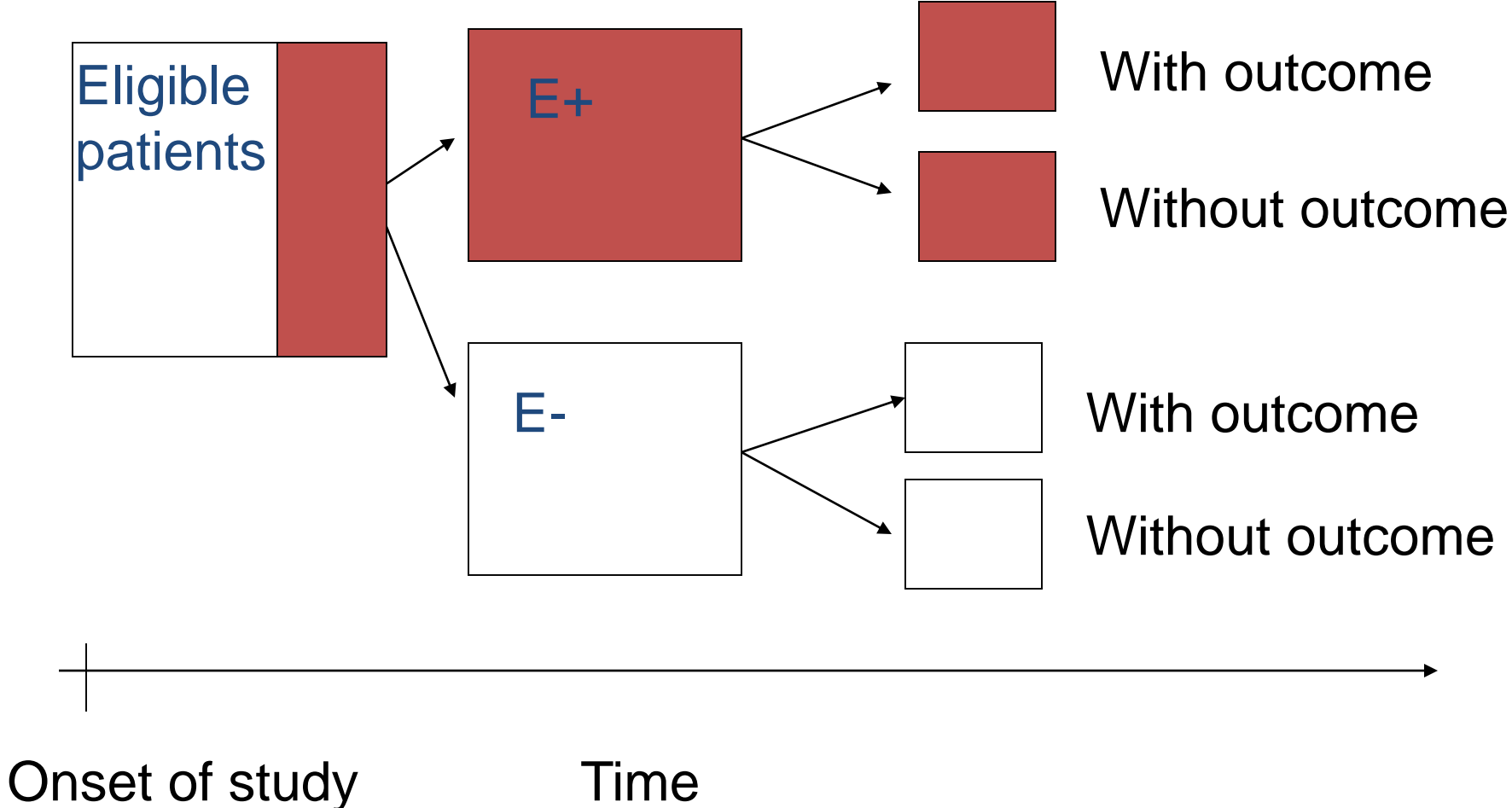
→ **Time (12 months)** →

Example: Cohort Follow-up

Exposed

UN-exposed

Time (12 months)

Cohort Studies

- <u>Definition</u>: groups, defined on the basis of some characteristics (often exposure and non-exposure) are prospectively followed to see whether an outcome of interest occurs

- <u>Comparison of interest:</u> Compare the proportion of persons with the disease in the exposed group to the proportion with the disease in the unexposed group (or compare rates)

- <u>Motivation:</u> If the exposure is associated with the disease, we expect that the proportion of persons with the disease in the exposed group (or rate of disease) will be greater than the proportion with disease in the unexposed group.

Cohort Studies

Prospective

Retrospective

now

past

future

now

Exposed

Diseased

Not diseased

Not Exposed

Diseased

Not diseased
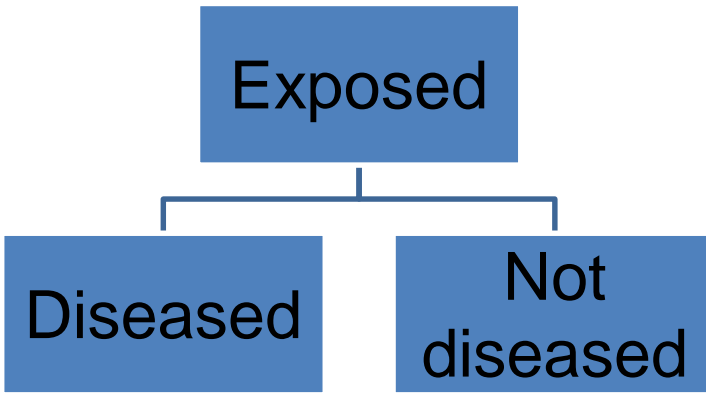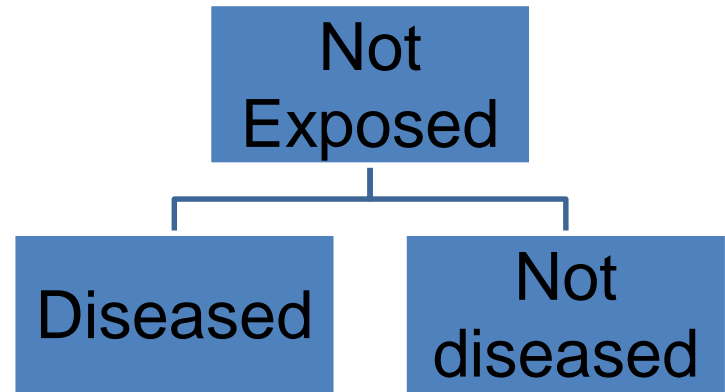
Prospective cohort studies

- Define sample free of the disease/outcome of interest, measure the exposure and classify to exposed vs unexposed at "baseline," then follow up to ascertain outcome

- Measure the proportion of outcome between the exposed and unexposed (Risk Ratio or Relative Risk) or rate (Rate Ratio)

Retrospective cohort studies

- Synonyms: **historical cohort study**, historical prospective study, non-concurrent prospective study

- Do not design retrospective cohort studies *a priori* – question always in retrospect

- Exposures and Outcomes have already occurred - data on the relevant exposures and outcomes already have been collected

Cohort study strengths

- May be used to define incidence / natural history

- Known temporal sequence

- Efficient in investigating <u>rare exposures</u>

- Permits study of multiple exposures AND outcomes

Some cohort study limitations

- Expensive

- Slow to find answers (time-consuming)

- Associations may be due to confounding (true with any observational study)

- Exposures assessed at baseline may be incomplete

- Disease with long pre-clinical phase may not be detected

- Sensitive to follow-up bias (loss of diseased subjects)

Type of Study?

Methods: Cigarette smoking data were collected on all household members during two private censuses in Washington County, Maryland. These two groups were followed up, one from 1963-1978 and the other from 1975-1994 for first-time diagnoses of rectal cancer.

Type of Study?

Methods: Cigarette smoking data were collected on all household members during two private censuses in Washington County, Maryland. These two groups were followed up, one from 1963-1978 and the other from 1975-1994 for first-time diagnoses of rectal cancer.

COHORT STUDY

Type of Study?

Methods: Ninety eight individuals 18-65 years of age were randomized to placebo or sertraline 25 mg/day for 2 days, followed by 50 mg from day 3 to 90, and buspirone 5 mg three times a day for 7 days, and 10 mg from day 8 to 90.

Type of Study?

Methods: Ninety eight individuals 18-65 years of age were randomized to placebo or sertraline 25 mg/day for 2 days, followed by 50 mg from day 3 to 90, and buspirone 5 mg three times a day for 7 days, and 10 mg from day 8 to 90.

CLINICAL TRIAL

CASE-CONTROL STUDIES

Case-control Studies

- <u>Definition</u>: compare various characteristics (past exposure) for cases (subjects with disease) to those of controls (subjects without the disease)

- <u>Comparison of interest:</u> Compare the proportion with the exposure in the cases to the proportion with the exposure in the control group.

- <u>Motivation:</u> If the exposure is associated with the disease, we expect that the proportion of persons with the exposure in the cases will be greater than the proportion with the exposure in the control group.

Case-control Studies

```
┌─────────────────┐                    ┌─────────────────┐
│   Cases with    │                    │ Controls without│
│    disease      │                    │    disease      │
└────────┬────────┘                    └────────┬────────┘
    ┌────┴────┐                            ┌─────┴────┐
┌───────────┐ ┌──────────────────┐   ┌───────────┐ ┌──────────────────┐
│Exposed in │ │Not Exposed in    │   │Exposed in │ │Not Exposed in    │
│   past    │ │     past         │   │   past    │ │     past         │
└───────────┘ └──────────────────┘   └───────────┘ └──────────────────┘
```

Case-control Studies

Present

```
            Cases with                          Controls without
             disease                                disease

   Exposed in past   Not Exposed in past   Exposed in past   Not Exposed in past
```

Past

Case-Control Studies (compared with cohort)

- More efficient for rare diseases

- Can evaluate multiple exposures

- Less expensive

- Can get answers more quickly

- Challenges of control selection

- Challenges of retrospective exposure assessment

Nested Case-Control Studies

- A case control study nested in a cohort study
- Controls selected either at baseline (case-cohort) or at the time the case occurs (nested)
- Advantage
  - Data on exposure are obtained before disease develops
  - Possibility of recall bias is thus eliminated.
  - Less expensive than expanding the analysis to include the entire cohort
  - Here the OR is a statistically unbiased estimate of relative risk

Type of Study?

Methods: Danish women with a first time MS discharge diagnosis from a neurological department at most 40 years old during the period 1998-2005, and an age and geographically matched healthy group. Information on number of full term pregnancies was elicited.

Type of Study?

Methods: Danish women with a first time MS discharge diagnosis from a neurological department at most 40 years old during the period 1998-2005, and an age and geographically matched healthy group. Information on number of full term pregnancies was elicited.

CASE-CONTROL STUDY

CASE-CONTROL STUDIES

Cross-Sectional Studies

- Prevalence studies
- All measurements of exposure and outcome are made simultaneously (snapshot)
- Disease proportions are determined and compared among those with or without the exposure or at varying level of the exposure
- Examine association – determination of associations with outcomes; generates hypotheses that are the basis for further studies
- Most appropriate for studying the associations between chronic diseases and and chronic exposure
- Sometimes useful for common acute diseases of short duration

Cross-Sectional Studies
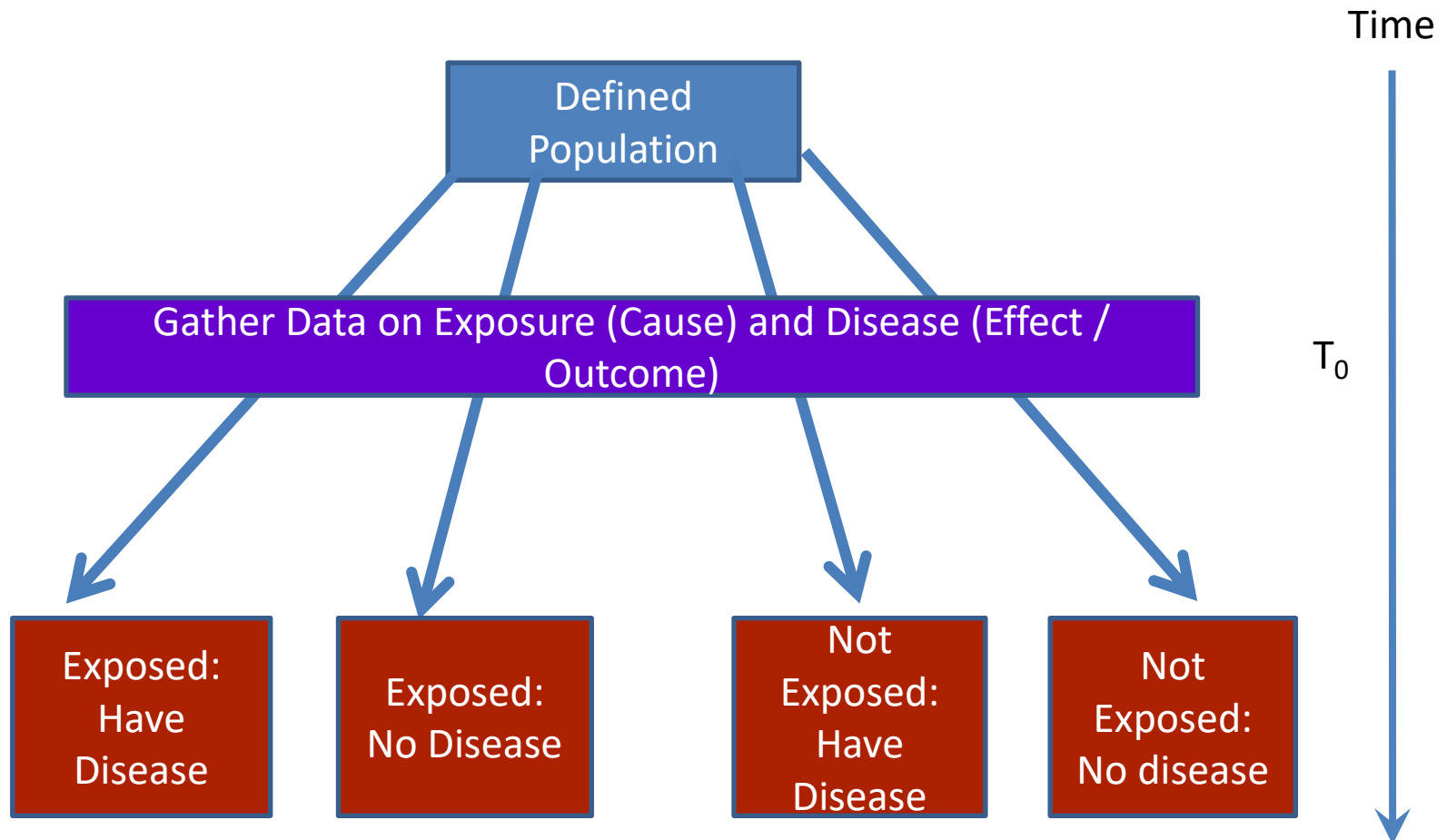
Cross-Sectional Studies
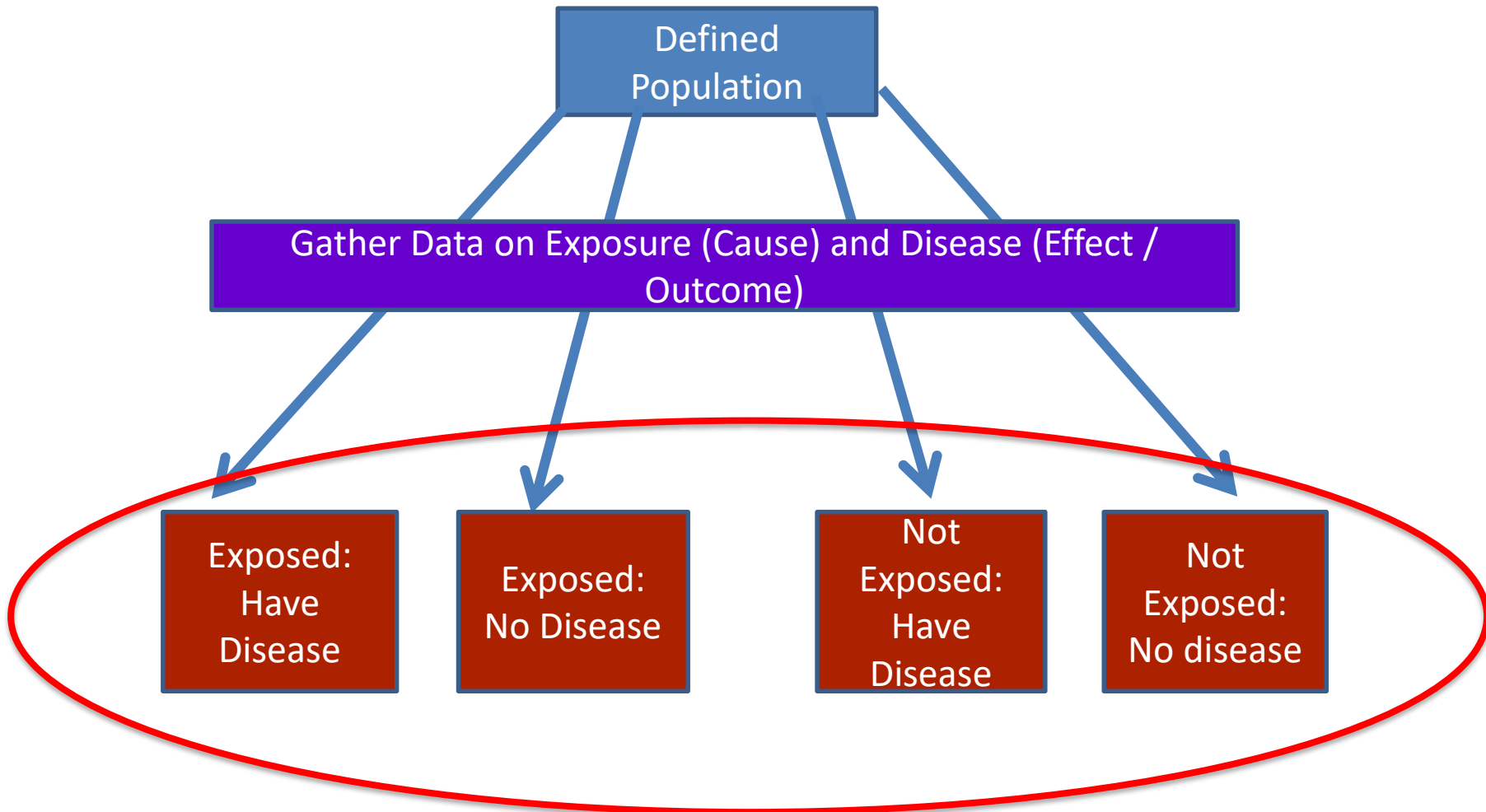
Ecological

- The unit of observation is the population or community

- Disease rates and exposures are measured in each of a series of populations

- Disease and exposure information may be abstracted from published statistics and therefore does not require expensive or time consuming data collection

Type of Study?

Methods: Two hundred children aged 9 to 12 years were recruited to evaluate the effect of body mass on foot structure. In addition to BMI, three reliable anthropometric measures were recorded: foot length, forefoot width, and navicular height.

Type of Study?

Methods: Two hundred children aged 9 to 12 years were recruited to evaluate the effect of body mass on foot structure. In addition to BMI, three reliable anthropometric measures were recorded: foot length, forefoot width, and navicular height.

CROSS SECTIONAL STUDY

Ecological Studies

CORRELATIONAL / ECOLOGICAL STUDIES

- Measures that represent characteristics of the entire population are used to describe disease in relation to some factor of interest.

- Presence of suspected risk factor can be measured in different populations and compared with the incidence of a particular disease.

Cancer Rates by Country
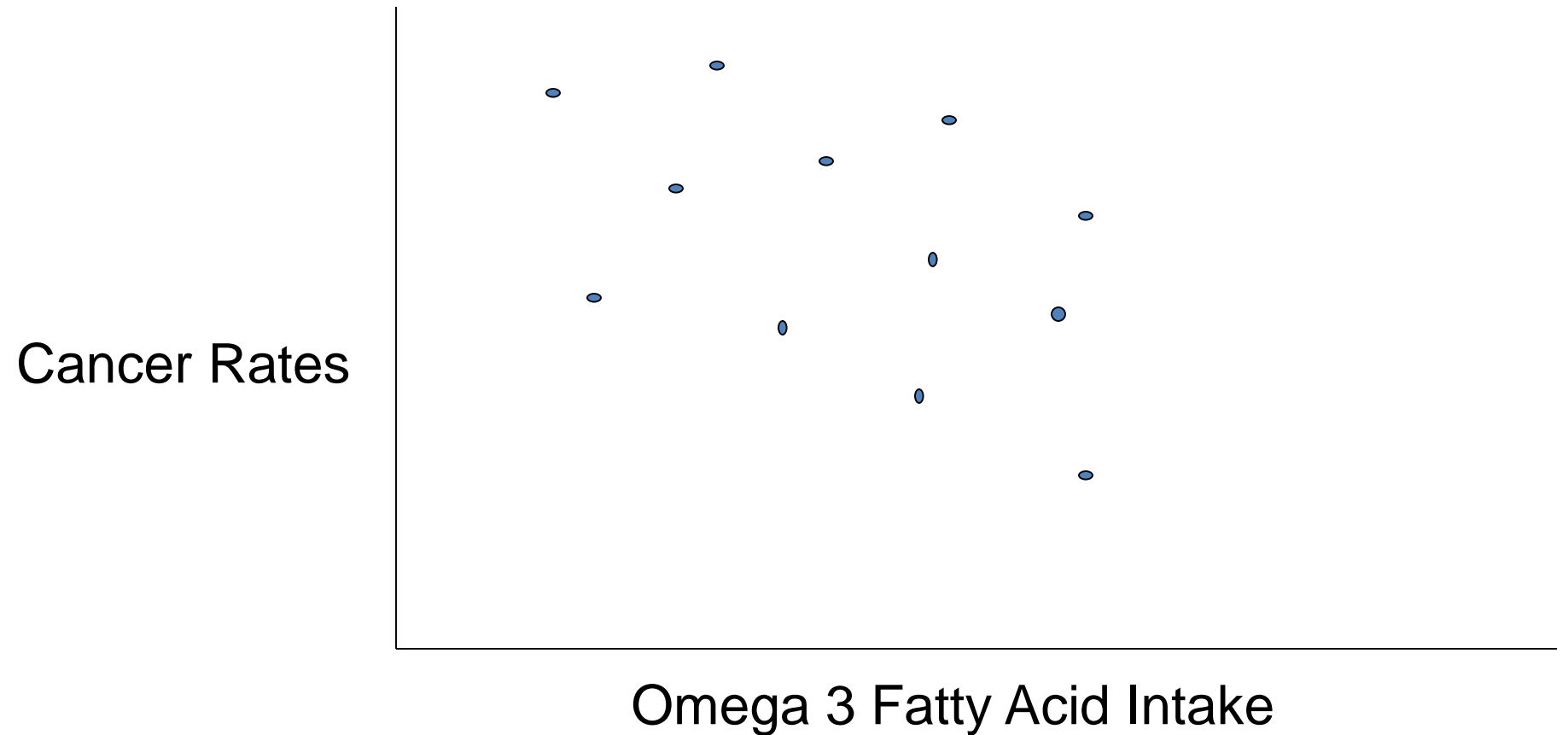
Cancer Rates

Omega 3 Fatty Acid Intake

Cancer Rates by Country

Cancer Rates by Country

Cancer Rates by Country

Cancer Rates

Omega 3 Fatty Acid Intake
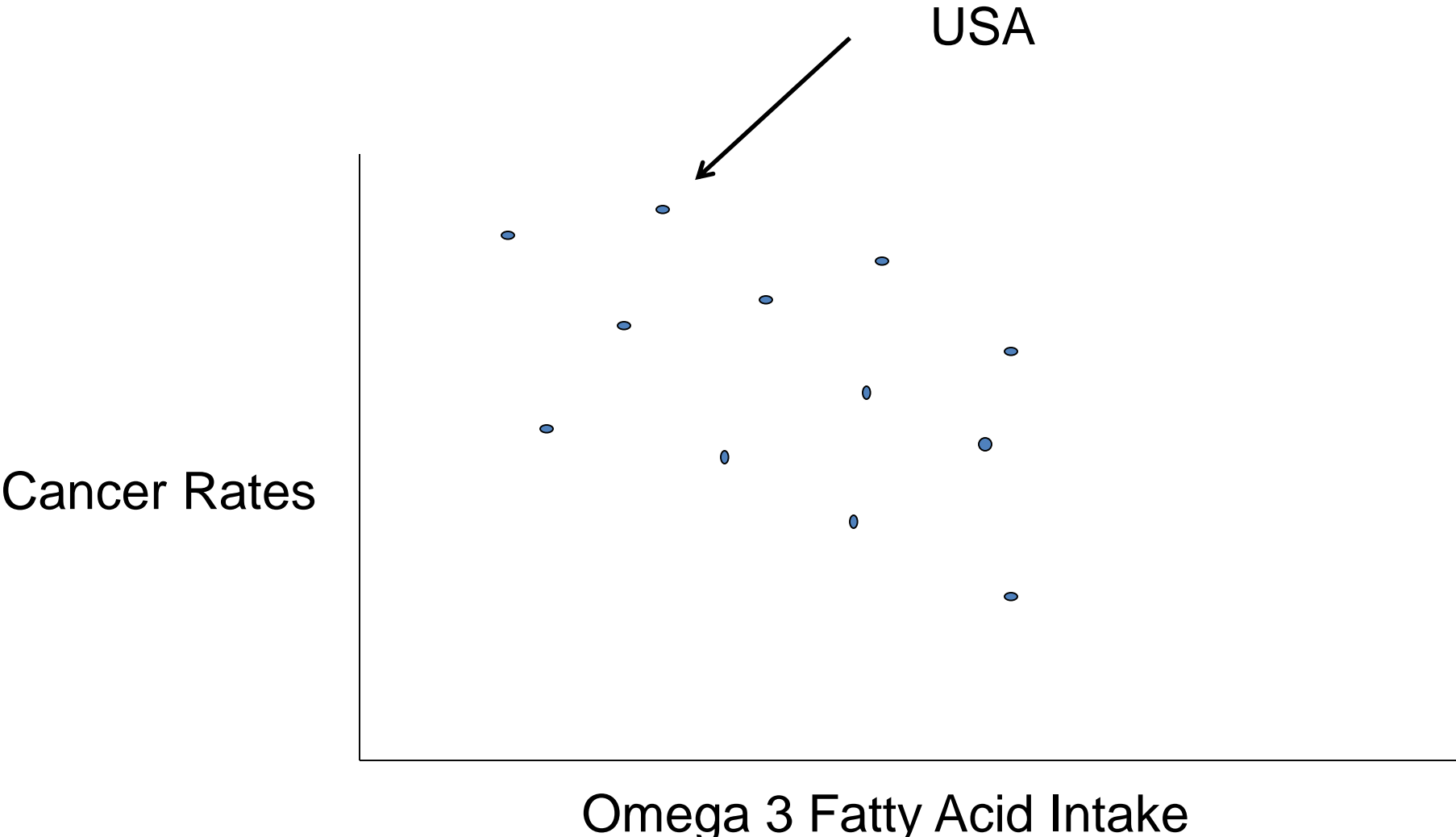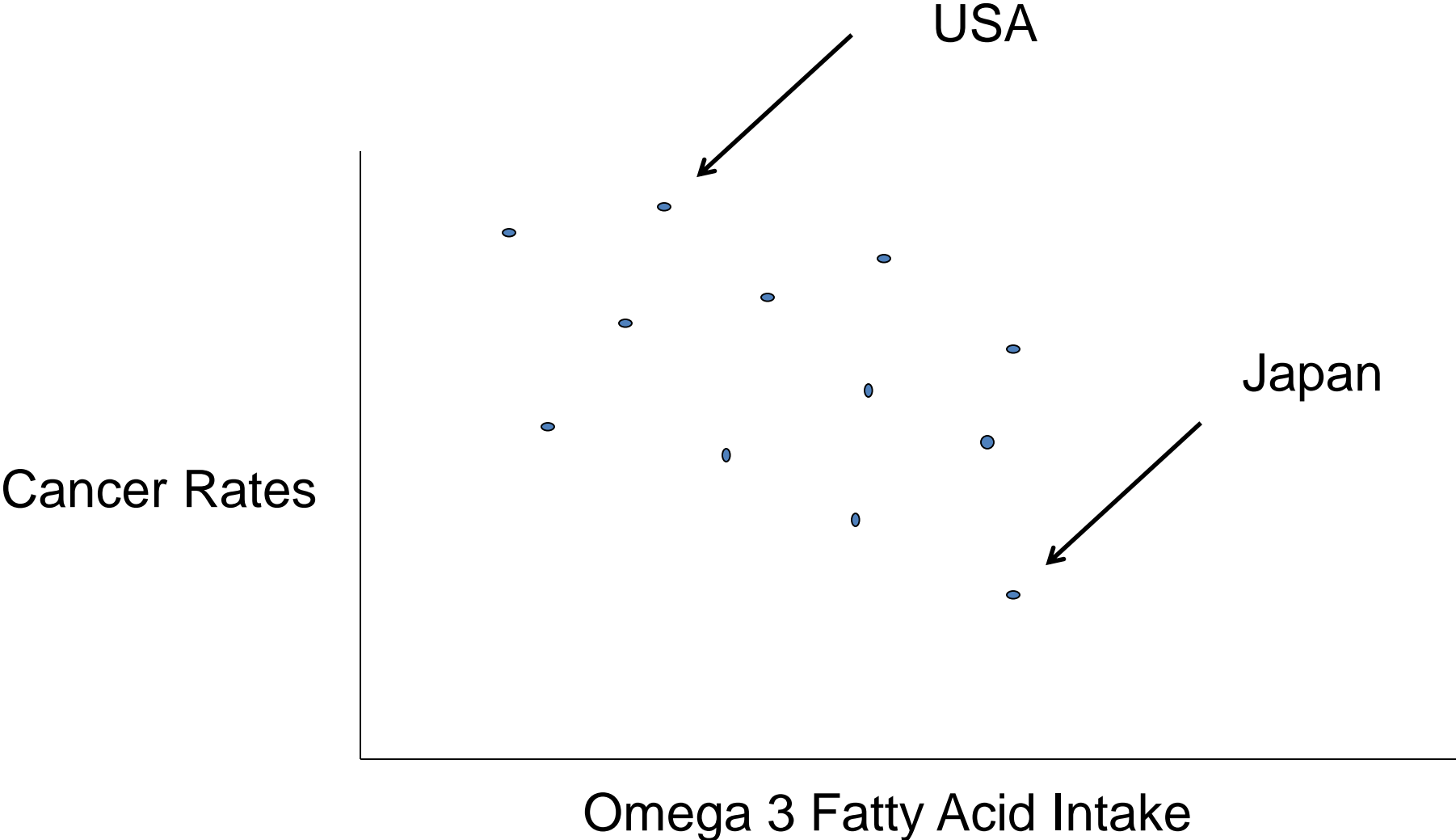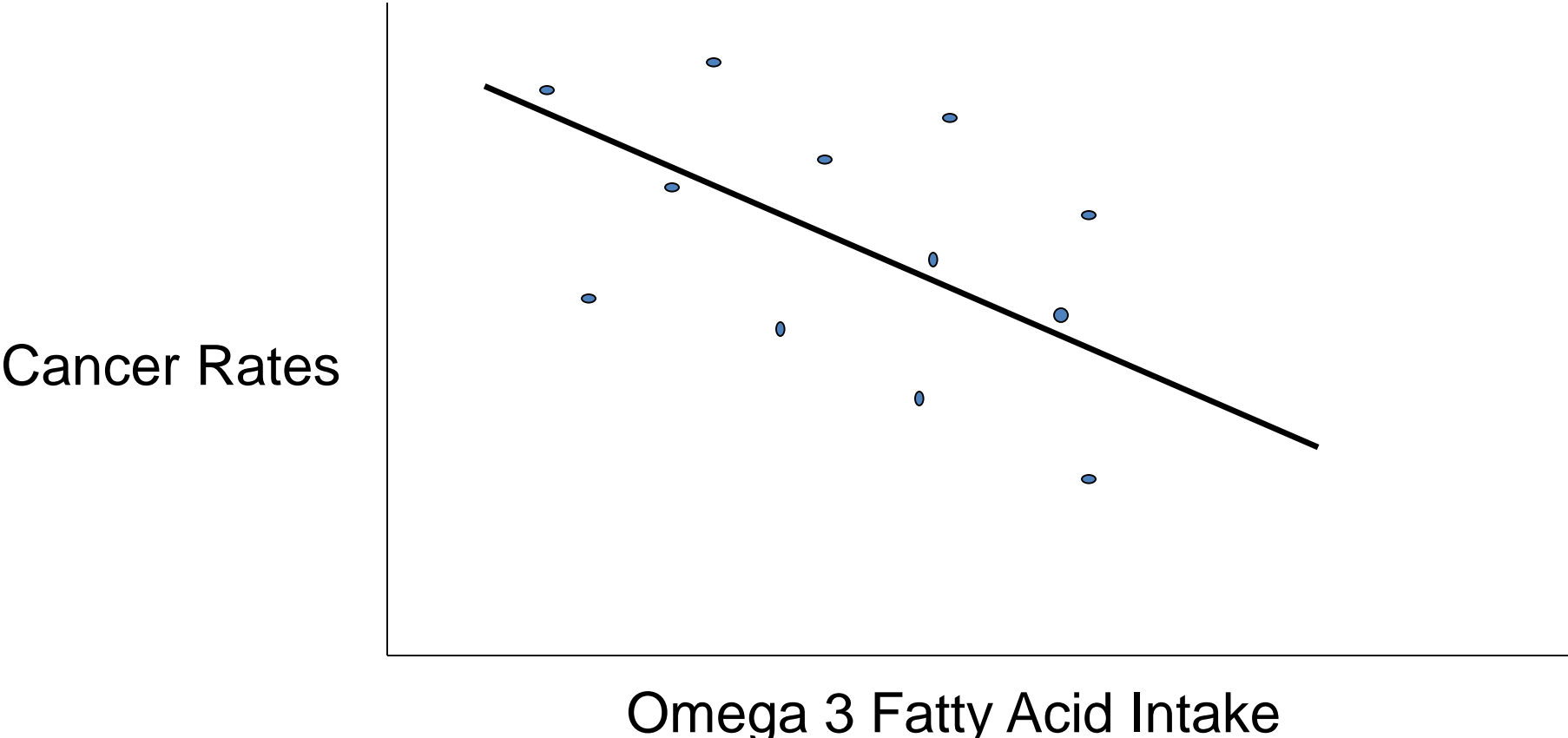
ECOLOGICAL STUDIES - LIMITATIONS

• Hypothesis generating- cannot establish causal relationship

• Unable to control the effects of potential confounding factors.

• Unable to link exposure with disease in a particular individual – *ecologic fallacy.*

ECOLOGIC FALLACY

Suspected risk factor and disease are associated at the population level, but not at the individual level.

Type of Study?

Methods: During the period 1995 to 2000, 81,132 lung cancer cases were reported in Texas. Researchers examined the association of metal air releases with the average annual age-adjusted primary and non-small cell lung cancer rates in the 254 Texas counties.

Type of Study?

Methods: During the period 1995 to 2000, 81,132 lung cancer cases were reported in Texas. Researchers examined the association of metal air releases with the average annual age-adjusted primary and non-small cell lung cancer rates in the 254 Texas counties.

CORRELATIONAL / ECOLOGIC STUDY

Type of Study?

Methods: A survey was performed in nine European countries, i.e. Austria, Belgium, Denmark, Iceland, the Netherlands, Norway, Portugal, Spain and Sweden, from October-December 2003, as a part of the Pro Children study. Data on usual intake of fruit and vegetables, and related correlates were collected by means of a self-administered questionnaire among 11-year-old school children.

Type of Study?

Methods: A survey was performed in nine European countries, i.e. Austria, Belgium, Denmark, Iceland, the Netherlands, Norway, Portugal, Spain and Sweden, from October-December 2003, as a part of the Pro Children study. Data on usual intake of fruit and vegetables, and related correlates were collected by means of a self-administered questionnaire among 11-year-old school children.

CROSS SECTIONAL STUDY

Measures of Association

Measures of Association

In general:

Cohort studies:

1. **Risk / Rate / Hazard Ratios (RR)**
2. Disease Odds Ratios (DOR or OR)

Case-control studies:

1. Exposure **Odds** Ratios (EOR or OR)
2. Risk Ratios (RR)

RELATIVE RISK

An estimate of the magnitude of an association between exposure and disease.

Indicates the likelihood of developing the disease for the exposed group relative to those who are not exposed.

ANALYSIS OF A COHORT STUDY

	Disease	No Disease
Exp +	24	50
Exp -	315	876

Relative risk = A / (A + B) = Risk ratio = (24/74) **= 1.23**
C / (C + D) / **(315/1191**

Null Hypothesis

- The risk of the outcome in the exposed persons is equal to the risk of the outcome in the unexposed persons.

Ho: RR = 1.0

Two-sided Alternate Hypothesis

- The risk of the outcome in the exposed persons is not equal to the risk of the outcome in the unexposed persons.

$$Ha: RR \neq 1.0$$

One-sided Alternate Hypothesis

- The risk of the outcome in the exposed persons is greater than (or less than) the risk of the outcome in the unexposed.

$$Ha: RR > 1.0$$

or

$$Ha: RR < 1.0$$

INTERPRETING RR

Relative risks between 1.0 and 2.0

RR – 1.0 = % increased risk

RR = 1.50

1.50 – 1.0 = 0.50 = 50% increased risk of outcome given exposure

INTERPRETING RR

Relative risks > 2.0

RR number = number of times increased risk

RR = 3.0 = 3 times increased risk of outcome given exposure

INTERPRETING RR

Relative risks < 1.0

1.0 – RR = % decreased risk

RR = 0.75

1.0 – 0.75 = 0.25 = 25% less risk of outcome given exposure

ANALYSIS OF A COHORT STUDY, Cont.

2. Evaluating the precision of the RR:

•The 95% confidence interval (CI) is a measure of precision.

Lower limit of 95% CI = RR x e $(-1.96 \ \sqrt{\text{var lnRR}})$
Upper limit of 95% CI = RR x e $(+1.96 \ \sqrt{\text{var lnRR}})$

•95% CI = (lower limit – upper limit)

Rate Ratio

Rate ratio = $\dfrac{\text{rate of outcome in E+}}{\text{rate of outcome in E-}}$ = RR

Interpretation: The rate of outcome in E+ is X times the rate of the outcome in the E-.

Rate Ratio

Rate ratio = 50/20 = 2.5

Interpretation: The rate of outcome in E+ is 2.5 times the rate of the outcome in the E-.

Analysis of a Case-Control Study

Odds

- Odds are another way of representing a probability

- The odds is the ratio of probability that the event of interest occurs to the probability that it does not.

- The odds are often estimated by the ratio of the number of times that the event occurs to the number of times that it does not.

Odds Ratios

- Relative risk requires an estimate of the incidence of the disease

- For **most** case control studies, we do not know the incidence of disease because we determine the number of cases and controls when the study is designed – we really don't know the underlying cohort

- For case control studies, generally use the odds ratio (OR)

Odds Ratio

$$\frac{p}{1-p} = \text{odds}$$

$$\frac{\left(\dfrac{p_1}{1-p_1}\right)}{\left(\dfrac{p_2}{1-p_2}\right)} = \frac{p_1(1-p_2)}{(1-p_1)p_2} = \text{odds ratio}$$

Odds Ratio Example

	CHD Cases	Controls	Total
Smokers	112	176	288
Nonsmokers	88	224	312
Total	200	400	600

- Case control study of 200 CHD cases and 400 controls to examine association of smoking with CHD

  (Note: now we are examining the probability of exposure)

- What is the odds of smoking among CHD cases?

  112/88=1.27

- What is the odds of smoking among controls?

  176/224= 0.79

Odds Ratio Example

	CHD Cases	Controls	Total
Smokers	112	176	288
Nonsmokers	88	224	312
Total	200	400	600

OR = 1.27 / 0.79 = 1.61

Interpretation:

The Cases' odds of exposure is 1.6 times that of controls.

Odds Ratio Example

	CHD Cases	Controls	Total
Smokers	112	176	288
Nonsmokers	88	224	312
Total	200	400	600

Another simple calculation:

(112 x 224) / (88 x 176) = 1.61

155

Odds ratio

- Odds ratio = <u>odds of exposure in case</u>
  odds of exposure in controls

OR=1 exposure is not associated with the disease
OR>1 exposure is positively associated with the disease
OR<1 exposure is negatively associated with the disease

Odds Ratio

- **Interpretation:** The odds of exposure among the diseased is X times higher/lower than the odds of exposure among the non-diseased.

OR vs. RR

- **OR:** The odds of <u>exposure</u> among the <u>diseased</u> is X times higher/lower than the odds of exposure among the non-diseased.

- **RR:** The risk of <u>disease</u> among the <u>exposed</u> is X times higher/lower than the risk of disease among unexposed.

Odds Ratios vs. Relative Risks

Odds ratio can be used to estimate the relative risk when in a case control study when:

1. Cases are representative of people with the disease in the population with respect to history of exposure AND

2. The controls are representative of the entire study population ("source population") with respect to history of exposure AND

3. The disease is rare

Odds ratio estimates relative risk when disease is rare

- When the disease is rare, the number of people with the disease (a and c) is small so that a+b≈b and c+d≈d

$$RR = \frac{a/(a+b)}{c/(c+d)} \approx \frac{a/b}{c/d} = \frac{ad}{bc} = OR$$

Odds Ratios for matched case control studies

- Often, cases are matched with a control based on age, sex, etc.

- For a matched study, describe the results for each pair

- Concordant pairs: both case and control exposed or both not exposed

- Discordant pairs: Case exposed/control unexposed or case unexposed/control exposed

Odds Ratios for matched case control studies

Controls

	Exposed	Unexposed
Exposed	a	b
Unexposed	c	d

Cases

OR is based on the discordant pairs:

OR = b/c

Measures of IMPACT

Risk

- RR and OR measure strength of the association

- How much of the disease can be attributed to the exposure? How much of the CHD risk experienced by smokers can be attributed to smoking?

- OR and RR do not address this.

Risk Difference

- Most often referred to as "attributable risk"
  - Refers to the amount of risk attributable to the exposure of interest
  - For example, in the birth cohort analysis, where exposure = prenatal care in the first 5 months

    $RD = R_1 - R_0 =$ Excess risk of preterm birth

    attributable to prenatal care

Absolute Excess Measures
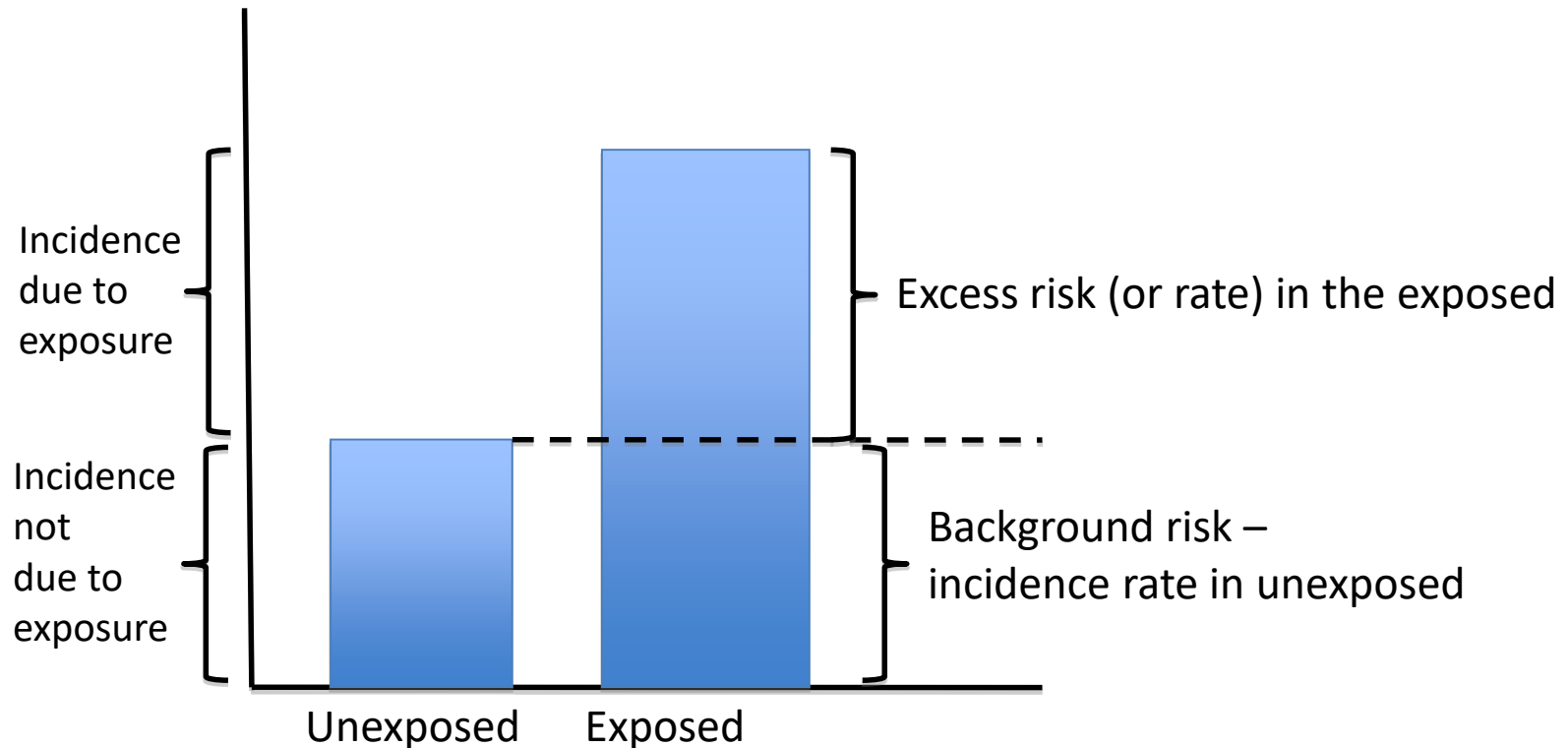
Incidence proportion (or rate)

Incidence due to exposure

Excess risk (or rate) in the exposed

Incidence not due to exposure

Background risk – incidence rate in unexposed

Unexposed   Exposed

If E is thought to cause D: Among persons exposed to E, what amount of the incidence of D is E responsible for?

Absolute Excess Measures

Incidence proportion (or rate)

Background Risk

Incidence due to exposure

Excess risk (or rate) in the exposed

Incidence not due to exposure

Background risk – incidence rate in unexposed
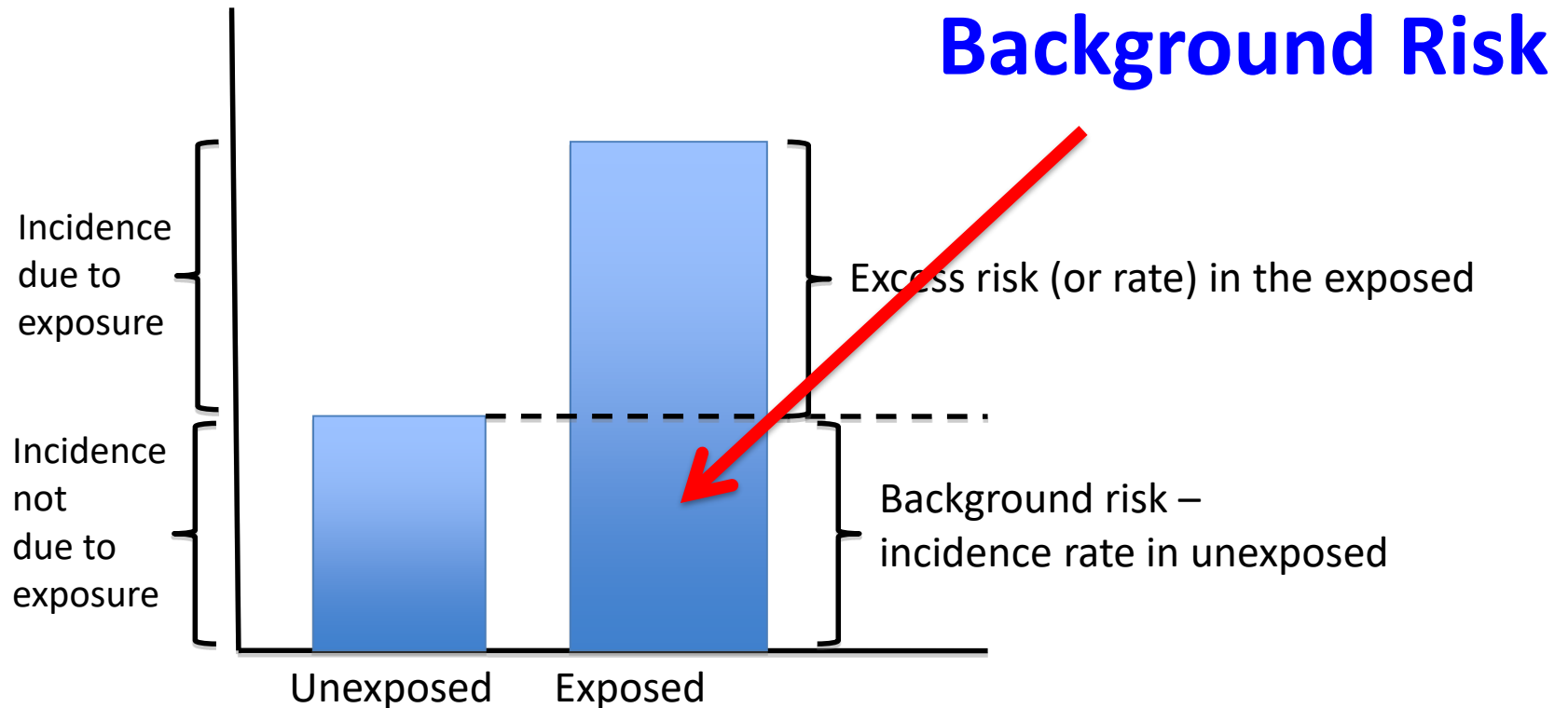
Unexposed   Exposed

If E is thought to cause D: Among persons exposed to E, what amount of the incidence of D is E responsible for?

Absolute Excess Measures

Incidence proportion (or rate)

Excess Risk
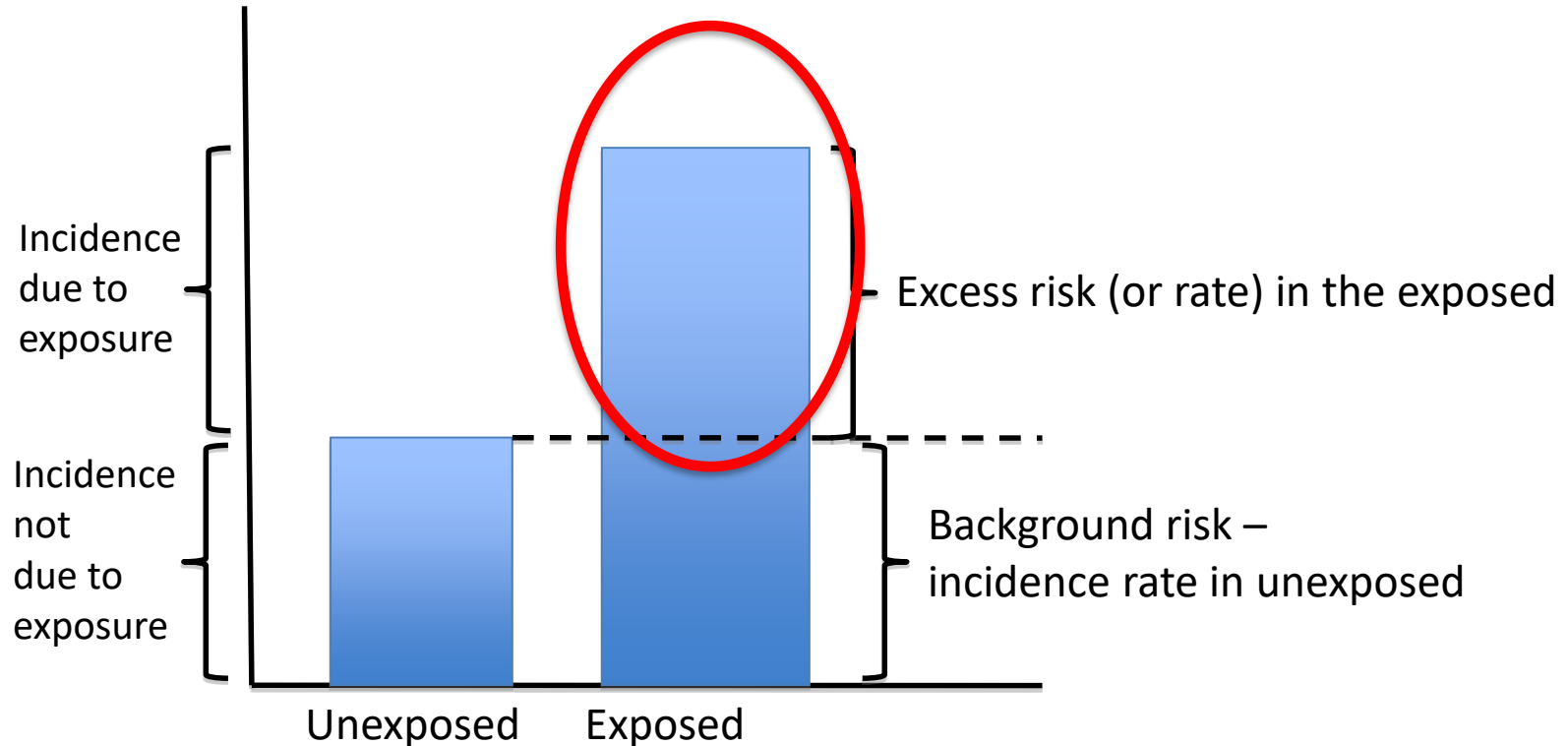
Incidence due to exposure

Excess risk (or rate) in the exposed

Incidence not due to exposure

Background risk – incidence rate in unexposed

Unexposed   Exposed

If E is thought to cause D: Among persons exposed to E, what amount of the incidence of D is E responsible for?

Example

Sleeping Position and Crib Death

	Crib Death		
Usual sleeping position	YES	NO	TOTAL
Prone	9	837	846
Other	6	1755	1761
Total	15	2592	2607

1-year risk prone = 9/846 = 10.64 per 1000
1-year risk other = 6/1761 =   3.41 per 1000

Risk difference = 10.64 per 1000 – 3.41 per 1000 = 7.23 per 1000

Added risk due to exposure

Attributable Risk Percent

$$\%AR_E = \frac{IP_1 - IP_0}{IP_1} \times 100$$

(Risk difference / Risk in Exposed) x 100

What proportion of occurrence of disease in exposed persons is due to the exposure?

Example

Sleeping Position and Crib Death

	Crib Death		
Usual sleeping position	YES	NO	TOTAL
Prone	9	837	846
Other	6	1755	1761
Total	15	2592	2607

1-year cumulative incidence prone = 9/846 = 10.64 per 1000
1-year cumulative incidence other = 6/1761 = 3.41 per 1000

Risk difference = 10.64 per 1000 – 3.41 per 1000 = 7.23 per 1000

Attributable risk percent = $\underline{10.64 \text{ per } 1000 - 3.41 \text{ per } 1000}$ x 100 = 68.0%
10.64 per 1000
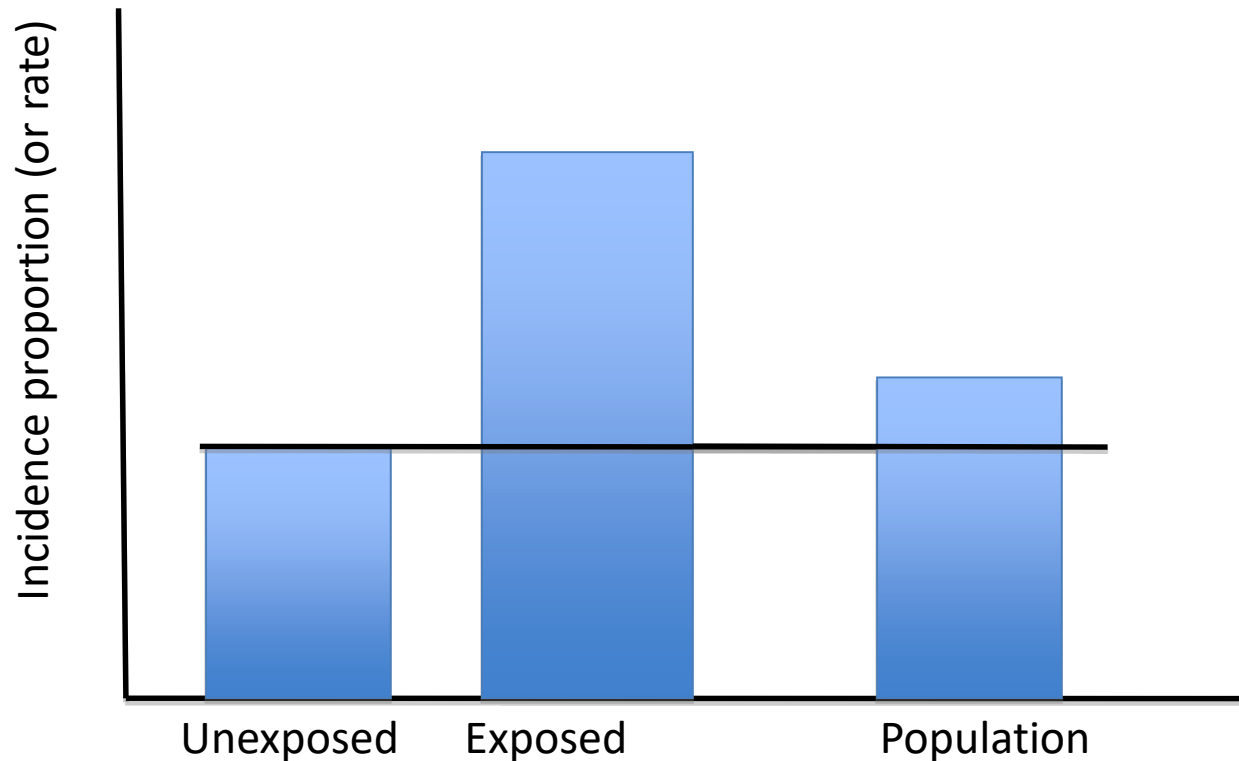
Population Attributable Risk

Should resources be allocated to controlling E or, instead, to exposures causing greater health problems in the population

Population Attributable Risk
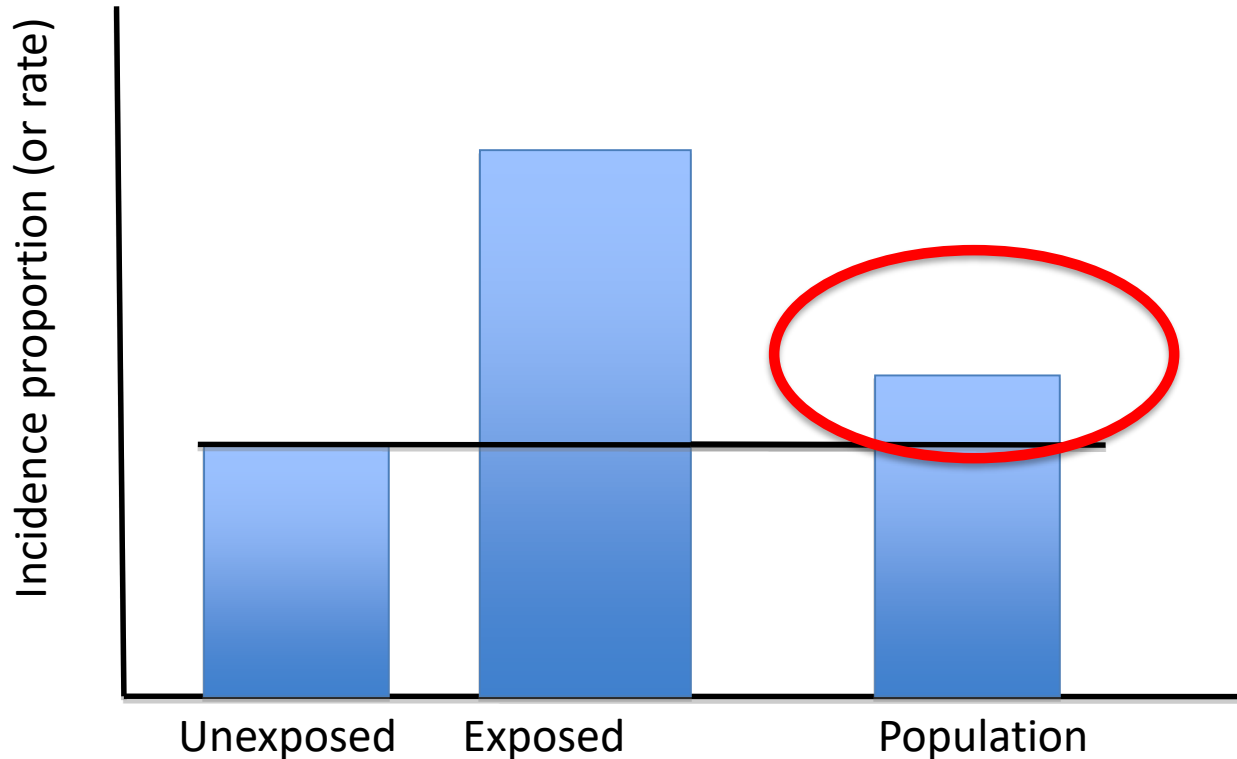
Should resources be allocated to controlling E or, instead, to exposures causing greater health problems in the population

Example

Sleeping Position and Crib Death

	Crib Death		
Usual sleeping position	YES	NO	TOTAL
Prone	9	837	846
Other	6	1755	1761
Total	15	2592	2607

1-year cumulative incidence total = 15/2607 = 5.75 per 1000
1-year cumulative incidence other = 6/1761 = 3.41 per 1000

Population attributable risk (PAR) = 5.75 per 1000 – 3.41 per 1000 =
= 2.35 per 1000

Example

Sleeping Position and Crib Death

	Crib Death		
Usual sleeping position	YES	NO	TOTAL
Prone	9	837	846
Other	6	1755	1761
Total	15	2592	2607

1-year cumulative incidence total = 15/2607 = 5.75 per 1000
1-year cumulative incidence other = 6/1761 = 3.41 per 1000

Population attributable risk percent (PAR) =
= 5.75 per 1000 – 3.41 per 1000 x 100 = 40.8%
   5.75 per 1000

Summary of Measures

- Absolute measures address questions about public health impact of an exposure
  - Excess risk in the exposed or population attributable to the exposure
- Relative measures address questions about etiology and relations between exposure and outcome
  - Relative difference in risk between exposed and unexposed populations

Causal Inference

The Epidemiologic Triad

HOST

AGENT

ENVIRONMENT

Agent

Vector

Host

Environment
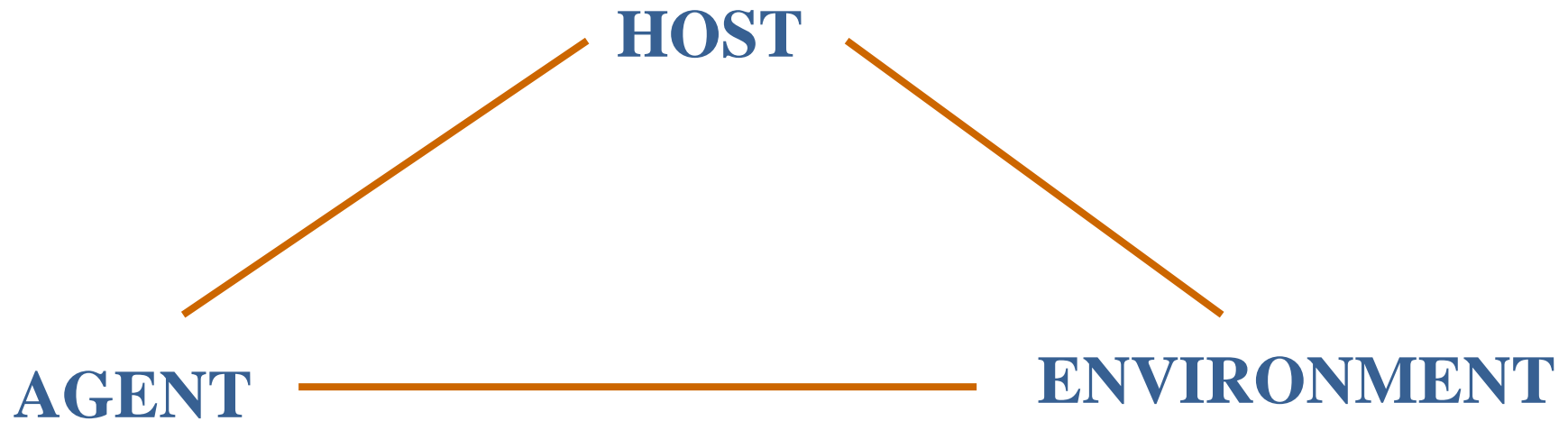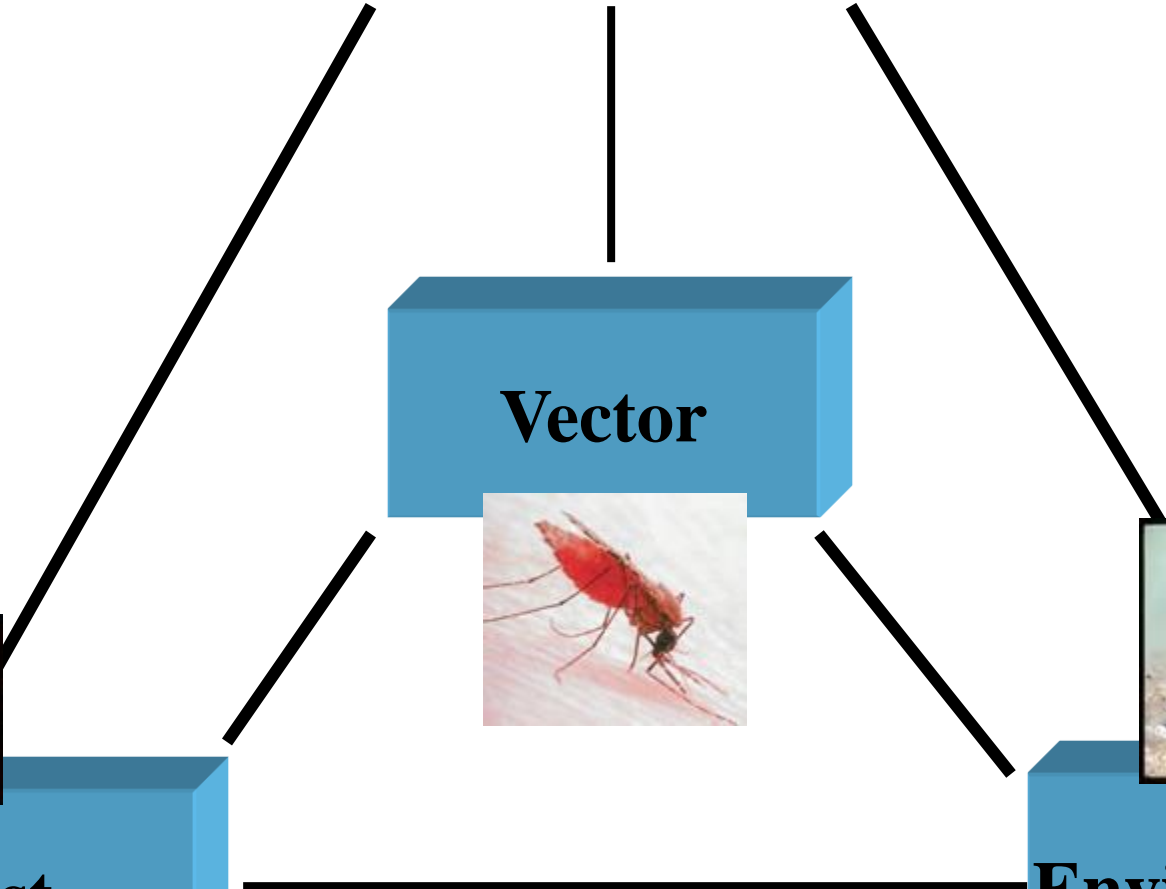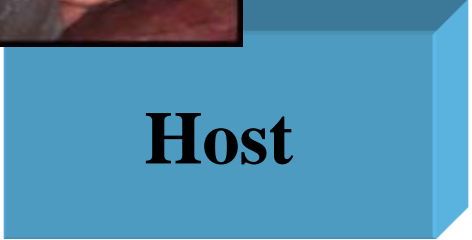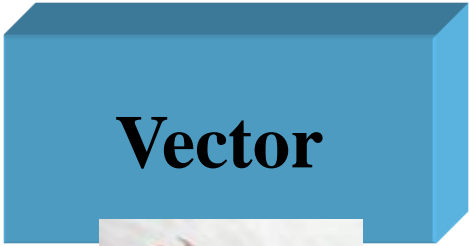
Causal Inference

- During 1950s -1960s epidemiologists developed a set of postulates for causal inferences regarding non-infectious diseases of unknown etiology

- Response to the discovery of association between smoking and lung cancer

- Sir Austin Hill came up with the best known criteria or guidelines in 1965

Hill "Criteria"

1. Strength of Association
2. Consistency
3. Specificity of the Association
4. Temporal relationship
5. Biological gradient
6. Biologic plausibility
7. Coherence
8. Experiment
9. Analogy

Disease Causation – 2 components

- Sufficient Cause
  - precedes the disease
  - if the cause is <u>present</u>, the disease <u>always</u> occurs

- Necessary Cause
  - precedes the disease
  - if the cause is <u>absent</u>, the disease <u>cannot</u> occur

From Study to Causation

- Associations between 'exposures' and outcomes identified in observational studies may or may not be 'causal'

- There is need to pay attention to valid assessment of exposure and outcome in order to think about causality
  - Reliability
  - Validity
    - External validity
    - Internal validity – three concepts are considered
      - Bias
      - Confounding
      - Chance (Random error)

Validity

- Suggests that a measure actually measures what it is expected to measure:
  - Accurate (free of systemic error or bias)
  - Precise (minimal variations; repeatability)

- The degree to which a measurement or study reaches a correct conclusion

- Two types of validity: Internal validity, External validity

Internal validity

- Is the extent to which the results of the study accurately reflect the true situation of the study population

- Is influenced by:
  - **Chance**
    - The probability that an observation occurred unpredictability without discernible human intention or observable cause

  - **Bias**
    - Any systemic error (not random or due to chance) in a study which leads to an incorrect estimate of the association between exposure and disease

  - **Confounding**
    - The influence of other variables in a study which leads to an incorrect estimate of the association between exposure and disease

External validity: generalizabilty

- The extent to which the results of a study are applicable to broader populations

  - Example: Do the study results apply to other patients?

- A representative sample is drawn from the population (usually randomly)

- Individuals have equal chance to participate in the study

- High participation rate

- Inference is made back to the population – but still may not apply to **other** populations

Random error

- Chance

- "That part of our experience that we cannot predict"

- Usually most easily conceptualized as sampling variability and can be influenced by sample size

Random error can be problematic, but . . .

- Influence can be reduced
    - increase sample size
    - improve precision of instrument

- Probability of an observation occurring by chance can be quantified (e.g., p-value or confidence interval width)

I. Bias - Definition

- Any systemic error (not random or due to chance) in a study which leads to an incorrect estimate of the association between exposure and disease or outcome

- Therefore:
    - **Bias is a systematic error that results in an incorrect (invalid) estimate of the measure of association**
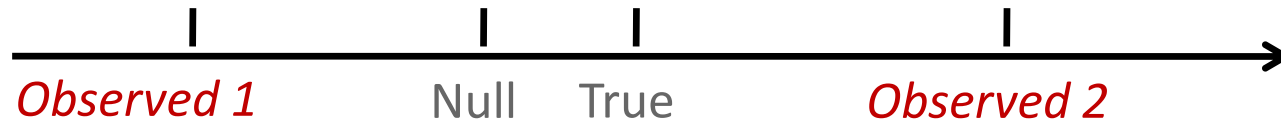
I. Bias - Definition

1. Can create spurious association when there is none (bias away from the null)
2. Can mask an association when there is one (bias towards the null)
3. Bias is primarily introduced by the investigator or study participants
4. Bias does not mean that the investigator is "prejudiced"
5. Can occur in all study types: experimental, cohort, case-control
6. Occurs in the design and conduct of a study
7. Bias can be evaluated but not necessarily "fixed" in the analysis phase
8. Three main types are <u>selection</u> and <u>information</u> bias and <u>confounding</u>

Direction of bias

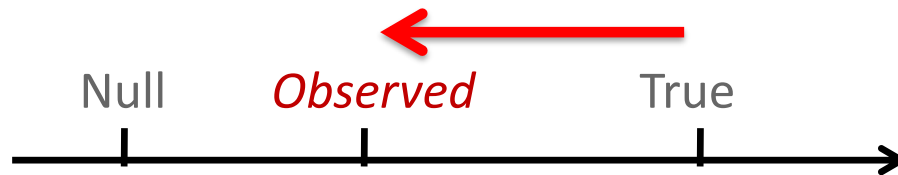- ***Bias towards the null*** – observed value is <u>closer to 1.0</u> than is the true value

Null	*Observed*	True

- ***Bias away from the null*** – observed value is <u>farther from 1.0</u> than is the true value

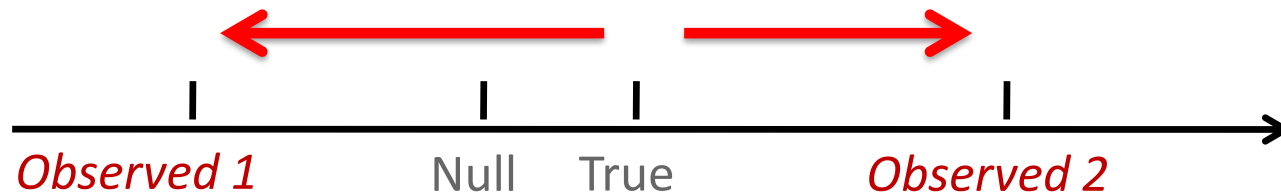| *Observed 1*	Null	True	*Observed 2*

Direction of bias

- ***Bias towards the null*** – observed value is <u>closer to 1.0</u> than is the true value

- ***Bias away from the null*** – observed value is <u>farther from 1.0</u> than is the true value

Types of bias

- **<u>Selection bias</u>**
  - Refusals, exclusions, non-participants
  - Failure to enumerate the entire population
  - Loss to follow up

- **<u>Information bias</u>**
  - Interviewer bias
  - Recall bias
  - Misclassification of exposure and outcome

- **<u>Misclassification (is part of information bias)</u>**
  - Non-differential
  - Differential

II. Selection bias

- Systematic error that occurs in the process of identifying (or retaining) study populations

- The error that occurs when losses to follow-up are is **not independent of exposure and outcome** (cohort study)

- Error due to systematic difference between those selected for study versus those not selected for the study (case-control study)

194

II. Selection bias- cohort study

Solutions:

– Minimize losses to follow-up!!!

II. Selection bias: case-control study

- **<u>Sources of selection bias</u>**

  – When controls do not reflect the population that gave rise to the cases

- The selection of cases and controls <u>must be **independent** of the exposure status</u>

  – Do controls in the study have higher or lower prevalence of exposure than controls not selected for the study?

II. Selection bias: case-control study

1. Occurs when controls or cases are more or less likely to be included in a study if they have been exposed –

   inclusion in the study is not independent of exposure

2. Results: relationship between exposure and disease observed among study participants is different from relationship between exposure and disease in eligible individuals who were not included

3. The odds ratio from a study that suffers from selection bias will incorrectly represent the relationship between exposure and disease in the overall study population

III. Information bias

- An error that arises from systematic differences in the way information on exposure or disease is obtained from the study groups

- Results in participants who are incorrectly classified as either exposed or unexposed or as diseased or not diseased

- Occurs after the subjects have entered the study

- Several types of observation bias: recall bias, interviewer bias, and differential and non-differential misclassification

III. Observation/Information bias

Recall bias

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

- Can result in over-or under-estimation of measure of association

III. Observation/Information bias

Recall bias

- Solutions:
  - Use controls who are themselves sick
  - Use standardized questionnaires that obtain complete information
  - Mask subjects to study hypothesis

III. Observation/Information bias

Interviewer bias

- Systematic difference in soliciting, recording, interpreting information

- Can occur whenever exposure information is sought when outcome is known (as in case-control) or when outcome information is sought when exposure is known (as in cohort study)

III. Observation/Information bias

- ## **<u>Interviewer bias</u>**

  - Solutions:

    - Mask interviewers to study hypothesis and disease or exposure status of subjects

    - Use standardized questionnaires, or standardized methods of outcome or exposure ascertainment

    - Use biomarkers to compare when possible

III. Observation/Information bias – Misclassification bias

- A type of information bias

- Error arising from **inaccurate measurement or classification** of study subjects or variables

- Subject's exposure or disease status is erroneously classified

- Happens at the assessment of **exposure** or **outcome** in both cohort and case-control studies

- Two types: **non-differential and differential**

A. Non-differential misclassification

- Inaccuracies with respect to disease classification are independent of exposure

- Inaccuracies with respect to exposure are independent of disease status

- The probability of exposure (or of outcome) misclassification is the same for cases and controls (or in study/comparison groups)

- Bias results towards the null - if the exposure has two categories, will make groups more similar

- Solution: Use multiple measurements and/or choose the most accurate sources of information

B. Differential Misclassification

- **Differential misclassification**

  - Probability of misclassification of disease or exposure status differs for exposed and unexposed persons (cohort) or presence of absence of exposure (case-control)

  - Probability of misclassification is different for cases and controls or for levels of exposure within cases and controls

  - Direction of bias is unknown, i.e. overestimation or underestimation of the true risk

  - Know that the observed RR or OR deviates from truth, but direction is **unknown**
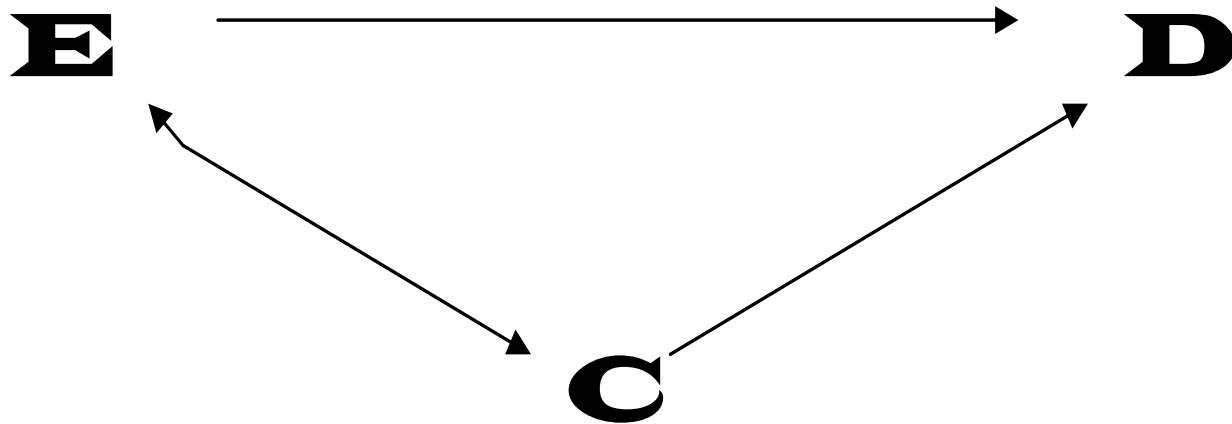
Confounding

Definition and Impact

- "A mixing of effects": the association between exposure and disease is distorted because it is mixed with the effects of another factor that is associated with the disease

- Result of confounding is to distort the true association toward the null or away from the null

Criteria for a variable to be a confounder

- The variable must be an **independent predictor** of disease

- The variable must be **associated (correlated)** with exposure

- The variable must not be **an intermediate link** in the **causal chain** between exposure and outcome

CONFOUNDING

$$E \longrightarrow D$$

$$E \nwarrow \qquad \nearrow D$$

$$\searrow C \nearrow$$

Example:

Smoking is a confounder of association between coffee consumption and lung cancer

Opportunities for confounding

- **In an experimental designs:**

  – No randomization
  – Residual confounding after randomization

- **In cohort and case-control studies:**

  – When comparison group differs by subject characteristics

  – When risk factors other than the exposure are distributed differently between the exposed and unexposed groups

  – There is residual confounding

Control for confounding- design phase

- **Randomization**
  - With sufficient sample size, randomization is likely to control for both known and unknown confounders- but not guaranteed

- **Restriction**
  - Restrict admissibility criteria for study subjects and limit entrance to individuals who fall within a specified category of the confounder

- **Matching**
  - **No so much a control for confoundin**g; more of a way to maximize efficiency

Control for confounding- analysis phase

- **Standardization:** by age, race, gender, or calendar time in order to make fair comparisons between populations

- **Restriction:** Restrict during data analysis

- **Stratified analysis**: a way of eliminating variation in the confounding factor – feasible with a small number of variables

- **Multivariate analysis:** To enable controlling for several potential confounders simultaneously

Effect modification

- Interaction

- The strength of the association between an exposure and disease differs according to the level of another variable.

- Modification of the relationship between exposure and a disease by a variable.

- If the association changes according to the level of that variable, then effect modification is present

Example: Antioxidant Intake and Esophageal Squamous-cell

- RR high vs. low (SMOKERS)   **= 0.4**
- RR high vs. low (NON-SMKs)   **= 0.9**
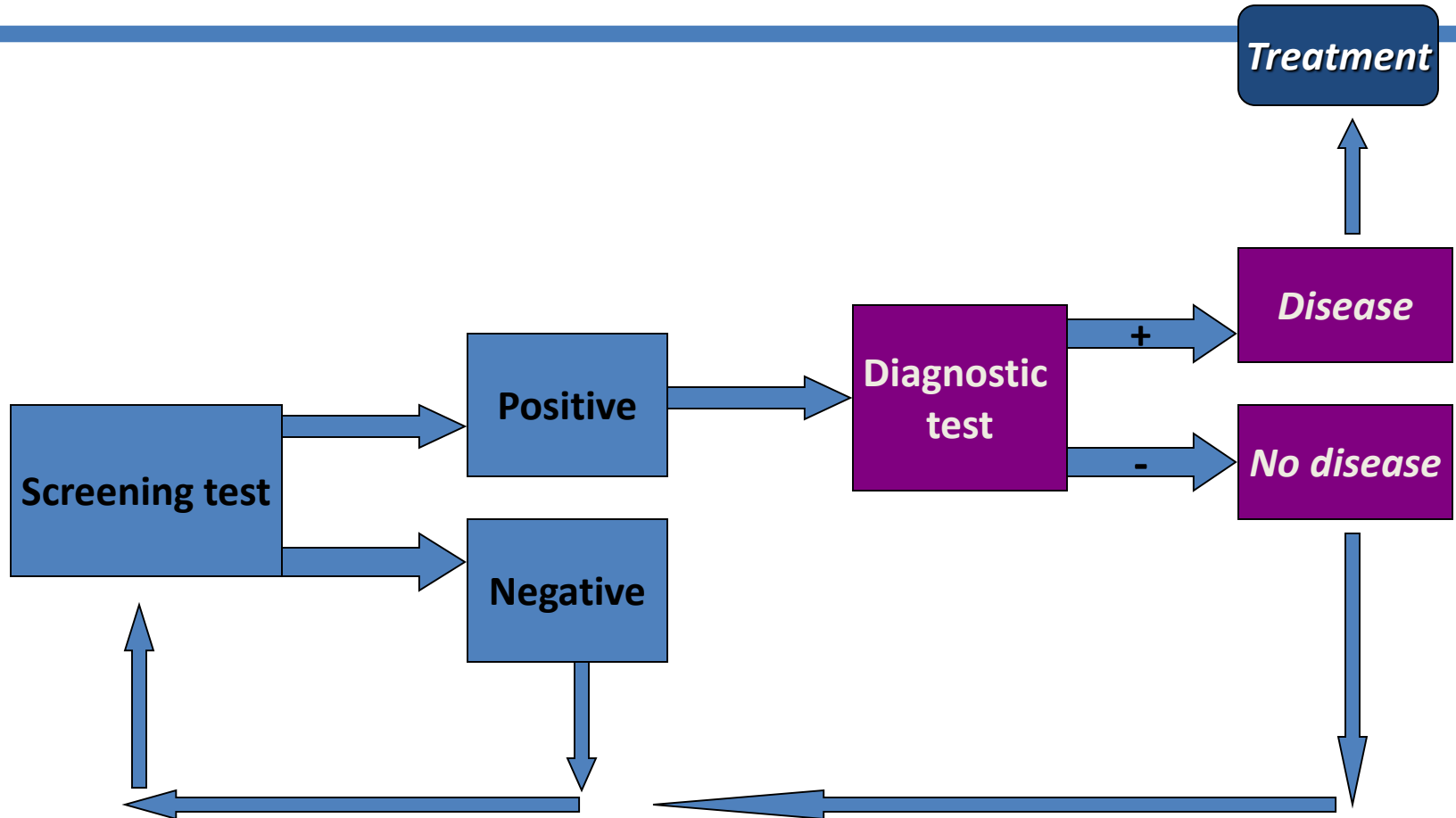
Example: Aspirin and Reye's Syndrome

- RR yes vs. no (youth)   **= 4.4**
- RR yes vs. no (adults)   **= 1.0**

SCREENING

Screening

- The application of one or more tests to determine those likely to have the disease from those unlikely to have the disease
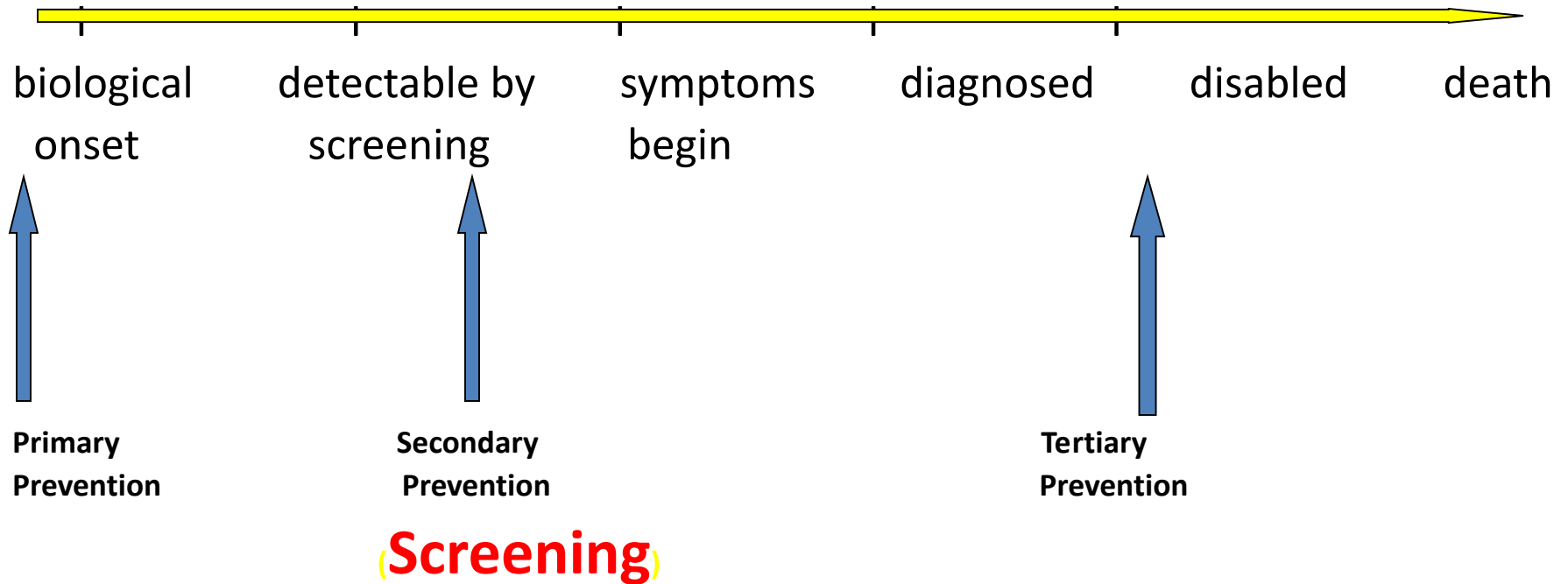- Two step process – screening followed by diagnosis

Two Step Process

Some examples

- Mammography – breast cancer

- Fecal occult blood test – colon cancer

- Pap smear for cervical cancer

- X-ray – lung cancer

- Blood pressure - hypertension

- Blood sugar – diabetes

- Prostate specific antigen – prostate cancer

Natural History of Chronic Disease and Types of Prevention

biological onset   detectable by screening   symptoms begin   diagnosed   disabled   death

Primary Prevention

Secondary Prevention

Tertiary Prevention

(Screening)

Some Fundamental Truths about Screening

▶ Screening is not error-free

▶ Screening tests will fail to identify some individuals with the disease, and falsely identify some without disease as needing further testing.

Accuracy

- Sensitivity and specificity measure the ability of a test to correctly identify diseased and nondiseased people

- **Sensitivity** refers to the proportion of people with the disease who test positive

- **Specificity** refers to the proportion of people without the disease who test negative

- A test with poor sensitivity will miss people who have the disease (false negatives) and a test with poor specificity will falsely identify healthy people as diseased (false positives)

- Gold standard is needed to assess those classified as test positive or test negative – "*diagnostic test*"

Sensitivity

- The probability of testing positive if the disease is truly present

Sensitivity = a / (a + c)

		True Disease Status	
		+	-
Results of Screening Test	+	a	b
	-	c	d

Specificity

- The probability of screening negative if the disease is truly absent

Specificity= d / (b + d)

		True Disease Status	
		+	-
Results of Screening Test	+	a	b
	-	c	d

Relationship between Sensitivity and Specificity

- Lowering the criterion of positivity results in an increased sensitivity, but at the expense of decreased specificity

- Making the criterion of positivity more stringent increases the specificity, but at the expense of decreased sensitivity

- The goal is to have a high sensitivity and high specificity, but this is often not possible or feasible

GOOD LUCK!!!