Epidemiology

Introduction

Mel and Enid Zuckerman College of Public Health
The University of Arizona

Basic Principles of Epidemiology CPH/EPID 573A

Learning Objectives

- Understand definitions & motivations of epidemiology
- Begin to consider the 3 epidemiologic questions to analyze information
- Become familiar with the topics covered in the course

Learning Objectives

Students will be able to:

- define Epidemiology
- give examples of how the field & principles of Epidemiology are useful in everyday situations.

Epidemiology

epi upon or befall

demo the people

ology the study of

study of what falls upon the people

Not: epidermis

Classical definition

Study of the **distribution** & **determinants** of **disease frequency** in human populations

Newer Definition

Study of the distribution & determinants of health-related states or events in specified populations & application of this study to the control of health problems

What Epidemiology Does

- Measurement of disease frequency
  - quantification of existence or occurrence of disease
- Distribution of disease
  - who is getting disease in the population, where & when is disease occurring
- Determinants of disease
  - final piece - is derived from other components

Epidemiologic Premise

- Human disease
  - does not occur at random
  - has causal and preventive factors

Such factors can be identified through systematic investigation of different populations in different places or different times

A Comparative Discipline

- Description of patterns & distributions
- Identification of associations (or disease occurrences)

Objective of Epidemiology

- The objective of epidemiology:
  to **JUDGE**
  whether an **ASSOCIATION**
  between exposure and disease

  is **CAUSAL.**

Epidemiology Questions

1. How many ?

2. Compared to what ?

3. Is it real ?

Progression of Reasoning

Suspicion

of an influence or association
of a factor & outcome

⇩

Formulation of hypothesis

⇩

Test hypothesis

Systematic collection & analysis of data
to test for statistical association

⇩

m01b

Progression of Reasoning continued

⇩

Assessment of the validity

Chance

Systematic errors (bias)

Effects of other variables (confounding)

⇩

Judgement

whether the association represents a cause & effect relationship

Purposes of Epidemiology

Describe

Explain

Predict

Control

m01b

m01b

Uses of Epi Tools

- Determine extent of disease in community
- Identify etiology or cause of a disease and the risk factors
- Determine the proportion of a disease attributable to a particular exposure

m01b

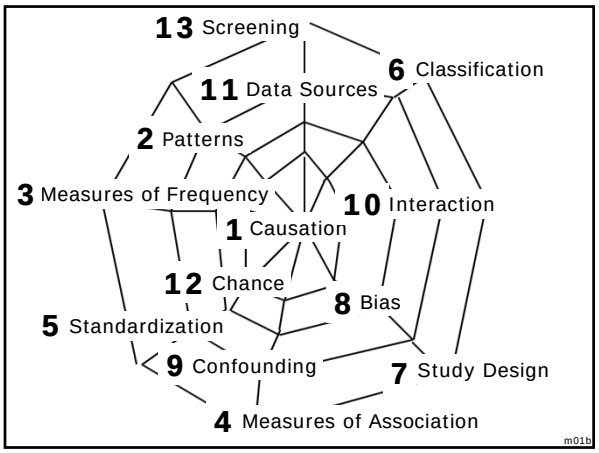
Uses of Epi Tools - more

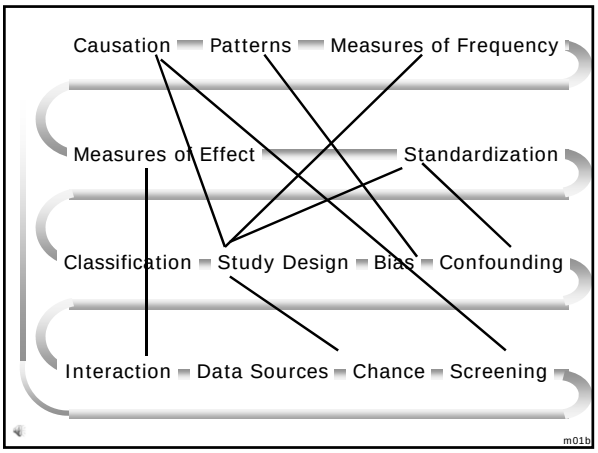
- Study the natural history and prognosis of disease
- Evaluate new preventive and therapeutic measures and modes of health delivery
- Provide foundation for developing public policy and regulatory decisions

Epidemiology Questions

1. How many?
2. Compared to what?
3. Is it real ?

m01b





History of Epidemiology

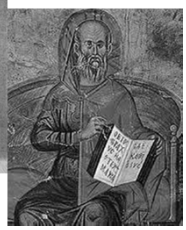
Learning Objectives

- Understand origins of epidemiology as a discipline
- Understand where to begin to learn more about the history of epidemiology

m02a

History of Epidemiology

400 BC

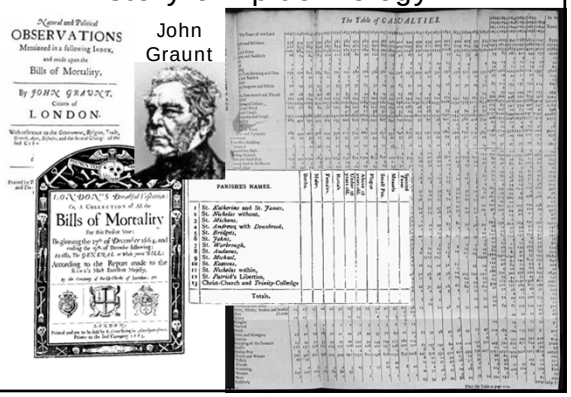


Two different artists interpretations of Hippocrates

Hippocrates

"airs, waters, and places" as causes of disease

History of Epidemiology



History of Epidemiology

Bernardino Ramazzini
the founder
of Occupational Medicine

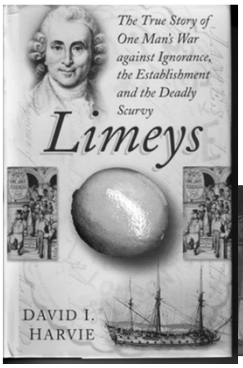
Bernardo Ramazzini

Disease of Workers
1700, 2nd edition 1713

History of Epidemiology

John Lind

1753
prospective study of scurvy



History of Epidemiology

Edward Jenner

Immunology & vaccinations

Late 1700's

History of Epidemiology

William Farr

1807 - 1883

Pioneer of medical statistics

While a trained physician interested in medical statistics, Dr. Farr was posted to the British Registrar General's Office as the Compiler of Abstracts tracking births & deaths to assure appropriate transfer of property rights of the landed gentry.

William Farr created the 1st vital statistics system along with the field of nosology from which the ICD *(International Classification of Diseases)* system developed. Along with other accomplishments, Dr. Farr developed the concept of surveillance.

History of Epidemiology

Ignaz Semmelweiss

1818 - 1865

Hungarian Physician showed that childbed fever was contagious & could be minimized by hand-washing.

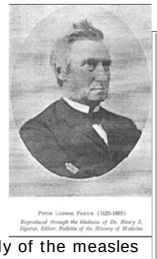
Most of his contemporaries did not understand his papers & did not believe his findings.

Only after Dr. Semmelweiss' death was the germ theory of disease developed, and he is now recognized as a pioneer of antiseptic policy and prevention of nosocomial disease.

History of Epidemiology

Peter Ludvig Panum

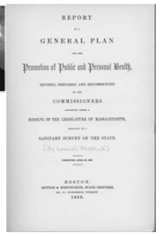
Danish Physiologist & Pathologist chosen by the government to undertake research of a measles epidemic in the Faroe Islands. As a result of his investigations he published a classic treatise titled "*Observations Made During the Epidemic of Measles on the Faroe Islands in the Year 1846*".



Panum's comprehensive study of the measles epidemic in a previously unexposed population, provided the basis of our understanding of the spread of an infectious disease.

History of Epidemiology

Lemuel Shattuck


REPORT
GENERAL PLAN
Promotion of Public and Personal Health.
COMMISSIONERS
SANITARY SURVEY OF THE STATE.
BOSTON:

In 1850, he published a Sanitation Report that established a model for state boards of health in Massachusetts (1869) and other parts of the United States. The recommendations in the report became the foundation of the US sanitation movement. Application of these recommendations led to an increase of life expectancy in Boston from 25 years in 1850 to more than 75 years in 2000. The proposed legislation included in the report is the first attempt to write a comprehensive public health code.

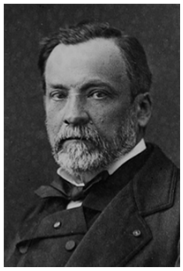
History of Epidemiology

Chronologically, we would discuss John Snow next. However, we will skip and come back to his contributions to Epidemiology.

History of Epidemiology

Louis Pasteur

1822 - 1895



Most infectious diseases caused by germs…

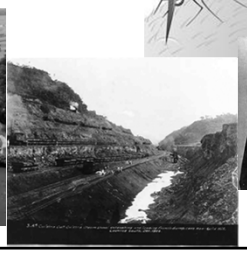
History of Epidemiology

Dr.Walter Reed

1851 - 1902

Mosquitos transmit yellow fever; Panama canal could be completed!

Walter Reed at UVa
1869



History of Epidemiology

Sir Richard Doll

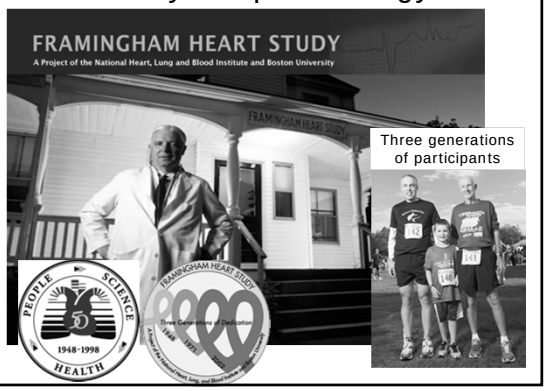
Sir Austin Bradford Hill

Two classic studies of smoking effects.

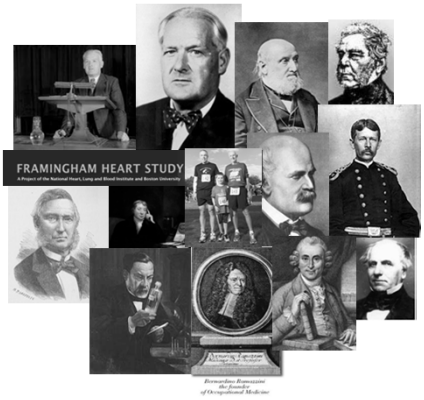

Sir Austin Bradford Hill

Doll used the case-control study method & compared the smoking history of a group of hospitalized patients with lung cancer with the smoking history of a similar group without lung cancer. Hill used a cohort study, categorizing a group of British physicians according to their smoking histories & then analyzing the causes of death among those who died, to see whether cigarette smokers had the highest incidence of lung cancer.

History of Epidemiology



Three generations of participants



History of Epidemiology

John Snow
1813 - 1858

London:
Cholera epidemics

year	died
1831	13,000
1848	21,000

Cholera in London

Death rates 1853-1854 by Water Company supplying each HH

Water Company	# of HH	Cholera Deaths	Deaths / 10,000 HH
Southwark & Vauxhall	40,046	1263	315
Lambeth	26,107	98	37
Rest of London	256,423	1422	59

HH = households

Source: J. Snow, *On the Mode of Communication of Cholera* (2nd ed) London: Churchill, 1855. Reproduced in *Snow on Cholera*. New York: Hafner, 1965

m02b

Cholera in London

Death rates 1853-1854 by Water Company supplying each subdistrict of London

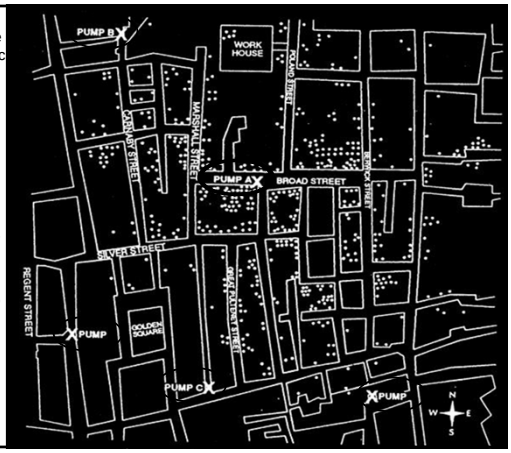
Water Company	Population in 1851	Cholera Deaths 1853-1854	Deaths /100,000 living
Southward & Vauxhall	167,654	192	114
Both	301,149	182	60
Lambeth	14,632	0	0

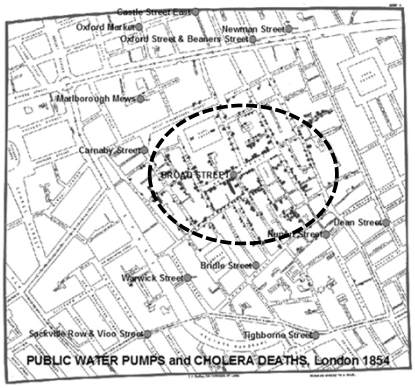
Source: J. Snow, *On the Mode of Communication of Cholera* (2nd ed) London: Churchill, 1855. Reproduced in *Snow on Cholera*. New York: Hafner, 1965

m02b

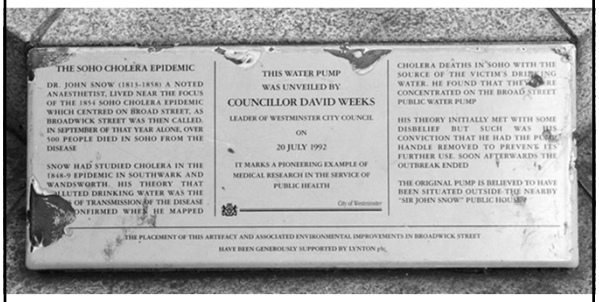
•Dots are
cholera c

Distribution of cholera cases in the Golden Square area of London August – September 1854





Broad Street Pump



Snow's strategy

- Clinical knowledge + observations on disease distribution
- Testable hypothesis considering alternative explanations
- Judgment that cause-effect relationship existed

% of smokers / nonsmokers in lung carcinoma patients and in control patients with diseases other than cancer

Group	# non-smokers	# Smokers	Probability Test
Males			
Lung carcinoma pts (646)	2 (0.3%)	647	P = 0.00000064 (exact method)
Control pts – diseases other than cancer (649)	27 (4.2%)	622	
Females			
Lung carcinoma pts (646)	19 (31.7%)	41	
Control pts – diseases other than cancer (649)	32 (53.3%)	28	$X^2 = 5.76$; n=1 0.01<P<0.02

1950 Hill & Doll study of lung cancer & smoking in England

Epidemiology: results into action!

Causation
Motivation & Definitions

50

Learning Objectives

- Understand definitions & motivations of causality as it relates to epidemiology

Epi Objective

A primary objective of epidemiology is to JUDGE

whether an ASSOCIATION between an exposure and a disease is

CAUSAL

Progression of Reasoning

- Suspicion of an influence or association of a factor and outcome
- Formulation of a hypothesis
- Testing the hypothesis
- Systematic collection and analysis of data to test for statistical association
- Assessment of the validity of the association:
  - Chance
  - Systematic errors (bias)
  - Effects of other variables (confounding)
- Judgment on whether the association represents a cause and effect relationship

Progression of Reasoning

- Suspicion of an influence or association of a factor and outcome
- Formulation of a hypothesis
- Testing the hypothesis
- Systematic collection and analysis of data to test for statistical association
- Assessment of the validity of the association:
  - Chance
  - Systematic errors (bias)
  - Effects of other variables (confounding)
- Judgment on whether the association represents a cause and effect relationship

Association ≠ Cause

- Association does not necessarily imply cause
  - Sunspots and economics
  - A rooster's crowing is associated with dawn, did the rooster cause the sun to rise?

Cause - definition

- "Something that brings about a result especially a person or thing that is the agent of bringing something about
  - » Merriam-Webster Dictionary
- "what makes something happen"
  - » Encarta Dictionary

http://encarta.msn.com/dictionary_1861595576/cause.html

Cause in Epidemiology

- An event, condition, or characteristic without which the disease would not have occurred
  - » KJ Rothman
- Something that makes a difference
  - » M Susser

Cause in Epidemiology

- A cause has 2 essential attributes
  - Association
  - Time order
- Causes can include
  - Host & environmental factors
  - Active agents and static factors
- Causes may be either positive or negative

Exposure

- Substances in the environment
  - Infectious
  - Carcinogenic
  - Physical
    - Radiation
    - Force (trauma)
    - Sound
  - Food or specific nutrients or lack of them

Exposure

- Host characteristics
  - Gender
  - Race
  - Age
  - Culture
  - Genes
  - Marital status
  - Personality

Outcomes

- Disease
  - Mortality
  - Illness
- Injury
  - Trauma
  - Dental caries
- Intermediates
  - Blood pressure
  - Blood sugar
  - White cell count
  - Cholesterol
- Functioning
  - Neurological function
  - Range of motion in joints
  - Vision
  - Hearing
- Quality of life indicators

Cause vs Risk Factor

Are these causes or risk factors?

Characteristics	High Risk Group for Breast Cancer	Low Risk Group for Breast Cancer
Country of Birth	North America, Northern Europe	Asia, Africa
Socioeconomic Status	High	Low
Marital Status	Never married	Ever Married

Causes

What are some causes of the following diseases and events?

- Influenza mortality
- Lung cancer
- Breast cancer
- Automobile fatality

Cause in Epidemiology

- A cause is sufficient when it inevitably produces or initiates a disease
- A cause is necessary if a disease cannot develop in its absence

History of Causation

65

Learning Objectives

- Understand origins of causal thought in epidemiology as a discipline

History of Causation

- Divine retribution
- Imbalance in body humors
  - caused by
    - Air
    - Water
    - Land
    - Stars
    - Spontaneous generation
- Miasma
  - Disease transmitted by miasmas or clouds clinging to earth's surface

Henle-Koch Postulates

- Germ Theory of Disease
  - microorganism must always be found with the disease
  - specificity of cause / one to one relationship between exposure and disease

Koch's Postulates

- The organism must be present in every case of the disease
- The organism must be able to be isolated and grown in pure culture
- The organism must, when inoculated into a susceptible animal, cause the specific disease
- The organism must be recovered from the animal and identified

Non-infectious Diseases

- What about non-infectious outcomes?
  - Cancer
  - Injury
  - Immune dysfunction
    - Asthma
    - Lupus
    - Rheumatoid arthritis
  - Mental Health outcomes

Web of Causation

- Shift from Henle-Koch Postulates
- Multiple causes
- Revisions for establishing causation in chronic diseases

Other Models Needed

- Web of causation
  - incorporates concepts of multifactorial causes
  - implies the cases of disease could be prevented by cutting only a 'few strands of the web'

Recent Controversies

- Causation cannot be established
- Causal criteria should be abandoned

Causation Models

74

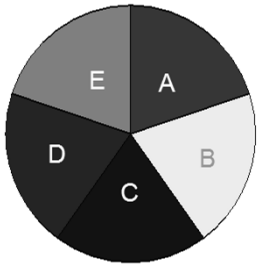
Learning Objectives

- Explore another model for conceptualizing the cause of disease

Models of causal relationships

- Necessary and sufficient
- Necessary but not sufficient
- Sufficient but not necessary
- Neither sufficient nor necessary

Causal Pies



from KJ ROTHMAN

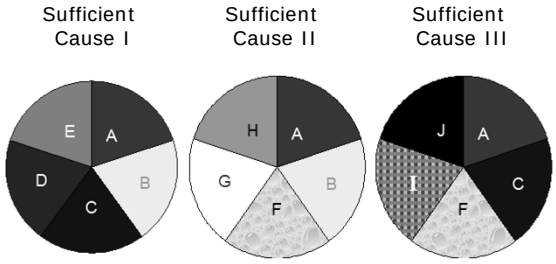
Causal Pies

- Sufficient cause
  - A set of conditions without any one of which the disease would not have occurred
  - This is one whole pie

Causal Pies

- Component cause
  - Any one of the set of conditions which are necessary for the completion of a sufficient cause
  - This is a piece of the pie

Causal Pies



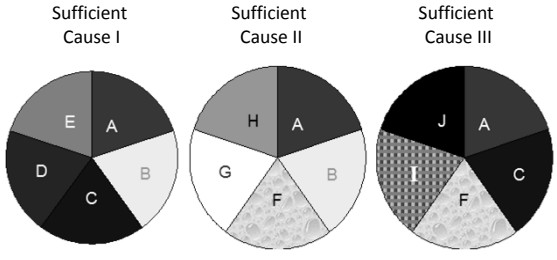
Causal Pies

- Completion of a sufficient cause is synonymous with occurrence of disease
  - although not necessarily diagnosis
- Component causes can act far apart in time

Causal Pies

- Necessary cause
  - A component cause that is a member of every sufficient cause

Causal Pies



A is a necessary cause
B - J are not necessary causes

Causal Pies

- Component cause
  - + causative exposure OR
  - lack of preventive exposure
- Blocking any component
  - Prevent causative exposure OR
  - Provide preventive exposure
- prevents completion sufficient cause
- therefore prevents disease by that pathway

Epidemiologic Causation Criteria

85

Learning Objectives

- Understand epidemiologic criteria for assessing causality

Systematic Approach

- Criteria developed by
  - 1964 Report of the Advisory Committee to the US Surgeon General on Smoking and Health
  - Sir Austin Bradford Hill – 1965
- None of criteria intended to be rigid, rather to serve as guidelines
- Emphasis on judgment & consensus

Hill's Guidelines

1. Temporal relationship
2. Strength of association
3. Dose-response
4. Consistency
5. Biologic plausibility / coherence
6. Specificity
7. Experiment
8. Analogy

Temporality

- Exposure to the causal factor must precede the disease in time
- This is the only one of Hill's criteria that everyone agrees with

Strength of Association

- Strength of association
  - Stronger associations are more likely to be causal

Strength of Association

- Larger measure of association
  ↑ likelihood 
  – exposure is causing the disease

Risk of lung cancer for heavy smokers is 20 times the risk for non-smokers

Strength of Association

- Strong associations
  – unlikely to be due to bias and confounding
- Weak associations
  – may be causal but
    - harder to rule out bias and confounding

Dose Response

- Biological gradient
- Increasing exposure →
- increasingly higher risk of disease
  - Lung cancer death rates rise with the number of cigarettes smoked:
    – Relative risk of lung cancer in smokers vs. non-smokers = 9
    Relative risk of lung cancer in heavy vs. non-smokers = 20
- Some exposures
  – No dose-response, but
  – "threshold effect"
  – below which:
    - no adverse outcomes

Consistency

- Other investigations
  – Different
    - populations
    - places
    - times
    - circumstances
    - investigators
  – Similar or same results

Consistency

- Smoking associated with lung cancer in at least
  – 29 retrospective and
  – 7 prospective studies
- Note: Sometimes there are good reasons why study results differ
  – For example, one study may have looked at low level exposures while another looked at high level exposures

Plausibility / Coherence

- Biologic plausibility
  – the association makes sense biologically
- Association does not conflict with current knowledge of natural history and biology of disease

Specificity

- Idea:
  - a single exposure should cause a single disease
- Hold-over from infectious diseases
- Many exceptions:
  - Exposure to cigarette smoking causes:
    - Cancer of many types: lung, breast, etc.
    - COPD
    - Intrauterine growth retardation
- If present, evidence of causality
  - but its absence does not preclude causation

Experiment

- Investigator controlled decrease in exposure
  - prevention,
  - treatment,
  - removal
    - results in less disease
- Controlled exposure
  - Fluoridated drinking water & dental caries
- Nice evidence when it can be gotten

Analogy

- A similar relationship observed with another exposure and/ or disease?
- Example
  - Thalidomide and Rubella
  - Birth defects
- Analogy for effects of similar substances on the fetus

Epidemiologic Causation Assessment

100

Learning Objectives

- Understand the process of assessing causality from the epidemiologic literature

Assessing Causality

- How do we actually assess whether a risk factor is indeed causal?
  - First, evaluation of individual studies

Hill's Criteria	Individual Study	Epi Literature
Temporal	X	
Strength of Assoc	X	
Dose-response	X	
Consistency		X
Biologic Plausibility / coherence	X	X
Specificity		X
Experiment	X	X
Analogy	X	X

Study Evaluation

- Focus on determining validity of the study and its methods
  - Collection of data
    - Temporal relationship demonstrated?
  - Analysis of data
    - Strength of association?
    - Dose-response?
  - Interpretation of data

Note: This is relevant to Project 2

Useful Criteria

- What is the purpose of the research?
- What is the study design?
- Who are the study subjects?
- What are the definitions of outcomes and exposures?
- What are potential sources of bias and how are they addressed?
- What is the statistical analysis, e.g. measures of association
- Are the conclusions justified?
- Are there alternate explanations?

Note: This is relevant to Project 2

Study Evaluation

- Valid statistical association?
  - … likely to be due to bias?
  - … likely to be due to confounding?
  - … likely to be due to chance?

Reviewing Epi Literature

- Individual studies valid?
- Studies consistent?
- Associations coherent?
- Literature include experimental data?
- Literature include analogous studies?

Evaluate

- Study population
  - 50 people recently diagnosed with Hepatitis C
  - 50 people without Hepatitis C
- Both groups interviewed
  - any body piercings within the past two months?
- Groups matched for age, sex, and race
- Findings
  - body piercing more likely among those people with Hepatitis C than those without the disease, after controlling for other factors.

Patterns / Descriptive Epi

109

Learning Objectives

- Understand how patterns in data lead to hypotheses about the causes of disease

110

Descriptive Epidemiology

- Descriptive statistics
  - information on disease patterns by characteristics of
    - Person
    - Place
    - Time

111

Person

What hypotheses can you generate from these data?

Mortality rates per 100,000 from diseases of the heart by age and sex (2000)

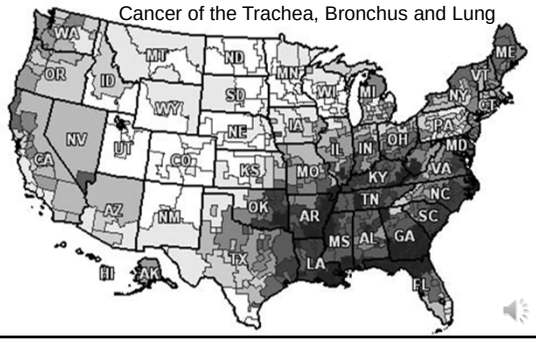
Age (in years)	Men	Women
25-34	10.3	5.5
35-44	41.6	17.2
45-54	142.7	50.3
55-64	378.6	160.4
65-74	909.2	479.9
75-84	2210.1	1501.5
85+	6100.8	5740.1

112

Place

What hypotheses can you generate from this map?

Cancer of the Trachea, Bronchus and Lung



113

Place- Migrant Studies

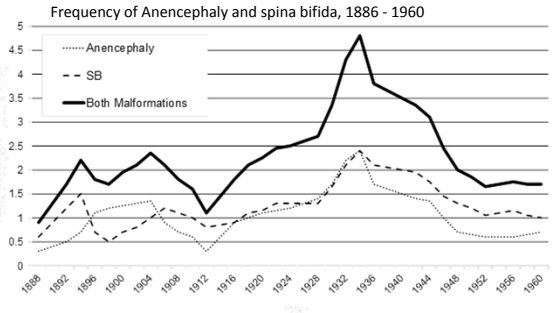
What hypotheses can you generate from these data?

Mortality rates (per 100,000) due to stomach cancer	
Japanese in Japan	58.4
Japanese Immigrants to California	29.9
Sons of Japanese Immigrants	11.7
Native Californians (Caucasians)	8.0

114

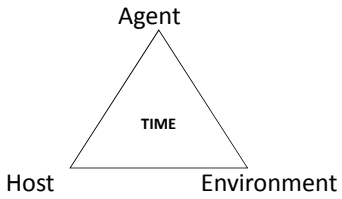
Time

What hypotheses can you generate from this graph?



115

Epidemiologic Triangle



external agent - susceptible host -
in an environment that brings them all together

116

Agent

- Infectious microorganism (bacterium, virus, parasite)
  - must be present for a disease to occur
- Non-infectious diseases
  - "Agent" broadened to include
    - Chemical
    - Nutrients
    - Physical causes of disease

117

Host

- Intrinsic characteristics that influence
  - Exposure
    - Age
    - Sex
    - Race
    - Lifestyle behaviors
  - Susceptibility
    - Age
    - nutritional status
    - psychological profile
    - genes
    - immunocompetence
  - Response to a causative agent

118

Environment

- Environment
  - Temp – ex: temp at which increase in mortality varies
  - Humidity
  - Altitude
  - Crowding
  - Housing
  - Neighborhood
  - Water / food supply
  - Air pollution
  - noise
- Affects survival of
  - agent
  - host
- Affects likelihood that the host and agent will come in contact

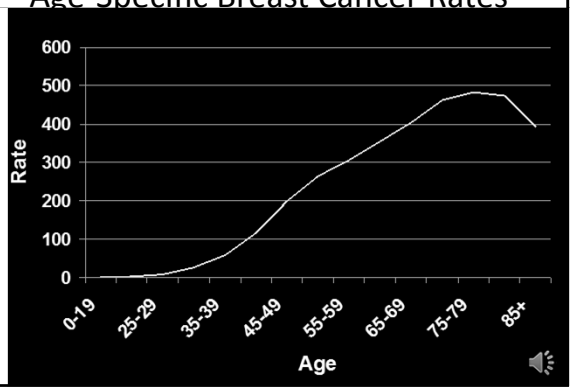
119

Person

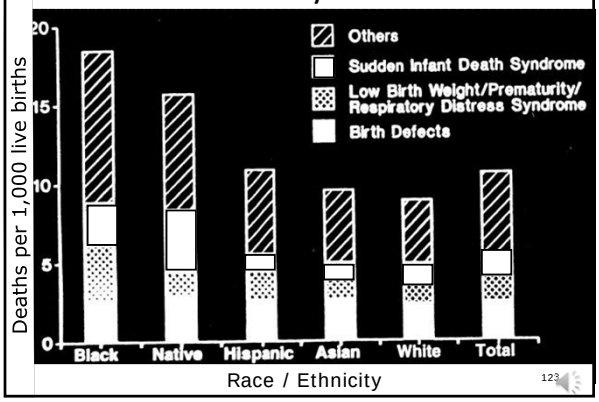
- Age
- Sex
- Race
- Marital status
- Nativity and migration
- Religion
- Socioeconomic status

120

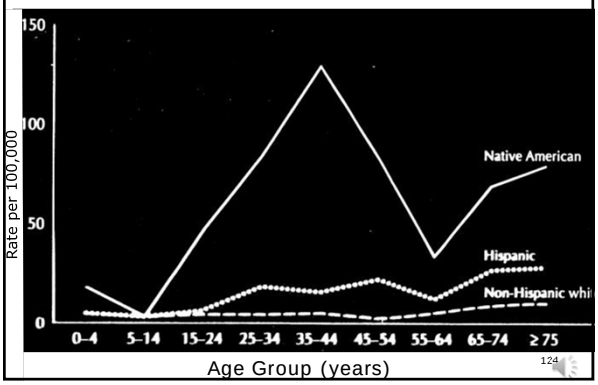
Age-Specific Breast Cancer Rates



Infant Mortality – CDC 1983



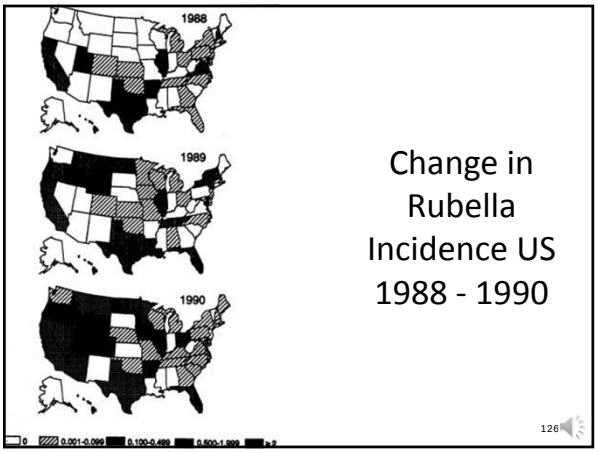
Pedestrian Death Rates



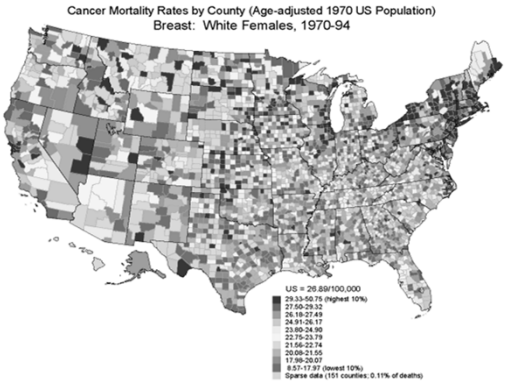
Place

- International variation
- Within country geographic variation
- Urban/rural differences
- Localized occurrences

125



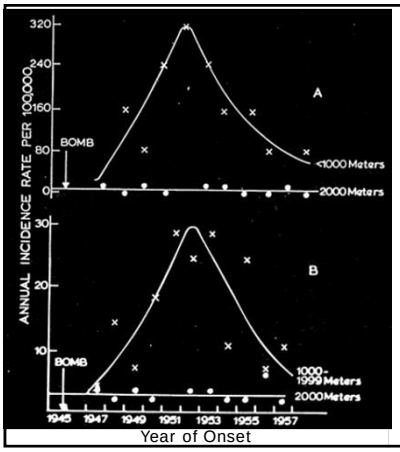
Geography of Breast Cancer



Time

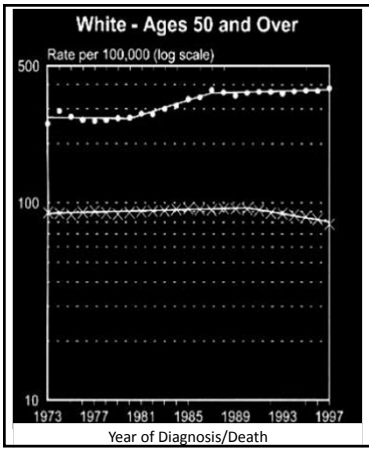
- Temporal variation
- Secular (long-term)
- Seasonal
- Cyclic
- Short-term
  - epidemic versus endemic

128



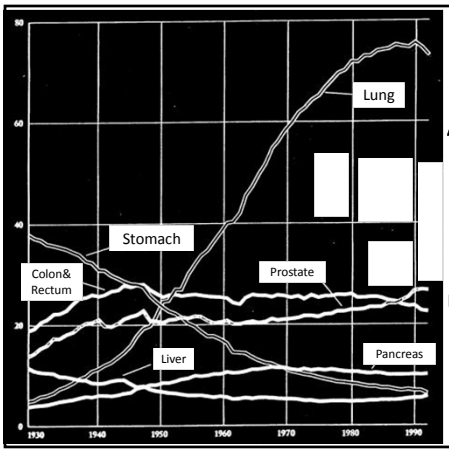
Annual Leukemia Incidence Among Atomic Bomb Survivors

129



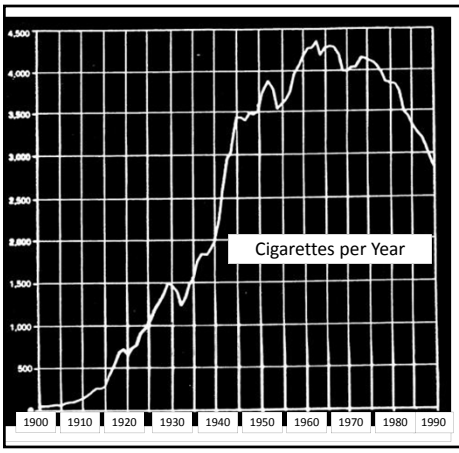
Breast Cancer Incidence & Mortality Rates

130



Age-adjusted Cancer Mortality Rates US 1930-1992

131

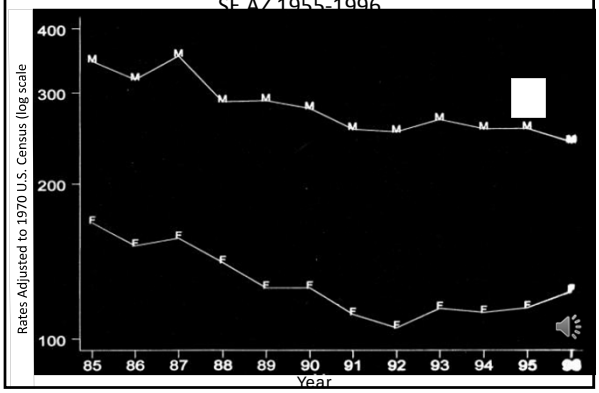


Per Capita Cigarette Consumption US Adults (18+ yrs) 1900 - 1990

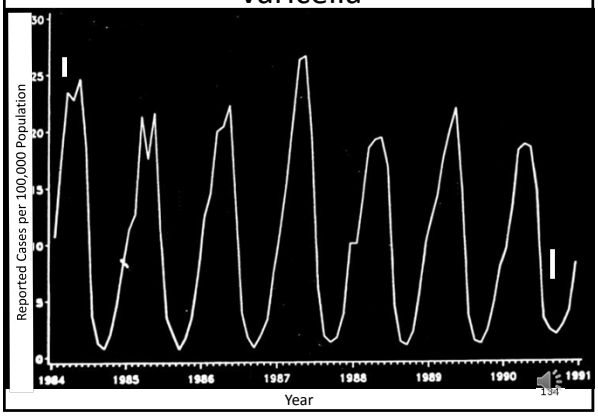
132

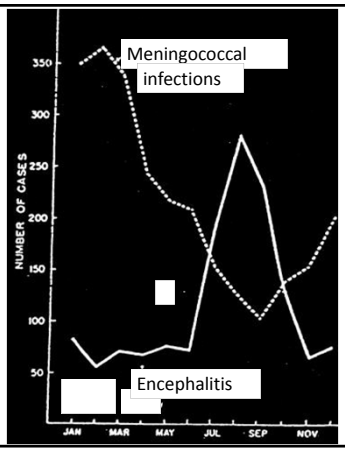
Squamous Cell Adjusted Rates

SE AZ 1955-1996

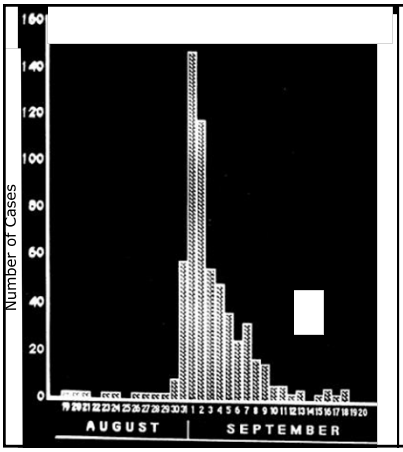


Varicella





Seasonal Variation of Diseases



Cholera Epidemic

Broad St & Golden Square

London 1800's

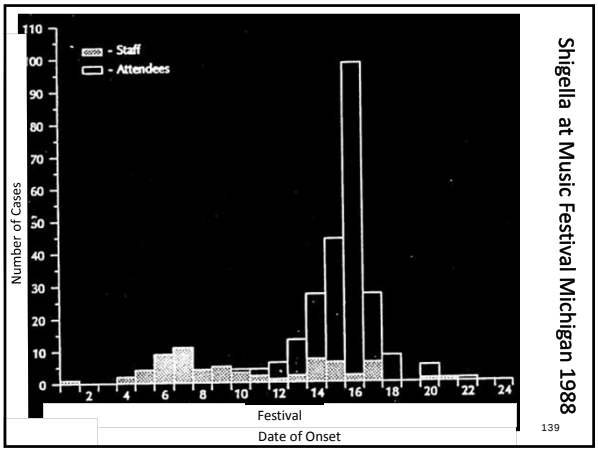
m08c

Epidemic vs Endemic Periods

- Endemic
  - Habitual presence of a disease within a geographic area
  - The usual occurrence of disease
- Epidemic
  - Time period when disease incidence exceeds what is normally expected
- Pandemic
  - Global epidemic

Epidemic vs Endemic Periods

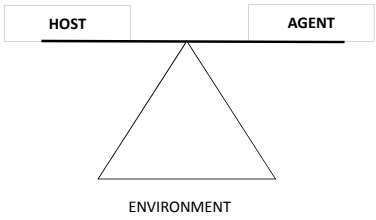
- Note: definition is broad
- No universally accepted levels
  - an "unusual" frequency of a disease above the usual or expected frequency for the population
  - compare past and present disease frequency
  - mathematical relationship
    - observed > expected
- Endemic rates (or expected frequency) can change over time
  - treatments
  - preventive measures developed
  - # susceptibles decrease



Shigella at Music Festival Michigan 1988

139

Ecologic Models of Disease



140

Ecologic Models of Disease

- Each model
  - a balance of forces determines an individual's health and response
- Change in any component may lead to detectable disease if no adequate reserve is available in another component

141

Pattern Detection

- Use Person, Place, Time to evaluate
  - how or whether
    - external agent
  - could combine or meet with
    - susceptible host
    - within an environment
  - that brings them all together

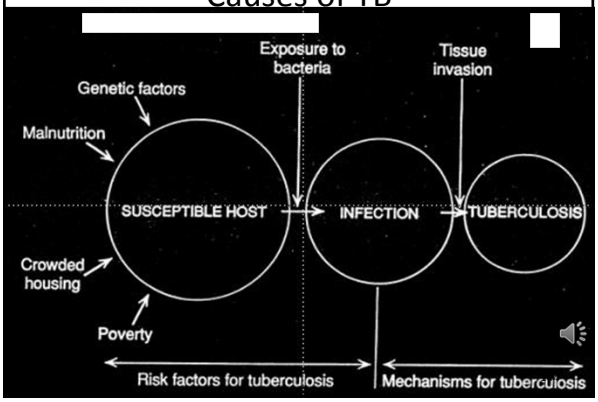
142

Example - TB

- Cause of tuberculosis using epidemiologic triangle
  - agent
  - host
  - environment

143

Causes of TB

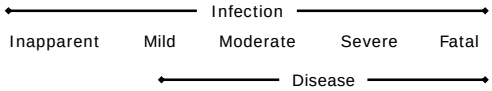


Disease Transmission

- Gradient of infection
- Spectrum of disease
- These concepts
  - allow determination of where to select cases
  - where to apply prevention strategies

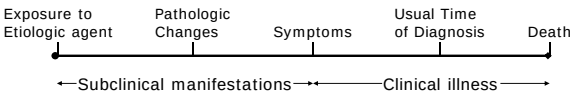
145

Gradient of Infection



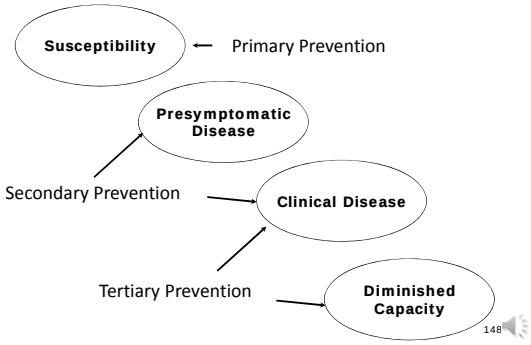
146

Spectrum of Disease



147

Spectrum of Prevention



148

Levels of Prevention

	Purpose	Examples
1º	• Prevent disease before it develops • Maintain health	• Smoking prevention programs • air pollution control • health education in schools
2º	• Early diagnosis & treatment • Disease in early stages • Restore or improve health	• Blood pressure screening for hypertension • vision screening in schools • case finding / referral for STDs
3º	• Reduce complications • Improve functioning & quality of life	• Rehabilitation programs for heart attack patients

149

Frequency Measures

Definitions & Concepts

150

Learning Objectives

- Understand the purpose and the calculation of measures of disease frequency

151

Epidemiology Questions

1. How many?

2. Compared to what?

3. Is it real?

152

How Many?

- How many people are affected?
- Resource allocation?
- Scope of the problem?

153

How Many?

- Describe
  - distribution or patterns of disease occurrence in human populations
- Identify disease determinants
- Measure frequency of disease (or outcome of interest) in the population

154

How Many?

- Define disease
- Count people with disease
- Determine who is at risk of being affected by the disease
  - Account for population at risk
- Account for the time frame for disease

155

Counts

- Most basic measure of disease frequency
- Good for
  - public health planners
    - to determine allocation of health care resources in a particular community

156

Frequency of AIDS

2 Hypothetical Cities

	# New Cases	Population
City A	58	25,000
City B	35	7,000

To Compare:

City A = 58/25,000 = 232/100,000

City B= 35/7,000 = 500/100,000

157

Counts

- Count data
  - A place to *start*
    - But not a place to *end*

158

Ratios

- Divide one quantity by another, without implying any specific relationship between the numerator and denominator
- Numerator is NOT a subset of the denominator
  - Person-doctor ratio
  - Person-hospital bed ratio
  - Male-Female ratio

159

Proportions

- Those in the numerator must also be included in the denominator
  - The numerator is a subset of the denominator
- 23/100 = 23% = 230 per 1,000

160

Population at Risk

- A central concern of epidemiology is to find and enumerate appropriate denominators in order to describe and to compare groups in a meaningful way

161

Choose a Denominator

What is the frequency of a drug side effect in this population?

Subject	# Treatments	# Reactions
individual 1	6 treatments	3 times had reaction
individual 2	4 treatments	1 reaction
individual 3	4 treatments	0 reaction
individual 4	1 treatment	0 reaction
individual 5	1 treatments	1 reaction

3/5 people who received the drug experienced a reaction

OR

5 reactions for every 16 treatments resulted in a reaction

162

Frequency of AIDS

2 Hypothetical Cities

	# New Cases	Population	Time Period
City A	58	25,000	1995
City B	35	7,000	1995-96

To Compare:

City A = 58/25,000 = 232/100,000

City B= 35/7,000 = 500/100,000

ty A = 58/25,000/1yr = 232/100,000/yr

ty B= 35/7,000/2yrs = 17.5/7000/1 yr = 250/100,000/yr

o compare rates use the same population unit (e.g. per 100,000 people) and time period (e.g. 1 year)

163

Rates

- Ratio in which there is a distinct relationship between the numerator and the denominator
- AND a measure of time is an intrinsic part of the denominator

164

Rates

$$\text{Rate} = \frac{\text{\# people with disease}}{\text{population at risk during the time period}}$$

165

Disease Frequency

- To quantify disease occurrence in a population
- Measures of disease frequency should account for:
  - # of individuals affected with the disease
  - Size of source population
  - Length of time the population was followed

166

Frequency Measures

Incidence & Prevalence

167

Learning Objectives

- Understand and distinguish between the two basic measures of disease frequency

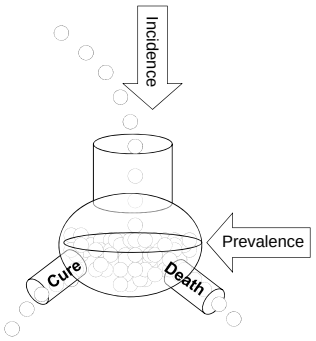
168

Basic Measures of Epidemiology

- Incidence
- Prevalence

169

Incidence & Prevalence



170

Incidence

- *New* disease occurring in a population or
- *New* disease developing in individuals at risk in a specific time period
  - Cumulative Incidence (or risk)
  - Incidence Rate (or incidence density)

171

Key Concepts

- New disease events
  - For diseases that can occur more than once, usually first occurrence of disease
- Population at risk (candidate population)
  - can't have disease already
  - should have relevant organs
- Time must pass for a person to move from health to disease

172

Cumulative Incidence

- Proportion of people who become diseased in a specific time period
- Is the likelihood or probability or RISK that an individual will develop the disease in the specified time period
- Assumes everyone in the population at the beginning of the time interval is followed for the entire interval

173

Cumulative Incidence

$$\text{Cumulative Incidence} = \frac{\text{\# people developing disease}}{\text{total population at risk}} \text{ per unit of time}$$

174

Cumulative incidence

Sudden Infant Death Syndrome (during 1st year of life)

Population	1,000 live births
Cases of SIDS	10
Cumulative Incidence	10/1,000 *or* 1% over one year

- Note that all live births are 'at risk'

175

Bacteriuria Among OC Users

- Population
  - Women
  - aged 16-49 years
  - free of bacteriuria at baseline
- Follow up 3years

176

Bacteriuria Among OC Users

	All women	Developed bacteriuria by 3yrs
Used OC	482	27
Non-user	1908	77
Total	2390	104

Cumulative Incidence of bacteriuria among OC users =

27/482 for time period = 5.6%

Women without bacteriuria and using OC have a 5.6% risk of developing bacteriuria within 3 years

177

Issues

- Cumulative incidence
  - calculation assumes that you have followed the entire population for the entire follow-up period
- Often you can't follow everyone for entire time period
- In a dynamic population, individuals enter population over time, become lost, etc.
- So length of follow-up is not uniform for all

178

Incidence Rate

- Measure of how fast new occurrences of the disease are arising in the population
- Accounts for varying time periods of observation of the 'at risk' population
- Is a true rate and requires the unit time as part of the denominator
- Incidence rates do not make assumption of complete follow-up

179

Incidence Rate

Incidence rate (IR) = $\frac{\text{\# new cases of disease}}{\text{person-time of observation in candidate population}}$

This measure is a true rate because it directly integrates time into the denominator.

Incidence rate also called incidence density*

*Multiple terms for the same concept… get used to it

180

Incidence Rate

Incidence rate (IR) = # new cases of disease
person-time of observation in candidate population

This measure is a true rate because it directly integrates time into the denominator.

Incidence rate also called incidence density*

*Multiple terms for the same concept… get used to it

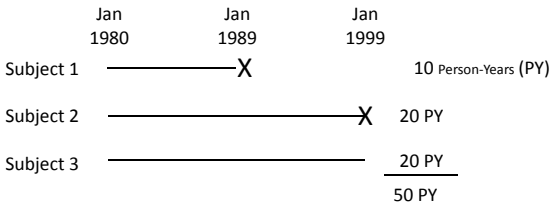
181

Person Time

- Painting a house
  - Hours worked:
    - Sue 3
    - Jim 4
    - Lisa 7
    - Jay 2
- Total Hours = 16
- How many person-hours?

182

Accrual of Person-Time



X = outcome of interest

Incident Rate = 2/50 PY

183

Accrual of Person-Time

- 100 people each followed 1 year = 100 py
- 10 people each followed 10 years = 100 py
- 50 people followed 1 year +
  25 people followed 2 years = 100 py
- Time unit for person-time
  - year, month or day
- Person-time unit
  - person-year, person-month, person-day

184

Incidence Rate

Example: Cohort study of risk of breast cancer among women with hyperthyroidism

- Followed 1,762 women ---> 30,324 py
- Average of 17 years of follow-up per woman
- Ascertained 61 cases of breast cancer
- Incidence rate = 61/30,324 py = 0.00201/y
  = 201 cases /100,000 py

185

Smoking & Stroke

Smoking Category	Cases of Stroke	Person-Years of Observation
Never smoked	70	395 594
Ex-smoker	65	232 712
Smoker	139	280 141
TOTAL	274	908 447

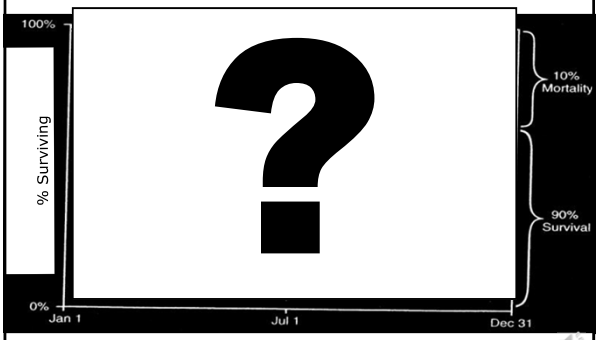
Incidence density for stroke among all =

274 / 908,447 = .0003016

or 30.16 stroke cases per 100,000 person-years

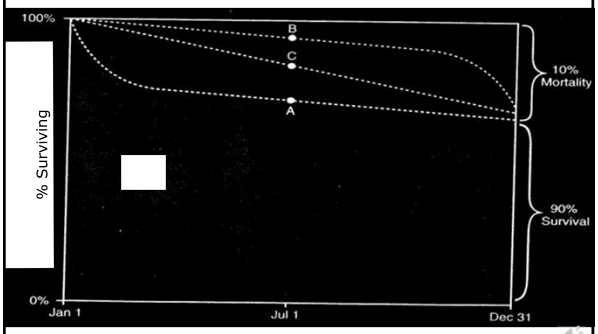
186

Importance of Time Interval



187

Importance of Time Interval



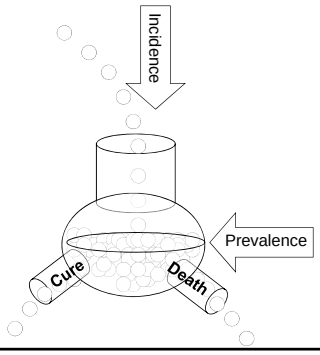
188

Prevalence

- Amount of disease present in a population
- Quantifies at a specific time the proportion of individuals in a population who have the disease
- Is a proportion and not a rate
- Estimates the probability that a person has a disease at that point in time

189

Incidence & Prevalence



190

Prevalence

191

Prevalence

$$\text{Prevalence} = \frac{\text{\# existing cases of disease}}{\text{\# in total population @ 1 time point or over a period}}$$

Example:

City A has 7000 people with arthritis on 1/1/1999

- Population of City A = 70,000
- Prevalence of arthritis on Jan 1st = .10 or 10%

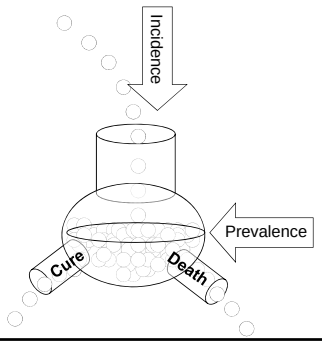
192

Point vs Period Prevalence

Prevalence	Numerator	Denominator
Point	Existing cases @ single time point	Total population
Period	Existing cases + new cases diagnosed over defined time period	Total population

193

Incidence & Prevalence



194

Prevalence and Incidence

- Prevalence depends on
  - incidence rate and
  - duration of disease
    - Duration: onset of disease to its termination

Low incidence with long duration → high prevalence

High incidence with short duration → low prevalence

195

Prevalence and Incidence

$$P / (1-P) = IR \times D$$

P = point prevalence of the disease
IR = incidence rate
D = duration of the disease

196

Requirements

- Conditions for equation to be true
  - Steady state
  - IR constant
  - Distribution of durations constant
- Note:
  - if prevalence of disease is low (<10%),
  - equation simplifies to:

$$P = IR \times D$$

197

Calculation

Figuring duration from prevalence and incidence

$$P = IR \times D$$

$$IR \times D = P$$

$$\frac{IR \times D}{IR} = \frac{P}{IR}$$

$$\frac{\cancel{IR} \times D}{\cancel{IR}} = \frac{P}{IR}$$

$$D = \frac{P}{IR}$$

198

Calculation

Figuring duration from prevalence and incidence

Lung cancer prevalence = 23/100,000 p

Lung cancer incidence rate = 45.9/100,000 py

$$D = \frac{P}{IR} = \frac{\underline{23/100,000}\ p}{45.9/100,000\ py} = 0.5 \text{ years}$$

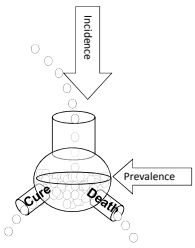
Conclusion: Individuals with lung cancer survive 6 months from diagnosis to death

199

Factors Influencing Prevalence

Increased by:

- Longer disease duration
- Prolongation of life without cure
- Increase in new cases
- In-migration of cases
- In-migration of susceptibles
- Out-migration of healthy people
- Improved diagnostic facilities/better reporting

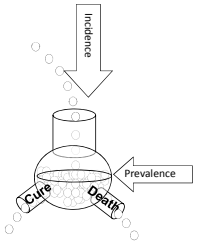


200

Factors Influencing Prevalence

Decreased by:

- High case-fatality rate
- Decrease in new cases
- In-migration of healthy people
- Out-migration of cases
- Improved cure rate



201

Exercises

Two surveys were done of the same community 12 months apart. Of 5,000 people surveyed the first time, 25 had antibodies to histoplasmosis. Twelve months later, 35 had antibodies, including the original 25. Calculate all estimates per 1,000 persons.

- Calculate the prevalence of people with histoplasmosis antibodies at the time of the second survey.
- Compare this prevalence with the 1-year cumulative incidence.

202

Other Frequency Measures

Learning Objectives

- Understand commonly used frequency measures

Specific Frequency Measures

Rate	Inc/Prev	Numerator	Denominator
Morbidity rate		New cases of nonfatal disease	Total population at risk
Mortality rate		# deaths from a disease (or all causes)	Total population
Case-fatality %		# deaths from a disease * 100	# cases of that disease

See Aschengrau p. 50 - 52

Specific Frequency Measures

Rate	Inc/Prev	Numerator	Denominator
Attack rate		# cases of a disease	Total population at risk, for limited period of observation
Disease rate at autopsy		# cases of a disease	# persons autopsies
Birth defect rate		# babies with a abnormality	# live births

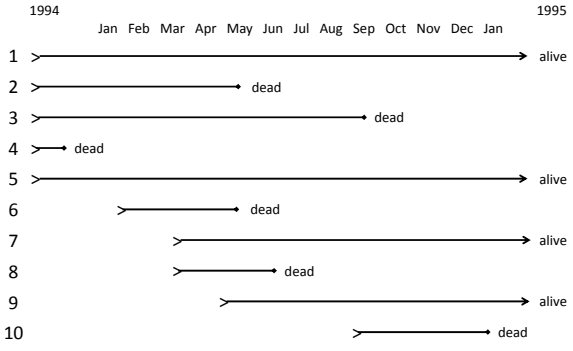
See Aschengrau p. 50 - 52

Name the frequency measure

1. % of infants enrolled in a day-care center who contracted impetigo during the course of an epidemic.

2. Percentage of potential army recruits rejected because of poor vision.

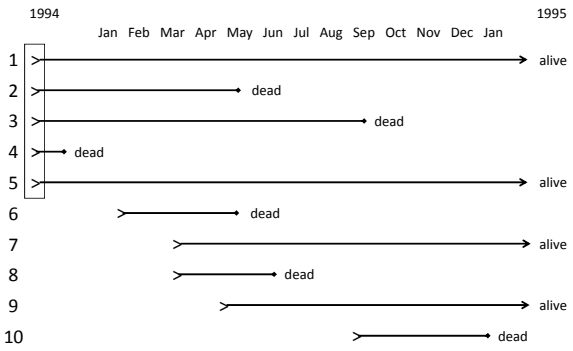
3. Percent of deceased males who are found to have prostate cancer at autopsy.



Assume that the 990 remaining individuals in the study did not become ill or die during the year of observation

Prevalence Practice

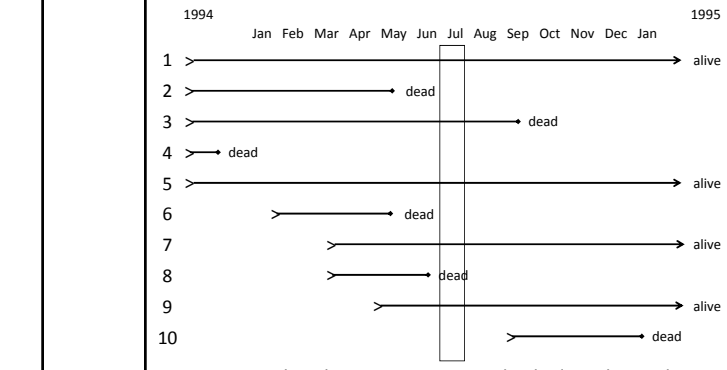
- Determine the Prevalence of the disease on:
  - January 1, 1994
  - July 1, 1994
  - December 31, 1994



Assume that the 990 remaining individuals in the study did not become ill or die during the year of observation

Prevalence Practice

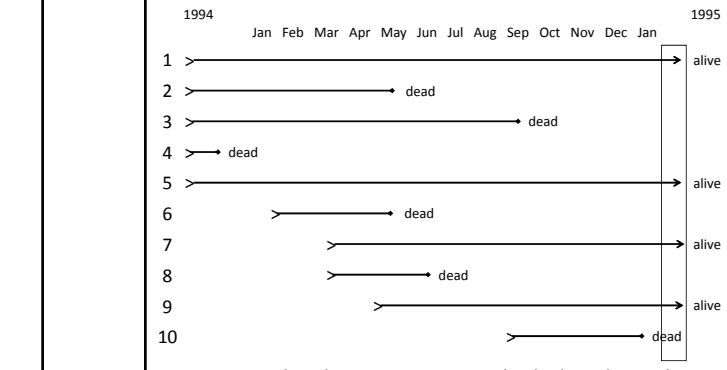
- Determine the Prevalence of the disease on:
  - January 1, 1994
  - July 1, 1994
  - December 31, 1994



Assume that the 990 remaining individuals in the study did not become ill or die during the year of observation

Prevalence Practice

- Determine the Prevalence of the disease on:
  - January 1, 1994
  - July 1, 1994
  - December 31, 1994



Assume that the 990 remaining individuals in the study did not become ill or die during the year of observation

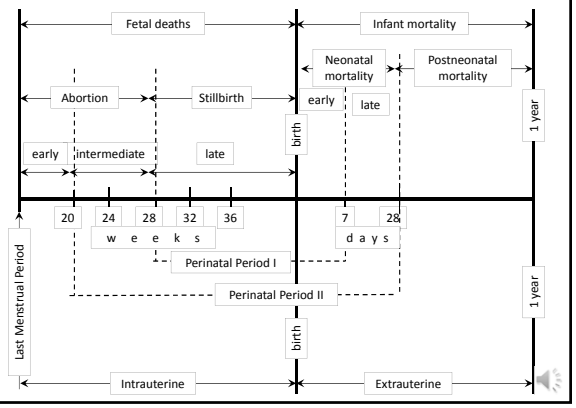
Prevalence Practice

- Determine the Prevalence of the disease on:
  - January 1, 1994
  - July 1, 1994
  - December 31, 1994

Other Rates

- Commonly used rates in public health
  - Infant mortality rates
  - Birth rates
- These can also be age and sex and race-specific rates

Early Life Mortality Measures



Measures of Association

Definition & Motivation

Comparing Risk

2 x 2 Tables

Learning Objectives

Understand the

- motivation &
- how to calculate measures of association
- the construction &
- usefulness of 2X2 tables

Epidemiology Questions

1 How many?

2 Compared to what?

3 Is it real?

Compared to What?

- 3 confirmed cases of disease in Phoenix

  Population of Phoenix: 1,321,045

3 / 1,321,045 = 2.27 per 1,000,000

If the disease is ***SMALLPOX …..***

Compared to what?

- 40 people with colds

  Stand on 1 foot 5, min/day

  - Average duration symptoms = 3 days

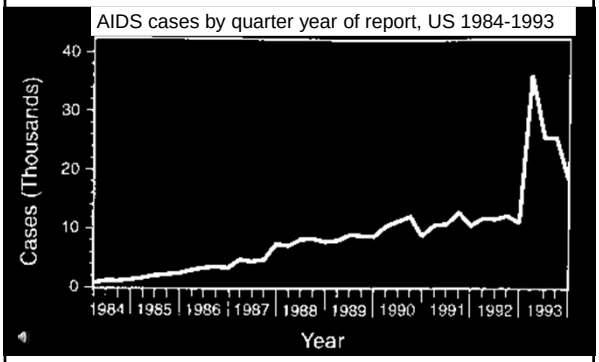
- 40 people with colds

  Usual treatment

  - Average duration symptoms = 6 days

Is it real?

AIDS cases by quarter year of report, US 1984-1993



From CDC: Update: Trends in AIDS diagnosis and reporting MMWR 43:827, 1994

Purpose: Measures of Association

Summarize relationship between exposure & disease

by comparing measures of disease frequency

Compared to what?

"Comparison is the essence of epidemiology"

- Aschengrau and Seage

Risk

$$\text{Risk} = \text{Incidence} = \frac{\#\text{ with Outcome}}{\text{Population at Risk}}$$

Comparing Risk

Exposed

$$R_e = I_e \frac{\#\ \textit{Exposed}\ \text{with Outcome}}{\textit{Exposed}\ \text{Population at Risk}}$$

Unexposed

$$R_u = I_u \frac{\#\ \textit{Unexposed}\ \text{with Outcome}}{\textit{Unexposed}\ \text{Population at Risk}}$$

Options for Comparison

1 Difference in risk between two groups:
exposed group & unexposed (comparison)

2 Ratio of risk between two groups:
exposed group & unexposed (comparison)

Difference Measures: Absolute Risk

Terms

- Risk difference
- Rate difference
- Incidence rate difference
- Cumulative Incidence difference

Attributable risk or rate

Difference Measures: Absolute Risk

Comparing risk of disease in exposed & unexposed by looking at difference -

subtracting one from the other

$$R_e - R_u$$

Risk Difference

Gives information on:

- absolute effect of exposure on disease risk
- public health effect of exposure ~

how much disease would be prevented if exposure were removed *assuming that the exposure causes the disease*

Risk Difference %

Disease associated with exposure as % of total

$$\frac{R_e - R_u}{R_T}$$

Risk Difference

People with hypertension:

Total people: 13,422
Non-fatal heart attacks: 117

People without hypertension:

Total people: 106,541
Non-fatal heart attacks: 125

Interpretation

The excess occurrence of non-fatal heart attack among these hypertensive women was 755 per 100,000

or

If hypertension causes non-fatal heart attacks, then 755 cases of non-fatal heart attack per 100,000 women could be eliminated if the hypertension were treated,

Comparing Risk

Population of 1,000

Disease	R_e	R_u	$R_e - R_u$
A	4/1000	2/1000	2/1000
B	102/1000	100/1000	2/1000
C	902/1000	900/1000	2/1000

Comparing Risk

Population of 1,000

Disease	R_e	R_u	R_e / R_u
A	4/1000	2/1000	2
B	102/1000	100/1000	1.02
C	902/1000	900/1000	1.002

Difference Measures:

Relative Risk

- Rate Ratio or Risk Ratio

  Compare:
  - risk of disease in exposed with
  - risk of disease in a comparison group *(usually unexposed)*

$$\frac{R_e}{R_u}$$

Relative Risk

Gives information on:

- relative effect of the exposure on disease
- how many times higher or lower disease risk is in the exposed compared to the unexposed
- commonly used in etiologic research

Relative vs Absolute Risk

	Annual Mortality Rate (per 100,000) Lung Cancer	Coronary Heart Disease
Non Smoker	10	413
Cigarette Smoker	140	669
RR	14.0 > > >	1.6
RD	130/100,000 per yr <	256/100,000 per yr

Interpretation

Cigarette smoking is a much stronger risk factor for lung cancer *(RR 14 vs 1.6)*

but,

elimination of cigarettes would prevent more deaths from CHD *(RD 256 vs 130)* assuming smoking is causally related to both diseases

Why is this so?

Population Risk Difference

- Excess disease occurrence associated with exposure in the total population
- Focus on most relevant exposures for health of a population
- Population Attributable Risk

Population Risk Difference

Formulas:

PRD = (RD) * (Pexp)

where Pexp = proportion of pop. exposed

RD = risk or rate difference

PRD = Rtotal – Runexp

where Rtotal = risk in total population

Runexp = risk among unexposed

Population Risk Difference %

Disease associated with exposure as % of total

$$\frac{R_{Total} - R_{unexposed}}{R_{Total}}$$

Comparing Risk

Exposed

$$R_e = I_e = \frac{\#\ \textit{Exposed}\ \text{with Outcome}}{\textit{Exposed}\ \text{Population at Risk}}$$

Unexposed

$$R_u = I_u = \frac{\#\ \textit{Unexposed}\ \text{with Outcome}}{\textit{Unexposed}\ \text{Population at Risk}}$$

2 X 2 Tables

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

Risk in the Exposed

$$R_e = I_e = \frac{\#\ \textit{Exposed}\ \text{with Outcome}}{\textit{Exposed}\ \text{Population at Risk}}$$

2 X 2 Tables

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

2 X 2 Tables

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

Risk in the Exposed

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

$$R_e = I_e = \frac{\#\ \textit{Exposed}\ \text{with Outcome}}{\textit{Exposed}\ \text{Population at Risk}} = \frac{a}{a+b}$$

Risk in the Unexposed

$$R_u = I_u = \frac{\#\ \textit{Unexposed}\ \text{with Outcome}}{\textit{Unexposed}\ \text{Population at Risk}}$$

2 X 2 Tables

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

2 X 2 Tables

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

Risk in the Unexposed

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

$$R_u = I_u = \frac{\#\ \textit{Unexposed}\ \text{with Outcome}}{\textit{Unexposed}\ \text{Population at Risk}} = \frac{c}{c+d}$$

Relative Risk

Just a little bit of algebra…

$$\frac{R_e}{R_u}$$

$$R_e = \frac{a}{a+b} \qquad R_u = \frac{c}{c+d}$$

$$\frac{R_e}{R_u} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

Relative Risk

$$\frac{R_e}{R_u} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

$$= \frac{a}{a+b} \times \frac{c+d}{c}$$

$$= \frac{a(c+d)}{c(a+b)}$$

2 X 2 Tables

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

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

Legionnaire's Disease

American Legion Convention

- 7/21 – 7/24 1976
- Delegates
- Family members
- Other legionnaires
- Pneumonia onset 7/20 – 8/30
- No new cases after 8/30

Legionnaire's Disease

		Developed Legionnaire's Disease: Yes	No	Total
Convention Status	Delegate	125	1724	1849
	Non-Delegate	3	759	762

What measure of risk can be calculated for each group?

What is the rate of Legionnaires' Disease among the delegates? among the non-delegates?

Cumulative Incidence

CI_{Del} *125/1849 = .068 in 41*

CI_{Non} *3/762 = .004 in 41 days*

Legionnaire's Disease

		Developed Legionnaire's Disease: Yes	No	Total
Convention Status	Delegate	125	1724	1849
	Non-Delegate	3	759	762

Cumulative Incidence

CI_{Del} *125/1849 = .068 in 41 days*

CI_{Non} *3/762 = .004 in 41 days*

Legionnaire's Disease

Calculate the rate ratio of Legionnaires' Disease among delegates compared to non-delegates

RR = 0.068 / 0.004 = 17

State in words the meaning of this rate ratio

Members of the Legionnaires who were delegates at the convention had a risk of pneumonia 17 times greater than those Legionnaires who were not delegates

Legionnaire's Disease

Calculate the rate difference of Legionnaires' Disease for delegates

RD = 0.068 - 0.004 = 0.064 or 64 / 1

State in words the meaning of this rate

There were 64/1000 more cases of pneumonia among delegates than among non-delegates at the Legionnaire's convention

Population Risk Difference

		Non Fatal Heart Attack		
		Yes	No	Total
Hypertension	Yes	117	13,305	13,422
	No	125	106,416	106,541
	Total	242	119,721	119,963

$PRD = (RD) \times (P_{exp})$

$= (R_e - R_u) \times (P_{exp})$

$=[(117/13,422) - (125/106,541)] \times (13,422/119,693)$

$= (0.00755) \times (0.112) = 0.00085$

$= 85 / 100,000$

Population Risk Difference

		Non Fatal Heart Attack		
		Yes	No	Total
Hypertension	Yes	117	13,305	13,422
	No	125	106,416	106,541
	Total	242	119,721	119,963

$PRD = R_{total} - R_{unexp}$

$= 242/119,963 - 125 /106,541$

$= 0.00202 - 0.00117$
$= 0.00085$
$= 85 / 100,000$

Interpretation

Hypertension results in an excess incidence of 85/100,000 non-fatal heart attacks in the total study population

o

if hypertension were eliminated, 85 /100,000 cases of non-fatal heart attacks could be eliminated among the total study population *assuming that hypertension <u>causes</u> heart attacks*

Population Risk Difference

- Depends on prevalence of exposure

What would the excess of non-fatal heart attack due to hypertension be if the prevalence of hypertension were 1% rather than 11.2%?

A prevalent risk factor that is "weak" (low RR) could account for more disease in a population than a stronger risk factor that is rarely present

Smoking and Lung Cancer*

Simple Rates

- Death rate from lung cancer in smokers 0.96 / 1,000 / year
- Death rate from lung cancer in non-smokers: 0.07 / 1,000 / year
- Prevalence of smoking in population: 56%

* Estimated data from Doll and Hill. Br J Med 1:1399-1410, 1964

Smoking and Lung Cancer*

Compared Rates

Rate Ratio: $\frac{0.96 \text{ / } 1{,}000 \text{ / yr}}{0.07 \text{ / } 1{,}000 \text{ / yr}} = 13.7$

Rate Difference:

0.96 / 1,000 / yr – 0.07 / 1,000 / yr

= 0.89 / 1,000 / yr

Population Risk Difference:

0.89 / 1,000 / yr x 0.56 = 0.50 / 1,000 / yr

* Estimated data from Doll and Hill. Br J Med 1:1399-1410, 1964

Measures of Association

Definition & Motivation

Comparing Risk

2 x 2 Tables

Classification

Motivation & Theory

268

Learning Objectives

- Understand the concept of classification systems and their relevance to epidemiology

269

Classification

- Classification
  - arrangement into groups of
    - objects,
    - concepts, or
    - information
  - based on their relationships or properties
- Fundamental to the quantitative study of any phenomenon
- Necessary for all scientific generalization
  - therefore, an essential element in statistical (and epidemiologic) methods

270

Purposes of Classification

- Summarize information
- Facilitate retrieval of information
- Describe structure & relation of items to similar items
- Simplify relationships so that generalizations can be made

271

The Diseases and Casualties This Year Being 1632

Abortive & stillborn	445	Drowned	34	Palsie	
Afrighted	1	Executed & prest to death	18	Piles	8
Aged	628	Falling sickness	7	Plague	8
Ague	43	Fever	1108	Planet	13
Apoplex, and meagrom	17	Fistula	13	Pleurisie & spleen	36
Bit with a mad dog	1	Flox & Small pox	531	Purples & Spotted fever	38
Bloody flux, scowring & flux	348	French pox	12	Quinsie	7
Brused, issues, sores & ulcers	28	Gangrene	5	Rising of the lights	98
Burnt & scalded	5	Gowt	4	Sciatica	1
Burst and rupture	9	Grief	11	Scurvey & itch	9
Cancer & wolf	10	Jaundies	43	Suddenly	62
Canker	1	Jawfain	8	Surfet	86
Childbed	171	Impostume	74	Swine pox	6
Chrisomes & infants	2268	Kil'd by several accident	46	Teeth	470
Cold & cough	55	King's evil	38	Thrush, sore mouth	40
Colick, stone, and strangury	56	Lethargie	2	Tympany	13
Consumption	1797	Lunatique	5	Tissick	34
Convulsion	241	Made away themselves	15	Vomiting	1
Cut of the stone	5	Measles	80	Worms	27
Dead in the street & starved	6	Murthered	7		
Dropsie & swelling	267	Over-laid/starved at nurse	7		

272

Acute Yellow Atrophy of the Liver

1900 Conditions included	*1968 ICD Code*
AYA due to poisons chloroform, carbon tetrachloride, etc.)	N. Code (Accidental poisonings)
AYA during pregnancy or puerperium	639 Toxemia of pregnancy and puerperium – complications of pregnancy
AYA – NEC	570 Acute necrosis of liver Diseases of liver, gallbladder and pancreas
Weil's disease	100 Leptospirosis – infectious diseases
Infective Hepatitis	070 Infectious hepatitis - infectious diseases
Alcoholic hepatitis	571.0 Alcoholic cirrhosis of liver – diseases of liver, etc.
Serum hepatitis	999.2 Serum hepatitis – complications of medical care
Hepatitis - NEC	573.0 Hepatitis – diseases of liver, etc.

273

Classification

- Assignment to predesignated classes on the basis of perceived common characteristics
  - Naturalness
  - Exhaustiveness
  - Usefulness
  - Simplicity

Last. A Dictionary of Epidemiology

274

Classification Schema

- Groups that have a characteristic in common
- Groups in which the items have several or many common properties but do not necessarily agree in any one characteristic

275

Types of Classification Systems

- Nature of etiologic agent
- Nature of disease process/pathology
- Body organ system affected
- Method of treatment
- Method of transmission
- Method of entry to or exit from body
- Factors influencing exposure/susceptibility of host
- Degree of incapacity

276

Classification

- Classification ≠ nomenclature
- Classification is a method of generalization

277

Classification Caveats

- "keyhole" effect
- Incorporating new information into the classification system
- Observer errors by omission or commission

278

Classification Systems

279

Learning Objectives

- Understand the concept of classification systems and their relevance to epidemiology

280

Nosology

1. A systematic arrangement, or Classification, of diseases

2. That branch of medical science which treats of diseases, or of the classification of diseases

   Origin: Gr. Disease: cf. F. Nosologie

- http://medical-dictionary.com/dictionaryresults.php

1. classifying of diseases: the branch of medicine concerned with the classification and description of known diseases
2. classified list of diseases: a completed classification of known diseases

- http://encarta.msn.com/dictionary_/nosology.html

281

ICD Codes

- International Classification of Diseases, Tenth Revision (ICD-10)
- World Health Organization (WHO)
- designed to promote international comparability in
  - collection,
  - processing,
  - classification, and
  - presentation of
- mortality statistics

282

ICD Codes

- Provides format for reporting causes of death on the death certificate
- Reported conditions translate into medical codes
  - using classification structure and
  - selection and modification rules
- Coding rules improve the usefulness of mortality statistics by
  - giving preference to certain categories,
  - consolidating conditions, and
  - systematically selecting a single cause of death from a reported sequence of conditions
- Single selected cause for tabulation is the underlying cause of death
  - other causes are nonunderlying causes of death
  - combination of underlying/nonunderlying is the multiple causes of death

283

ICD Codes

Chapter	Codes	Title
I	A00–B99	Certain infectious and parasitic diseases
II	C00–D48	Neoplasms
III	D50–D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
IV	E00–E90	Endocrine, nutritional and metabolic diseases
V	F00–F99	Mental and behavioural disorders
VI	G00–G99	Diseases of the nervous system
VII	H00–H59	Diseases of the eye and adnexa
VIII	H60–H95	Diseases of the ear and mastoid process
IX	I00–I99	Diseases of the circulatory system
X	J00–J99	Diseases of the respiratory system
XI	K00–K93	Diseases of the digestive system
XII	L00–L99	Diseases of the skin and subcutaneous tissue
XIII	M00–M99	Diseases of the musculoskeletal system and connective tissue
XIV	N00–N99	Diseases of the genitourinary system
XV	O00–O99	Pregnancy, childbirth and the puerperium
XVI	P00–P96	Certain conditions originating in the perinatal period
XVII	Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities
XVIII	R00–R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
XIX	S00–T98	Injury, poisoning and certain other consequences of external causes
XX	V01–Y98	External causes of morbidity and mortality
XXI	Z00–Z99	Factors influencing health status and contact with health services
XXII	U00–U99	Codes for special purposes

284

ICD Codes

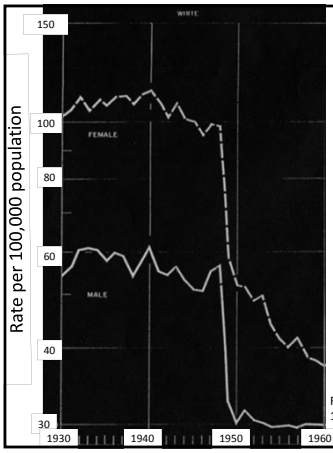
Chapter	Codes	Title
I	A00–B99	Certain infectious and parasitic diseases
II	C00–D48	Neoplasms
III	D50–D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
IV	E00–E90	Endocrine, nutritional and metabolic diseases
V	F00–F99	Mental and behavioural disorders
VI	G00–G99	Diseases of the nervous system
VII	H00–H59	Diseases of the eye and adnexa
VIII	H60–H95	Diseases of the ear and mastoid process
IX	I00–I99	Diseases of the circulatory system
X	J00–J99	Diseases of the respiratory system
XI	K00–K93	Diseases of the digestive system
XII	L00–L99	Diseases of the skin and subcutaneous tissue
XIII	M00–M99	Diseases of the musculoskeletal system and connective tissue
XIV	N00–N99	Diseases of the genitourinary system
XV	O00–O99	Pregnancy, childbirth and the puerperium
XVI	P00–P96	Certain conditions originating in the perinatal period
XVII	Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities
XVIII	R00–R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
XIX	S00–T98	Injury, poisoning and certain other consequences of external causes
XX	V01–Y98	External causes of morbidity and mortality
XXI	Z00–Z99	Factors influencing health status and contact with health services
XXII	U00–U99	Codes for special purposes

285

ICD Codes

- **(J00–J06) Acute upper respiratory infections**
- (J00) Acute nasopharyngitis (common cold)
- (J01) Acute sinusitis
- (J02) Acute pharyngitis
  - (J02.0) Streptococcal pharyngitis
    - Strep throat
  - (J02.8) Acute pharyngitis due to other specified organisms
  - (J02.9) Acute pharyngitis, unspecified
- (J03) Acute tonsillitis
- (J04) Acute laryngitis and tracheitis
  - (J04.0) Acute laryngitis
  - (J04.1) Acute tracheitis
  - (J04.2) Acute laryngotracheitis

286



Drop in death rates for diabetes among 55-to 64-year-old men and women, United States, 1930-1960, due to changes in ICD coding

From US Public Health Service publication No. 1000, series 3, No. 1.

287

DSM-IV / DSM-5 Codes

- Diagnostic and Statistical Manual of Mental Disorders, Fourth Revision
- Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision (May 2013)

288

Rheumatoid Arthritis Criteria*

American Rheumatism Association	New York
1. Morning stiffness	1. History of episode of three painful limb joints**
2. Joint tenderness or pain on motion	2. Swelling, limitation, subluxation, or ankylosis of three limb joints (must include a hand, wrist, or foot and symmetry of one joint pair and must exclude distal interphalangeal joint, fifth proximal interphalangeal joint, 1st metatarsophalangeal joints, and hips
3. Soft-tissue swelling of one joint	
4. Soft-tissue swelling of a second joint (within 3 months)	
5. Soft-tissue swelling of symmetrical joints (excludes distal interphalangeal joint)	
6. Subcutaneous nodules	
7. X-ray changes	3. X-ray changes (erosions)
8. Serum positive for rheumatoid factors	4. Serum positive for rheumatoid factors

289

Rheumatoid Arthritis Criteria*

* A score of
  - 3 or 4 points: "probable" rheumatoid arthritis
  - 5+ points: "definite"

** Count each joint group (e.g. proximal interphalangeal joints as one joint, scoring each side separately

From O'Sullivan JB, Cathcart ES. The prevalence of rheumatoid arthritis. Ann Intern Med 76:573. 1972.

290

Classification

Birth & Death Certificates

291

Learning Objectives

- Understand the concept of classification systems and their relevance to epidemiology

292

Death Certificates

1. Complete reporting of cases
2. Accurate diagnosis by the physician
3. Current and accurate census data are needed to provide the estimated population at risk in the denominator
4. Completion of medical records and death certificates by physicians who understand that the data might be useful for research purposes.
5. Medical terminology and definitions should be similar in different geographical areas
6. Technicians who collate and process the data should be well trained in nosology

293

Death Certificates

Term	Definition	Comment
Underlying cause of death (synonyms: primary, proximate)	The disease or injury that initiated the train of morbid events resulting in death, or the circumstances or violence that produced the fatal injury	In the absence of the underlying cause, the patient would be alive today
Immediate cause death	The disease, injury or complication that directly precedes death	The ultimate and final consequence of the underlying cause

294

Death Certificates

Term	Definition	Comment
Intervening cause(s) of death	Other conditions that contribute to death and are a result of underlying cause	The conditions are listed in physiologic sequence
Mechanism of death (synonym: mode)	A physiologic derangement or biochemical disturbance produced by a cause of death	The means by which cause exerts its effect
Manner of death	Explanation of how the cause of death arose	Natural or unnatural

295

Death Certificates

Part I

A. Immediate Cause

Due to, or as a consequence of

B. Intermediate Cause

Due to, or as a consequence of

C. Intermediate Cause

Due to, or as a consequence of

D. Underlying Cause

Part II OTHER SIGNIFICANT CONDITIONS: Conditions contributing to death but not resulting in the underlying cause of death in Part I

296

Death Certificates

Part I

A. Chronic ischemic heart disease

Due to, or as a consequence of

B. Atherosclerotic coronary artery disease

Due to, or as a consequence of

C.

Due to, or as a consequence of

D.

Part II OTHER SIGNIFICANT CONDITIONS:

Hypertension, diabetes mellitus

297

Death Certificates

Part I

A. Upper gastrointestinal hemorrhage

Due to, or as a consequence of

B. Undetermined natural cause(s)

Due to, or as a consequence of

C.

Due to, or as a consequence of

D.

Part II OTHER SIGNIFICANT CONDITIONS:

298

Death Certificates

Part I

A. *Pneumocystis carinii* pneumonia — Approximate interval Between Onset and Death

Due to, or as a consequence of

B. Acquired immunodeficiency syndrome

Due to, or as a consequence of

C. Human immunodeficiency virus infection

Due to, or as a consequence of

D.

Part II OTHER SIGNIFICANT CONDITIONS:

Risk factor: Intravenous drug abuse

299

Certificate Exercises

- Can be found in the Discussion area of the course D2L site

300

Take Home Message

- How are cases defined?
  - Who counts
  - Who doesn't count
  - Why?
  - How does this affect conclusions?

301

Sources of Data

Learning Objectives

- Understand how to locate sources of data &
- Evaluate its quality & utility

Objectives

- Criteria for assessment of
  - Quality &
  - Utility of data sources
- Identify government sources of data using
  - Internet
  - Library

Types of data

- Primary data
- Secondary data

Individual data
identified at individual level

Aggregate data
identifiable only for groups of people
not for individuals in the group

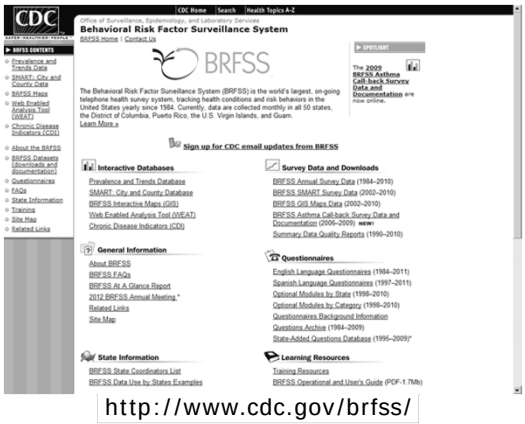
Criteria for Evaluation

- Epidemiologic Data
  - Quality
    - Who collected the data?
    - How was the data collected?
    - Completeness of population coverage
  - Utility
    - Nature of the data
    - Availability of the data
    - Completeness of population coverage
- Value & limitations

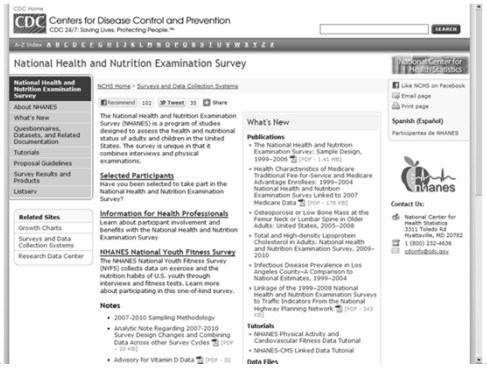
General Sources of Data

- Vital statistics
- Case registries
- Physician records
- Hospital and clinic records
- Surveys of the general population
  - NCHS – see webpage, next slide
  - BRFSS – see webpage, slide after next
  - NHANES - see webpage, third slide

http://www.cdc.gov/nchs/default.htm



http://www.cdc.gov/brfss/



http://www.cdc.gov/nchs/nhanes.htm

Morbidity & Mortality Statistics

- Disease reporting
  - Communicable diseases, cancer registries
- Data accumulated as a byproduct of insurance or medical plans
- Tax-financed public assistance and medical care plans
  - VA, Armed forces, state & federal medical care plans
- Hospitals & clinics

Morbidity & Mortality Statistics

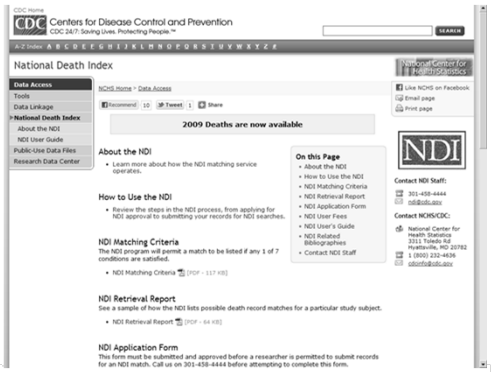
- Absenteeism records
  - Industry, schools
- Pre-employment and periodic physical exams
  - schools/industry
- Case-finding programs
- Selective service records
- Morbidity surveys on population samples

Epidemiologic Data Sources

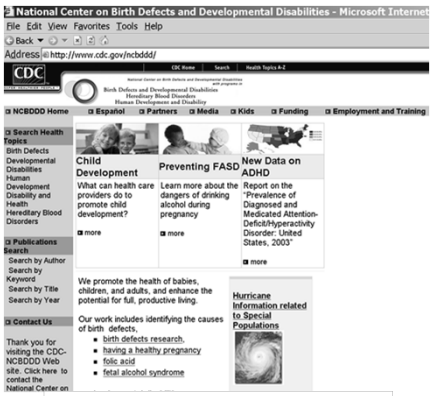
Data	Description	Availability	Population Coverage	Value & Limitations
Mortality statistics	Data from registration of vital events	Annually, from vital registration systems & political subdivisions	Complete	Useful for studying mortality
Medical data from birth records	Data on congenital abnormalities, complications of pregnancy, birth weight, etc	Annually, from vital registration systems & political subdivisions	Complete	Routinely available; some aspects may be incompletely reported

- see NDI webpage, next slide

- see National Centers on Birth Defects and Developmental Disabilities webpage, slide after next



http://www.cdc.gov/nchs/ndi.htm#about

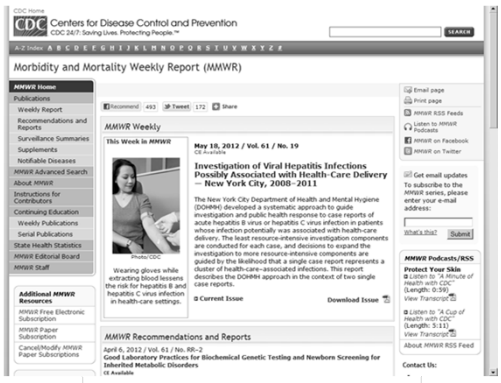


http://www.cdc.gov/ncbddd/

Epidemiologic Data Sources

Data	Description	Availability	Population Coverage	Value & Limitations
Reportable disease statistics	Based on physician reports, new cases of notifiable diseases	Weekly reports for the United States	Complete?	Useful for detection of outbreaks
Data on diseases treated in special clinics	Dependent on nature of clinic;	Generally not available without special approval	Not determinable	Counts of patients treated; difficult to estimate prevalence population denominator unknown

- see Morbidity & Mortality Weekly Reports webpage, next slide



http://www.cdc.gov/mmwr/

Limitations of Hospital Data

- Hospital admissions
  - Selective on:
    - Personal characteristics
    - Severity of disease
    - Associated conditions
    - Admission policies
- Hospital records - not designed for research
  - Incomplete, illegible, missing
  - Variable in diagnostic quality
- Populations at risk (denominators) ?

Useful Websites

- Population Data (Census)
  - http://www.census.gov
- NCHS
  - http://www.cdc.gov/nchs
- CDC
  - http://www.cdc.gov
  - http://www.cdc.gov/mmwr
- Cancer
  - http://seer.cancer.gov
- Health care atlas
  - http://www.dartmouthatlas.org

Useful Websites

- Agency for Toxic Substances & Disease Registry
  - http://www.atsdr.cdc.gov/
- US Dept Health & Human Services Data & Stats
  - http://www.hhs-stat.net/
- EPA Toxics Release Inventory
  - http://www.epa.gov/triexplorer/
- National Health Interview Survey
  - http://www.cdc.gov/nchs/nhis.htm
- National Cancer Institute
  - http://www.cancer.gov/
- National Institute of Environmental Health Sciences
  - http://www.niehs.nih.gov/home.htm

Useful Websites

- National Institutes of Health
  - http://www.nih.gov/
- National Toxicology Program
  - http://ntp-server.niehs.nih.gov/
- Arizona Health Statistics
  - http://www.hs.state.az.us

Study Design

Introduction

322

Learning Objectives

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

323

Evolution of Understanding Cause

1. Someone notices a pattern / available data used to examine patterns [Descriptive studies] hypotheses are generated
2. Additional data collected to test hypotheses [Analytic: observational studies]

Laboratory work

3. Treatment / prevention options developed and tested [Analytic: Experimental studies]

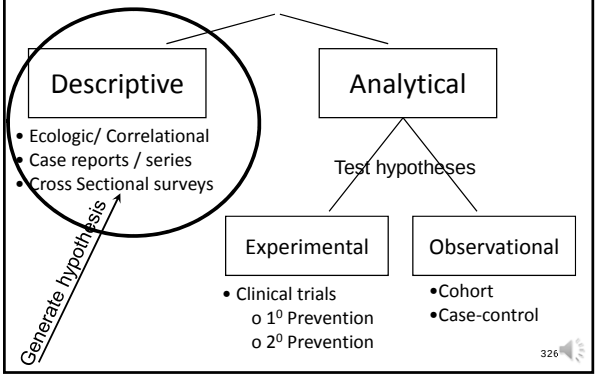
324

Scientific Method

1. Observation (descriptive studies)
2. Tentative description: *hypothesis*
  – consistent with observations
3. Hypothesis generates predictions
4. Test predictions (analytic studies)
  – Experimentation
  – Further *observations*
5. Modify the hypothesis according to results
6. Repeat 3 - 5 as necessary

325

Epi Study Types



326

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

327

Epi Study Designs

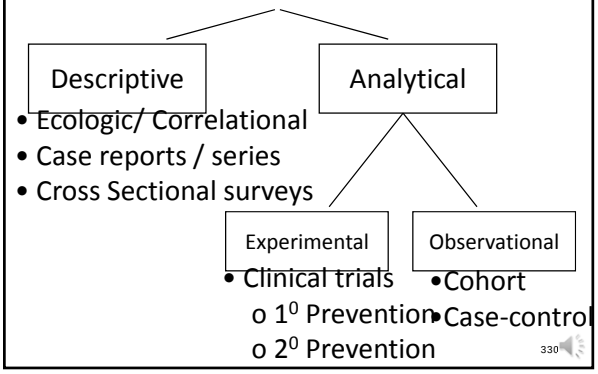
328

Learning Objectives

- Understand the role of descriptive studies in epidemiology

329

Epi Study Types



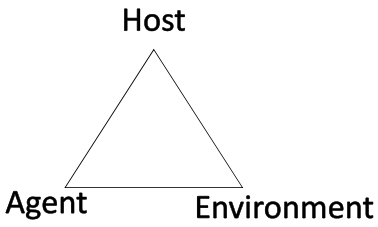
330

Descriptive Epidemiology

- Describe patterns of disease by
  - person,
  - place, and
  - time
- Person: Who is getting the disease?
  - Sex, religion, race, educational level etc
- Place
  - Mapping, cluster analysis, migration studies
- Time
  - Changes over time, seasonal effects

331

Descriptive Studies



332

Descriptive Studies

- Usually
  - Early
  - Quick
  - Easy
- Hypothesis generation
- Evaluation of trends

333

Case Reports – Case Series

- Describe experiences
  - single patient
  - group of patients with similar diagnosis

334

Minamata Disease

- Minamata City, Japan 1956
  - Patient with unusual neurologic symptoms
    - Sensory disorder in distal portion of extremities
    - Cerebellar ataxia
    - Bilateral concentric constriction of visual field
    - Central disorder of ocular movement
    - Central hearing impairments
    - Central disequilibrium

335

Minamata Disease

- CNS poisoning by Methylmercury
  - produced as by-product of acetaldehyde manufacture
  - methylmercury contaminated fish & shellfish

336

Toxic Oil

- Spain 1981
  - 19,828 cases
  - 315 deaths
    - Cough
    - Dyspnea
    - Pleuritic chest pain
    - Headache
    - Fever
    - Bilateral pulmonary infiltrates

337

Toxic Oil

- 51 unlabelled containers of "cooking oil"

	Cases	Comparison Group
% Eating Rapeseed oil	*100%*	*6%*

338

Kaposi's Sarcoma

Borkovic, S.P.; Schwartz, R.A. Kaposi's sarcoma presenting in the homosexual man -- a new and striking phenomenon! Ariz-Med. 1981 38(12): 902-4

PUBLICATION TYPE: Case-Reports; Journal-Article

339

Case Reports / Case Series Advantages

- Hypotheses generating
- Evidence of epidemic or unusual event

340

Case Reports / Case Series Caveats

- Cannot test hypothesis
  - Which basic epi question cannot be answered?

341

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

342

Ecologic Studies

- Ecologic / correlational
  - Data on whole populations
    - Skin cancer incidence vs latitude
    - Average dietary fat intake and breast cancer incidence
  - Measure of association is correlation coefficient

343

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

344

Ecologic Studies - Caveat

- Are the individuals with the exposure the same individuals who have the outcome?

345

Descriptive Studies

- Descriptive
  - Ecologic / correlational
  - Case reports / case series
  - Cross-sectional surveys
- Analytical

346

Cross Sectional Surveys

- Cross-sectional surveys
  - Exposure status and disease status determined at the *same time*
  - "Snapshot"
  - Hypothesis generating

347

Cross-sectional Survey – Advantages

- Quick
- Cheap
- Provides "snapshot" of the health experiences of the population
- Can provide prevalence of outcome in study population

348

Cross-sectional Survey – Caveats

- Difficult to determine temporal relationship

349

Design Characteristics

1. # of observations
2. Data collection methods
3. Timing of data collection

4. Unit of observation
5. Availability of subjects
6. Method of defining "*study population*"
7. Measure of association

350

Descriptive Studies

- Usually:
  - Early
  - Quick
  - Easy
- Hypothesis generation
- Evaluation of trends

351