This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike License. Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.

Copyright 2008, The Johns Hopkins University Sukon Kanchanaraksa, and Marie Diener-West. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided “AS IS”; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Biostatistics and Epidemiology within the Paradigm of Public Health

Sukon Kanchanaraksa, PhD
Marie Diener-West, PhD
Johns Hopkins University

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Section A

Biostatistics and Epidemiology within the Paradigm of Public Health

Steps in the Paradigm of Public Health

- Define the problem
- Measure its magnitude
- Understand the key determinants
- Develop intervention/prevention strategies
- Set policy/priorities
- Implement and evaluate

Quantitative Methods

- **Epidemiology** and **biostatistics** are the basic sciences of public health
- Public health investigations use **quantitative methods**, which combine the two disciplines of epidemiology and biostatistics
- Epidemiology is about the understanding of disease development and the methods used to uncover the etiology, progression, and treatment of the disease
- Information (data) is collected to investigate a question
- The methods and tools of biostatistics are used to analyze the data to aid decision making

Epidemiology

- **Epidemiology** is the study of the distribution and determinants of health, disease, or injury in human populations and the application of this study to the control of health problems
  - Examples
    - National and local surveillance system (cancer, AIDS, occurrence of *E. coli* O157:H7 outbreak)
    - Cohort study to investigate the association of cell phone use and the development of brain tumors
    - Survey of individuals who took Cox-2 inhibitors

Biostatistics

- **Statistics** is the science and art of dealing with variation of data in order to obtain reliable results and conclusions
- **Biostatistics** is the application of statistics to problems in the biological sciences, health, and medicine
  - Examples
    - Computing age-adjusted cancer incidence rates to determine trends over time and locality
    - Calculating statistical measures of the risk of developing brain tumors following cell phone use after adjusting for possible confounding variables
    - Quantifying the relationship between use of Cox-2 inhibitors and quality of life

Role of Quantitative Methods in Public Health

1. Address a public health question
   - Generate a **hypothesis**
     - Based on scientific rationale
     - Based on observations or anecdotal evidence (not scientifically tested)
     - Based on results of prior studies
   - Examples of a hypothesis
     - The risk of developing lung cancer remains constant in the last five years
     - The use of a cell phone is associated with developing brain tumor
     - Vioxx increases the risk of heart disease

2. Conduct a study

- **Survey study** is used to estimate the extent of the disease in the population
- **Surveillance study** is designed to monitor or detect specific diseases
- **Observational studies** investigate association between an exposure and a disease outcome
  - They rely on "natural" allocation of individuals to exposed or non-exposed groups
- **Experimental studies** also investigate the association between an exposure, often therapeutic treatment, and disease outcome
  - Individuals are "intentionally" placed into the treatment groups by the investigators

3. Collect data

- Numerical facts, measurements, or observations obtained from an investigation to answer a question
- Influences of temporal and seasonal trends on the reliability and accuracy of data
- Examples
  - The number of lung cancer cases from 1960–2000 in the United States
  - The number of deaths from cardiovascular diseases in Whites and African Americans from 2000–2004
  - The number of people with heart attacks among individuals having used Vioxx before 2004

4. Describe the observations/data

- **Descriptive statistical methods** provide an exploratory assessment of the data from a study
  - Exploratory data analysis techniques
  - Organization and summarization of data
    - Tables
    - Graphs
    - Summary measures

5. Assess the strength of evidence for/against a hypothesis; evaluate the data

- **Inferential statistical methods** provide a confirmatory data analysis
  - Generalize conclusions from data from part of a group (sample) to the whole group (population)
  - Assess the strength of the evidence
  - Make comparisons
  - Make predictions
  - Ask more questions; suggest future research

Role of Quantitative Methods in Public Health

6. Recommend interventions or preventive programs

- The study results will prove or disprove the hypothesis, or sometimes fall into a grey area of "unsure"
- The study results appear in a peer-review publication and/or are disseminated to the public by other means
- As a consequence, the policy or action can range from developing specific regulatory programs to general personal behavioral changes

Six Examples

- Examples of how quantitative methods are useful in addressing public health problems

Probability of Breast Cancer?

- The National Cancer Institute estimates that women have an average lifetime risk of 13.2 percent (often expressed as "1 in 8") of being diagnosed with breast cancer at some time in their lives
- The chance that a woman will never develop breast cancer is 86.8 percent (expressed as "7 in 8")
- What is a probability?
- Suppose there are two women working in the same office
  - What is the chance (probability) that both women develop breast cancer over their lifetimes?

Source:
http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_key_statistics_for_breast_cancer_5.asp?sitearea=

Relative Risk of Breast Cancer?

- Having **one** first-degree relative (mother, sister, or daughter) with breast cancer approximately doubles a woman's risk of developing breast cancer (as compared to women having no first-degree relatives with breast cancer)
- Having **two** first-degree relatives increases her risk fivefold
- What is the probability or "risk"?
- What is the "relative risk"?

Source:
http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_key_statistics_for_breast_cancer_5.asp?sitearea

Comparison of Mortality Rates?

- Johns Hopkins University is in the state of Maryland
  - The death rate in Maryland was about 805/100,000 in 2002
- Disney World is in the state of Florida
  - The death rate in Florida was about 1,004/100,000 in 2002
- Does the higher death rate in Florida mean that Florida is a "riskier" place to live than Maryland?
- What other factors must be considered?

Source: http://www.cdc.gov/nchs/data/nvsr/nvsr53/nvsr53_05.pdf

Chances and Family History

- Does chance of having cancer depend on family history?
  - You and your best friend just turn age 50
  - During a routine visit to your doctor, you both had a PSA blood test and both scores resulted in the same value of 5 ng/ml
  - You have no family members with a history of prostate cancer—but your best friend has a family history
- Given the moderate PSA level, is the probability of having prostate cancer the same for the two of you?

Findings from a Clinical Trial of Vioxx Treatment?

- The drug Vioxx is effective in reducing pain
- On September 30, 2004, Merck announced a worldwide withdrawal of Vioxx from the market
- A 12-week clinical trial study of Vioxx compared to Naproxene in 5,500 patients was first published in 2003
  - It reported that "five patients taking Vioxx had suffered heart attacks during the trial, compared with one taking naproxene, a difference that did not reach statistical significance"

Findings from a Clinical Trial of Vioxx Treatment?

- Subsequent follow-up by the FDA found that:
  - "Eight people taking Vioxx suffered heart attacks compared with one taking naproxene … The difference was statistically significant …"
- Can the addition of three deaths have such a great impact on the finding?
- What is the meaning of **statistical** versus **clinical significance**?

Role of Quantitative Methods in Public Health

- Address a public health question
- Conduct a study
- Collect data
- Describe the observations/data
- Assess strength of evidence for/against a hypothesis; evaluate the data
- Recommend interventions or preventive programs

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Section B

Descriptive vs. Inferential Statistical Methods

Hypothesis and Study Design

- **Hypothesis:** seat belts save lives
- **Study design:** cross-sectional study of fatality outcome and seat-belt use of victims of motor vehicle accidents during a one-month time period in a large city

Description: Effect of Seat Belt Use on Accident Fatality

	Seat Belt	
Driver	Worn	Not Worn
Dead	10	20
Alive	40	30
Total	50	50
Fatality rate	10/50 (20%)	20/50 (40%)

Your Conclusion?

- What is your conclusion?
  - The fatality rate is:
    - **40%** in the group of drivers who **did not** wear seat belts
    - **20%** in drivers who **did** wear seat belts
  - Seat belts appear to save lives

The Inferential Questions of Interest

- The inferential questions of interest are:
  - Are results applicable to the population of all drivers? (generalization)
  - Does wearing seat belts save lives? (assess strength of evidence)
- Is the fatality rate of those not wearing seat belts higher than the fatality rate of those wearing seat belts? (comparison)
- How many lives can be saved by wearing seat belts? (prediction)
- Do other variables influence the conclusion?
  - For example: the age of driver, alcohol use, type of car, speed at impact (ask more questions)

Speed at Impact

	Speed at Impact			
	<= 30 Miles per Hour		> 30 Miles per Hour	
Driver	Seat Belt Worn	Seat Belt Not Worn	Seat Belt Worn	Seat Belt Not Worn
Dead	3	2	7	18
Alive	27	18	13	12
Total	30	20	20	30
Fatality rate	10%	10%	35%	60%

How Does This Influence Your Conclusion?

- How does this influence your conclusion?
  - The fatality rate is 10% at low-impact speeds regardless of seat-belt use
  - The fatality rate at high impact speeds is:
    - **60%** in drivers **not wearing** seat belts
    - **35%** in drivers **wearing** seat belts