Descriptive studies

*Objectives:

- ✓ Identify the types and concepts of different study designs.

- ✓ Recognize the strength and limitation of several study designs.

- ✓ Describe and contrast advantages and disadvantages of each study design

- ✓ Identify applications in clinical and in non-clinical settings.

- ✓ Identify common sources of bias for each design

- ✓ Identify and characterize basic designs in non-experimental epidemiologic research, including cohort and case-control studies, and variants of these designs such as nested case-control studies, case-cohort studies.

- ✓ Distinguish between individual-level measures and population-level measures in epidemiologic research.

*Types of epidemiological Studies:

A. Experimental

- Study factor is manipulated by the investigator

B. Observational –

- No manipulation of study factor by the investigator

- ■ Descriptive
- ■ Analytic

*Types/designs of descriptive studies:

- **Populations (groups):**

-ecologic (aggregate) studies

- **Individuals**

- Case reports
- Case series
- Cross sectional surveys

Table 3.1. Types of epidemiological study

Type of study	Alternative name	Unit of study
Observational studies		
Descriptive studies		
Analytical studies		
Ecological	Correlational	Populations
Cross-sectional	Prevalence	Individuals
Case-control	Case-reference	Individuals
Cohort	Follow-up	Individuals
Experimental studies	*Intervention studies*	
Randomized controlled trials	Clinical trials	Individuals
Cluster randomized controlled trials		Groups
Field trials		
Community trials	Community intervention studies	Healthy people Communities

Case Report (case studies):

- ✓ The case report is the basic type of descriptive studies.

- ✓ It describes the experience of one case.

- ✓ It represents the first clue in the identification of a new disease or adverse effect of exposure.

- ✓ Describe the experience of a single case or a group of patients with similar diagnosis

- ✓ Document unusual medical occurrences.

- ✓ Generally report a new or unique finding

  E.g. previously un-described disease

  E.g. unexpected link between diseases

  E.g. unexpected new therapeutic effect

E.g. adverse events

*The objectives of case report are:

- ✓ To describe clinical observation
- ✓ To describe the variation of disease
- ✓ To describe results of new treatment

*Example of a case report

1. In 1961, pulmonary embolism has been described in 40-years, pre-menopausal female after 5 years of using oral contraceptive pills. This phenomenon was considered interesting as pulmonary embolism usually occurs among older subjects, post menopausal.

2. **Formulation of a new hypothesis**

- ■ Adeno-carcinoma of the vagina was reported in a young girl.
- ■ This tumor is not only rare, but also the usual victim is over 50 years of age.
- ■ Documents showed that this girl was exposed to estrogen during fetal life.
- ■ This unusual event and exposure leads the investigator to formulate a hypothesis that:
- ■ Foetal exposure to estrogen may be responsible for the occurrence of this tumor.

*Limitations:

- ❖ Not a true epidemiologic design
- ❖ No appropriate comparison group
- ❖ Can't be used to test for presence of valid statistical association, since it's based on the experience of one person **(presence of any risk factors, maybe purely coincidental).**

*2- Case series report:

- It describes the experience of a group of similar cases having an unusual disease or event.
- Case series describe characteristics of a number of patients with a given disease.
- The patients are identified during *routine surveillance*.
- It can also generate a hypothesis.

Example of case series

- In 1980, Pneumocystitis Carinii has been described among 5 young adult homosexual men.

This disease was known to affect only immune-compromised subjects.

- Similarly, Kaposi sarcoma, a disease of old age has been reported among young adult homosexual men.

■ **Advantages**

¤We can aggregate cases from disparate sources to generate hypotheses and describe new syndrome .Example **(hepatitis, AIDS)**

¤ Used as an early means to identify the beginning or presence of an epidemic.

■ **Limitations**

❖ We cannot test for statistical association because there is no relevant comparison group

❖ Based on individual exposure {may simply be coincidental

❖ Not a true epidemiologic design

3-Correlation study (Ecological study):

✓ The correlation study is referred to as ecological study because the analysis is at the level of an entire population rather than at the individual level.

- ✓ The correlation studies measures the association between two *quantitative variables.*
- ✓ It uses data from entire populations to compare disease frequencies in relation to supposed harmful (or beneficial) exposures during the same period of time or at different points in time.
- ✓ They are quick and inexpensive as they use already available data like vital statistics, censuses and national health surveys.
- ✓ It is useful for generating hypotheses

ɸ **Aims of ecologic study**
- To generate etiologic hypothesis
- To evaluate effectiveness of population interventions.

ɸ **Limitations of correlation studies:**
a. They cannot be used for testing the hypothesis i.e. never prove causation.

b. It is impossible to link exposure and disease in a particular individual because data are for groups, Data represent average exposures rather than individual exposures, so we cannot determine a dose-response relationship

c. Lack of the ability to control for the effects of the **confounding factors**. These are factors other than the studied factor (exposure) that disturb the relation between the studied exposure and disease. **However,** the possible association between meat consumption and cancer colon may be due to other factors such as decrease intake of vegetables, high socioeconomic condition or increase intake of fat.

d. Caution must be taken to avoid drawing **inappropriate conclusions, or ecological fallacy**

e. Cannot establish an association between exposure and outcome (can only suggest)

f. Uses measures that represent characteristics of entire populations

g. It describes outcomes in relation to age, time, utilization of services, or exposure

h. Can't link exposure –disease relationship at the individual level

i. Inability to control for confounding factors

j. Uses average exposure levels rather than actual levels of exposure

■ **ADVANTAGES**

    a. We can generate hypotheses for case-control studies and environmental studies

    b. We can target high-risk populations, time-periods, or geographic regions for future studies

    c. Cheap, quick and simple(generally make use of secondary data)

4- Cross sectional study (prevalence studies):

    ✓ An "observational" design that measures existing disease (D) and current exposure levels (E) at a single point in time (a cross-section of the population)

    ✓ In the cross-sectional study the exposure (E) and the disease (D) status are assessed simultaneously (at the same time) among individuals in a well-defined population.

    ✓ It shows the association between exposure and disease

    ✓ Each subject is assessed once at point in time

    ✓ Useful for investigating exposures that are fixed characteristics of individuals, such as ethnicity or blood group.

    ✓ It is quick and cheap

    ✓ "Snapshot Studies"

    ✓ Used to study conditions that are relatively frequent with long duration (chronic conditions)

- ✓ It's include surveys
- ✓ Can conduct repeated cross-sectional studies to measure changes in a population.
- ✓ Can measure attitudes, beliefs, behaviors, personal or family history, genetic factors, existing or past health conditions, or anything else that does not require follow-up to assess.
- ✓ Cross-sectional studies are the source of most of what we know about the population other than vital statistics
- ✓ Cross-sectional studies are carried out for public health planning and for etiologic research.

Examples of cross-sectional studies:

- Assess the presence of IHD in relation to physical exercises.
- Assess the presence of obesity in relation to diabetes mellitus.
- In 2007-2008, almost one in five children older than 5 years was obese. (*Health, United States, 2010*; data from the National Health and Nutrition Examination Survey)
- 35% (~7.4 million) of births to U.S. women during the preceding 5 years were mistimed or unwanted (2002 National Survey of Family Growth, Series 23, No. 25, Table 21

Table 5-3. Cross-sectional survey of coronary heart disease (CHD) among white male farm owners, aged 40–74 years, by occupational physical activity

	Number examined	Number with CHD	Prevalence rate	Age-adjusted prevalence rate
Not physically active	89	14	157.2/1000	126/1000
Physically active	90	3	33.3/1000	36/1000
Total	179	17	95.0/1000	87/1000

Source: J. R. McDonough et al., Coronary heart disease among Negroes and Whites in Evans County, Georgia. *J. Chronic Dis.* 18:443, 1965.

- **Can be used as a type of analytic study for testing hypothesis, when;**
  1) Current values of exposure variables are unalterable over time
  2) Represents value present at initiation of disease
     1. E.g. eye color or blood group
  3) If risk factor is subject to alterations by disease, only hypothesis formulation can be don

■ **Limitations**:
  a. Cannot determine whether exposure preceded disease (It faces the chicken egg dilemma).
  b. Data deals with survivals. Those who died or cured are not included
  c. Not suitable for studying highly fatal diseases or a disease with short duration of expression
  d. It is not suitable for rare diseases.
  e. It considers prevalent rather than incident cases, results will be influenced by survival factors

■ **Advantages**
  ■ They cut across the general population, not simply those seeking medical care
  ■ Good for identifying prevalence of common outcomes, such as arthritis, blood pressure or allergies
  ■ Provides information for planning and evaluation of health services
  ■ Formulate Hypothesis.

*Survey: is an investigation that uses a "structured and systematic gathering of information" from a sample of "a population of interest to describe the population in quantitative terms

- Designed to be **representative** of the entire country

- Surveys can be carried out by one or multiple modes, particularly household interviews, telephone, or mail.
- The subjects of a survey can be members of the general public, patients, health-care providers, or organizations

National surveys in USA:

National Surveys

- A large set of national epidemiological surveys are conducted by the National Center for Health Statistics (NCHS) --- the Federal Government's principal vital and health statistics agency.

- The NCHS is part of the Center for Disease Control and Prevention (CDC) — under the U.S. Dept. of Health and Human Services.

National Surveys

- National Health Interview Survey (NHIS) Principal source of information on the health of the civilian, non-institutionalized household population of the U.S.

- National Health and Nutrition Examination Survey (NHANES) Collects information about the health and diet of people in the U.S.

National Surveys

- **National Maternal and Infant Health Survey (NMIHS)**
  Collects data on study factors related to low birthweight, stillbirth, infant illness, and infant death.

- **Behavioral Risk Factor Surveillance System**
  Collects data on prevalence of personal health behaviors among adults associated with premature morbidity and mortality.

National Surveys

- **National Health Care Survey (NHCS)**
  Collects a wide range of data on health care use, impact of medical technology, and quality of care provided.

- **National Immunization Survey (NIS)**
  Collects information on the immunization coverage and health care of children across the U.S.

National Surveys

- **National Survey of Family Growth (NSFG)**
  Collects data on factors affecting pregnancy and women's health in the U.S.

- **National Mortality Followback Survey (NMFS)**
  Aggregate sample of death certificates across the U.S. for targeted research purposes.

Other Sources of National Data

- **National Vital Statistics System**
  - --- Birth data
  - --- Mortality data
  - --- Fetal death data
  - --- Marriages and divorces
  - --- Linked births/infant deaths
  - --- National death index

- **U.S. Census Bureau**

Example1: calculate the Prevalence of CHD among men over 60 years

Smoking	CHD		Total
	Yes	No	
Yes	10	30	40
No	8	152	160
total	18	182	200

Example2: prevalence of anemia among school children

Sex	Anemic	Non-anemic	Total
Male	20	300	320
Female	60	120	180
Total	80	420	500