Definition of Epidemiology

- **Epidemiology**
  - The study of the distribution and determinants of disease frequency in human populations (Hennekens & Buring, 1987 (H & B, 1987)).
  - Concerned with the distribution and determinants of health and diseases, morbidity, injuries, disability, and mortality in populations (Fris & Sellers, 1996).
  - … and the application of this study to control of health problems. (Last 1995)
  - “Epidemiology involves both science and public health practice” (CDC. Principles of Epidemiology 2nd Ed).

Study Design in Epidemiology

- Basic design strategies used in epidemiologic research can be categorized according to whether the investigator focus on:
  - Describing the distributions of disease
  - Elucidating its determinants
Study Design in Epidemiology (cont.)

Descriptive epidemiology

- It is concerned with the distribution of disease, including consideration of what populations do or do not develop a disease, in what geographic locations it is most or least common, and how the frequency of occurrence varies over time.
- Information on each of these characteristics can provide clues that help to formulate an epidemiologic hypothesis consistent with existing knowledge of disease occurrence.

Analytic Epidemiology

- Focuses on the determinants of a disease by testing the hypotheses formulated from descriptive studies.
- The ultimate goal is judging whether a particular exposure causes or prevents disease.

Each descriptive and analytic study design has its unique strengths and limitations.

Study Design in Epidemiology (cont.)

- Descriptive studies
  - Populations (correlational studies)
  - Individuals
    - Case reports
    - Case series
    - Cross sectional surveys
- Analytic studies
  - Observational studies
    - Case-control studies
  - Cohort studies - retrospective and prospective
  - Intervention studies (clinical trials)
Descriptive studies (cont.)

- They are concerned with describing the general characteristics of the distribution of a disease, particularly in relation to **person, place and time**.

- **Person**: age, sex, race, marital status, occupation, life-style (food, drugs, exercise)

- **Place**: variations among countries or within countries (urban vs. rural)

- **Time**: seasonal patterns, frequency today compared to the past.

**Correlational Studies (populations)**

- Uses data from entire populations to compare disease frequencies between different groups during:
  - **the same period of time**
  - **in the same population at different points in time**

(i.e. Positive correlation between per capita daily consumption of fluid and rates of colon cancer for different countries have suggested that various dietary components may be risk factors for colon cancer (Armstrong & Doll, 1975 in H & B, 1987))

(i.e. Comparison between cardiovascular disease mortality rates in the US in two periods. The decline observed in 1977 related to the occurrence in 1968 raised numerous hypotheses: decline in deaths was due to prevention, decline in deaths was due to better treatment (NIH, 1979 in H & B, 1987)).
Correlational studies are useful for the formulation of hypotheses, but they cannot be used to test them because of a number of limitations inherent in their design.

- Since correlational studies refer to whole populations rather than to individuals, it is not possible to link an exposure to occurrence of disease in the same person.

Descriptive studies (cont.)

Meat consumption example: not possible to know if those who develop cancer were the ones with the highest meat intakes, only that on average populations with the highest per capita consumption of meat also have the highest rates of the disease. Pops. also differ on other colon cancer risk factors: diet high in saturated fat or low in fiber. Thus, difference between countries may be due not to varying levels of meat consumption, but rather to the independent effects of other risk factors. These studies are Hypothesis generators.

Descriptive studies (cont.)

- Case Reports and Case Series (individuals)
  - Case Report. It is the most basic type of descriptive study of individuals, consisting of a careful, detailed report by one or more clinicians of the profile of a single patient.
  - Case Series. It describes characteristics of a number of patients with a given disease.
Descriptive studies (cont.)

- **Case Report.** Example: In 1961 a case report was published of a 40-year-old premenopausal woman who developed a pulmonary embolism 5 weeks after beginning to use an oral contraceptive preparation to treat endometriosis. Since this problem is more common in older postmenopausal women, the investigator postulated that the drug may have been responsible. The crucial question was whether women who developed this problem were more likely to have used contraceptives than those who did not develop the disease. An adequate sample, a comparison group and considering other factors was necessary. (Jordan, 1961 in H & B, 1987)

Descriptive studies (cont.)

- **Case Series.** Routine surveillance programs often use accumulating case reports to suggest the emergence of new diseases or epidemics. Example: In 1981 a cluster of 5 young, previously healthy homosexual men were diagnosed as having *Pneumocystis carinii* pneumonia at 3 hospitals in L.A. during a 6-month period. This disease had been seen in old people with suppressed immune system. This unusual circumstance suggested that they were suffering from an Acquired Immunodeficiency Syndrome (AIDS) and raised the hypothesis that some aspect of sexual behavior could be related to risk of the disease. A study that evaluated if the risk of disease is different among individuals exposed or not exposed to a factor was necessary. (Langmuir, 1983 in H & B, 1987)

Descriptive studies (cont.)

**Cross-Sectional Surveys.**

The status of an individual with respect to the presence or absence of both exposure and disease is assessed at the same point in time.
Descriptive studies (cont.)

Cross-Sectional Surveys. Example:

Health Interview Survey (HIS) is a national cross-sectional study that periodically collects extensive information by questionnaire from a representative sample throughout the US. Then frequency of various diseases and other health outcomes are calculated and examined in relation to several risk factors (age, sex, race, SES, smoking, etc.). These data is basic to public health administrators in assessing health status and needs. Since exposure and disease are assessed at the same point in time these surveys cannot always distinguish whether the exposure preceded the development of the disease or whether the presence of the disease affected the individual’s level of exposure. These studies are useful for raising the question of the presence of an association rather than for testing a hypothesis.

Comparison in studies

All study designs involve some implicit (descriptive) or explicit (analytic) type of comparison of exposure and disease status.

In a case report a clinician observes a feature of a single case, compares with what he/she has seen on other cases (implicit comparison with the “expected”), and formulates a hypothesis.

Analytic studies (cont.)

In analytic study designs the comparison is explicit, since the investigator assembles groups of individuals for the specific purpose of systematically determining whether or not the risk of the disease is different for individuals exposed or not exposed to a factor of interest.

It is the use of an appropriate comparison group that allows testing of epidemiologic hypotheses in analytic study designs.
Analytic studies (cont.)

- **Analytic studies** can be divided into two broad categories and the major difference lies in the role played by the investigator.
  - **Observational studies** (case-control & cohort studies).
  - **Intervention studies** (clinical trials).

Observational studies.
Case-control studies.

- A case group or series of patients who have a disease of interest and a control, comparison, group of individuals without the disease are selected for investigation, and the proportions with the exposure of interest in each group are compared.
Observational studies.
Case-control studies (cont).

Case-control studies.
- This design offers a solution to the difficulties of studying diseases with very long latency periods, since investigators could identify affected and unaffected individuals and then look backward in time to assess their antecedent exposures rather than having to wait a number of years for the disease to develop.

Observational studies.
Cohort studies.

Cohort studies.
Subjects are classified on the basis of the presence or absence of exposure to a particular factor and then followed for a specified period of time to determine the development of disease in each exposure group.

The feature that distinguishes a prospective from a retrospective cohort is simply and solely whether the outcome of interest has occurred at the time the investigators initiates the study.

Observational studies.
Cohort studies.

Cohort studies.
- Prospective. At the beginning of the study, the groups of exposed and unexposed subjects have been assembled, but the disease has not yet occurred, so that the investigator must conduct follow-up during an appropriate interval to ascertain the outcome of interest.
Observational studies. 
Cohort studies.

**Cohort studies.**

- **Retrospective.** The investigation is initiated at a point in time after both the exposure and disease have already occurred, but the investigator goes back to the records (i.e., factory records) when the exposure started and then ‘follow’ the information from there to the occurrence of the disease.

---

**Timing of case control, prospective and retrospective cohort studies in relation to exposure and outcome.** (Hennekens & Buring, 1987)

- **Case control study:**
  - Exposure
  - Disease

- **Prospective cohort study:**
  - Exposure
  - Disease

- **Retrospective cohort design:**
  - Exposure
  - Disease

= Present
= Absent
= to be determined

---

**Strengths of the case-control study design.**

(Hennekens and Buring, 1987.)

- **Strengths**
  - Relatively quick and inexpensive
  - Well-suited to evaluate diseases with long latent periods
  - Optimal for the evaluation of rare diseases
  - Can examine multiple etiologic factors for a single disease
Limitations of the *case-control study* design.  
(Hennekens and Buring, 1987.)

- Limitations.
  - Inefficient for the evaluation of rare exposures
  - Cannot compute incidence rates of disease in exposed and non-exposed, unless study is population based
  - Temporal relationship between exposure and disease may be difficult to establish
  - Prone to bias, in particular selection bias and recall bias.

Strengths of the *cohort study* design.  
(Hennekens and Buring, 1987.)

- Strengths
  - Is the particular value when the exposure is rare.
  - Can examine multiple effects of a single exposure
  - Can elucidate temporal relationship bet. exposure & disease.
  - If prospective, minimizes bias in the ascertainment of exposure
  - Allows direct measurement of incidence of disease in the exposed and non-exposed groups

Limitations of the *cohort study* design.  
(Hennekens and Buring, 1987.)

- Limitations.
  - Inefficient in the evaluation of rare diseases
  - If prospective, can be extremely expensive and time consuming
  - If retrospective, requires the availability of adequate records
  - Validity of results can be seriously affected by losses to follow-up.
Intervention studies

*Intervention studies* (clinical trials).

Also referred to as experimental studies may be viewed as a type of prospective cohort study, because participants are identified on the basis of their exposure status and followed to determine whether they develop the disease.

The distinguishing feature of the intervention design is that the investigator allocate the exposure and then follow the subjects for the subsequent development of disease.

**Conclusion**

- While distinctions are not clear-cut, in general *descriptive studies* are useful primary for describing patterns of disease occurrence and for allowing the formulation of etiologic hypotheses.

- Similarly, while data from *analytic studies* can certainly be used to generate additional research questions, their chief contributions is to test epidemiologic hypotheses.