

EPIDEMIOLOGICAL STUDIES

((Design strategies in epidemiological research))

Introduction :-

Research can be defined as (the systemic search for information and new knowledge to solve a problem) .

The daily practice of medicine need the knowledge about who is likely to develop a particular condition and under what circumstances , what is the best treatment , what is the prognosis , and who we can prevent such illness .

Health care providers must be able to identify persons who are at high risk to reduce that risk . This type of knowledge emerges in many cases from epidemiological researches .

Epidemiology is the basic scientific tool which public health professional use to health status , diseases , medical interventions and the effectiveness of healthy delivery services . It is a perspective in which systematic observation is the basis on which grounded inferences can be made about observed –phenomena .

Aims of medical (epidemiological) research .

- a-monitoring and surveillance of health and disease .**
- b-Establishing causes of disease or factors associated with death or disease .**
- c-Detecting diseases .**
- d-preventing death or disease .**
- e-Evaluating treatments for disease .**
- f- Evaluating health services and programs .**

Learning objectives .

- a-To under stand the concepts of different study designs .**
- b-To learn about the advantages and disadvantages of each study design .**
- c-To appropriately use a study design in research projects .**

It is important to :-

- a-Determine the exposure and outcome precisely .**
- b-Remember that the choice of any study design based on :**
 - 1-Features of exposure and outcome .**
 - 2-Time and resources available .**
 - 3-Results from previous studies .**
 - 4-Gaps in the knowledge that remain to be filled .**
- c-Take care about sampling method . (proper sampling procedure) .**

((NO STUDY DESIGN IS FIT FOR ALL TYPES OF STUDIES

))

2x2 table to asses the relationship between an exposure and outcome .

OUTCOME

| | +ve | -ve | |
|------------------------|------------|------------|--------------------------------|
| Exposed +ve | a | b | a+b |
| Non-exposed -ve | c | d | c+d |
| total | a+c | b+d | Pop.number (n)(a+b+c+d) |

The basic design strategies used in epidemiological research can be broadly categorized into :-

1- Descriptive studies . (describe occurrence of outcome) .

a- popular (correlation studies) .

b-individual :- divided .

1-CASE-SERIES .

2-CASE-REPORT .

3- CROSS-SECTIONAL .

2-Analytic studies :-descriptive the association between the exposure and outcome , divided into .

a- observational studies .

1- CASE-CONTROL STUDY .

2-COHORT STUDY .

b-Interventional study . (CLINICAL TRIAL) .

Descriptive studies .

a-Popular studies , (correlation studies) .

b-Individual studies , case-report , case-series , and cross-sectional .

so the descriptive studies , describe the pattern of disease occurrence in relation to variables such as person , place , time .

DESCRIPTIVE EPIDEMIOLOGY

Three groups of variables are commonly used in descriptive epidemiology. These are:

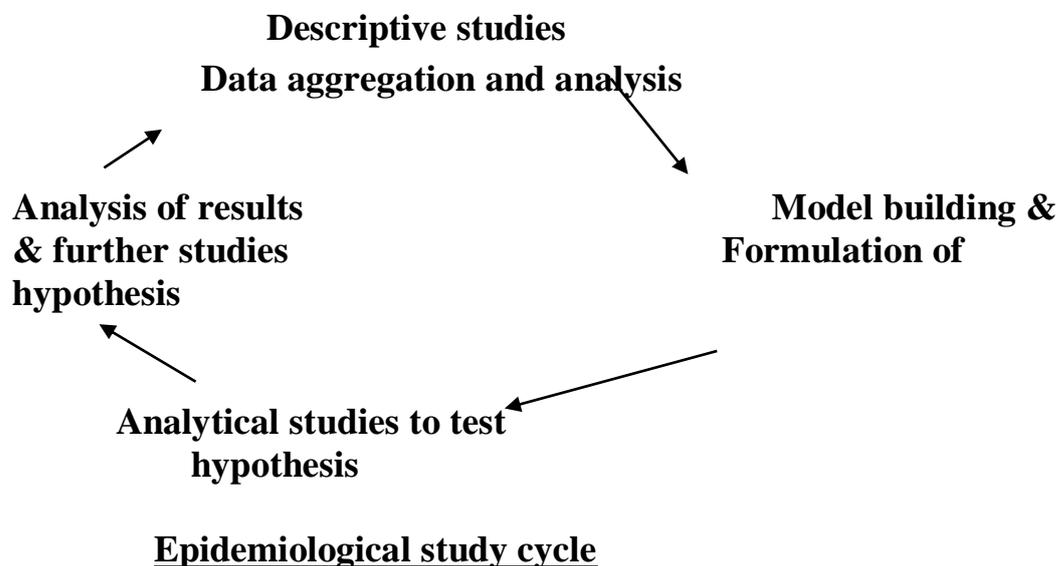
A. Characteristics describing the persons affected such as age, sex, marital status, education, occupation, habits, genetics and ethnic groups.

B. Characteristics describing the place where persons were found affected. The distribution of the disease may have international, national (limited to one country, continental, local: only part of a country or urban-rural pattern?

C. Characteristics describing the time in which persons were found affected. Does the distribution follows secular trend (over many years and decades), seasonal trend (within the same year), recurrent

pattern or the occurrence of disease after special events, e.g., raining.
?

It is evident that variables related to person, place and time are not necessarily those which have causal association with the disease being examined. They are simply the variables that are measured. An association of a particular disease with a given place might be due to the type of people inhabiting that place. Also, an association of disease with a characteristic of person might be due to environmental factors peculiar to that place of residence rather than the person characteristic itself. The descriptive associations are helpful in hypothesis formulation. Testing of hypothesis needs other methods (analytical methods).



Examples

Distribution of disease with age

It is well known that the distribution of disease is very variable with age. Actually, age is an important confounding variable and must be considered and controlled for (there are many ways to control for the effect of age) when the distribution of disease is examined in relation to other variables. The variation of disease distribution with age may be explained as follows:

1. Accuracy of diagnosis. Disease is less likely to be ascertained in extreme age groups. This leads to under estimation of certain causes of death in the very young and the very elderly people. In some instances, basic population data are lacking on such extreme age groups.
2. Variation in intensity and duration of exposure to risk factors.

3. **Variation in immunity and susceptibility.** People are not constant in their immunological response and immunological status and are not necessarily similar to each other in that respect.
4. **The type of epidemiological parameter used, i.e., incidence.**
Prevalence or mortality. For example, in a disease with constant incidence with age and negligible mortality and incomplete recovery, the age specific prevalence increases with age because cases once developed tend to accumulate over time thus raising the prevalence.
5. **Bimodality.** In some instances the distribution of disease frequency with age may have more than one peak (bimodal) as in case of the incidence of lymphomas leukemia's, testicular cancer and tuberculosis with age. This bimodality may suggest the heterogeneity of data and the possibility that we are dealing with two disease entities rather than with one disease. For example, the first peak in the incidence rate of tuberculosis in young children is definitely primary(exogenous) tuberculosis. On the other hand, the peak late in life is mainly secondary (endogenous) tuberculosis.
6. **Cohort effect.** Some times a high frequency of disease in a given age group might reflect an old exposure rather than a current exposure to a etiological factors.
6. **Ageing or biological clock.** Some times, people become very aging and lose the ability to carry out even simple tasks, yet they have no apparent disease. They probably follow a preceded biological clock, which determines the life span.

Distribution of disease with marital status

In many studies it was reported that death rates and suicidal rates are higher among non-married people (single, widowed and divorced) than they are in married people. This is true for both males and females. Such variation might be difficult to explain but two explanations are possible:

- a. **Marriage stabilizes life and reduces the risk of exposure to hazardous behavior.** Married people may feel more responsible not only for their lives but for the care and life of their spouses and children. They may avoid certain risky behaviors.
- b. **People who are unmarried are actually not healthy to start with and they prefer not to marry.** The higher risk of death and suicide among them is perhaps related to their poor health to start with rather than marital status (being unmarried). But this is not completely true – at least not in some countries such as Germany, where it was shown, that widowed wives live considerably longer than married wives while the opposite effect was observed for

males.

Interpretation of association of disease distribution with place

The following criteria are essential to demonstrate an association of disease distribution with place:

- 1. High frequency rates of the disease are observed in all ethnic groups living in that place.**
- 2. Similar people who inhabit other places do not show high frequency rates of the disease.**
- 3. Healthy people entering the place become affected by the disease at a rate similar to that of the indigenous population.**
- 4. People who leave the place and move to other places do not experience high frequency rates of the disease.**
- 5. Species other than man may show similar pattern of the disease.**

When these criteria are fulfilled, it might be that the physical, chemical, biological or social environment of that place contain etiological factors of the disease. More intensive research is required to identify such factors.

Interpretation of disease variation with time

Secular changes

Distribution of disease with time may follow long term changes (secular changes or trends). The changes occur over years or decades. Examples are the changes in cancer, cardiovascular disease and AIDS. Such secular changes which show a clear rise in the disease frequency with time (years or even decades) as has been shown with the rise in mortality rate due to lung cancer in some European countries during the twentieth century could be explained as follows:

- 1. The rise indicates real increase in the incidence of the disease in response to:**
 - a. Massive exposure to disease agents.**
 - b. Change in life style of the people and**
 - c. Failure of adaptation to social change.**
- 2. The rise is artificial due to**
 - a. Improved diagnosis of disease which lead the identification of cases which were previously missed.**
 - b. Change in classification of disease.**
 - c. Improved recording of cases.**
 - d. Ageing of the population/ change in population at risk**

Seasonal changes

A change of disease frequency within the year reflects a change in population immunity (susceptibility), change in the environmental situation in favor of disease agent development or multiplication, and its transmission to new host or both.

Data provided by descriptive studies are essential for :

1-Public health administrators :

a-Which population or subgroup are most or least affected .

b- Efficient allocation of resources and preventive programs .

2-Epidemiologist : Identification of risk factors and possible etiology of the disease .

1-Correlation study . (Ecological study) .

Describe the disease in the entire population in relation to the factor of interest such as age , sex , calendar time , utilization of health services , consumption of certain food or medication --- etc .

E.G= To describe mortality from coronary heart disease (CHD) in 5 countries (population) with per capita cigarette consumption in one year . ----

The descriptive measure of association in correlation study is (correlation coefficient) (r) which ranges from (-1) to(+1) if r= -ve means inverse association (may be preventive) .

If r=+ve means positive association (may be causal) . If r= 0 means no association . -----

Advantages :

1-Quick and inexpensive .

2-Use already available data .

3-Usually used as a first step in investigation a possible exposure – outcome relation-ship .

Limitations :

1-In ability to determine the temporal relation –ship between exposure and outcome .

2-Lack the ability to control for the effect of confounder .(العوامل المربكه)

3-Represent average exposure level rather than actual individual level .

4-Formulate the hypothesis but can not test it .

THE PRESENCE OF CORRELATION DOSE NOT NECESSERLY IMPLY THE PRESENCE OF A VILD STATISTICAL ASSOCIATION .

2-CASE REPORT STUDY : (تقرير حاله واحده)

Describe the experience of a single patient . A condition develops in single individual and draws the attention of the clinician or researcher . It is the first step in disease recognition . EX, Kaposi sarcoma in healthy homosexual adult . (REFLECTING UNUSUAL FEATURES OF DISEASES) .

3- CASE SERIES STUDY :

Describe the experience of a group of patients with similar diagnosis . (COLLECTION OF INDIVIDUAL CASE REPORT)

Advantages (report and series study) .

- 1- Recognition of new disease EG, AIDS . (INCIDENCE) .
- 2-Formulation of hypothesis concerning possible risk factor .

Limitation :

- 1-Based on experience of one or few patients only .
- 2- Lack of comparison group . (لا توجد مجموعه للمقارنه)
- 3-Formulate the hypothesis but can not test it .

4- CROSS- SECTIONAL STUDY (PREVALENCE STUDY) OR (PREVALENCE SURVEY) .

A- SHORT SINGLE OBSERVATION CARRIED OUT IN ONE GROUP OF POPULATION (AT THE SAME TIME)

B-LONGITUDINAL SURVEY OF FOLLOW UP SURVEY .

In this study(A) the exposure and the outcome are assessed simultaneously or (at the same time) among individual in well defined population . It is as if we were taking a photo-graph or a slice through the population at a point in time .

EX, A researcher wishes to investigate a possible association between cigarette smoking and CHD in a certain population .

146 young adults were randomly selected from city x . smoking history was taken and an ECG performed for evidence of CHD .

From 146 total population , 15 have had CHD , 131 have not CHD , 86 have had smoking and 60 not smoking . (13) have had CHD and smoking .

1- Draw 2x2 table and calculate the prevalence of CHD .

2- Calculate the prevalence of disease CHD in person with exposure (smoking) . $(a/a+b)$

3-calculate the prevalence of disease in person without exposure (smoking) . $(c/c+d)$
OR , (Compare the prevalence of exposure in person with disease to the prevalence of exposure in person without the disease) . $(a/a+c)$, $(b/b+d)$.

| | CHD | NO CHD | TOTAL |
|------------|-----|--------|-------|
| Smoking | 13 | 73 | 86 |
| No smoking | 2 | 58 | 60 |

| | | | |
|--------------|-----------|------------|------------|
| Total | 15 | 131 | 146 |
|--------------|-----------|------------|------------|

Advantages :

- 1-Quick , inexpensive and less time consuming .**
- 2-Provide information about the frequency and characteristic of the disease .**
- 3-Provide information about the prevalence of the disease .**

limitation :

- 1-Not determine the temporal relation ship between exposure and outcome .**
- 2- Not determine prognostic factor from risk factor .**
- 3-Liable for information bias (recall or interviewer bias)**
- 4-Formulate the hypothesis but can not test it .**
- 5-selecrion survival .. due to the exclusion of deaths from the study , while it is possible that many deaths could be attributed to the health problem under study .**
- 6-chiecken-egg dilemma .anxiety –angina pectores , or anxies person develop the angina .**

From example above , the fact that those with CHD had a higher level of smoking than those without smoking dose not necessarily lead to the conclusion that smoking increases the risk of CHD . This may well be so , but it is only by demonstrating increased CHD in people with previously elevated smoking that a causal inference about the relationship may be drawn .

NOT :- Studies based upon(N) entire population or a sample of entire pop , these studies describe or identify health problems at a point in time or over a short period of time .

2. Epidemiological surveys (household surveys)

These surveys are specially designed and carried out for answering specific questions. They include different types of epidemiological studies but two main types are commonly used.

Cross-sectional surveys or studies.

A-These are based on a single observation usually carried out in a short time. They are characterized by the following:

- a. They usually measure prevalence of disease or related outcome.**
- b. They suggest hypotheses.**
- c. They are not useful for diseases of short duration. A single observation may miss cases.**
- d. Their results are difficult to interpret because of seasonal variation and cohort effect.**
- e. They are relatively quicker and cheaper to do.**

f. Although they are essentially descriptive, they can be modified to estimate incidence of disease and to test hypotheses. A case-control design can be made within the context of a cross-sectional study.

B-Longitudinal or follow up surveys or studies.

These are based on repeated observation of the study population over a defined period of time. They start with a base-line data provided by initial cross-sectional study.

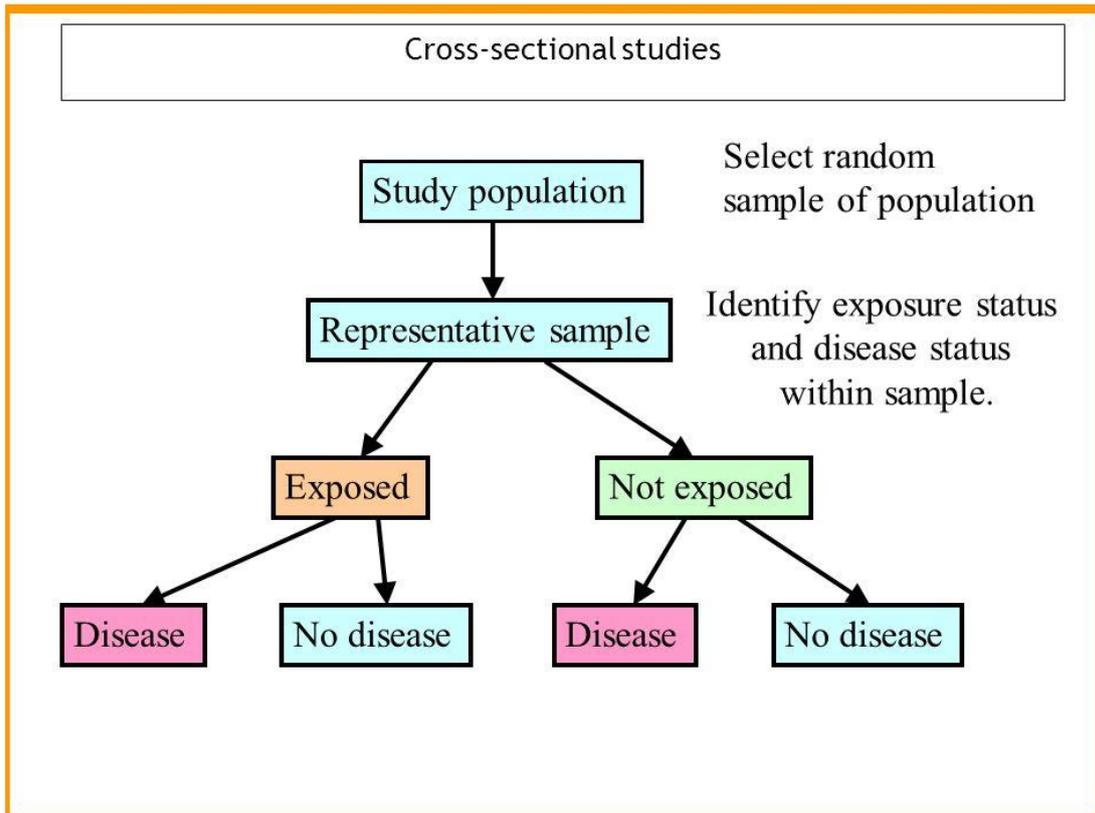
- a. They measure incidence of disease or related outcome.**
- b. They suggest hypotheses.**
- c. They are relatively more expensive and difficult to organize.**
- d. They are not useful for diseases of rare occurrence.**
- e. The results are easier to interpret.**
- f. They can be useful to determine seasonal variation of disease and other health related outcomes.**

Advantages of epidemiological (household) surveys

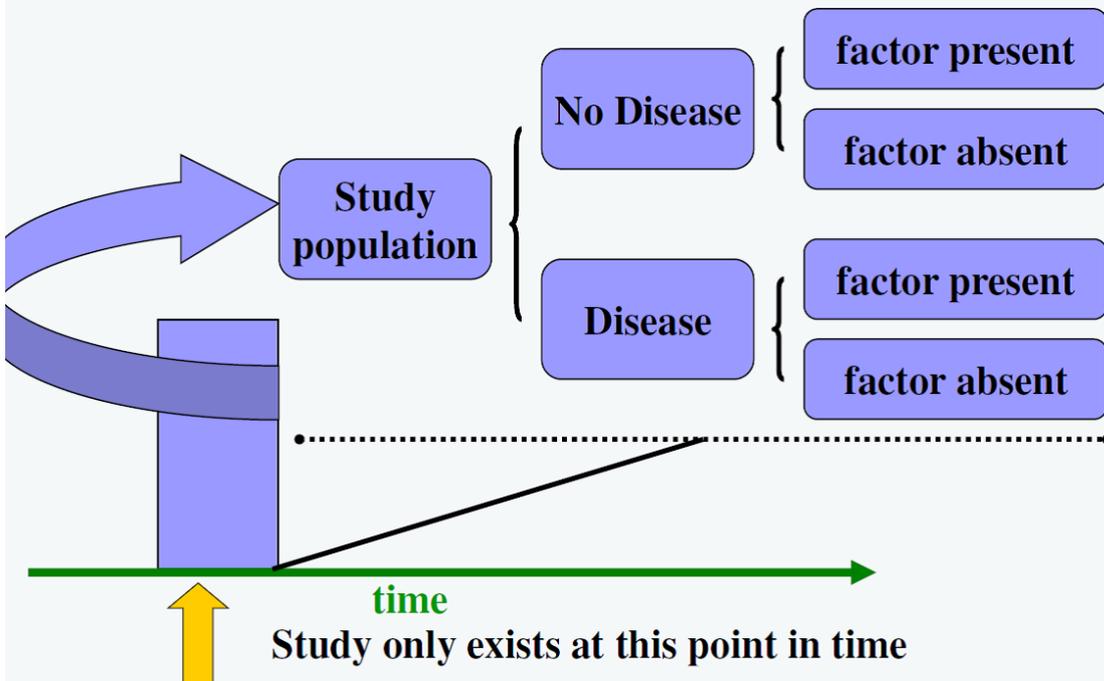
- a. They serve group diagnostic purposes.**
- b. They identify population selective action in response to illness.**
- c. They help in assessing total population health needs.**
- d. They are particularly useful in rural areas where other data sources may not be available.**
- e. They are useful in evaluating health care effectiveness and in planning health care for population.**
- f. They identify services non-users.**
- g. They help in identifying factors which determine utilization variation.**

Problems associated with household surveys

- a. Diagnostic accuracy. Diagnosis is based on population reporting which is technically questionable.**
- b. Memory errors. People tend to remember important events.**
- c. Over and under-reporting and non-response.**
- d. Social distance between interviewer and interviewee.**
- e. Sex of interviewer.**
- f. Self-reporting and reporting for others.**
- g. Cost and time required to achieve the work.**



Cross-sectional Design



Cross-Sectional Studies

- Participants of different ages studied at the same time.



Longitudinal Studies

- One group of people studied over a period of time.

