MEDICAL UNIVERSITY – PLEVEN
FACULTY OF PUBLIC HEALTH

DEPARTMENT OF PUBLIC HEALTH SCIENCES

DAY 1 INTERNSHIP

EPIDEMIOLOGY – DEFINITION AND SCOPE. BASIC CONCEPTS. MEASURING DISEASE FREQUENCY. COMPARING DISEASE OCCURRENCE.

Prepared by:
Assoc. Prof. Dr. Mariela Kamburova, MD, PhD

Definition and scope of epidemiology

- **Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control of health problems**

  John Last, 1988

Definition and scope of epidemiology

Epidemiology has three main objectives:

1. **To identify the causes of different diseases**
2. **To describe the distribution and the magnitude of health-related problems in human populations**
3. **To provide data for planning and implementing health promotion and disease prevention programmes in human populations**

Historical development, definition and scope of epidemiology

Epidemiology is also concerned with:

4. **The study of the course and outcome /natural history/ of diseases in individuals and groups /clinical epidemiology/**
5. **Evaluation of effectiveness and efficiency of interventions and health services**

Basic concepts in epidemiology

POPULATION - group of people, sharing common characteristics - industry workers, hospital patients, military recruits for a given year

POPULATION AT RISK - that part of a population which is susceptible to disease and from which the new cases could arise

Basic concepts in epidemiology

RISK GROUP - population group, that has higher frequency of the risk factors and higher probability of disease occurrence

Risk - probability of disease occurring

Basic concepts in epidemiology

RISK FACTOR - personal and life style characteristics, environmental factors, genetic or congenital characteristics, that increase the probability of disease occurrence and that have to be prevented.

Basic concepts in epidemiology

Risk factors have different contribution to health in human populations:

- **life-style factors - about 50%**
- **biological and genetic factors - about 20%**
- **environmental factors - about 20%**
- **factors related to health services - about 10%**

Basic concepts in epidemiology

EXPOSURE - specific factor that can be measured quantitatively by level and dose

/often used as synonym for risk factor/

EXPOSED GROUP - group of persons exposed to the influence of the factor /with a negative or positive effect/ under study

NONEXPOSED GROUP - the group that is not exposed to the influence of the factor under study

Prevalence -P

Measures the frequency of existing cases in a defined population :

- at a given point in time which is **Point Prevalence**
- during a specified period of time which is **Period prevalence**

Point prevalence

Measures the frequency of existing cases at one moment, in cross-sectional studies.

$$\frac{\text{Number of existing cases at a given point in time}}{\text{population at risk at the same point in time}} \times 10^{n}$$

Period prevalence

Measures the number of cases at the beginning of the period plus the newly developed cases, devided by the population at risk during that period.

$$\frac{\text{number of registered cases /old and new/ during a given period}}{\text{population at risk during the same period}} \times 10^n$$

Period prevalence

Prevalence is increased by:

- longer duration of the disease
- lower case-fatality
- medical technology, improving survival of patients
- increase in new cases due to changes in risk factors or improved disease diagnostic
- in-migration of cases
- out-migration of healthy people

Period prevalence

Prevalence is decreased by:

- shorter duration of disease
- high case-fatality
- improved cure rate of cases
- decrease in new cases
- in-migration of healthy people
- out-migration of cases

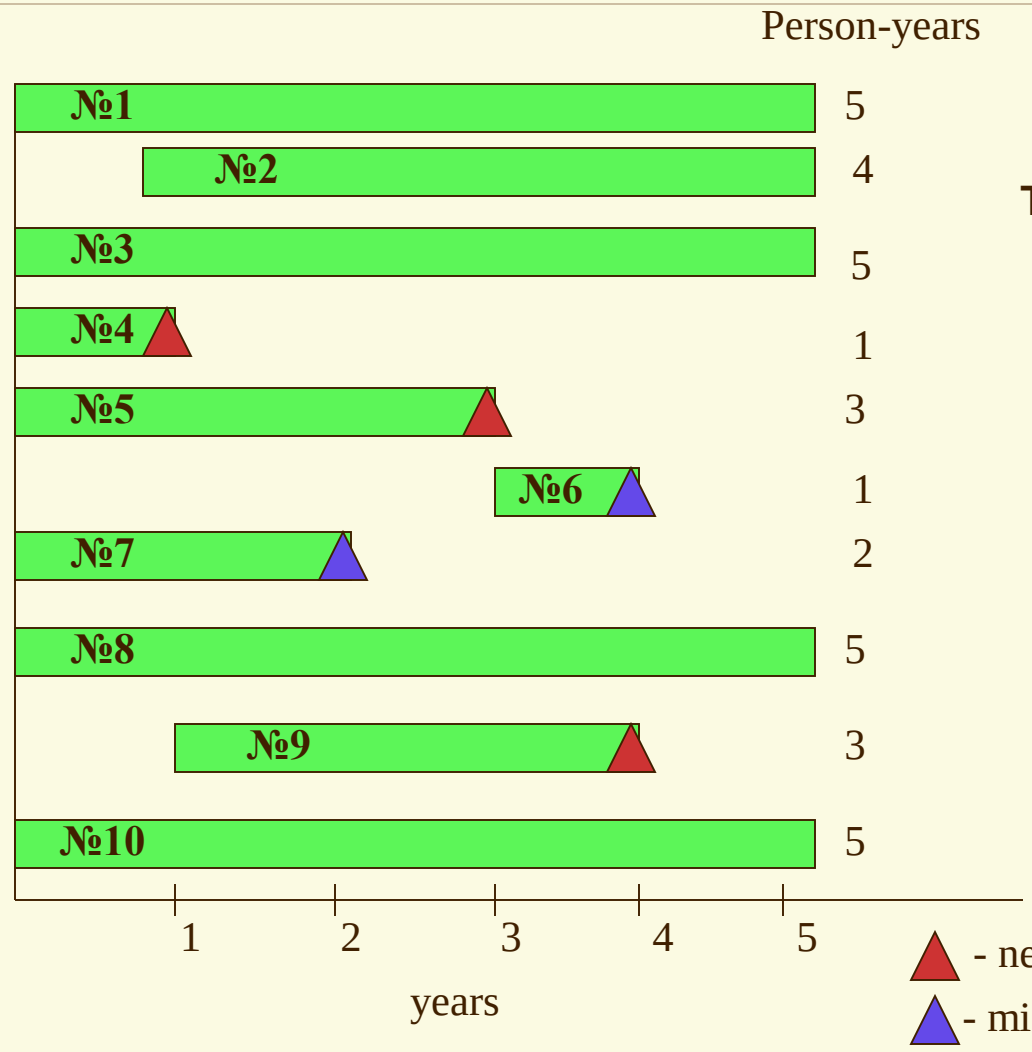
Incidence rate - I

Measures the number of new cases of disease that develop in a population at risk during a specified time period.

$$\frac{\textbf{number of new cases of a disease during a period}}{\textbf{sum of the individual time at risk for each person in the population at risk}} \times 10^n$$

The units of measurement always include a dimension of time /person-year, month,day/

Incidence in an open cohort



Total time at risk= 34 person-years

N of new cases = 3

Incidence

I = (3/34) x 1000 =

88 per 1000 pyrs.

Cumulative incidence - CI

Quantifies the frequency of newly developed cases in a closed cohort over a period of time

It is a measure of the risk of individuals in the population getting the disease during the specified period

$$\frac{\text{number of new cases of a disease during a period}}{\text{population at risk at the beginning of the period}} \times 10^n$$

Comparing disease occurrence

- **Epidemiological process begins with measuring the occurrence of disease in human populations.**
- **The next essential step is the comparison of disease occurrence in two or more groups of people whose exposure have differed.**
- **Those can be exposed and nonexposed individuals or exposed people who have different levels and duration of exposure.**

Comparing disease occurrence

Comparison can be absolute and relative.

ABSOLUTE COMPARISON - indicates on an absolute scale how much grater the frequency of disease is in the exposed group compared with the nonexposed.

RELATIVE COMPARISON - indicates how much more likely exposed group is to develop a disease than the nonexposed.

Measures of absolute comparison
RISK DIFFERENCE /RD/

/Excess risk, attributable risk of exposed/

It measures the absolute effect of the exposure or the excess risk of disease in exposed group compared with nonexposed.

$$RD = I_e - I_o = CI_e - Ci_o$$

Risk difference indicates the number of cases of the disease among the exposed group that can be attributed to the exposure itself.

Etiological fraction /EF/

It measures the proportion of the disease among exposed attributable to the exposure. Etiological fraction estimates the proportion of disease in exposed that could be prevented by eliminating the exposure.

$$EF = \frac{I_e - I_o}{I_e} \times 100 = \frac{CI_e - CI_o}{CI_e} \times 100$$

Etiological fraction /EF/

Example:

By eliminating smoking as a risk factor, 92% of lung cancer cases among exposed could be prevented, or ***92% of lung cancer incidence among exposed are due to the exposure.***

Population attributable risk /PAR/

It measures the proportion of disease in the total study population which is attributable to the exposure. Indicates the preventable proportion of the disease in the total population if eliminate exposure.

$$PAR = \frac{\mathbf{I_p - I_o}}{\mathbf{I_p}} \times 100 = \frac{\mathbf{CI_p - CI_o}}{\mathbf{CI_p}} \times 100$$

Population attributable risk /PAR/

$$PAR = \frac{I_p - I_o}{I_p} \times 100 = \frac{CI_p - CI_o}{CI_p} \times 100$$

Example:

Lung cancer incidence in the total population will be reduced by 83% if eliminate smoking.

RELATIVE COMPARISON
Relative risk /RR/

- Relative comparison quantifies the strength of association between exposure and disease
- **Relative risk** is a measure that can be calculated in **cohort studies** as the ratio of the incidence /cummulative incidence/ of disease among exposed divided by the corresponding incidence of disease in nonexposed.

$$RR = \frac{I_e}{I_o} = \frac{CI_e}{CI_o}$$

RELATIVE COMPARISON

Relative risk /RR/

$$RR = \frac{I_e}{I_o} = \frac{CI_e}{Ci_o}$$

Relative risk indicates the likelihood of developing the disease in the exposed relative to nonexposed /how many times the risk of developing disease in exposed is grater compared to nonexposed/.

Example:

$$RR = \text{——} = 12$$

12 times grater is the risk of exposed to develop Lung cancer in comparison with nonexposed.

RELATIVE COMPARISON

Relative risk /RR/

$$RR = \frac{I_e}{I_o} = \frac{CI_e}{Ci_o}$$

RR = 1 - there is no associaction, no effect

RR > 1 - positive association, risk factor

RR < 1 - inverse association, protective factor

RELATIVE COMPARISON
Odds ratio /OR/

- In case-control studies it is not possible to calculate the Incidence or Cummulative incidence as the size of the population at risk is not known because the participants are selected on the basis of disease status. Thus, RR can not be calculated as well.
- To estimate the magnitude of an association in case-control study we calculate ODDS RATIO. Odds ratio is very simmilar to the relative risk and has the same meaning and interpretation.

RELATIVE COMPARISON
Odds ratio /OR/

Disease → **Exposure ↓**	Disease Yes	Disease No	**Total**
Exposure Yes	a	b	a + b
Exposure No	c	d	c + d
Total	a + c	b +d	a+b+c+d

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

TYPES OF EPIDEMIOLOGICAL STUDIES

Epidemiological study

Epidemiological study is a scientific investigation to reveal the frequency and the distribution of disease in human populations and the relationship of disease to different potential risk factors.

- **Definition of the research question**
- **Formulation of hypothesis**
- **Testing the hypothesis in an appropriate study design**

Types of epidemiological studies. Observational vs. Experimental studies

Observational studies -

allow nature to take it course: the investigator measures and analyze but does not intervene and does not have control over the exposure or the progress of disease.

Experimental studies -

the investigator actively intervenes to change a disease determinant /exposure or behaviour/ or the progress of a disease through the intervention. The investigator is controlling the experimental situation.

Observational studies

- Descriptive - **they are limited to a description of the occurrence of disease or disease-related phenomena in a population according to *basic group characteristics, geographic location and time.***
- Analytical - **they analyse the relationships between health status and other variables and explain the observed pattern of occurrence of disease**
  - **ecological studies - populations**
  - **cross-sectional studies - individuals**
  - **cohort studies - individuals**
  - **case-control studies - individuals**

Experimental studies

- Randomized controlled trials - **an epidemiological experiment to study a new preventive or therapeutic regimen in groups of patients**
- Field trials - **an experiment that involve disease-free people considered to be at risk and the intervention is applied to each person individually**
- Community trials - **an experiment in which the intervention is applied to communities rather than individuals**

Descriptive studies

- **Describe and compare the patterns of disease occurrence in and between the populations in relation to person, place and time.**
- **They answer the following questions:**

1. **Who is getting the disease? /What are the basic characteristics of people who have the disease?/**
2. **Where the disease occurs? / What is the geographical distribution of the disease?/**
3. **When the disease occurs? / What is the pattern of disease occurrence in time ?/**

Descriptive studies

Description of disease pattern in relation to person:

1. ***Demographic characteristics*** **- age, sex, race, marital status**
2. ***Socio-economic characteristics*** **- education, occupation, income, religion**
3. ***Personal habits*** **- smoking, diet**
4. ***Biological characteristics*** **- Hb, Er, Leuc**
5. ***Genetic characteristics*** **- blood group, HLA-system**

Descriptive studies

Description of disease pattern in relation to place:

1. ***International comparisons*** **of different countries; the registration methods of event occurrence are of great importance for the validity of comparison**
2. ***National comparisons*** **of regions within countries**
3. ***Comparisons of small areas*** **within regions - urban/rural, areas within a city**

Descriptive studies

Description of disease pattern in relation to time:

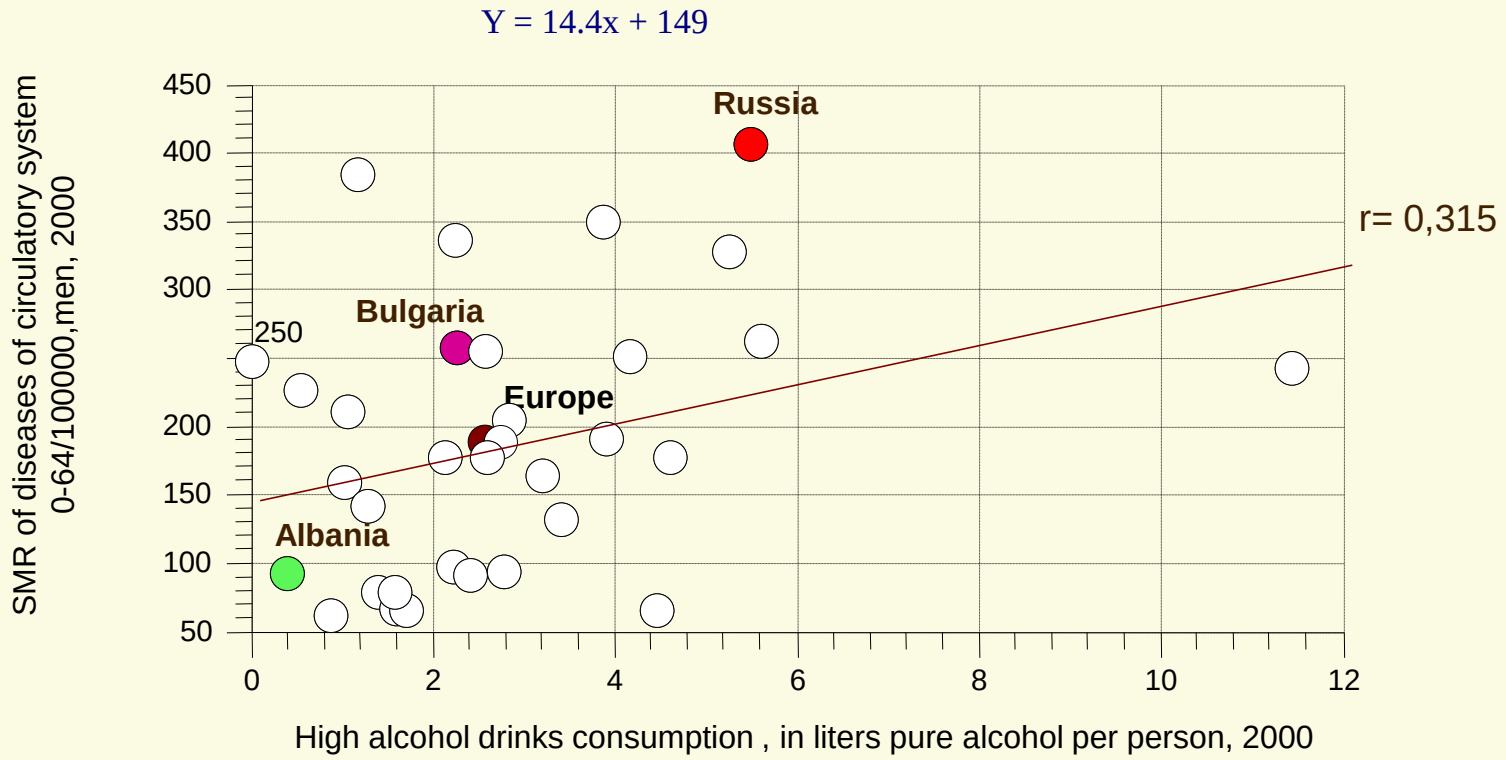
1. ***Short term changes*** **in disease occurrence - increases or decreases in disease incidence that are measured in hours, days, weeks or months; epidemics**
2. ***Recurrent /Periodic/ time trends*** **- seasonal variation, short-term periodical variation**
3. ***Long-term /Secular/ time trends*** **- progressive increase or decrease in disease occurrence that is manifested over years or decades**

Ecological /correlational/ epidemiological studies

Observational studies in which the units of study and analysis are populations or groups. Comparisons of disease occurrence are made between populations in different countries at the same time or in the same population at different times.

Rely on data available from routine national statistics; can be done quickly and inexpensively

Ecological study on relationship of alcohol consumption and cardiovascular mortality in men



Ecological /correlational/ epidemiological studies

Results are difficult to interpret since it is seldom possible to examine directly the various potential explanations for findings. As the study rely on data collected for other purposes information on different exposures and on some important population characteristics nay be not available. Thus the confounding effect can not be controlled.

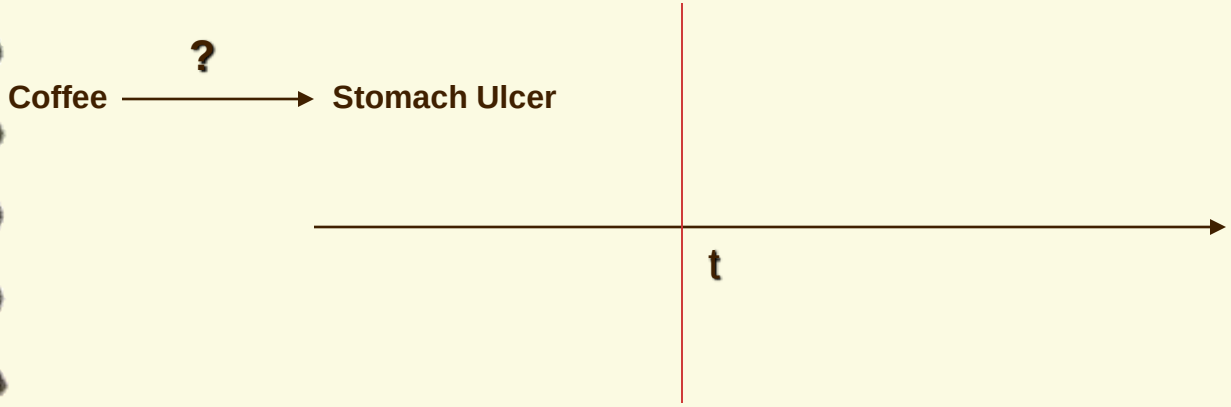
The main limitation is the possibility of *ecological bias* - inappropriate conclusions about the existing association between exposure and the disease at the individual level are drawn on the basis of the observed association at an aggregate /group/ level

Cross-sectional /prevalence/ epidemiological studies

Measure the prevalence of disease at a particular moment and the data are collected directly from the study subjects in a short period of time.

Data are collected on distribution of risk factors, health services utilization, health needs, self-perceived health status and other variables.

Cross-sectional study on stomach ulcer frequency and coffee consumption



Coffee + —— Stomach ulcer +

Coffee + —— Stomach ulcer -

Coffee - —— Stomach ulcer +

Coffee - —— Stomach ulcer -

Cross-sectional /prevalence/ epidemiological studies

- Carried out on representative samples
- Measure the exposure and the effect at the same time and it is not possible to determine the whether the exposure preceded or resulted from the disease
- Data collection rely on well-trained researchers and standardized methods

Cross-sectional /prevalence/ epidemiological studies

Advantages:

1. **Relatively quick, easy and economical to conduct. Prevalence of disease, risk factors frequency, health status and health services needs of the population are determined**
2. **Useful for investigating exposures that are fixed characteristics of the individuals**
3. **Can be used as a screening tool for detecting unknown cases of disease**
4. **Formulate etiological hypotheses but not able to test them**

Cross-sectional /prevalence/ epidemiological studies

Disadvantages:

1. **Not suitable for measurements of time relationship and for proving causality**
2. **Strict requirements to sampling methods and standardization of methodology and techniques of data collection**
3. **Not suitable for studying diseases with high case-fatality rate**

ANALYTICAL EPIDEMIOLOGICAL STUDIES. COHORT STUDIES. CASE-CONTROL STUDIES.

ANALYTICAL STUDIES

- **The objective – to test epidemiological hypotheses**
- **The subject of interest - the individual within the population**

Cohort studies

Case-control studies

COHORT STUDIES
/Follow-up studies, prospective studies/

- The cohorts are identified prior to the occurrence of the disease under study
- The study groups are observed over a period of time /follow-up/ to determine the frequency of the disease among them
- The study proceeds forward – from cause to effect

COHORT STUDIES

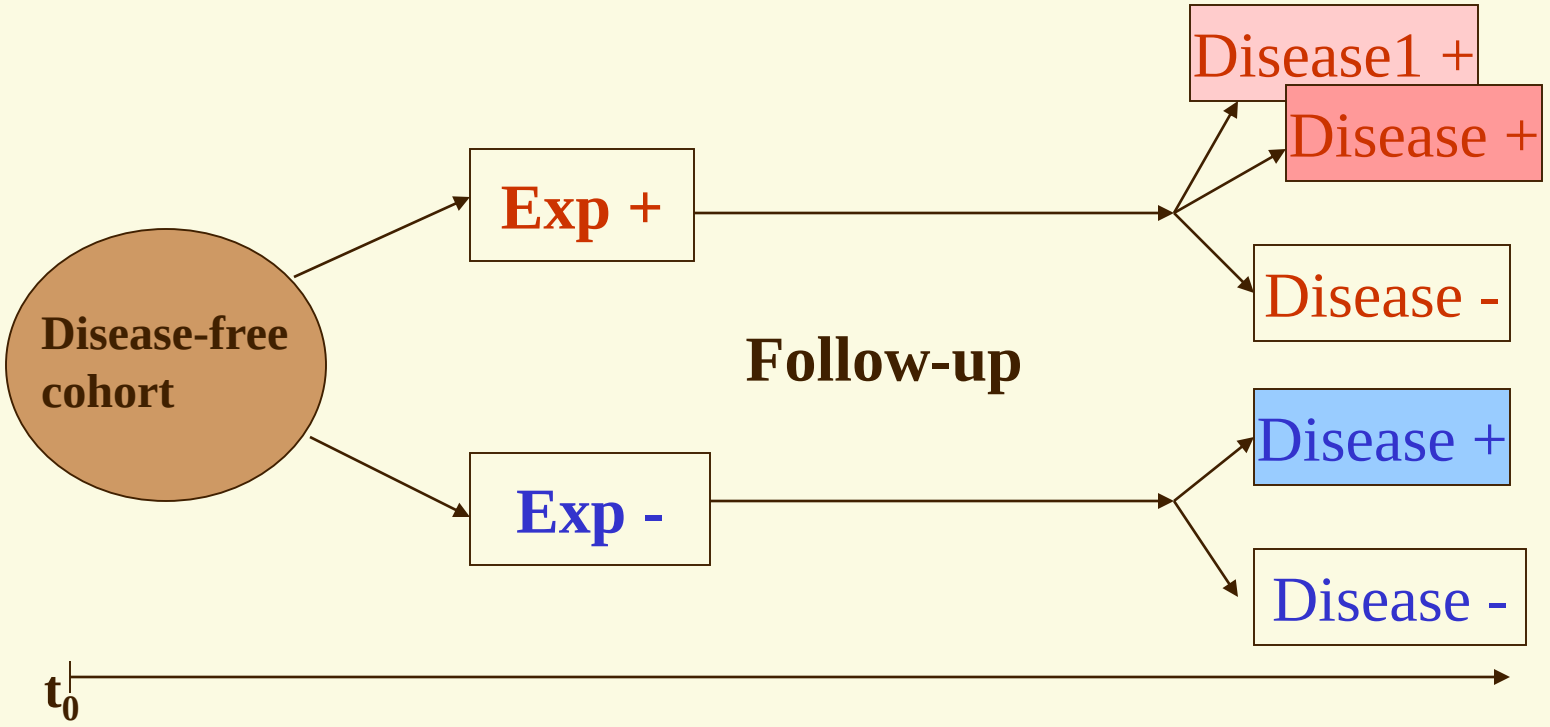
- Cohort – group of people who share a common characteristic or experience within a defined time period

e.g. age, occupation, exposure to a drug treatment, pregnancy, etc.

TYPES OF COHORT STUDIES

1. **Prospective cohort studies /current CS/ -**

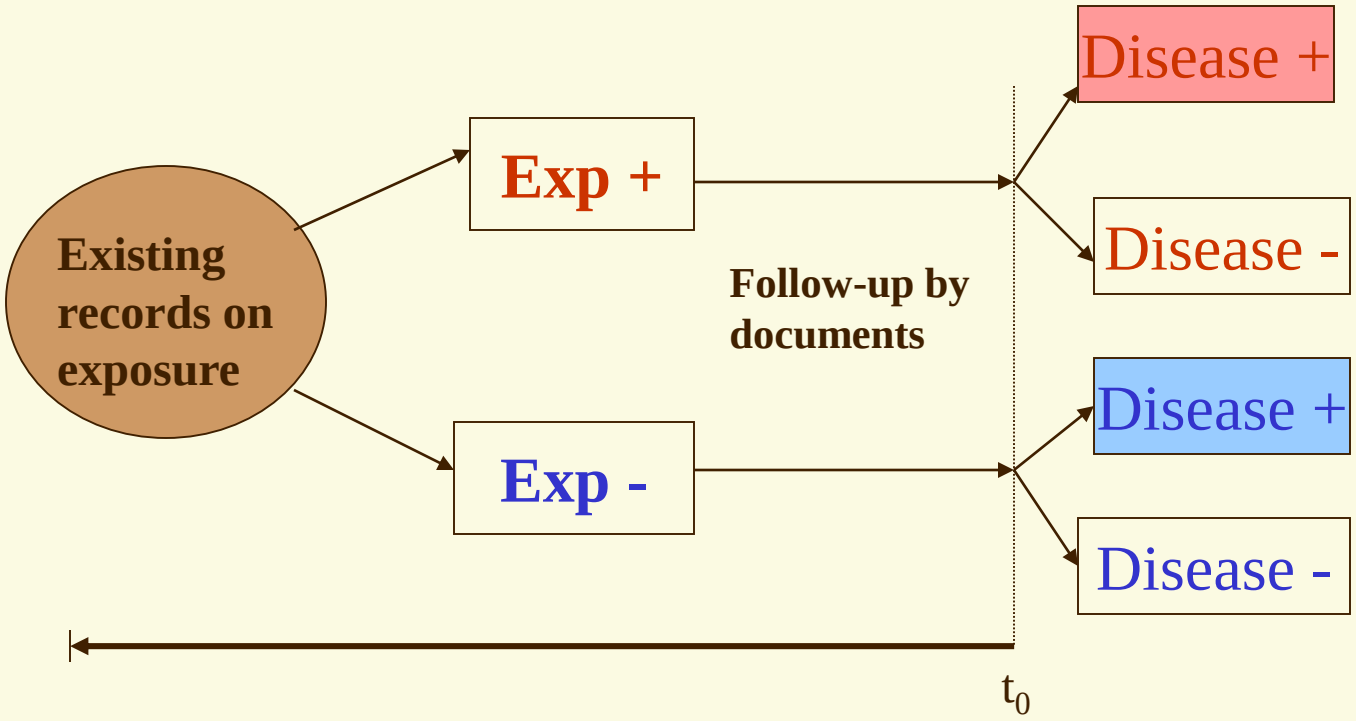
   the outcome has not yet occurred at the time the study begins



TYPES OF COHORT STUDIES

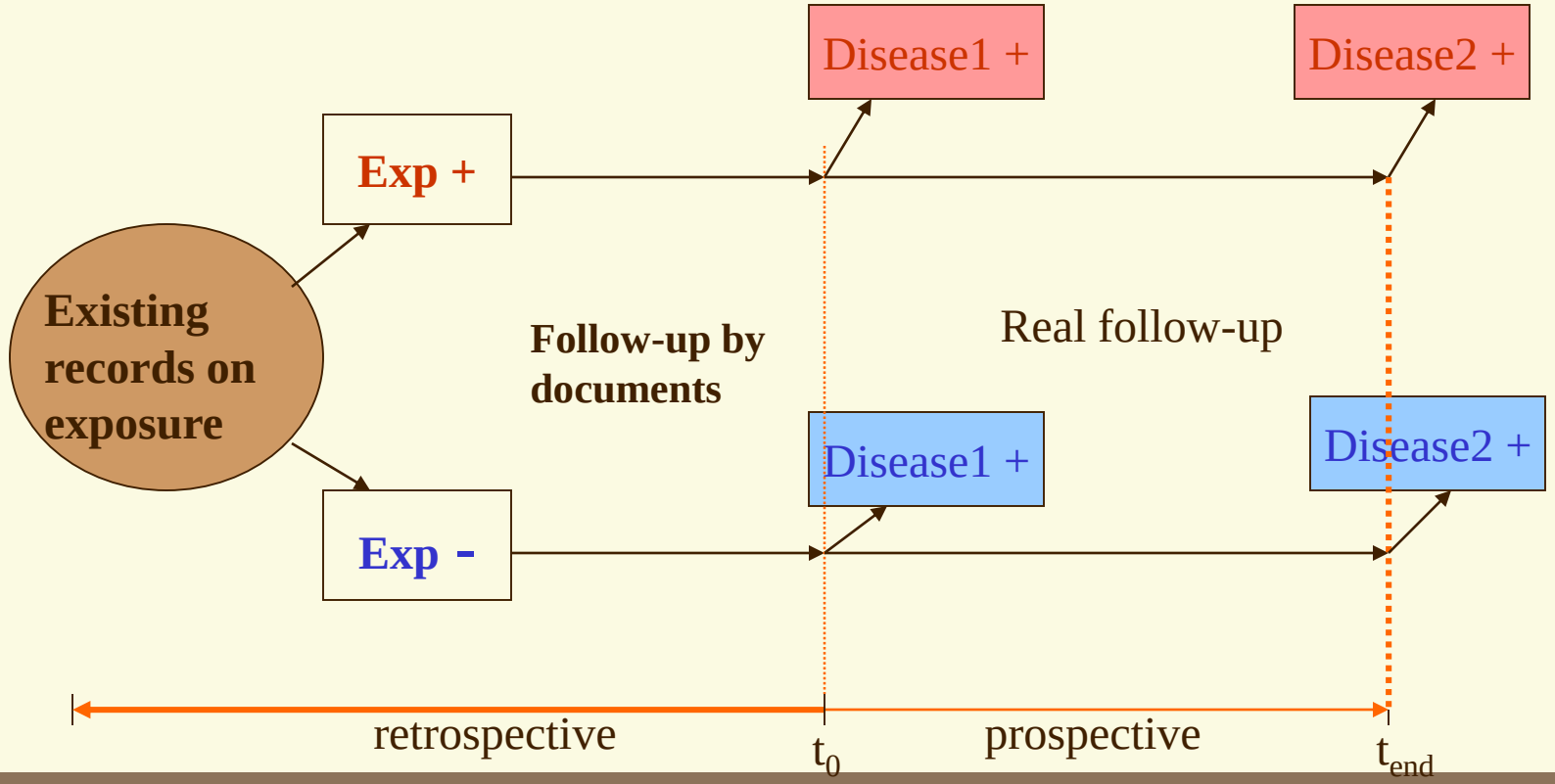
2. Retrospective cohort studies /historical CS/ -

the outcomes and the exposure have all occurred before the start of the study



TYPES OF COHORT STUDIES

3. Ambispective cohort studies – combination of retrospective and prospective cohort studies



CONDUCTING OF COHORT STUDIES

2. **Follow-up:** the groups are followed under the same identical conditions over a period of time to determine the outcome

- **Sufficient time period**
- **Procedures for following-up the groups:**
  - **Periodical medical examination**
  - **Review of medical records**
  - **Routine surveillance of death records**
  - **Mailed questionnaires, telephone interviews, periodical home visits**
- **Losses to follow-up – due to death, migration, withdrawal** **May bias the results!**

Try to achieve 95% follow-up of the cohort

CONDUCTING OF COHORT STUDIES

3. Measuring the disease frequency in exposed and non-exposed groups –

direct measurement of Incidence Rate or Cumulative Incidence

4. Analysis –the measure of association is RR

$$\text{RR} = \frac{Ie}{I_0} = \frac{CIe}{CI_0}$$

All measures of absolute and relative comparison of disease incidence can be calculated

ADVANTAGES AND DISADVANTAGES OF COHORT STUDIES

+ Incidence can be calculated
+ Several possible outcomes related to exposure can be studied simultaneously
+ Direct estimation of RR
+ Dose-response ratios can be calculated
+ Minimum bias

ADVANTAGES AND DISADVANTAGES OF COHORT STUDIES

- Large number of participants
- Not suitable for rare diseases
- Long duration /10-30 yrs./
- Difficult to follow-up the cohort
- Administrative problems - lack of funds, staff, extensive record keeping
- Expensive
- Ethical problems

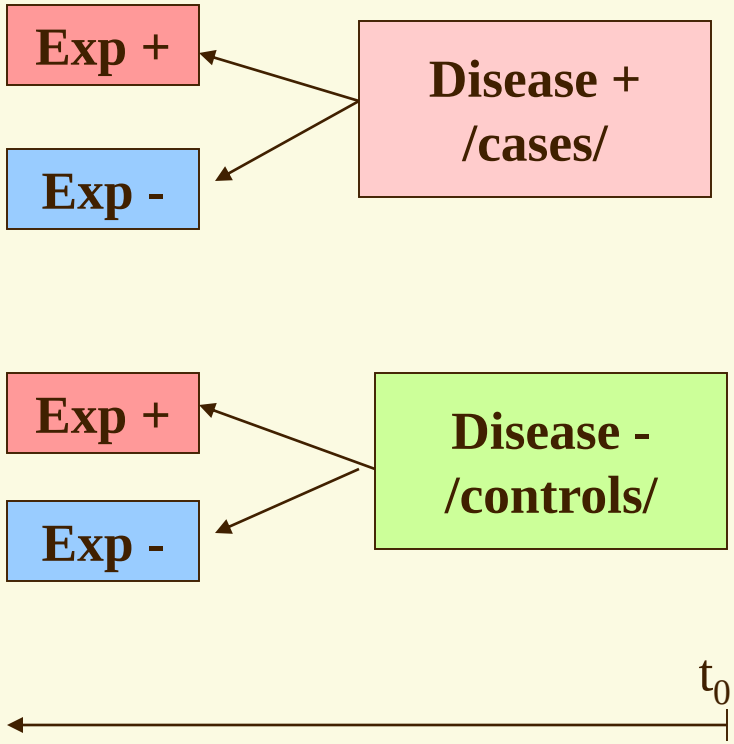
CASE-CONTROL STUDIES

CASE-CONTROL STADIES

- Both exposure and outcome /disease/ have occurred before the start of the study
- The study proceeds backwards from effect to the cause
- It uses a control group

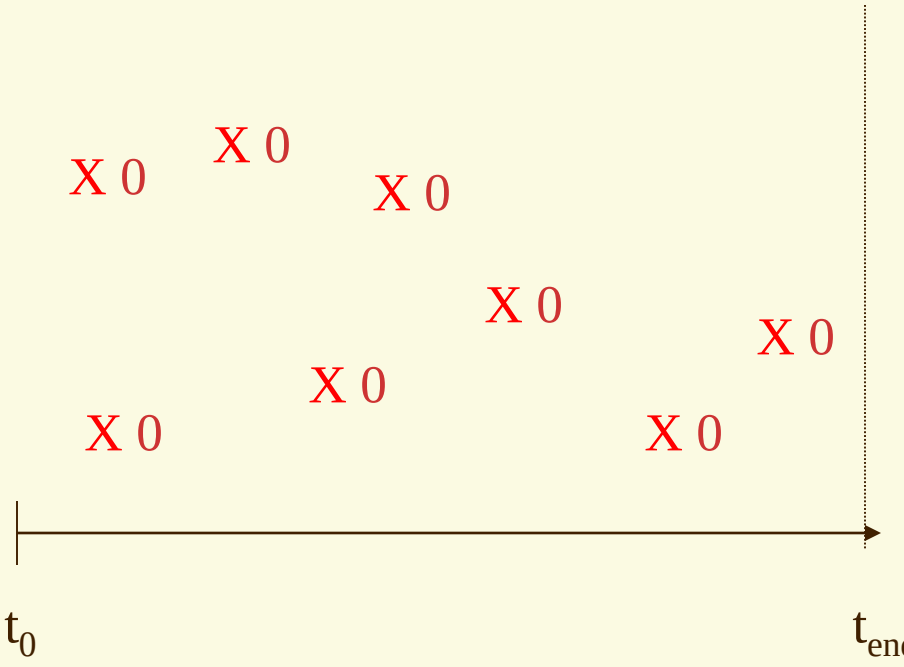
TYPES OF CASE-CONTROL STADIES

- Retrospective case-control studies



TYPES OF CASE-CONTROL STADIES

- Prospective case-control studies



X – cases

0 - controls

CONDUCTING CASE-CONTROL STUDY

- **Selection of cases:**
  - Clear definition of a case, diagnostic criteria
  - New cases of disease instead of prevalent cases
- **Selection of controls – crucial point!**
  - As similar to the cases as possible, except for the absence of disease under study
- **Matching – prevents confounding, ensures comparability of the groups**

  Do not use too many criteria for matching – do not overmatch!

CONDUCTING CASE-CONTROL STUDY

- Measurement of exposure – bias is possible if the information for the groups is differently collected, analyzed and reported in relation to their disease status
- Investigator bias
- Measurement bias
- Recall bias

CONDUCTING CASE-CONTROL STUDY

Analysis – to find out an estimation of disease risk associated with the exposure

Measure of association is Odds Ratio.

ADVANTAGES AND DISADVANTAGES OF CASE-CONTROL STUDIES

+ Relatively easy to carry out
+ Rapid and inexpensive
+ Require less subjects
+ Suitable for rare diseases
+ Allows the study of several different exposures
+ Risk factors can be identified
+ Ethical problems minimal

ADVANTAGES AND DISADVANTAGES OF CASE-CONTROL STUDIES

- Problem of bias – selection bias, recall bias, investigator bias
- Selection of controls may be difficult
- Incidence can not be measured
- Prone to confounding
- Problem of representativeness of cases and controls
- Not suitable for evaluation of therapy or disease prevention

MAIN DIFFERENCES BETWEEN CASE-CONTROL AND COHORT STUDIES

CASE-CONTROL

- Proceeds from effect to cause
- Starts with the disease
- Tests whether the exposure occurs more frequently among cases than among controls
- Usually the first approach to the testing of a hypothesis
- Less study subjects
- Quick
- Rare disease
- Indirect estimation of risk- OR
- More exposures
- Inexpensive

COHORT

- Proceeds from cause to effect
- Starts with the exposure
- Tests whether the disease occurs more frequently in exposed than in non-exposed group
- Reserved for testing of precisely formulated hypotheses
- Larger number of subjects
- Long period of time
- Rare exposure
- Measures IR or CI and RR
- More outcomes
- Expensive