THE STATE UNIVERSITY OF MEDICINE AND PHARMACY "NICOLAE TESTEMIŢANU"

Department of Management and psychology

Descriptive and Observational Studies

Spinei Larisa,
PhD. DHMS.,professor

Globa Nina,
MPHM, assistant professor

2020

Discussion topics

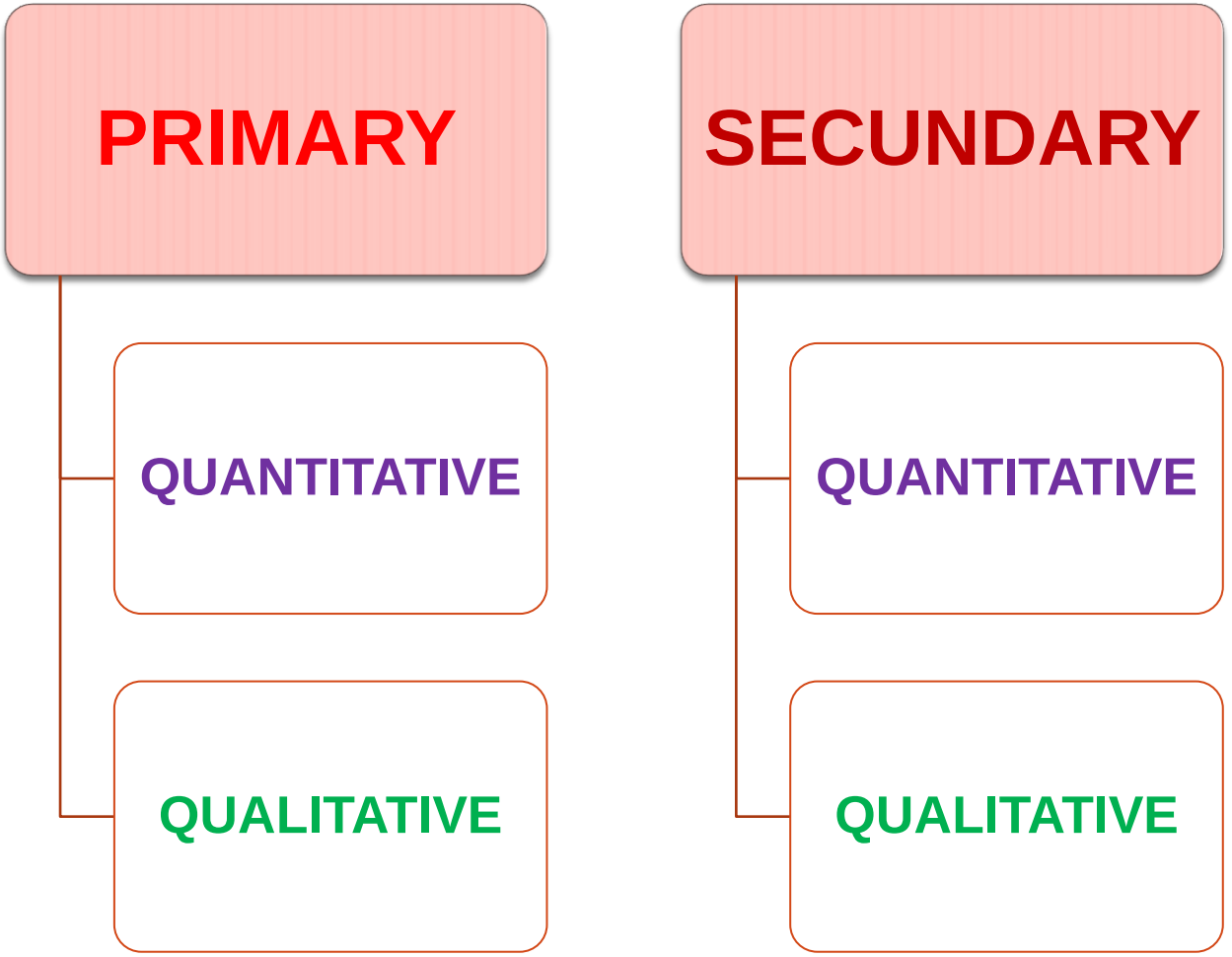
1. **Classification of study designs.**
2. **Descriptive studies.**
3. **Cohort Studies.**
4. **Cases – Control Studies.**

Types of research

1. **Primary – made by the researcher from a practical point of view.**

2. **Secondary - analysis of the results of previous research from published literature and unpublished results.**

Types of research



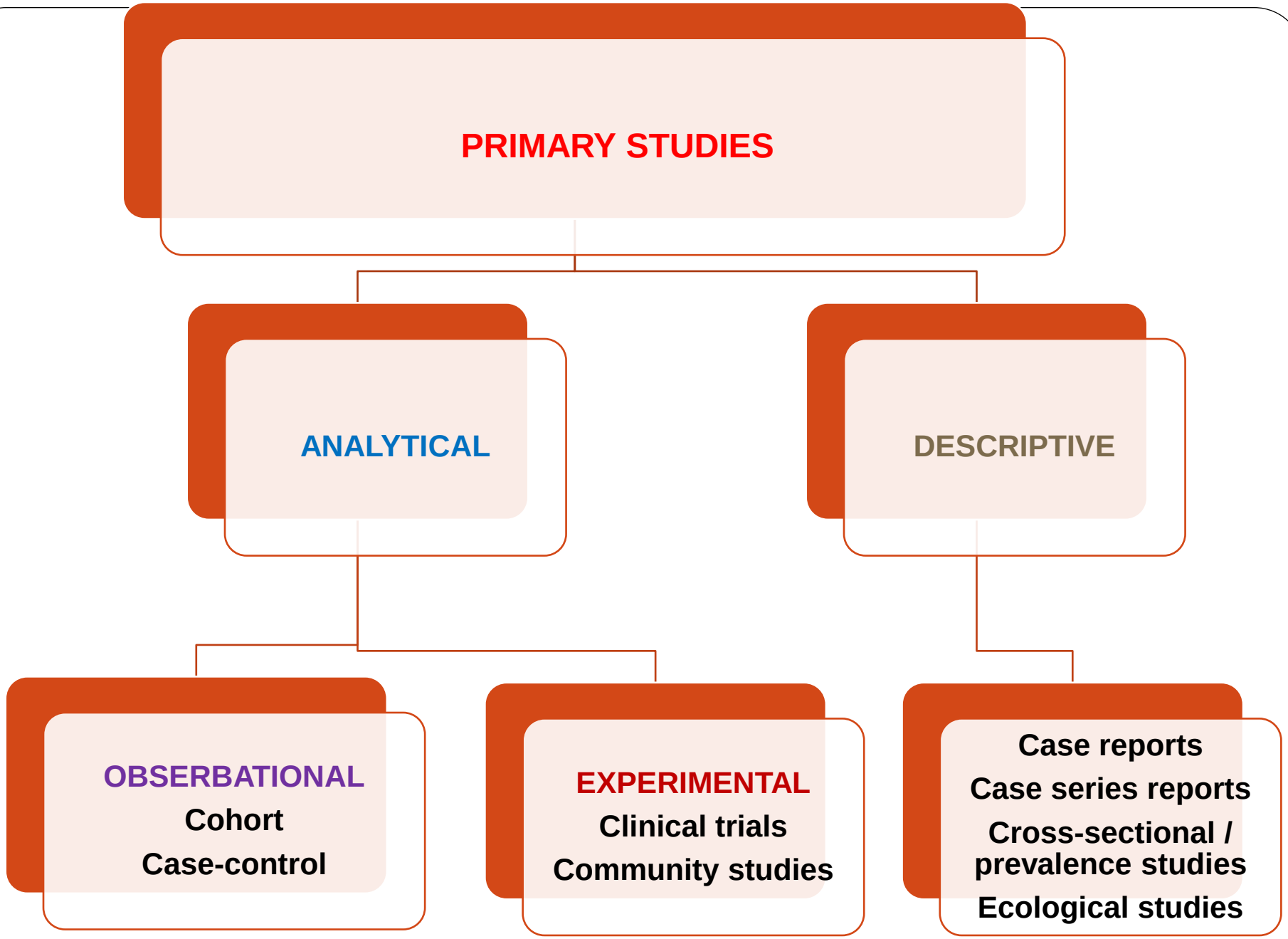
The basic types of research

- **Qualitative – used to understand the nature or quality of phenomenon. The questions are:**

What? Who? How? When? Why?

- **Quantitative – used to understand the magnitude of an occurrence or an association. The question is:**

How much?


PRIMARY STUDIES
ANALYTICAL
DESCRIPTIVE
OBSERBATIONAL
Cohort
Case-control
EXPERIMENTAL
Clinical trials
Community studies
Case reports
Case series reports
Cross-sectional / prevalence studies
Ecological studies

Types of research

- **Descriptive study** is limited to a description of the occurrence of a disease or other phenomenon . Do not have a comparison group. Generate hypothesis.

Answer What? Who? When? Where?

- **Analytical study** analyses relationships between health status and other variables. Have a comparison group. Test hypothesis.

Answer Why? How?

Classification of research methods

- **Observational/ Non-experimental -** allow nature to take its course: the investigator measures but does not intervene
- **Experimental / Interventional –** involve an active attempt to change a disease determinant – such as an exposure or a behaviour - or the progress of a disease through treatment, and are similar in design to experiments in other sciences.

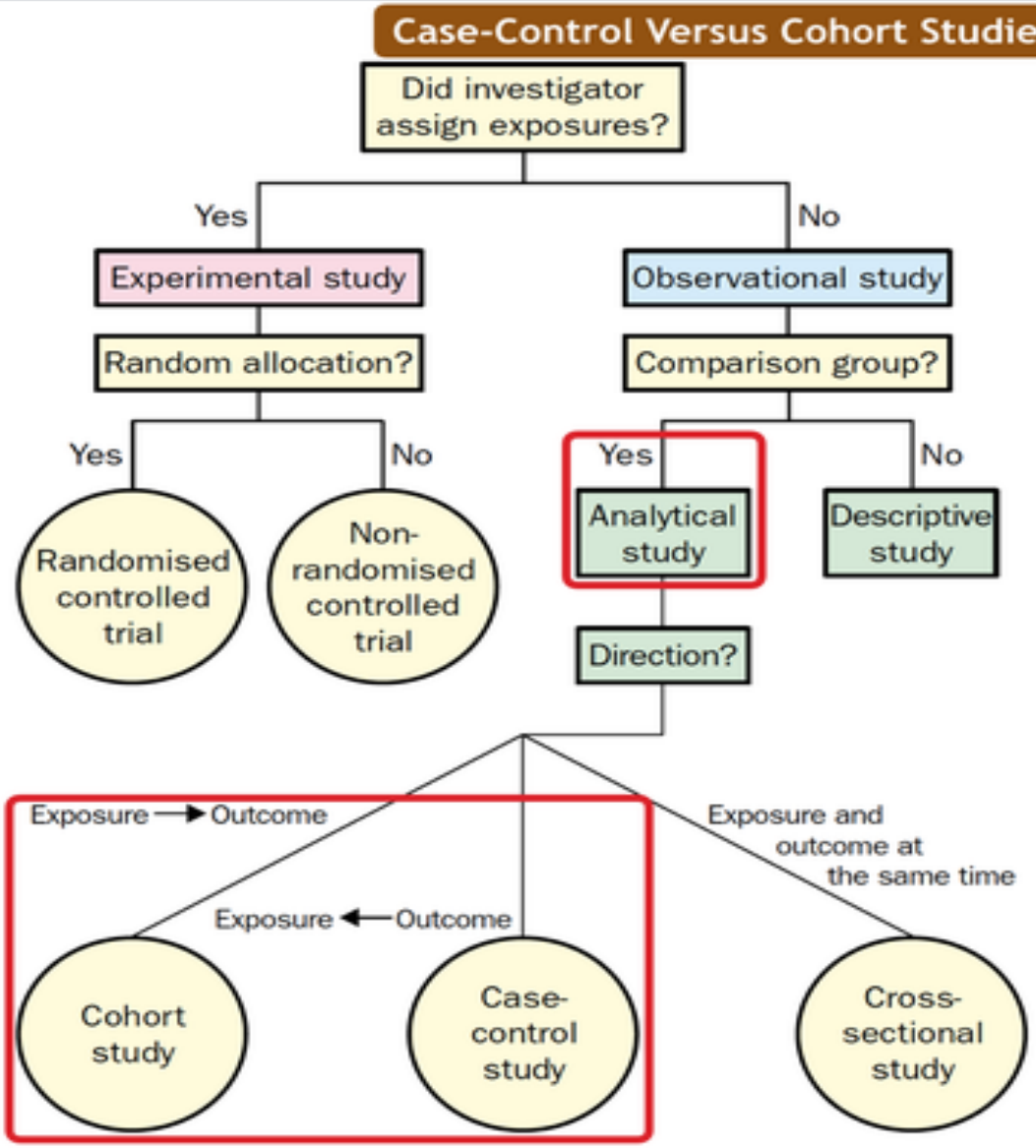


Figure 1: **Algorithm for classification of types of clinical research**

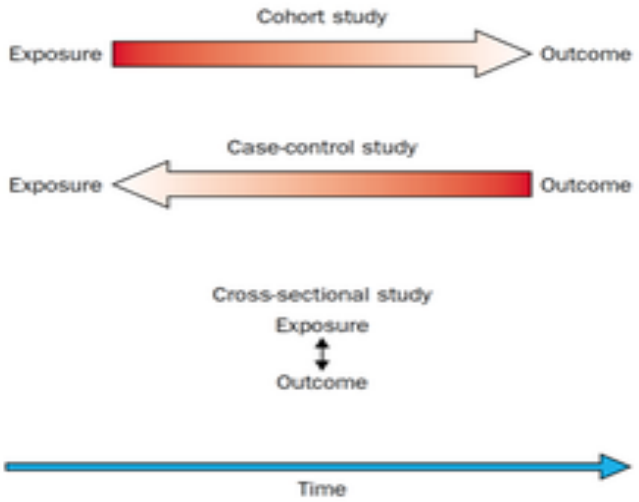
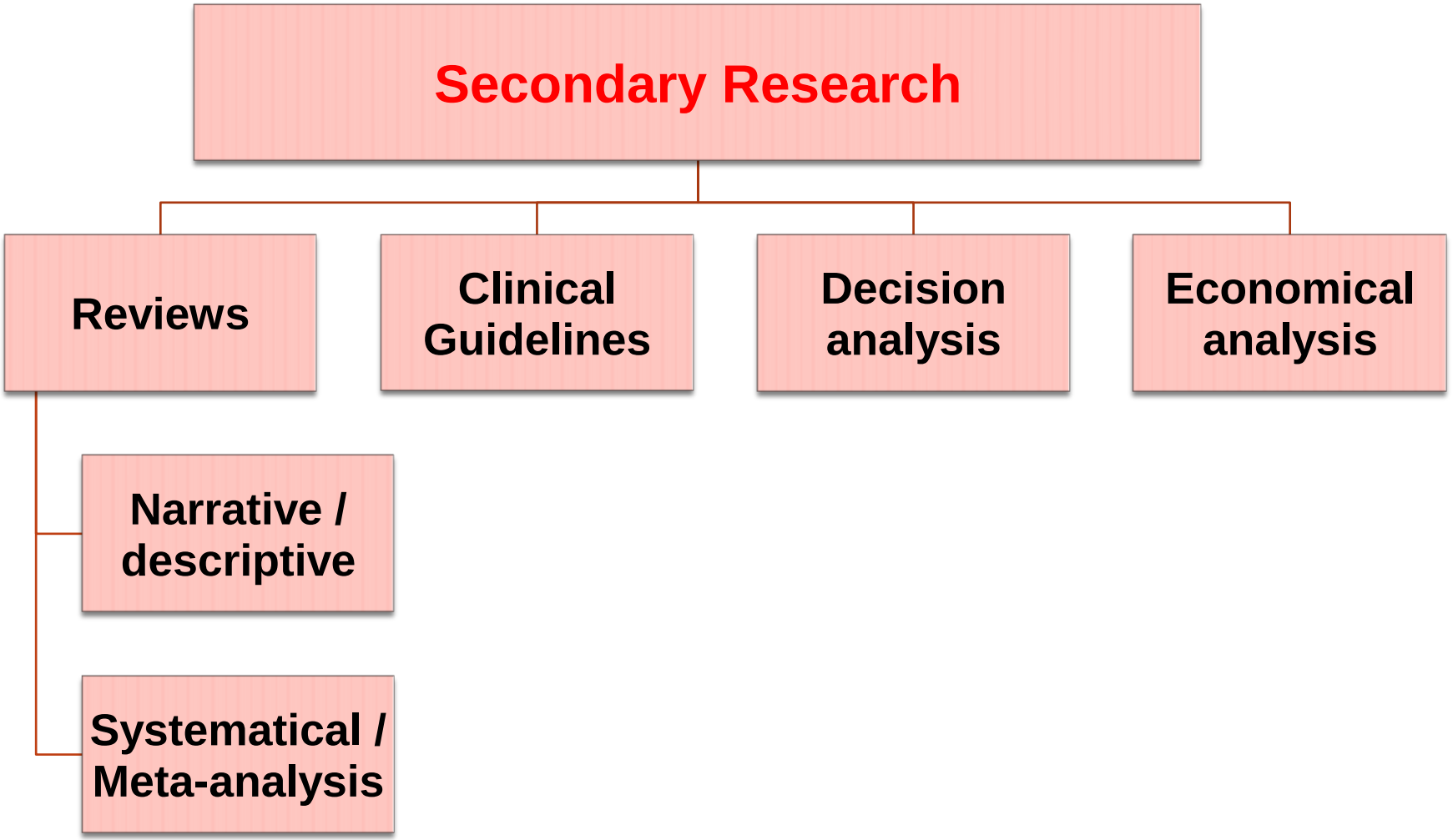


Figure 2: **Schematic diagram showing temporal direction of three study designs**

Modified from Grimes, David A., and Kenneth F. Schulz. "An overview of clinical research: the lay of the land." The lancet 359.9300 (2002): 57-61. https://pdfs.semanticscholar.org/e37a/8c9cb7b3814494128351e1051c068673092b.pdf


Secondary Research
Reviews
Clinical Guidelines
Decision analysis
Economical analysis
Narrative / descriptive
Systematical / Meta-analysis

2. Descriptive studies (DS)

Individual:

- **Case reports**
- **Case series**
- **Cross - sectional/ prevalence studies**

Aggregate/ Group:

- **Ecological studies**

Descriptive studies allow

- Identification of health problems
- Establish the hierarchy of health problems according to the existing risk factors
- Establish the frequency of the disease in the population
- Determine the severity of the disease at the population level
- Determine the social impact of the disease on the population

Characteristics

- **DS describe models of occurrence of disease or of exposure to risk factors depending on Person, Place, and Time.**
- **DS must answer the following questions: What? Who? When? Where?**
- **Generally involve people seen over a relatively short time.**

Characteristics

- **DS is used for generation of hypotheses which can be investigated for ex. during a case-control or cohort study .**
- **This study doesn't include control subjects (a comparison group) but**

We can do 3 types of comparisons:

1. **Demographic**
2. **Geographic**
3. **Time comparisons**

Demographic characteristics

Who is affected?

Characteristics of person:

- **age,**
- **sex,**
- **social status, marital status,**
- **economic status,**
- **personal history,**
- **living and working conditions,**
- **hereditary characteristics etc.**

Geographic characteristics

- **Where is the problem?**

The place:

- **geographical area,**
- **country,**
- **district,**
- **city,**
- **village,**
- **organization etc.**

Time characteristics

- **When is the problem?**

Time:

- **Year, period of years,**
- **Season – summer or winter,**
- **Months,**
- **Weeks,**
- **Days, etc.**

Temporal changes

- **Short-term fluctuations**

  epidemics

- **Cyclic changes**

  seasonal changes

- **Secular trends**

  changes over several years

Descriptive / case study– case series studies

A case or case – series report is a simple descriptive account of interesting characteristics observed in a person (case study) or group of patients (case –series study)

Case report study

- **Describe the experience of one or several patients(less than 5) with similar diagnosis**
- **Describe the way in which clinicians identify unusual characteristics of some diseases**
- **They may be the first clues in identifying new diseases, or the effects of an exposure**
- **They represent more than 1/3 of published medical articles, but the result can not be generalize because of scientific base.**

Case-series studies

- **Number of cases is more than 5 and less than 100**
- **Reporting of dates of patients group with similar diagnosis**
- **May generate hypothesis and describe new symptoms and syndromes**

Some limits:

- **Subjective way of subjects selection**
- **Can not test hypothesis, because do not have a comparison group**

Cross – Sectional Studies (CSS)

- **Purpose – to learn about a characteristics of the population at one point in time like a photo "snap shot"**
- **Analyze data collected from a group of subject at one time rather that over a period of time**
- **CSS are designed to determine "What is happening" right now.**

Cross – Sectional Studies (CSS)

- Subjects are selected and information is obtained in a short period of time.
- Because they focus on a point in time, they are sometimes called prevalence studies
- Survey and polls are cross-sectional studies
- CSS are used in all fields of medicine.

Cross-sectional studies

- The presence of a disease and a risk factor in a given population
- No reference is made to their past or future evolution
- Cross-sectional approaches measure the outcome and exposure simultaneously in a well-defined population

Advantages of CSS

- CSS are best for determining the status quo of the disease or condition
- CSS are relatively quick to complete
- May be relatively inexpensive as well
- Are indicated to identify the prevalence of common diseases

Disadvantages of CSS

- **They provide only a "snapshot in time" of the disease or the process, which may result in misleading information if the research question is really one of disease process.**
- **A common problem with survey research is obtaining sufficiently large response rates; many people asked to participate in a survey decline because they are busy, not interested, and so forth.**
- **The people who agree to participate may not be representative or similar to the entire population**
- **Another issue is subjective - the way questions are posed to participants; if questions are asked in a leading or emotionally way the responses may not truly represent the participants' feelings or opinions.**

Ecological / correlation studies

- They use measurements that represent the characteristics of the whole population to describe the results in relation to some factors of interest (age, time, use of services, exposures, etc.).

Advantages:

- They can generate hypotheses for analytical studies
- They can focus populations at risk for certain periods of time for geographical regions for future studies

Ecological / correlation studies

Disadvantages:

- **Data being for groups, cannot be linked to outcome and exposure to individuals**
- **Cannot be controlled for confounders**
- **Data are average exposures and not individual exposures**
- **Cannot determine a dose-response relationship**

Advantages and Disadvantages of Descriptive Studies

Advantages:

- **They are easy to do and to write**
- **The observations may be extremely useful to investigators designing a study to evaluate causes or the explanations of the observations**
- **They are less expensive**
- **Offer a scientifically base for health planning, delivery and evaluation of health services**

Disadvantages:

- **Are susceptible to many possible biases related to the subject selection and characteristics observed.**
- **Are hypothesis – generating and not conclusive.**
- **The temporal relationship between exposure and outcome is difficult to determine**
- **Do not have comparison group etc.**

Observational studies

- **Measures the strength of the epidemiological association between an Exposure and an Outcome and tests epidemiological hypotheses**

Direction

- **The question:**

What do we start the research with? - with the Exposure or with the Result

FORWARD – from Exposure to Result (cohort)

BACKWARD – from Result to Exposure (case – control)

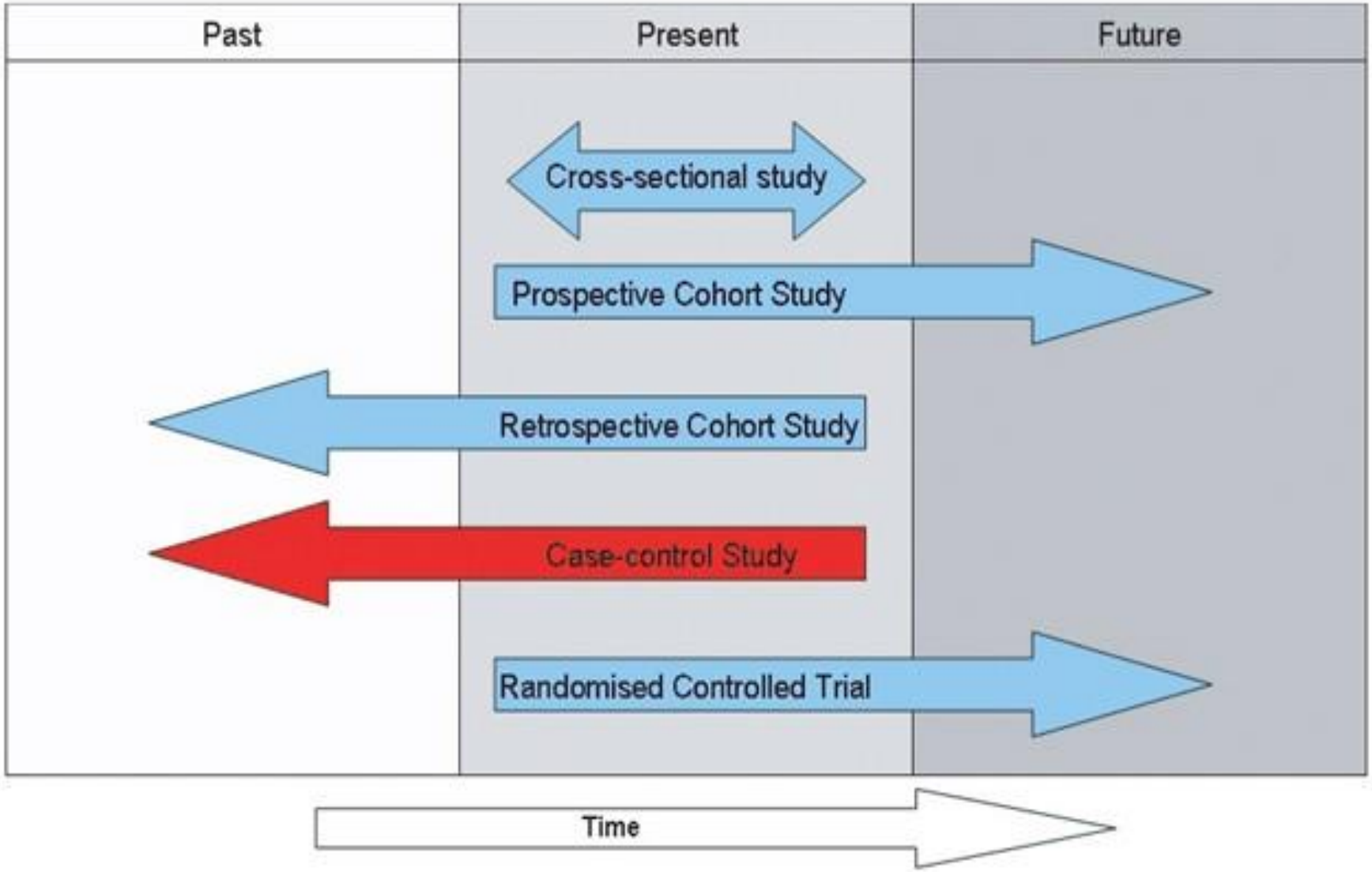
TEMPORARY SEQUENCE

- **The question: “**

Did Result appear at the time of initiating the research or not?

Prospective studies – the Result occur after the study start (cohort)

Retrospective studies - the Result occur before the start of study (case-control, cohort - as a component)


Past
Present
Future
Cross-sectional study
Prospective Cohort Study
Retrospective Cohort Study
Case-control Study
Randomised Controlled Trial
Time

3. Cohort Studies (CS)

- **A cohort is a group of people who have something in common and who remain part of the group over an extended time.**
- **CS ask the question "What will happen?"**
- **The direction in cohort studies is forward in time and this study is also called prospective.**

Design of Cohort Study

In this study the investigator selects a group of EXPOSED individuals and a group of NON-EXPOSED individuals and follows up both groups over a certain period to compare the incidence of RESULTS in the two groups.

If the positive association exists between the EXPOSURE and the RESULT we would expect that the proportion of the exposed group in whom the result develops would be greater that the proportion of the non-exposed group in whom the result develops.

The cohort study start with

- **HYPOTHESIS**

- **Ho – null hypothesis**

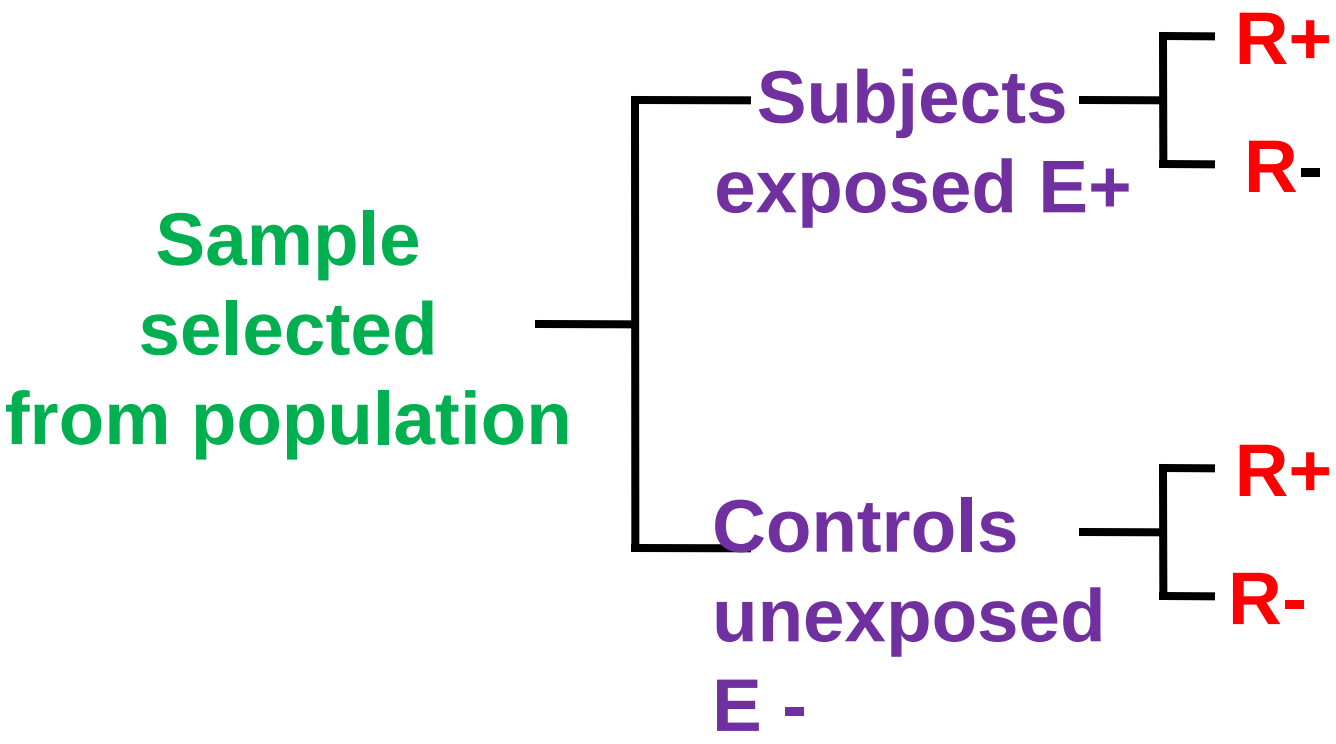
There is no association between exposure and outcome

- **Ha - Alternative hypothesis**

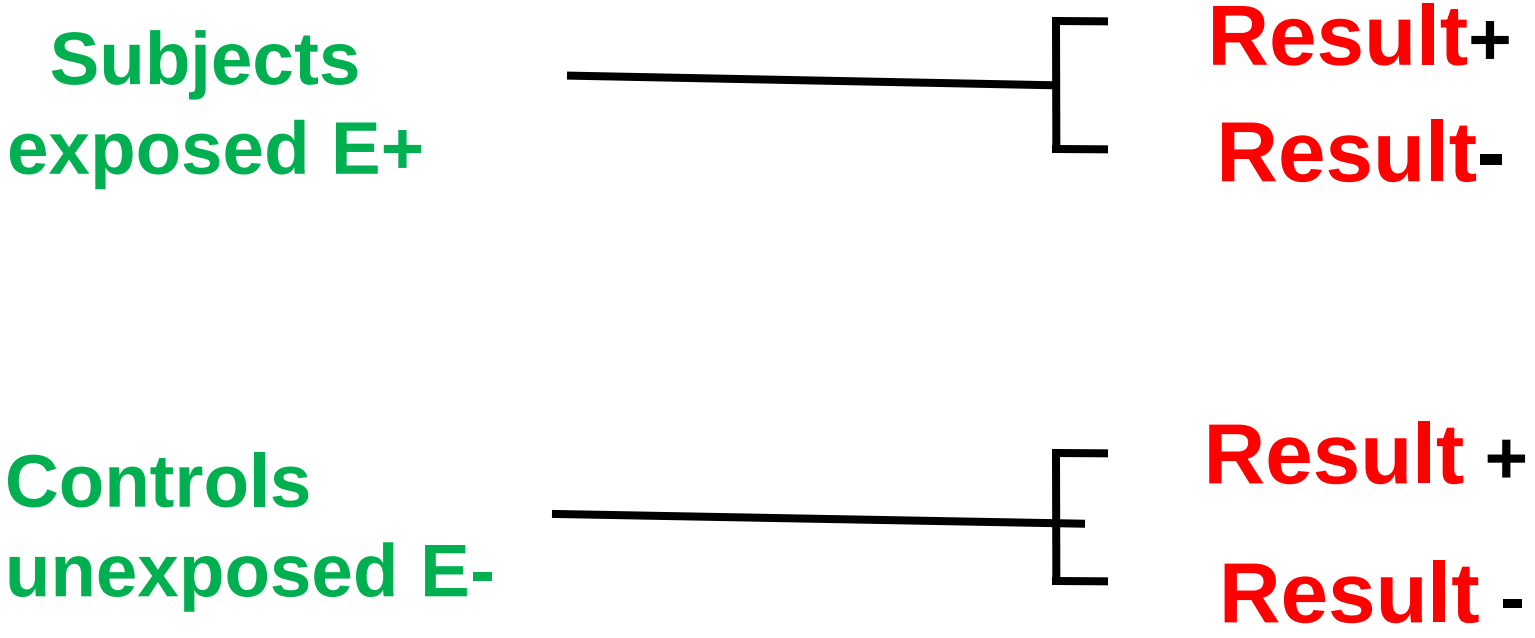
Cohort study

- We want to know if excessive sugar consumption will lead to the development of caries

The Diagram of Cohort Study Design Type 1



The Diagram of Cohort Study Design Type 2



Cohort Study

1. **Prospective cohort**

- **Cohort identification based on current Exposure**
- **Tracking the cohort in the future for a result**

2. **Retrospective cohort**

- **Identifying the cohort based on past exposure**
- **Cohort follow-up after Exposure**
- **The result has already appeared**

Measurement instruments

- Questionnaires
- Laboratory tests
- Instrumental investigations
- Physical measurements
- Observation sheets
- Medical records

The contingency 2x2 table

	WITH RESULTS	WITOUT RESULTS	TOTAL
EXPOSED	a	b	a+b (m_1)
NON-EXPOSED	c	d	c+d (m_0)
TOTAL	a+c (n_1)	b+d (n_0)	a+b+c+d (t)

Design of the Cohort Study

- **Then follow to see whether**

	Result develop	*Result does not develop*	*Totals*	*Incidence rates*
1. Exposed	**a**	**b**	**a+b**	**a / a+b**
2. Non Exposed	**c**	**d**	**c+d**	**c / c+d**

Indicators in the Cohort Study

- **Risk of disease in exposed:**

$R_1 = a / (a + b)$

- **Risk of disease in unexposed:**

$R_o = c / (c + d)$

Indicators in the Cohort Study

The Relative Risk (determine whether there is an association between exposure to a factor and development of a outcome)

$$\text{Relative Risk} = \frac{\text{Risk in exposed}}{\text{Risk in non-exposed}}$$

$$\text{Relative Risk} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{a\,(c+d)}{c\,(a+b)}$$

Relative risk

- RR = R1/Ro

How many times?

Risk in exposed is higher than risk in non-exposed

Risk difference

- **Define as a difference between individual risk in exposed and individual risk in non-exposed**

- **RD= R1 – Ro**
- **How much?**
- **Risk in exposed is higher than risk in non-exposed**

ATTRIBUTABLE RISK PERCENT

- **The attributable risk percent, expresses how many % of the effect present in exposures can be explained by the established exposure**

$$AR\% = [(R_1 - R_0)/R_1] \times 100$$

The RISK in Population

- **Express the frequency of those exposed to the risk factor in the studied group**

$$Rp = (a + b) / (a + b + c + d)$$

- **Excess of risk in the population (Attributable Risk)**

$$Rap = Rp - Ro$$

Other very important indicators

CI -confidence interval, for RR

SA -strength of association

DETERMINING OF CI

- $CI = RR^{(1 \pm z / x)}$

$$x^2 = \frac{(t-1)[(a \times d) - (b \times c)]^2}{n_1 \times n_0 \times m_1 \times m_0}$$

- For 95% veracity , probability, z = 1,96

$$CI_{sup.\ lim.} = RR^{(1 + z / x)}$$
$$CI_{inf.\ lim.} = RR^{(1 - z / x)}$$

Interpreting the Relative Risk

- **If RR = 1 Risk in exposed equal to risk in non-exposed (no association)**
- **If RR >1 Risk in exposed is greater than risk in non-exposed (positive association, possibly causal, harmful factor)**
- **If RR < 1 Risk in exposed is less than in nonexposed (negative association, possibly protective factor)**

RESULT EVALUATION

RR	Result
0.0 – 0.3	Strong Protection Factor
0.4 – 0.5	Moderate Protection Factor
0.6 – 0.9	Low protection factor
1.0 - 1.1	Indifferent Factor
1.2 – 1.6	Low Risk
1.7 – 2.5	Moderate Risk
>2.5	High Risk

Interpreting the RR according CI

- **For RR with a value greater than 1 and CI with values close to the calculated RR that does not include the value 1 we can decide that there is a positive association between the risk factor and the result**
- **For RR values higher than 1 but CI that includes the value 1 it can be concluded that the studied risk factor is indifferent (no matter how high its calculated value)**

Interpreting the RR according CI

- **For RR with a value less than 1 and CI with values close to the calculated RR that does not include the value 1 we can decide that there is a negative association between the risk factor and the result**
- **For RR values less than 1 but CI that includes the value 1 it can be concluded that the studied risk factor is indifferent (no matter how high its calculated value)**

Advantages of Cohort Study

1. CS are the design of choice for studying the causes of a condition, the course of the disease, or the risk factors because they are longitudinal and follow a group of subjects over a period of time
2. Allow the direct measure of incidence
3. They possess the correct time sequence to provide strong evidence for possible causes and effects.
4. In well- designed CS, investigators can control many sources of bias related to patient selection and recorded measurements
5. Can be identified different effects of one exposure
6. Very useful in rare expose – cohort type 2.

Disadvantages of Cohort Study

1. Many years are need for this type of study, long duration, usually not less than 5 years.
2. This make such studies costly, expensive.
3. They make it difficult for investigators to argue causation because other events occurring in the intervening period may have affected the outcome.
4. Diagnostic criteria and the definition of the disease may change over time

Disadvantages of Cohort Study

- The CS is especially vulnerable to problems associated with patient follow-up, particularly patient attrition (patients stop participating in the study) and patient migration (patient move to other communities)
- Can not be used for rare disease
- When is retrospective, need a qualitative records.

4. The Case – Control Studies (CCS)

- **CCS begin with the absence or presence of an outcome (result, disease) and then look backward in time to try to detect possible causes (risk factors, expose)**
- **The *cases* in CCS are individuals selected on the basis of some disease or outcome**
- **The *controls* are individuals without disease or outcome.**
- **CCS is characterize as studies that ask "What happened ?"**
- **They are called also retrospective studies because of the direction of the inquiry**
- **CCS are longitudinal because inquiry covers a period of time**

The case – control studies

- **Direction – backward in time**
- **Temporary Sequence - retrospective**

The case-control study start with

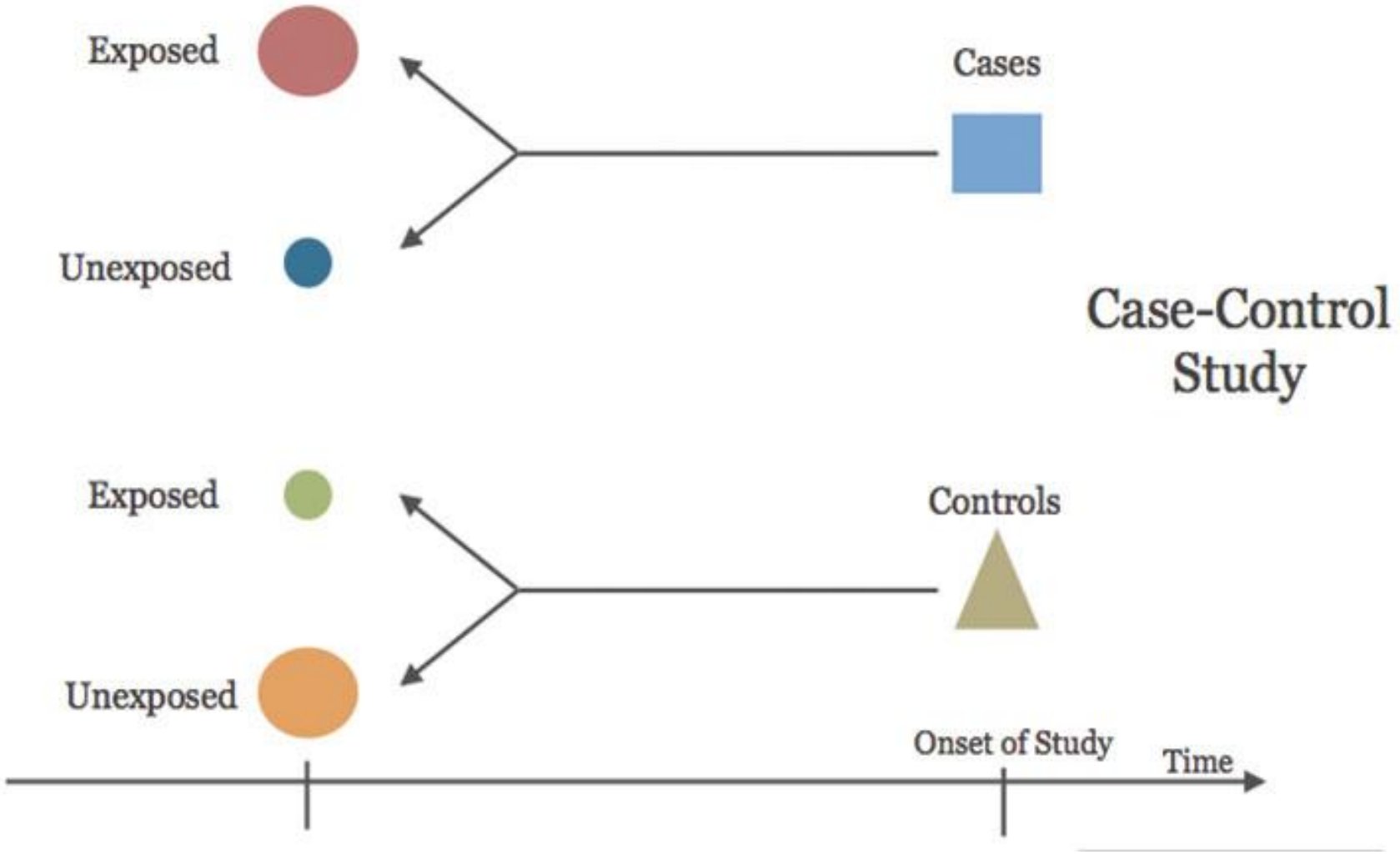
- **HYPOTHESIS**

- **Ho – null hypothesis**

The proportion of cases exposed to the studied potential risk factor is equal to the proportion of non-exposed control persons

- **Ha - Alternative hypothesis**

Diagram of Case - Control Study Design



Traditional versus modern

- **Traditional vision**

A case control study as an alternative of cohort study

- **Modern vision**

A case-control study as a part or result of cohort study

The methodology of case-control studies

- **Definition of disease**
- **Definition of "case"**

Based on symptoms or syndromes, results of laboratory and functional examinations

This allows to divide the study population in two groups: cases and control

Sources:

- ❑ **Patients from hospital**
- ❑ **Patients from primary care**
- ❑ **Medical records of specific diseases**
- ❑ **Death certificates**

Requirements for “case” selection

- The cases should be representative for study population
- Wrong selection of cases give wrong results
- The criteria for defining cases make it possible to correctly select people with and without disease
- Sources must be efficient and correct
- "Cases" with incidence are more preferred than "cases" with prevalence
- Selective study may be preferred to full study if the required population can be easily found

Requirements for “control” selection

- The "cases" and "control" come from the same population and have the same probability of selection
- Control group selection is the hardest part in case-control studies
- Sources for "control" depend on sources of "cases”

Requirements for “control” selection

- The criteria for "control" must be comparable to all the criteria used to select "cases", except that "control persons" must not have the health problem studied.
- Matching of cases with controls, better is individual matching for one case – similar control or controls, according some factors – age, sex, place of residence, race etc.

Sources for “control” group

- Tax lists
- Voting lists
- Driver's license
- National Register
- Hospitals
- Death records
- Friends / relatives

Data collecting methods

- **Personal interview**
- **Telephone interview**
- **Questionnaires**
- **Statistical forms**
- **Medical records**
- **Use of data from other studies**

Calculating the Indicators in the Case - Control Study

- **In the CCS we don't know the incidence in the exposed group or the incidence in the nonexposed group because we start with disease people (case) and nondisease people (controls).**
- **In a CCS we can not calculate the Relative Risk directly**
- **We have another measure of association, called *Odss ratio,* with which we estimate the relative risk**

The contingency 2x2 table

	WITH RESULTS	WITOUT RESULTS	TOTAL
EXPOSED	a	b	m_1
NON-EXPOSED	c	d	m_0
TOTAL	n_1	n_0	t

Design of Case – Control Studies

	Case With disease	Controls Without disease
Exposed	a	b
Not exposed	c	d
Total	a+c	b+d
Proportion Exposed	$\frac{a}{a+c}$	$\frac{b}{b+d}$

Indicators

- The odds to have disease in exposed group

[a / (a + b)] : [b / (a + b)] = a / b

- The odds to have disease in non-exposed group

[c / (c + d)] : [d / (c + d)] = c /d

- **The odds ratio**

a / b : c /d = a*d / b*c

Indicators

- **The odds of an event can be defined as the ratio of the number of ways the event can occur to the number of the event cannot occur**
- **The odds that a case was exposed**

$$a/(a+c) : c/(a+c) = a/c$$

- **The odds that a control was exposed**

$$b/(b+d) : d/(b+d) = b/d$$

The Odds Ratio

- In CCS the *odds ratio* is defined as the ratio of the odds that the cases were exposed to the odds that the controls were exposed
- $\frac{a}{c} : \frac{b}{d}$ or $\frac{a*d}{b*c}$

The formula expresses how many times the chance of exposure of cases was higher than the chance of exposure of controls

RESULT EVALUATION

OR	Result
0.0 – 0.3	Strong Protection Factor
0.4 – 0.5	Moderate Protection Factor
0.6 – 0.9	Low protection factor
1.0 - 1.1	Indifferent Factor
1.2 – 1.6	Low Risk
1.7 – 2.5	Moderate Risk
>2.5	High Risk

Other indicators

CI -confidence interval, for OR

AR -attributable risk

SA -strength of association

DETERMINING OF CI

- $CI = OR^{(1 \pm z / x)}$

$$x^2 = \frac{(t-1)\,[(a \times d) - (b \times c)]^2}{n_1 \times n_0 \times m_1 \times m_0}$$

- For 95% veracity , probability, z = 1,96

$$CI_{sup.\ lim.} = OR^{(1 + z / x)}$$
$$CI_{inf.\ lim.} = OR^{(1 - z / x)}$$

Attributable risk (AR)
Attributable risk in population

- **AR = [(OR - 1) / OR]**

- **ARP = Po (OR -1) / P (OR - 1) + 1**

where :

- **Po - prevalence of E in controls (control group)**
- **P - prevalence of E in the general population**

Attributable risk percent

- **The effect of a factor in the exposed group**
- **What proportion of the disease in the exposed group is due to exposure?**

- **AF (AR%) = [(OR-1) / OR] x 100%**

Interpreting the Odss Ratio

- If OR = 1 Risk in exposed equal to risk in non-exposed (no association)

- If OR >1 Risk in exposed is greater than risk in non-exposed and CI does not include value 1 there is a positive association, possibly causal, harmful factor)
- If OR < 1 Risk in exposed is less than in non-exposed and CI does not include value 1, there is a negative association, possibly protective factor)

Advantages of Case-Control studies

- Are especially appropriate for studying rare disease or events
- Are generally the quickest and less expensive studies
- Can identify various causes of a disease
- Test hypothesis

Disadvantages of CCS

- From all study methods, they have the largest number of possible biases or errors, and they depend completely on high - quality existing records
- Data availability for CCS sometimes requires compromises between what researches wish to study and what they a able to study
- One of the greatest problem in a CCS is selection of an appropriate group.

- **Thank you!!!**