Overview of study designs

Dr Munir Abu-Helalah
MD,MPH,PHD

Associate Professor of Epidemiology and Preventive Medicine
Consultant Clinical Preventive Medicine

Important notes are in red boxes.

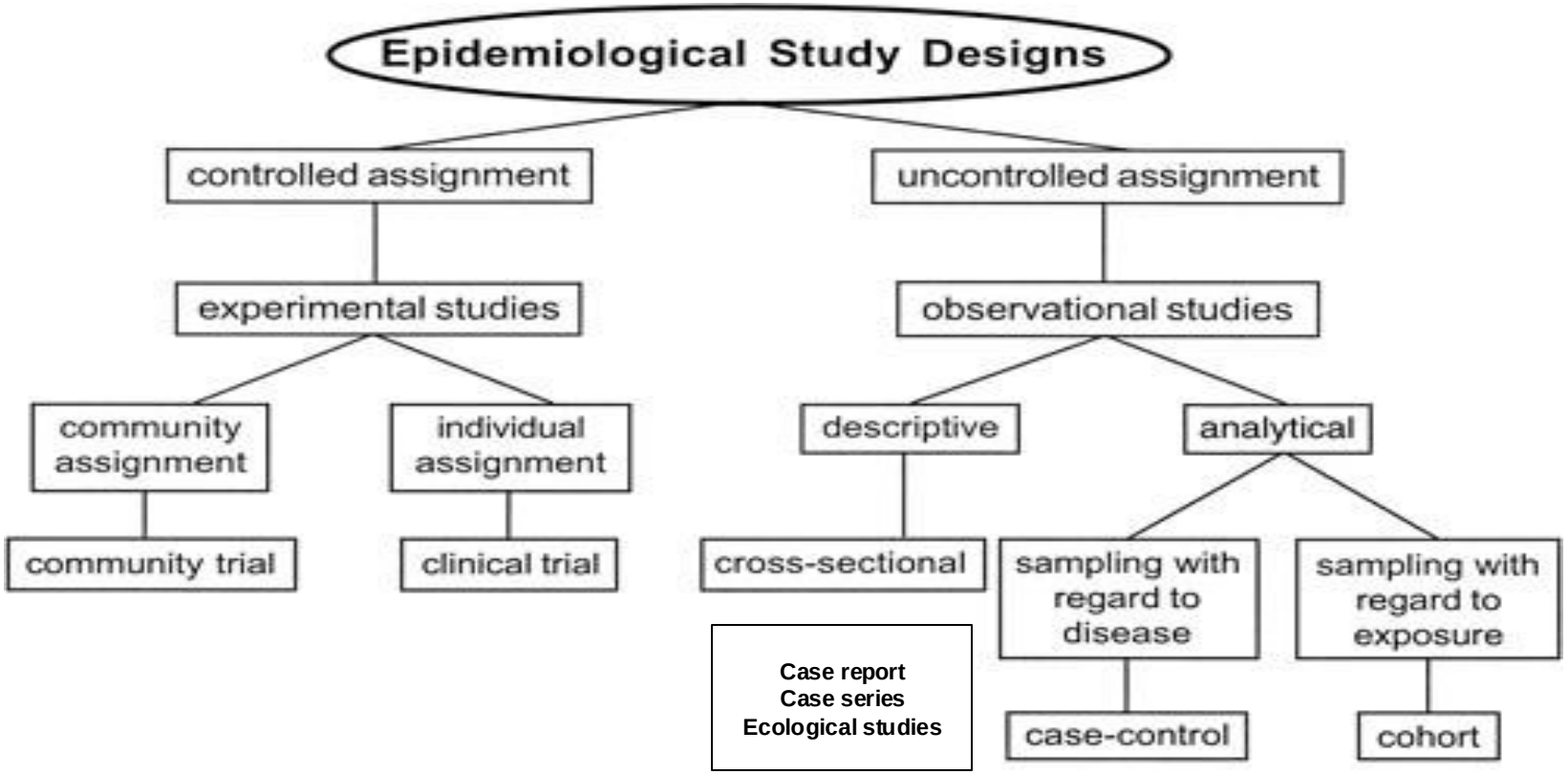
- **Understand when to use each type of study design.**
- **Know how to differentiate between different types of studies.**

Part 1

Descriptive studies

Study design: Definition

A study design is a specific plan or protocol for conducting the study, which allows the investigator to translate the conceptual hypothesis into an operational one.



Source: Waning B, Montagne M: *Pharmacoepidemiology: Principles and Practice*: http://www.accesspharmacy.com

Observational epidemiology

a. Descriptive

Case reports and case series

Descriptive analysis (Person place time)

Ecological (correlational)

Cross-sectional

b. Analytical

Case Control

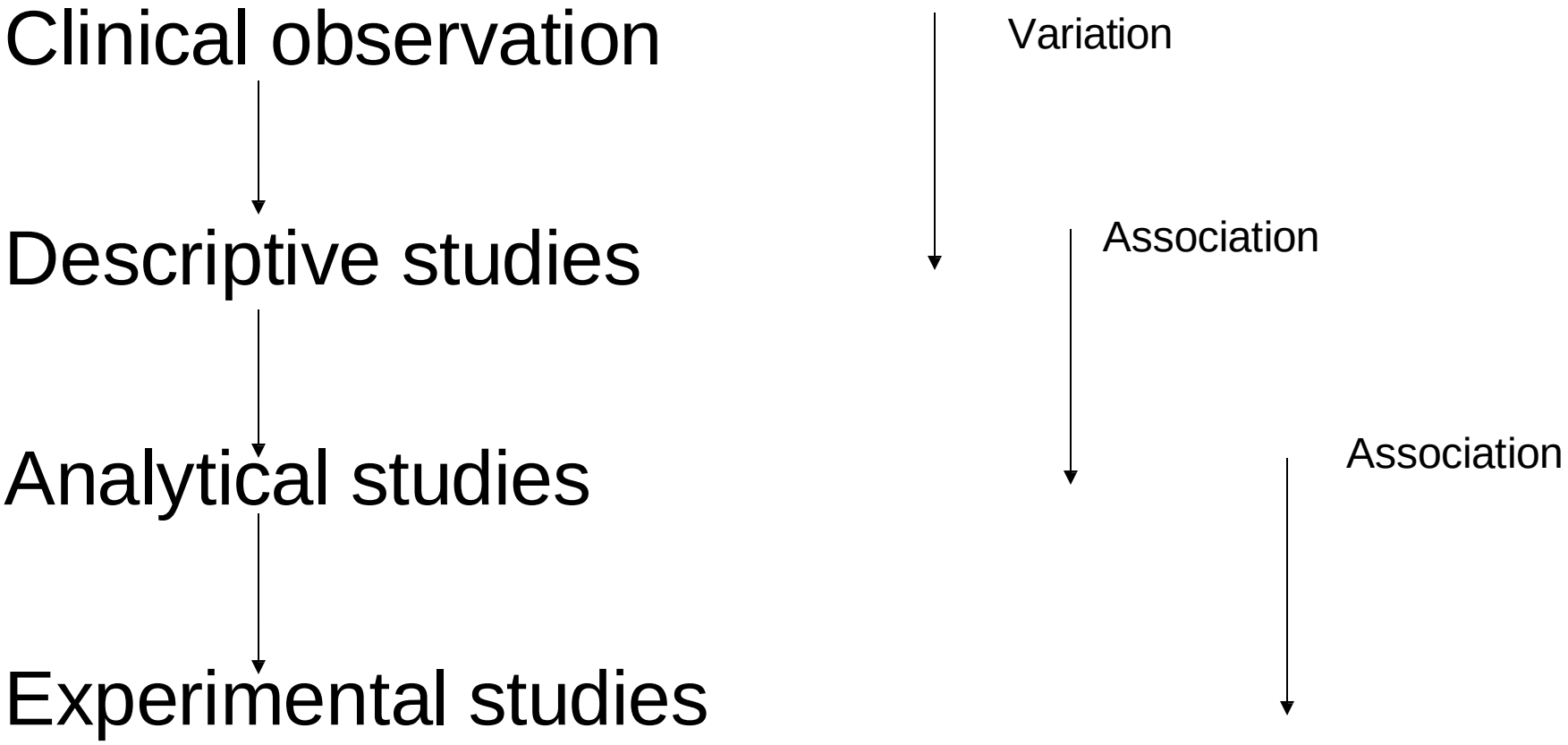
Cohort

Observational epidemiology

- Descriptive studies: provide insight, data, and information about the course or patterns of disease or drug use problems in a population or group.

- Analytical studies are used to test cause–effect relationships, and they usually rely on the generation of new data.

Epidemiological studies



Prospective vs. retrospective studies

Comparison of Retrospective and Prospective Approaches	
Retrospective	**Prospective**
Inexpensive to conduct	Expensive to conduct
Completed in a shorter time period	Completed over a longer time period
Easier to access a larger number of subjects	More difficult to access subjects and usually requires a larger number of subjects
Allows results to be obtained more quickly	Exposure status and diagnostic methods for disease may change
Useful for studying exposures that no longer occur	Loss of subjects from the study over time may be substantial
Information and data may be less complete and inaccurate	Information and data may be more complete and accurate
Subjects may not remember past information	Direct access to study subjects enhances reliability of data

Case Reports and Case Series

Case report is detailed report by one or more clinicians of the profile of a single patient.

Example: 1961; pulmonary embolism 5 weeks after use on oral contraceptive.

Question: Are women who develop pulmonary embolism more likely to have used oral contraceptives than women who did not develop the disease?

Case Series describes the characteristics of a number of patients with a given disease.

Application: Routine surveillance activities (accumulated case reports). Striking clustering of cases may suggest emergence of new diseases or epidemics

Example: 5 Previously healthy homosexual men were diagnosed to have Pneumocystis carinii pneumonia at three Los Angeles hospitals during a six month period (1980-1981).

Case report and case series

- **Clinician finds unusual features of a disease or effects of a drug, or the patient's medical history, that lead to the formulation of a new research question or hypothesis**

Case-series:
Clinical case series

- Usually a coherent and consecutive set of cases of a disease (or similar problem) which derive from either the practice of one or more health care professionals or a defined health care setting, e.g. a hospital or family practice.

Case series: Limitations

Usually we cannot estimate the prevalence or incidence rate

- **Breast cancer registry in Jordan: We cannot provide incidence or prevalence rates without:**

1. **Population size**
2. **Time- period of data collection**
3. **All cases of breast cancer are registered**

<u>No control group for comparison</u>

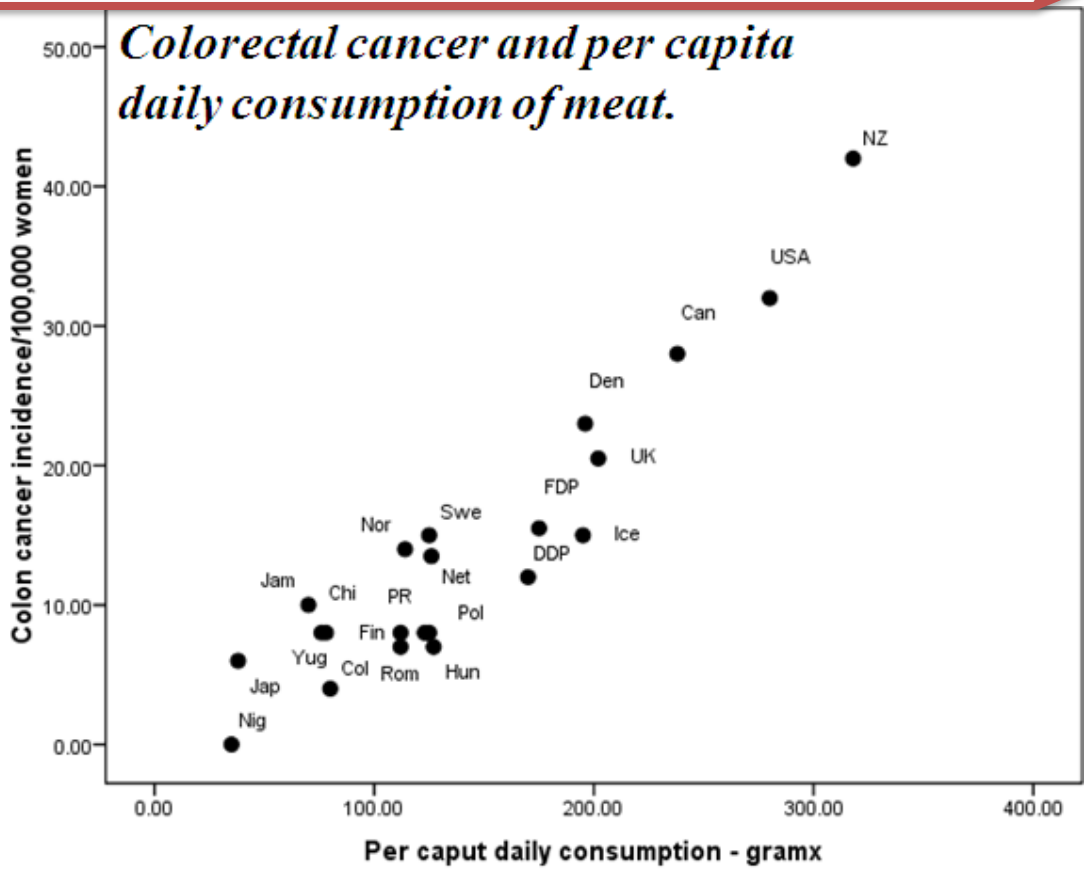
Ecological studies

Are studies in which information on the characteristics and/or exposures of individual members of the population groups are generally not obtained. Existing statistics are used to compare the mortality or morbidity experience of one or more populations with some overall index exposure. care is needed to avoid the 'ecological fallacy' where inappropriate conclusions are made from ecologic data

Ecological studies

- In ecological studies the unit of analysis is some aggregate individuals rather than individual persons
- Geographic areas or time period are often used as a basis for defining aggregates
- The analysis centers on determining whether the ecological units with a high frequency of exposure are also unit with a high frequency of disease (+ve correlation) or a low frequency of disease (- ive correlation)

What type of study does this figure represents?



Adapted from: *Int. J. Cancer 15:617, 1973*

Ecological (Correlational studies)

Disadvantages:

1. It is unable to control for confounding factors. This is often referred to as 'ecological fallacy', where two variables seem to be correlated but their relationship is in fact affected by cofounding factor(s).
2. It cannot link exposure with disease in individuals as those with disease may not be expose.
3. Its use of average exposure levels masks more complicated relationships with disease.
4. Its units of study are populations not individuals. Therefore, the disease rates linked with population characteristics and the association observed at group level does not reflect association at individual level.

CROSS-SECTIONAL STUDY DESIGN

- **Sometimes called *prevalence studies*.**
- **They are studies of total populations or population groups in which information is collected about the present and past characteristics, behaviors, or experiences of individuals.**
- **There are a number of advantages in performing a cross-sectional study. These studies involve a single data collection and, thus, are less expensive and more expedient to conduct.**

CROSS-SECTIONAL STUDY DESIGN

- **Emphasis is on differences between groups at one point in time.**
- **They provide a one-time glimpse at the study population, showing the relative distribution of conditions, diseases, and injuries—and their attributes—in a group or population.**

Cross-sectional (or prevalence) studies

Are studies in which a defined population is surveyed and their disease or exposure status determined at one point in time

- **The prevalence rates of disease in the whole population as well as in those with and without the exposure under investigation can be determined**
- **Cross-sectional studies are generally not suitable for a disease which is rare or of short duration as few people will have the disease at any one point in time**

Cross-sectional studies

- More effective in identifying chronic diseases and problems
- Less effective in identifying communicable diseases of short incubation periods and short durations.

Cross-sectional (or prevalence) studies

- **It is often difficult to separate cause and effect as the measurement of exposure and disease at any one point in time**

- **Because of this limitation, cross-sectional studies are useful when investigating exposures which do not change e.g genetic characteristics such as ABO blood group and HLA**

- **Cross-sectional studies are often used as an initial exploration of a hypothesis prior to conducting a case-control or follow-up study**

Cross-sectional study

Chemotherapy	**Outcome**		
	With pain	Without pain	Total
Yes	664	556	1220
No	879	1088	1967
Total	1543	1644	3187

Prevalence of pain among chemotherapy = 664/ 1220 = 54.4%

Prevalence of pain among no chemotherapy = 879 / 1967 = 44.7%

Prevalence Rate Ratio (PRR) = = 54.4 / 44.7 = 1.22

Cross-sectional studies

- Seasonal variations of disease are not well represented in cross-sectional studies except if the duration of the study allows such comparison
- In the example below, studying RTA in October would not provide a valid result for incidence of RTA in whole year and does not allow identifying seasonal variations in the RTA
- Road traffic accidents by month of accident, Slovenia, average 2003-2006

Cross-sectional studies: advantages

- **Quick**
- **Many associations can be studies**
- **Data on all variables is only collected once.**
- **Sample size depends on the question**
- **Standard measures used**
- **Prevalence estimated**
- **The prevalence of disease or other health related characteristics are important in public health for assessing the burden of disease in a specified population and in planning and allocating health resources.**
- **Good for descriptive analyses and for generating hypotheses**

Surveys

A survey may be defined as a collection of information from all individuals or a sample of individuals chosen to be representative of the population from which the are drawn

Understanding when to use surveys is enough.

Types of information collected by surveys

- Morbidity prevalence
- Mortality
- Detailed risk factors or behavioral information
- Knowledge, attitudes, and practices
- Physical signs (paralysis, splenomegaly, malnutrition)
- Serological or laboratory tests

Characteristics of survey

- representative if sample chosen correctly
- Single point in time –snapshot
- Provide more in depth information than surveillance or chart reviews
- Usually performed by a limited number of personnel

  specially trained to perform surveys
- Can sometimes be expensive, time consuming to perform
- Cannot be used to monitor change unless repeated at a later time therefore may be better for situational analysis than

  for ongoing monitoring of a problem or a programme

Survey

Key Concepts of survey design:

1. Primary data
2. Communication
3. Sample
4. Representative

TYPE OF MEASUREMENT

- *Attitudes*: What people feel
- *Knowledge*: What people know
- *Beliefs*: What people think is true: their beliefs
- *Behaviours*: What people do or have done
- *Evaluation*: Peoples perception of thing are/were

Case-control studies

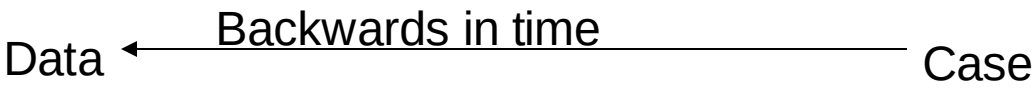
Are studies in which a group of people with a particular disease (the cases) are compared with a group of people without the disease (the controls). The purpose of the comparison is to determine whether, in the past, the cases have been exposed more (or less) often to a specific factor than the controls

- This type of study is done to identify factors that could be responsible for the development of a disease or drug use problem.

CASE-CONTROL STUDIES

- The direction of time

- Cases identified now
- Data on past events collected



CASE-CONTROL STUDY DESIGN

- Designed to assess association between disease occurrence and exposures (e.g., causative agents, risk factors) suspected of causing or preventing the disease.

- **Rare disease with commonest factors, cases-control.**
- **Rarest factors, you use cohort.**

Case-control studies

- A group of people with a disease are compared to a group without the disease from the same population.
- Compare exposure to risk factors in both groups
- Able to look at many different possible risk factors
- Able to study diseases with a long latency period
- Most common analytic study design seen in the medical literature today

Case-control studies

- In general, the cases included in a case-control study include people with **one** specific disease only
- But, a case-control study can provide information on a wide range of possible **exposures** that could be associated with that particular disease
- Useful for the study of rare diseases
- Not suitable for the study of rare exposure
- Relatively small and inexpensive
- Takes a relatively short time to complete
- Can test current hypotheses
- Cannot measure disease incidence

CASE-CONTROL STUDIES

- Cases have the disease of interest

Eg. Cerebral palsy

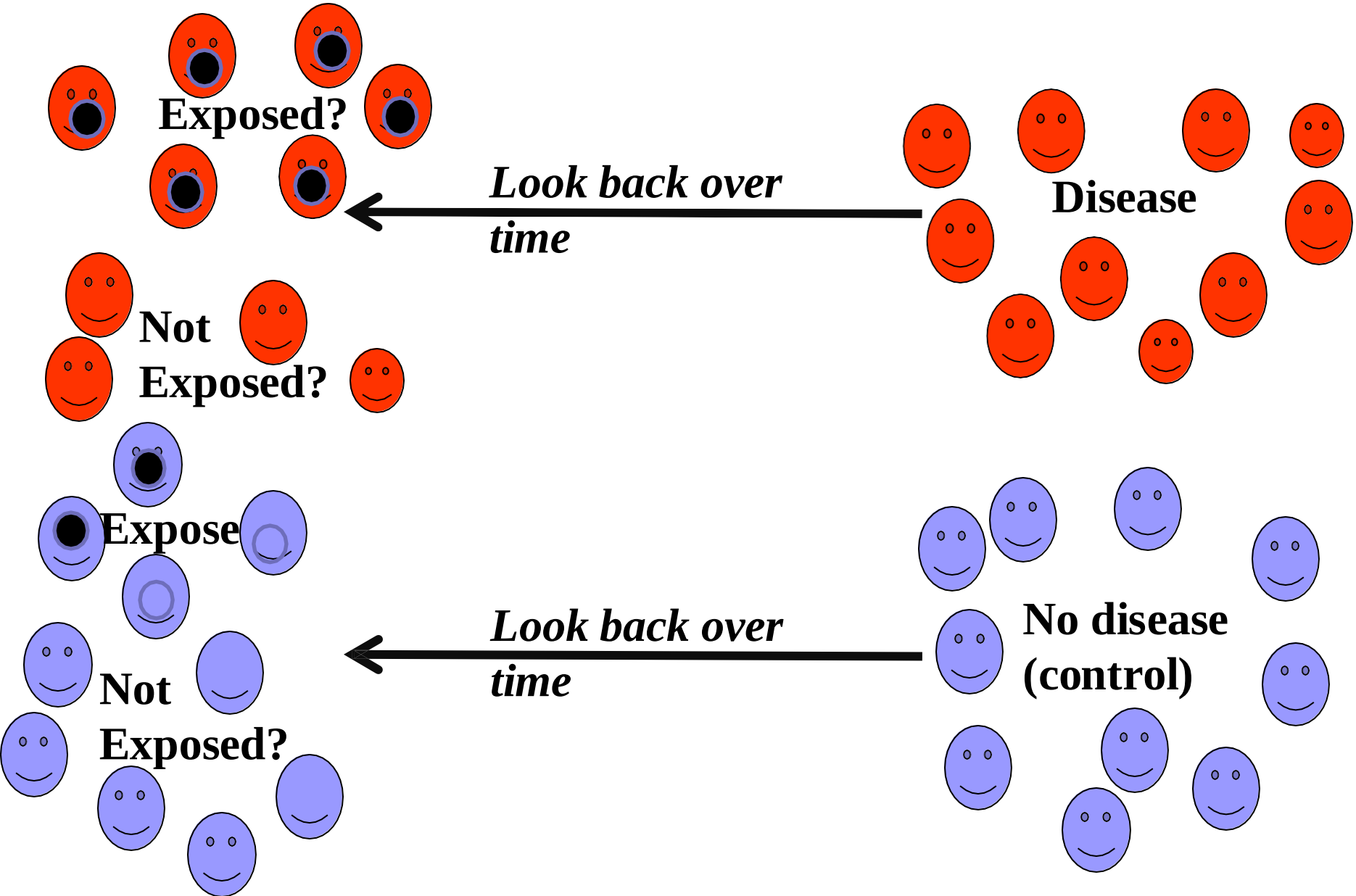
- Controls do not have the disease

Eg. Healthy babies born at the same time

Case-control study: challenges

- Selecting cases
  - Eligibility
- Selecting controls
  - Representativeness
- Exposure assessment
  - Accurate

Case control studies



CASE-CONTROL STUDIES

Strengths

- Suited to study disease with long latency periods, but can be used in outbreaks investigations
- Optimal for rare diseases
- Efficient in terms of time and costs: relatively quick and inexpensive
- Allows for evaluation of a wide range of possible causative factors that might relate to the disease being studied
- Odds ratio estimated

CASE-CONTROL STUDIES

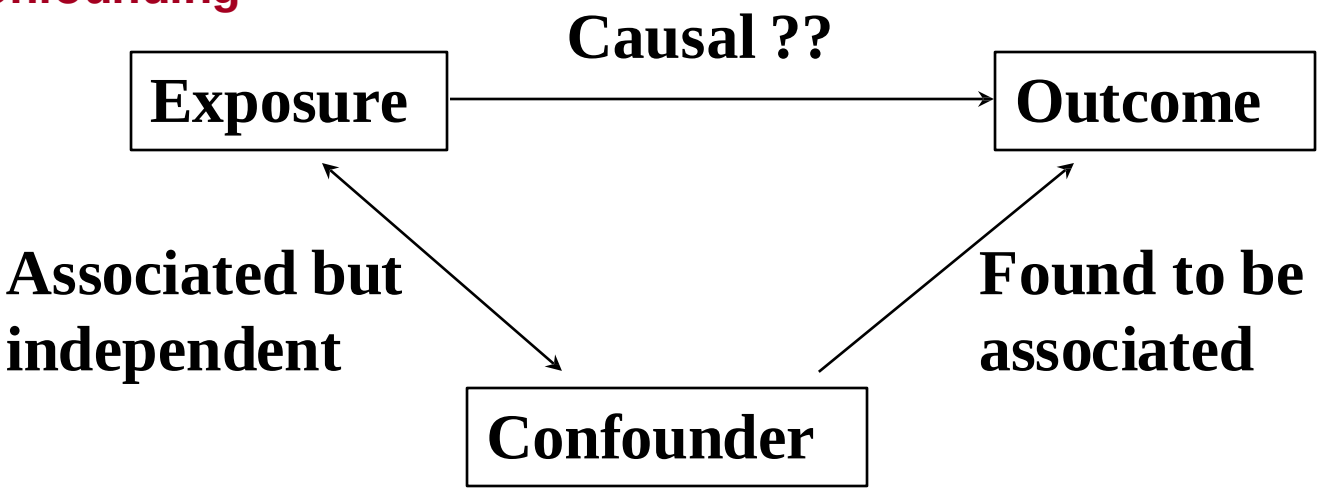
Limitations

- Very susceptible to bias (especially selection and recall bias) as both the disease and the exposure have already occurred when participants enter the study. Cases and controls might not be representative of the whole population
- We cannot calculate incidence or prevalence rate of disease
- We cannot be certain that exposure came before disease
- Choice of controls difficult
- Controls do not usually represent non-exposed population
- Past records incomplete
- No absolute risk estimates

Confounding

A confounding factor is one that is associated with the exposure and that independently affects the risk of developing the outcome, but that is not an intermediate link in the causal chain between the exposure and the outcome under study

Matching - often used in case-control studies to decrease confounding



Cohort studies

Cohort (or follow-up) studies

- **Are studies in which people are identified and grouped with respect to whether or not they have been exposed to a specific factor.**

- **The groups are followed up over time to determine whether the incidence of a particular disease is any greater (or less) in the exposed group than in the non-exposed group.**

Cohort study examples:

- Life expectancy of cerebral palsy children
- Fine needle breast biopsy and breast cancer
- Aspirin intake and colorectal cancer

Cohort study: Primary purposes

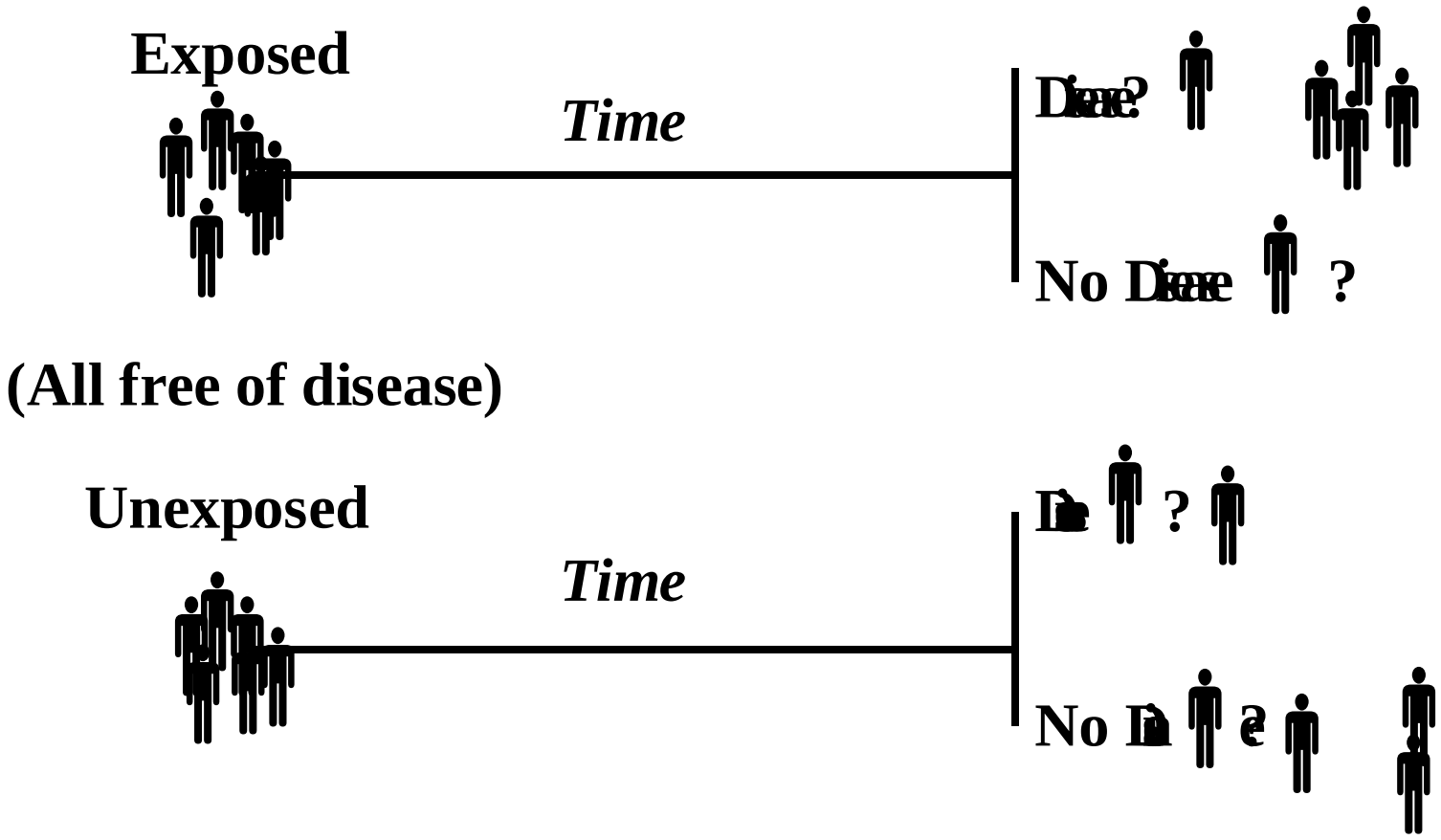
- **Descriptive (measures of frequency)**

– To describe the incidence rates of an outcome over time, or to describe the natural history of disease

- **Analytic (measures of association)**

– To analyze associations between the rates of the outcomes and risk factors or predictive factors

Cohort studies



COHORT STUDY DESIGN

- This design is the best observational one for establishing cause–effect relationships. Prevention and intervention measures can be tested and affirmed or rejected.
- Cohort studies take into account seasonal variation, fluctuations, or other changes over a longer period.
- Objective measures of exposure, such as biological markers, are preferred over subjective measures.

COHORT STUDY DESIGN
Strengths

- We can measure incidence of disease in exposed and unexposed groups
- Can get a temporal (time related) sequence between exposure and outcome as all individuals must be free of disease at the beginning of the study.
- Good for looking at effects of rare exposures.
- Allows for examination of multiple effects of a single exposure.
- Not open to bias as much as other types of study
- Direct calculation of the risk ratio or relative risk is possible.
- Provide information on multiple exposures

COHORT STUDY DESIGN

Limitations:

- Not efficient for rare diseases
- Can be expensive and time-cosuming
- Large sample
- Drop-out biases

  If study goes over many years, can get considerable loss to follow up. This can 'dilute' results or lead to bias, and therefore the validity of result can be seriously affected

- Locating subjects, developing tracking systems, and setting up examination and testing processes can be difficult.
- Changes over time in diagnostic methods, exposures, or study population may lead to biased results.

Cohort study: Example

Hypertension as a risk factor for spontaneous intracerebral hemorrhage

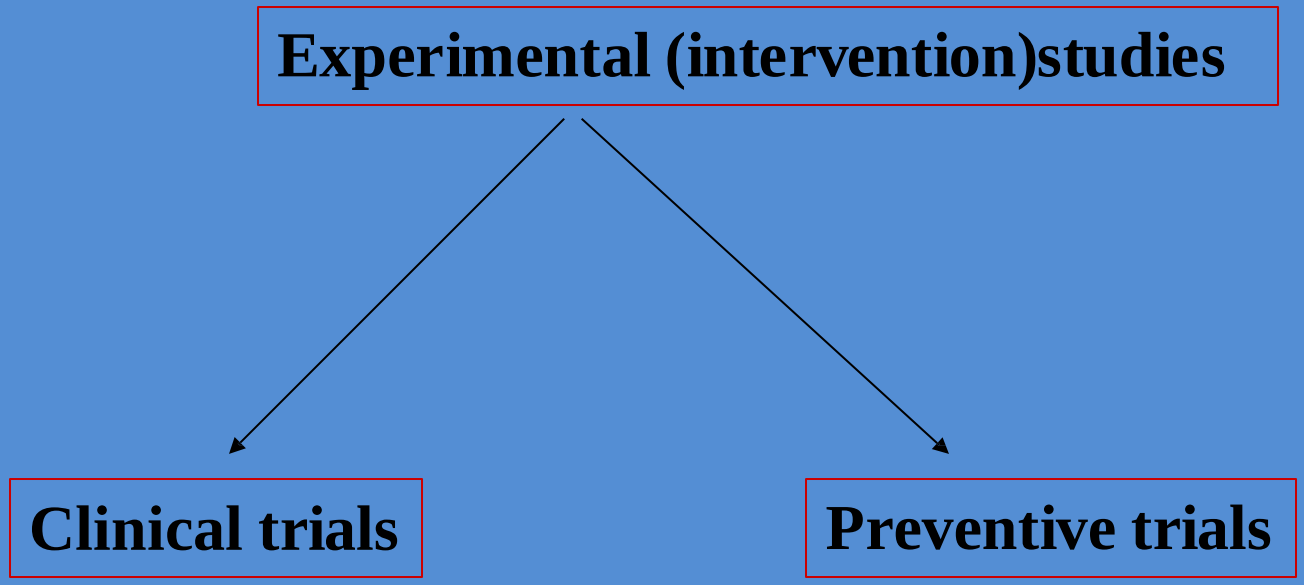
COHORT STUDY DESIGN: Summary

- In general, can investigate the effect of only a limited number of exposure
- Useful for investigating a range of outcomes associated with only one exposure
- Useful for study of rare exposure
- Not suitable for the study of rare diseases
- Follow-up studies are often large and expensive
- May take many years to complete
- Cannot test current hypotheses
- Can measure disease incidence

Experimental Study Design

A study in which a population is selected for a planned trial of a regimen, whose effects are measured by comparing the outcome of the regimen in the experimental group versus the outcome of another regimen in the control group.

Experimental studies (Intervention)



Experimental Study Design

Different from observational designs by the fact that there is manipulation of the study factor (exposure), and randomization (random allocation) of subjects to treatment (exposure) groups.

Why experimental study design?

- Limitations of theory
- Previous disasters

Clofibrate:

Successfully lowers cholesterol

Treated group: reduced CHD incidence, but higher all causes mortality

- Spontaneous improvements
- Importance of small effects

Clinical trials

- **Individuals with particular disease are randomly allocated into experimental or control groups. randomization is used to ensure that both groups are comparable with respect to all other factors except for the one under investigation.**

- **The experimental group is given the agent being tested and the control group is given either an agent in current use or a placebo**

- **Ideally both patients and the observers should be 'blind' to the treatment being given. This in order to reduce bias.**

Clinical trials

- **Are studies of the effect of a specific treatment on patients who already have a particular disease**

- **They are used to evaluate the efficacy of a preventive or therapeutic agent in the treatment or prevention of a disease**

- **“The most definitive tool for evaluation of the applicability of clinical research” - 1979 NIH release.**

Clinical trials

- **Assessment of each subject must involve bias free accurate measure of outcome**

- **Both groups are followed over a defined period of time when the outcome is then measured in both groups.**

What trials assess

- Drugs
- Surgery
- Type of management
- New services

Purpose of Control Group

- To allow discrimination of patient outcomes caused by test treatment from those caused by other factors
  - **Natural progression of disease**
  - **Observer/patient expectations**
  - **Other treatment**
- Fair comparisons
  - **Necessary to be informative**

Randomized allocation

- Like tossing a coin
- Avoids choosing
- Permits fair comparison

Types of outcomes

- Death
- Clinical measurement
- Symptoms
- Quality of life
- Psychological wellbeing

The need for blinding

- Open
- Single blind
- Double blind
- Triple blind

Definitions

- **Single Blind Study**: A clinical trial where the participant does not know the identity of the treatment received
- **Double Blind Study**: A clinical trial in which neither the patient nor the treating investigators know the identity of the treatment being administered.
- **Triple Blind study: Biostatisticians is also blinded**

Definitions

- **Placebo:**
  - Used as a control treatment

    1. An inert substance made up to physically resemble a treatment being investigated

    2. Best standard of care if “placebo” unethical

    3. “Sham control”: Faked surgical intervention with the patient's perception of having had a regular operation

Definitions

- **Adverse event:**
  - An incident in which harm resulted to a person receiving health care.
  - **Examples:** Death, irreversible damage to liver, nausea
  - Not always easy to specify in advance because many variables will be measured
  - May be *known* adverse effects from earlier trials

Adverse Events

- Challenges
  - Long term follow-up versus early benefit
  - Rare AEs may be seen only with very large numbers of exposed patients and/or long term follow-up
- Example – COX II inhibitors
  - Vioxx & Celebrex
  - Immediate pain reduction versus longer term increase in cardiovascular risk

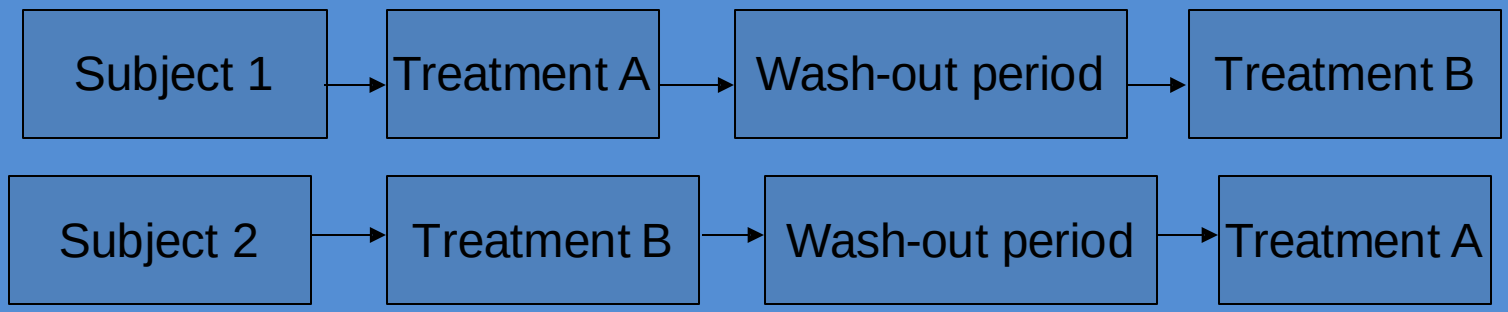
Cross-over clinical trial

Each patient gets both treatments

Half get A then B

Half get B then A

Wash-out period in between



Cross-over clinical trial

- Cross-over design
- Patient as own control

-Reduce variations

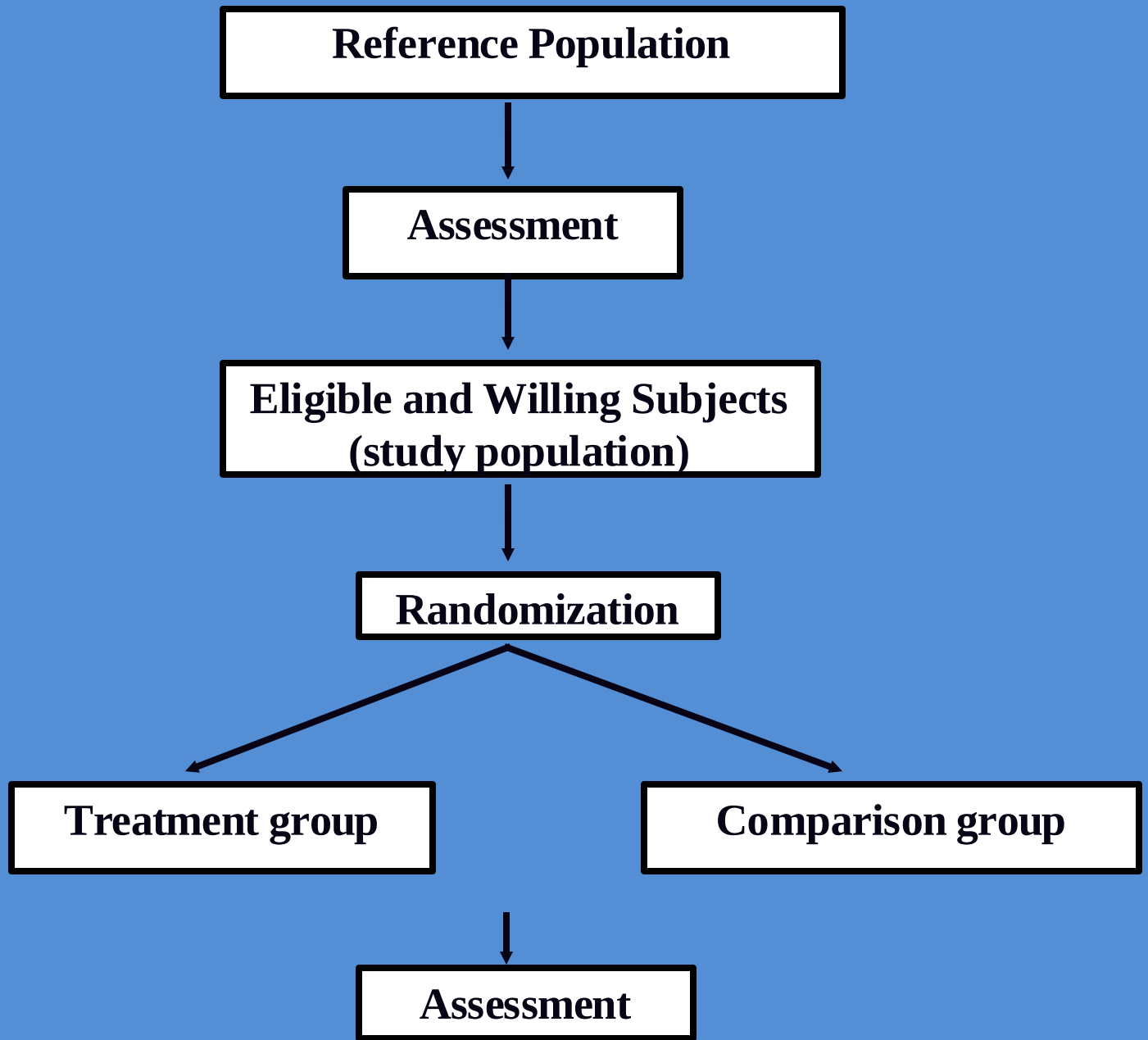
-Much smaller sample size

Requirements: Carry over period(s)

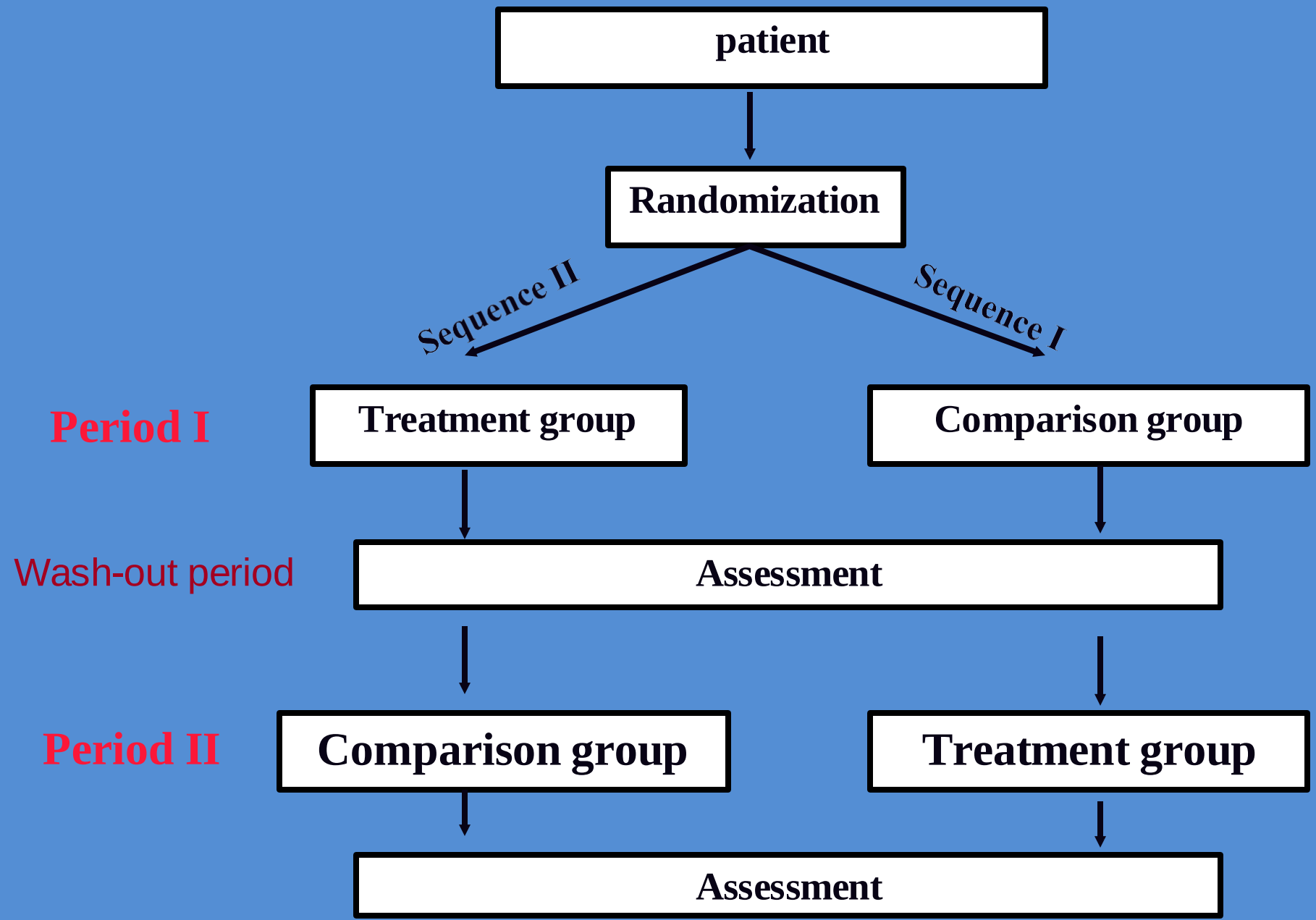
Key elements of RCTs

- Selection of subjects
- Comparison group
- Randomization
- Allocation of treatment
- Blinding (single, Double blind design/placebo)
- Intention to treat analysis in which the treatment and control groups are analyzed with respect to their random allocation, regardless of what happened subsequently
- Ethical considerations

Parallel Design



Crossover Design



Preventive trials

Are studies of the effect of a possible preventive measure on people who do not yet have a particular disease. Another type of preventive trial is a study of the effect of a possible preventive measure on whole communalities.

Preventive trials

- **The risk of developing any particular disease among the people who are free from disease is small. Because of this, preventive trials usually require a greater number of subjects than clinical trials, and are therefore more expensive**
- **This expense limits their use to the study of preventatives of extremely common or extremely severe diseases**
  e.g. vaccination to prevent whooping cough
  vaccination to prevent poliomyelitis
- **When a disease occurs rarely, it is more efficient to study those people thought to be at high risk of disease , e.g. vaccine to prevent Hepatitis B**

Preventive trials

- As in clinical trials, the preventatives should be given so that the individuals who do and do not receive the preventative are as comparable as possible. This is often difficult.

- In some types of trials the preventative have to be administered to communities rather than individuals, e.g. water fluoridation to prevent dental caries

Results of a trial to determine whether A vaccine could prevent whopping cough

	No. with Whooping cough	No. without Whooping cough
Number vaccinated 3801	149(4%)	3652(96%)
Number not vaccinated 3757	687(18%)	3070(82%)

Community Trials

- A community participates in a behavioral intervention, nutritional intervention, a screening intervention, etc
- Intervention: Any program or other planned effort designed to produce changes in a target population.
- *Community* refers to a defined unit, e.g., a county, state, or school district.
- Communities are randomized and followed over time.
- Determine the potential benefit of new policies and programs.

Examples:

- A community-level intervention for tobacco control might combine a school curriculum for youth to prevent initiation of smoking
- A media campaign aimed at reducing smoking rate