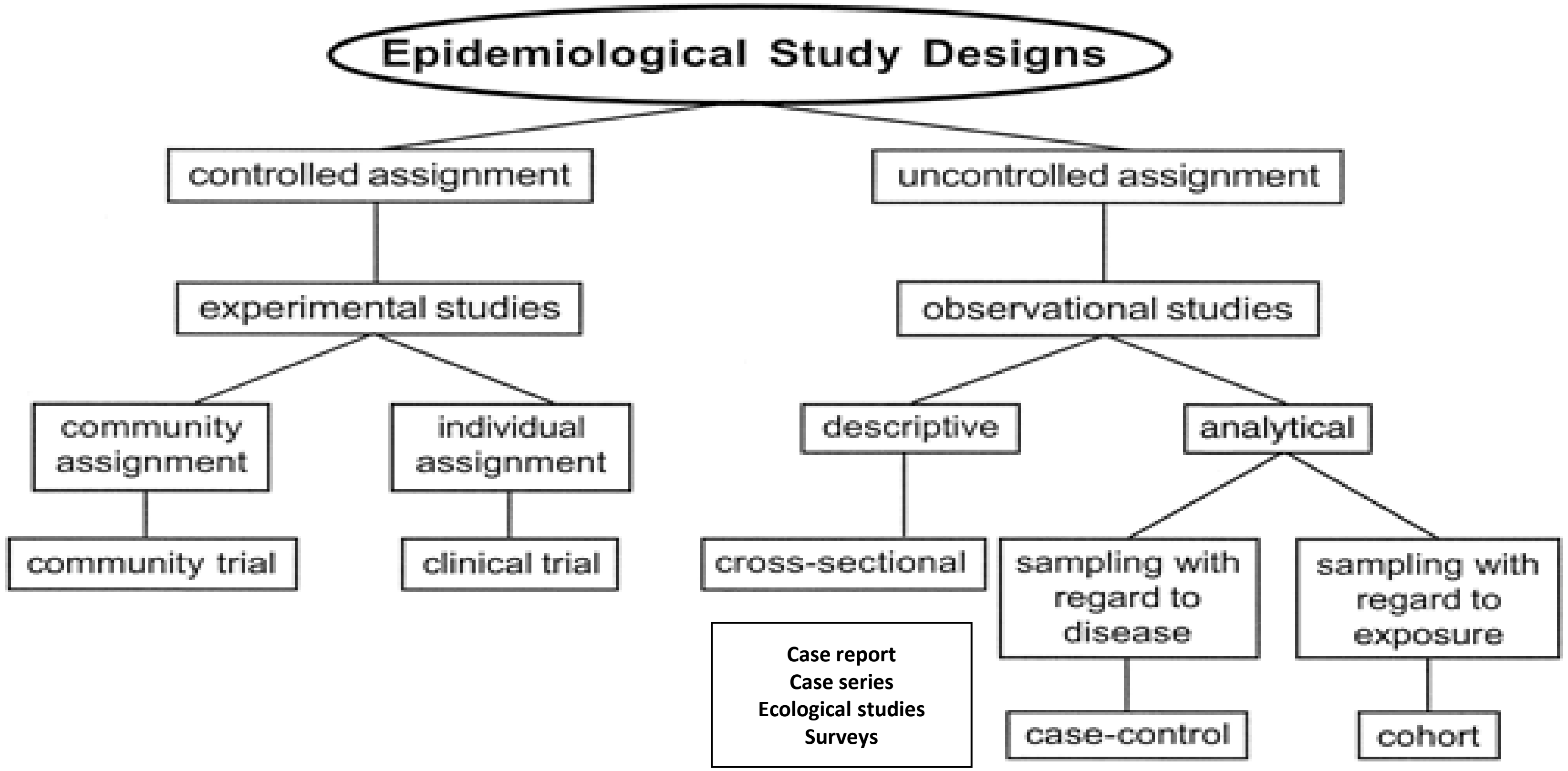
Overview of study design part 2: Analytical experimental studies

Dr Munir Abu-Helalah

MD,MPH,PHD

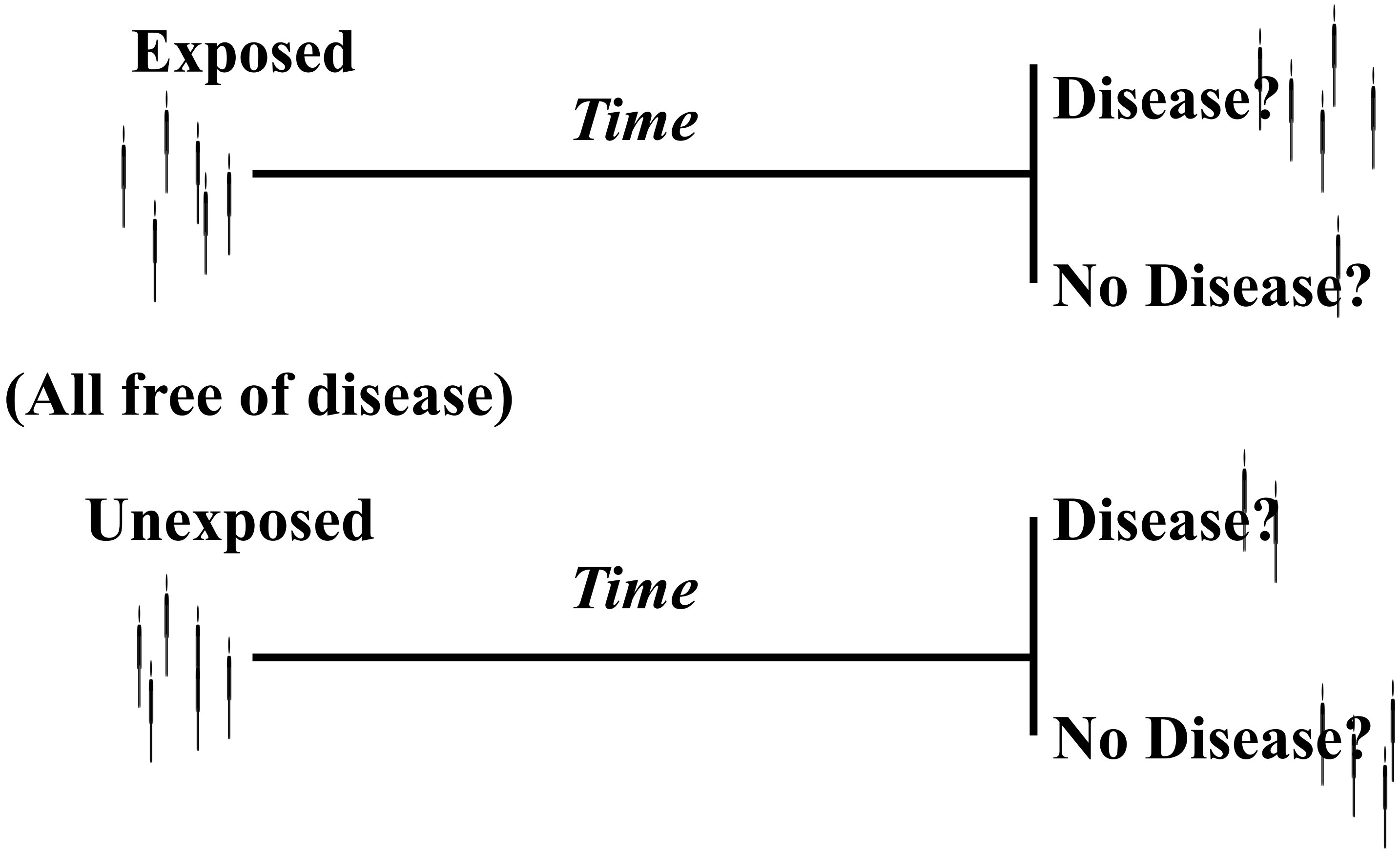
Associate Professor of Epidemiology and Preventive Medicine



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

Copyright © The McGraw-Hill Companies, Inc. All rights reserved.

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.**

- **The starting point is the risk factor!**

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 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 consider 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/diseases 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

In study risk factors, we start with what is rare!

- Rare disease: we conduct case control study starting with cases
- Rare risk factor: we conduct a cohort study starting with rare risk factors

Calculation of the relative risk

Cohort study

	Disease Present	Disease absent	
Exposure Present	a	b	a+b
Exposure absent	c	d	c+d
Total	a+c	b+d	a+b+c+d

Measuring the association between risk factor and diseases

Relative risk

$$\text{Relative Risk}\,(RR) = \frac{\text{Risk in the exposed}}{\text{Risk in the non exposed}}$$

- **RR=1**
  There is no association between exposure and disease.

- **RR>1**
  Exposure is associated with an *increase* of the frequency of the disease.

- **RR<1**
  Exposure is associated with a *decrease* of the frequency of the disease.

	Disease Present	Disease absent	
Exposure Present	a	b	a+b
Exposure absent	c	d	c+d
Total	a+c	b+d	a+b+c+d

Risk in the exposed=(a)/(a+b)

Risk in the non exposed=(c)/(c+d)

$$Relative\ Risk\ (RR) = \frac{a/(a+b)}{c/(c+d)}$$

Physical Activity and Incident Cognitive Impairment in Elderly Persons

ARCH INTERN MED/VOL 170 (NO. 2), JAN 25, 2010

Background: Data regarding the relationship between physical activity and cognitive impairment are limited and controversial. We examined whether physical activity is associated with incident cognitive impairment during follow-up.

Methods: As part of a community-based prospective cohort study in southern Bavaria, Germany, 3903 participants older than 55 years were enrolled between 2001 and 2003 and followed up for 2 years. Physical activity (classified as no activity, moderate activity [<3 times/wk], and high activity [≥3 times/wk]), cognitive function (assessed by the 6-Item Cognitive Impairment Test), and potential confounders were evaluated. The main outcome measure was incident cognitive impairment after 2 years of follow-up.

Cohort study

Physical activity	**Cognitive impairment**		Total
	Yes	No	
Moderate	10	990	1000
None	100	900	1000
Total	110	1880	2000

Risk of outcome in exposed (not active) = 100/1000 = 10%

Risk of outcome in non-exposed (active)=10/1000 =1%

Relative risk 10%/1%=10 =

Design of cohort studies

1. **Research question must be clear**
2. **Set the sample size**
3. **Set the follow-up period (immediate, short term and long term)**
4. **Specify study group Sample must be representative of the population you are studying**
5. **All participants should be free of the outcome (disease) at the beginning of the study**
6. **Must be able to get correct information about exposure status easily**
7. **Measure the outcome**
8. **Comparison group must be as similar as possible to exposed group**
9. **Put measures in place to reduce loss to follow up if possible**

COHORT STUDY DESIGN

- Measurement of exposures should be based on intensity, duration, regularity, and variability.
- Some exposures are acute, one-time episodes never repeated in a subject's lifetime.
- Other exposures are long term, such as cigarette smoking or use of oral contraceptives.
- Exposures may also be intermittent.

COHORT STUDY DESIGN
Retrospective cohorts

- Uses information on prior exposure and disease status.
- All of the events in the study have occurred and conclusions can be drawn more rapidly.
- Costs can be lower
- May be the only feasible one for studying effects from exposures that no longer occur, such as discontinued medical treatments.
- The main disadvantage of a retrospective cohort study is that the investigator must rely on existing records or subject recall.

Retrospective cohort

- Smoking and type II DM

- We start from the year 2002 and follow up for 20 years until 2022.
- In the year 2002 we split the files into: Medical notes of smokers versus medical note for non-smokers
- Both groups should not have diabetes or impaired glucose profile at baseline
- Then, we measure the incidence of Type II DM in the smoking and no-smoking groups.
- The follow up was completed in the past, therefore, we call it a retrospective cohort study.

Ambidirectional Cohort

- **Data collected both retrospectively and prospectively on the same cohort to study short and long term effect of exposure**

- **If medical notes in the previous example were incomplete in 2002 but more complete and accurate data are available since 2015.**

- **From the year 2015 until date, the follow-up is in the past, if we continue for additional 12 year. This means a combination of retrospective and prospective data.**

COHORT STUDY DESIGN
Midpoint analysis

- Occurs when, at a defined point in time in the study, all data collected to that point are analyzed so a decision can be made to stop or continue the study.

Framingham Heart Study

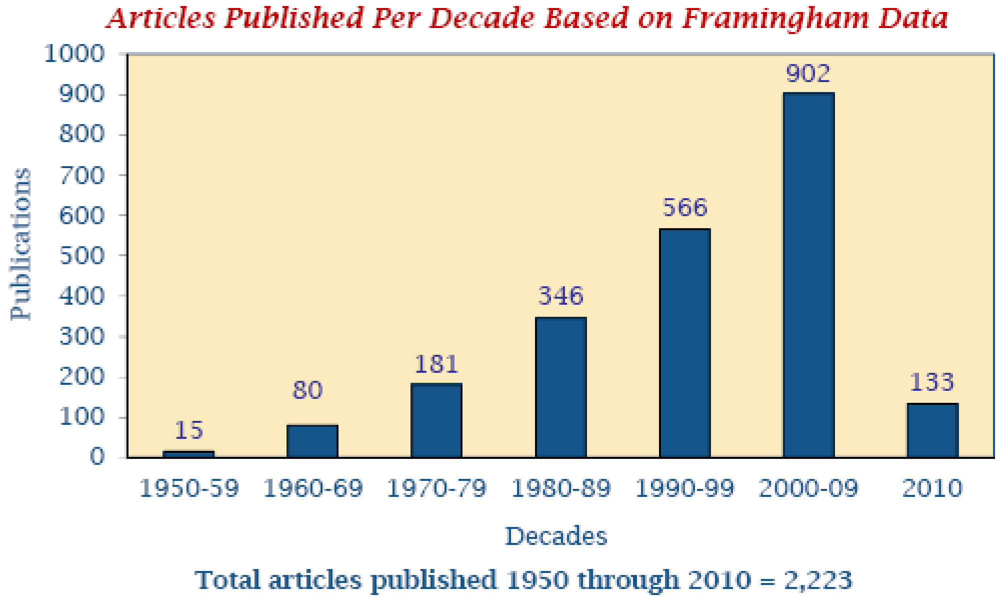
Approximately 5100 residents of this Massachusetts community are followed for > 30 years.
Selected because of a number of factors has permitted assessment of the effects of a wide variety of factors on the risk of numerous diseases

- **stable population,**
- **had a number of occupations and industries represented**
- **had a single, major hospital that was utilized by the vast majority of the population**
- **prepared annually updated population lists that would facilitate follow-up,**

Diseases studied included:

- **coronary heart disease**
- **rheumatic heart disease**
- **congestive heart failure**
- **angina pectoris**
- **intermittent claudication**
- **stroke**
- **gout**
- **gallbladder disease**
- **a number of eye conditions**

The Framingham Heart Study



http://www.framinghamheartstudy.org/risk/index.html

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
- Can measure disease incidence

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

Data $\xleftarrow{\text{Backwards in time}}$ Case

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.

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

Design of case control studies

- **Comparability:Two groups must be as similar to each other as possible so selection of controls is very important. Controls must be as similar as possible to cases – except that they do not have the outcome (disease).**

- **Outcome (disease) must be very clearly defined. (Diagnostic criteria must be clear)**

- **Use objective data about exposure status wherever possible, to reduce the risk of bias**

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

CASE-CONTROL STUDY DESIGN

- **Data Analysis**
- Data collection and analysis are based on whether the case-control study involves a matched or unmatched design. The measure used typically in case-control studies is the odds ratio.
- **Odds ratio (OR):** odds of a particular exposure among people with a specific condition divided by the corresponding odds of exposure among people without the condition under study

Odds Ratio

The word "odds" means the chances of an event to happen. The Odds of an event is the *ratio* of the event to happen over the event not to happen.

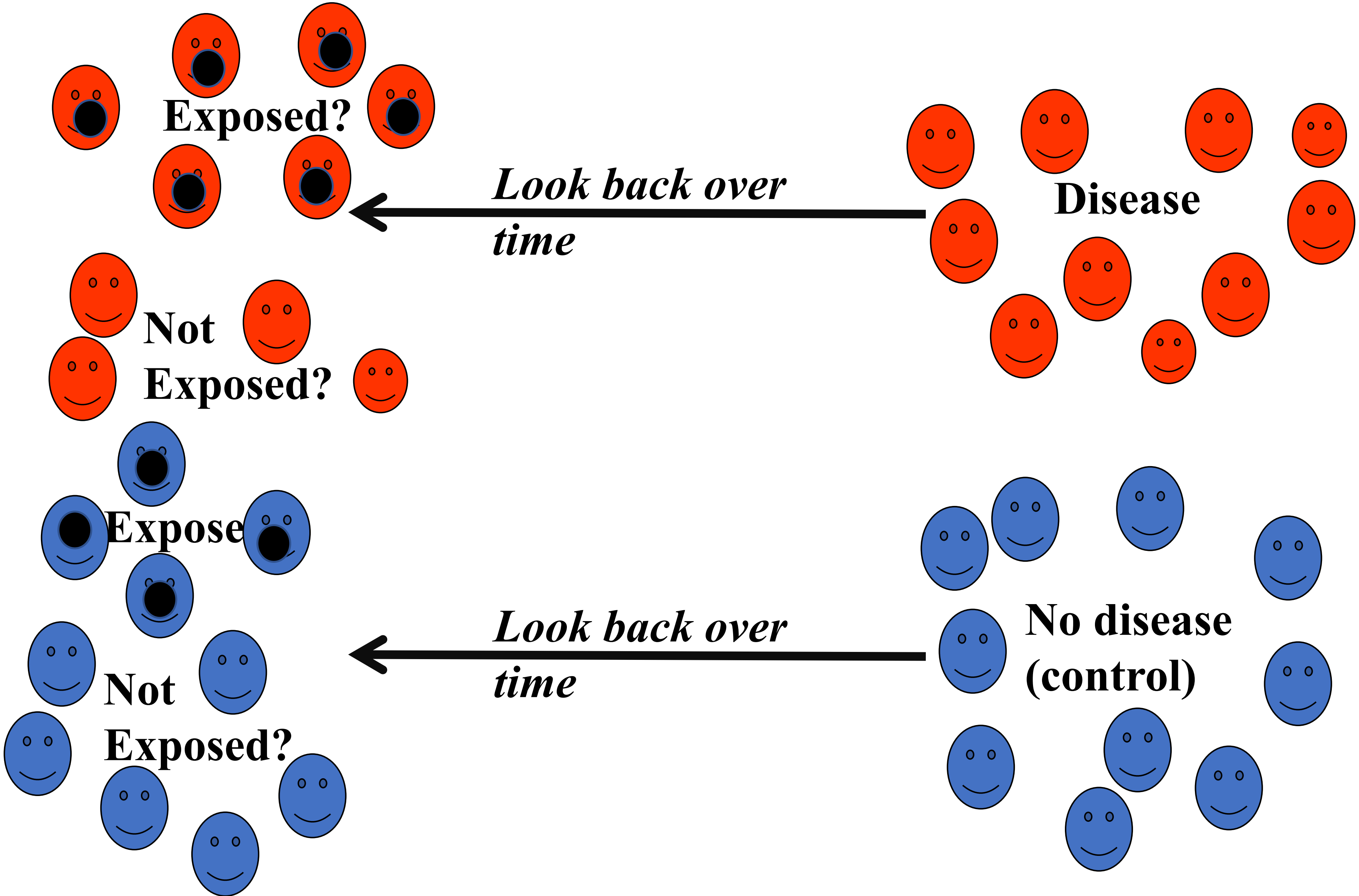
$$Odds(A) = \frac{probability(A\ happens)}{probability(A\ does\ not\ happen)} = \frac{prob(A)}{1 - prob(A)}$$

$$prob(A) = \frac{Odds(A)}{1 + Odds(A)}$$

Odds Ratio (OR)

$$OR = \frac{\text{Odds of exposure}_{\text{cases}}}{\text{Odds of exposure}_{\text{controls}}}$$

Case control studies



Case-control study

	Disease Present	Disease absent	
Exposure Present	a	b	a+b
Exposure absent	c	d	c+d
Total	a+c	b+d	a+b+c+d

Odds of being ill in exposed=a/b

Odds of being ill in non exposed =c/d

Odds ratio (OR)=Odds in exposed/Odds in non exposed

= OR=(a/b)/(c/d)

$$Odds\ Ratio\,(OR) = \frac{ad}{cb}$$

Case-control study

Early life exposure to diagnostic radiation and ultrasound scans and risk of childhood cancer: case-control study

BMJ 2011;342:d472

Objective To examine childhood cancer risks associated with exposure to diagnostic radiation and ultrasound scans in utero and in early infancy (age 0-100 days).
Design Case-control study.
Setting England and Wales.
Participants 2690 childhood cancer cases and 4858 age, sex, and region matched controls from the United Kingdom Childhood Cancer Study (UKCCS), born 1976-96.
Main outcome measures Risk of all childhood cancer, leukaemia, lymphoma, and central nervous system tumours, measured by odds ratios.

Case-control study: example

Radiation	**Case**	**Control**	Total
Yes	140	165	305
No	1550	5693	7243
Total	1690	5858	7548

Odds of outcome in exposed = 140 / 165 = 0.85

Odds of outcome in non-exposed = 1550 / 5693 = 0.27

Outcome odds ratio = (a/b) / (c/d) = 0.85/0.27=3.1

CASE-CONTROL STUDIES

Methods of data collection

Case-note review: Completeness

Postal questionnaire: response rate

Interview: Detailed information

Obtaining cases and controls for case control studies

Study	Source of cases	Source of controls
PROM (premature rupture of membrane)	Hospital patients	Hospital patients
Rheumatoid arthritis	Outpatient clinic	Other outpatient clinic
Cervical screening	GP register	GP register

Bias

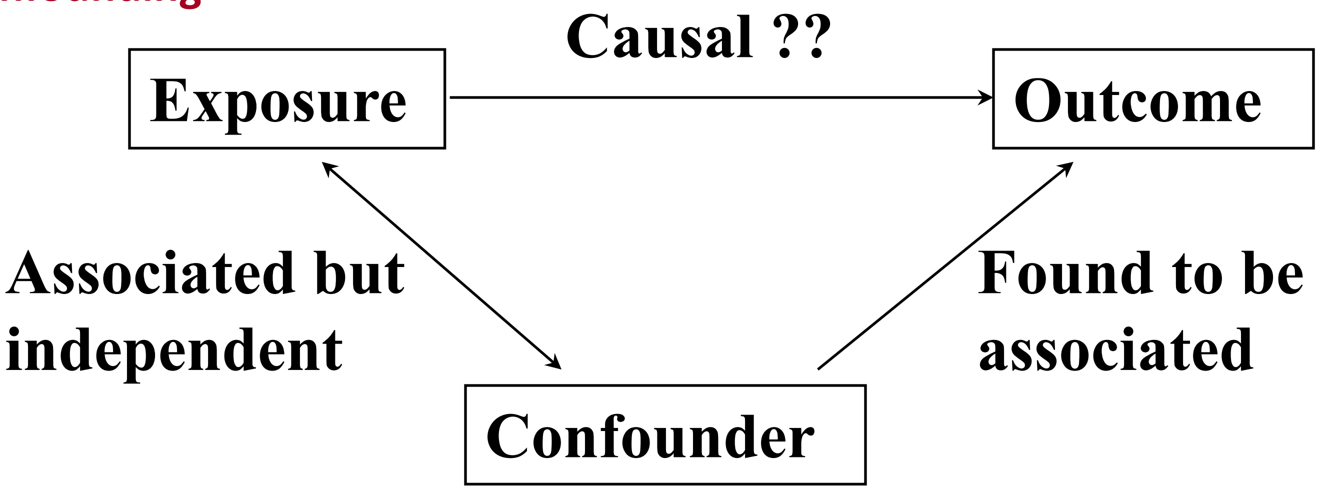
Bias is any systematic error in an epidemiological study that results in an incorrect estimate of the association between exposure and risk of the outcome

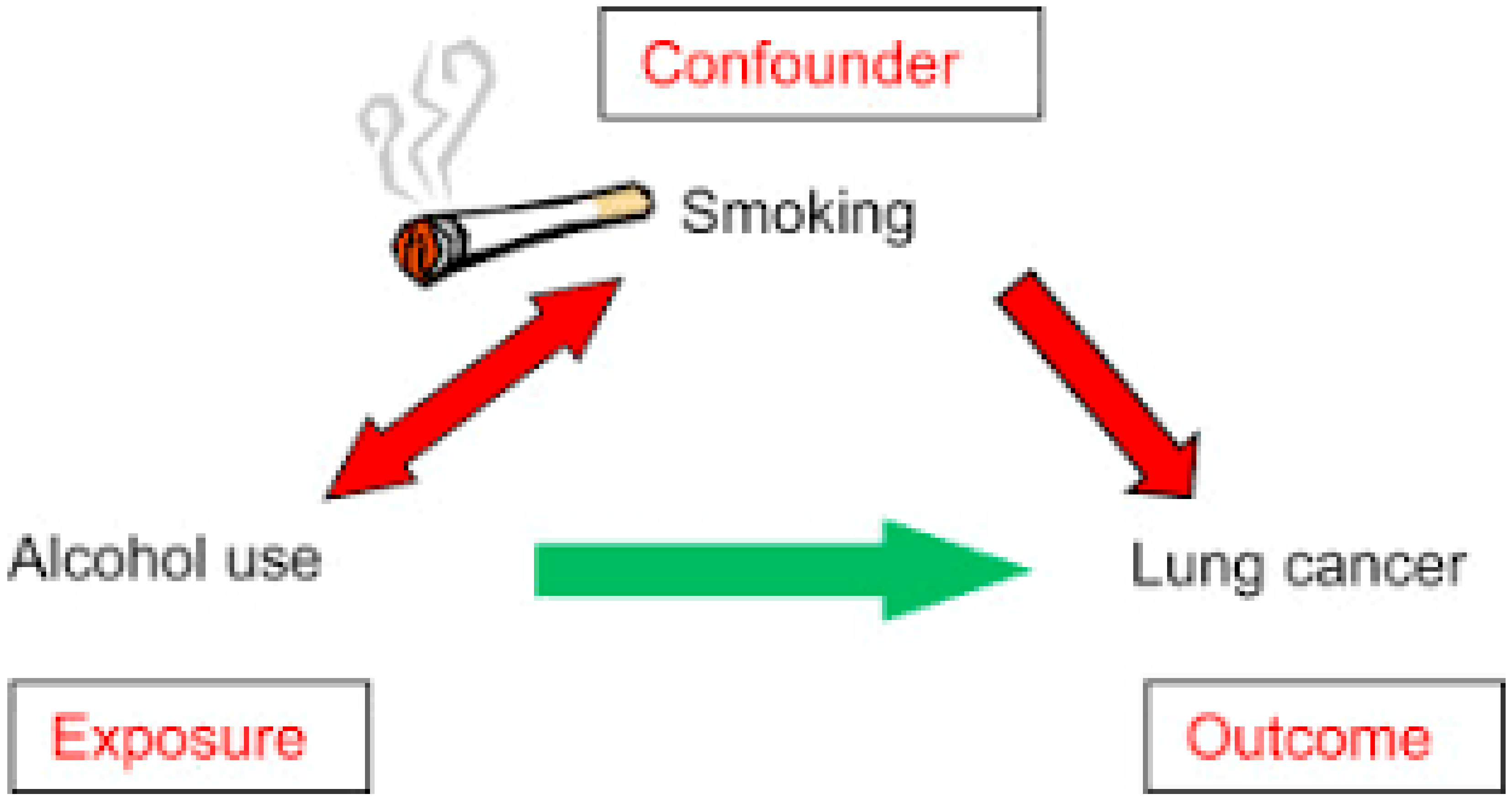
- **Selection bias:** inappropriate controls
- **Observation bias**
  - **Subject and recall bias:** eg recall bias of mothers with cerebral palsy babies
  - **Interviewer bias: blind if possible**
  - **Misclassification**

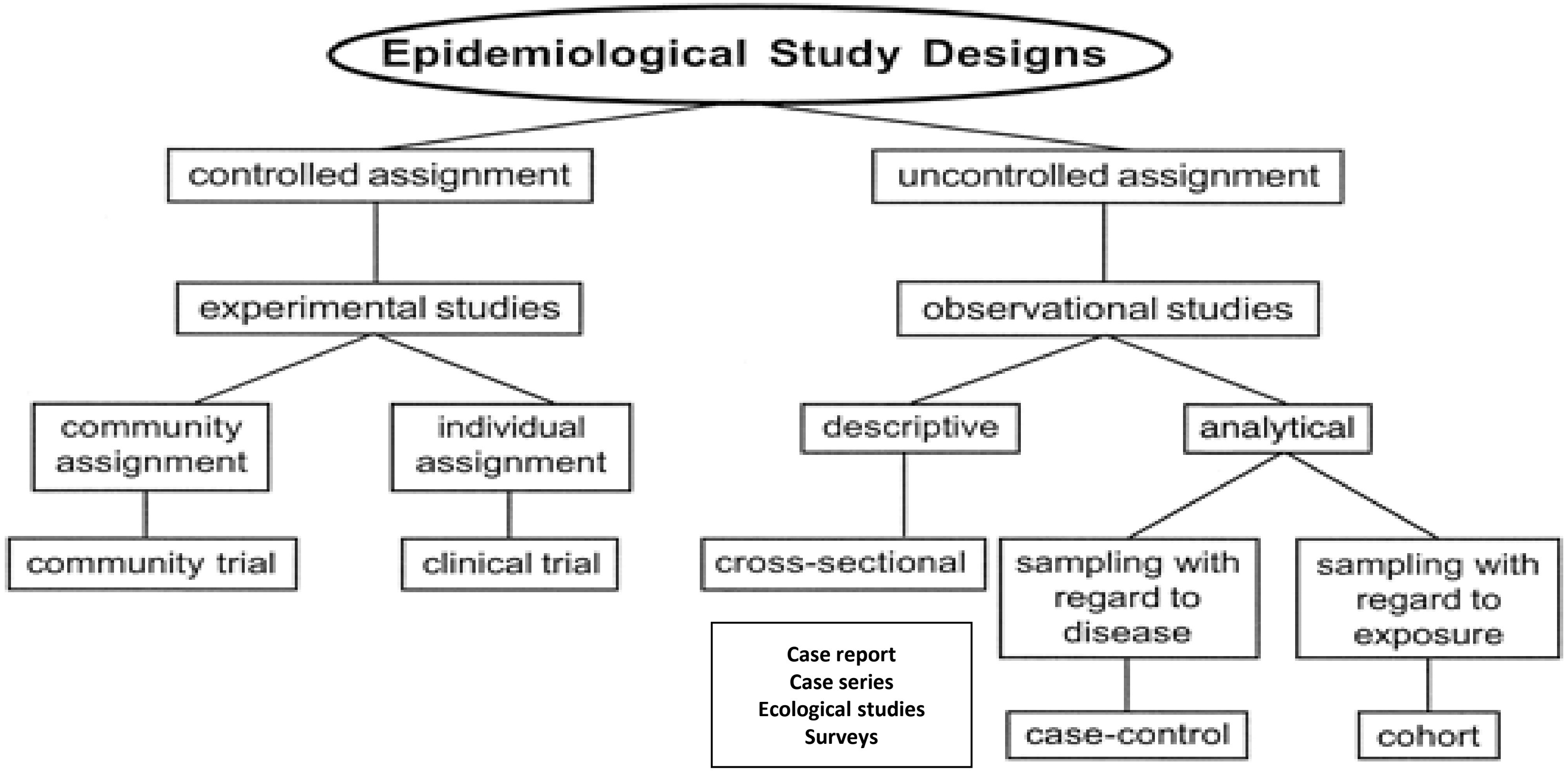
Confounding

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




Confounder
Smoking
Alcohol use
Lung cancer
Exposure
Outcome



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

Copyright © The McGraw-Hill Companies, Inc. All rights reserved.

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 (intervention)studies

Clinical trials

Preventive trials

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(if not available approved treatment)**

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

Clinical trials

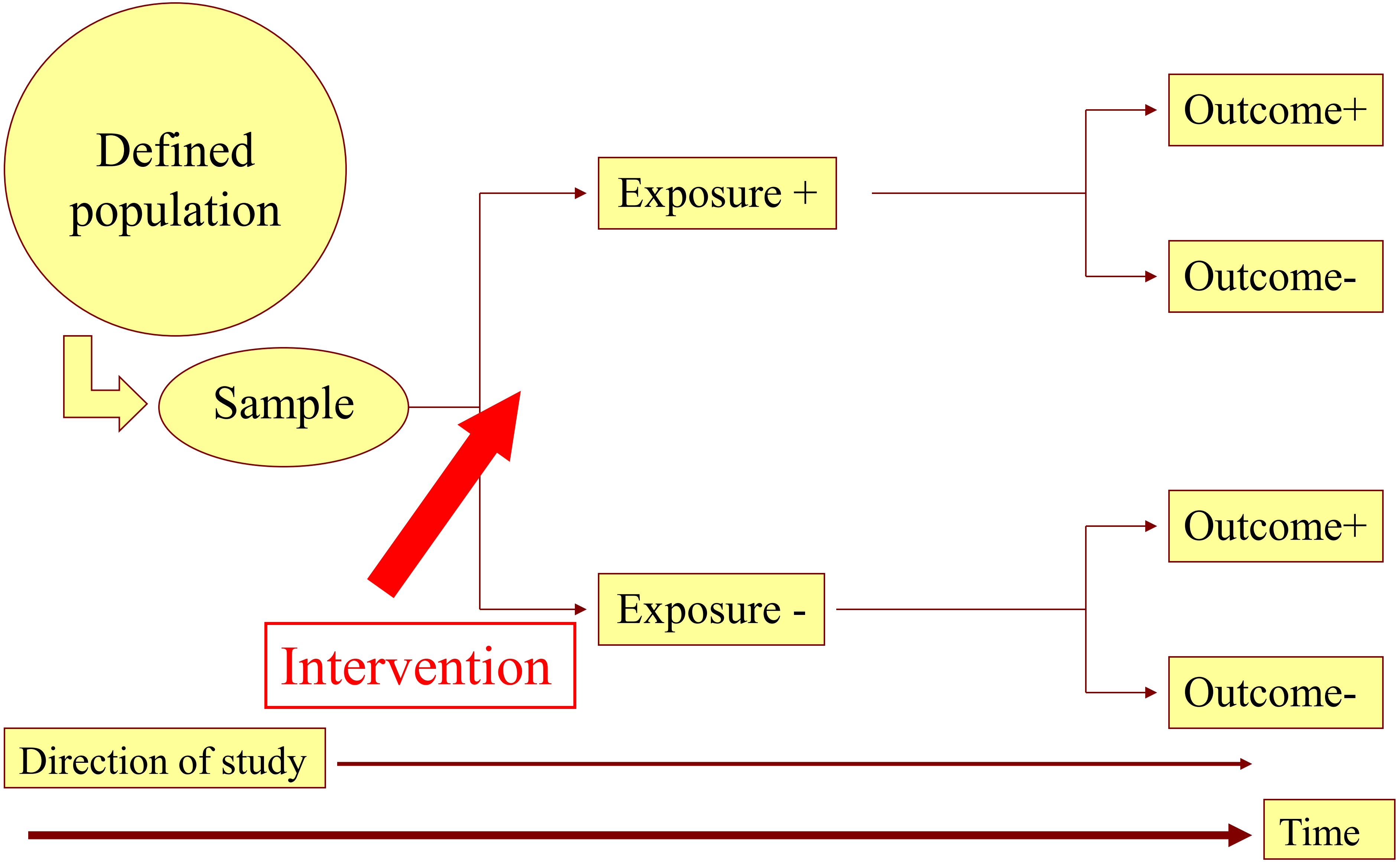
- **Are studies of the effect of a specific treatment on patients who already have a particular disease**

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

What trials assess

- Drugs
- Surgery
- Type of management
- New services

Clinical trial



RCT Disadvantages

- **Large trials (may affect statistical power)**
- **Long term follow-up (possible losses)**
- **Compliance**
- **Expensive**
- **Public health perspective ?**
- **Possible ethical questions**
- **As above, may take a long time.**
- **Must be ethically and laboriously conducted.**
- **Requires treatment on basis (in part) of scientific rather than medical factors. Patients may make some sacrifice**

Clinical trial: Study design

It is also related to:

- **Status of existing knowledge**
- **Occurrence of disease**
- **Duration of latent period**
- **Nature and availability of information**
- **Available resources**

Defining the patients

- Diagnostic features
- Eligibility criteria (inclusion and exclusion)

Assessing the outcome

- Clinically relevant
- Easily measured
- Accurately measured

Types of outcomes

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

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

Summary of trial design

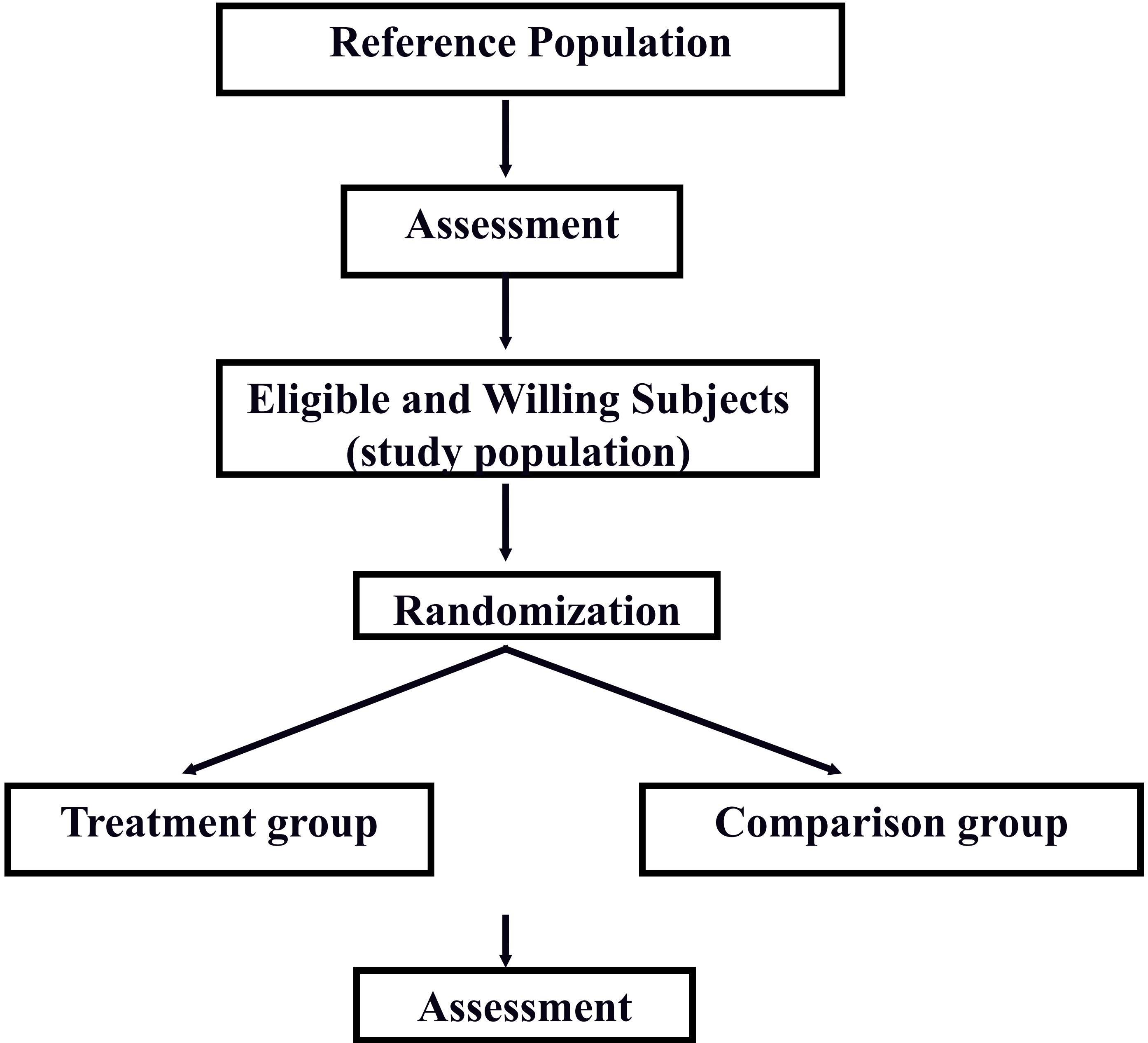
- Specify the treatment
- Define study group
- Random allocation
- Blinded outcome assessment
- Fair interpretation

Clinical trial

Common problems

- Too few patients
- Failed randomization
- Patients lost to follow-up
- Flawed analysis-interpretation
- Power of study: not big enough

Parallel Design



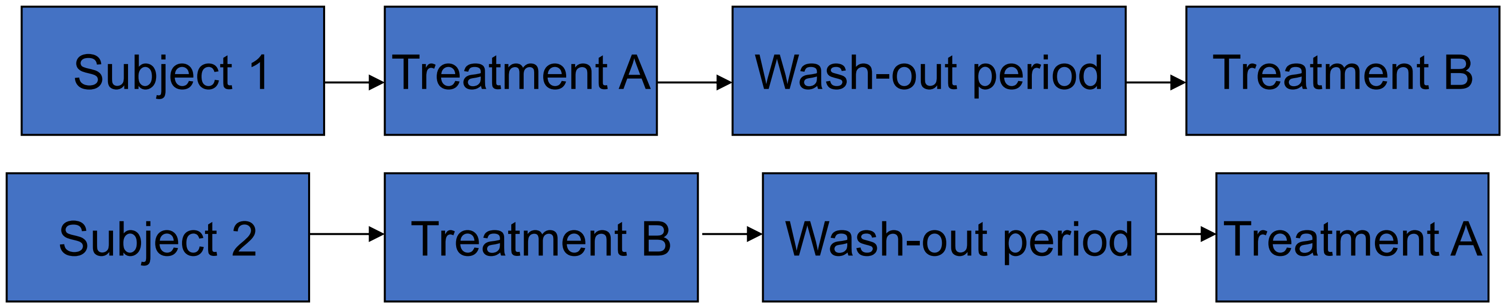
Cross-over clinical trial

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

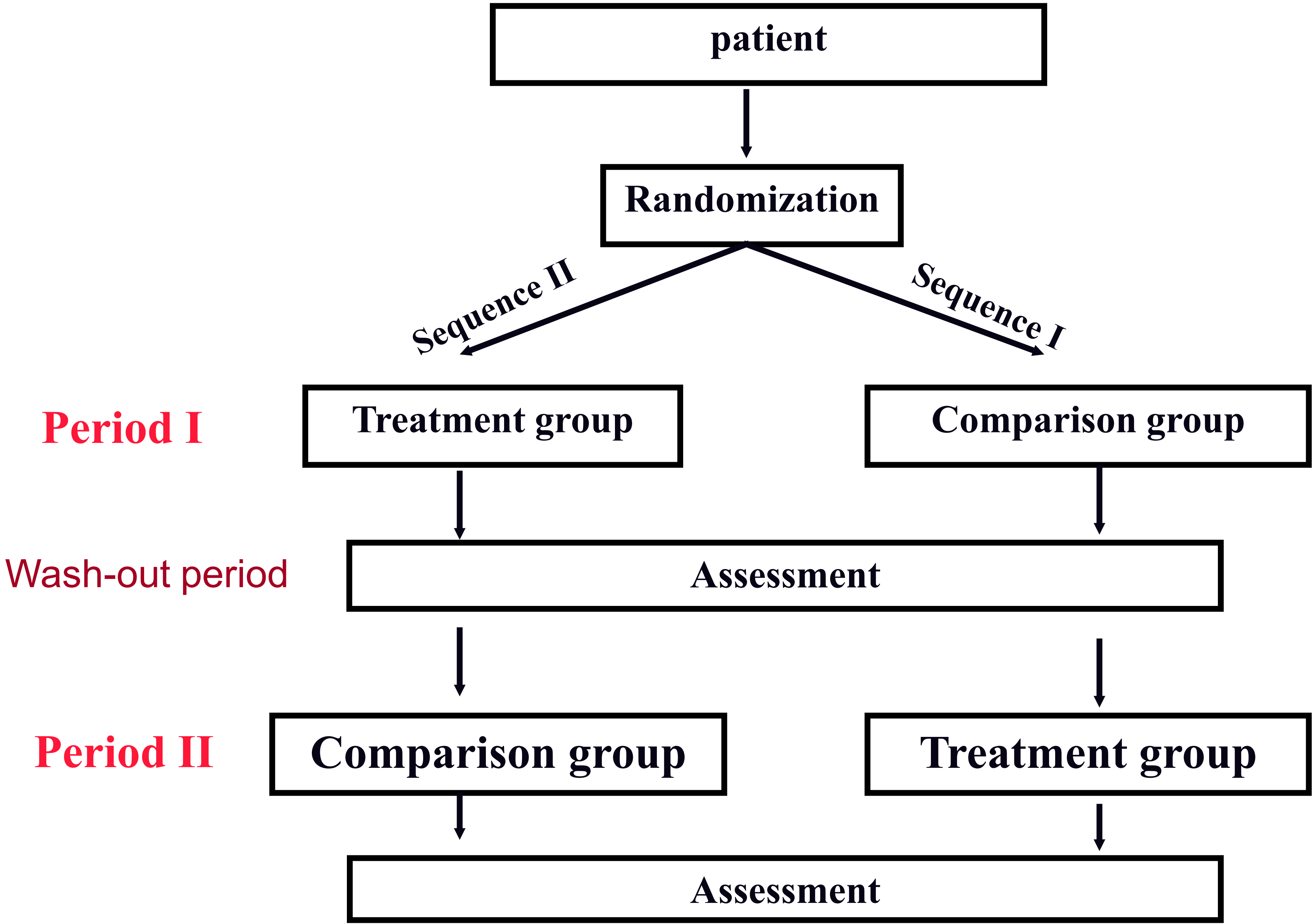
-Much smaller sample size

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

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 community

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 whooping 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**

Examples

- Smoking cessation interventions for secondary schools
- Medical Research participation interventions: one for JU and another intervention for JUST

- Increasing fluoride level within acceptable limits in all drinking water sources in Aqaba and comparing with Irbid, keeping this as they are.

Primary outcome: dental cases incidence for children younger than the age of 5.