Research Methodology

Study Design

Alireza Amirabadizadeh

PhD by research

Research Institute for Endocrine Sciences

Shahid Beheshti University of Medical Sciences

Study Design

Qualitative

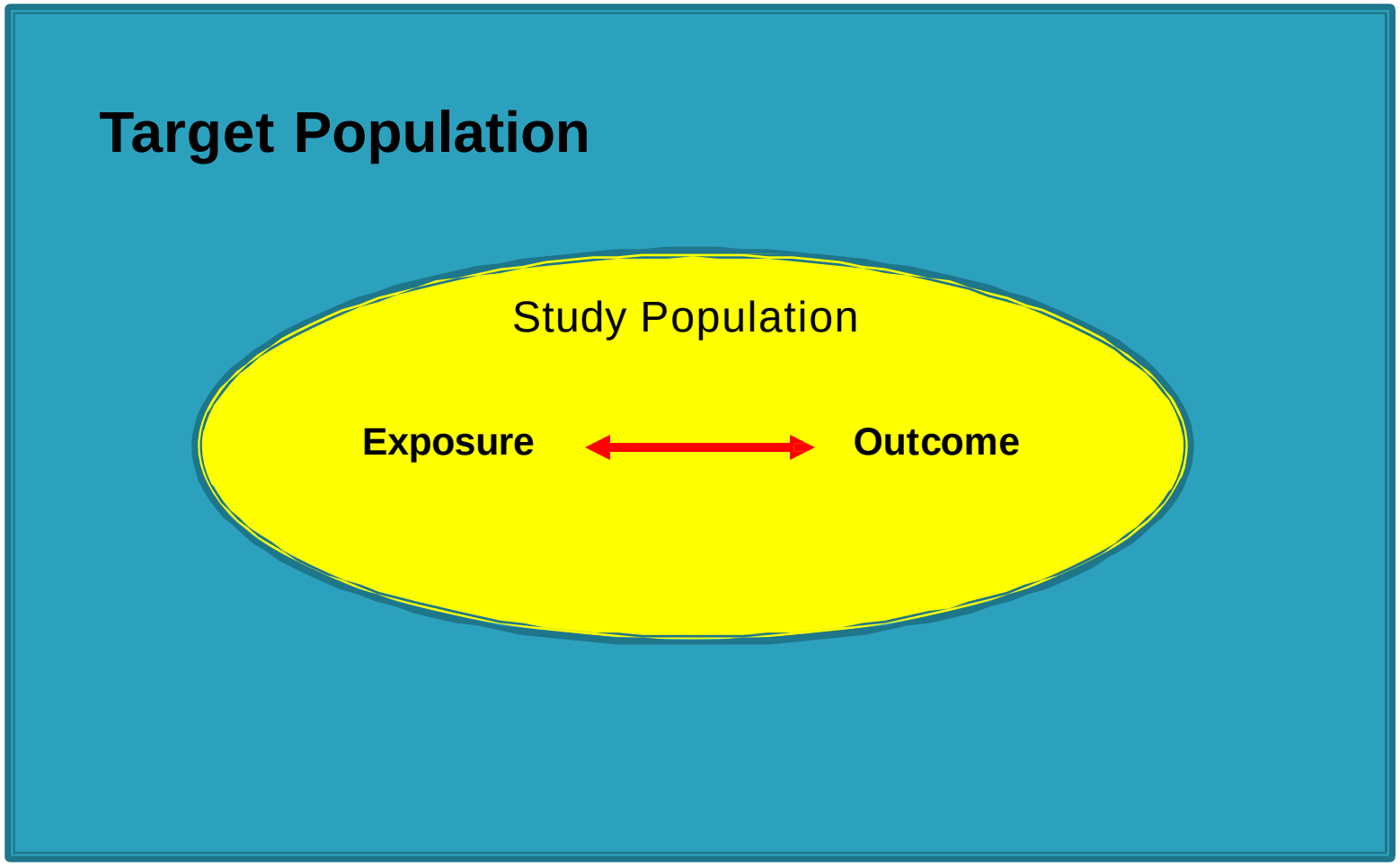
Quantitative

- Observational
- Experimental

Study Design

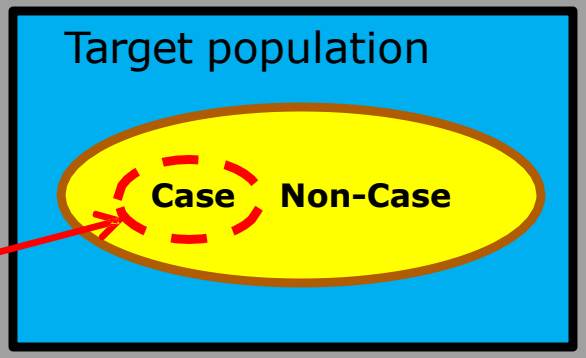
- *Observational*
  - Descriptive
  - **Analytic**
    - **Cross- Sectional**
    - **Case-Control**
    - **Cohort**
- *Experimental* *(Randomized Control Trial - RCT)*

Analytic Studies



Cross-Sectional Study

or prevalence study



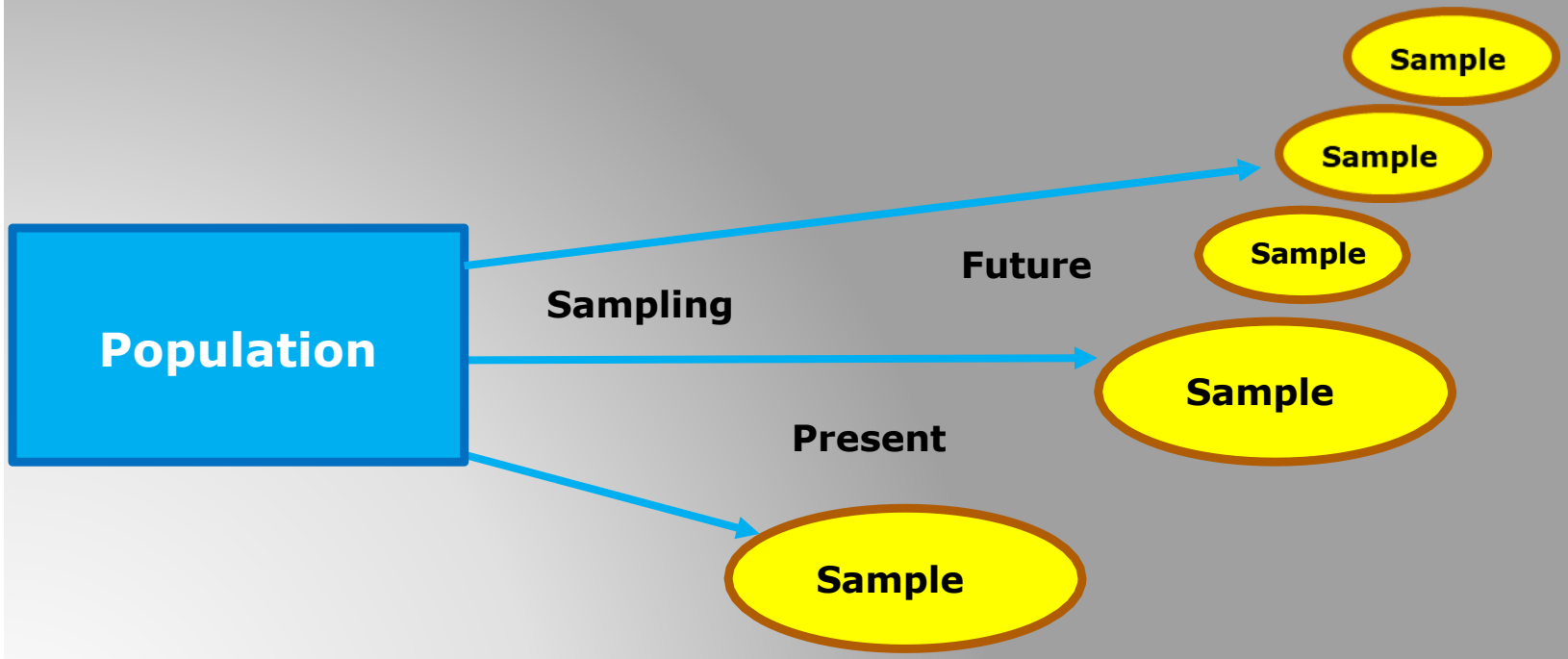
3 important questions to consider:

- ❑ Definition of the Population
- ❑ Definition of Case
- ❑ Definition of risk factors

Prevalence of Dis. = No. of disease+ / No. of subjects

Prevalence of R.F. = No. of risk factor+ / No. of subjects

Trend Design in Cross-Sectional studies



Advantages

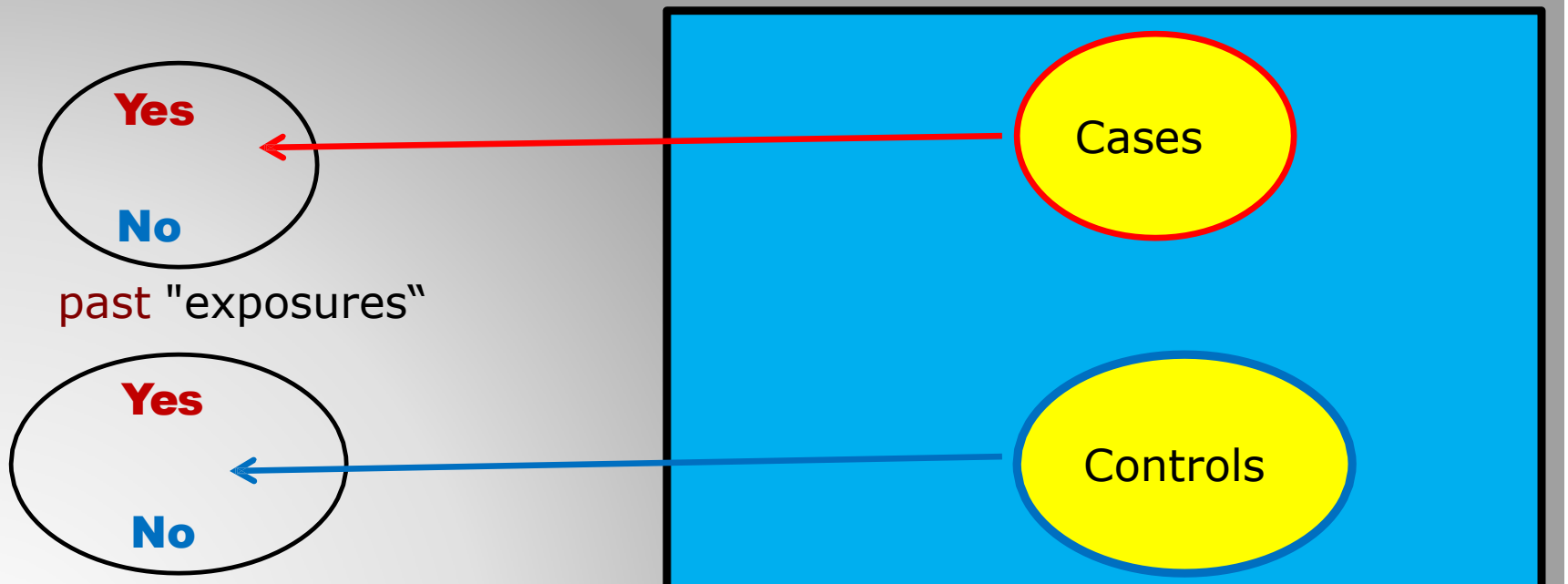
- Useful for descriptive studies
- Rapid, inexpensive, can provide *analytic clues*.
- Less prone to error about exposure recall bias

Disadvantages

- Prone to sample distortion bias.
- Unable to sort out what came first exposure or outcome
- Prone to seasonal and time to time variations

cross-sectional studies

Case-control Study



Selection of cases

First Step

- Precise definition of ‘case’.
- Inclusion / Exclusion criteria.
- How are cases to be identified? How recruited?

Selection of Controls

Second Step

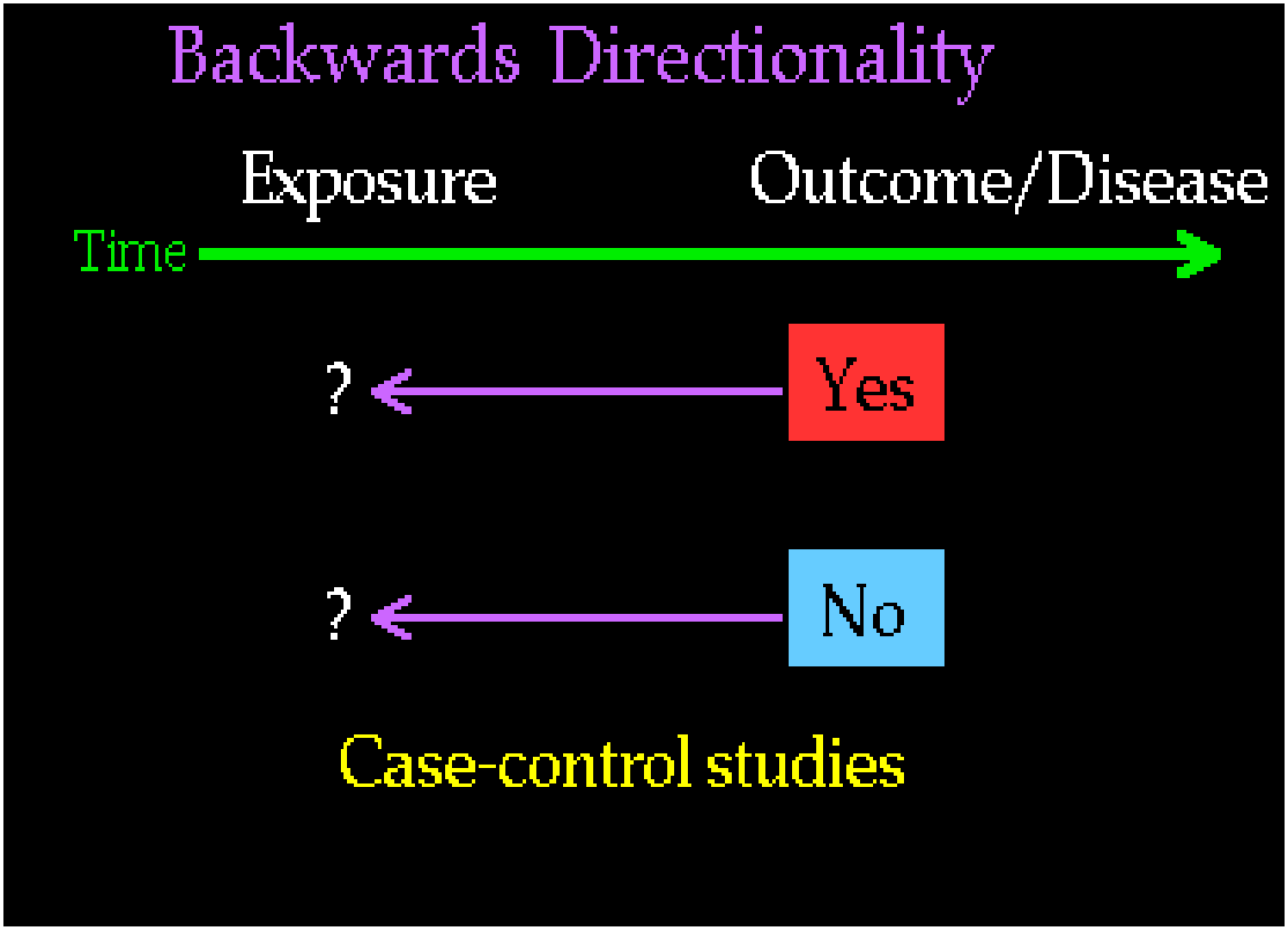
- Source (hospital patients without disease; neighborhood controls; random sample of population; sibs).
- Inclusion / exclusion criteria.

Controls must be related to the same population as the cases are.

Collection of information

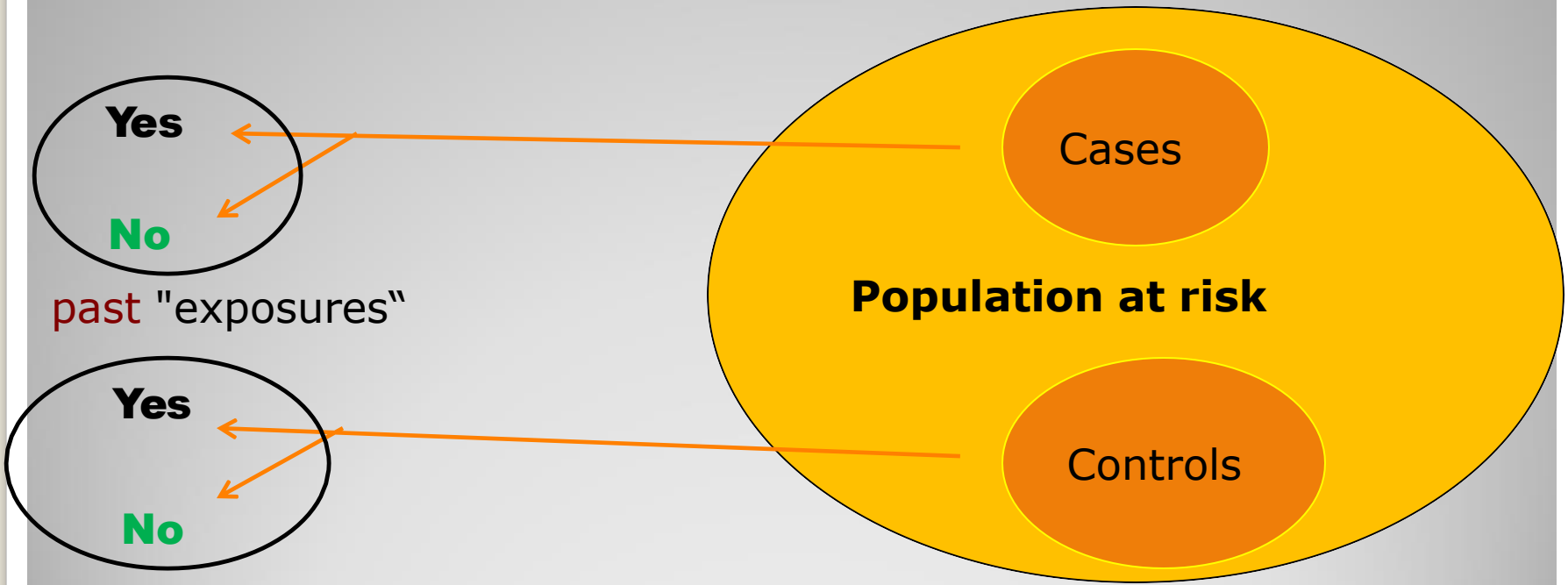
Third Step

- Identify risk factor of interest
- Method of collection of information (questionnaire; medical records; employment records)
- Same procedure to be used for cases and controls
- Interviewer should be unaware who is a case and who a control.


Backwards Directionality
Exposure
Outcome/Disease
Time
?
Yes
?
No
Case-control studies

Case-control Study

Compare { people who ***get the disease***
people who ***do not get the disease***



ADVANTAGES

- Relatively cheap compared to cohort studies
- Relatively quick
- Useful for study of rare diseases.
- No ethical problems
- Useful for diseases with long latent period.

Disadvantages

- Estimate of disease incidence cannot be done
- At times difficult to measure exposure accurately
- Open to selection bias.
- Difficult to interpret.

Case-control Studies

Cohort Study

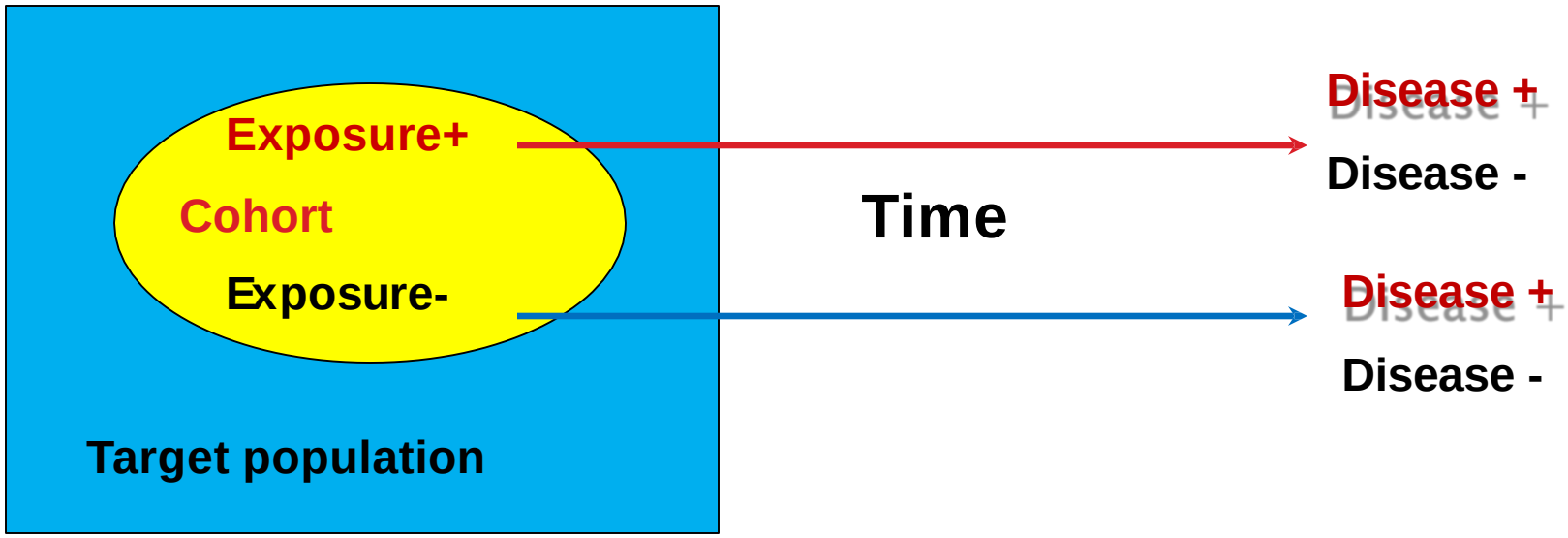
- A major limitation of cross-sectional surveys and case-control studies is difficulty to determine if exposure or risk factor preceded the disease or outcome.

- Cohort Study:

is the Key Point: 

Presence or absence of risk factor determine before outcome occurs.

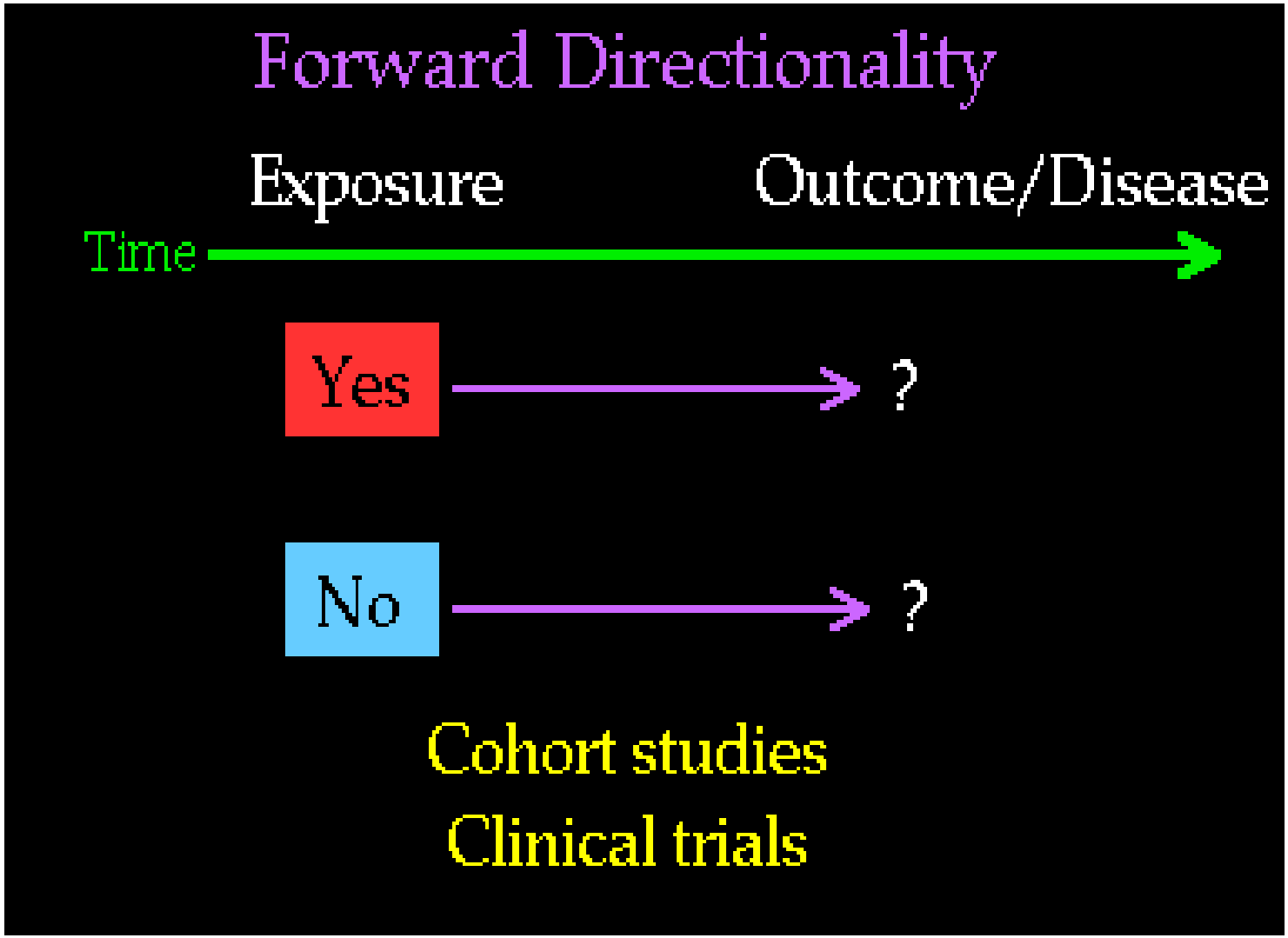
Cohort Study



Cohort studies

- Forward looking study (Prospectively or Retrospectively)
- Incidence study

- starts with people free of disease
- assesses exposure at “baseline”
- assesses disease status at “follow-up”


Forward Directionality
Exposure
Outcome/Disease
Time
Yes
?
No
?
Cohort studies
Clinical trials

Indication of a cohort study

- When there is good evidence of exposure and disease.
- When exposure is rare but incidence of disease is higher among exposed
- When follow-up is easy, cohort is stable
- When ample funds are available

Elements of cohort study

- Selection of study subjects

  (A defined population)

- Obtaining data on exposure

- Follow up to detect outcome

Selection of study subjects

- General population
  - Whole population in an area
  - A representative sample

- Special group of population
  - Selected group
    - occupation group / professional group
  - Exposure groups
    - Person having exposure to some physical, chemical or biological agent
      - e.g. X-ray exposure to radiologists

Types of Cohort Study

- Prospective cohort study
- Retrospective (historical) cohort study
- Combination of Retrospective and Prospective cohort study.

Cohort studies

Strengths

- We can find out incidence rate and risk
- More than one disease related to single exposure
- can establish cause - effect
- good when exposure is rare
- minimizes selection and information bias

Weaknesses

- losses to follow-up
- often requires large sample
- ineffective for rare diseases
- long time to complete
- expensive
- Ethical issues

Results of a Case-Control Study

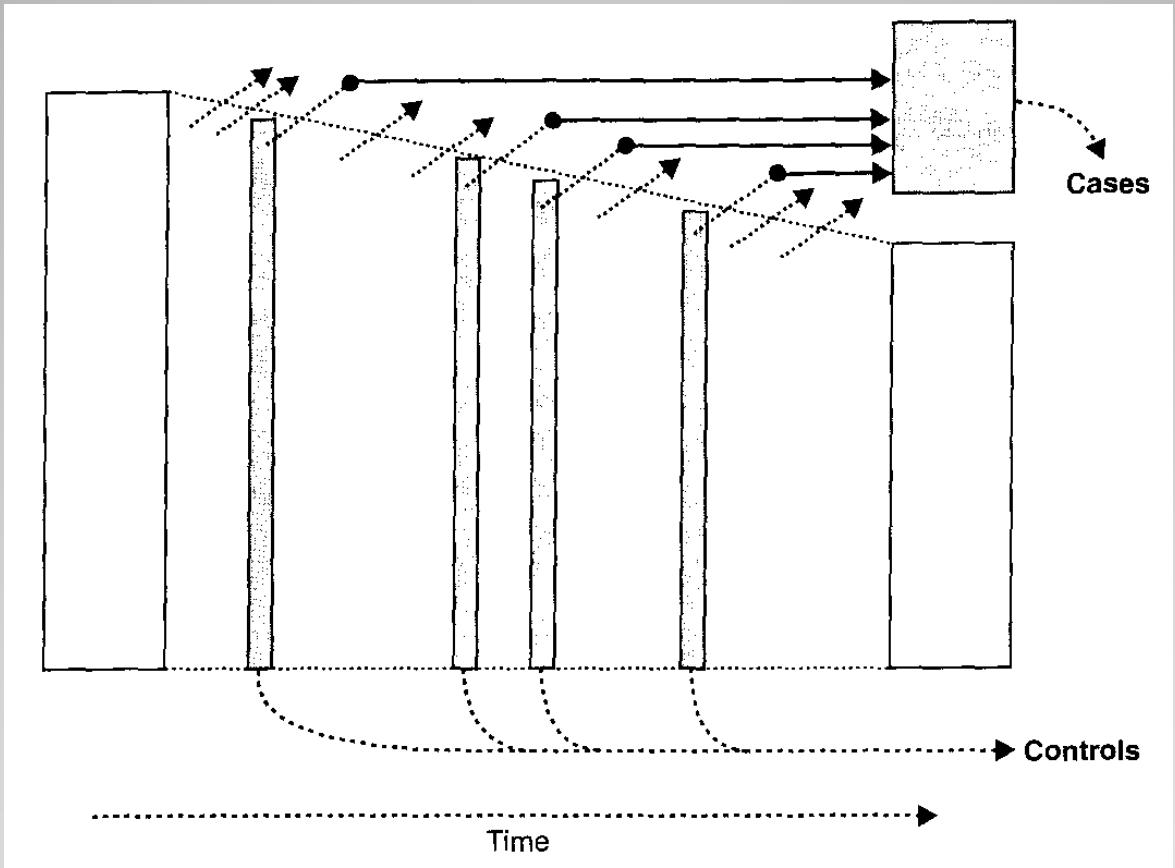
Risk factor	Disease	
	Yes (cases)	No (controls)
Yes	a	c
No	b	d
Total	N1	N2

N1 and N2 are fixed numbers

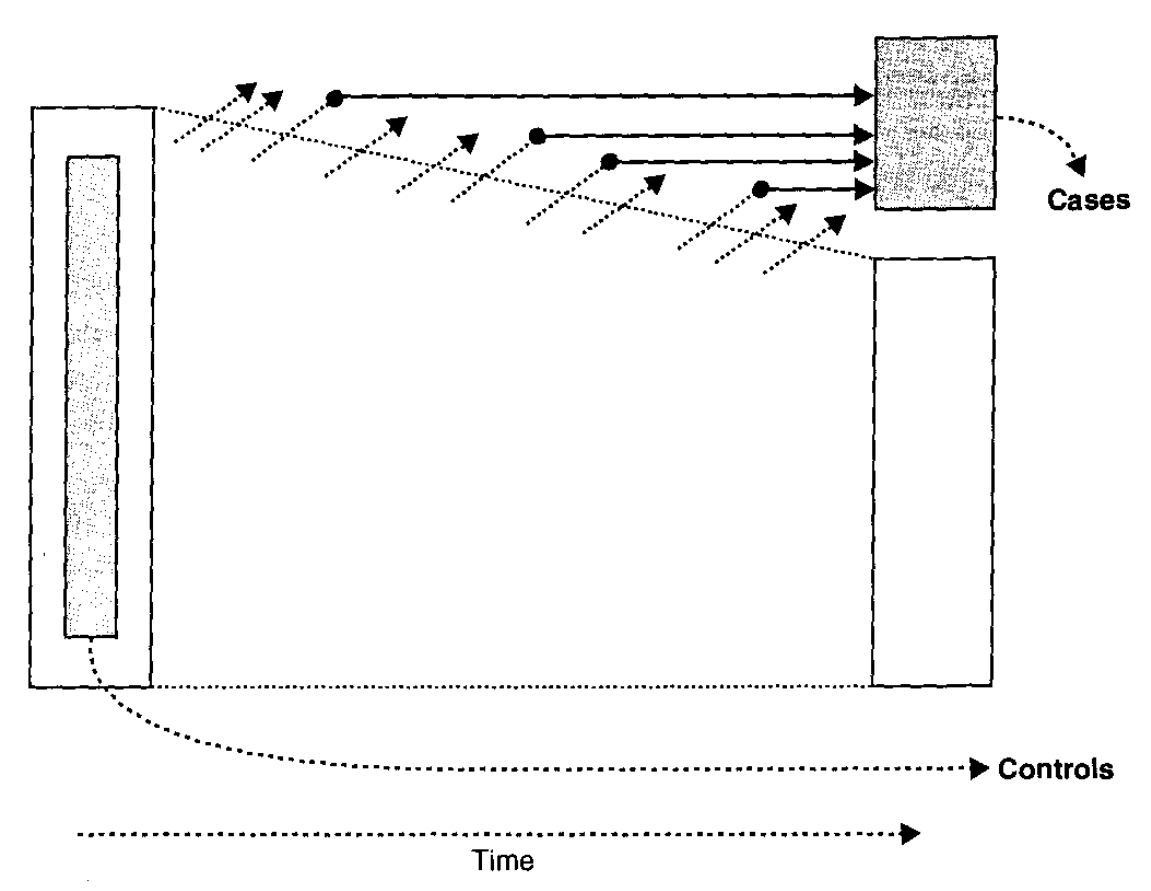
Result of cohort study

Risk factor	Disease	
	Yes	NO
Yes	A	B
NO	C	D
Total	N1	N2

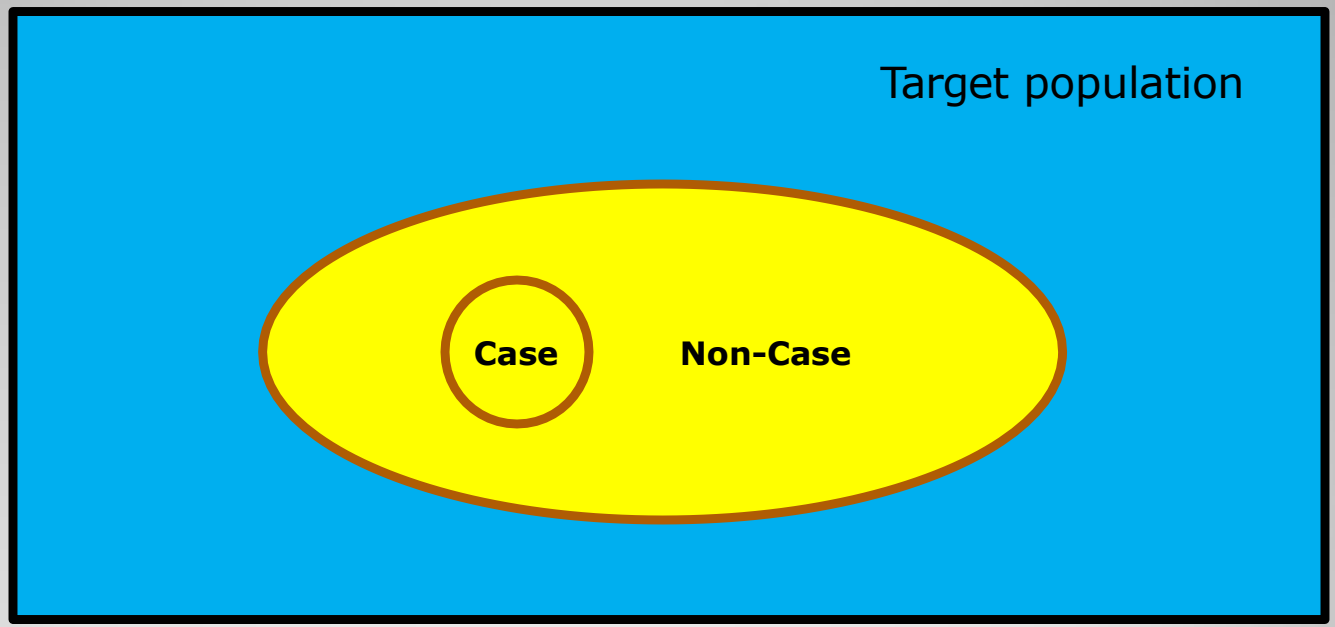
Nested case-control study



Case-cohort study

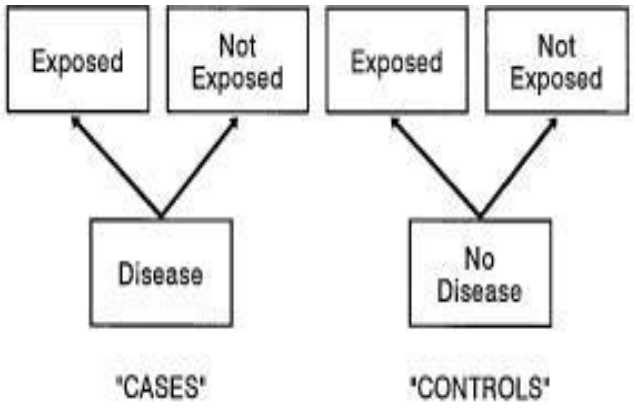


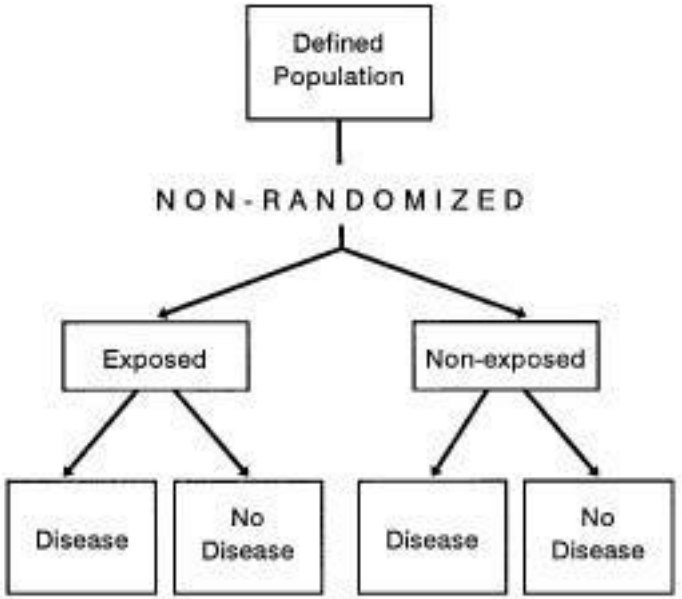
Cross-Sectional Case-Control Study

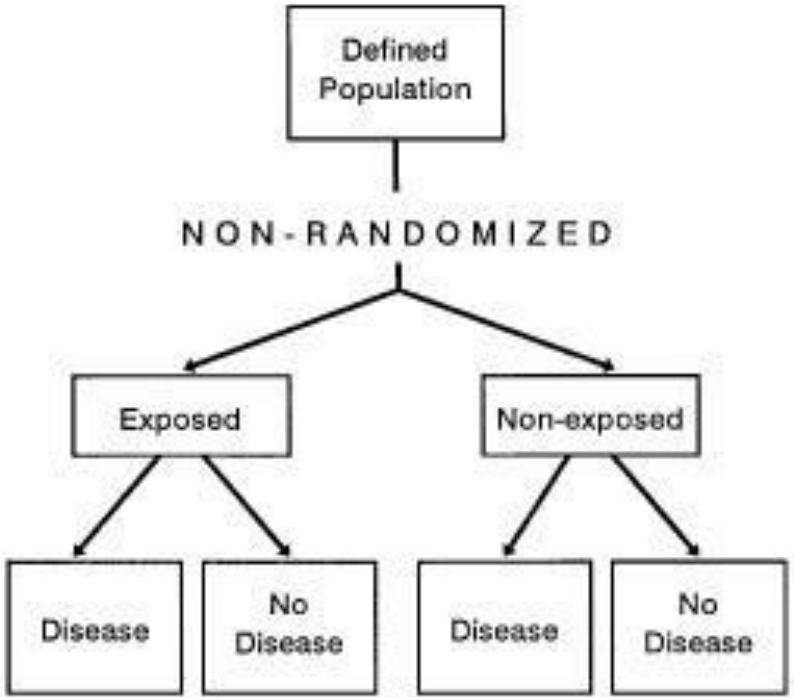


Measuring Associations between EXPOSURE and OUTCOME

Consider three kind of study designs:







How Do We Determine Whether a Certain Disease Is Associated with a Certain Exposure?

How can we determine whether an excess risk is associated with each of the food items?

	Disease +	Disease -	Total
Exposure +	a	b	a+b
Exposure -	c	d	c+d

Excess risk can be calculated in the two following ways:

- 1. The ratio of the risks (or of the incidence rates):

$$\text{Risk Ratio} = \frac{\text{Risk of disease in exposed group}}{\text{Risk of disease in unexposed group}}$$

Excess risk can be calculated in the two following ways:

- 1. The ratio of the risks (or of the incidence rates):

$$\text{Risk Ratio} = \frac{\text{Risk of disease in exposed group}}{\text{Risk of disease in unexposed group}}$$

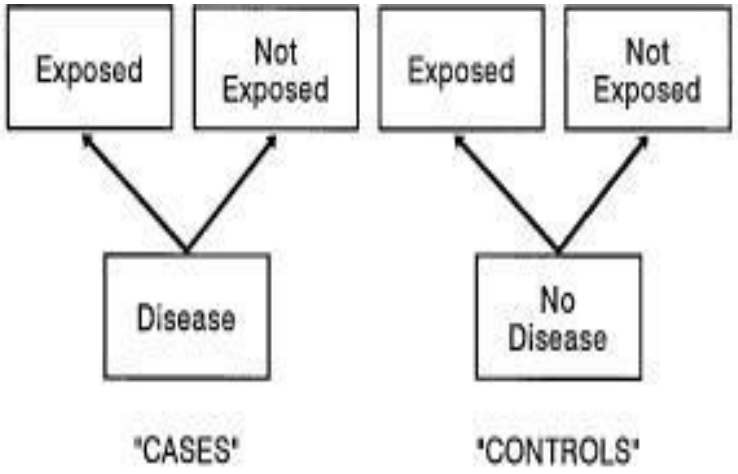
- 2. The difference in the risks (or in the incidence rates):

Risk Difference = (Risk of disease in exposed) – (Risk of disease in non exposed)

If RR = 1	Risk in exposed equal to risk in nonexposed (no association)
If RR > 1	Risk in exposed greater than risk in nonexposed (positive association; possibly causal)
If RR < 1	Risk in exposed less than risk in nonexposed (negative association; possibly protective)

Relative Risk in Case-Control Studies

- The incidence can't be derived from case-control studies since
  - **Begin with diseased people (cases) and non-diseased people (controls)**
- Therefore, can't calculate relative risk directly But, **we can use another method called an odds ratio**



	Disease +	Disease -
Exposure +	a	b
Exposure -	c	d
Total	a+c	b+d

Odds:

- The chance of something happening to the chance of it not happening
- Odds= P/ 1-P
- An odds is a special type of ratio, one in which the numerator and denominator sum to one.

Example:

- Suppose we are betting on a horse, which has a 60% probability of winning the race *(P).* The horse therefore has a 40% probability of losing (1 - *P*).
- If these are the probabilities, what are the *odds* that the horse will win the race?

Odds ratio in a cohort study

	Develop Disease	Do Not Develop Disease
Exposed	a	b
Not Exposed	c	d

$$\text{Odds Ratio} = \frac{\text{Odds that an exposed person develops disease}}{\text{Odds that a non-exposed person develops disease}}$$

$$A = \frac{a/b}{c/d} = \frac{ad}{bc}$$

- Odds ratio can be obtained from either a cohort or a case-control study and can be used instead of the relative risk.

Odds ratio in a case control study

	Cases (With Disease)	Controls (Without Disease)
History of Exposure	a	b
No History of Exposure	c	d

$$\text{Odds Ratio} = \frac{\text{Odds that a case was exposed}}{\text{Odds that a control was exposed}}$$

$$B = \frac{a/c}{b/d} = \frac{ad}{bc}$$

Cohort Study

1-year incidence of acute MI in individuals with sever SBP (≥180mmHg) and normal SBP (<120mmHg)

Blood Pressure Status	Myocardial Infarction				
	Number	*Present*	*Absent*	*Probability*	*Probability Odds$_{dis}$*
Severe hypertension	10,000	180	9820	180/10,000 = 0.0180	180/(10,000 − 180) = 180/9820 = 0.01833
Normal	10,000	30	9970	30/10,000 = 0.0030	30/(10,000 − 30) = 30/9970 = 0.00301

$$RR = \frac{\frac{180}{10{,}000}}{\frac{30}{10{,}000}} = \frac{0.0180}{0.0030} = 6.00$$

$$OR = \frac{\frac{180}{9820}}{\frac{30}{9970}} = \frac{0.01833}{0.00301} = 6.09$$

Incidence is low ⇨ RR ≈ OR

Cohort Study

Local reactions to influenza vaccine

	Local Reaction				
Group	*Number*	*Present*	*Absent*	*Probability*	*Probability Odds$_{dis}$*
Vaccine	2570	650	1920	650/2570 = 0.2529	650/(2570 − 650) = 650/1920 = 0.3385
Placebo	2410	170	2240	170/2410 = 0.0705	170/(2410 − 170) = 170/2240 = 0.0759

$$RR = \frac{\frac{650}{2570}}{\frac{170}{2410}} = \frac{0.2529}{0.0705} = 3.59 \qquad OR = \frac{\frac{650}{1920}}{\frac{170}{2240}} = \frac{0.3385}{0.0759} = 4.46$$

Incidence is high ⇨ RR ≠ OR

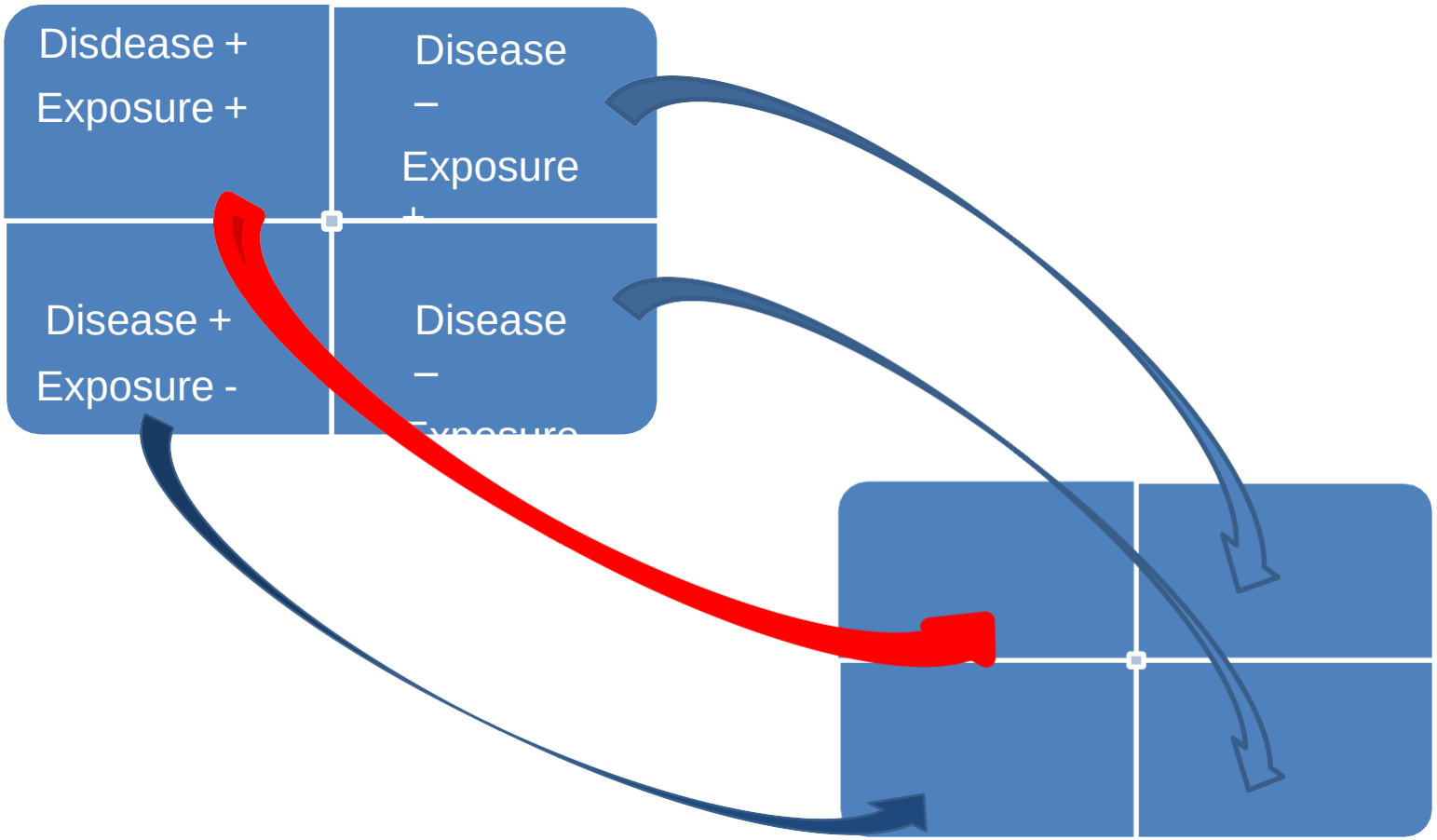
What about 95% CI of RR or OR?

- It should not include "1".
- Then the p-value would be <0.05.

Selection Bias

the way in which cases and controls, or exposed and nonexposed individuals, are selected such that an apparent association is observed

Reference Population



Study Sample

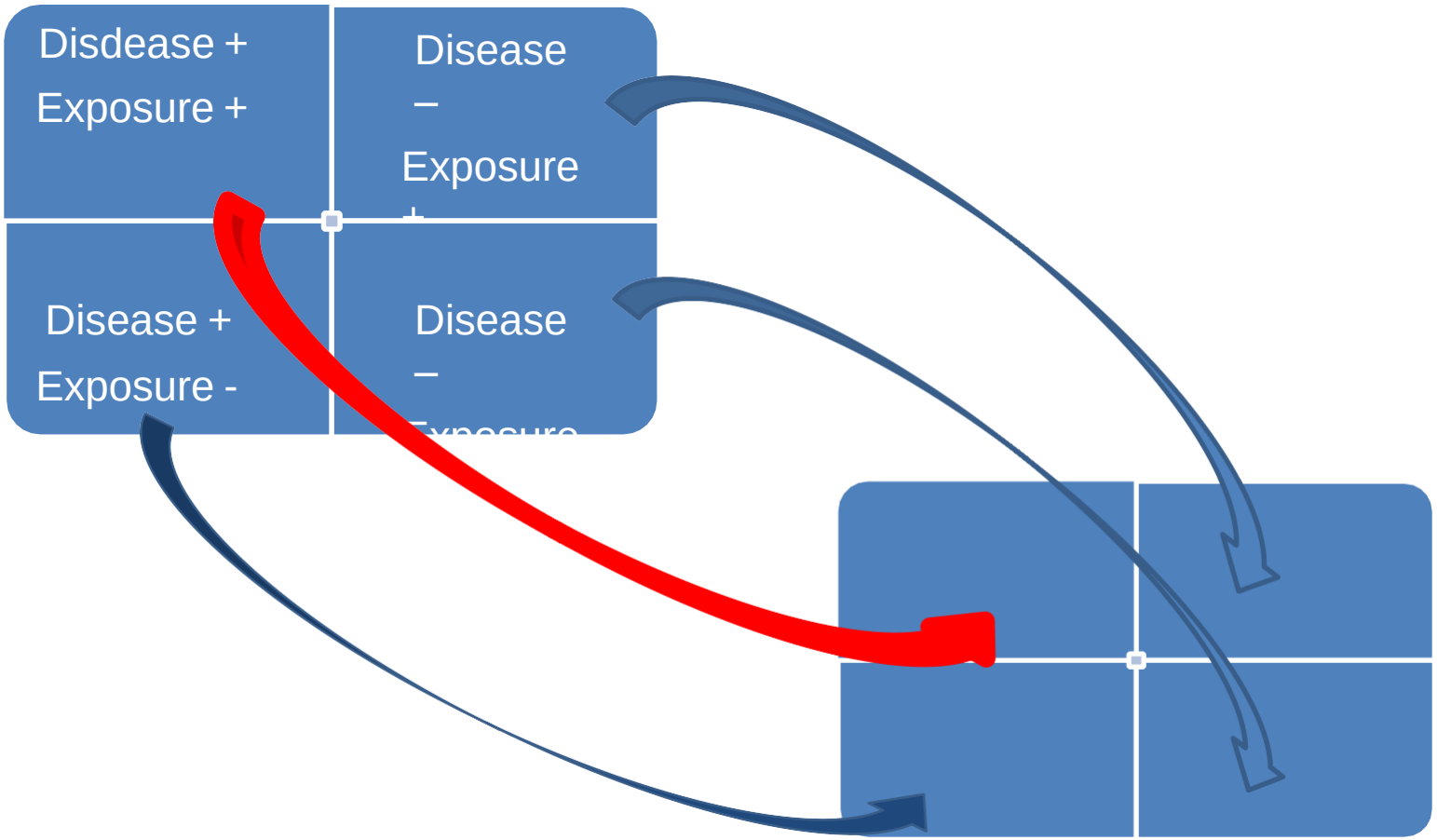
Selection Bias

- Non-response Bias
- Exclusion Bias
- Berksonian Bias
- Healthy worker effect
- Differential losses to follow-up

Selection Bias

the way in which cases and controls, or exposed and nonexposed individuals, are selected such that an apparent association is observed

Reference Population

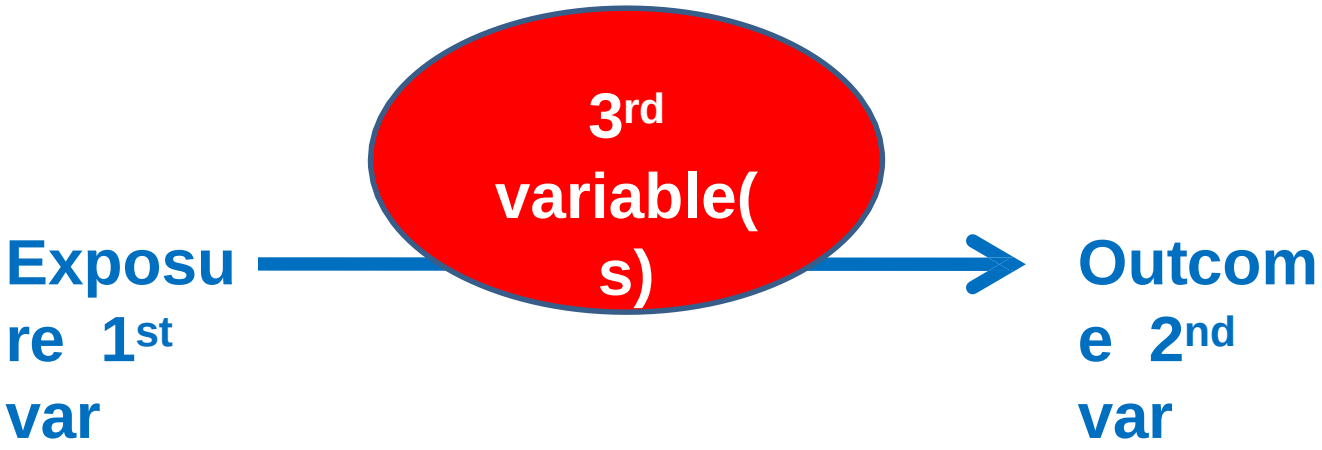


Study Sample

Considering

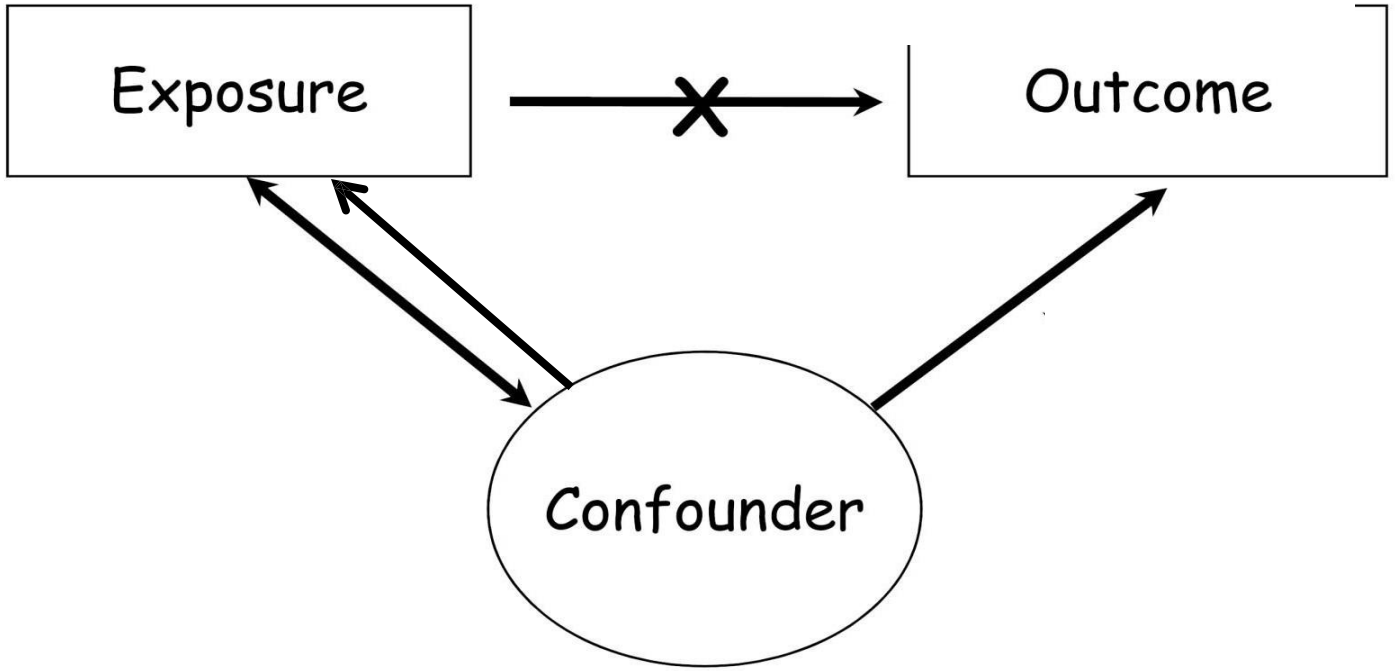
3rd Factor(s)

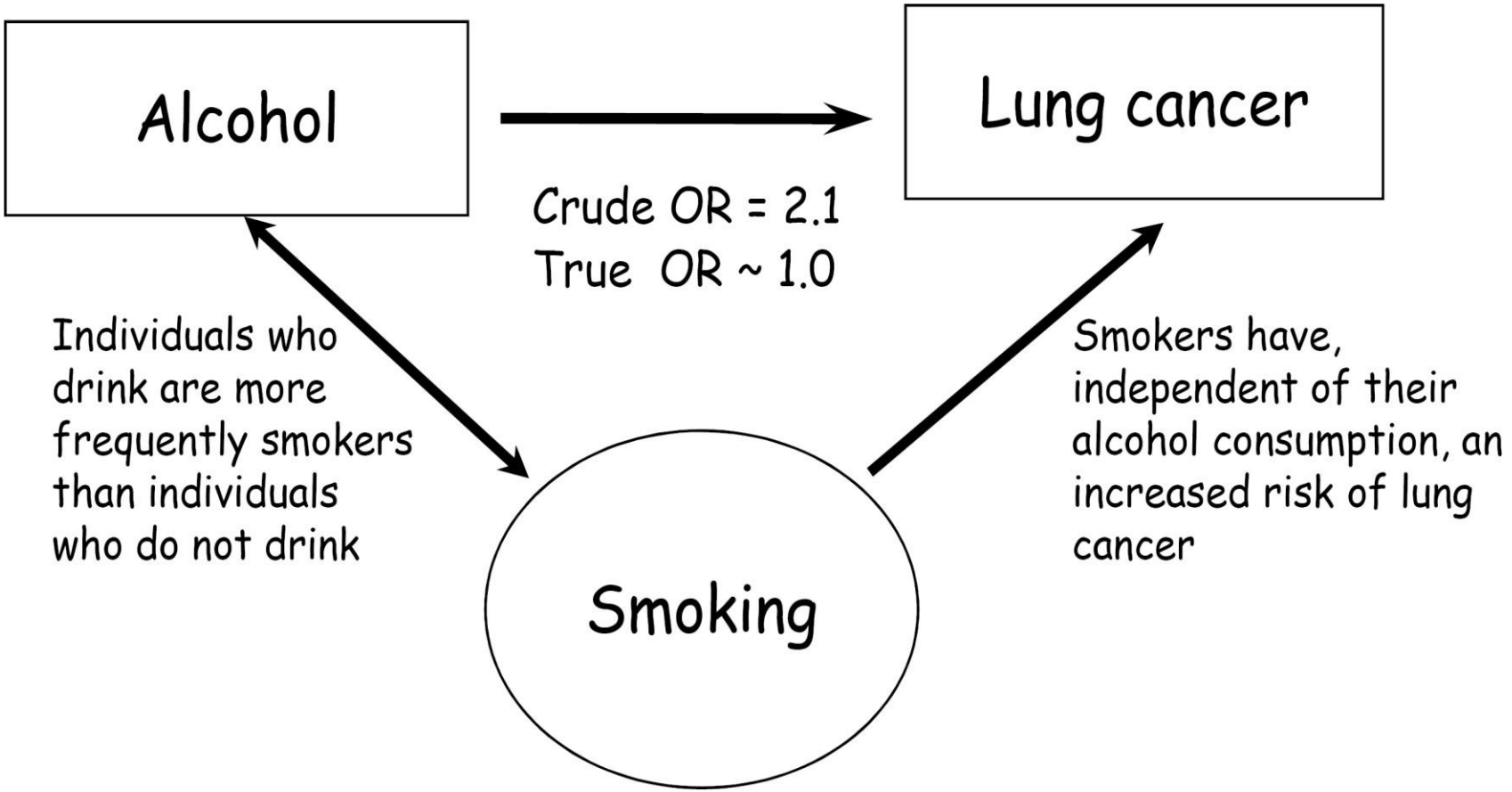
in Causality



CONFOUNDING

A confusion of effect




Alcohol
Lung cancer
Crude OR = 2.1
True OR ~ 1.0
Individuals who drink are more frequently smokers than individuals who do not drink
Smokers have, independent of their alcohol consumption, an increased risk of lung cancer
Smoking

Control of confounding

IN DESIGN

- Randomization
- Restriction
- Matching

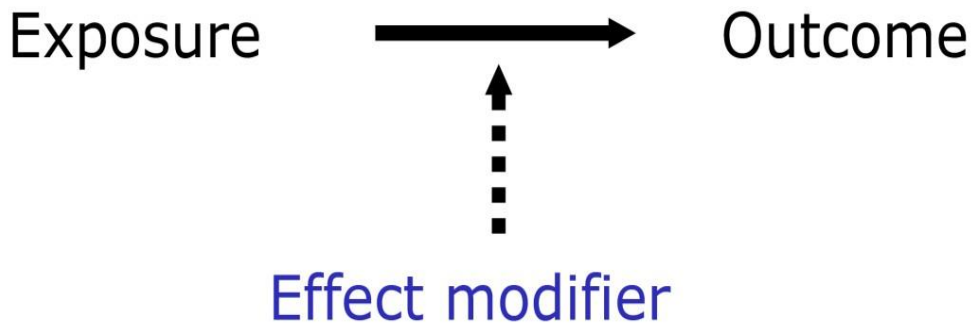
IN ANALYSIS

- Standardization
- Stratification
- Multivariate analysis

The most applicable method

Some points in:

EFFECT MODIFICATION

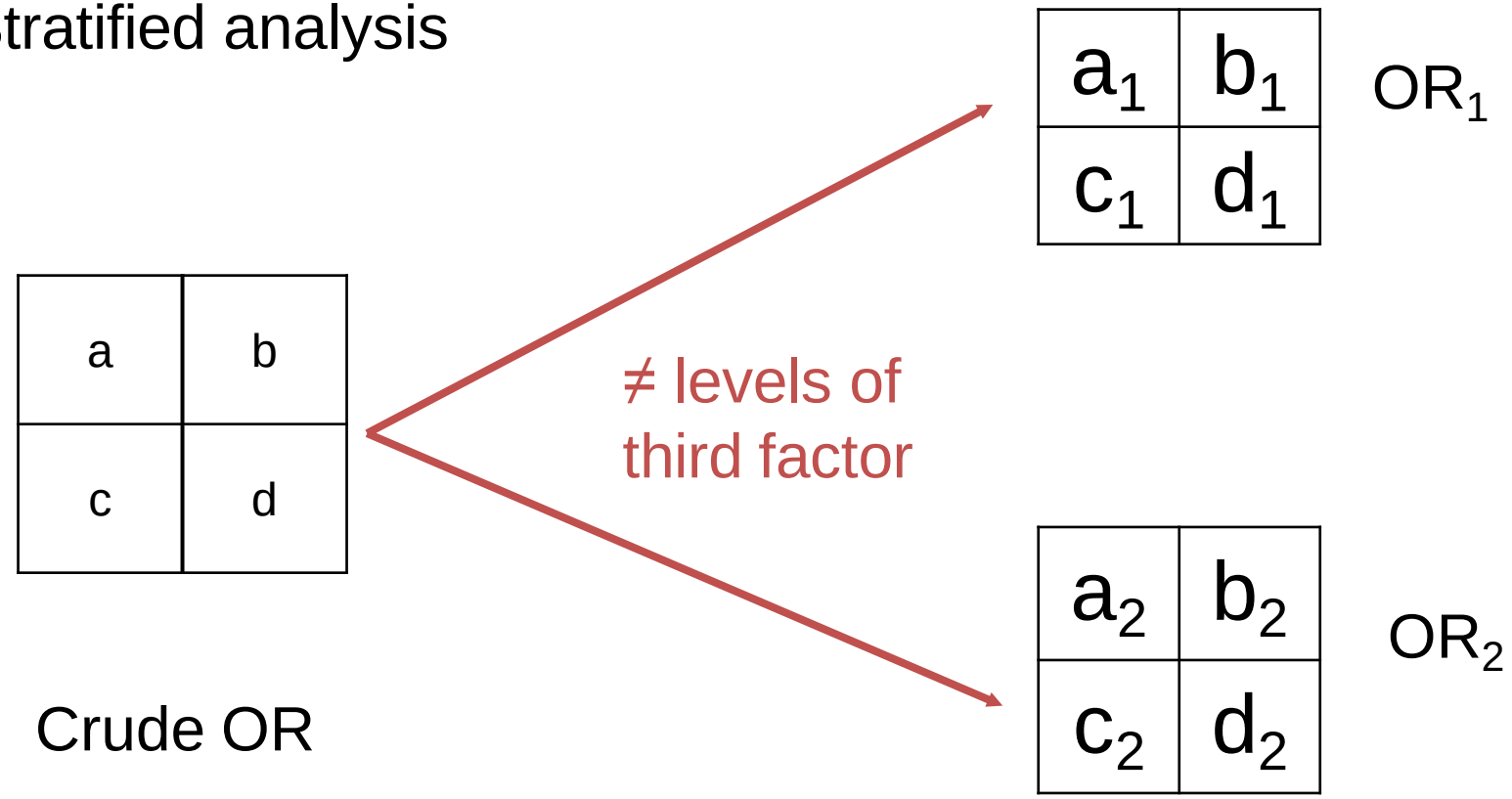


- The effect of one factor on outcome is modified by levels of another factor
- Important to present and discuss
- A factor may be both a confounder and an effect modifier

Strategy to take into account a third factor in data analysis

1) Crude analysis

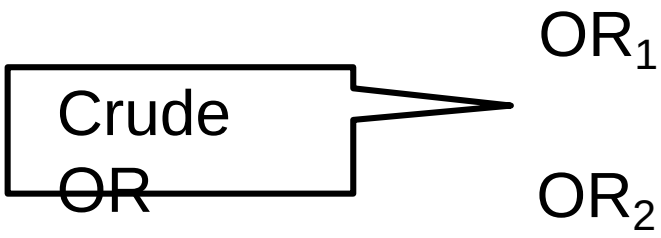
2) Stratified analysis



Strategy to take into account a third factor in data analysis

5a)

$OR_1 \neq OR_2$



Third factor = Effect modifier

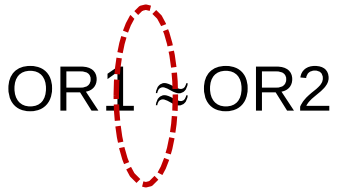
If it is clinically important

Don’t compute an adjusted OR

Report Stratum-specific results of the association between exposure and outcome

Strategy to take into account a third factor in data analysis

5b)



OR_1
OR_2

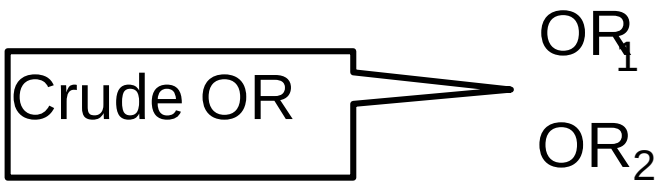
Crude OR

Computation of adjusted OR

Strategy to take into account a third factor in data analysis

5c)

$OR1 \approx OR2$



$OR_{ad} \approx OR_{Crude}$

Third factor = no role

Use crude OR to measure the association between exposure and outcome

- **Clinical Trial** **(on patients)**

- Field Trial (on healthy people)

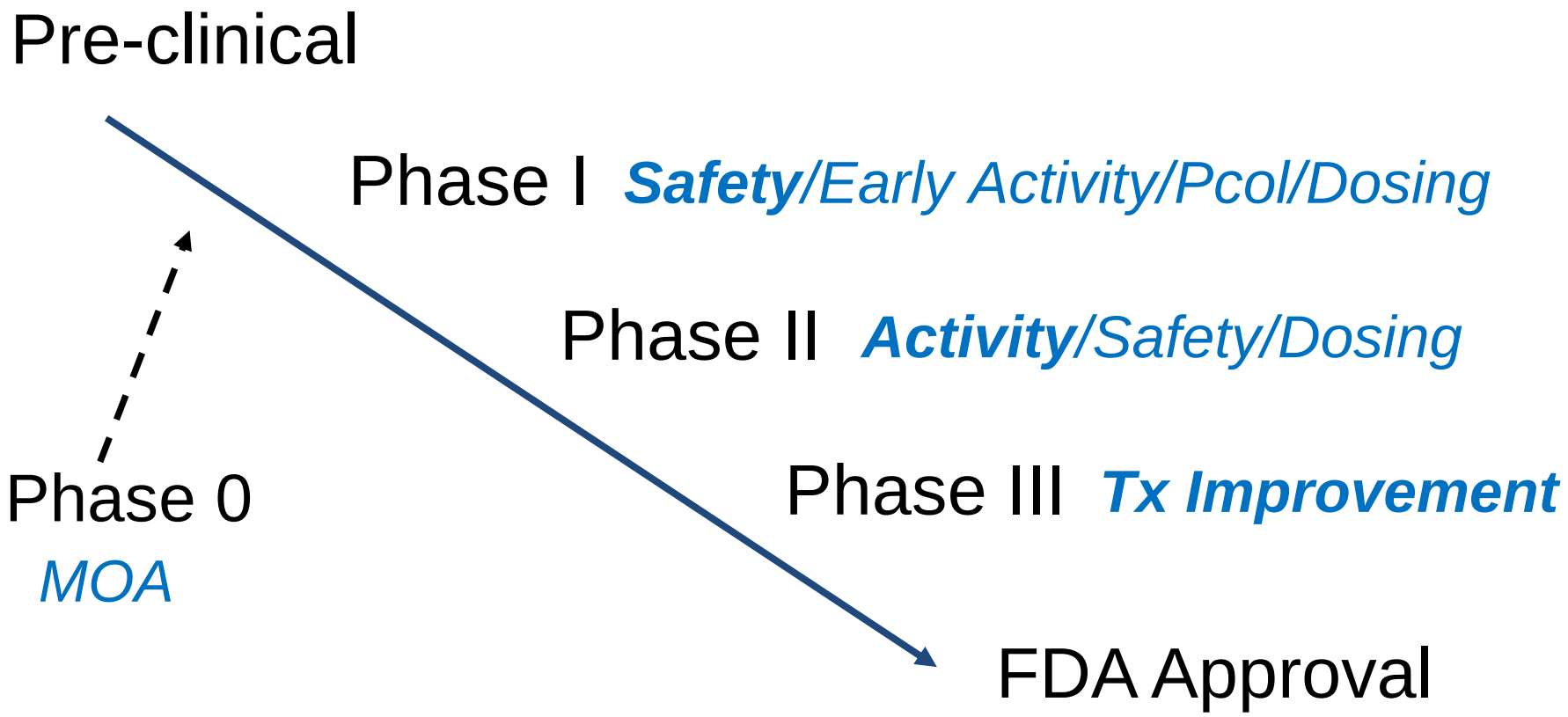
- Community Trial (on communities)

Clinical Trial Center
Tehran University of Medical Sciences

Phases of Drug Development

	Phase 1	Phase 2	Phase 3	Phase 4
No. of Participants	15-30	<100	100 to thousands	Several hundreds to several thousands
Purpose	First in humans Find safe dose	Determine efficacy	Compare new agent with standard treatment	Post –market Long-term safety and efficacy

Overview of Clinical Drug Development



Typical Study Design Features

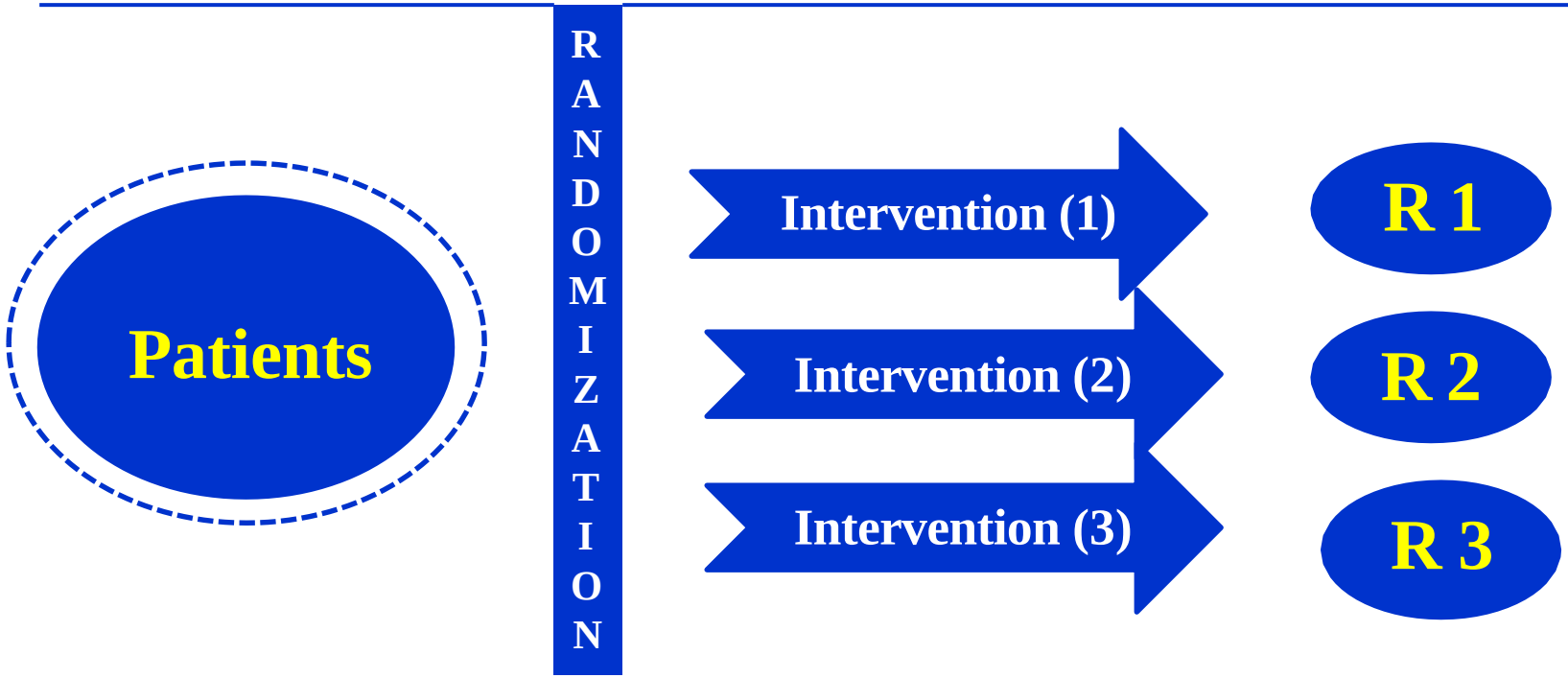
- Treatment sequences
  - e.g. single, parallel, crossover, withdraw, survival
- Blinding/masking
  - e.g. open label, single blind, double blind, double dummy
- Control
  - dose response, active, placebo
- Methods of assigning treatment
  - e.g. randomization +/- stratification

SINGLE ARM TRIALS

- Mostly in phase II clinical trials

PARALLEL GROUP DESIGNS

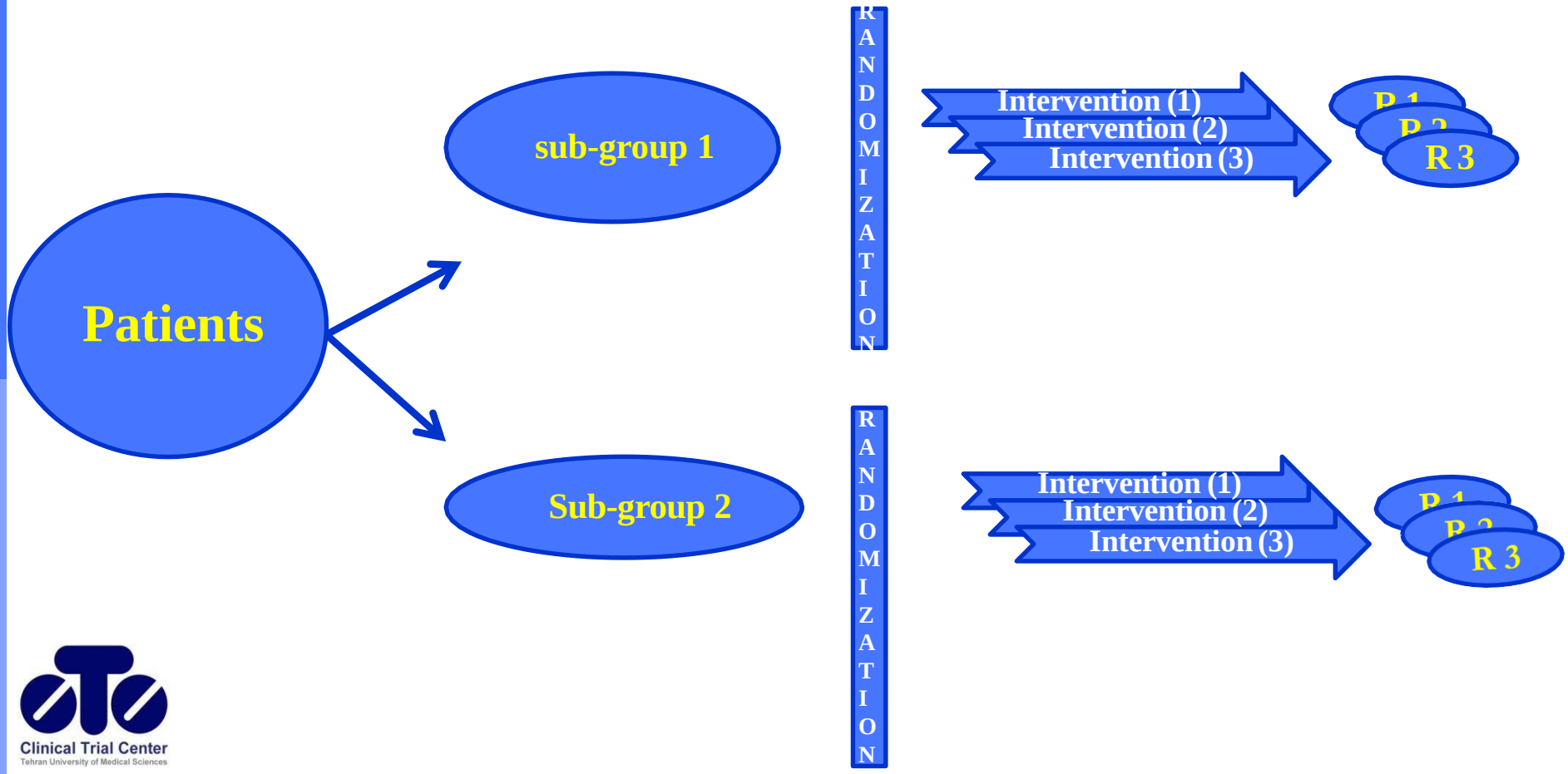
"gold-standard" of clinical research.



- There are as many groups as study treatments under comparison.
- Each patient is assigned to only one of the treatment groups through randomization.
- All treatment groups are treated and evaluated simultaneously

Parallel Group Design(Cont.)(Stratified Design)

- with a stratified randomization considering some prognostic factors as sub-experimental factors.

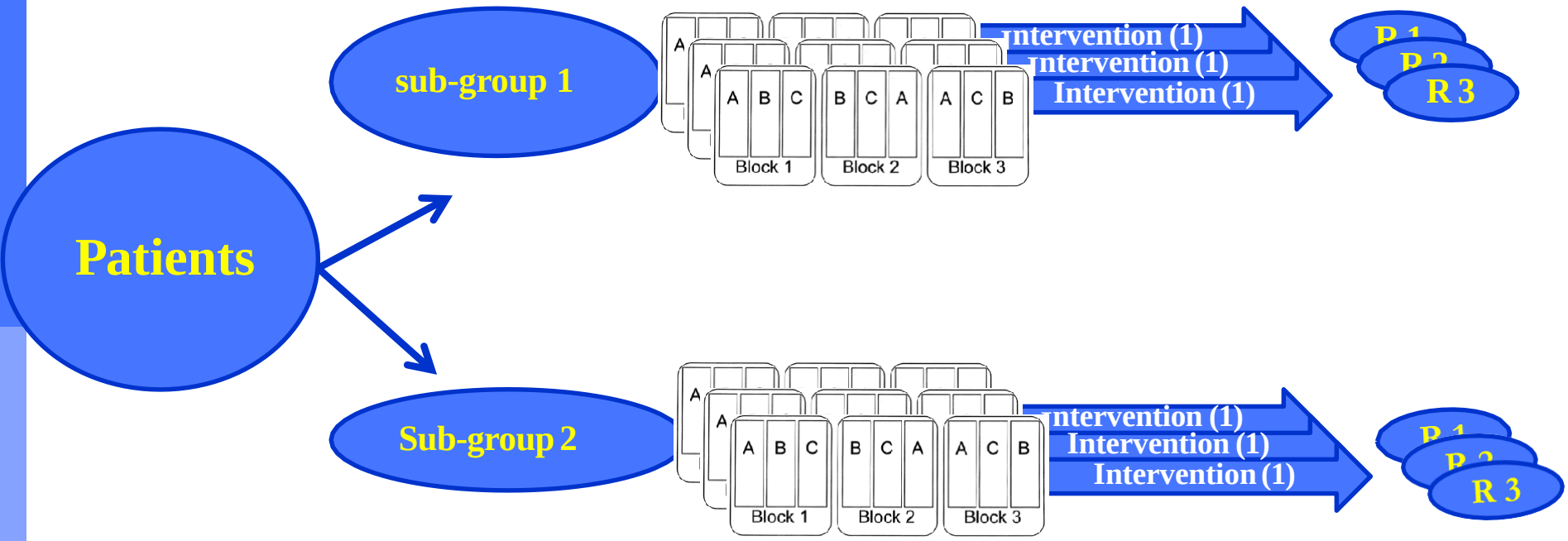


Clinical Trial Center
Tehran University of Medical Sciences

Parallel Group Design(Cont.)

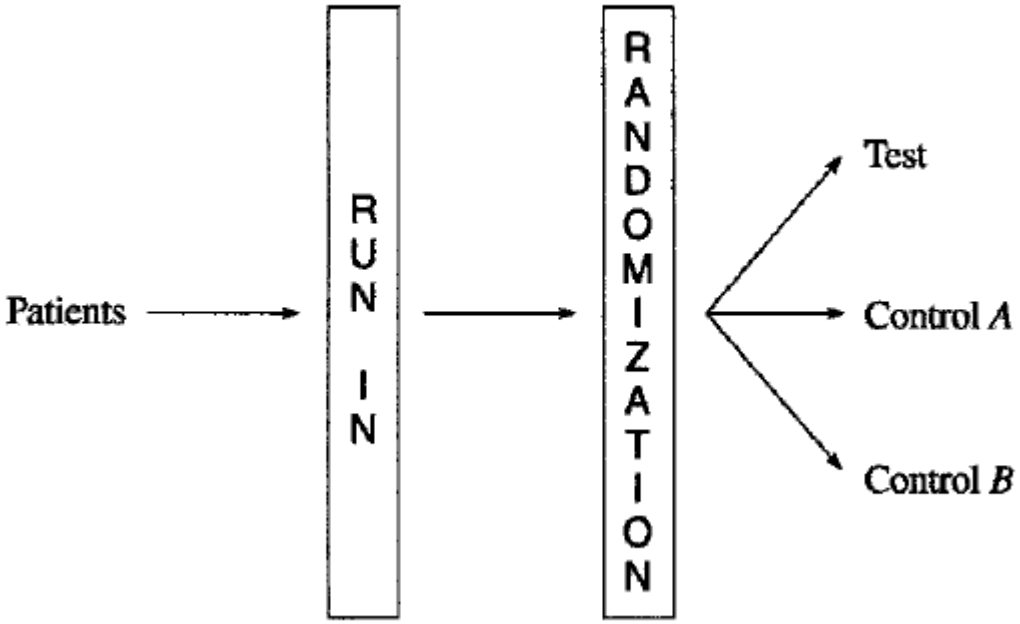
(Randomized Block Design)

"Matched" PARALLEL GROUP DESIGN



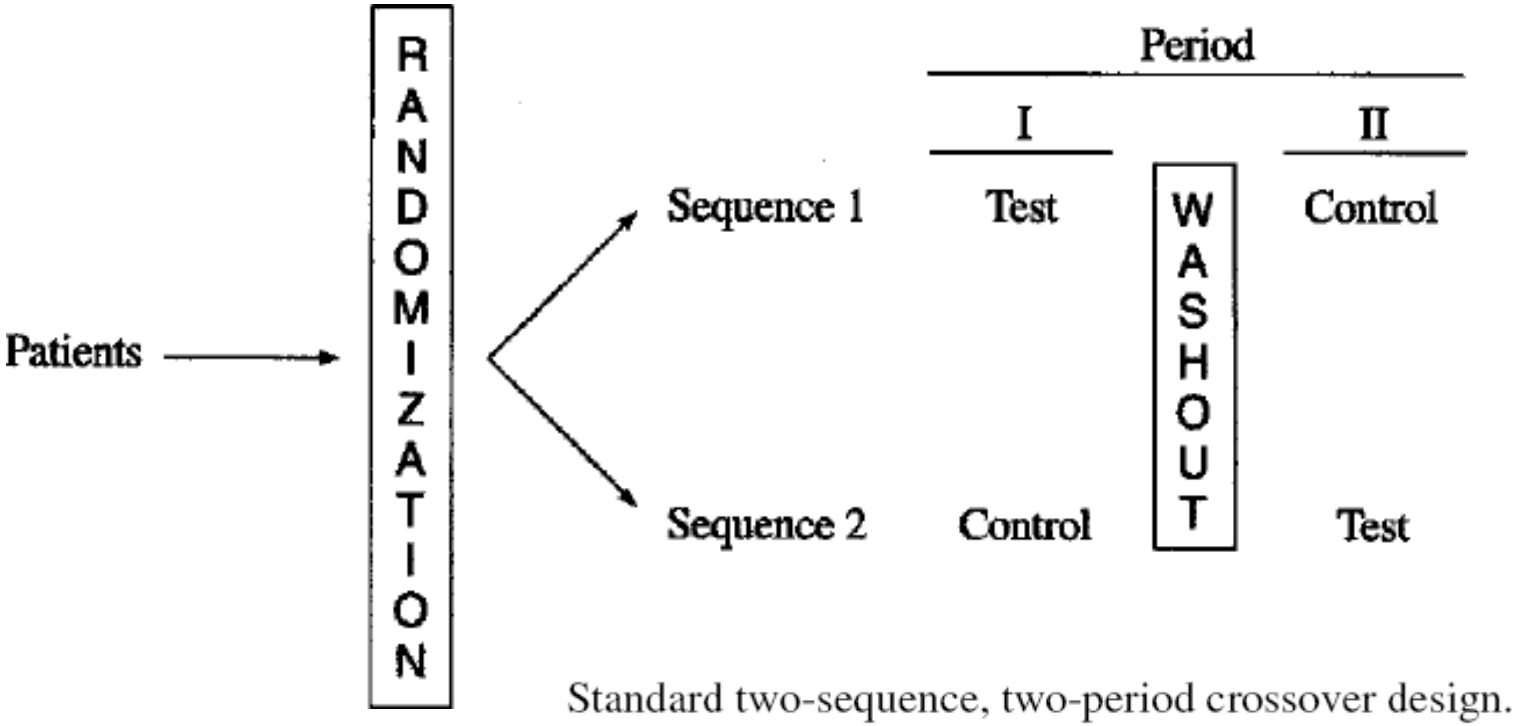
Clinical Trial Center
Tehran University of Medical Sciences

Run-in Period



- Before patients enter a clinical trial, a run-in-period of placebo, no active treatment, dietary control, or active maintenance therapy is usually employed prior to randomization.
- A run-in period is usually employed based on a single-blind fashion.

CROSSOVER DESIGNS



Standard two-sequence, two-period crossover design.

- **Sequence (Period) Effect?**
- **Carryover Effect?**

Clinical Trial Center
Tehran University of Medical Sciences

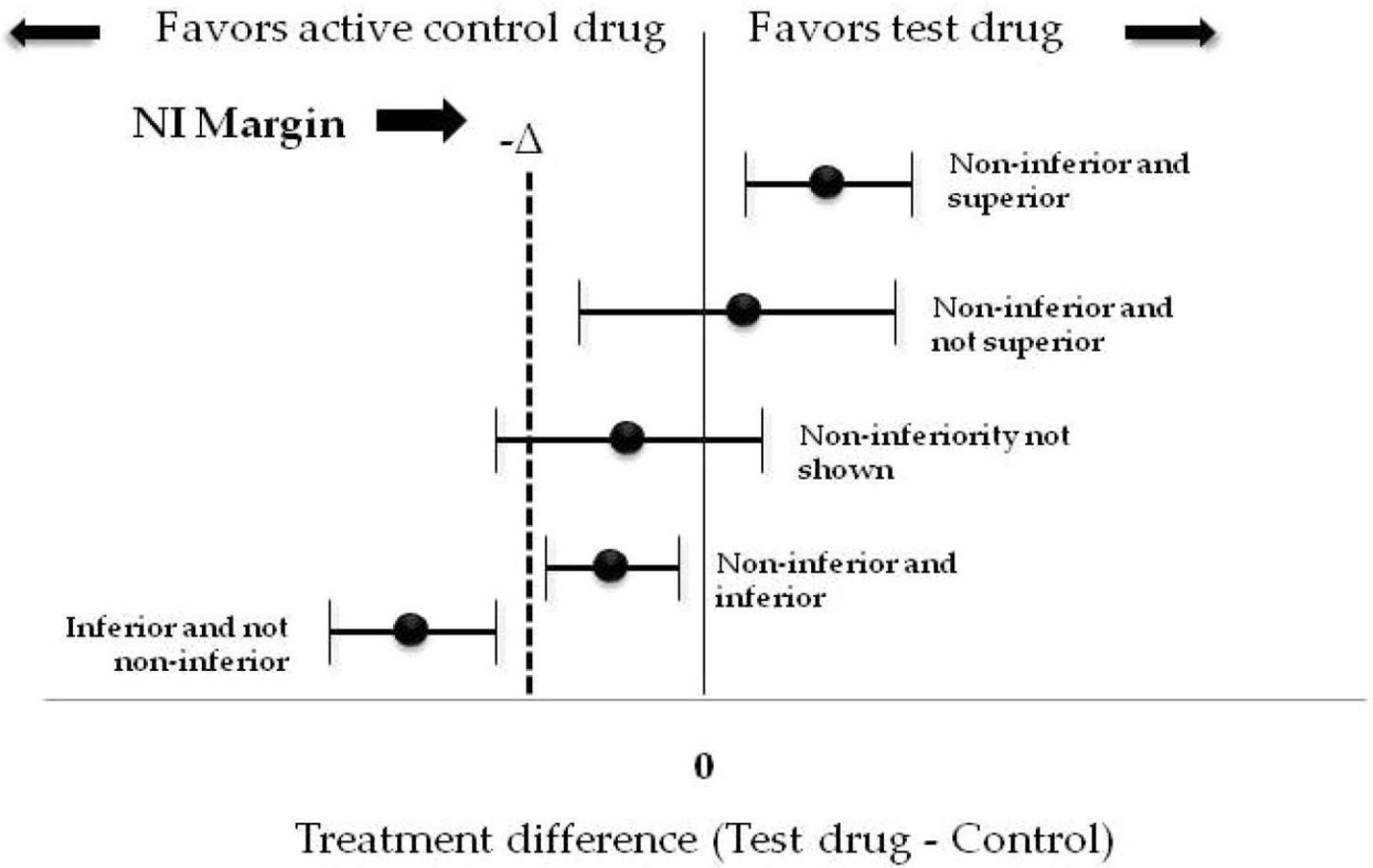
Equivalence/Non-inferiority Vs. Superiority

- Sometimes, the goal is not to show that the new treatment is better, but that the new treatment is 'equivalent' to the control.

- If the CI lies strictly within [-Δ, +Δ]the two treatments are called 'equivalent.' But the amount of Δis more important in equivalency/non-inferiority than superiority.

- Non-inferiority is different from equivalence. In an equivalence trial, the desired conclusion is that two products are the same or 'not unacceptably different' from each other. In a non-inferiority trial, by contrast, the aim is to show that a new product is not unacceptably worse than an older one.

Equivalence/Non-inferiority Vs. Superiority



Meta analysis

Why are Systematic Reviews Necessary?

- The large amount of medical literature requires clinicians and researchers alike to rely on systematic reviews in order to make an informed decision.

- Systematic Reviews minimize bias. “A systematic review is a more scientific method of summarizing literature because specific protocols are used to determine which studies will be included in the review.”

Why are Systematic Reviews Necessary?

- “The volume of published material makes it impractical for an individual clinician to remain up to date on a variety of common conditions. This is further complicated when individual studies report conflicting conclusions, a problem that is prevalent when small patient samples and retrospective designs are used.

Characteristics of Systematic Reviews

- Two possible approaches:
  Or qualitative synthesis

  statistical synthesis of data (meta-analysis) if appropriate and possible

Literature Search

- List of popular databases to search
  - Pubmed/Medline
  - Embase
  - Cochrane Review
  - ISI Web of Science
  - SCOPUS

- https://www.riskofbias.info/
- https://www.cochranelibrary.com/
- https://www.crd.york.ac.uk/prospero/