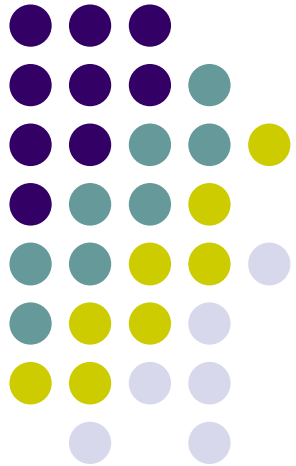
Study Design

Jieyun Yin, Ph.D

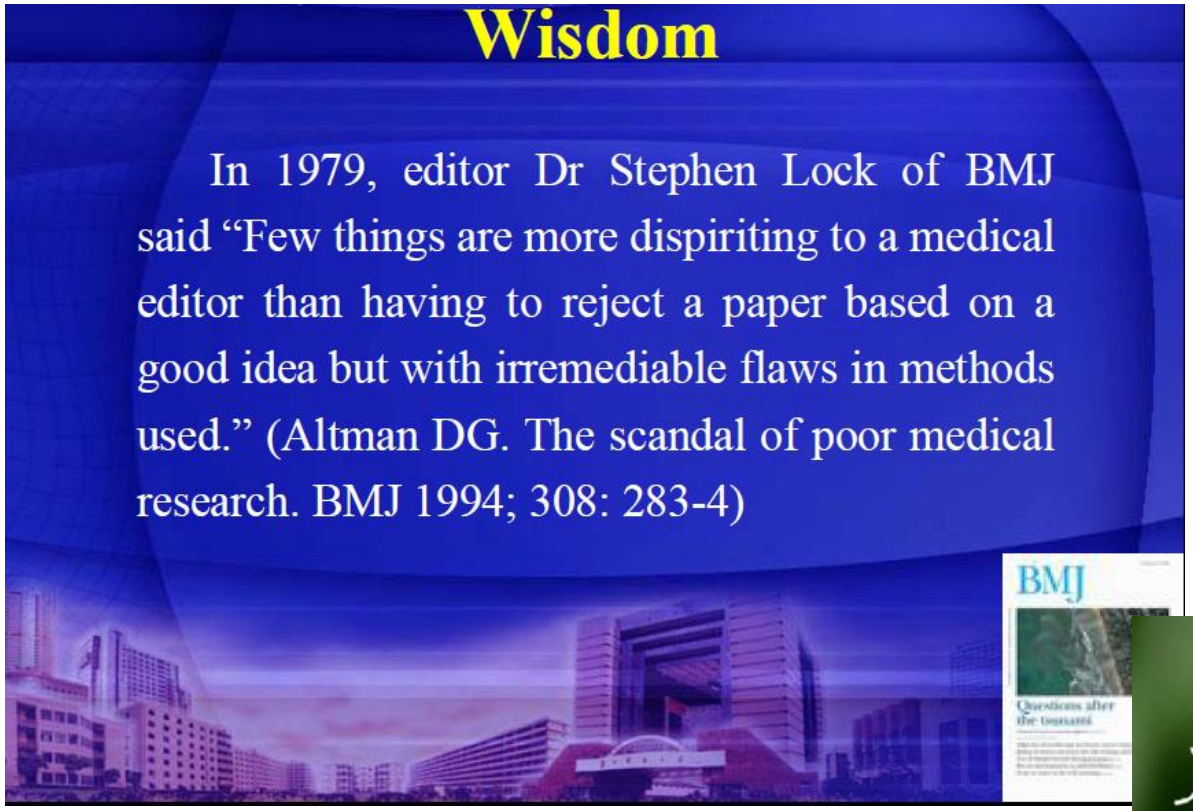
Email: *445229002@qq.com*



Medical statistics

- Medical statistics is the science of collecting, summarizing, presenting and interpreting data, and of using them to estimate the magnitude of associations and test hypothesis in the medical fields.

Make your whole year's plans in spring

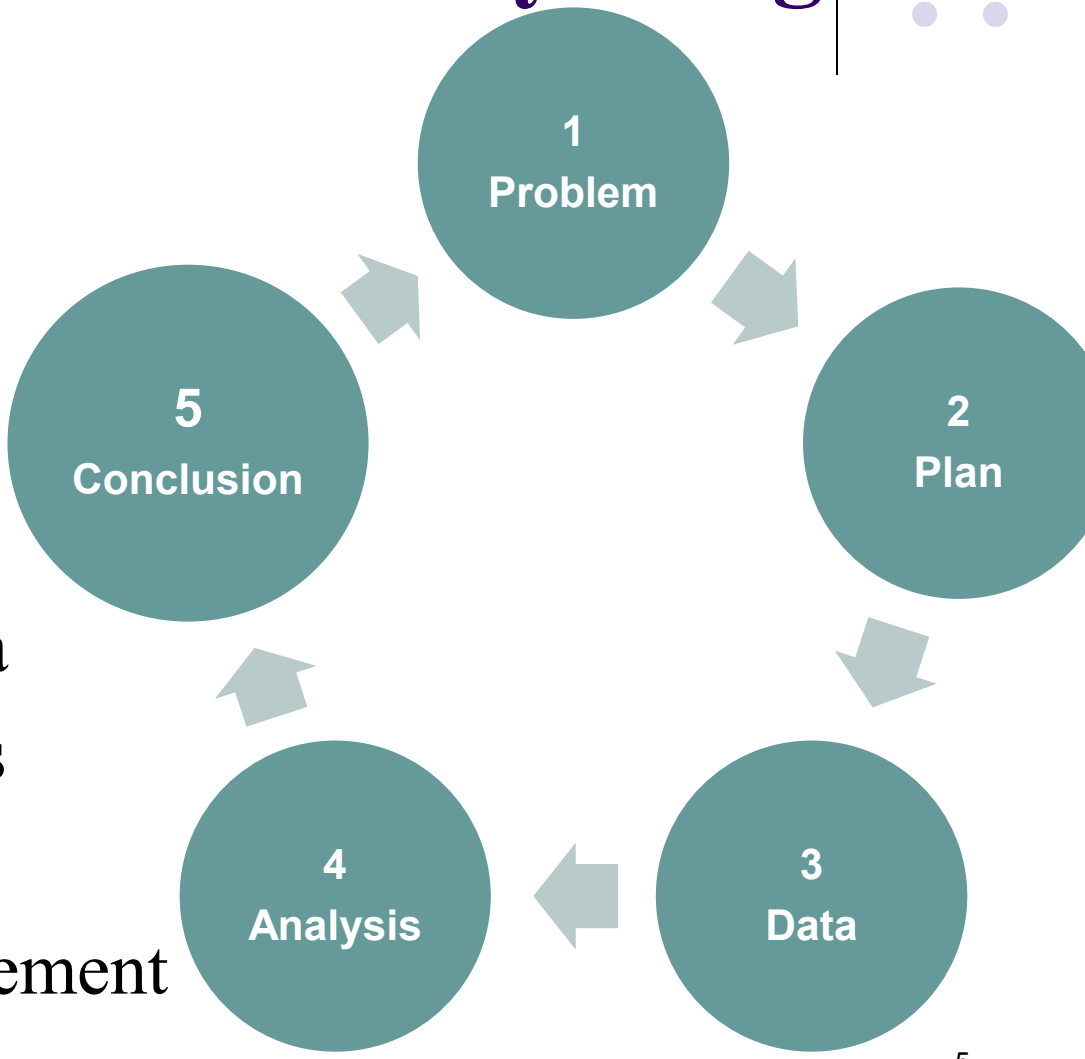


Contents

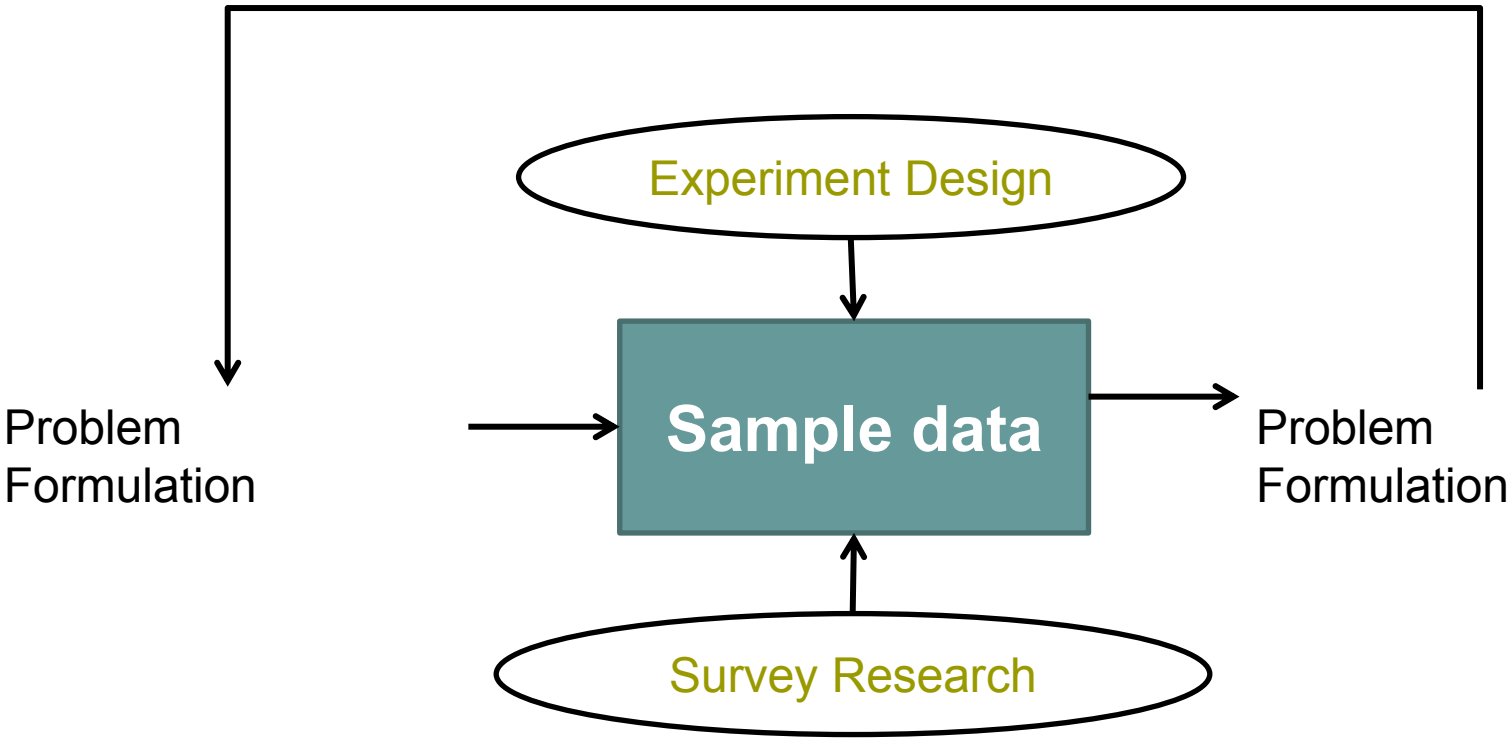
1. Introduction to the study design
2. Survey study
3. Experiment study
4. Basic principles of experiment study
5. Common Types of study design

1. Introduction to the study design

- Including
  research design
  professional design
- Should consider
  - Define subjects
  - Way of getting the data
  - Data input and analysis
  - Presenting
  - Time and funds arrangement

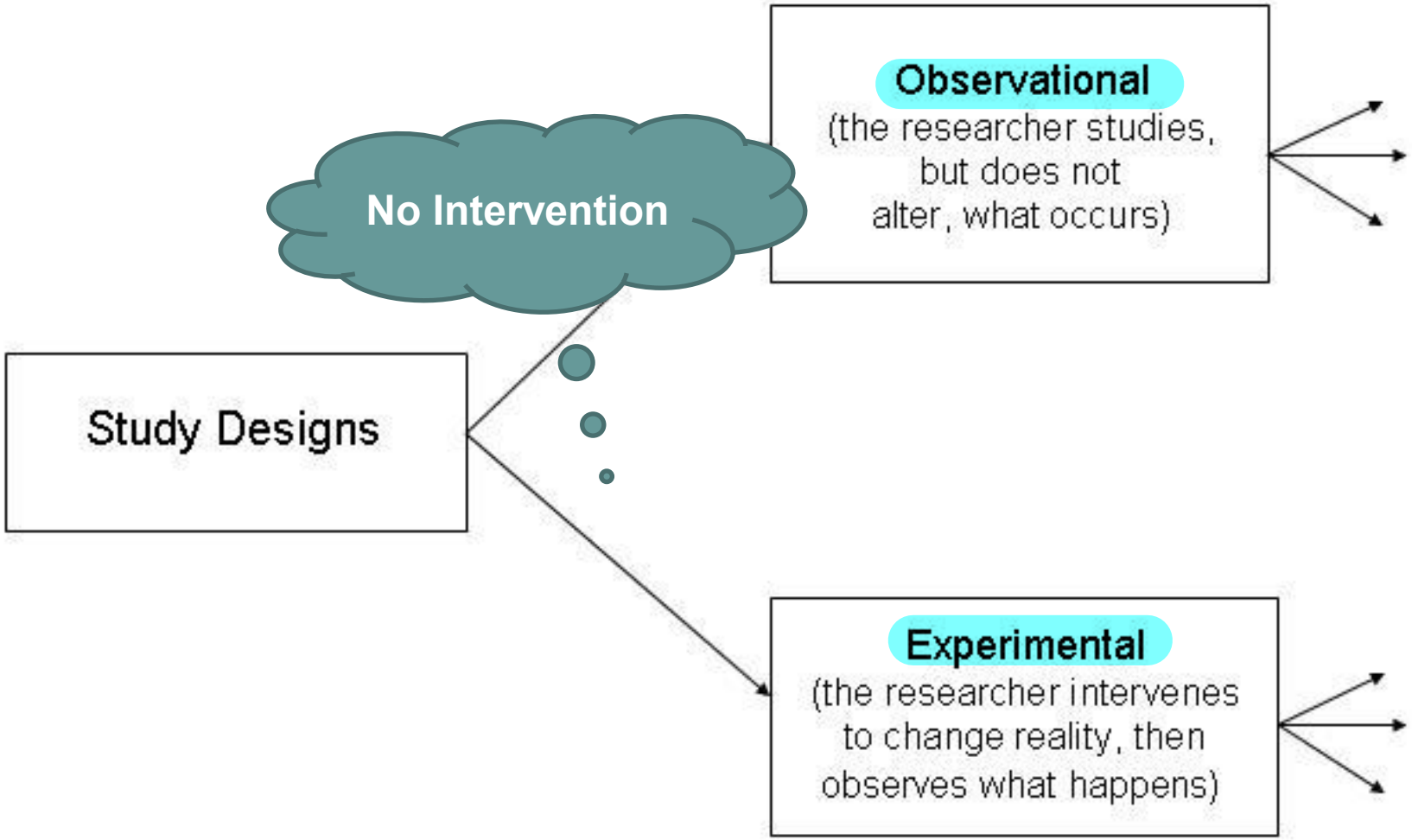


2. Types of Study Design

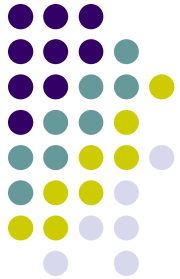


Also called investigation, observational study

Types of Study Design



Types of Study Design (Terms)

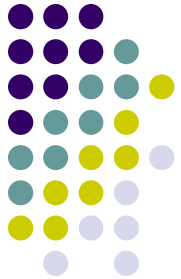


- **Experiment study**: is characterized with adding experiment or treatment factor on the subject.

- **Survey study**: is only to carry out an investigation or personal interview on the subject. It does not add the treatment on the subject just the passive collection of the regarding information about the subject.

- **Observational studies**: the researcher studies, but does not (or cannot) alter, what occurs. Observational studies are used in describing the health status of a population, recording indicators of morbidity and health. This is used in planning health services, developing health policies, etc. Observational studies are also widely used in research on the causes of human disease when it would be unethical to undertake an experiment.
- **Experimental studies**: the researcher intervenes to change reality, then records the results. Experiments form the mainstay of clinical research on the effectiveness of therapies, using the randomized clinical trial.

Examples of observational studies



- a survey of drinking habits among students;
- a researcher who joins a biker gang to study their lifestyle (note, as long as the researcher does not try to change their behavior, it's an observational study);
- taking blood samples to measure blood alcohol levels during Monday morning lectures (yes, you are intervening to take the blood, but you are not trying to change the blood alcohol level: it's just a measurement).

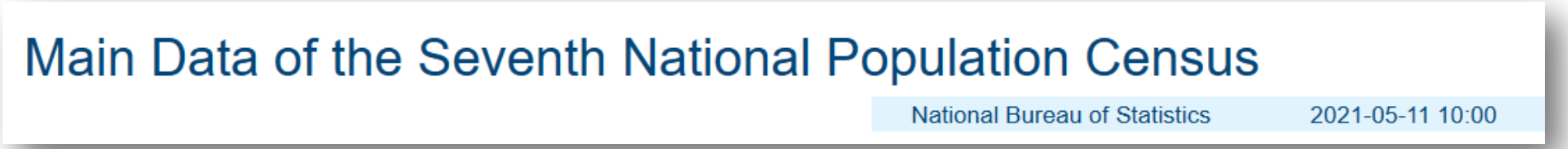
Examples of experiments

- plying a law student with beer to see whether lawyers argue better when drunk;
- encouraging bikers in one group to stop smoking those funny-looking cigarettes to see whether they get less belligerent;
- warning one group of students that you are going to take blood alcohol levels next Monday to test for alcohol, and comparing their levels to another group that you did not warn.

3. Survey Study

I. In term of the range of survey, it includes three types

- Census: is a survey involving all subjects (population) whom you want to investigate on
- **Sampling survey**: obtain the sample from the population at random (most common)
- Typical survey: carry out a intensive survey for a serious public event such as SARS, bird flu cases and Ebola virus cases.



Survey Study

2. In term of the study time

- Retrospective study (case control study)
- Cross-sectional study
- Perspective study-cohort study

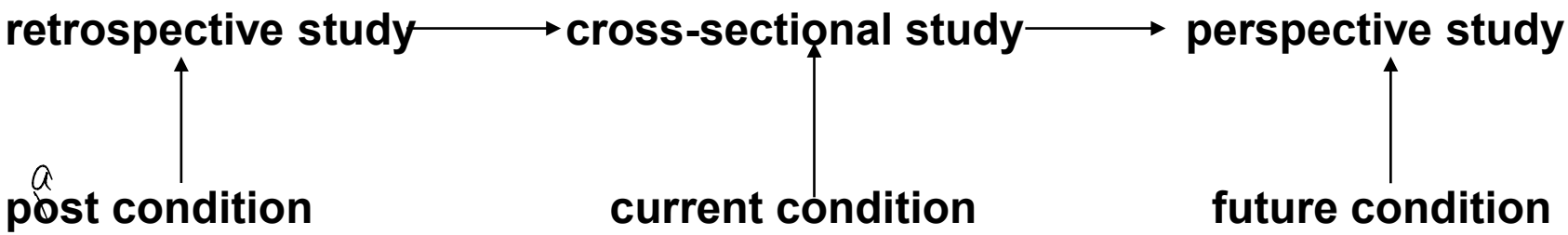
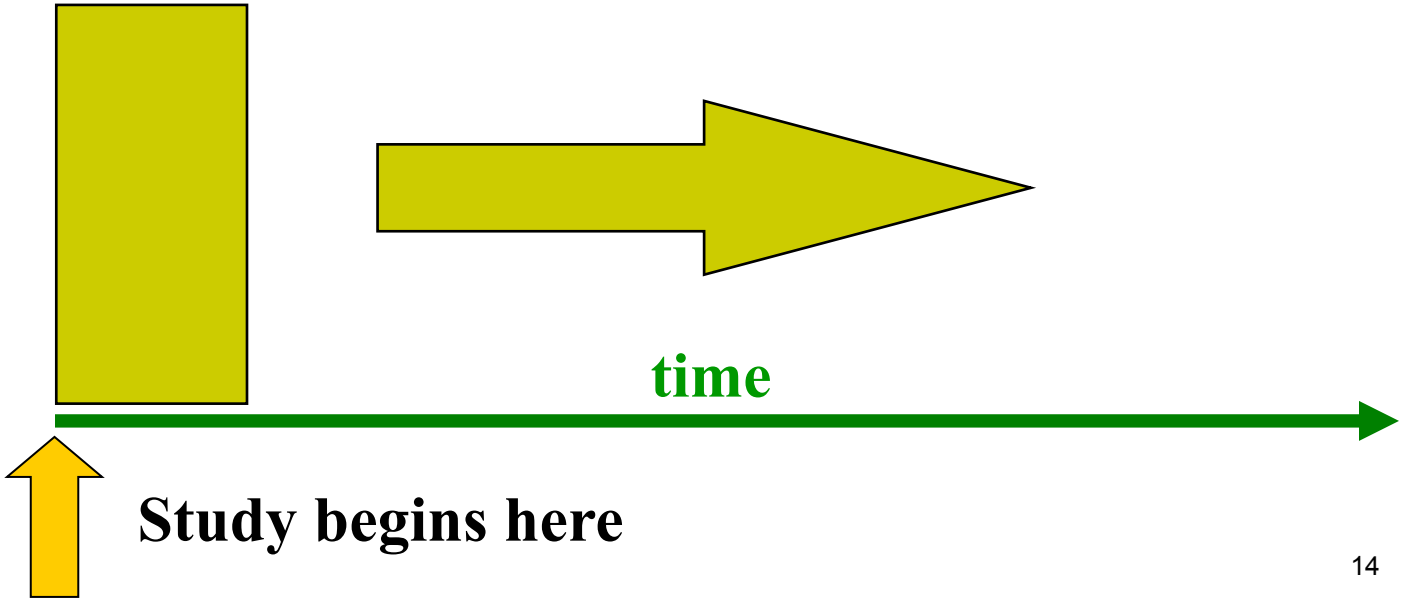


Fig. 1 Illustration of survey study

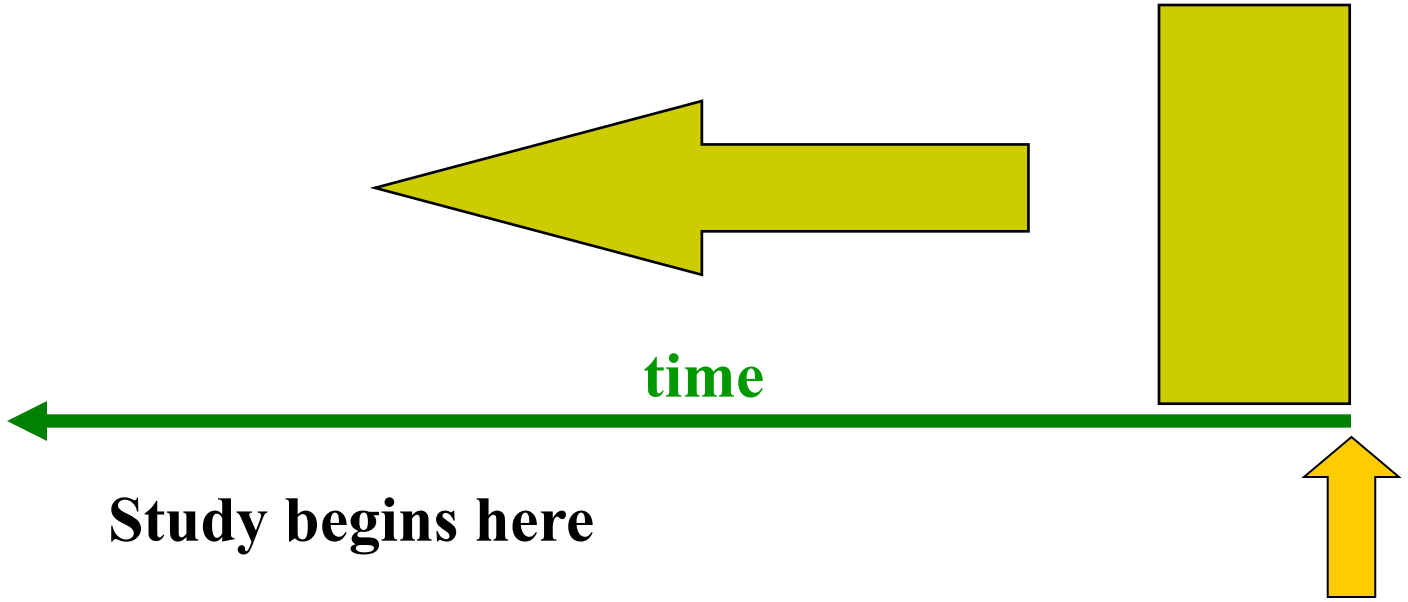
Timeframe of Studies

- **Prospective Study - looks forward, looks to the future, examines future events, follows a condition, concern or disease into the future.**



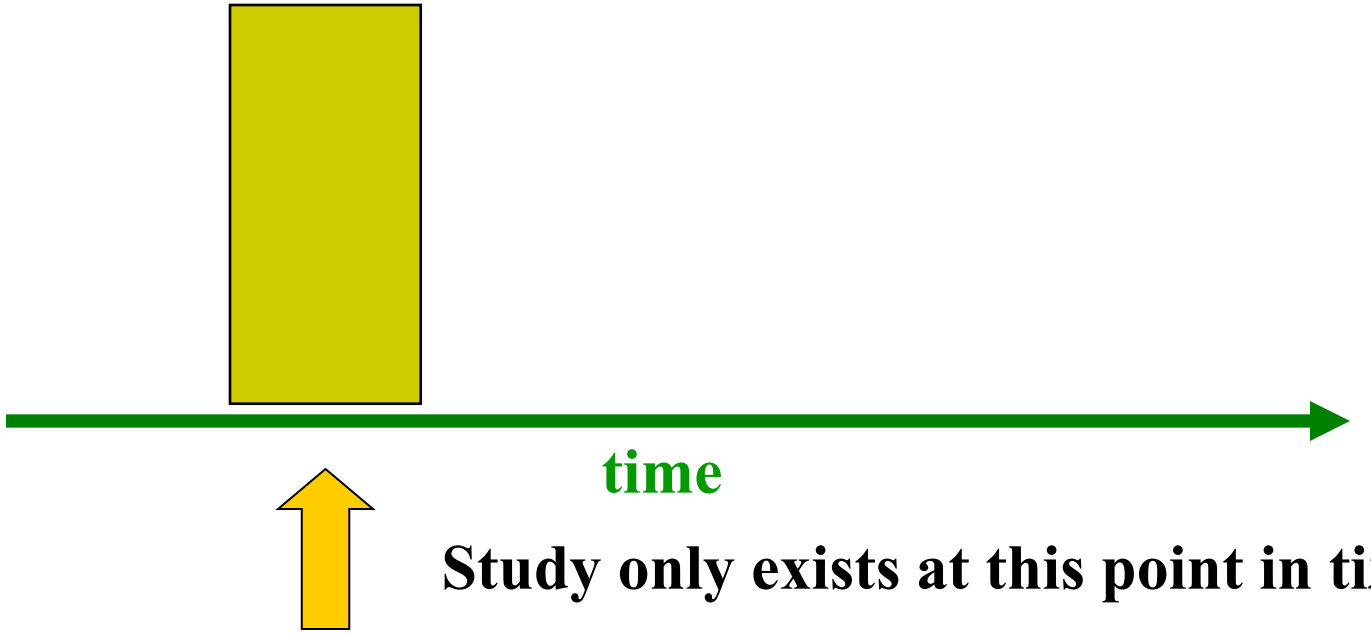
Timeframe of Studies

- **Retrospective Study - "to look back", looks back in time to study events that have already occurred**

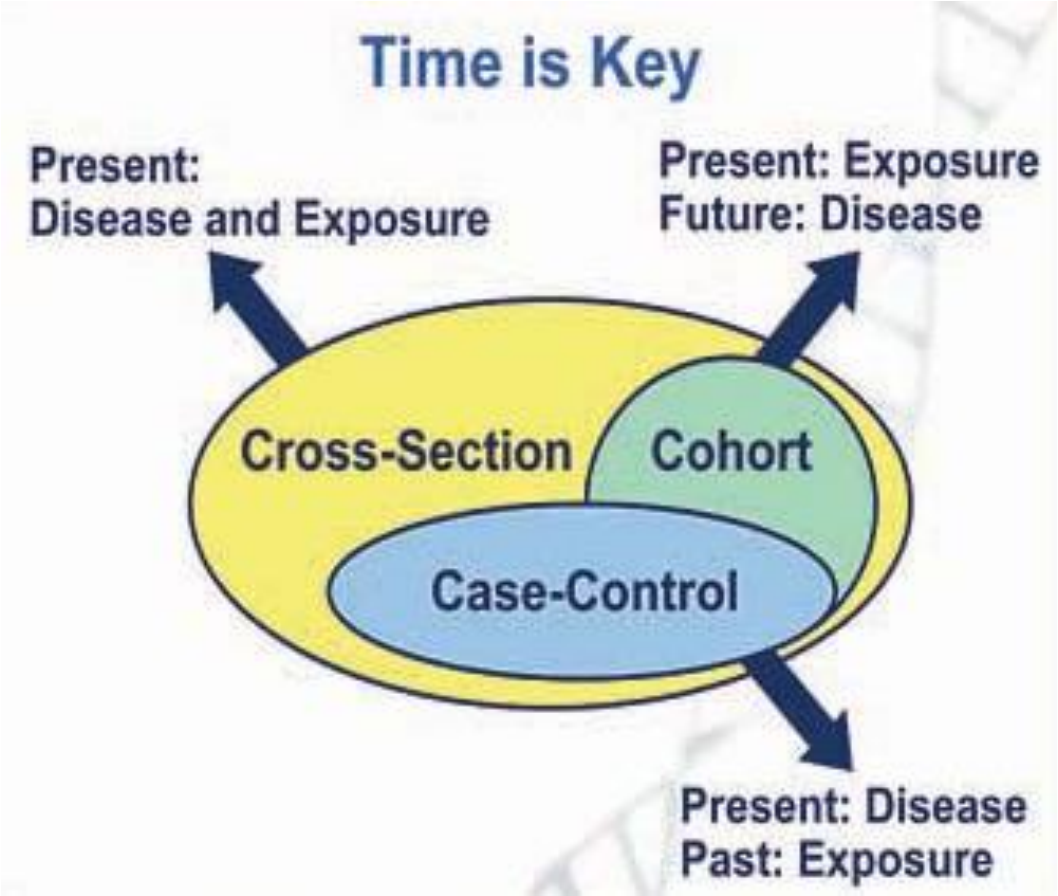


Cross-sectional studies

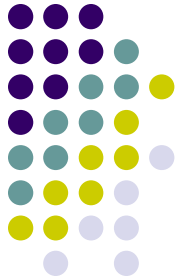
- An “observational” design that surveys exposures and disease status at a single point in time (a cross-section of the population)



Observational Studies and Timeframe



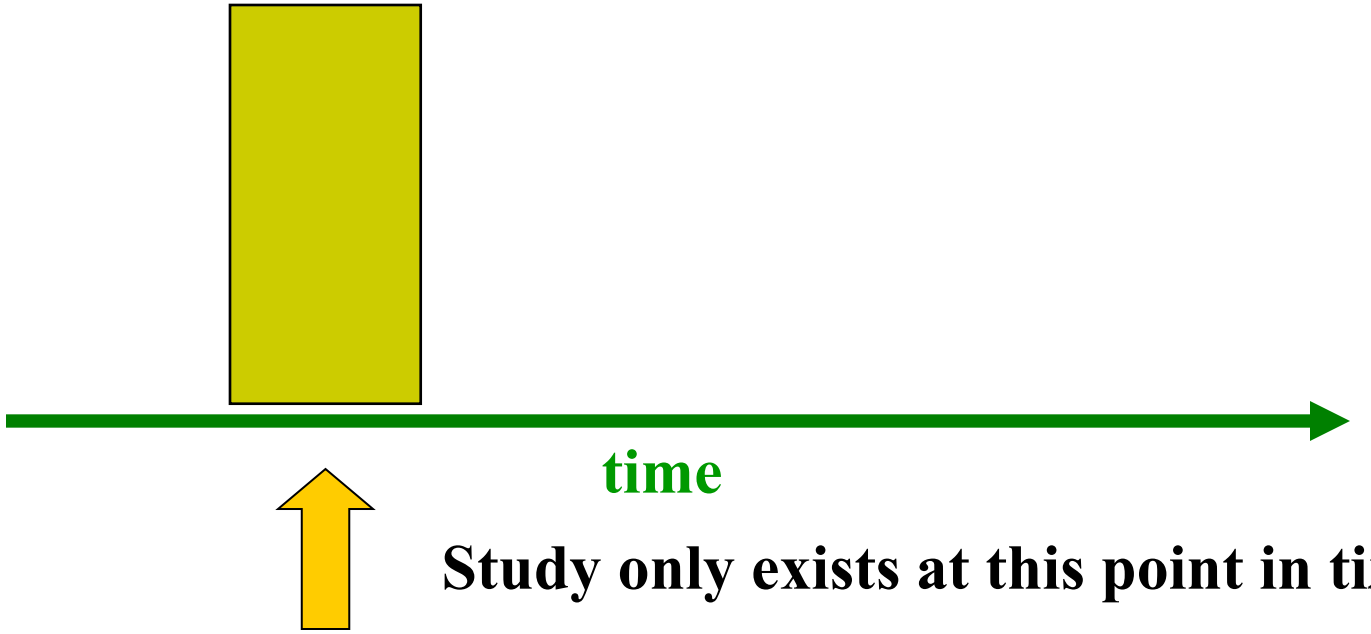
3.3 Cross-sectional study



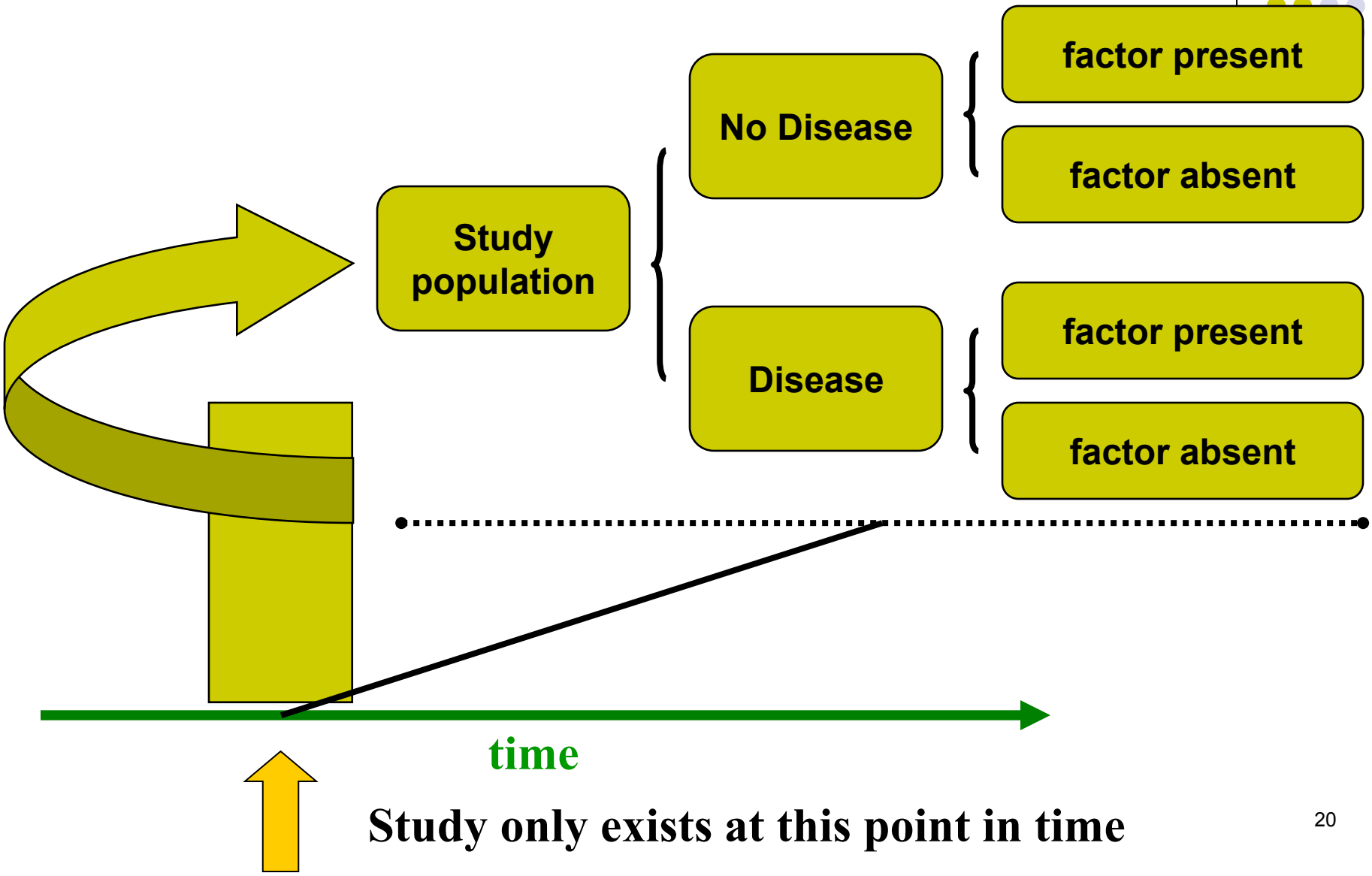
- Only investigate the current information of a sample
- The subjects in the sample automatically are divided into 2 groups, case and control according the definition for the 'case'
- Use OR as an index to estimate the relationship between the factor and disease
- Biggest shortcoming is that even if we find the association between the factor and disease, we do not know which is the cause, or which is the result.

Cross-sectional studies

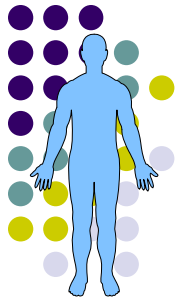
- An "observational" design that surveys exposures and disease status at a single point in time (a cross-section of the population)



Cross-sectional Design

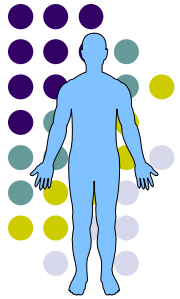


Cross-sectional Studies



- Often used to study conditions that are relatively frequent with long duration of expression (nonfatal, chronic conditions)
- It measures prevalence, not incidence of disease
- Example: community surveys
- Not suitable for studying rare or highly fatal diseases or a disease with short duration of expression

Cross-sectional studies



- **Disadvantages**
  - **Weakest observational design, (it measures prevalence, not incidence of disease). Prevalent cases are survivors**
  - **The temporal sequence of exposure and effect may be difficult or impossible to determine**
  - **Usually don't know when disease occurred**
  - **Rare events a problem. Quickly emerging diseases a problem**

4. Experiment Study

The researcher intervenes to change reality, then records the results. Experiments form the mainstay of clinical research on the effectiveness of therapies.

Components of experiment study

- An experiment study consists of three essential components:
- experiment subject
- experiment treatment
- effective index/outcome

Experiment subject

- Animals
- Human

Requires them homogenous!

Experiment subject-animals

- Can be performed under strict conditions

The mice were synchronized to standard lighting conditions of 12 hours of light followed by 12 hours of darkness for 2–3 weeks and then were randomly assigned either to remain in this lighting regimen (group A) or to undergo experimental chronic jet lag produced by 10 days of serial 8-hour advances of light–dark cycle every 2 days (group B). In experiment 2 only, a third group of mice (group C) underwent chronic jet lag and concurrent meal timing from circadian time (CT) 12, ie, subjective time of onset of darkness (activity) to CT 24, i.e., subjective time of onset of light (rest). Meal timing was shown to entrain circadian expression of clock genes in the liver of mice with ablated suprachiasmatic nuclei *(27),* a condition that also disrupts circadian rhythms of activity and temperature *(27).*There were 60 control

Experiment subject-animals

- Can be treat with toxic and harmful substance
- Have relatively short lifespan and gestational period
- Can obtain tissues directly after dissection
- extrapolate results of animal to humans should be cautious

Ethical approval!

Experiment subject-human

- Patients
  - ✓ Outpatient
  - ✓ Inpatient

Different in severity of illness, compliance

- Healthy individuals **informed consent**
- Volunteer: bias

informed consent ! 许可 approval

Ethical approval!

Retracted article

See the retraction notice

Randomized Controlled Trial > Am J Clin Nutr. 2012 Sep;96(3):658-64.

doi: 10.3945/ajcn.111.030775. Epub 2012 Aug 1.

β-Carotene in Golden Rice is as good as β-carotene in oil at providing vitamin A to children

Guangwen Tang [1], Yuming Hu, Shi-an Yin, Yin Wang, Gerard E Dallal, Michael A Grusak, Robert M Russell

Affiliations – collapse

Affiliation

1 Carotenoids & Health Laboratory, USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA 02111, USA. guangwen.tang@tufts.edu

Experiment subject-human

- Sensitive to intervention
- ✓ i.e., Resistant hypertension

Resistant hypertension is defined as a blood pressure that remains above goal despite concurrent use of three antihypertensive agents of different classes taken at maximally tolerated doses, one of which should be a diuretic (the diuretic should be selected based upon kidney function)

Experiment Treatment

- Refers to the major treatment or intervention factor which we are concerned with
- Non-treatment factor means confounding factors which will make effect on our study outcome, but it is not our main study factor.

Effective Index/Outcome

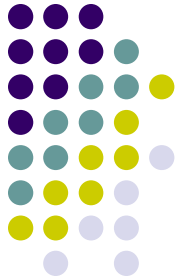
- Precise
- Accurate
- Objective

Measured by instruments as far as possible

5. Principles of experiment study

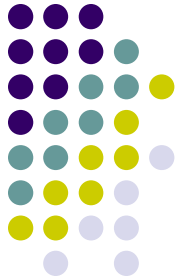
- ✓ Three rules
  - Randomization
  - Control
  - Replication

5.1 Randomization



- To balance the confounding factors between the two groups.
- An important methodology is randomization, which means that let the study subjects have the same chance to enter the experiment group or control.

randomization



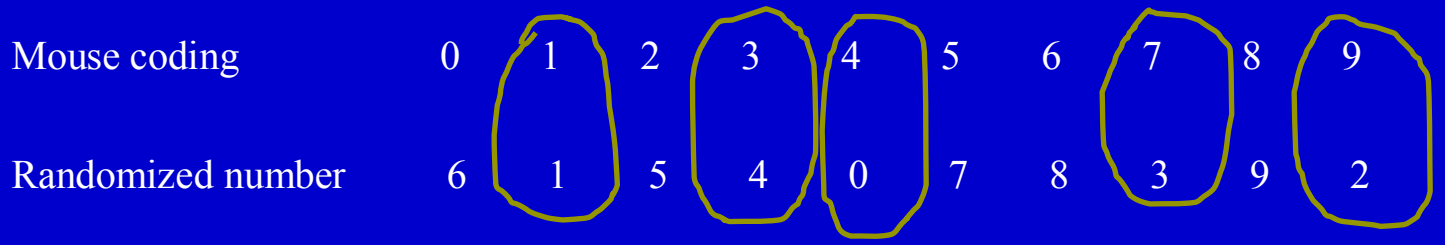
- randomization means randomly allocating experimental units to treatment groups. (Or, equivalently, randomly allocating treatments to experimental units.)
- Also means random sampling

Randomization?

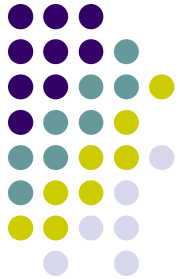
Example 1

- Now we have 10 white mice, how do you assign them on average to the two groups using completely randomized design
- Drawing straw :firstly, coding the 10 mice as 0 to 9 by writing number on the back of mouse using the color pen. Then write down 0-9 number on the strip papers, respectively. Curl it up and put them in a bag or an envelope. Sway it and make the lots mix round. We take out the first 5 lots and regard them as the experiment group, the remains will be enter into control.

- Randomized number table :After coding, we excerpt a series of randomized number (0~9) in order beginning from 4th row in randomized number table and put them under the mouse coding number (note, cancel the number over 9). If the randomized number is less than 5, we will assign the corresponding mouse into experiment group (1, 3, 4, 7, 9), then the remains will be in the control.



Defining Confounding – The more traditional way



- A confounding variable is a risk factor* for the outcome

AND

- Associated with the exposure

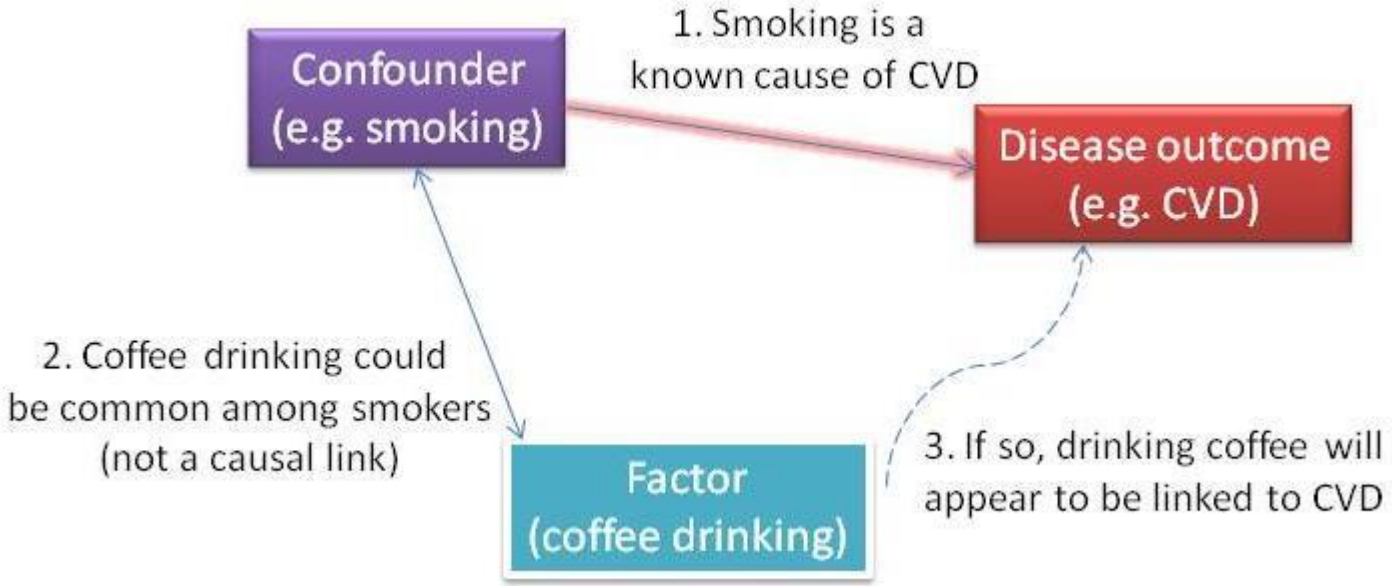
BUT

- Not an intermediate variable in the causal pathway between exposure and outcome

* Causal or a marker (surrogate) of the causal factor

Confounding

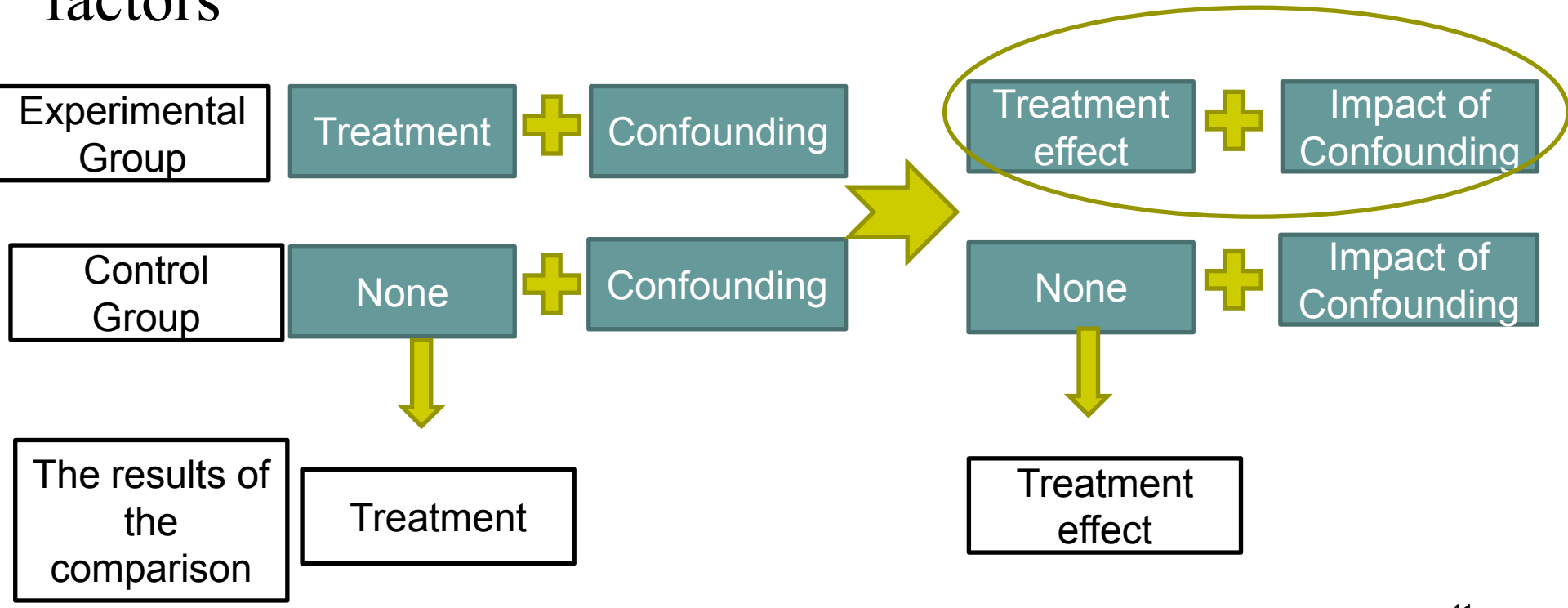
- could coffee be causing the heart disease?



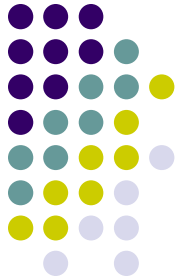
•Confounder is associated with the exposure.

Confounding

- In the experiment study, it needs the control group to balance or counteract the influence of confounding factors

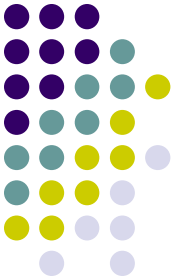


Prevention of Confounding during Analysis Phase



- Stratified Analysis
  - Compare the exposed and unexposed groups (or cases and controls) within homogenous categories of confounding variables. Each stratum provides an unconfounded estimate of the r/ship. An adjusted estimate can then be computed.
- Multivariable Analysis
  - A statistical tool that uses mathematical models to assess and control confounding (logistic, linear, cox proportional hazards)

Prevention of Confounding in the Design Phase



- Control (Matching)
- Randomization
- Replication

5.2 Control = minimizing variability

- As we have seen, in the ideal experiment, everything would be the same in the treatment groups except for the treatment itself.
- Control means trying as hard as you can to make that true: minimizing variability.

5.3 Repetition

- Our study needs a certain number of subjects (sample size) to keep our results stable. So replication refers to having multiple independent experimental units in each treatment group.
- we need to repeat our experiments, because a positive experiment result only from a few subjects may be produce mistake or error.
- Repeated measurement

- 2016年5月2日，韩春雨作为通讯作者在（Nature Biotechnology）发文，他的团队发明了一种新的基因编辑技术——NgAgo-gDNA，是诺贝尔奖的热门；
- 中国生物学界轰动，多家高校邀请讲座；
- 被列入国家“中青年科技创新领军人才”候选人；
- 越来越多的实验室表示无法重复韩春雨所报道的实验结果
- 2017年8月3日，《Nature Biotechnology》发表题为《Time for the data to speak》社论，并宣布撤回韩春雨团队的论文

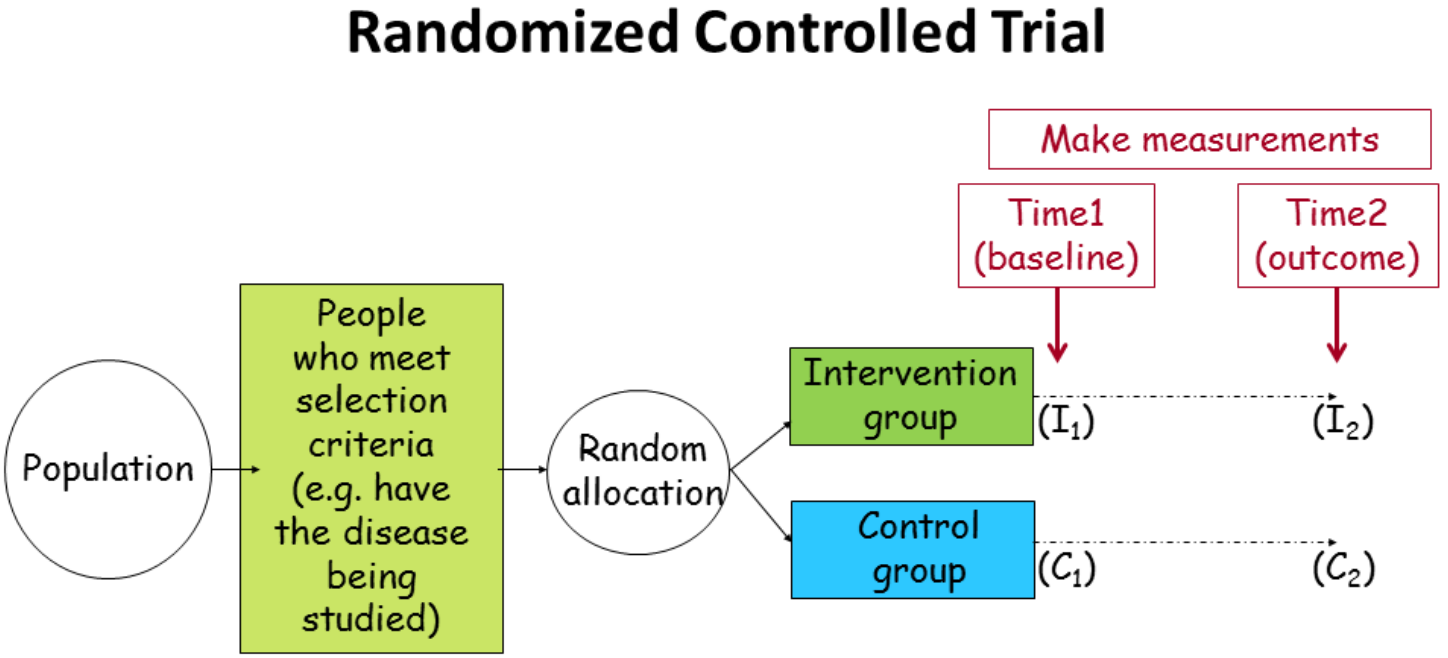
6. Types of experimental design

- **Randomized controlled trial**
- **Completely randomized design**
- **Paired design**
- **Randomized block design**
- Cross-over design
- Factorial design
- Repeated measurement design/within-subject design

Randomized Controlled Trials (RCTs)

the "gold standard" of research designs

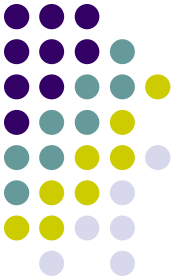
provides most convincing evidence of relationship between exposure and effect



Statistic = difference in outcome scores = $(I_2-I_1) - (C_2-C_1)$
This shows the experimental change minus the change in the control group.

Note: random allocation should make $I_1 = C_1$, but these values are included in analyses to correct for any minor differences.

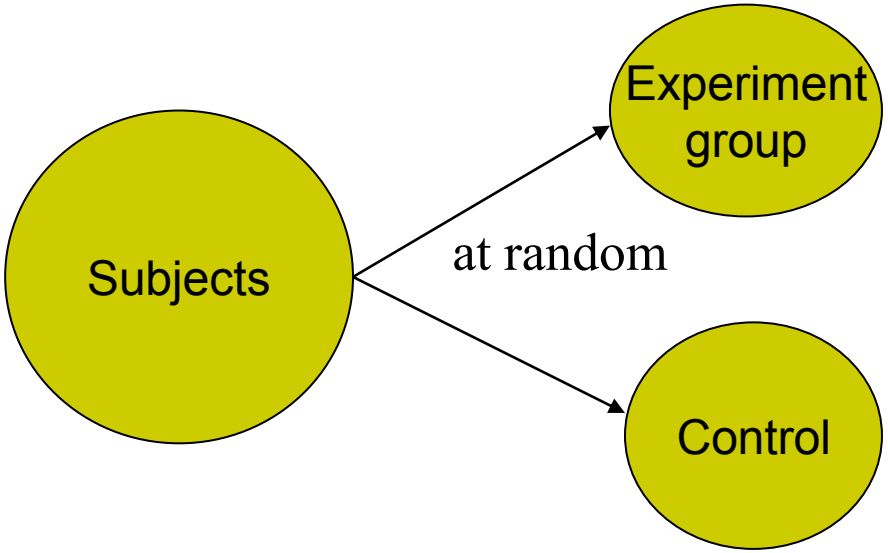
Randomized Controlled Trials



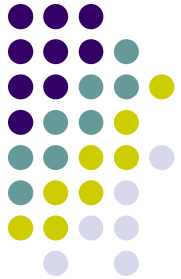
- **Disadvantages**
  - **Very expensive**
  - **Not appropriate to answer certain types of questions**
    - **it may be unethical, for example, to assign persons to certain treatment or comparison groups**

Completely Randomized Design

- It means that we assign the subjects to experiment group or control in random.
- We have concrete methods such as drawing straw or randomized number table.

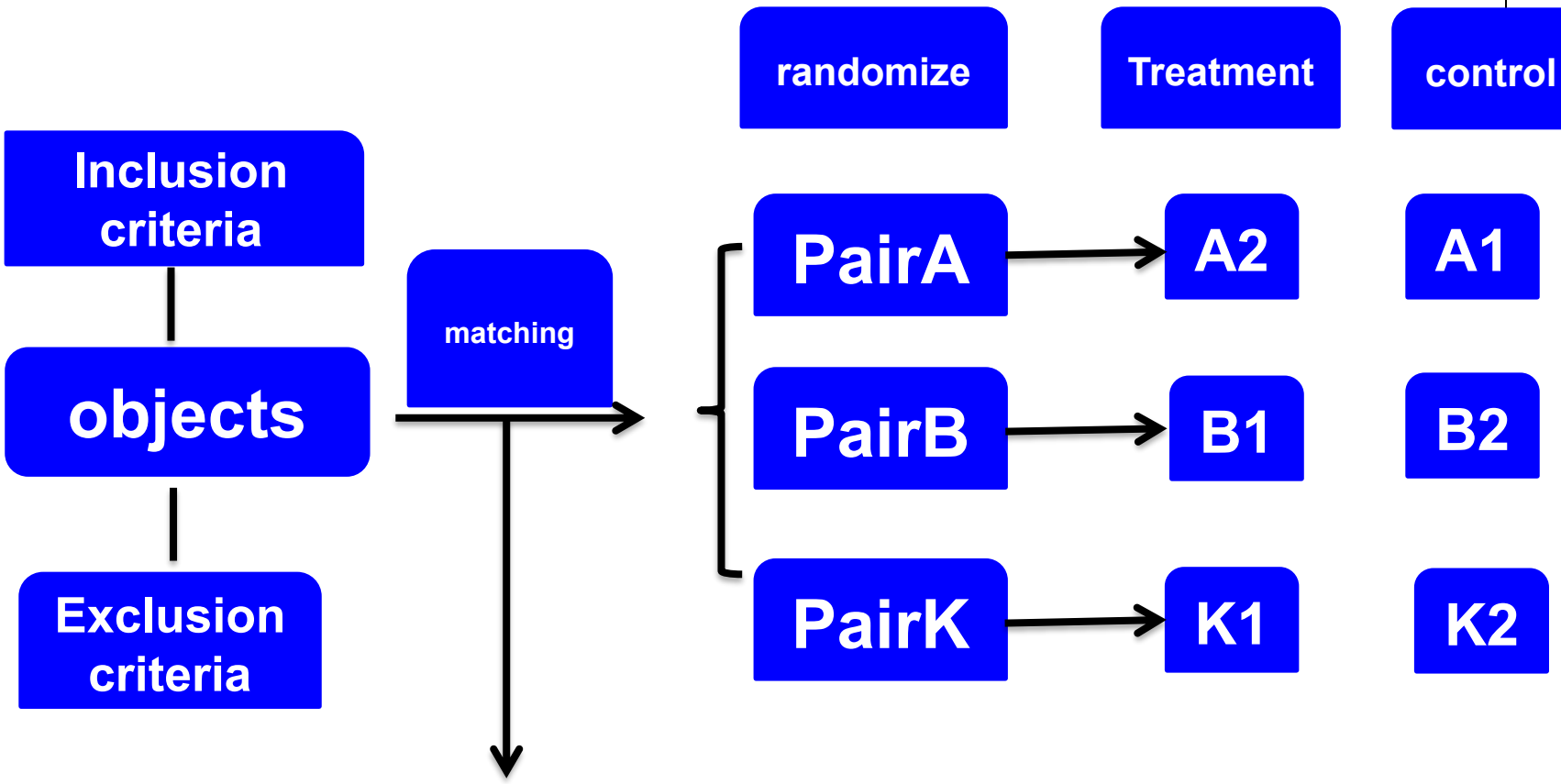


Paired Design1



- Different parts of the same object
- Measured the same objects at two different time points
- Measured the same objects with two different methods

Paired Design2



Matching condition：sex、age、BMI and etc

advantage：smaller standard error, higher statistical efficiency and smaller sample size required than CRS

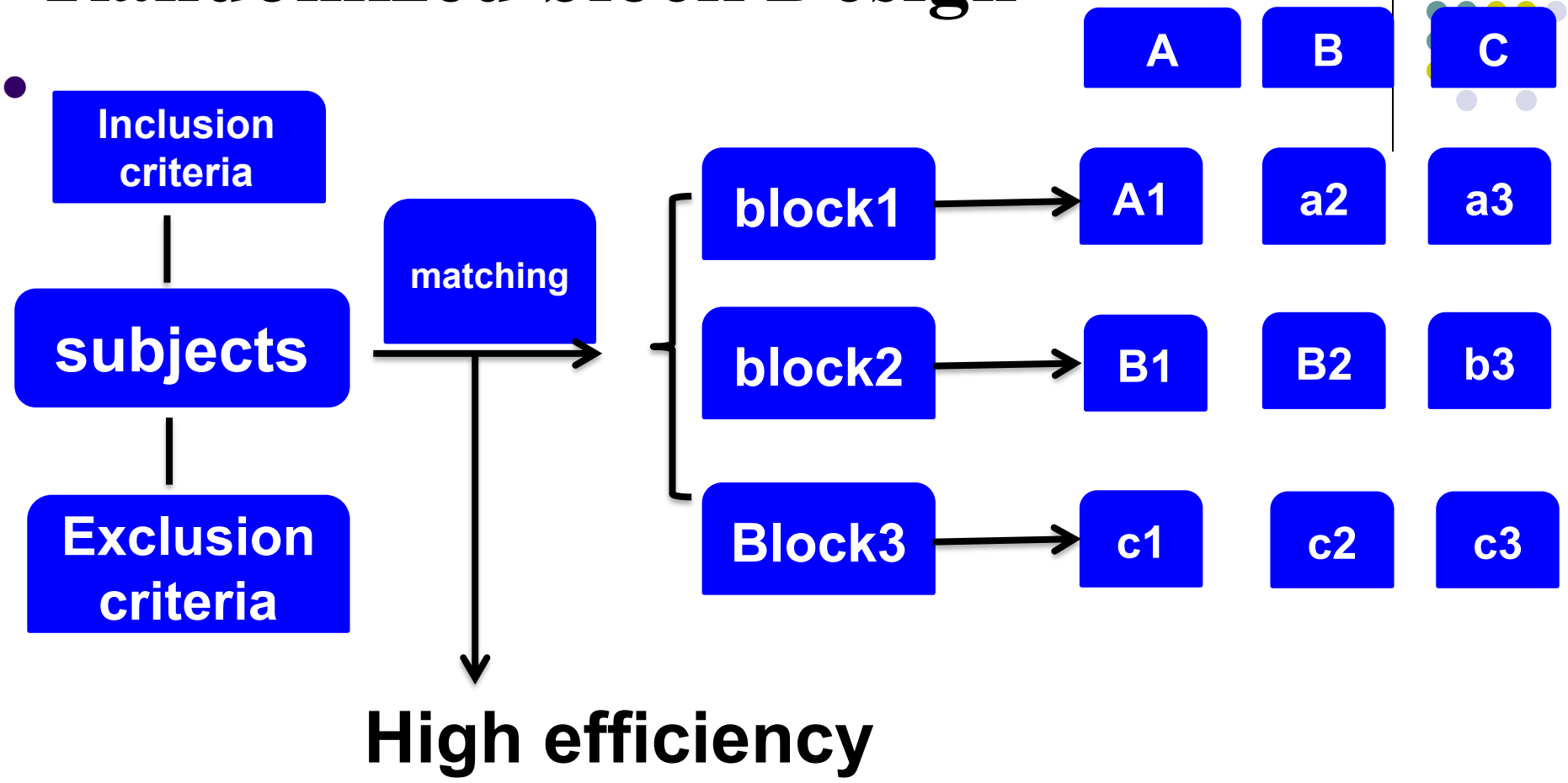
disadvantage：complicated, poorly matching may weaken the efficiency

Example 2

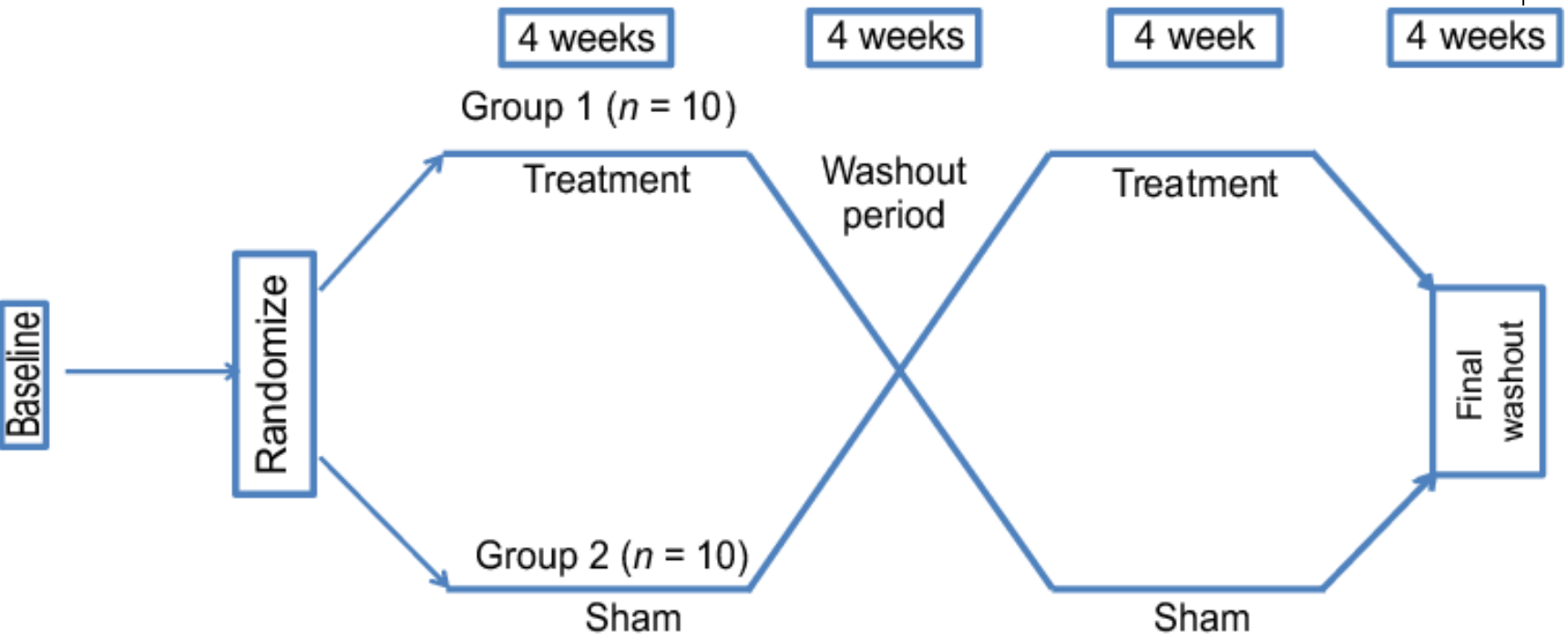
- Now we have 10 pairs of white mouse who are matched by the condition of same sex, ±10g. Please assign the pairs into experiment group and control

Mouse pair	A1-A2	B1-B2	C1-C2	D1-D2	E1-E2	F1-F2	G1-G2	H1-H2	I1-I2	J1-J2
randomized number	1	2	7	4	18	0	15	13	5	12
experiment group	A2	B1	C2	D1	E1	F1	G2	H2	I2	J1
control	A1	B2	C1	D2	E2	F2	G1	H1	I1	J2

Randomized block Design



Cross-over design



Cross-over design

Advantages

- Reduced influence by confounders since patients serve as their own controls
- Reduced variability in the outcome(s) being measured, thus increasing the precision of estimation
- Smaller sample sizes required
- Having the opportunity to receive both treatments can sometimes be attractive for subjects

Disadvantages

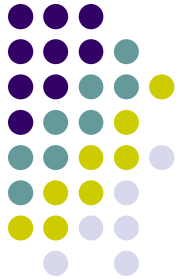
- Cannot be done when the subjects can only receive one treatment
- Assumption of no carryover effects (from washout period) is difficult to sometimes accurately test
- May take longer than a randomized clinical trial since patients have to cross over into each arm after an appropriate washout period
- Can be subject to period effects where differences in the effectiveness of an intervention can occur due to the passage of time. For example:
  - Development of tolerance
  - Resistance
  - Dropouts
  - Changes in the disease process being evaluated or treated

Factorial design

- 2×2 factorial design

		Independent Variable 2	
		Level 1	Level 2
Independent Variable 1	Level 1	Dependent Variable	Dependent Variable
	Level 2	Dependent Variable	Dependent Variable

Factorial design



- Advantage
- 1. more efficient mainly due to the smaller sample size required (up to one-half) compared with two separate two-arm parallel trials.
- 2. A factorial design is the only design that allows testing for interaction
- 3. Reduced costs, reduced recourses and management needs are found due to the fact that a smaller sample will be required compared with two separate trials.

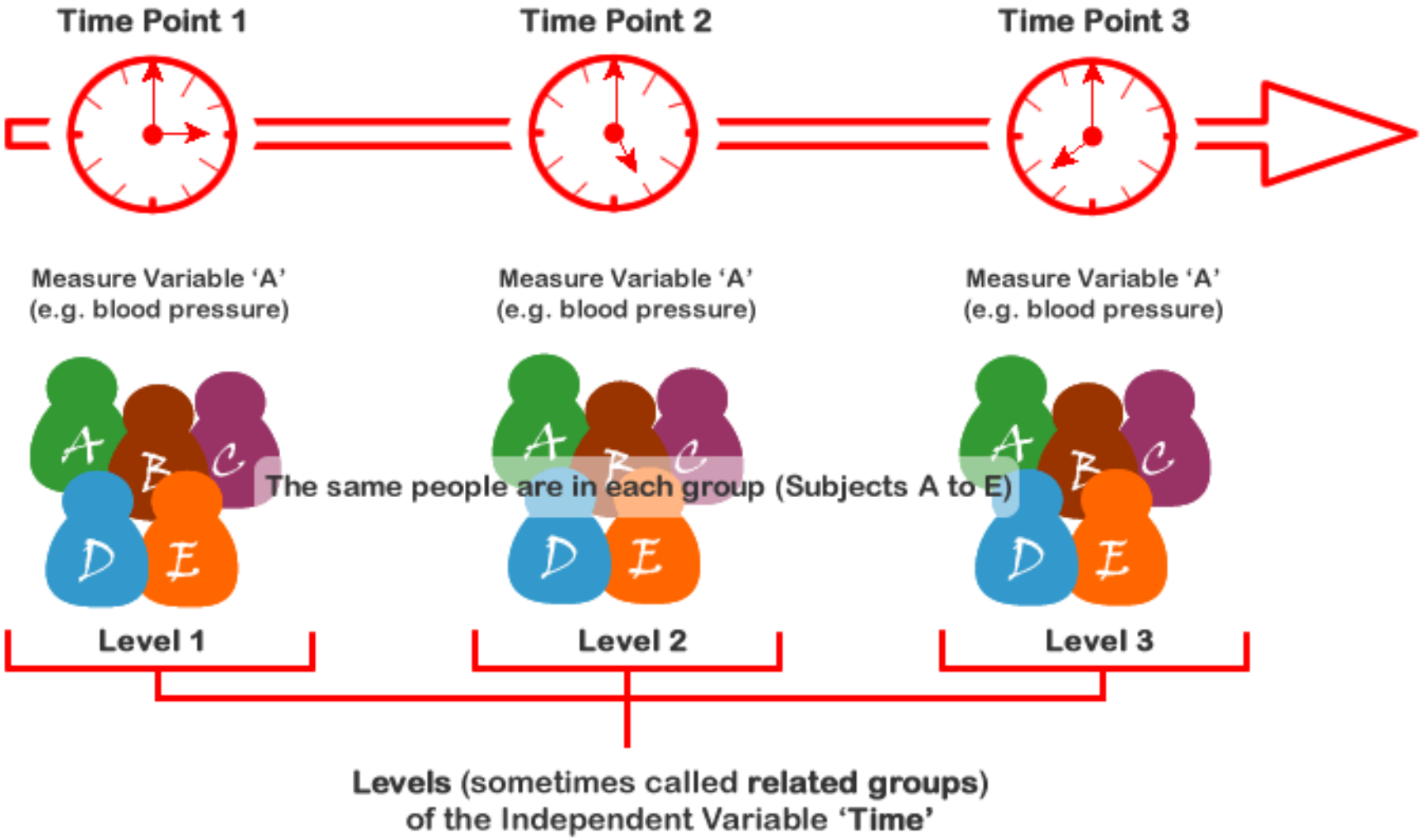
Factorial design

- disAdvantage
- 1. A factorial design may require extra time, compliance, and management of applying two treatments at the same time.
- 2. Data analysis and randomization may be a little more complex because participants must be allocated to four arms either in one (A, B, C, and D) or two stages (first intervention and comparator, and then second intervention and its comparator

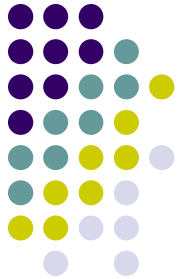
Factorial design

- disAdvantage
- 3. Appropriateness and acceptability/tolerance of the combined intervention on biologic and scientific grounds must be explored and determined
- 4. If interaction is expected, but there is no intention to detect the interaction, the factorial has no sample size advantages compared with two separate two-arm parallel trials.

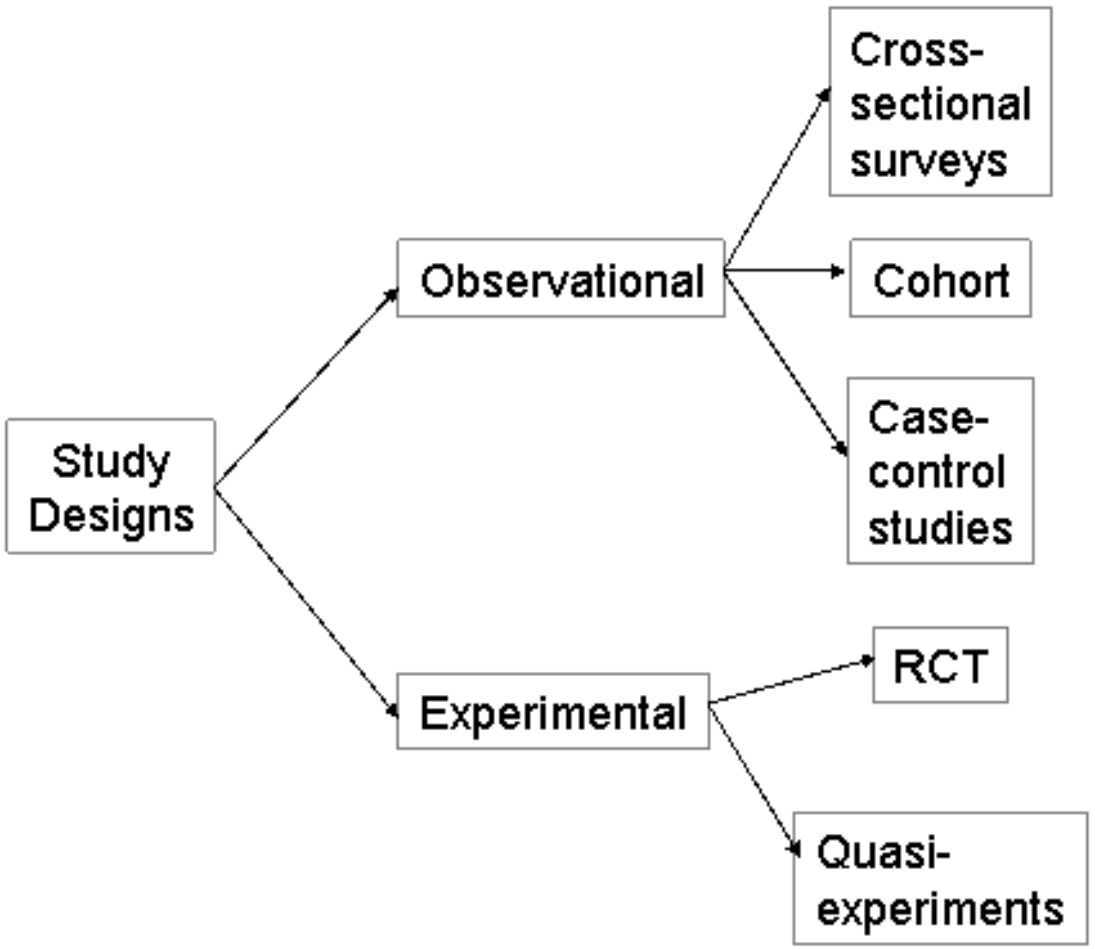
Repeated measurement design



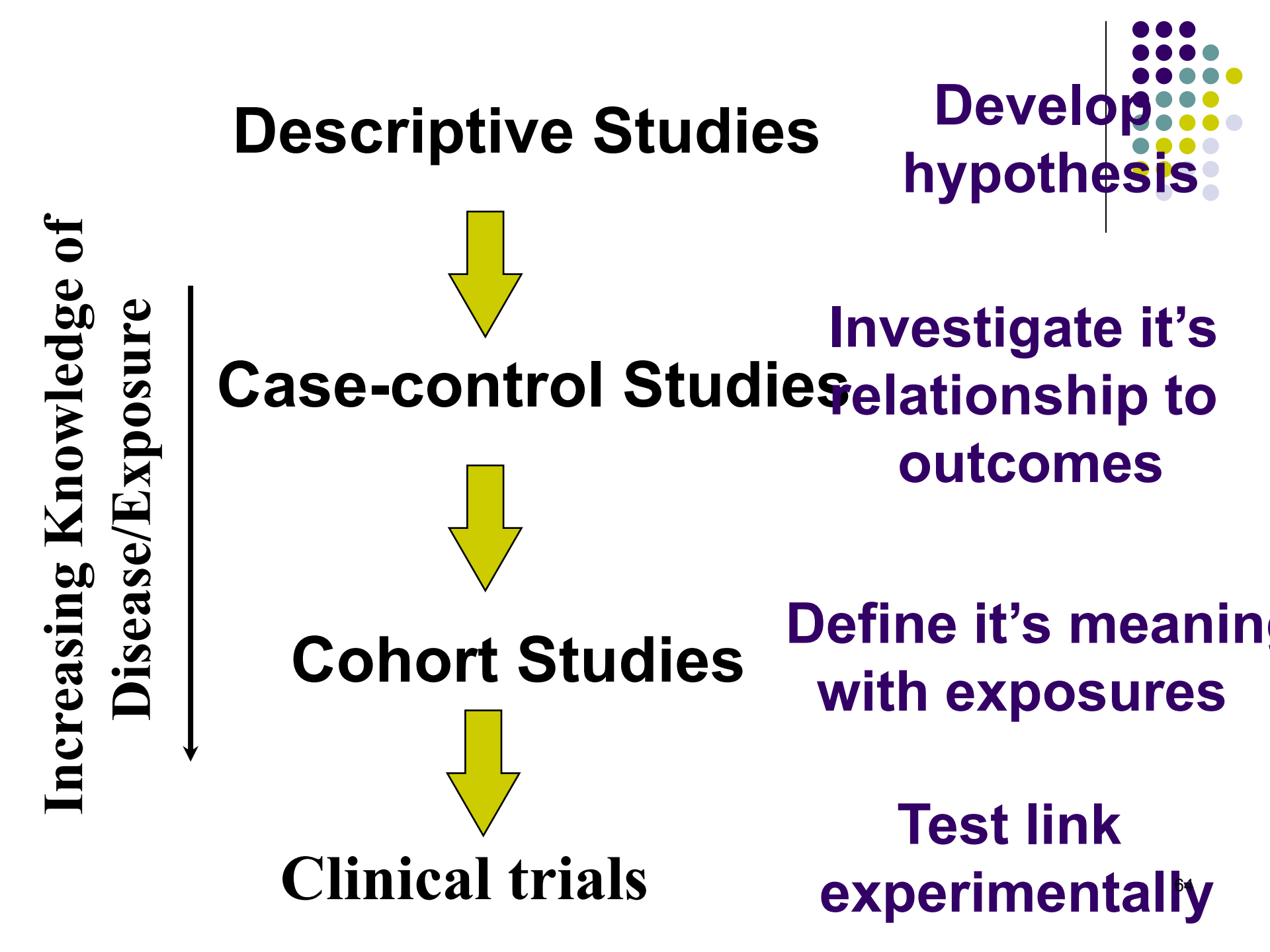
How Do I Remember Which Design is Which?



When someone says….	I think….
Case report	One person, one case. Like a police report of an incident.
Case-control	Cases AND controls – the "case" people have the disease, the "control" people don't. We are going to look back in time to figure out what made them different.
Cohort	A group of people. Perhaps your class has been referred to as the MD20xx cohort because you started at the same time and continue through medical school together. A cohort study is the same way – recruited at the same time and observed over a long time.
Cross-sectional	Cross-section of a tree (or anything else of your choosing). It's a one-time slice to get a better view.
RCT	The major type of experimental study; people are randomly sorted into control and experimental groups to try something out.



Advantages	Disadvantages
Quick; can cover whole population, giving representative information whether or not people are seeking care	Based mainly on self-report (biases?); diagnostic information usually inaccurate; can't establish causal sequence
Prospective, so can establish causal sequence; can estimate incidence	Time-consuming; costly; attrition of cohort?
Relatively cheap way of focusing on causal factors	Requires recall of past events (inaccurate?); controls not equivalent to cases
Controls for all main forms of bias; good for both etiological and evaluative research	Ethical concerns in etiological applications; Often uses selected populations: issue of generalizability?
May be more practical than RCT: can use "natural experiments"	Allocation bias often significant (exp'tal and control groups not equivalent)


Descriptive Studies
Develop hypothesis
Increasing Knowledge of Disease/Exposure
Case-control Studies
Investigate it's relationship to outcomes
Cohort Studies
Define it's meanin with exposures
Clinical trials
Test link experimentally

Have a nice Day!

Have a nice Day!