Statistical Design in Animal Research

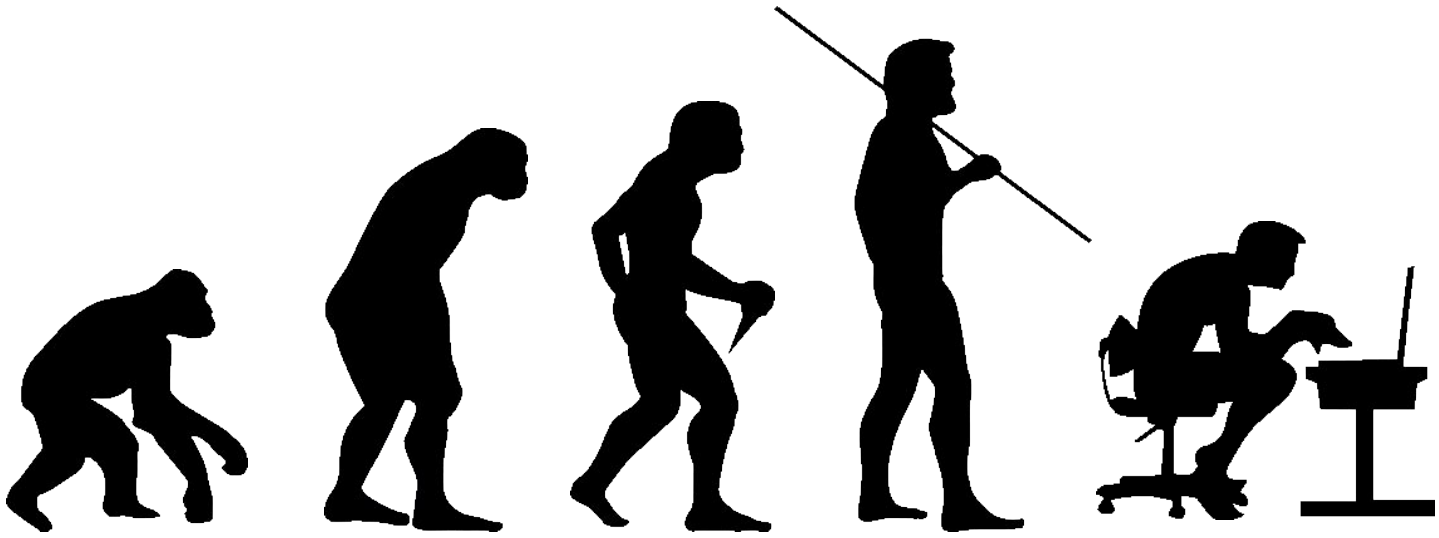
Important Principles for Experimental Researchers.

Servan Grüninger
MSc Biostatistics (UZH) and MSc Computational Science & Engineering (EPFL)

@SGruninger, www.servangrueninger.ch

Servan Grüninger, MSc UZH & MSc EPFL
Applied Statistics Group, Dept. of Mathematics, UZH
servan.grueninger@math.uzh.ch
@SGruninger, www.servangrueninger.ch

My academic evolution



Political Science & Law

Biology &
Neuroinformatics

Biostatistics &
Computational Science

1. Introduction

Good statistics is like strong coffee: bitter at times, always sobering

The science of

- collecting
- analysing
- interpreting
- presenting

data.

Good statistics is like strong coffee: bitter at times, always sobering

- collecting (how do you get data in a rigorous manner?): study design, sampling methods, data cleaning etc.
- analysing (how can you decompose your data?): visual inspection, statistical modeling, descriptive statistics, parameter estimation etc.
- interpreting (what can you *actually* say based on your data): model testing, inferential statistics, sensible(!) hypothesis testing etc.
- presenting (how can you convey your findings accurately to others?): descriptive statistics, graphs, tables etc.

Statistical design focuses on collecting, analysing and interpreting data

- collecting (how do you get data in a rigorous manner?): study design, sampling methods, data cleaning etc.
- analysing (how can you decompose your data?): visual inspection, statistical modeling, descriptive statistics, parameter estimation etc.
- interpreting (what can you *actually* say based on your data): model testing, inferential statistics, sensible(!) hypothesis testing etc.
- presenting (how can you convey your findings accurately to others?): descriptive statistics, graphs, tables etc.

Statistics is (mostly) an auxiliary science

«Statisticians are second-class mathematicians, third-rate scientists and fourth-rate thinkers. They are the hyenas, jackals and vultures of the scientific ecology: picking over the bones and carcasses of the game that the big cats, the biologists, the physicists and the chemists, have brought down.»

- Stephen Senn, Dicing with Death

But: It’s still a science (not a commodity)

- «I've been studying statistics for over 40 years & I still don't understand it. The ease with which non-statisticians master it is staggering» - Stephen Senn
- Don’t expect to become proficient in statistics in 90 minutes.
- Don’t treat statistics as a commodity or a “problem-solving machine”
- Statistics is immensely diverse and multifaceted.
- Know what you want: Many problematic uses of statistics stem from researches not knowing what their scientific(!) question is.

Goals for Today – What Can Be Achieved

- Raise awareness about statistical design and its links to sound scientific practices
- Give you an idea about the most relevant keywords, ideas and tools in the field
- Give you the tools to recognise problems, dig deeper, and convince your peers about the importance of statistical sound designs

-> Good statistical design might seem tedious and daunting at first, but it saves time, money, animal and human lives!

Goals for Today – What *Cannot* Be Achieved

- Make you proficient in study design, experimental execution, and statistical evaluation
- Cover all topics relevant for reproducibility and sound statistical work
- Give you any mathematical formulas for specific statistical problems

-> For more and related topics: LTK 11 (score sheets), LTK 13 (statistics), LTK 17 (systematic reviews), LTK 22 (reproducibility)

If in doubt, talk to a statistical consultant!

University of Basel

- http://ceb-institute.org/consulting

University of Bern

- http://www.ctu.unibe.ch/services/consulting_services/statistical_consulting___ctu_sprechstunde/index_eng.html
- http://www.imsv.unibe.ch/service/index_eng.html

Université de Fribourg

Université de Neuchâtel

- https://www.unine.ch/statistics/en/home/staff/slobodeanu-radu.html

Université de Genève

- https://www.unige.ch/gsem/en/research/institutes/rcs/consulting-service-in-statistics/
- http://www.unige.ch/math/en/camas

If in doubt, talk to a statistical consultant!

University of Zurich

- For MNF & MeF https://www.math.uzh.ch/consulting/?id=animalwelfare
- For VetMed https://www.vetepi.uzh.ch/en/services.html
- 3R-Competence Center: https://www.tierschutz.uzh.ch/de/3R/3R-Resources--Links.html

ETH

- https://www.math.ethz.ch/sfs/consulting/statistical-consulting-appointment.html

EPFL

- https://bbcf.epfl.ch/

UNIL/CHUV

- https://www.iumsp.ch/fr/cepic/consmeth

Contact your statistician during planning phase

Remember!

«To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.»

Ronald A. Fisher

4. Overview of design and analysis of experiments

4. Overview of design and analysis of experiments

- A. Types of studies
- B. Planning a study: design & analysis
- C. Types of data
- D. Exploratory vs. confirmatory research
- E. Pre-Registration

-> 4.A. to 4.C. and 4.E. are adapted from BME321 slides produced by Prof. Dr. Reinhard Furrer, Dr. Bernadetta Tarigan and me.

-> 4.D. is adapted from personal slides used in LTK 13 and LTK 22.

2.A. Types of studies

- Two types of studies:

  Observational study

  - Merely observe
  - Natural environment

  Experimental study

  - Manipulations or interventions
  - Controlled/standardized environment
- In statistical semantic: study ≠ experiment

2.A. Types of studies

- Two types of studies:

  Observational study

  - Merely observe
  - Natural environment

  Experimental study

  - Manipulations or interventions
  - Controlled/standardized environment

- In statistical semantic: study ≠ experiment

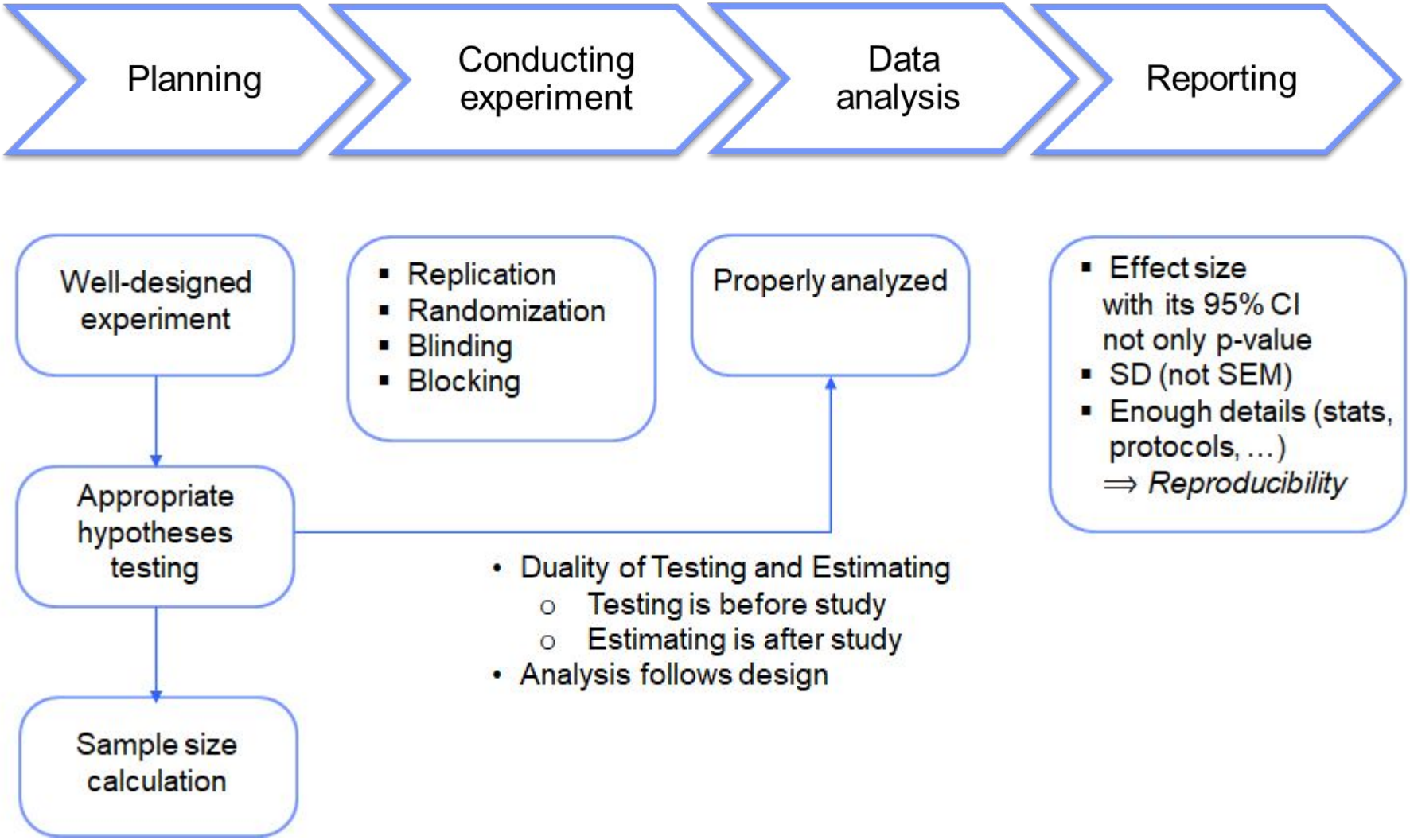
linear association = correlation ≠ causal effect

Experimental vs Observational

	Experimental	Observational
Situation	Controlled: *The settings you wish are the ones you get*	Given: *What you observe is what you get*
Statistical Analysis	Typically straight forward (ANOVA families)	Rather difficult (multivariate techniques)
Purpose	**To infer causality (if properly set up)**	To infer association

Your treatment is the only cause of the effect you observe in the outcome

Details of the four stages of an experiment



2.B. Planning a study: four basic questions

	Four basic questions	Details
Objective	What is **the question we want to ask**? Can it **be measured**?	What population (experimental unit)? What is an appropriate measure/outcome? **What type of outcome?**
Design	How to **sampling** from the defined population? When, where and how will you **get the data**?	Observational/Experimental? What sampling/randomization technique? How much data (replications)? Blinding? Blocking?
Analysis	How to describe the data? What do you think **the data are telling you**?	Descriptive statistics (tables & graphs) Inferential statistics (models/tests)
Reporting	How to **communicate** the finding(s)?	Report SD (not SEM) Report CI (not only p-value) Reproducibility?

What does it mean to design an experiment?

Design refers to

> the process of planning the experiment such that appropriate data will be collected and analyzed by proper statistical methods, resulting in valid and objective conclusions

It means **to plan properly** so that we can answer the four basic questions

Design	1. **What question and data** do you collect? → Objective
	2. How do you **collect** the data? → Randomization
Analysis	3. How do you **properly analysis** the data?
	4. What **conclusion** do you expect?

Planning

Recall that analysis (test/model and reporting) follows design

Thus we focus on the first two questions:

1. **Objective**
2. **Randomization process of the EU's**

How EU's are randomized into treatment

1. **CRD** (Completely Randomized Design)
2. **RCBD** (Randomized Complete Block Design)
3. **Split Plot** (combination CRD + RCB)

Linear models:

outcome = treatment + error

➔ **ANOVA** family
(continuous outcome)

PREPARE before you ARRIVE

- Follow PREPARE guidelines when planning your experiments:
  - Smith et al. 2017, PREPARE: guidelines for planning animal research and testing

- Follow some basic rules for effective statistical practice when preparing data collection and analysing it:
  - Kass et al. 2016, Ten Simple Rules for Effective Statistical Practice

- Follow ARRIVE guidelines when reporting your experiments (updated in 2020)
  - Percie du Sert et al. 2020, Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0

PREPARE

Planning **R**esearch and **E**xperimental **P**rocedures on **A**nimals: **R**ecommendations for **E**xcellence

PREPARE covers 15 topics:

Formulation of the study

1. Literature searches
2. Legal issues
3. Ethical issues, harm-benefit assessment and humane endpoints
4. Experimental design and statistical analysis

Dialogue between scientists and the animal facility

5. Objectives and timescale, funding and division of labour
6. Facility evaluation
7. Education and training
8. Health risks, waste disposal and decontamination

Methods

9. Test substances and procedures
10. Experimental animals
11. Quarantine and health monitoring
12. Housing and husbandry
13. Experimental procedures
14. Humane killing, release, reuse or rehoming
15. Necropsy

4.C. Different types of outcome variables

Name	*Description*	*Examples*
Qualitative / categorical variables		
Nominal	unordered categories	- weight gain yes/no - success/failure
Ordinal	Ordered categories, but no fixed distance between categories and categories don't have to be of same size	- Small/medium/ large weight gain
Quantitative / numerical variables		
Discrete	ordered numbers with fixed step-size (values between steps *not* possible)	- count data (e.g. number of animals in a cage)
Continuous	ordered numbers without fixed step-sizes (all values are possible)	- weight gain in g - blood flow rate

Focus on quantitative / numerical variables

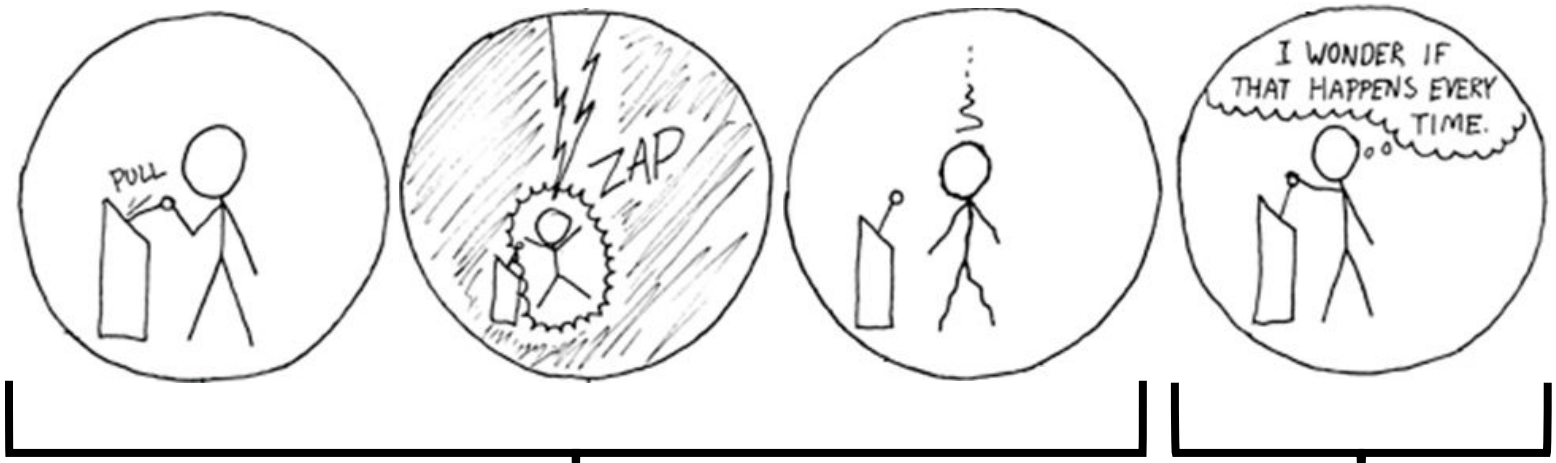
Name	*Description*	*Examples*
Interval Scale		
Discrete	ordered numbers with fixed step-size (values between steps not possible), no fixed zero	- money in your bank account
Continuous	ordered numbers without fixed step-sizes (all values are possible), no fixed zero	- temperature in degrees Celsius
Ratio Scale		
Discrete	ordered numbers with fixed step-size (values between steps not possible), fixed zero	- money in your wallet
Continuous	ordered numbers without fixed step-sizes (all values are possible), fixe zero	- temperature in degrees Kelvin

Type of outcome determines the analysis

Primary outcome (examples(!) of comparisons)	**Analysis (examples(!) of models/tests)**
Continuous (comparing means)	**ANOVA tables (F tests)**
Categorical (comparing proportions)	Contingency tables (χ^2 tests)
Time-to-event (comparing survival curves)	Survival analysis (logrank tests)

4.D. Exploratory vs. confirmatory research

Why the difference matters



Explore

Confirm

Image: adapted from xkcd comics (https://xkcd.com/242/)

The four stages of an experiment

Before you start with these stages, you need to know whether you conduct exploratory or confirmatory research!

Two types of experimental study

It depends on the general objective

Exploratory experimental study

- To learn something new or to generate hypothesis
- Discover as much as possible about the phenomenon under investigation
- Under given time, feasibility, resource and other constraints

Confirmatory experimental study

- To test a specific hypothesis
- To confirm a previous finding (which was often derived from an earlier exploratory experiment)

Exploratory vs. confirmatory research

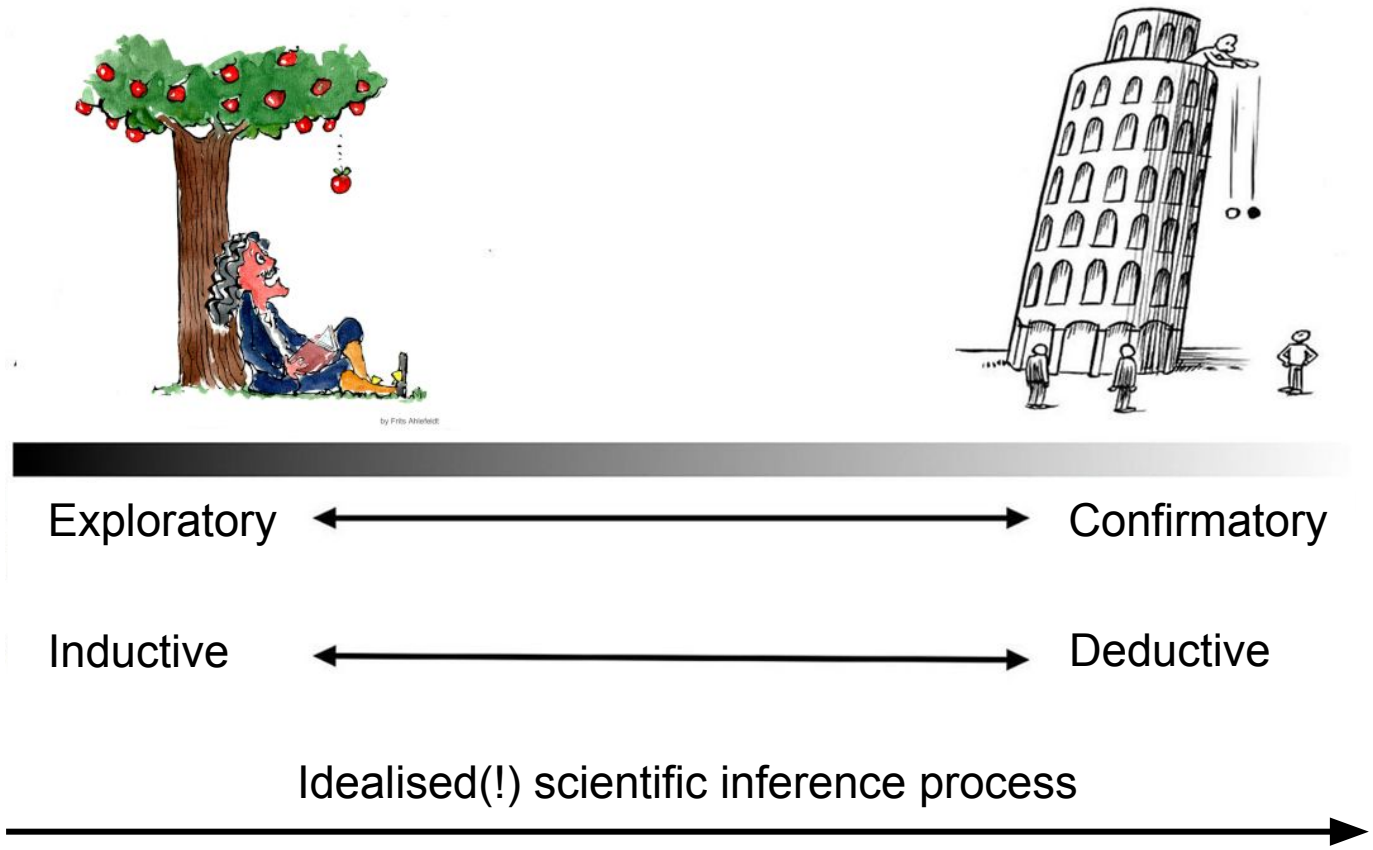


Image: adapted from Dirk-Jan Hoek and Frits Ahlefeldt
see also: Wagenmakers et al. 2012, An Agenda for Purely Confirmatory Research

Distinction is important

Distinction between the two is important

- Design features
- Subsequent data analysis/testing (statistical models)
- Sample size calculation

Design features

Design feature	Exploratory	Confirmatory
Subjects	Heterogeneous	Homogeneous
Environment	Varied	Standardized
Treatment factors	Many	Few
Factor levels	Many	Few
Outcomes	Many	Few
Hypotheses	Loose & flexible	Narrow & predefined
Replicates	Yes	Yes
Randomisation	Yes	Yes
Blinding	Yes	Yes
Blocking	Yes	Yes

4.E. Preregistration

Preregistration = register the methods and proposed analysis of your study before you conduct it

Is it beneficial? YES!

Why? Because you get a **peer review (assessment)** of your methods and analysis before you conduct your study

It means:

- you get feedback on your methods and proposed analysis, to avoid poor statistical design
- you can avoid doing the same thing as somebody else
- you can avoid somebody else might claim your study
- you can increase your chance to publish negative results
- increase transparency for science
- reduce the risk of reporting bias
- avoid duplication

Register Reports

Register Reports at Center for Open Science (COS)

https://www.cos.io/initiatives/registered-reports

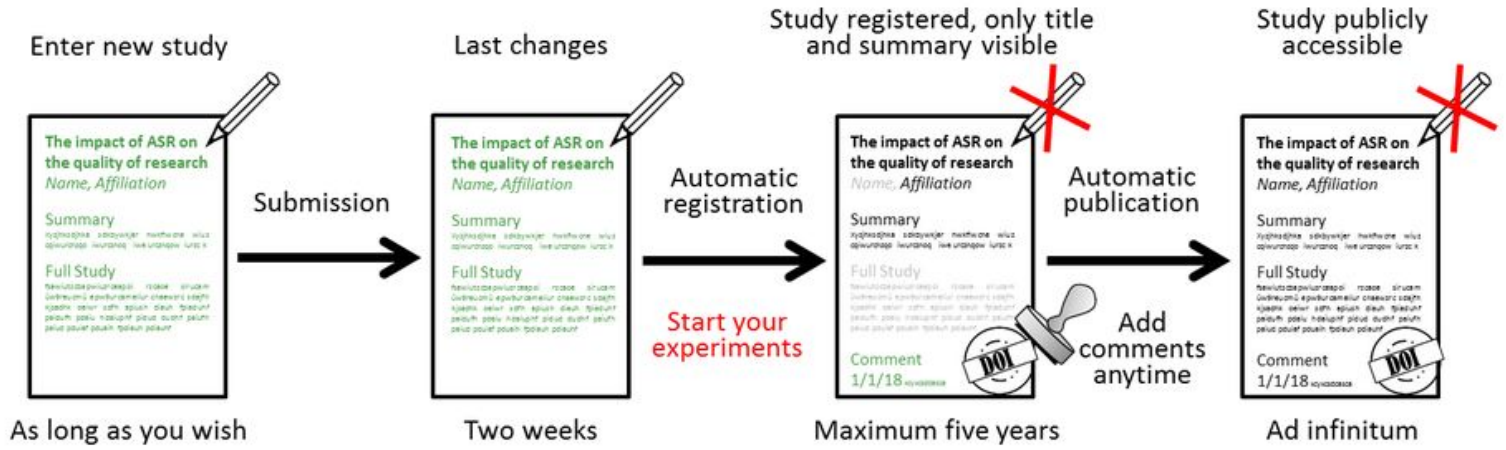
Participating Journals

“Currently, **over 300** journals use the Registered Reports publishing format either as a **regular submission option** or as part of a single **special issue**. Other journals offer **some features** of the format. This list will be updated regularly as new journals join the initiative.”

Animal Study Registry

https://www.animalstudyregistry.org

Operated by the German Centre for the Protection of Laboratory Animals (Bf3R) at the German Federal Institute for Risk Assessment (BfR)



PreclinicalTrials.eu

https://www.preclinicaltrials.eu

The register is open for all animal studies.

We especially encourage researchers to register their confirmatory studies.

This register is created with support of the Transnational AllianCe for regenerative Therapies In Cardiovascular Syndromes (TACTICS) alliance.

Therefore the original focus is on the field of cardiac regenerative therapy.

The design of the registration form is discussed with members of the TACTICS alliance.

However, the register is open for studies from all fields of biomedical science.

Furthermore, registration is not limited to a certain country or continent.

5. How to plan, execute and analyse a study

5. How to plan, execute and analyse an experiment

A. What does it mean to design an experiment?

B. Do you even know what your question is? How to transform scientific hunches into solid study designs.

C. What could possibly go wrong? Biases and how to prevent them.

-> Slides are adapted from joint slides from LTK 13 (created together with Prof. Dr. Reinhard Furrer & Dr. Bernadetta Tarigan) and from personal slides from LTK 22 (held together with Dr. Maike Heimann and Dr. Phil Bugnon).

5.A. What does it mean to design an experiment?

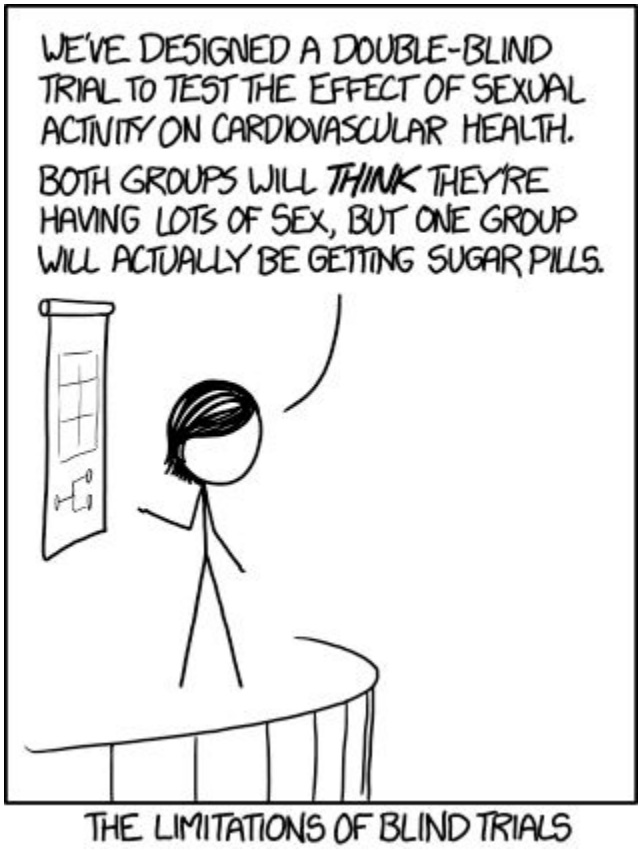
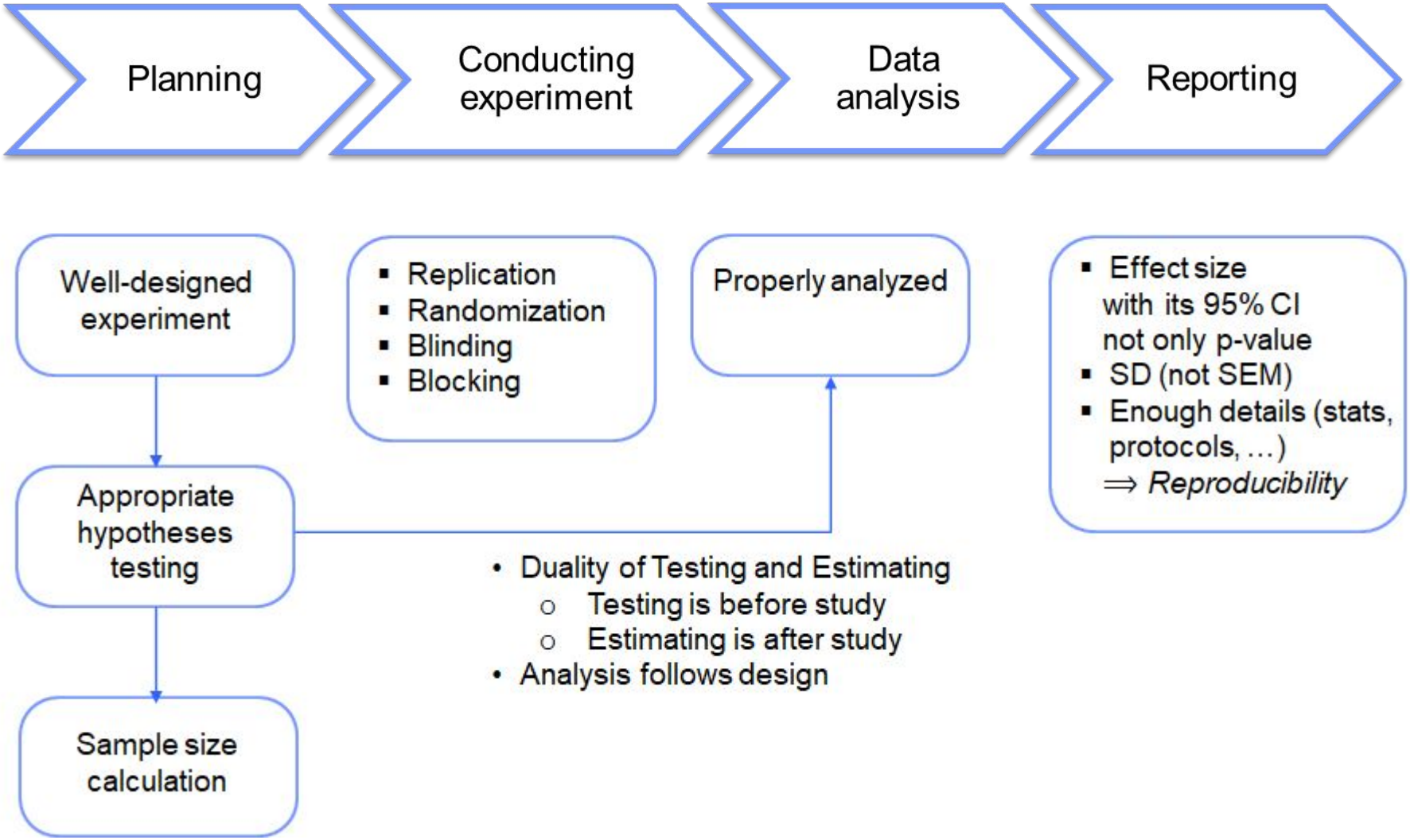


Image: xkcd comics (https://xkcd.com/1462/)

Details of the four stages of experiment



What does it mean to design an experiment?

Fundamental experimental design equation:

Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”)

Conceptually easy, but details depend on many factors:

- Exploratory or confirmatory experiment
- Knowledge about factors and effect sizes
- Available resources

Characteristics of well-defined experiment

Characteristics	How to do it
Clear objective	PICO-B method
Clear definition of EUs	Think about the smallest unit to which you can apply a different treatment
Unbiased	Randomized, Blinding
High precision (low variability)	Replication, Blocking
Able to estimate uncertainty	Replication, Randomized
Wide range of applicability	Blocking (deliberate variation)
Simple	Protect against mistakes

Characteristics of well-defined experiment

Characteristics	How to do it
Clear objective	PICO-B method
Clear definition of EUs	Think about the smallest unit to which you can apply a different treatment
Unbiased	Randomized, Blinding
High precision (low variability)	Replication, Blocking
Able to estimate uncertainty	Replication, Randomized
Wide range of applicability	Blocking (deliberate variation)
Simple	Protect against mistakes

5.B. Do you even know what you're asking?

How to transform scientific hunches into solid study designs?

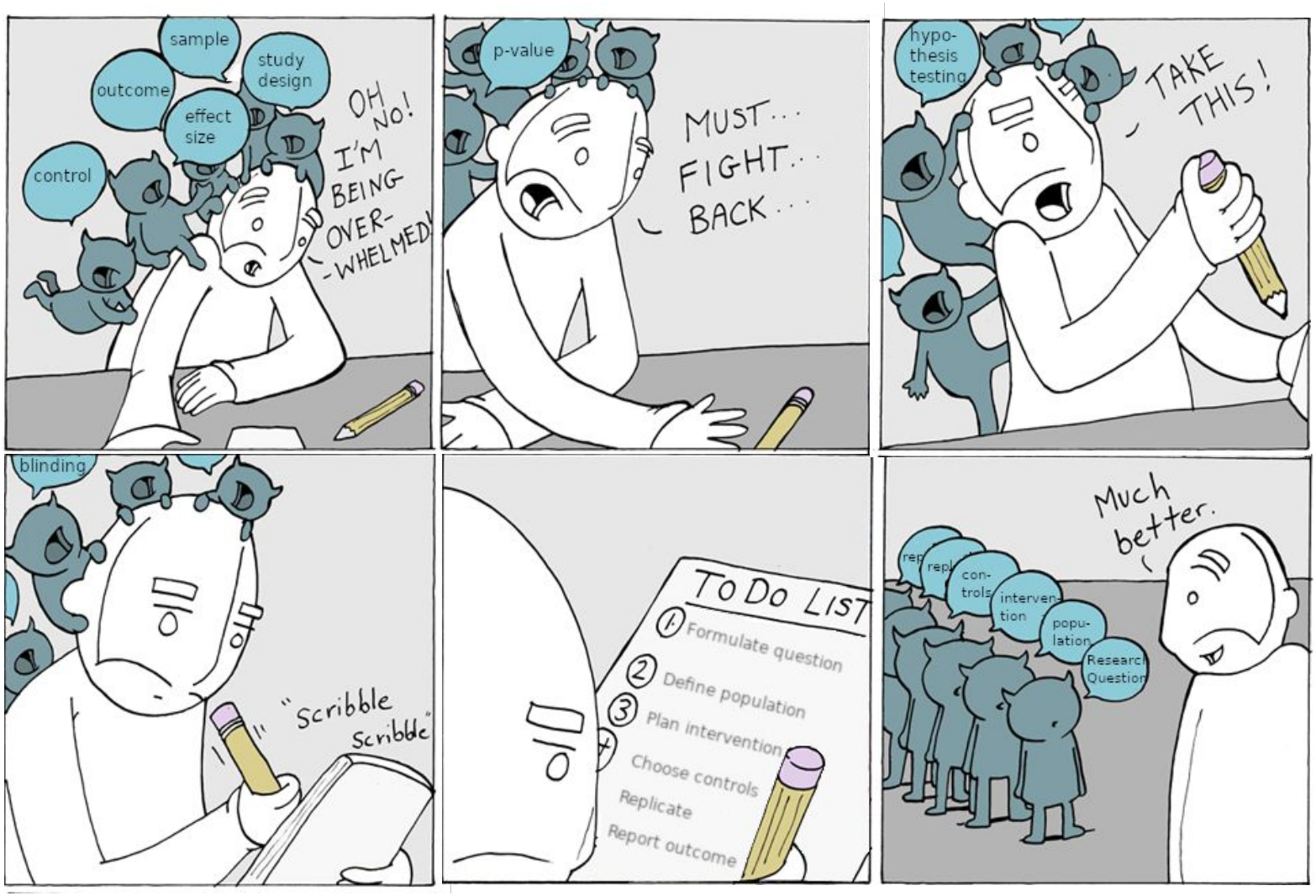


Image: adapted from Lunarbaboon (http://www.lunarbaboon.com/comics/overwhelmed.html)

Good planning is key

- clear objective
- well-defined experim. unit
- simple design
- methods correspond to design
- pre-registration

How to formulate experimental objectives?

***PICO-B* method**	
Population	*Experimental Units (EUs)*
Intervention	*Predictor*
Control	*Predictor*
Outcome	*Response*
Blocking	*Confounders*

How to formulate experimental objectives?

***PICO-B* method**	
Population	*Experimental Units (EUs)*
Intervention	*Predictor*
Control	*Predictor*
Outcome	*Response*
Blocking	*Confounders*

What is “N”?

Types of units to consider

- **Biological unit** (BU) of interest: the entity about which you make inferences, e.g.
  - litter of mouse
  - individual mouse
  - organs
  - parts of organs (such as brain areas)

- **Experimental unit** (EU): the entity that is randomly and independently assigned to one of the treatment levels., e.g.
  - a BU of interest
  - groups of BU
  - parts of BU
  - a sequence of observation of a BU

- **Observational unit** (OU): the entity on which measurements are taken.

adapted from: Lazic (2016). Experimental Design for Laboratory Biologists. Chapter 3, p. 96

What is “N”? -> The experimental unit (EU)

The EU is the unit which has **to be replicated** in an experiment!

Sample size “N” is the **number of proper replications** aka the number of EU.

Increasing the number of OU does **not** increase the sample size unless EU = OU (so-called “pseudo-replication”).

The experimental unit (EU)

How to check whether you identified the correct experimental unit?

1. The EU were independently randomised -- following the chosen randomization scheme such as CRD, RCBD, Split-Plot, etc. -- to the treatment conditions.

2. The treatment must be independently applied to each EU without spill-over to other EUs.

3. The EUs must not **systematically** influence each other **with regard to the experimental outcome.**

For more details, see, Lazic et al. (2018), What exactly is “N” in cell culture and animal experiments? and the NC3R’s entry on experimental units: https://eda.nc3rs.org.uk/experimental-design-unit

How to formulate experimental objectives?

***PICO-B* method**	
Population	*Experimental Units (EUs)*
Intervention	*Predictor*
Control	*Predictor*
Outcome	*Response*
Blocking	*Confounders*

Blocking

Blocking involves dividing the experiment into **a series of (roughly) homogeneous mini-experiments** according to predefined and experimentally relevant criteria.

Reasons to block:

- practical: easier to manage because of technical constraints such as lab equipments, operators, etc.
- statistical: to reduce variability

When you **know** about sources of variability, you should make sure that your intervention and control groups are homogeneous with regard to these sources **and include them into your analysis** in order improve your statistical efficiency.

What use are blocks?

Fundamental experimental design equation without blocking (but assuming proper randomisation):

Outcome	=	Treatment Effect + Large Noise (aka “Random Error”)

The “noise” (or “error”) of your model describes the amount of imprecision of of your outcome measure. Can be reduced by **prudently** including biological or technical factors as blocking factors in the design and the analysis.

Outcome	=	Treatment Effect + Blocking Factors (biological or technical) + Small Noise (aka “Random Error”)

Don’t overdo it with blocks!

The following will often lead to disaster:

Outcome	=	Treatment Effect + Biological Factor 1 + Biological Factor 2 + Biological Factor 3 + Biological Factor 4 + Biological Factor 5 + Biological Factor 6 + Technical Factor 1 + Technical Factor 2 + Technical Factor 3 + Technical Factor 4 + Technical Factor 5 + Technical Factor 6 + Small Noise (aka “Random Error”)

Reasons:

- Risk to introduce dependencies
- Risk to introduce colliders
- Risk of overfitting
- Instead of gaining power, you start losing power (i.e. need more experimental units to answer your question).

Rule of thumb: Add blocking factors that you know have an effect on your outcome variable, randomise the rest.

PICO-B for concrete examples

PICO-B		Example 1	Example 2 Metabolic study
Population	*Experimental Units (EUs)*		
Intervention	*Predictors*		
Control	*Predictors*		
Outcome	*Response*		
Blocking	*Confounders*		

Example 1: A fictional but realistic example

You work with a specific mouse model for a severe bowel disease.

The animal research committee demands that you administer pain killers to your animals.

However, those pain killers of which you know the effect on your mice are out of the question because they would interfere with your research question.

You are left with one drug that has proven to be effective in other bowel disease models but which has never been tested in your specific model.

You neither know whether it is effectively reducing pain nor whether its effect could interfere with your scientific results.

Hence, you are tasked to conduct a pilot study to assess the effects of this drug on a range of different biological parameters.

How do you proceed?

PICO-B for concrete examples

PICO-B		Example 1	Example 2 Metabolic study
Population	*Experimental Units (EUs)*	Mice (specific bowel disease model) of both sexes	
Intervention	*Predictors*	Pain killer	
Control	*Predictors*	No pain killer (negative control) Standard pain killer (positive control)	
Outcome	*Response*	- Severity of bowel disease (e.g. MEIC) - Pain level - Weight loss	
Blocking	*Confounders*	- (Sex) - Weight	

Example 2: Metabolic study

Exploratory questions	Confirmatory questions
Does the types of carbohydrate intake affect rodent metabolism (for how long, with regard to which metabolism measures)?	Does slowly absorbed carbohydrate diet (SAC diet) increase the body fat percentage by at least 15% compared to rapidly absorbed carbohydrate diet (RAC diet) on male mice consume in duration of 38 weeks?

PICO-B for concrete examples

PICO-B		Example 1	Example 2 Metabolic study
Population	*Experimental Units (EUs)*	Mice (specific bowel disease model) of both sexes	Mice or rats: -of different strains -both sexes
Intervention	*Predictors*	Pain killer	Diet 1 / diet 2
Control	*Predictors*	No pain killer (negative control) Standard pain killer (positive control)	Normal-diet
Outcome	*Response*	- Severity of bowel disease (e.g. MEIC) - Pain level - Weight loss	- Weight gain - Body fat (%) - Energy intake - Insulin AUC - Glucose AUC
Blocking	*Confounders*	- (Sex) - Weight	- (Sex) - Weight

5.C. What could possibly go wrong?

Biases and how to prevent them

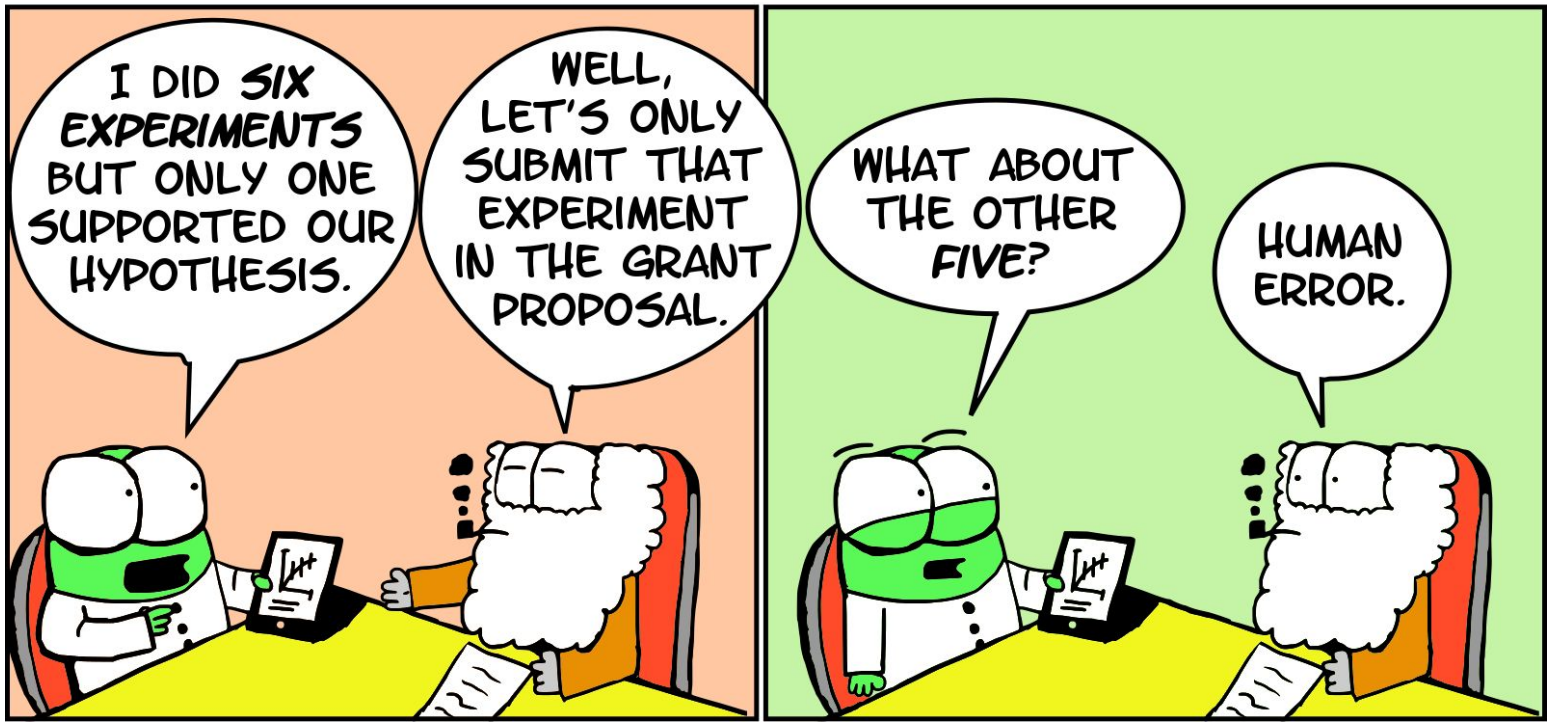


Image: The Upturned Microscope
(https://theupturnedmicroscope.com/comic/logical-fallacies-confirmation-bias)

How to prevent biases during the experiment?

- clear objective
- well-defined experim. unit
- simple design
- methods correspond to design
- pre-registration

- Replication
- Prevent biases
- Correct randomization
- Appropriate blinding

Recall: characteristics of well-defined experiment

Characteristics	How to do it
Clear objective	PICO-B method
Clear definition of EUs	Think about the smallest unit to which you can apply a different treatment
Unbiased	Randomized, Blinding
High precision (low variability)	Replication, Blocking
Able to estimate uncertainty	Replication, Randomized
Wide range of applicability	Blocking (deliberate variation)
Simple	Protect against mistakes

Why these characteristics?

The fundamental equation

Fundamental experimental design equation:

Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”)

- **Treatment effect(s)** are the effect(s) caused by the manipulations or interventions of the experimenter.
- **Random Error** is not a mistake but the **variation** in the outcome that cannot be explained or attributed to the treatment effect or other effect(s), also called the experimental error.
- **Biological effects** are differences that arise from intrinsic properties of the samples and are not actively manipulated by the experimenter
- **Technical effects** are the properties of the experimental system that can influence the outcome: of little interest but may affect the outcome

Any study is subject to uncertainty/error

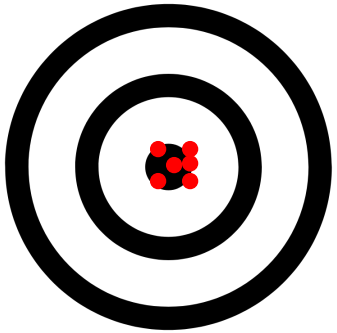
Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”) + Bias (aka “Systematic Error”)

	Source of uncertainty/error	
	Random error almost unavoidable	**Systematic error / bias** can be eliminated
Implication	*Imprecision*	*Low trueness*

Source: ISO 5725-1:1994(en) Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions (https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:en)

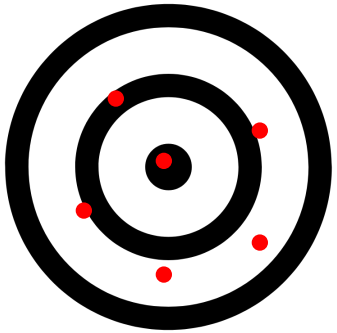
Imprecision and low trueness

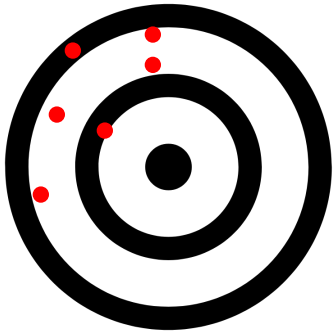
high trueness
high precision



low trueness
high precision

high trueness
low precision





low trueness
low precision

large variability = large random error → low precision
large bias = large systematic error → low trueness

Any study is subject to uncertainty/error

Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”) + Bias (aka “Systematic Error”)

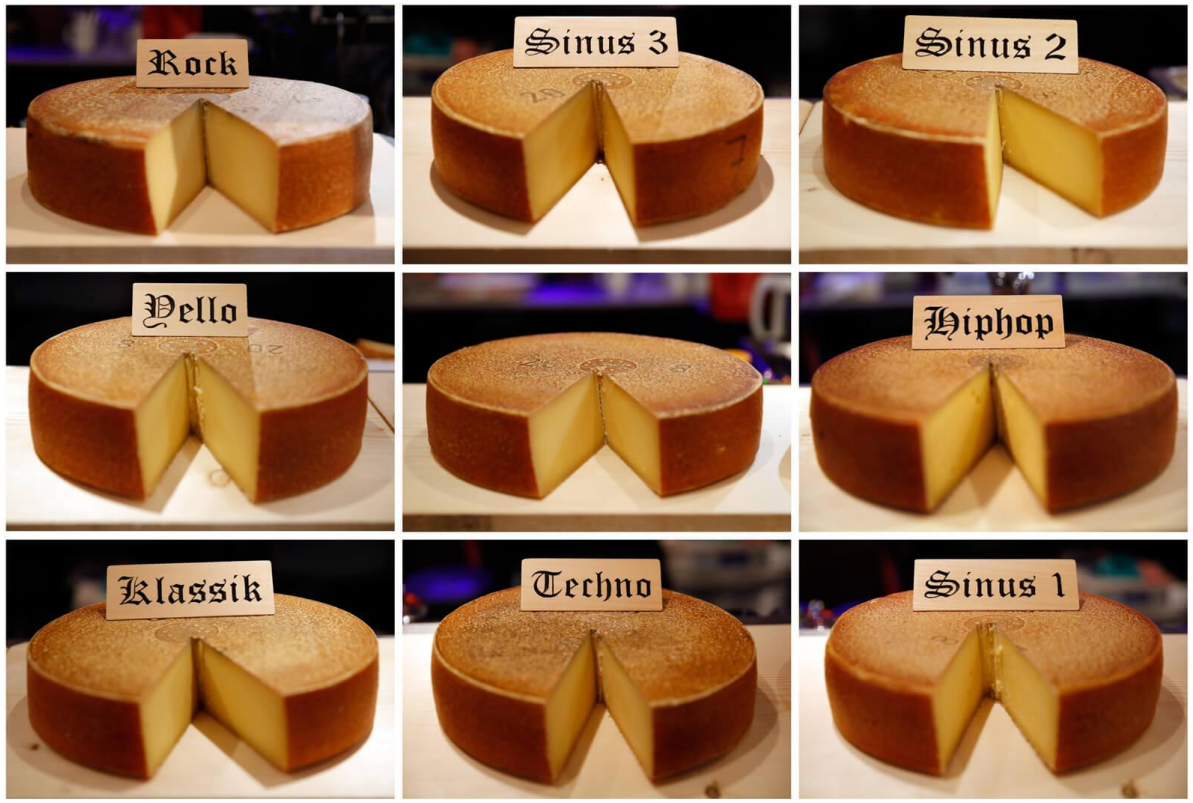
	Source of uncertainty/error	
	Random error almost unavoidable	**Systematic error / bias** can be eliminated
Implication	*Imprecision*	*Low trueness*

-> to measure imprecision, we need replicates!

Source: ISO 5725-1:1994(en) Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions (https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:en)

Without replicates, the results stink

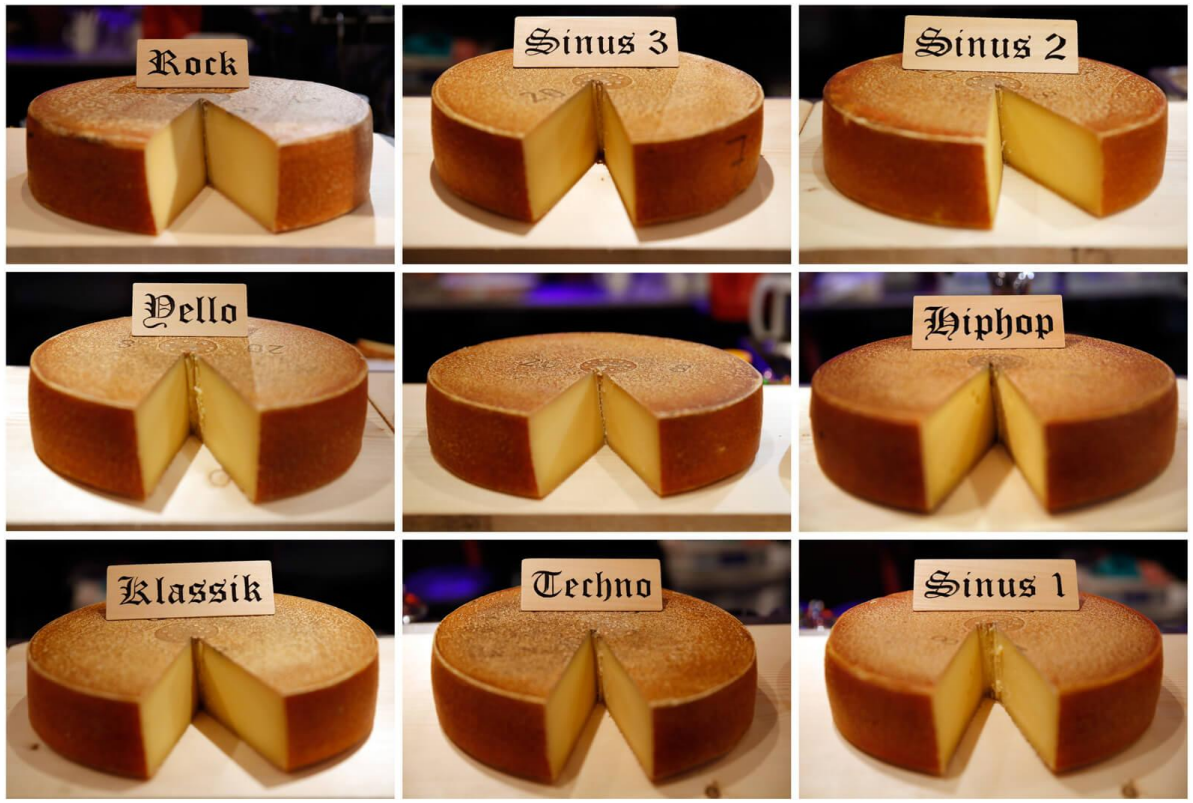
Can hip-hop music make cheese tastier? Yes, says science!



Newly Swissed (https://www.newlyswissed.com/cheese-in-sound/)

Without replicates, the results stink

And to push the idea even further, the project applied eight different types of sounds, one each per cheese wheel:



Newly Swissed (https://www.newlyswissed.com/cheese-in-sound/)

Any study is subject to uncertainty/error

Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”) + Bias (aka “Systematic Error”)

	Source of uncertainty/error	
	Random error almost unavoidable	**Systematic error / bias** can be eliminated
Implication	*Imprecision*	*Low trueness*

-> to remove bias we must properly plan and conduct study!

Source: ISO 5725-1:1994(en) Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions (https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:en)

Page 81

Bias (aka systematic error)

- Bias is any **variation that systematically occurs** in the result
- In other word, a deviation from the truth in the result
- Bias leads to low trueness estimates of the treatment effects to the true effects
- Bias may be introduced at the design of analysis phase of a study
- We consider a couple of biases that can happen at the design phase

see also: van der Worp et al. 2010, Can Animal Models of Disease Reliably Inform Humans?
Catalogue of Bias (https://catalogofbias.org/biases/confounding/)

Selection bias

It is a bias caused by non-random allocation of EUs to treatment groups

Examples:

- Allocating the less healthy animals to the high-dose group
- Allocating the more healthy subjects to intervention group
- Allocating the males to the intervention, the females to the control group

Solution: randomization

Performance bias

It is a bias caused by differences in care given to the subjects across treatment groups.

It can happen in two scenarios:

1. If researchers provide -- intentionally or unintentionally -- unequal care to subjects in different groups
2. If subjects in different groups behaved differently

Examples:

- The technician in laboratory animal gives *unequal husbandry care* for the mice in the intervention group than in the control group
- The animals were *not randomly housed* during the experiment
- A study in a weight-loss trial of a special counselling programme compared to a usual care in general practice. As participants in the control group were *disappointed* for being offered usual care instead of the new helpful program, the study concluded that their reaction to disappointment may introduce performance bias.

Solution: blinding, randomization

Observer/detection bias

It is a bias caused when the person assessing the outcome knows which treatment group the subject is assigned

Examples:

- Was the outcome assessor blinded?

  When assessing animal behavior, it is human nature to want to see a positive effect in your experiment

- Were animals selected at random for outcome assessment?

Solution: blinding, randomization

Attrition bias

Attrition means a reduction or decrease in numbers

Attrition occurs when participants leave during a study

Systematic differences between people who leave the study and those who continue can introduce bias into a study's results – this is attrition bias

Examples:

- Were incomplete outcome data adequately addressed?
- Study of psychosocial factors among patients with cardiac conditions showed that those who fully completed the study differed in clinical and psychosocial features from those who dropped out before the study ended. Such differential attrition could have biased the study's results.

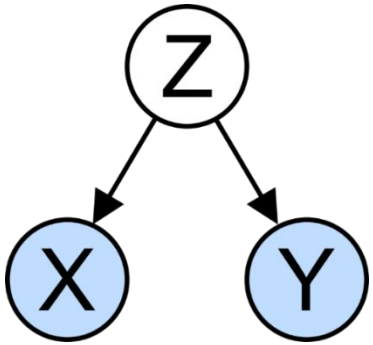
Solution:

- use the so-called "intention-to-treat" analysis (as opposed to "as-treated" analysis)
- take into account dropout rate in sample size calculation

Confounding bias

It is a bias due to another factor that distorts the relationship between treatment and outcome

Confounder is a variable that influences both the dependent variable and independent variable, causing a spurious association



Examples:

- Recall the metabolism study where the treatment levels are two types of certain diets and the outcome is the body weight gain. We can think that the treatment effect could be confounded by the initial weight.
- Can you think of an example?

Solution: blocking, randomization

Summary of biases

Type of Bias	Description	Solution
Selection	Bias caused by non-random allocation of EUs to treatment groups	Randomization
Performance	Bias caused by differences in care given to subject across treatment groups	Blinding, randomization
Observer/ Detection	Bias caused when the person assessing the outcome has knowledge of treatment assignment	Blinding
Attrition	Bias caused when participants drop out from a study and loss to follow-up across treatment groups	Intention-to-treat analysis, add dropout rate in sample size calculation
Confounding	Bias due to another factor that distorts the relationship between treatment and outcome	Blocking (reducing variability), randomization

What about your unknown knowns and unknown unknowns?

You will most likely never know all factors that influence your outcome variable.

Page 89

What about your unknown knowns and unknown unknowns?

Known Knowns Factors of which you are aware and of which you know the effect	**Unknown Knowns** Factors of which you are not aware but of which somebody else knows the effect
Known Unknowns Factors of which you are aware but of which you don't know the effect	**Unknown Unknowns** Factors of which you are not aware and of which nobody knows the effect
-> Handle with blocking & including factors in design and analysis	**-> Handle with randomisation and blinding**

Any study is subject to uncertainty/error

Outcome	=	Treatment Effect + Biological factors + Technical factors + Noise (aka “Random Error”) + Bias (aka “Systematic Error”)

	Source of uncertainty/error	
	Random error almost unavoidable	**Systematic error / bias** can be eliminated
Implication	*Imprecision*	*Low trueness*
Improved by	**Replications** **(least needed)** **Blocking**	**Randomization** of assignment **Blinding** **Blocking**

Source: ISO 5725-1:1994(en) Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions (https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:en)

Assessment of bias (risk of bias)

There are multiple helpful tools to help you assessing possible risks of bias, for example:

- Cochrane RoB page
- RoB tools
- Risk of Bias (RoB) tool in animal study: SYRCLE

A systematic assessment of different risk of bias tools can be found in Page et.al (2017)

Have we removed all the biases?

It’s impossible to control for all potential sources of bias on your outcome - simply because you might not even know what is influencing your outcome (“unknown knowns” and “unknown unknowns”).

Solution: randomize

Randomization

Assign experimental units to treatment groups by chance alone
→ independently from each other
→ following a pre-defined probabilistic assignment rule (often: identically with the same probability/chance)

Randomization ensures that – on average – the only systematic difference between the groups is the treatment.

How to randomize?

- Never do it by “hand”
- Use computer to randomize your experimental units

Type of randomization:

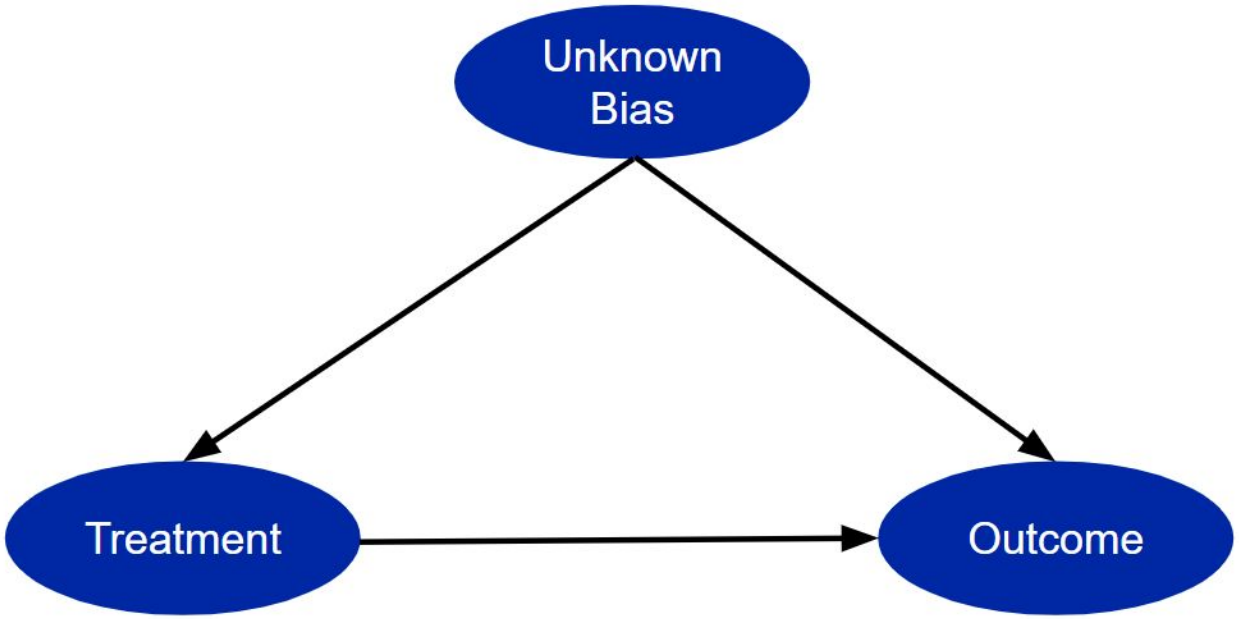
- Simple randomization:
  - strictly simple (might be unbalanced)
  - balanced
- Stratified/block randomization

Randomization - some common misunderstanding

- Randomization is **not** about “balancing” co-variables or factors.
- You **can** include co-variables and factors into your analysis even if you haven’t randomized.
- It is important **what** you randomize: You should randomize the treatment allocation to your experimental unit.
- Randomization “breaks” the link between a co-variable or a factor and your treatment. It does **not** break the link between a co-variable or a factor and your outcome.

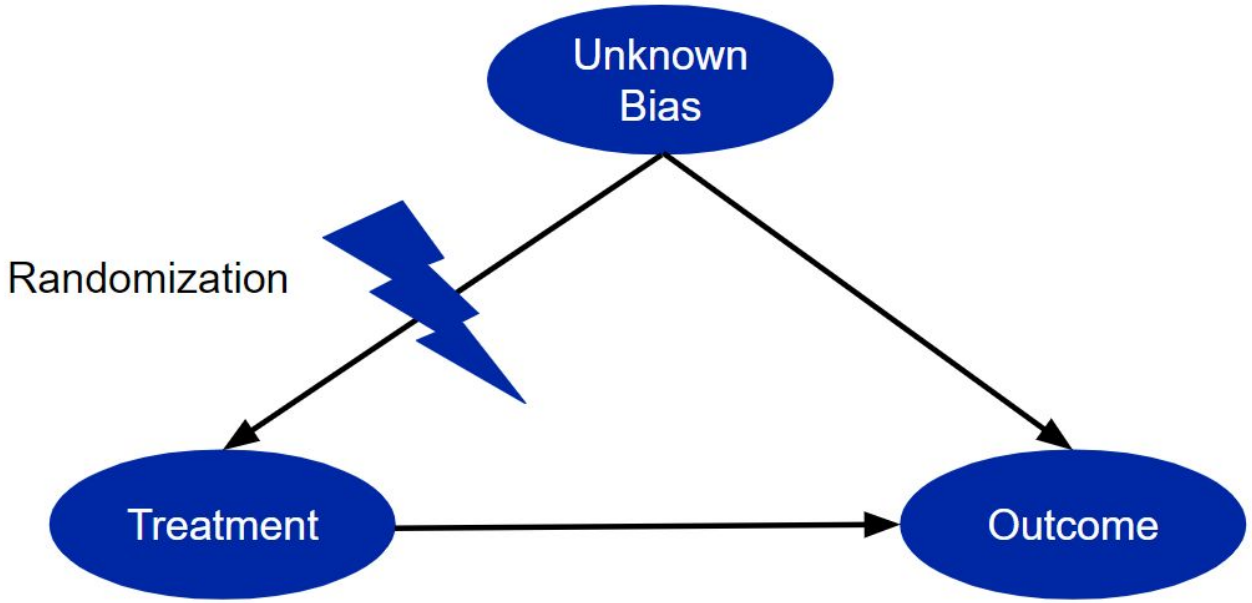
The inferential power of randomization

In this setup, you can't distinguish between the effect of the unknown bias and the treatment on the outcome. In this setup, you can't distinguish between the effect of the unknown bias and the treatment on the outcome.



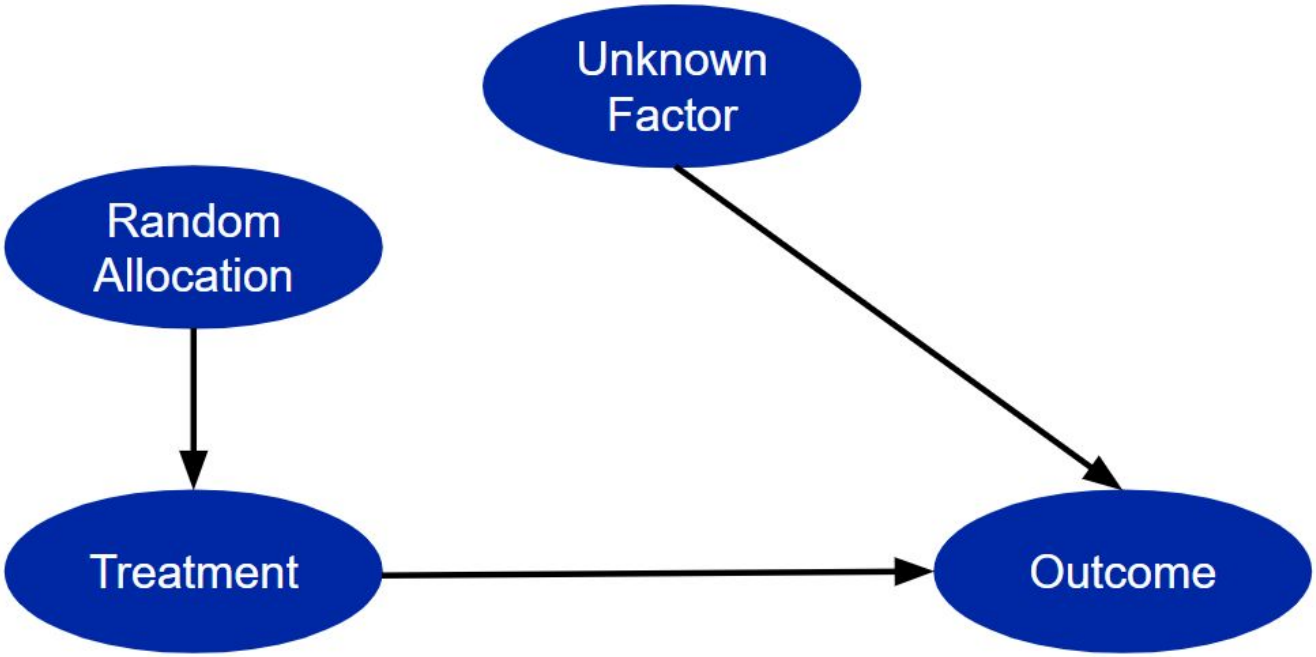
The inferential power of randomization

Randomization **of the treatment allocation to the experimental unit** breaks the link between the unknown bias and the treatment.



The inferential power of randomization

In this setup, you the effect of the unknown factor onto your outcome is still present, but you can distinguish between its effect on the outcome and the effect of the treatment on the outcome. Hence, the **bias** turned into a factor / co-variable that is **independent** of your treatment allocation.



Randomization

Assign experimental units to treatment groups by chance alone
→ independently from each other
→ identically with the same probability/chance

Randomization ensures that – on average – the only systematic difference between the groups is the treatment.

How to randomize?

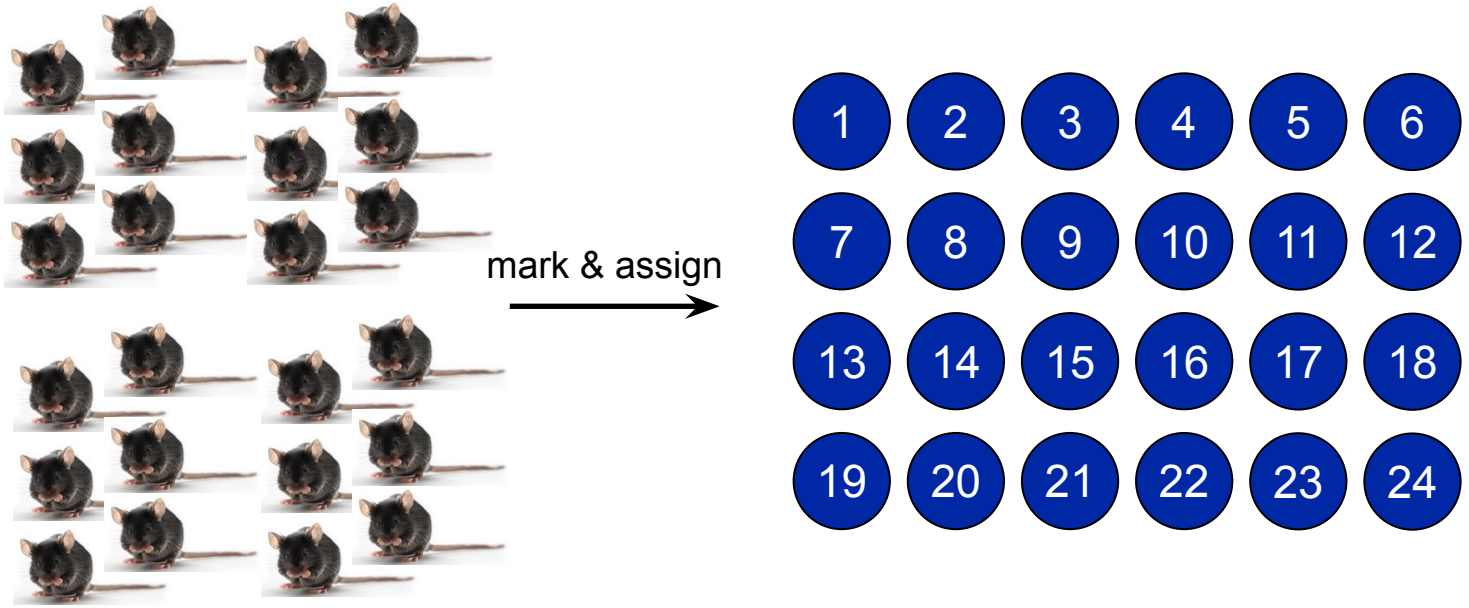
- Never do it by “hand”
- Use computer to randomize your experimental units

Type of randomization:

- Simple randomization:
  - strictly simple (might be unbalanced)
  - balanced
- Stratified/block randomization

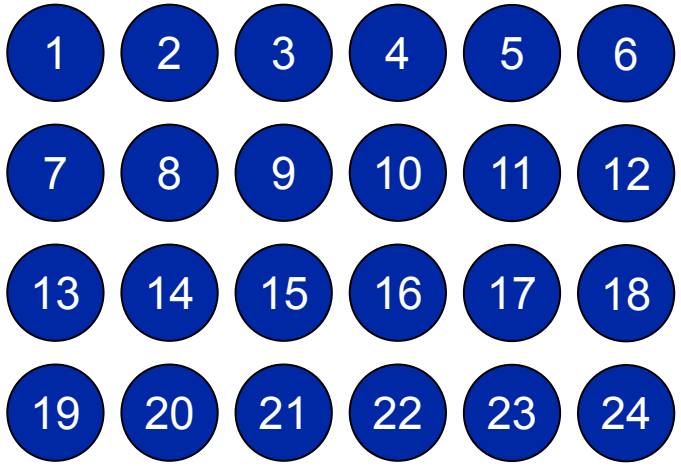
Simple randomization

- Remember example 2 (metabolic study) with 3 treatment groups and 1 control group; assume you have 24 mice.
- Goal: Assign each mouse randomly to one of the four groups



Simple randomization

Take each mouse and randomly assign it to one of the four groups



random

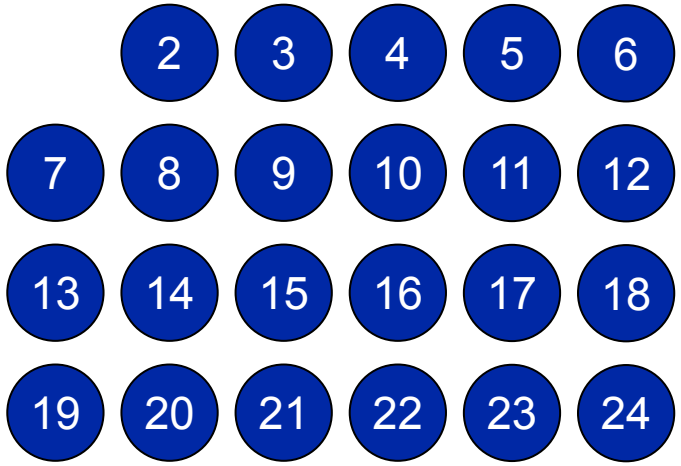
group assignment

A(=1)	B(=2)	C(=3)	D(=4)

In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

Take each mouse and randomly assign it to one of the four groups



random

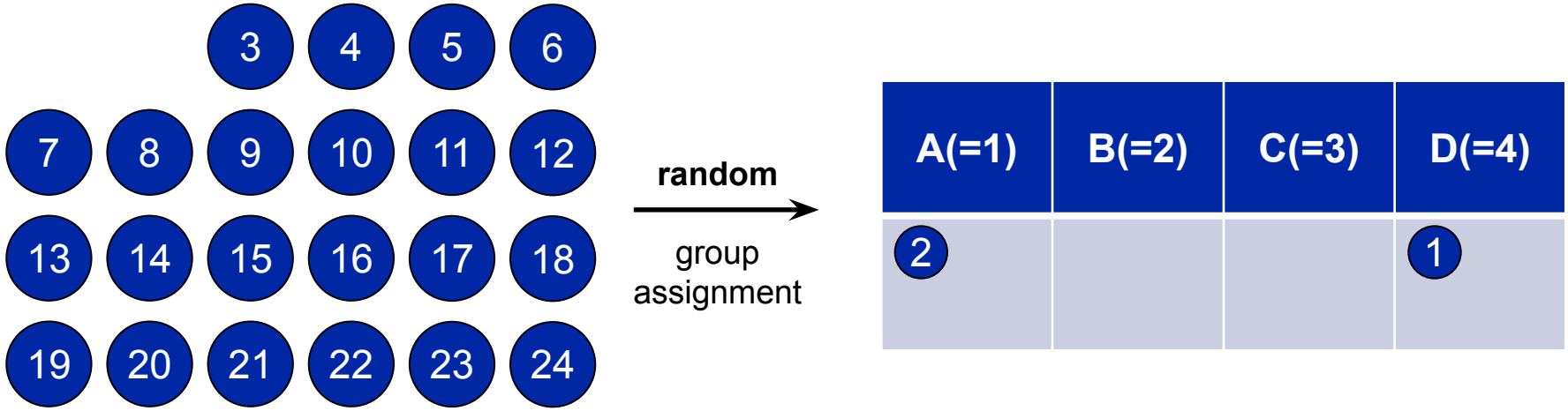
group assignment

A(=1)	B(=2)	C(=3)	D(=4)
			1

In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

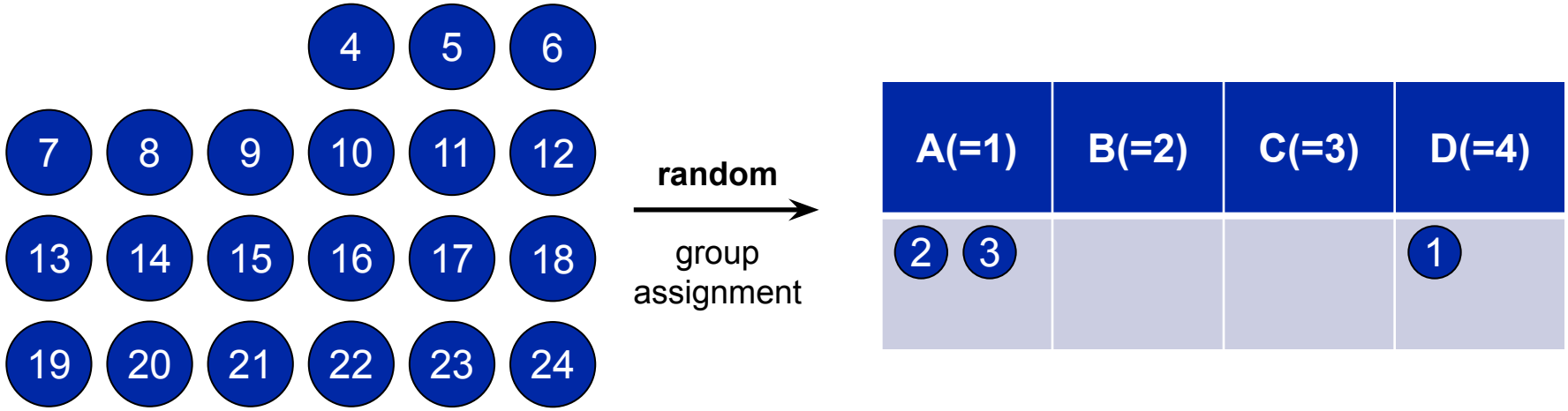
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

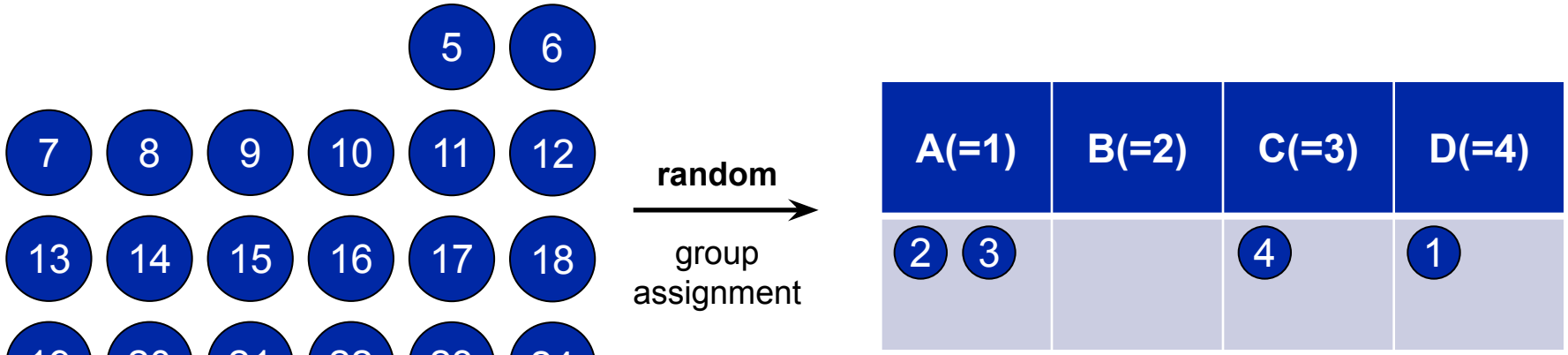
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

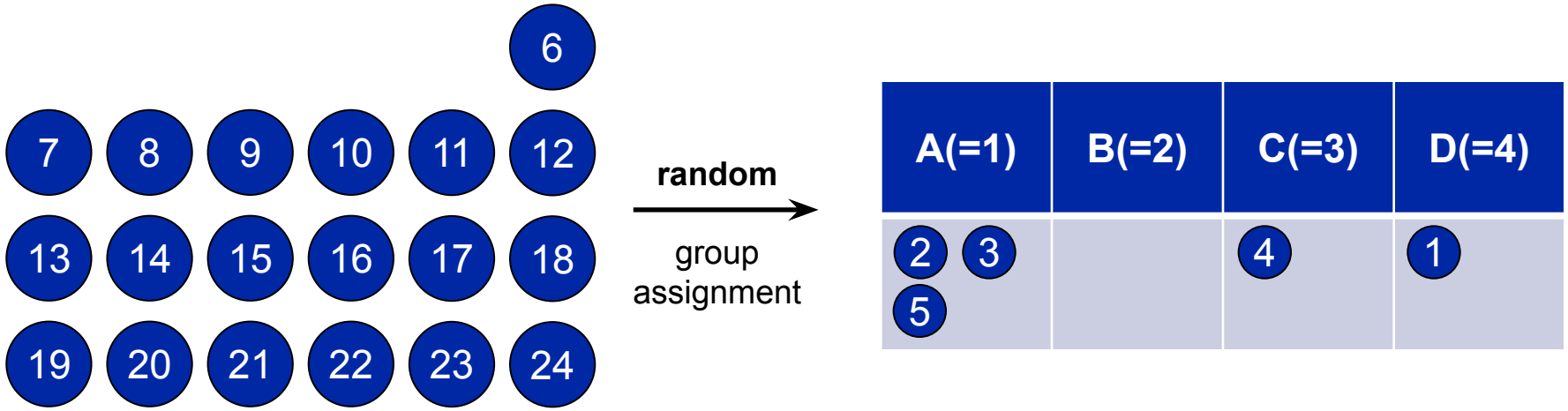
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

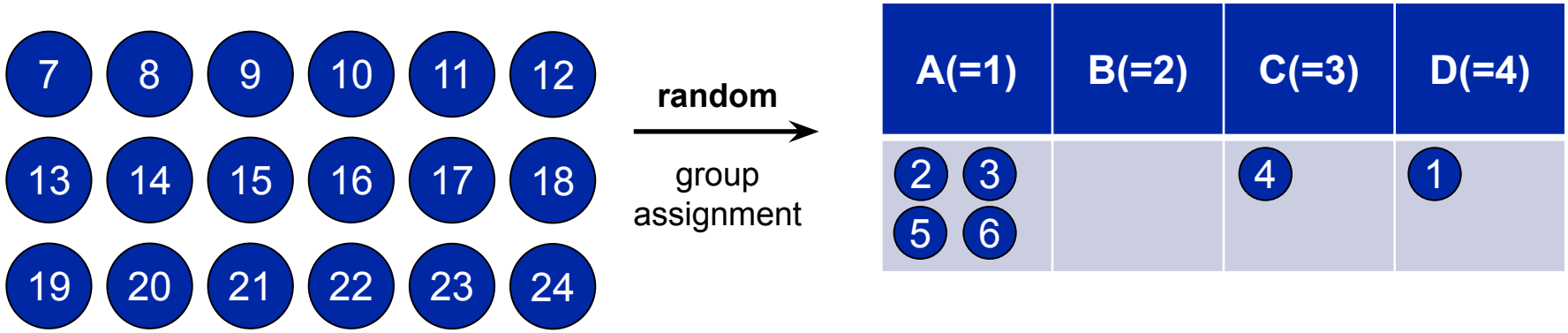
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

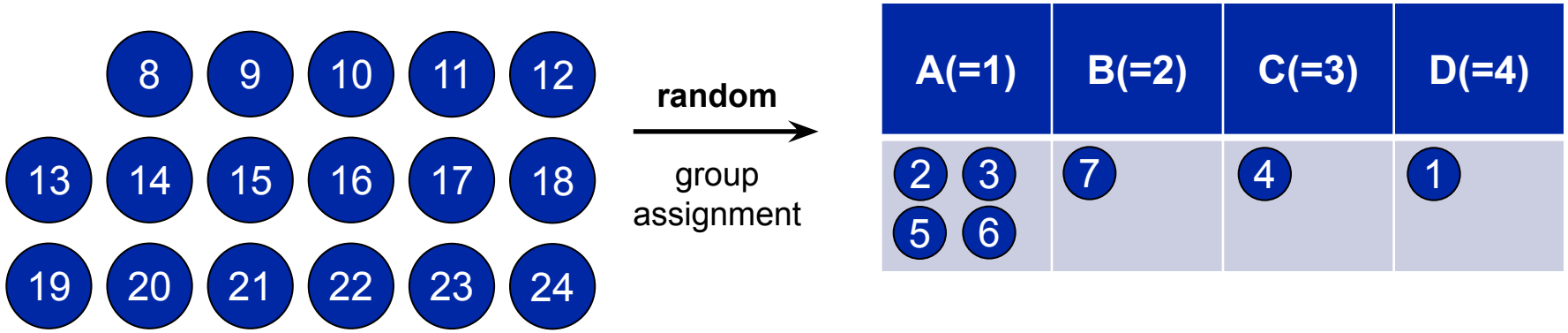
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

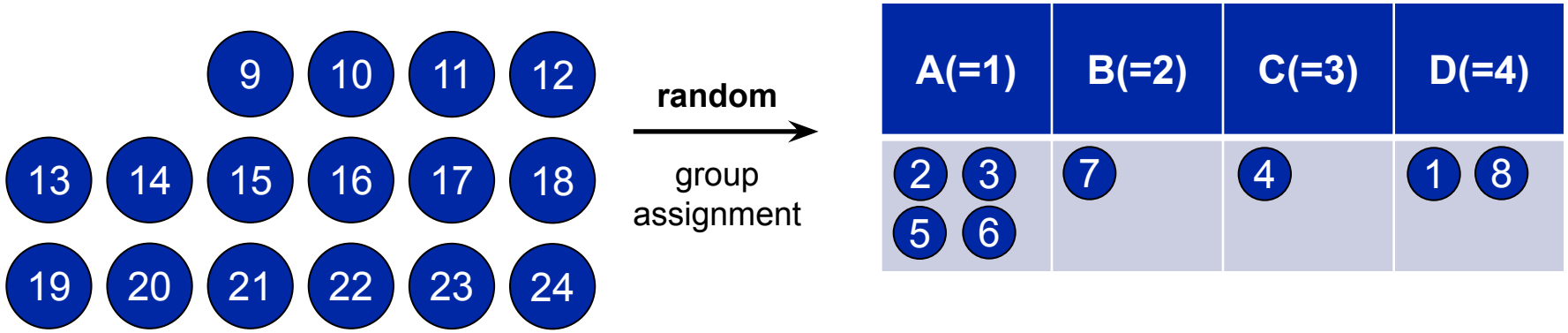
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

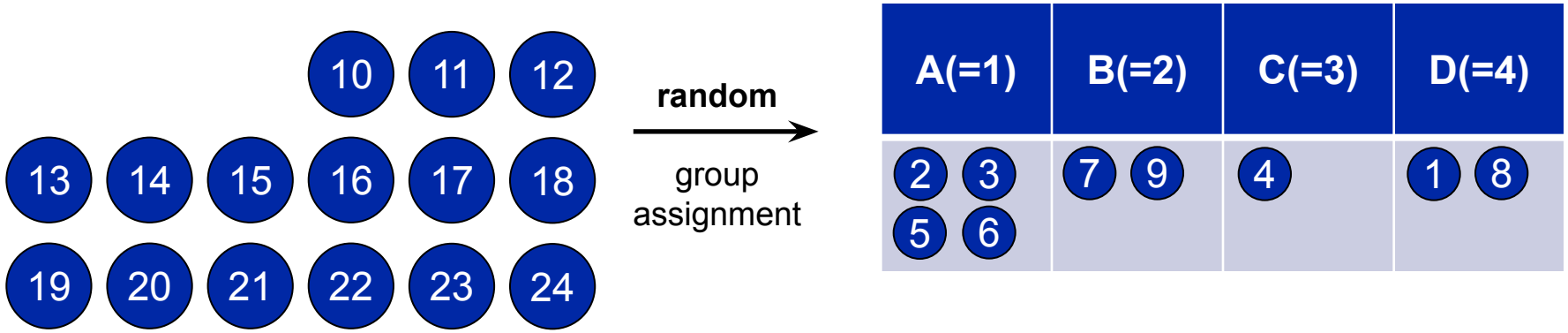
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

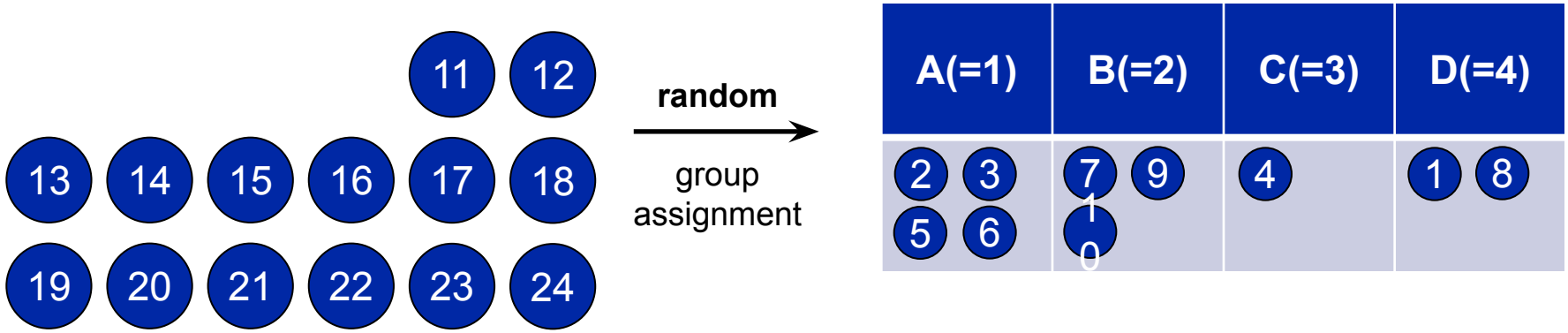
Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

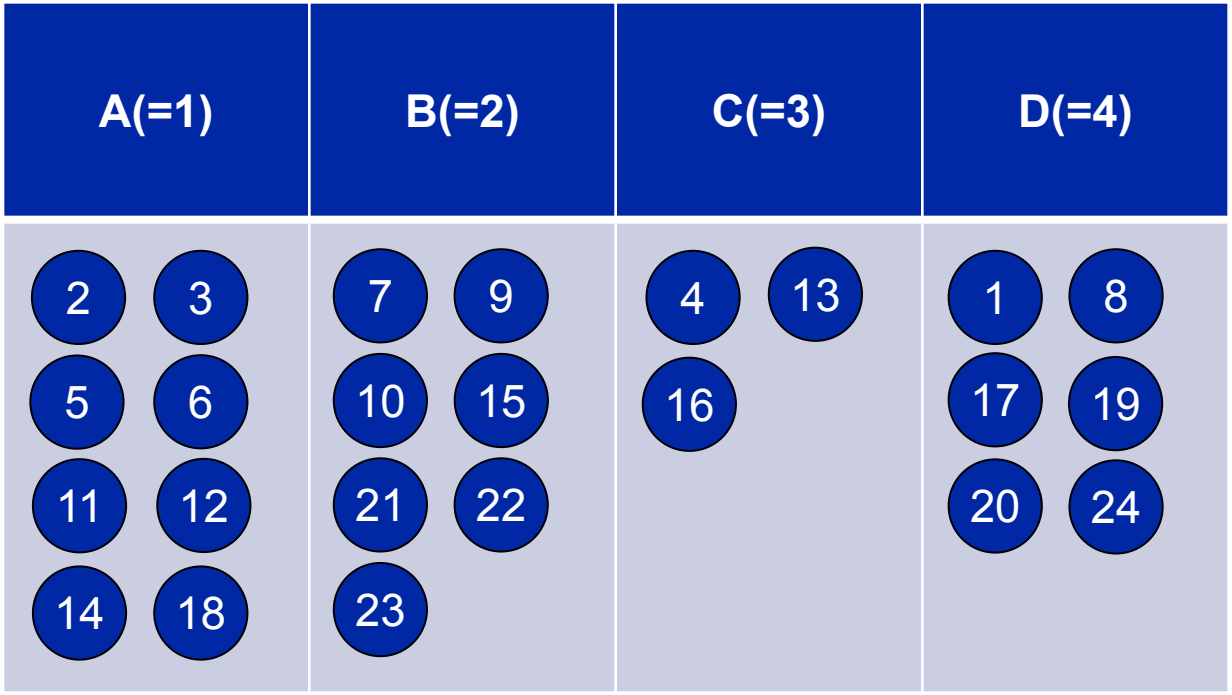
Simple randomization

Take each mouse and randomly assign it to one of the four groups



In R:
For each mouse: `sample.int(4,size=1,replace=F)`

Simple randomization

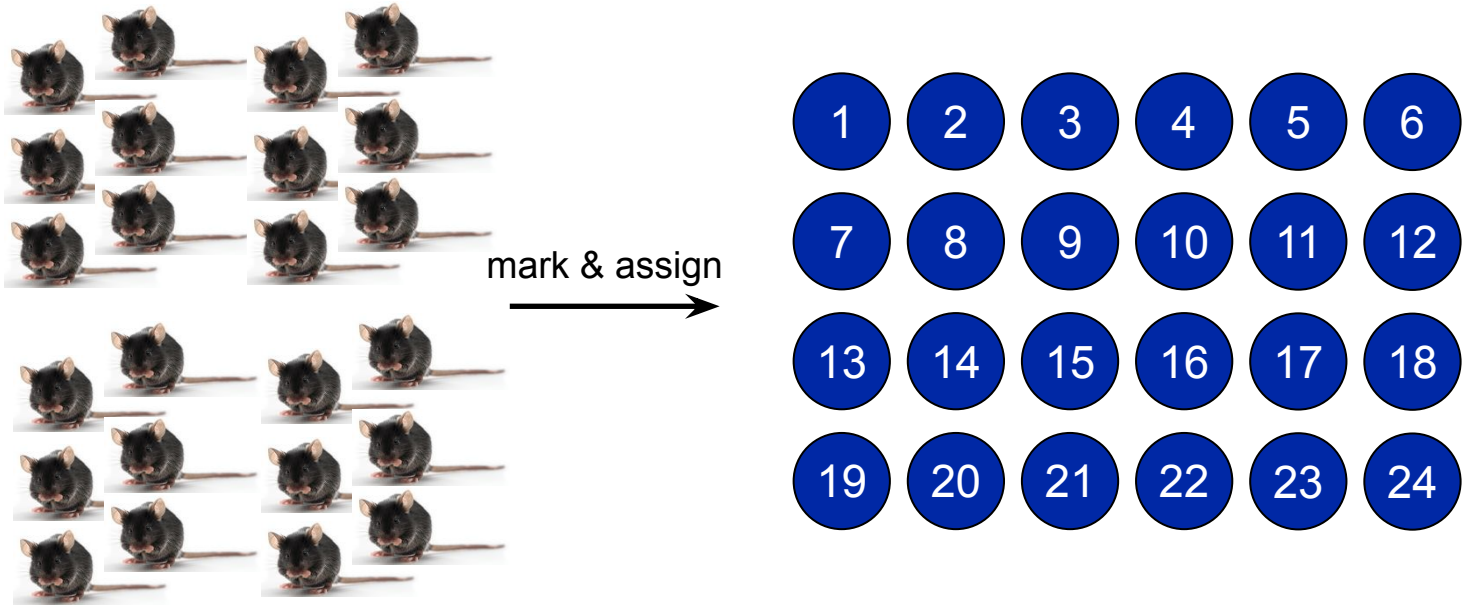


In R:

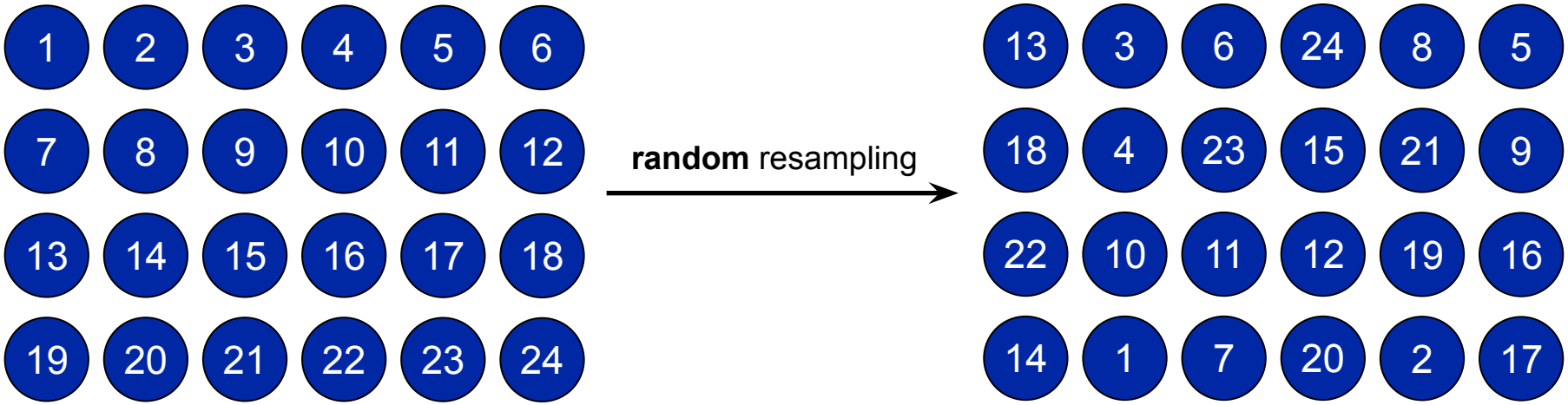
For all mice together: `sapply(1:24,function(x) sample.int(4,size=1,replace=F))`

Balanced randomization

- Remember example 2 (metabolic study) with 3 treatment groups and 1 control group; assume you have 24 mice.
- Goal: Assign each mouse randomly to one of the four groups, but ensure that group sizes are equal



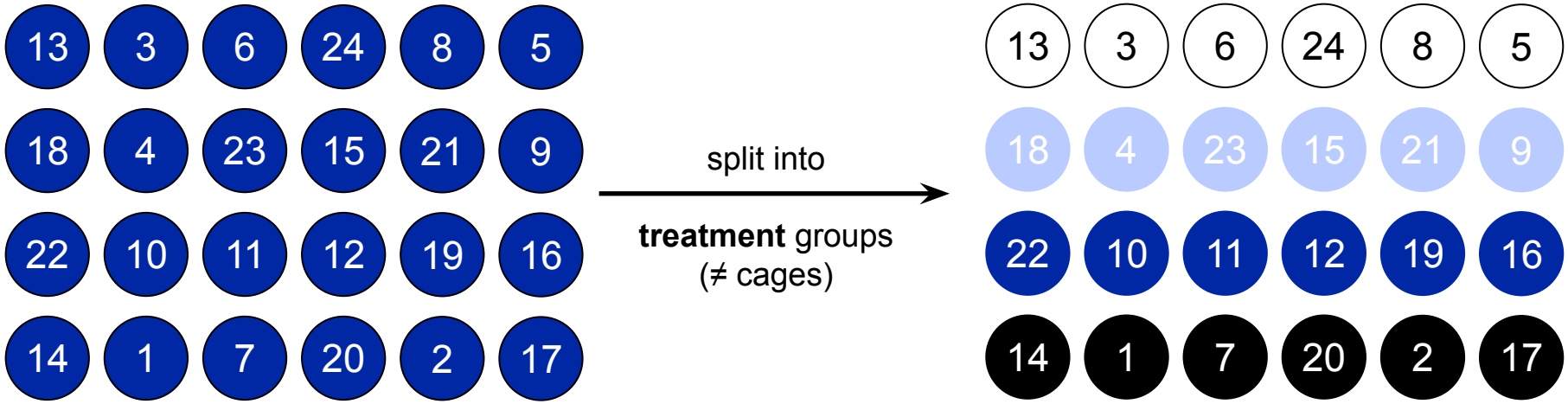
Balanced randomization



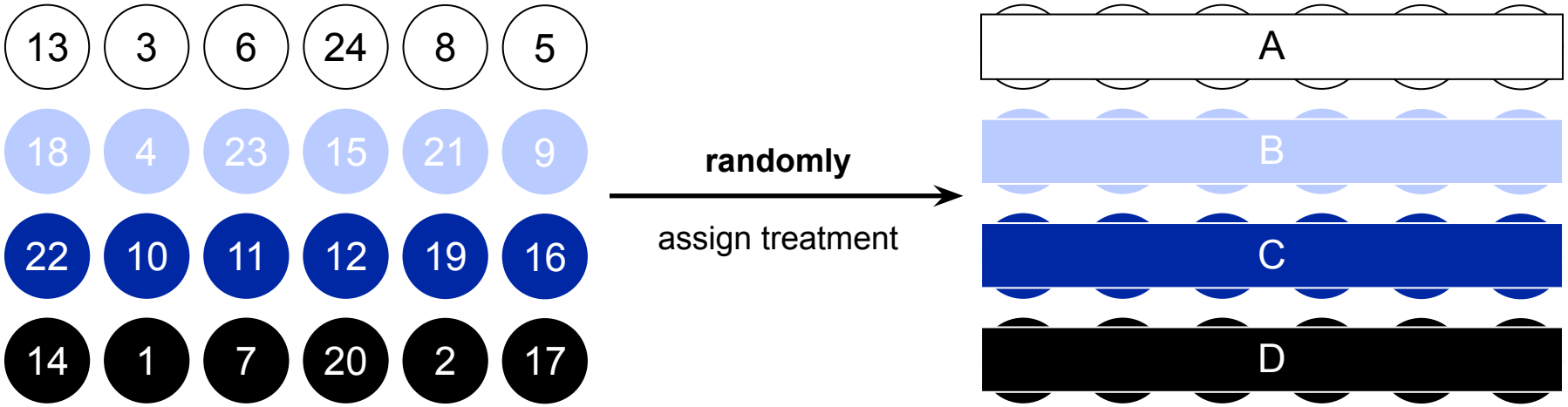
In R:

```
sample.int(n=24,size=24,replace=F)
```

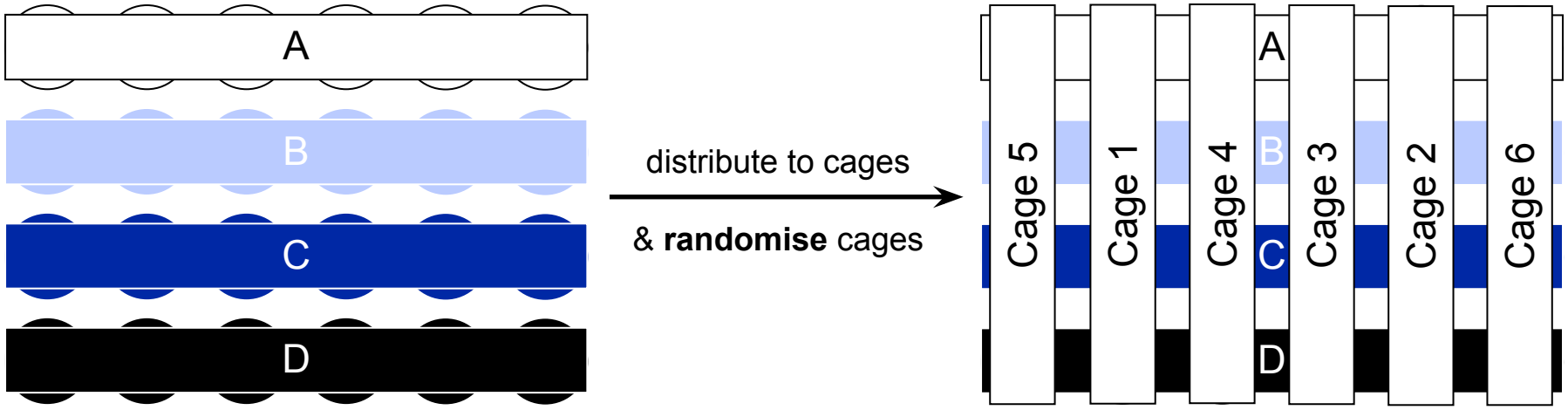
Balanced randomization



Balanced randomization

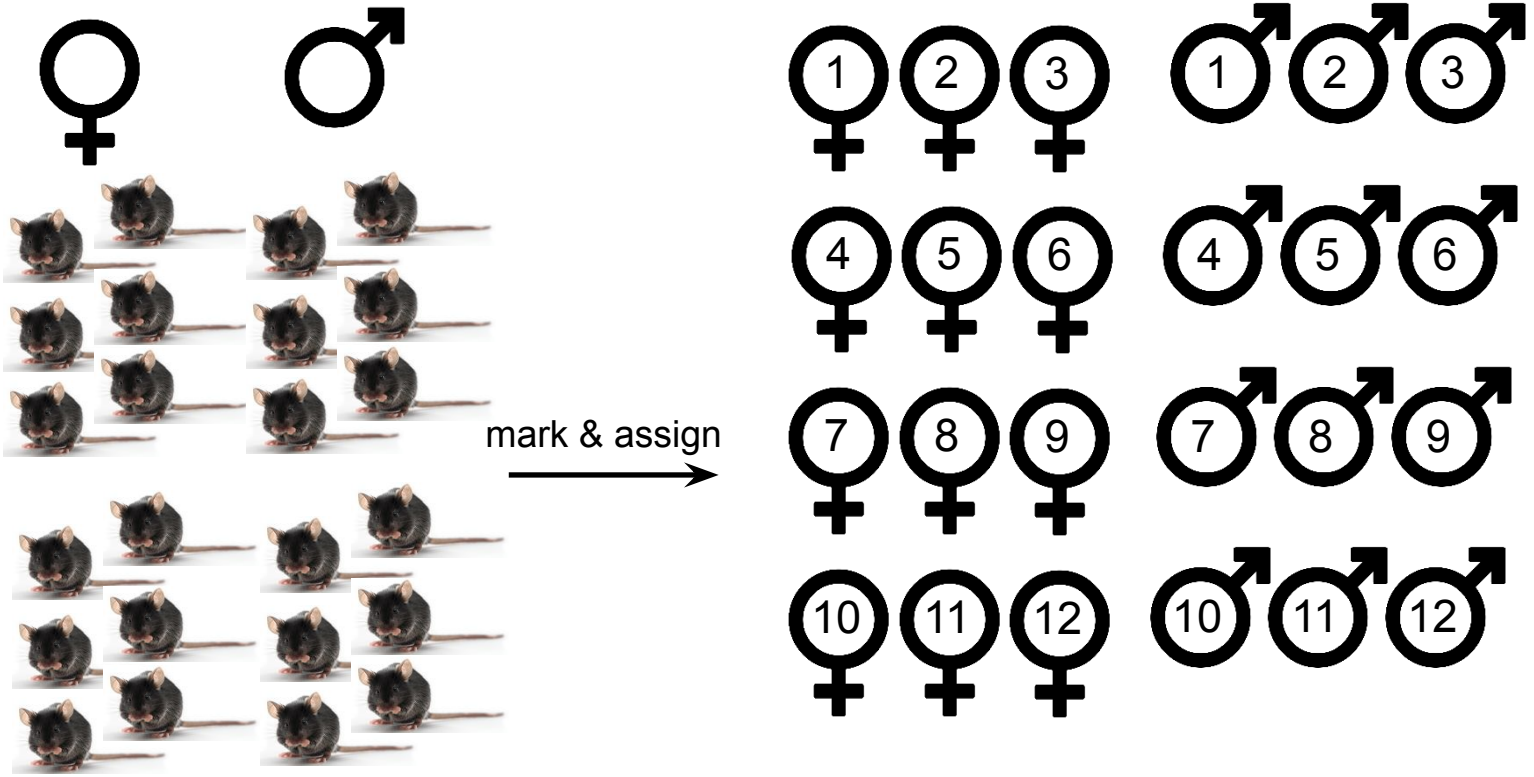


Balanced randomization

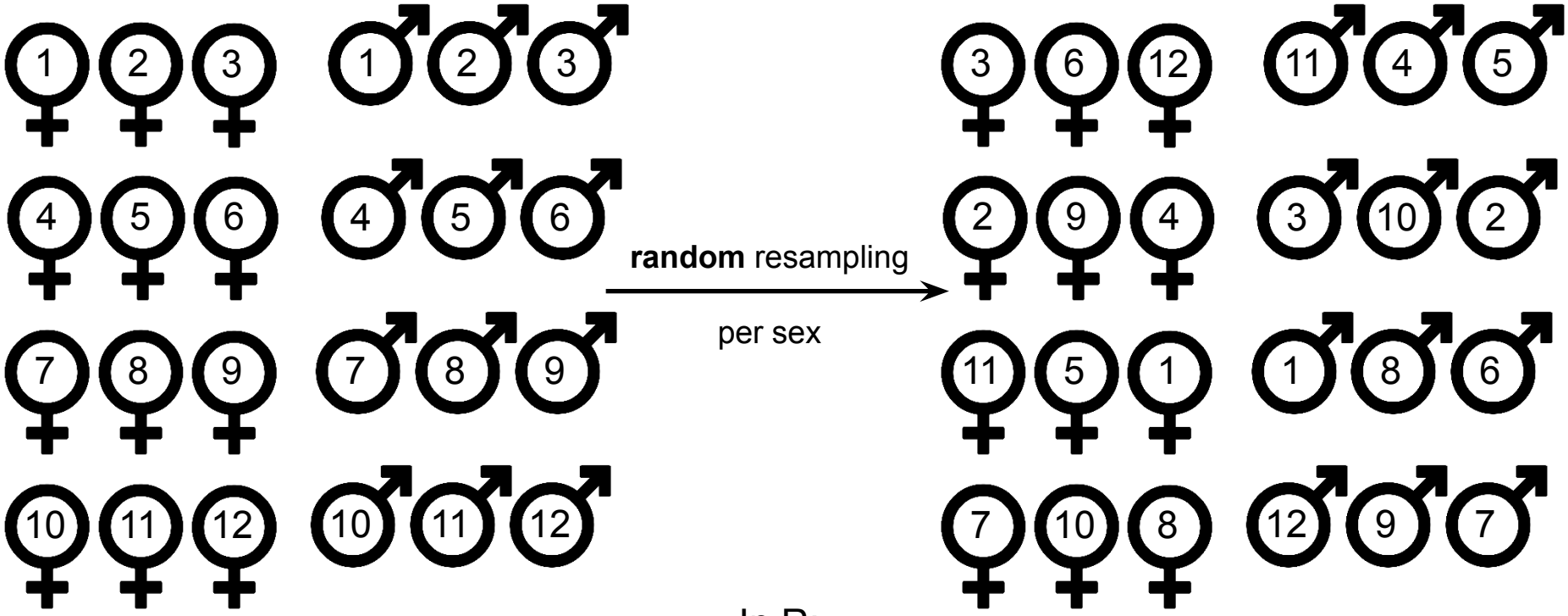


Stratified randomization

- Remember example 2 (metabolic study) with 3 treatment groups and 1 control group; assume you have 24 mice, 12 female and 12 male.
- Goal: Assign each mouse randomly to one of the four groups, but ensure that each group contains equal amount of both sexes (and all groups are equally sized)



Stratified randomization

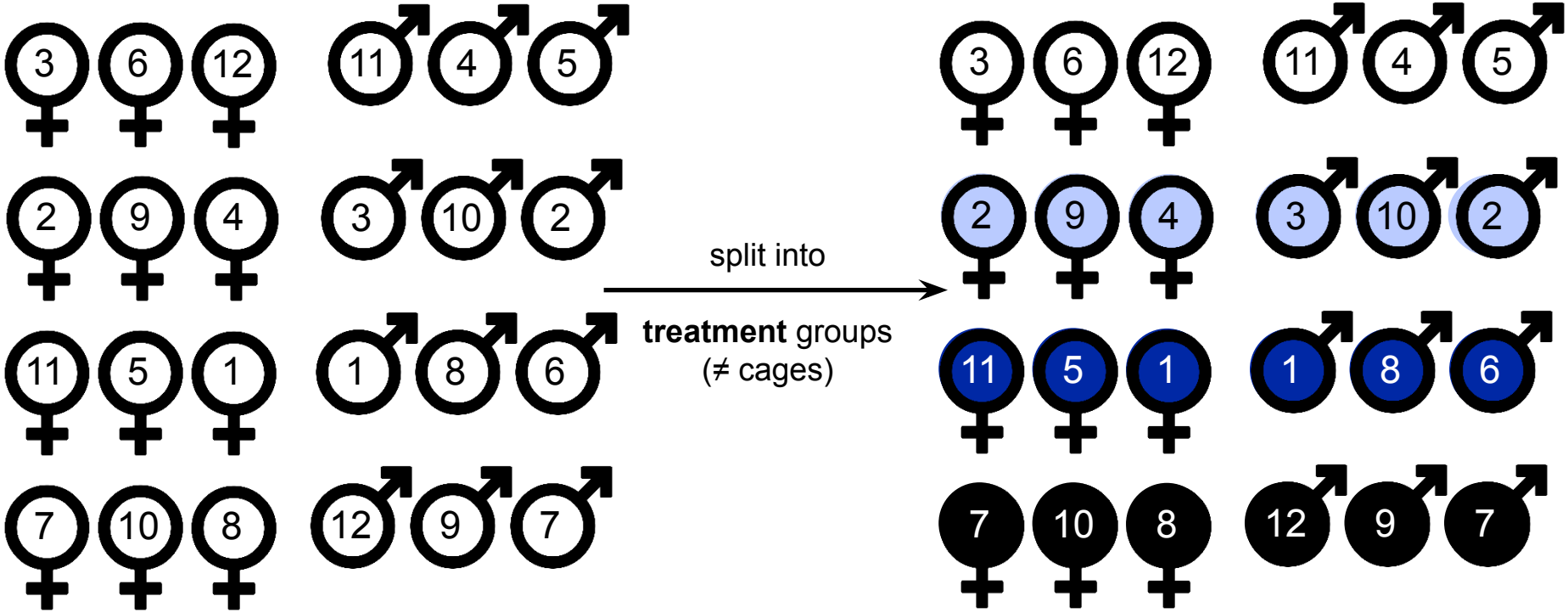


In R:

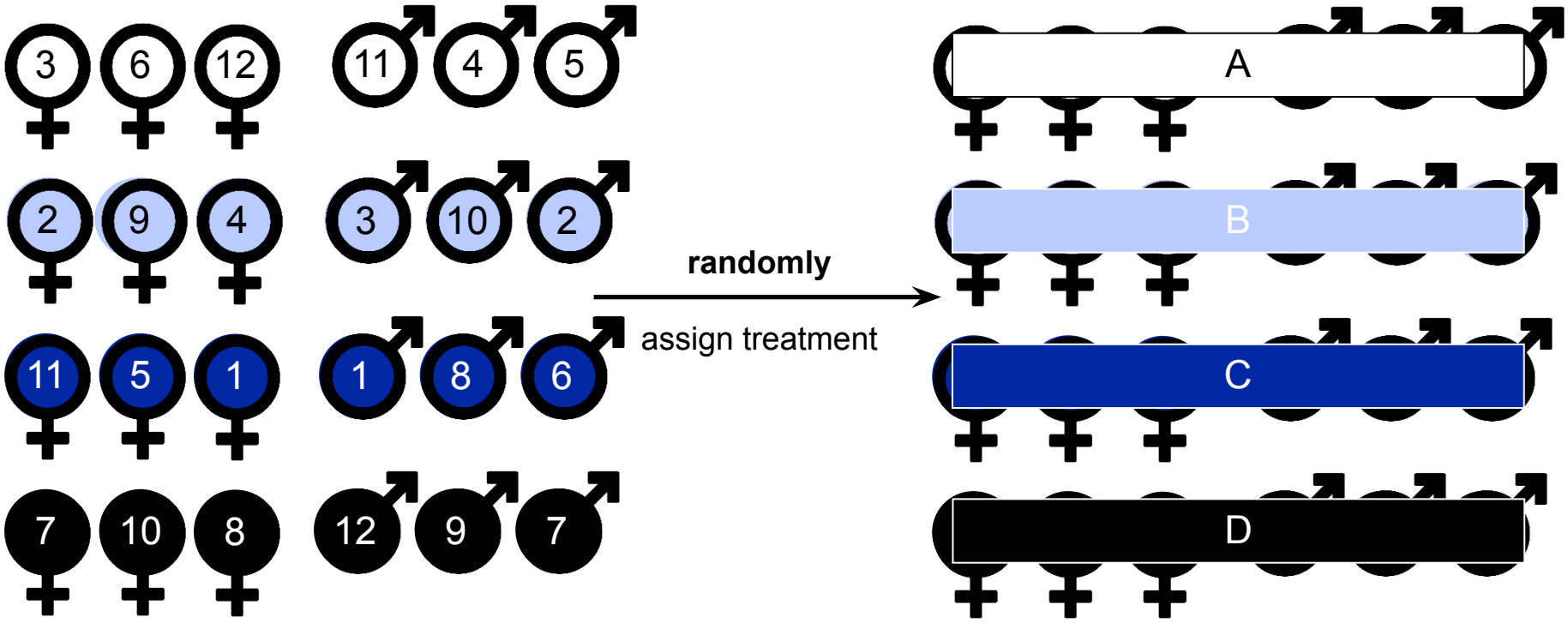
For each sex: `sample.int(n=12,size=12,replace=F)`

Stratified randomization

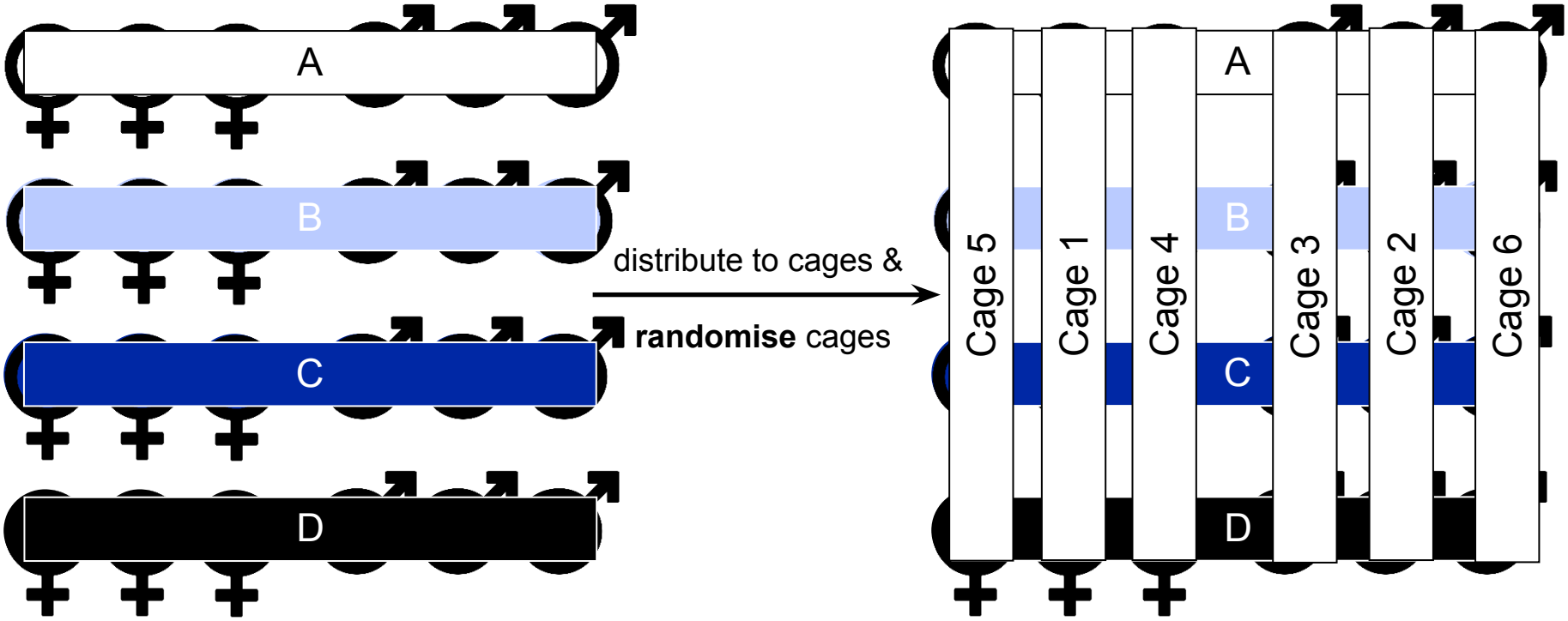
Stratified randomization (aka block randomization) is basically simple balanced randomization applied on the blocks/strata



Stratified randomization



Stratified randomization



Randomization – What if it's not so simple?

- In some cases, your treatment has to be allocated to the all animals of the same cage because co-housing of animals of different groups would lead to bias

- In these cases, you need to randomly distribute animals to cages and then randomly allocate one of the interventions to the cages.

- Be aware that in this case, the experimental unit is the cage, not the individual animal!

- This needs to be taken into consideration when calculating sample size and analysing the experiments. Two examples of how to do this:
  - Calculate sample size & analyse results on per-cage levels
  - Use hierarchical/nested/clustered designs

Design and analyse on a per-cage basis

Calculate sample size and analyse results on per-cage basis:

1. Estimate all relevant statistics (e.g. empirical mean, variance, biologically relevant effect size etc.) on the cage level with a given number of animals per cage.
2. Calculate number **of cages** that are needed to achieve the desired power.
3. When conducting a statistical test, use the mean measurement of each cage as the basis of your test.

- Advantage: Easy and straightforward, standard methods can be used
- Disadvantage: Interpretation of effect size is not straightforward; you lose a lot of information; in most cases, you do not address the biologically relevant question

Use hierarchical/nested/clustered designs

Calculate sample size and analyse results on a per-animal basis:

1. Estimate all relevant statistics (e.g. empirical mean, variance, biologically relevant effect size etc.) on the level of the individual animal.
2. Calculate number of animals that are needed to achieve the desired power.
3. Correct sample size for correlation of animals that are held within the same cage.
4. Analyse the data **including the cluster effect**!

- Advantage: Usually more powerful and insightful.
- Disadvantage: More complex -> Approach statistician for support!

Blinding

Conceal information about treatment groups from people involved in an experiment.

Blinding prevents a lot of systematic errors from biasing your results.

Who should be blinded?

- Experimenters administering treatments
- Experimenters assessing outcome
- Caregiver, animal facilities staff
- Others who interact with the animals

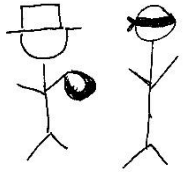
What information should be concealed?

- The experimental group when assessing the outcome
- The future treatment group of subjects during pre-treatment interventions such as inducing a lesion
- Previous values when multiple observations are made on the same experimental units or observational units
- The values of other outcomes from the same experimental units or observational units
- Information that provides clues about treatment groups

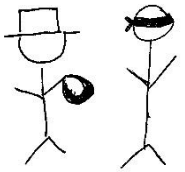
Blinding – It's easy!

1. Plan your experiment in an unblinded manner.

2. Ask a colleague to randomly assign your mice into different treatment groups and to give you a **coded** group allocation so that you do not kno which mice belong to the control group and which belong to the treatment group.

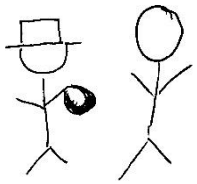


3. Ask a colleague to code all substances used for interventions etc. so that you do not know whether you are applying a treatment or a control substance.

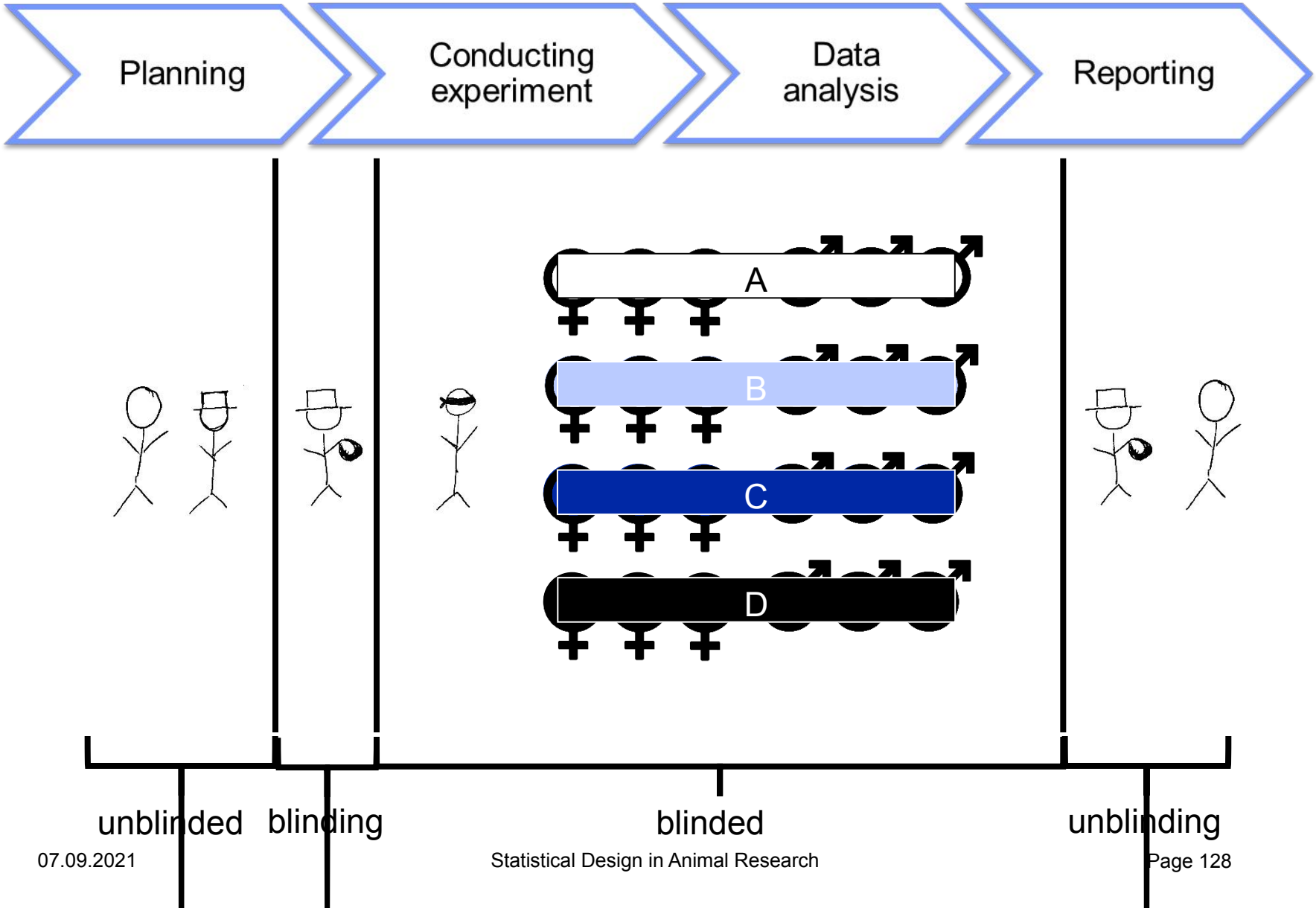


4. Keep blinding until the end of the experiment (including data analysis)!

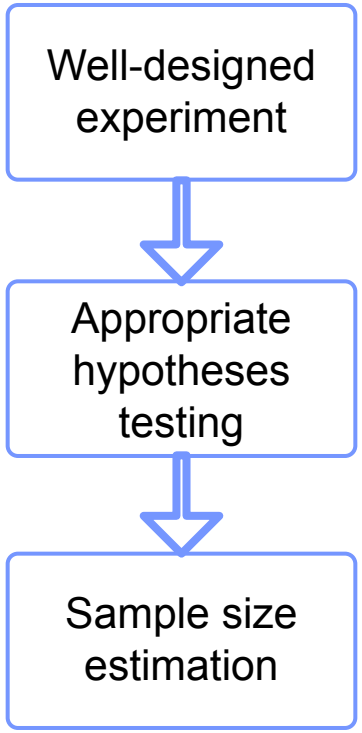
5. Unblind for reporting



Blinding – It's easy! (cont.)



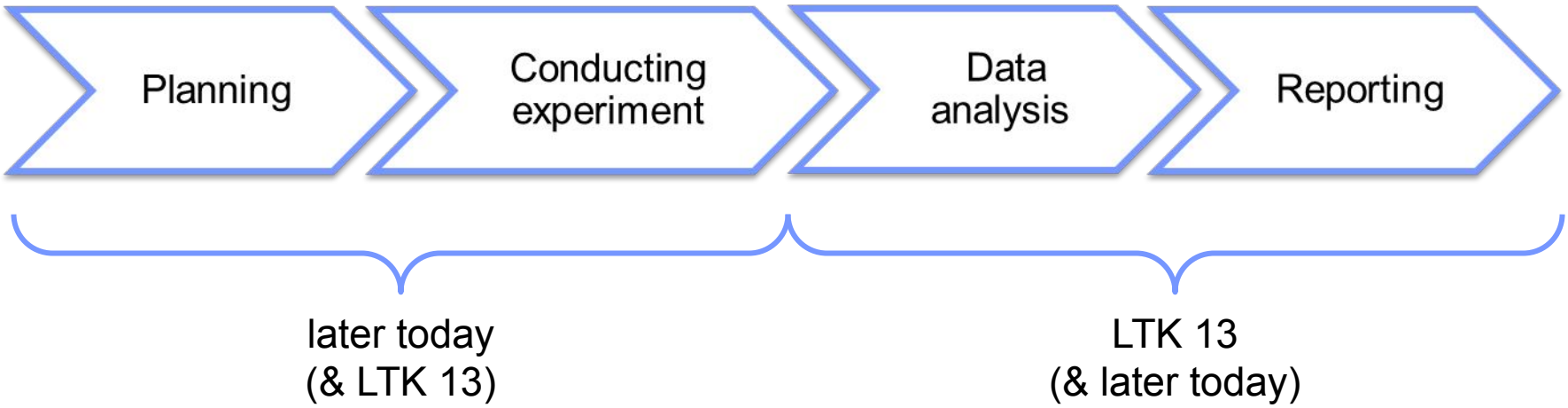
Recap: Planning for success



It provides the <u>strongest evidence</u> in support of <u>causal inference</u>

- Clear and focused hypotheses (PICO-B)
- Identification of source of variability/uncertainty
- Ensure unbiasedness (randomization, blinding)
- Choosing a good primary outcome
  - Objectively measured (biomarker vs subjective evaluation)
  - Measure close to where the action is
  - Unlikely to be missing or censored
  - High reliability
  - Normal distributed with constant variance
- Equal size of each group (balanced-design)
- Correct identification of experimental unit
- Wide range of applicability → blocking: deliberate variation
- Keep it simple but not simpler

Details of the four stages of experiment



Guidance document for animal research applications on
www.servangrueniger.ch/studydesign

Literature - How to plan, execute and analyse an experiment

Lazic, S.E. (2016). Experimental Design for Laboratory Biologists. Cambridge University Press. Chapter 3.

Lazic SE, Clarke-Williams CJ, Munafò MR (2018) What exactly is 'N' in cell culture and animal experiments?. PLOS Biology 16(4): e2005282. https://doi.org/10.1371/journal.pbio.2005282

Bate & Clark (2014), The Design and Statistical Analysis of Animal Experiments, 2014 https://www.cambridge.org/core/books/design-and-statistical-analysis-of-animal-experiments/BDD758F3C49CF5BEB160A9C54ED48706

Van der Worp, H. B., Howells, D. W., Sena, E. S., Porritt, M. J., Rewell, S., O'Collins, V., & Macleod, M. R. (2010). Can animal models of disease reliably inform human studies?. PLoS med, 7(3), e1000245.