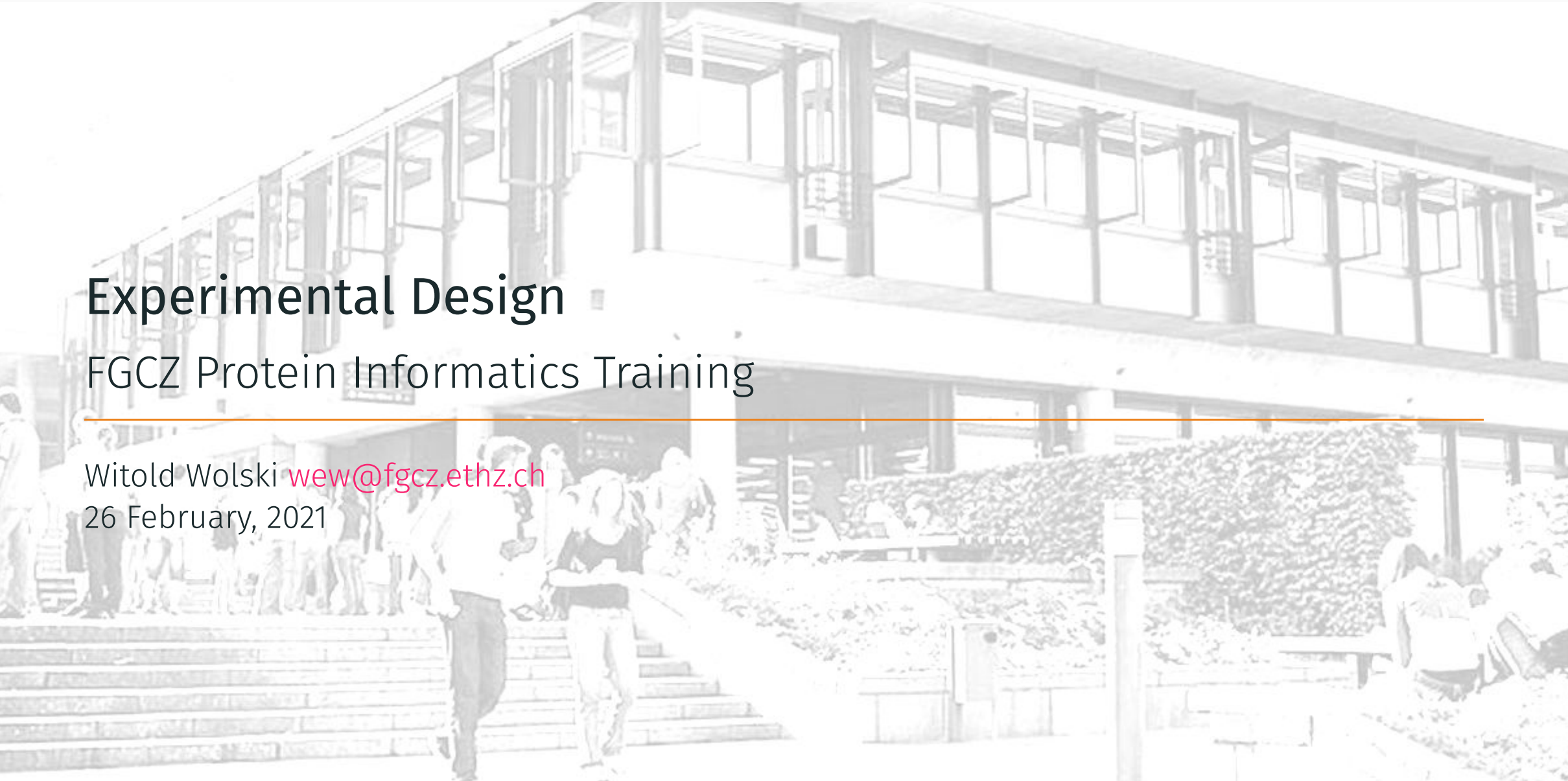


Experimental Design

FGCZ Protein Informatics Training

Witold Wolski wew@fgcz.ethz.ch

26 February, 2021


design

Overview

- Scientific method
- Causality and Bradford Hill criteria
- Study Types
- Randomized Controlled Trial
- Types of Bias
- Confounders
- Ethical Considerations
- Sample size estimation
- Experimental Designs

Scientific Method

1. **Question**
2. **Hypothesis**
3. **Experiment**

   ordered investigation that attempts to prove or disprove a hypothesis
   - must show if hypothesis is supported or not.
   - results must be measurable
   - experiment must be repeatable
4. **Observation**

   make observations about results of experiment
5. **Analysis**

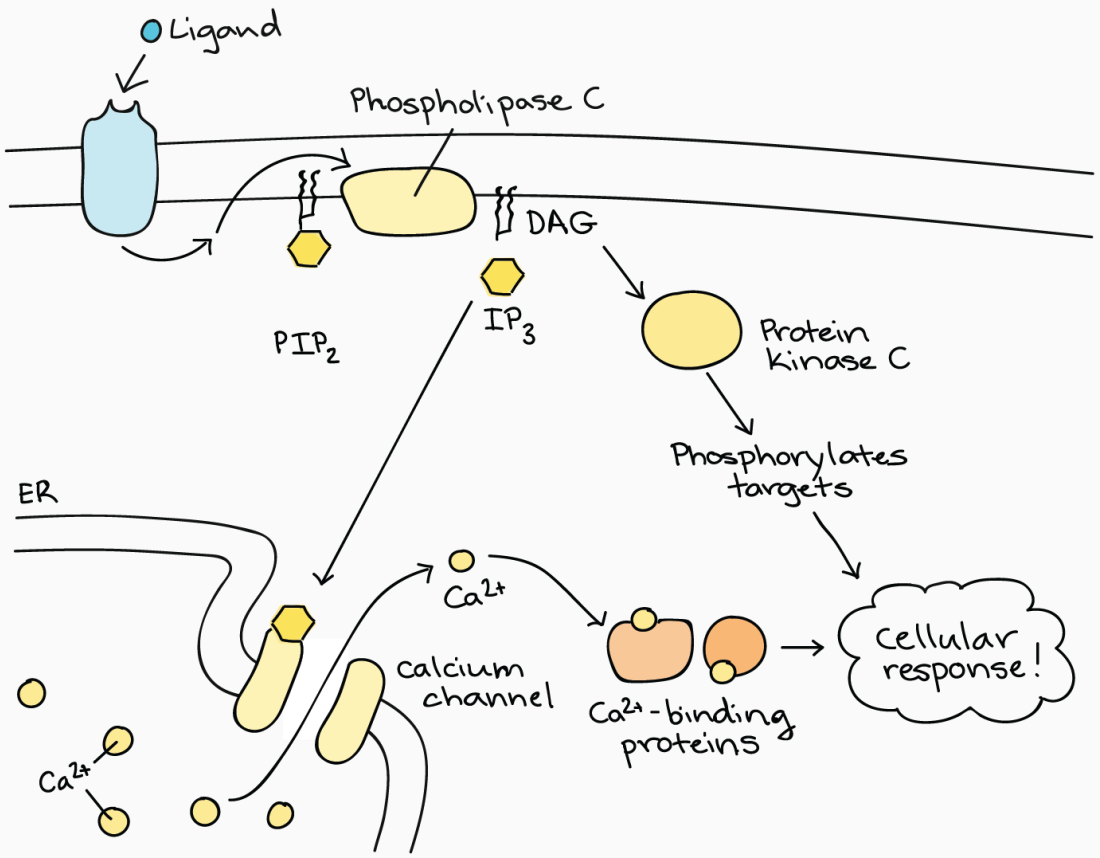
   run test
6. **Conclusion**

   significant result

Question - Causality

Assess whether a particular **agent**

caused or influenced a particular **outcome**



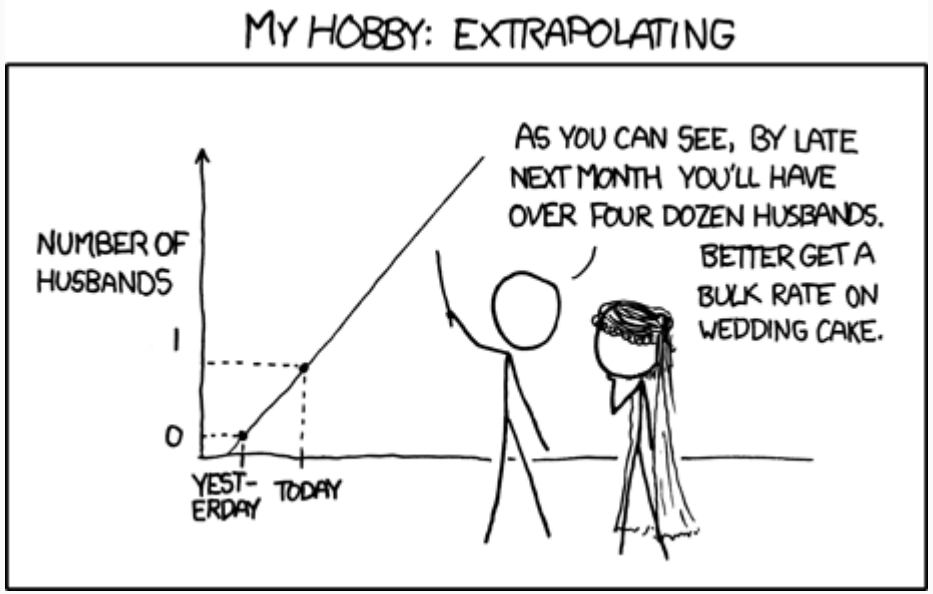
Causality - Bradford-Hill Criteria

1. **Strength of association** – (effect size) the greater the effect compared with those not exposed to the agent the more plausible is the association
2. **Consistency** – (reproducibility) does it happen in other groups of people – both men and women, different countries
3. **Specificity** – no other likely explanations
4. **Temporality** – effect follows cause, and if expected delay, effect must occur after that delay
5. **Biological gradient** – the stronger the agent the greater the effect
6. **Plausibility** – is there a possible biological mechanism that could explain the effect
7. **Coherence** – do different types of study result in similar conclusions – e.g., controlled trials and observational studies
8. **Experiment** "Occasionally it is possible to appeal to experimental evidence".

Sir Austin Bradford Hill CBE FRS was an English epidemiologist and statistician, pioneered the randomized clinical trial (use of streptomycin in treating tuberculosis) and, together with Richard Doll, demonstrated the connection between cigarette smoking and lung cancer (case-control study).

Bradford-Hill Criteria

Plausibility



Specificity

Explaining Hill Criteria using xkcd

The Bradford Hill Criteria Don't Hold Up

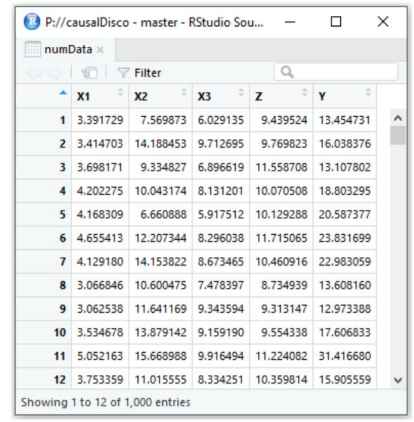
Question - Causality - Bradford-Hill Criteria

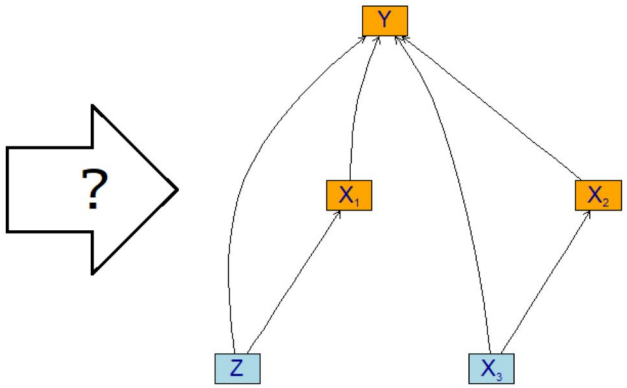
- Does establishing all Bradford-Hill criteria prove cause and effect?
- Which Bradford-Hill criteria do you aim to meet in your study?

- Protein Quantification
  - Strength of association : Control, Treatment $5\mu g$, Treatmant $10\mu g$
  - Consistency : Different Genetic backgrounds, Control and Treatment
  - Temporality : Time course data
  - Specificity : e.g. test for interactions.

Question - Causality

- Can we infer causal models directly from data?
- Would you prefer?
  - to have dataset with 4000 proteins and 2 conditions
  - a dataset with 400 proteins and more than > 100 conditions?





CausalDisco an overview of methods (implemented in R) to infer causal networks from data

Judea Pearl "The Book of Why: The New Science of Cause and Effect"

Study Types

“Test your servants for ten days: Give us nothing but vegetables to eat and water to drink. Then compare our appearance with that of the young men who eat the royal food, and decide what to do with us based on how we look.”

Bible (see Daniel 1:1–16).

Bregman, Rutger. Utopia for Realists (p. 206). Bloomsbury Publishing. Kindle Edition.

Study Types

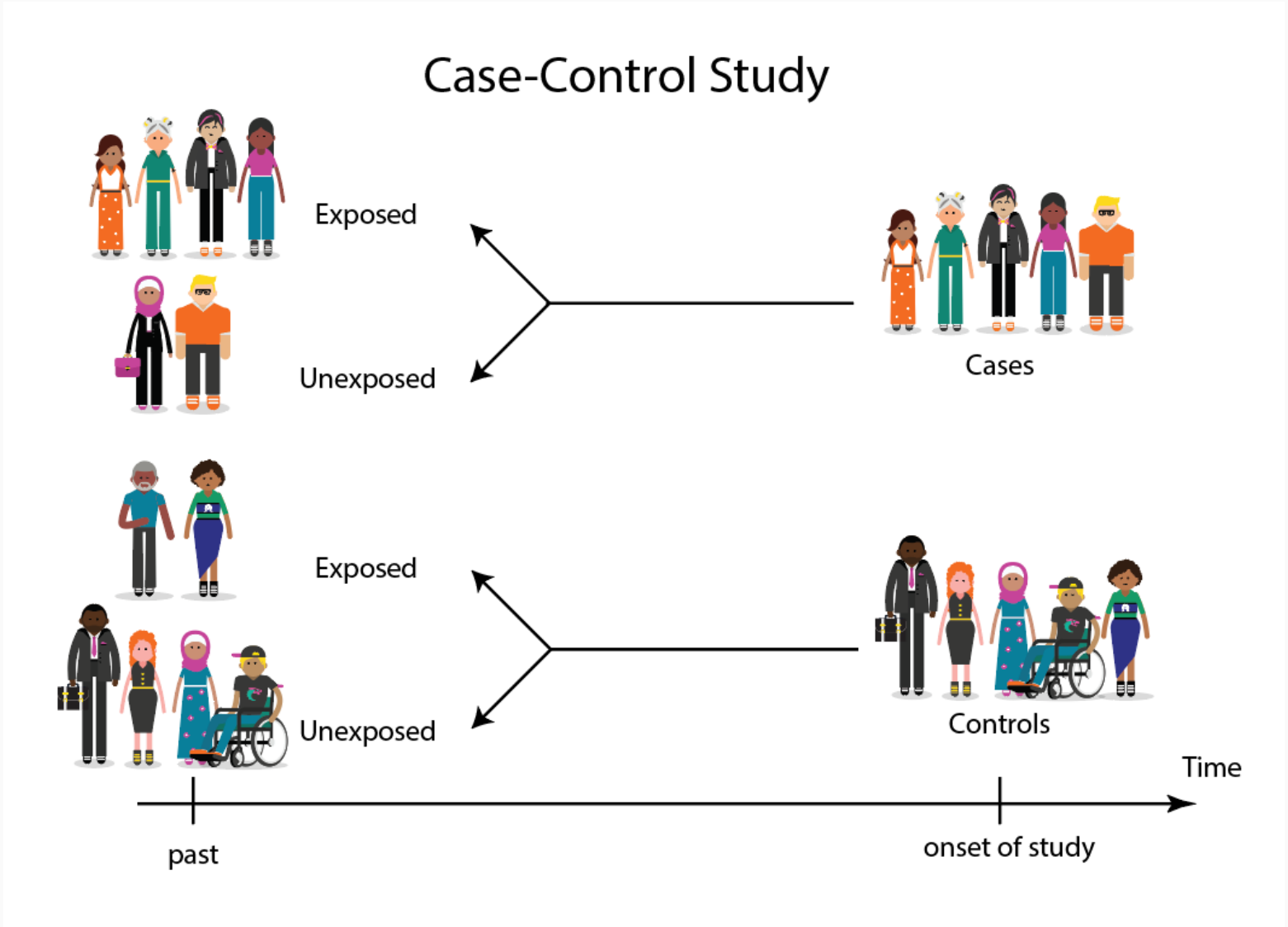
Observational study

- Case-control study - individuals with a specific characteristic (disease and similar individuals without disease)
- Cross-sectional study - aim to provide data on the entire population under study.
- Longitudinal study - repeated observations of the same variables (e.g., people) over short or long periods of time. They can also be structured as longitudinal **randomized experiments**, e.g. Mice aging.

Experimental Intervention

- randomized controlled trials (randomize subjects into two groups, placebo and treatment group)

Study Types - Case control study



Study Types - Pros and cons

Observational study

pros:

- May require less resources
- Less ethical considerations (e.g. smokers)
- Good if outcome of interest is rare

cons:

- difficult to determine causality
- no randomization or blinding
- The exposure status is not determined by the researcher

Experimental Intervention

pros:

- More validity
- Can determine causality
- Randomized and blinded

cons:

- may require more resources
- Ethical concerns for certain exposures
- Difficult if outcome studied is rare (e.g. Vaccination)

Randomized controlled trial - Randomization

A randomized controlled trial is a type of scientific experiment that aims to reduce certain sources of **bias** when testing the effectiveness of new treatments.

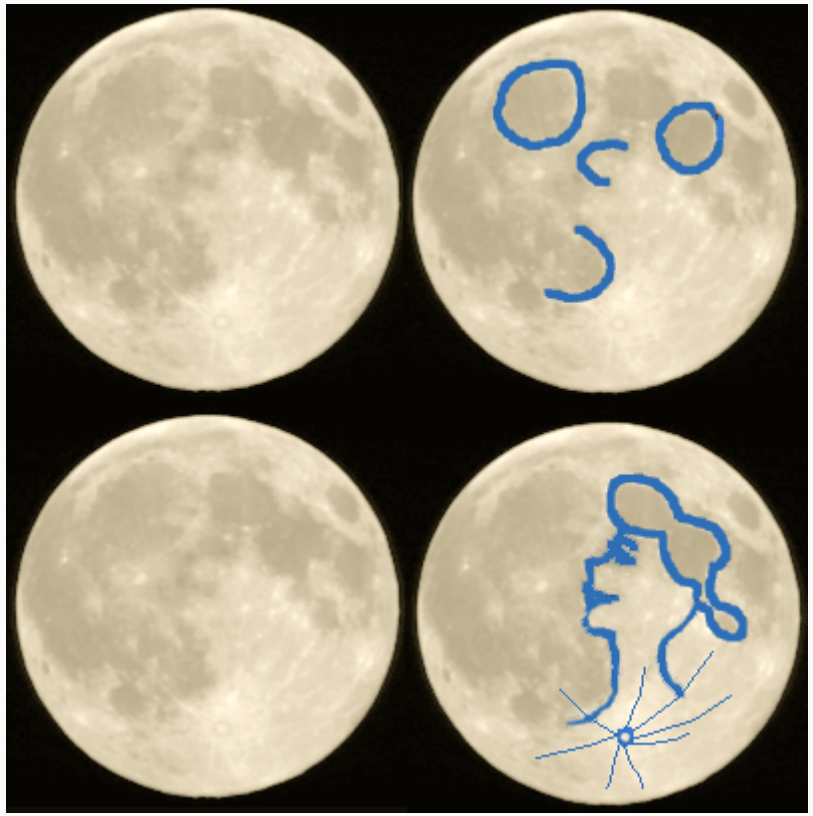
- **selection bias**
  - sample obtained is *not representative* of the population intended to be analyzed.
- **allocation bias**
  - *systematic difference* in how participants are assigned to treatment groups and comparison groups
  - **Randomly allocate subjects to two or more groups**

https://en.wikipedia.org/wiki/Randomized_controlled_trial
https://en.wikipedia.org/wiki/Bias

Bias

- Confirmation bias,
  tendency to favor information that confirm hypothesis *

- Funding bias,
  bias relative to the commercial interests of a study's financial sponsor
- Publication bias,
  bias towards publication of certain experimental results.



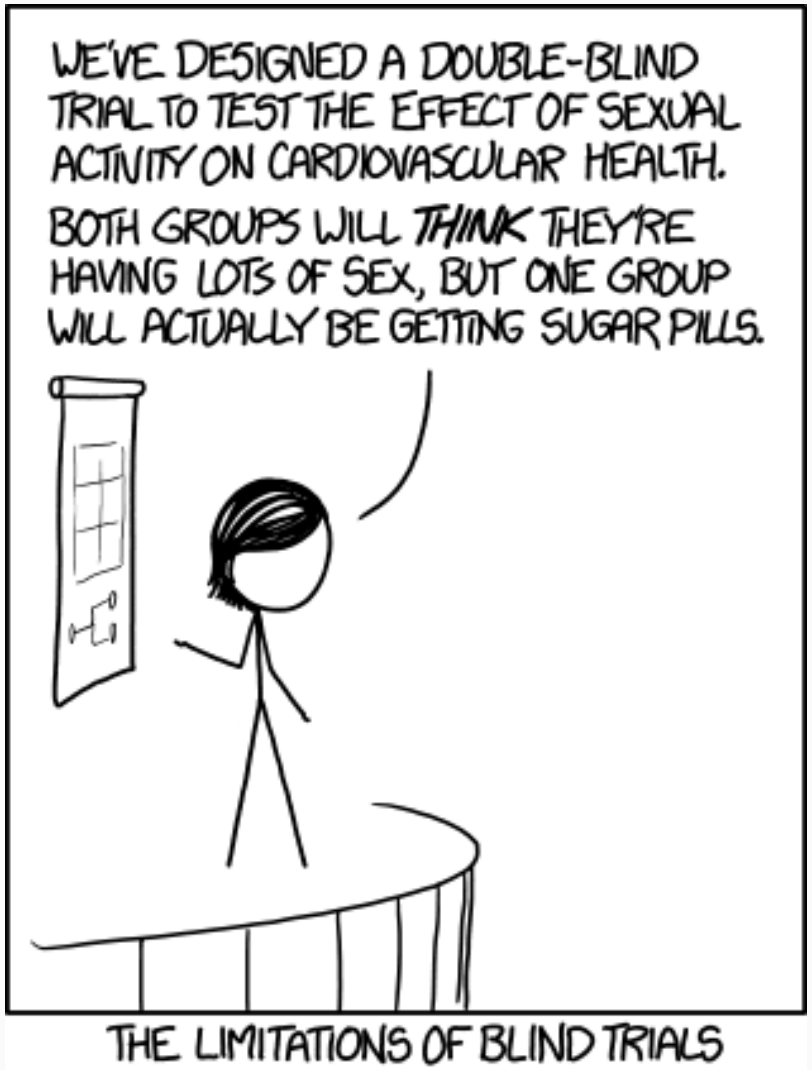
* Do you examine an experiment for a technical problem which fits your hypothesis?

** Over-interpretation of improvements of new methods.

Randomized controlled trial - Blinding

- Information which may influence the participants is withheld until after the experiment is completed
- A blind can be imposed on any participant of an experiment, including **subjects, researchers, technicians, data analysts, and evaluators**

- In clinical Trials double blind trials should be used where possible
  - **Single blind** – either patient or evaluator blind.
  - **Double blind** – both patient and evaluator blind

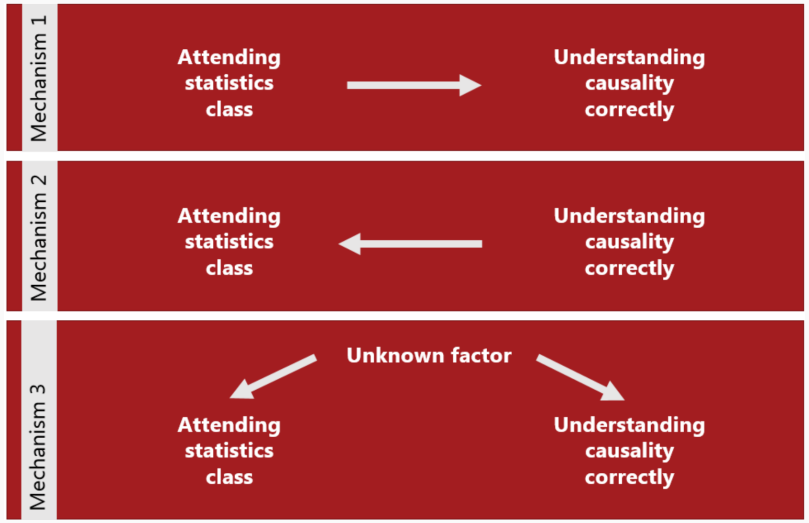


Confounders

Correlation does not imply causation but causation may imply correlation.

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

Reichenbach's common cause principle:
A correlation occurs due to one of the three possible mechanisms

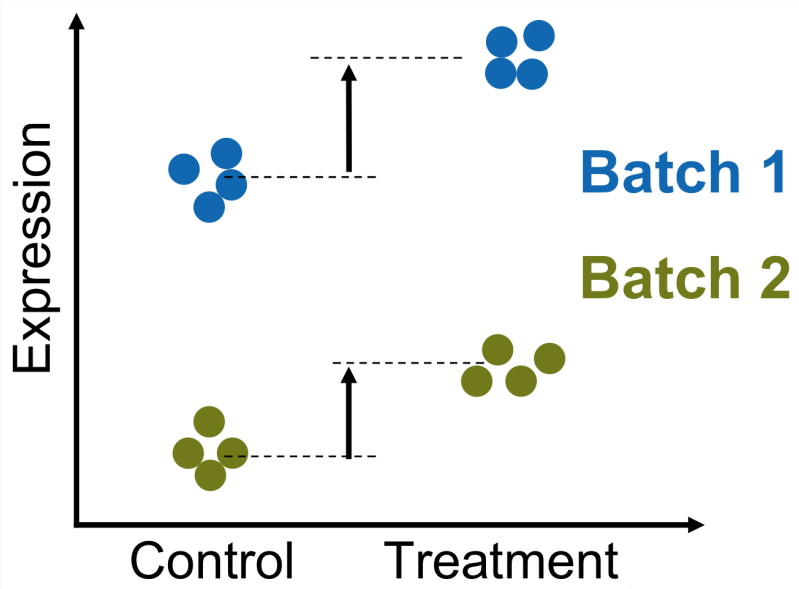


Mediators, confounders, colliders – a crash course in causal inference

Controlling for Confounders

- Randomize over biological and technical co-variates.
  e.g. run Id, age.
- Avoid batch effects:
  - process all samples in parallel
- block confounders
  complete block design, i.e.,
  all treatments are present in each batch in equal numbers.

https://doi.org/10.1021/acs.jproteome.0c00536]]



How to avoid confounding by batch

https://doi.org/10.1021/pr8010099

Controlling for Confounders

Queue Generator

Use block random queue when processing samples to *block* for changes in the chromatographic column over time.

Example:

# of samples in each condition 3 x A, 3 x B, 2 x C, 2 x D.
Blocks : ABCD, BCDA, BA

http://fgcz-ms-shiny.uzh.ch:8080/queue_generator10/

Confounders - Reporting

Document possible co-variates, e.g.,

- Human subjects:
  - age
  - bmi
  - gender
  - ...
- Biochemical/Technical:
  - batch
  - run Id
  - instrument
  - ...

- Report your research so that it can be reviewed, reproduced and analysed.
- Do not over-interpret your results.

Over-interpretation and misreporting of prognostic factor studies in oncology: a systematic review
http://www.equator-network.org/
https://www.ncbi.nlm.nih.gov/pubmed/29873743

Ethical considerations

- It is agreed that it is **unethical** *to conduct research which is badly planned or executed*.
- It is unethical to perform a trial which:
  - has *many more subjects* than are needed to reach a conclusion.
  - has little prospect of reaching any conclusion,
    e.g. because of *insufficient numbers of subjects* (or some other aspect of poor design)
- The local ethics committee has discretion on how it will supervise **non-interventional** studies
  - US - Institutional Review Board (IRB)
  - EU - ethics committees

Declaration of Helsinki (1964+amendments))

Types of error

A **type I error** (false positive) occurs when
the null hypothesis (H0) is true, but is rejected.
The *type I error rate* or **significance level** (p-Value)
is the probability of rejecting the
null hypothesis given that it is true.

A **type II error** (false negative) occurs when
the null hypothesis is false,
but erroneously fails to be rejected.
The *the type II error rate* is denoted by the Greek letter β
and is related to the **power of a test** (which equals $1 - \beta$).

For a given test, the only way to reduce both error rates
is to **increase the sample size**, and this may not be feasible.

		reality	
		H_0 = true	H_0 = false
conclusion	H_0 is not rejected	OK	type II error
	H_0 is rejected	type I error	OK

Sample size calculation

Run pilot experiment and measure the coefficient of variation (CV) or σ^2 of replicates:

- technical
- *biochemical*
- **biological**

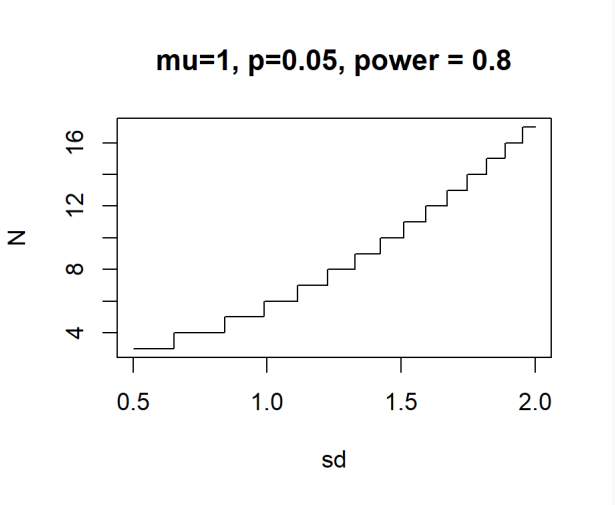
- *biological variance >> bio-chem+tech variance*
  - Only provide biological replicates
- *bio-chem+tech variance >> biological variance*
  - **improve sample handling and preparation**
  - choose different technology
  - buy better instrument

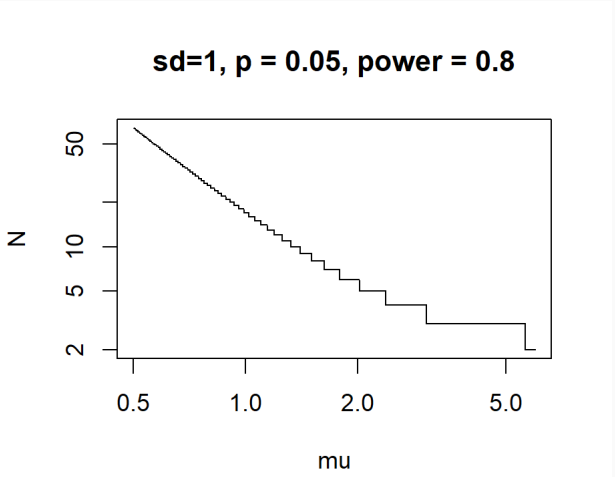
measure 3 tech replicates, 3 tech+bio-chem replicates, 3 tech+bio-chem+biological replicates

Sample size calculation

For each statistical test,
there exists a unique relation between:

- desired smallest detectable effect size μ
- sample variance σ^2
- sample size N
- critical p-value p_0
- statistical power





Top - greater standard deviation requires larger sample sizes, Bottom - smaller effect size requires larger sample sizes

Experimental Design

The key technical issue is whether comparisons are made **between** or **within** subjects.

Between

- Parallel Group Design
  k - groups, n_i patients in group i receive treatment i
- Factorial Design
  - more than one factor. combining treatments, e.g. A and B to same patient.

Within (repeated/paired)

- In series design
  each patient all k treatments in same order
- crossover design
  each patient all k treatments in different order

Are combinations of both possible?

Ronald Fisher (1890 - 1962) "The Arrangement of Field Experiments" (1926) and "The Design of Experiments" (1935).

Factorial Design vs Parallel group

Parallel (one factor)

- drug A
- drug B
- plac

Factorial (2 or more factors)

- Factor A
  - drug A
  - plac A
- Factor B
  - drug B
  - plac B

40 patients, Placebo, drug A and drug B. How would you allocate these patients?

Ronald Fisher pioneered the use of factorial experiments instead of the one-factor-at-a-time method.

Factorial Design vs Parallel group

Parallel (one factor)

- drug A (13)
- drug B (13)
- plac (14)

Compare:

- 13 drug A vs 14 plac
- 13 drug B vs 14 plac

Factorial (2 or more factors)

- drug A + drug B (10)
- plac A + drug B (10)
- drug A + plac B (10)
- plac A + plac B (10)

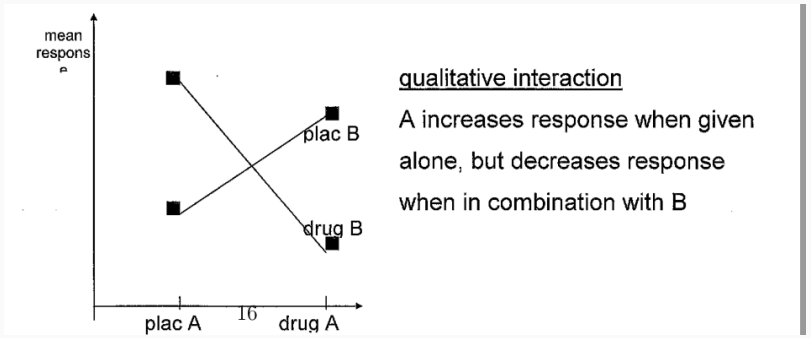
Compare:

- 20 drug A vs 20 plac A
- 20 drug B vs 20 plac B

Factorial Design - Interactions

Factorial Designs are efficient at evaluating the effects and possible interactions of several factors (independent variables).

- No interaction
- Quantitative interaction
- Qualitative interaction



[Medical Statistics - University of Sheffield]

https://en.wikipedia.org/wiki/Design_of_experiments

Summary

Did you observe bias in research? What is the type of your study? What are potential confounders? Which Bradford Hill Criteria are you examining and how? What is the Design of your experiment?