Choice of design and methods

Forskerinitierte kliniske studier: fra idé til publikasjon

Maiju Pesonen

Clinical Trial Unit, Oslo University Hospital (OUS)

Oslo Center for Biostatistics and Epidemiology (OCBE), University of Oslo

maiju.pesonen@medisin.uio.no

Oslo universitetssykehus

UiO : University of Oslo

Why: research question

Experiments conducted to answer a clinical research question

Are concomitant intra-articular glucocorticoid injections of joints with active arthritis in JIA patients starting Tumour Necrosis Factor inhibitor treatment increase the proportion of JIA patients reaching sustained, inactive disease, when compared to the control group not receiving joint injections?

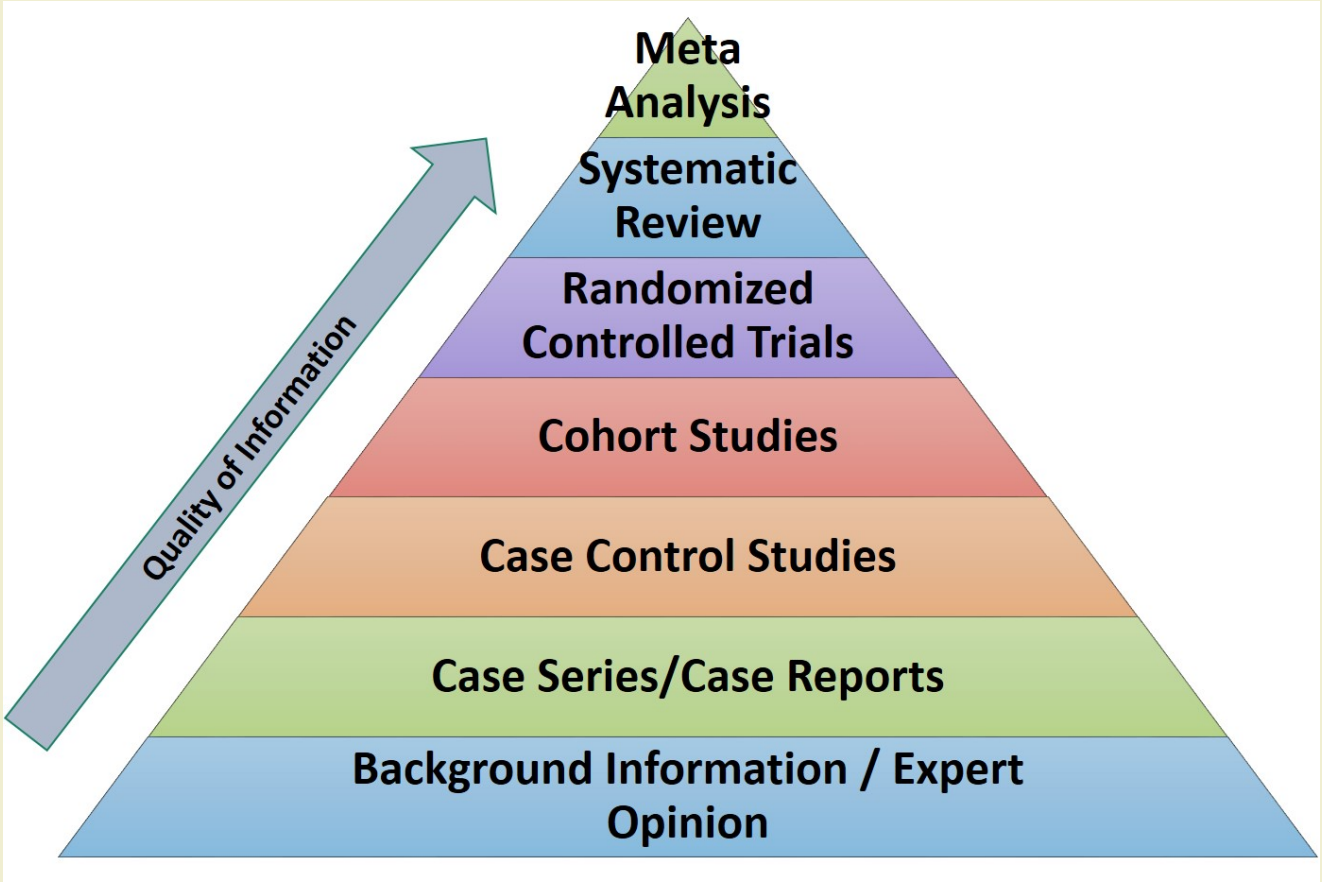
Is antibiotic therapy beneficial in children with pneumonia who do not have a highly suspicious bacterial infection, such as in lobar pneumonia?

Why: research question

Experiments conducted to answer a clinical research question

- The objective of our work
- Well defined and specific
- Supported by data
- It affects all subsequent decisions and tasks

RCTs in the context



RCTs in the context

- Pyramid of evidence: ranking of types of studies based on the strength of the evidence they provide
- RCTs are considered the gold standard
- Strongest evidence in establishing causality

RCTs in the context

- Pyramid of evidence: ranking of types of studies based on the strength of the evidence they provide
- RCTs are considered the gold standard
- Strongest evidence in establishing causality
- IF the RCT is correctly planned and correctly analysed!

What is a RCT?

A study to assess the efficacy of a new treatment for some conditions

What is a RCT?

A study to assess the efficacy of a new treatment for some conditions

Randomized Controlled Trial

What is a RCT?

A study to assess the efficacy of a new treatment for some conditions

Randomized Controlled Trial

$\rightarrow$ Controlled: the new treatment is given to the treated group, and another treatment is given to the control group

What is a RCT?

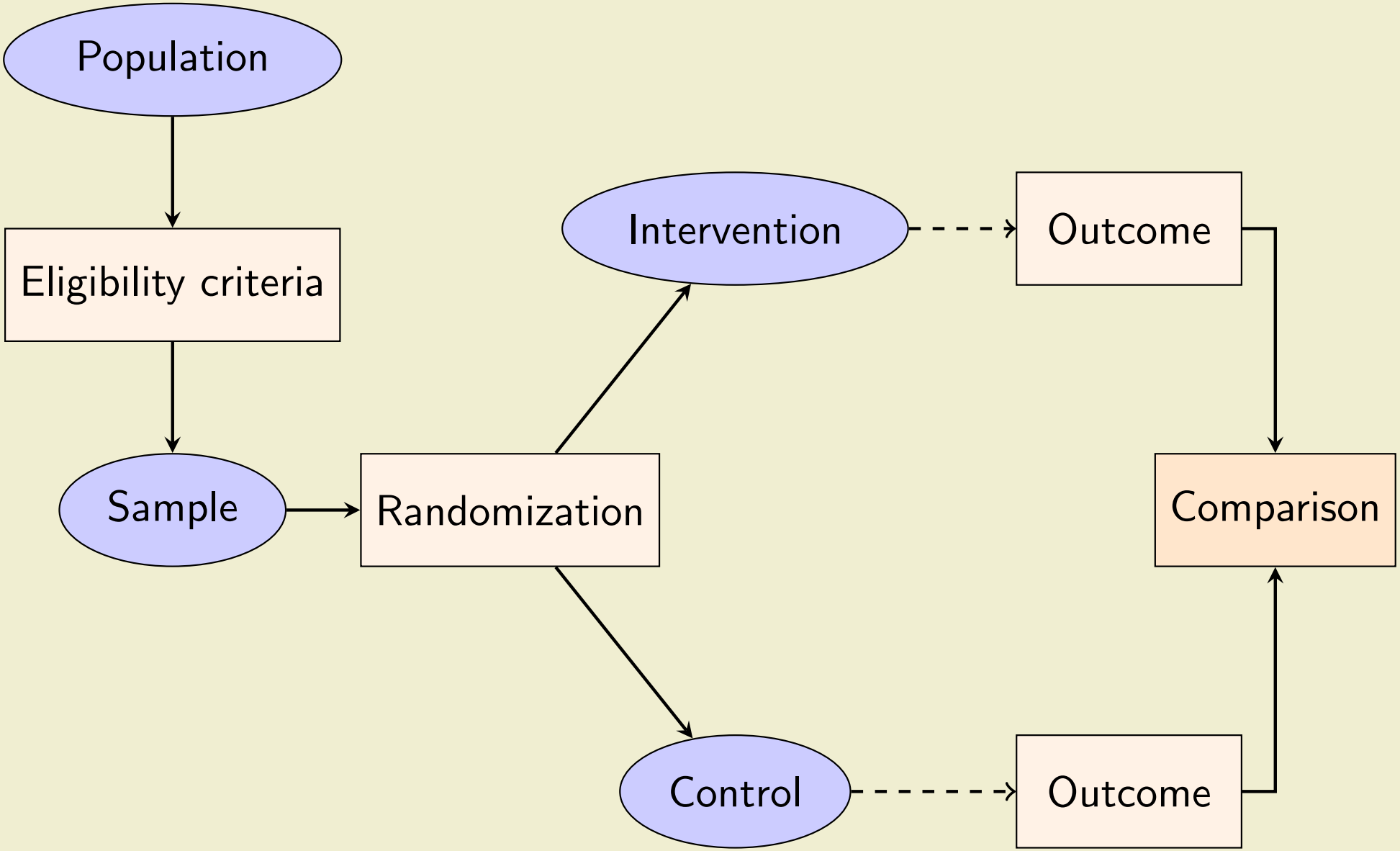
A study to assess the efficacy of a new treatment for some conditions

Randomized Controlled Trial

→ Controlled: the new treatment is given to the treated group, and another treatment is given to the control group

→ Randomized: Patients are allocated to one of the two groups by randomization

Structure of an RCT



Main ingredients: PICO

Population

Intervention

Controls

Outcome

P I C O: Population

A sample of eligible patients

P I C O: Population

A sample of eligible patients

- Type and stage of the disease
- Age
- Gender
- Type of surgery/intervention received before
- ...

P I C O: Population

A sample of eligible patients

- Type and stage of the disease
- Age
- Gender
- Type of surgery/intervention received before
- ...

$\rightarrow$ To generalize results to all patients who are similar

P I C O: Intervention and controls

- Intervention = treatment being tested
  drug, practice, surgical intervention, diagnostic tool, ...

P I C O: Intervention and controls

- Intervention = treatment being tested
  drug, practice, surgical intervention, diagnostic tool, ...
- Control = treatment used as comparison
  standard of care or placebo

P I C O: Intervention and controls

- Intervention = treatment being tested
  drug, practice, surgical intervention, diagnostic tool, ...
- Control = treatment used as comparison
  standard of care or placebo
- We need controls to detect differences only due to treatment allocation, all the rest is kept unchanged

P I C O: Intervention and controls

- Intervention = treatment being tested
  drug, practice, surgical intervention, diagnostic tool, ...
- Control = treatment used as comparison
  standard of care or placebo
- We need controls to detect differences only due to treatment allocation, all the rest is kept unchanged

$\rightarrow$ Allocation at random!

Randomization techniques

- Simple randomization
- Random permuted blocks ("block randomization") to avoid size imbalance across groups
- Stratified randomization to avoid imbalance in prognostic factors
  - Choose 2-3 factors at most
  - Important prognostic factors measured at baseline
  - Will have to be accounted for in the analysis

P I C O: Outcome

- Primary outcome ($\boldsymbol{X}$) determines the result of the study: usually one, max two or three
- Secondary outcomes in addition: can be many

P I C O: Outcome

- Primary outcome ($\boldsymbol{X}$) determines the result of the study: usually one, max two or three
- Secondary outcomes in addition: can be many
- How is the outcome measured in practice?
  medical device, assessment by nurse, patient reported, ...

P I C O: Outcome

- Primary outcome ($\boldsymbol{X}$) determines the result of the study: usually one, max two or three
- Secondary outcomes in addition: can be many
- How is the outcome measured in practice?
  medical device, assessment by nurse, patient reported, ...
- Side effects? Adverse events?

P I C O: Outcome

- Primary outcome ($\boldsymbol{X}$) determines the result of the study: usually one, max two or three
- Secondary outcomes in addition: can be many
- How is the outcome measured in practice?
  medical device, assessment by nurse, patient reported, ...
- Side effects? Adverse events?
- Assessment bias: Blinding!

Some more definitions

- New international trial regulatory guidance (ICH E9): addendum on Estimands and Sensitivity Analyses in Clinical Trials
- Estimand: precise description of the (treatment) effect that the study wants to find

Some more definitions

- New international trial regulatory guidance (ICH E9): addendum on Estimands and Sensitivity Analyses in Clinical Trials
- Estimand: precise description of the (treatment) effect that the study wants to find
- This is the well-defined research question

Some more definitions

- New international trial regulatory guidance (ICH E9): addendum on Estimands and Sensitivity Analyses in Clinical Trials
- Estimand: precise description of the (treatment) effect that the study wants to find
- This is the well-defined research question
- Has to contain information on:
  - Population
  - Treatment conditions
  - Outcome variable
  - Strategies for handling intercurrent events
  - Population level summary measure for outcome

Some more definitions

- New international trial regulatory guidance (ICH E9): addendum on Estimands and Sensitivity Analyses in Clinical Trials
- Estimand: precise description of the (treatment) effect that the study wants to find
- This is the well-defined research question
- Has to contain information on:
  - Population
  - Treatment conditions
  - Outcome variable

  } PICO
  - Strategies for handling intercurrent events
  - Population level summary measure for outcome

Handling of intercurrent events

- Intercurrent events: events that occur after the intervention and that can preclude the observation of the outcome variable or affect its measurement
- Can be related to disease or intervention, or completely unrelated
- Examples: patients who discontinued treatment or took rescue medication

Handling of intercurrent events

- Intercurrent events: events that occur after the intervention and that can preclude the observation of the outcome variable or affect its measurement
- Can be related to disease or intervention, or completely unrelated
- Examples: patients who discontinued treatment or took rescue medication
- Several possible strategies:
  - Treatment policy: event is irrelevant, the outcome is used regardless
  - Hypothetical: as if the event didn't occurr
  - Composite: occurrence of event is part of the outcome definition
  - While on treatment: the treatment effect is of interest only before the occurrence of an intercurrent event
  - Principal stratum: target population is set where event did not occurr

Population level summary measure for outcome

- How to present the results
- How to show difference between treatments
- Depends on the outcome definition:
  - Binary outcome $\rightarrow$ risk difference, risk ratio, odds ratio, 1-risk ratio, ...
  - Continuous outcome $\rightarrow$ mean difference, median difference, mean difference in AUCs, median ratio, ...
  - Count/rate outcome $\rightarrow$ incidence rate ratio
  - Survival/time-to-event outcome $\rightarrow$ hazard ratio, 1-risk ratio

How many patients do I need?

How many patients do I need?

- Not too few: we want to detect something
- Not too many: ethical reasons and waste of resources

How many patients do I need?

- Not too few: we want to detect something
- Not too many: ethical reasons and waste of resources
- Power calculations: formulas and software

How many patients do I need?

- Not too few: we want to detect something
- Not too many: ethical reasons and waste of resources
- Power calculations: formulas and software
- Needs to be done beforehand and can not be changed
- Relies on choices and assumptions you have to make

What do I need for sample size calculation?

- Research question: clear idea of what we want to estimate and how, including the design of the study
- Choice of endpoint: binary, continuous, time to event,...
- Idea of statistical method you will use: t-test, comparing proportions, regression,...

But I also need some numbers...

- Effect size: effect of the treatment, we need a size of the effect that is clinically relevant to detect
- Variation in the data: estimate of standard deviation in the outcome variable
- Power: how strongly we avoid false negatives, ability to detect a difference if there is one.
- Significance level: how strongly we avoid false positivies, do not detect a difference if there is none.

But I also need some numbers...

- Effect size: effect of the treatment, we need a size of the effect that is clinically relevant to detect

$$\rightarrow d$$

- Variation in the data: estimate of standard deviation in the outcome variable

$$\rightarrow \sigma$$

- Power: how strongly we avoid false negatives, ability to detect a difference if there is one.

$$\rightarrow 1 - \beta$$

- Significance level: how strongly we avoid false positivies, do not detect a difference if there is none.

$$\rightarrow \alpha$$

Formula for sample size calculation

$$n = \frac{(Z_\alpha + Z_\beta)^2 \cdot \sigma^2}{d^2}$$

- Higher effect size fewer patients needed
- Higher variation more patients needed

Where do I find these numbers?

- Good news: power and significance are often set to standard levels
- Power set to 80% or 90%
- Significance level set to 5%

Where do I find these numbers?

- Good news: power and significance are often set to standard levels
- Power set to 80% or 90%
- Significance level set to 5%
- Bad news: effect size and standard deviation are difficult to find

Where do I find these numbers?

- Good news: power and significance are often set to standard levels
- Power set to 80% or 90%
- Significance level set to 5%
- Bad news: effect size and standard deviation are difficult to find
- Some options: literature, pilot studies, ...

When does it get more complicated?

- More than one primary outcome
- More than one look at the results: interim analysis
- More than one arm
- Superiority vs non-inferiority
- More complex designs

Choice of design

Some designs

- Parallel arms
- Crossover trials
- Cluster randomized trials
- Factorial design
- Platform trials
- ...

Crossover design

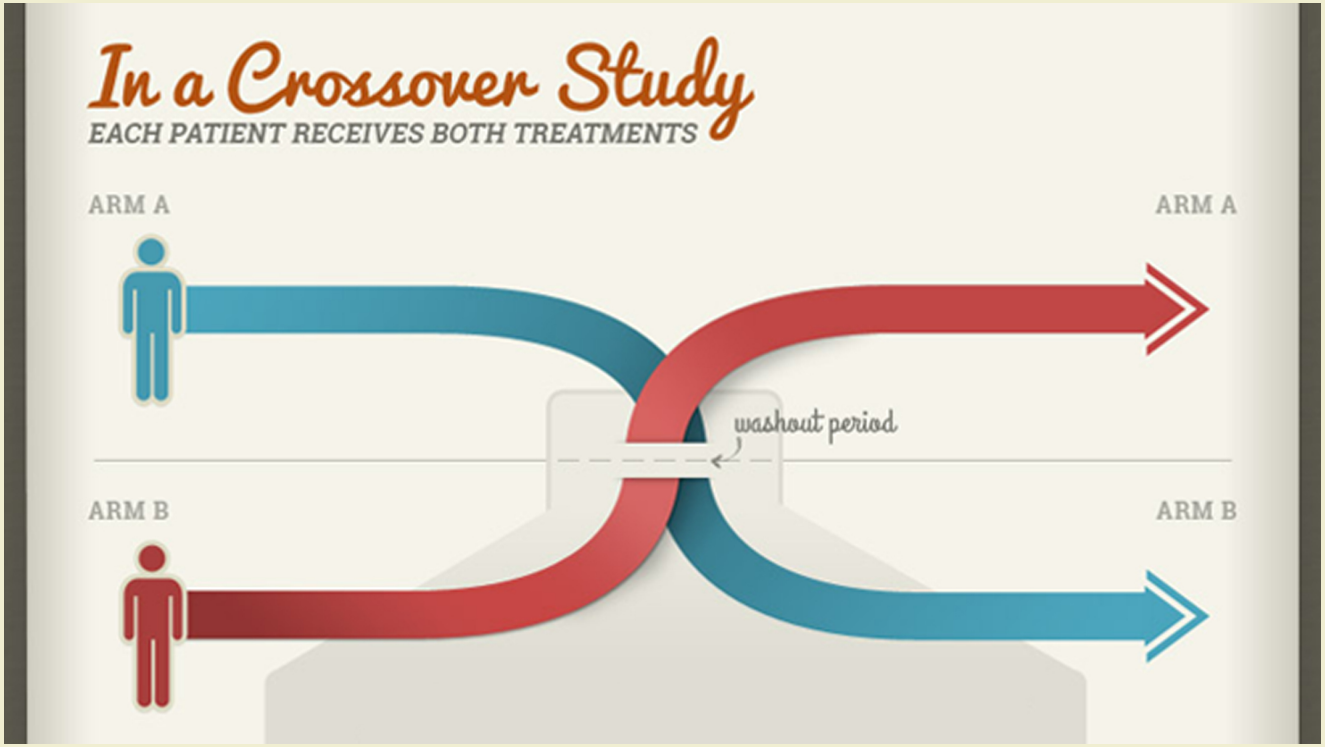


Figure: https://blog.lillytrialguide.com/clinical-trial-design-parallel-crossover-studies/

Crossover design

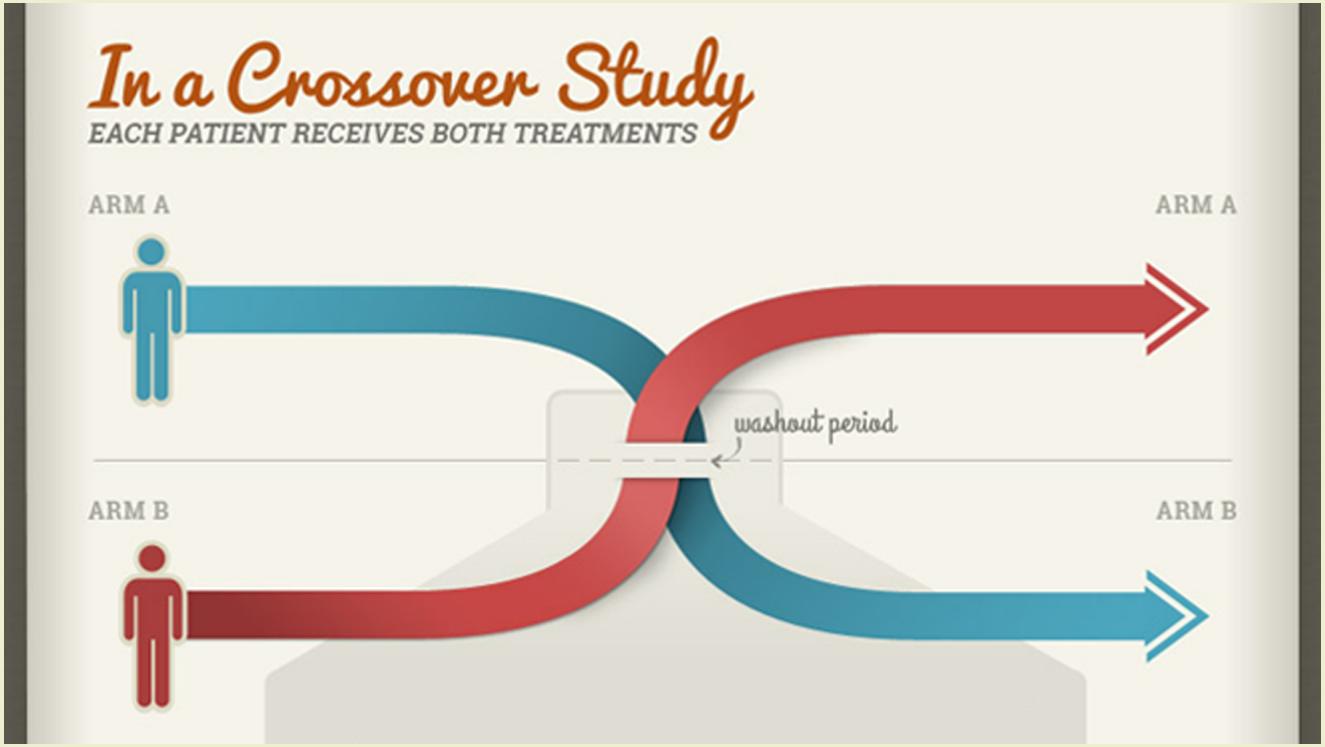


Figure: https://blog.lillytrialguide.com/clinical-trial-design-parallel-crossover-studies/

- Carry-over effects: wash-out period
- Time effect: chronic, stable conditions
- Efficient
- Strong assumptions

Cluster-randomized design

- Randomize groups (clusters) instead of individuals
- Examples: schools, hospitals, regions,...

Cluster-randomized design

- Randomize groups (clusters) instead of individuals
- Examples: schools, hospitals, regions,...
- Needs high sample size
- Heterogeneity vs within cluster interactions
- Pragmatic, quick to recruit
- Good for non- drug interventions

Factorial design

	Control	Treatment 1
Control	Control	Treatment 1
Treatment 2	Treatment 2	Treatment 1 + Treatment 2

Factorial design

	Control	Treatment 1
Control	Control	Treatment 1
Treatment 2	Treatment 2	Treatment 1 + Treatment 2

- One control, two interventions
- Possible to assess interaction
- Efficient if no interaction
- Difficult to interpret if there is an interaction

Summary

- Clear research question
- Estimand (PICO + intercurrent events + summary measure for outcome)
- Type of outcome
- Type of test (superiority vs non-inferiority vs ...)
- Number of patients
- Design: is parallel enough or do we need some other structure?

Summary

- Clear research question
- Estimand (PICO + intercurrent events + summary measure for outcome)
- Type of outcome
- Type of test (superiority vs non-inferiority vs ...)
- Number of patients
- Design: is parallel enough or do we need some other structure?

$\rightarrow$ Come and talk to us!