DECIDE

Introduction to Health Interventions, Policy and Services

Assessment of Efficacy and Effectiveness of Healthcare Interventions – Clinical Trials

Luís Filipe Azevedo

lazevedo@med.up.pt

Summary

- Experimental studies in humans – The randomized clinical trial – RCT
- Phases of clinical trials
- Randomized clinical trials – Types and methodological principles
- Brief notes on the statistical analysis of clinical trials
- Critical appraisal of randomized clinical trials

HEADS

Methods in HTA

- **Primary data methods**
  - Primary data methods involve **collection of original data**, ranging from more scientifically rigorous approaches for determining the causal effect of health technologies, such as randomized controlled trials (RCTs), to less rigorous ones, such as case series. These study designs can be described and categorized based on multiple attributes or dimensions:
    - Experimental studies
    - Quasi-experimental studies
    - Observational studies

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Box III-13. Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard**	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or *n*-of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, *n*-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or *n*-of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non -randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

*** Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.**

**** As always, a systematic review is generally better than an individual study.**

Experimental studies in humans

The Randomized Clinical Trial – RCT

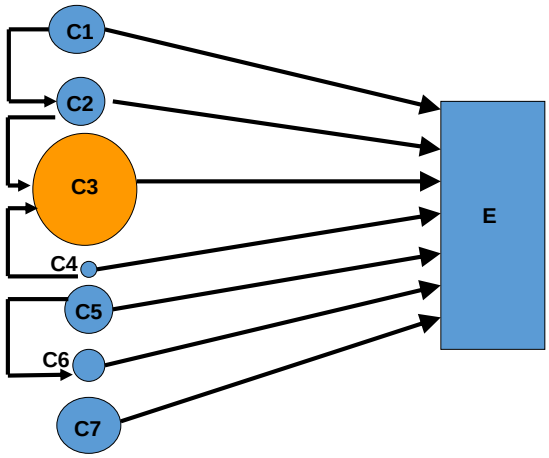
Causality

HEADS

Assessment of causality

- Causes
- Exposure
- Factor
- Independent variable

(E.g.: Risk factors, prognosis factors, therapeutic interventions)



- Effect
- Status
- Result
- Outcome

(E.g.,: Disease, Death, Cure)

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health.Research.

Study design

Question: **What is the effect of a therapeutic intervention?**

- **Study design**
  - ***Experimental study (study with [1] direct control of the intervention and [2] randomisation)***
    - Comparison (Control group)
    - Direct control of the intervention (Experimental or intervention group)
    - Randomisation
    - Control of inespecific effects (Placebo effect, Hawthorne effect, others)
    - Blinding
- **Selection of participants**
  - Inclusion/exclusion criteria (affect the generalisability of the study's conclusions – external validity)
- **Selection of data collection methods**
  - Outcome variables (Clinically relevant variables used for group comparison – e.g., mortality, morbidity, quality of life, etc.)

The Revised CONSORT Statement for Reporting Randomized Trials

Methods		
Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected.
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered.
Objectives	5	Specific objectives and hypotheses.
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors).
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.
Randomization		
Sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification).
Allocation concealment	9	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.
Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated.
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses.

Phases of Clinical Trials

Phases I, II, III and IV

Clinical Trials

- **Phase I Clinical Trials**
  - Small number of participants (~20-80)
  - Asessment of safety, dose safety range, pharmakokynetics,main effects and main adverse reactions
  - Usually “single arm” and in healthy volunteers
- **Phase II Clinical Trials**
  - Higher number of participants (~40-100)
  - Efficacy and safety assessment in specific indications
  - Can be single arm or with randomisation. Usually conducted in patients with specific diseases

Clinical Trials

- **Phase III Clinical Trials**
  - Larger samples (usually >100-200 participants)
  - Asessment of efficacy and efficiency; adverse reactions monitoring
  - Encompass randomisation and comparison with other treatments or placebo. Conducted in patients with specific diseases
- **Phase IV Clinical Trials**
  - Post-commercialisation trials
  - Efficacy and safety monitoring in different populations
  - Less frequently conducted

Clinical Trials

	Phase 0 "Exploratory"	Phase I	Phase II	Phase III	Phase IV
Description	First-in-man early trial to determine if drug engages its expected target	Initial safety evaluations, determine safe dosage range, identify common side effects, study toxicity profile of the drug	Begin to explore efficacy while maintaining safety	Final confirmation of safety and efficacy	Any trials conducted after FDA approval of the drug
Number of subjects	10-15 healthy volunteers	20-80 healthy volunteers	100-300 volunteers with the targeted medical condition	1,000-3,000 subjects with the targeted medical condition	Number of subjects depends on trial endpoints
Dose	Single, low dose (<1% of dose calculated to produce a clinical effect)	• Single dose • Single ascending dose • Multiple ascending dose	Multiple dose trials, often conducted against placebo	Multiple dose trials, ascending doses	Variable
Endpoints	Not expected to show clinical effect or significant adverse effects. Helps to choose between competing chemical analogs for further study.	Escalation of dose ends when unacceptable side effects occur; the previous dose is considered the maximum tolerated dose.	Explores clinical effects against the targeted condition, and reveals the less-common side effects	Confirms clinical efficacy of the drug against the targeted condition and evaluates safety and side effects	Confirms clinical efficacy and safety and explores other possible drug uses; may be required as a condition of drug approval
Timing	Can be conducted with prior approval while final IND review is pending	Together with Phase 0 trials, first clinical trials conducted in an IND process	Conducted after report to FDA of results of Phase I trials	Conducted after report to FDA of results of Phase II trials	Conducted after release of the drug by the FDA for marketing

Randomized Clinical Trials – Types and Methodological Principles

Experimental studies

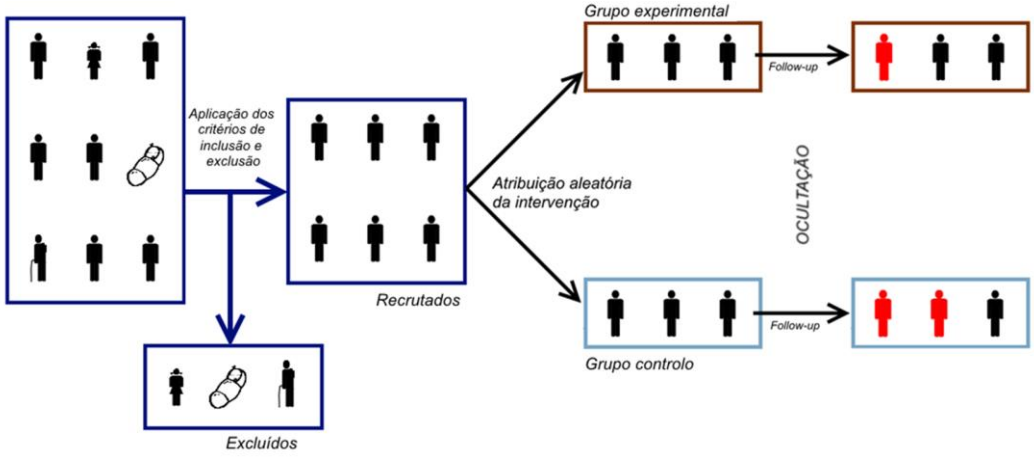
HEADS
PhD PROGRAMME IN HEALTH DATA SCIENCE

- **Basic concepts:**
  - Randomization
  - Stratification (blocking)
- **Types of randomization:**
  - Complete randomized design
  - Random permuted blocks design
  - Randomized cross-over design
  - Randomized stratified (blocked) design
  - Factorial designs
  - Randomized clusters design

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health.Research.

HEADS

Study design

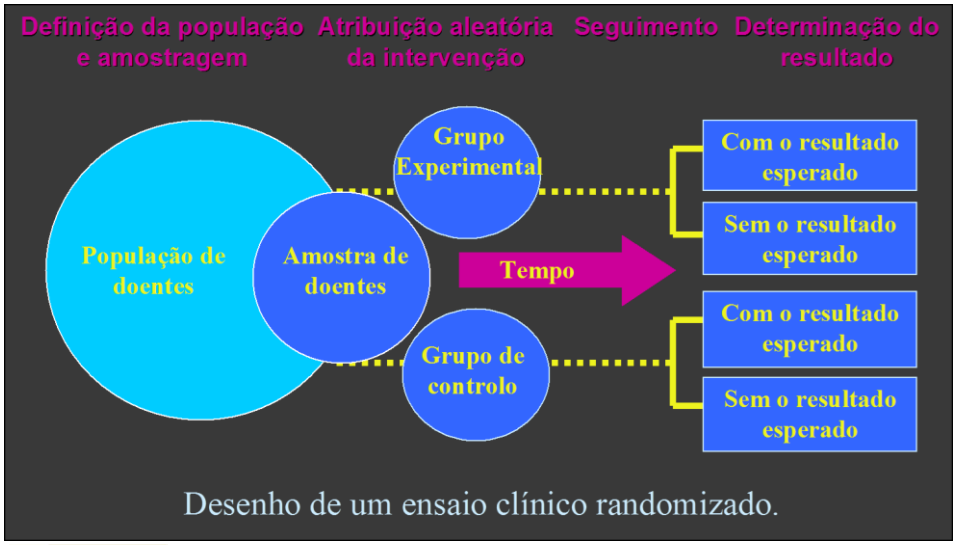


MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

HEADS

Study design

Parallel design



Desenho de um ensaio clínico randomizado.

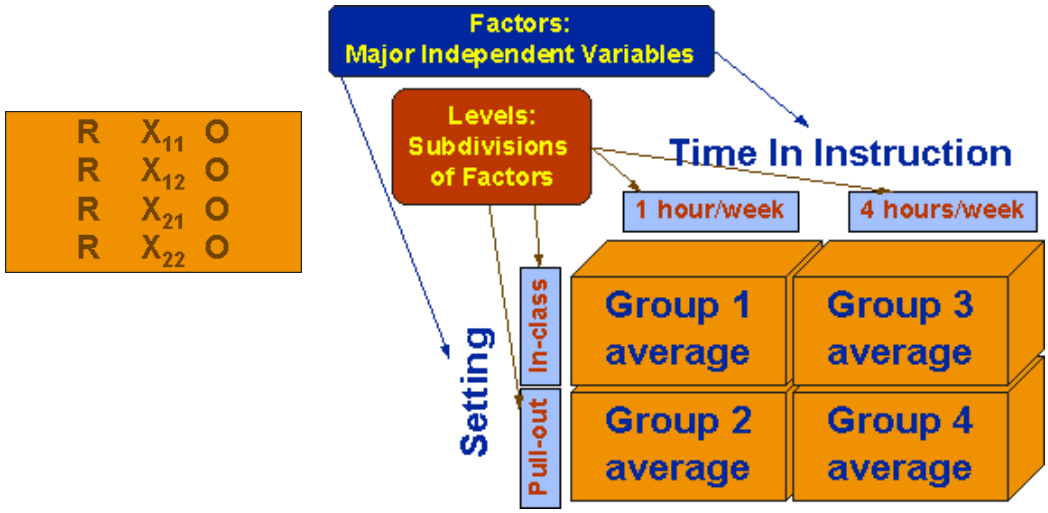
MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Study design

Randomized factorial design

HEADS

R X_{11} O
R X_{12} O
R X_{21} O
R X_{22} O



MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Adapted from: Trochim, William M. The Research Methods Knowledge Base, 2nd Edition. Internet WWW page, at URL: http://www.socialresearchmethods.net/kb/ (version current as of October 20, 2006).

Effects of a polypill (Polycap) on risk factors in middle-aged individuals without cardiovascular disease (TIPS): a phase II, double-blind, randomised trial

The Lancet, Volume 373, Issue 9672, 2009, 1341 – 1351. http://dx.doi.org/10.1016/S0140-6736(09)60611-5

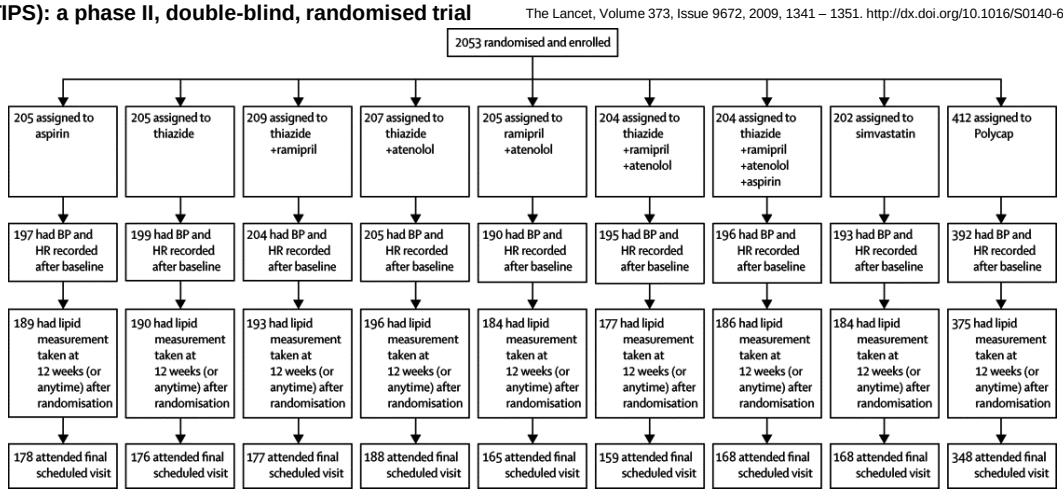
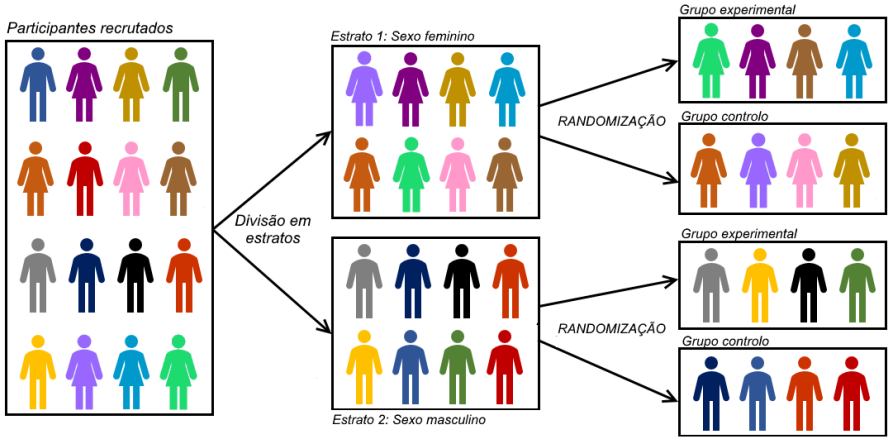


Figure 1 Trial profile The number of people screened for eligibility was not recorded. BP=blood pressure. HR=heart rate.

Since we were testing the effects of five active pharmacological components (three agents to lower blood pressure, statin, and aspirin: Polycap [Cadila Pharmaceuticals, Ahmedabad, India]), a full factorial design would require 32 cells. Such a design was not practical. Therefore, we identified five questions (see above) that were most relevant and could be addressed by randomly assigning individuals to one of nine groups (figure 1,table 1).

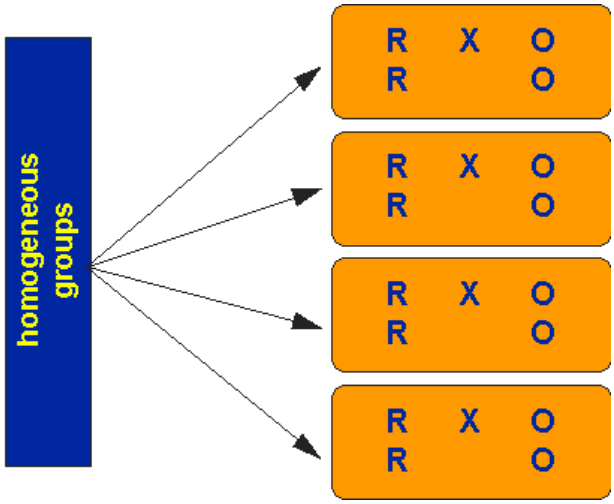
Study design

Experimental studies with stratification



Study design

Experimental studies with stratification

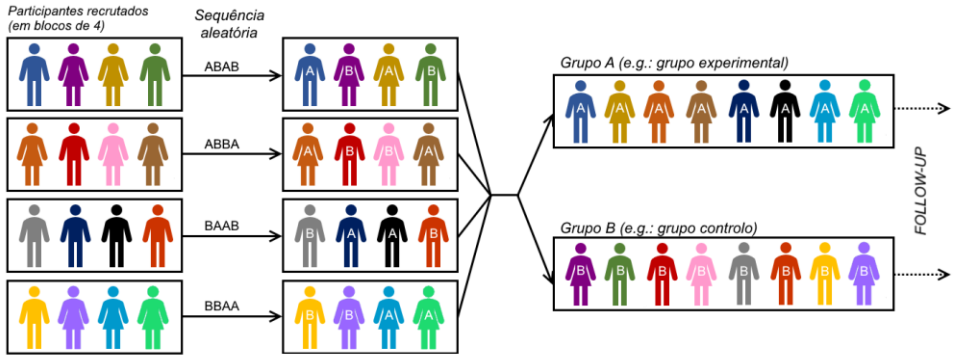


Adapted from: Trochim, William M. The Research Methods Knowledge Base, 2nd Edition. Internet WWW page, at URL: http://www.socialresearchmethods.net/kb/ (version current as of October 20, 2006).

Study design

HEADS

Random permuted blocks design



MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Study design

HEADS

Random permuted blocks design

Randomized Permuted Blocks

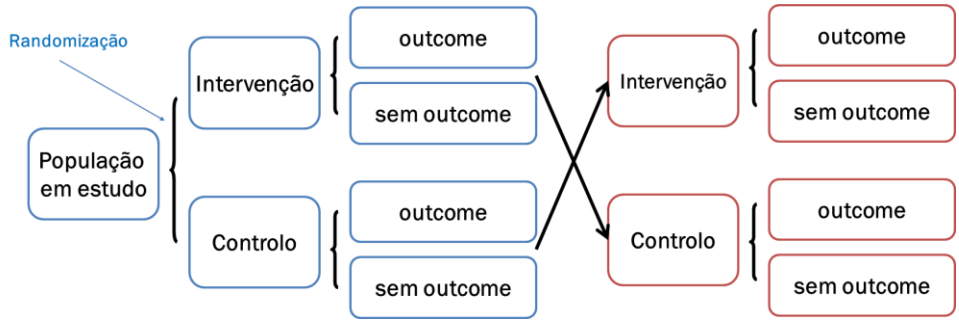
ABAB	ABBA	AABB
BBAA	BAAB	BABA

16-05-2014 59

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

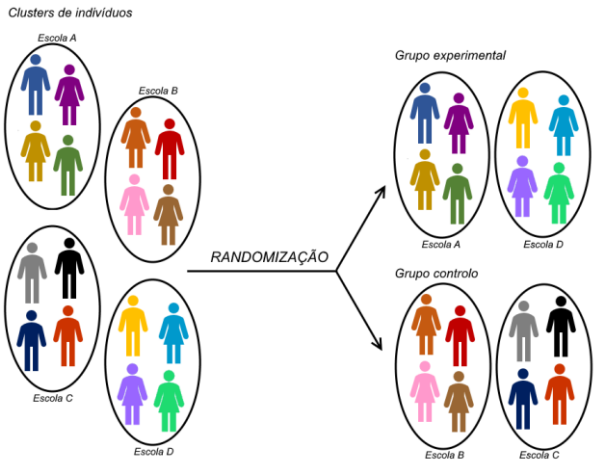
Study design

Randomized cross-over design



Study design

Randomized clusters design



Study design

Randomized clusters design

Lancet. 1986 May 24;1(8491):1169-73.

Impact of vitamin A supplementation on childhood mortality. A randomised controlled community trial.

450 villages in northern Sumatra were randomly assigned to either participate in a vitamin A supplementation scheme (n = 229) or serve for 1 year as a control (n = 221). 25 939 preschool children were examined at baseline and again 11 to 13 months later.
(…)
Among children aged 12-71 months at baseline, mortality in control villages (75/10 231, 7.3 per 1000) was 49% greater than in those where supplements were given (53/10 919, 4.9 per 1000) (p less than 0.05).

Pragmatic experimental studies in a real-life context

- Externalities included in the total set of determinants of a treatment response
- Results assessed according to what is relevant for decision-making in daily clinical practice
- Participants corresponding to those patients attending the clinical practice (and not to highly-selected patients)

Pragmatic experimental studies in a real-life context

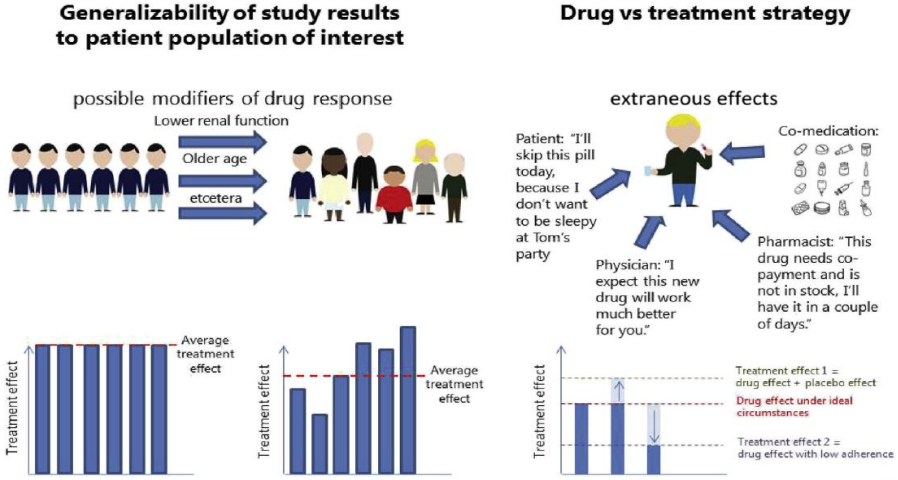
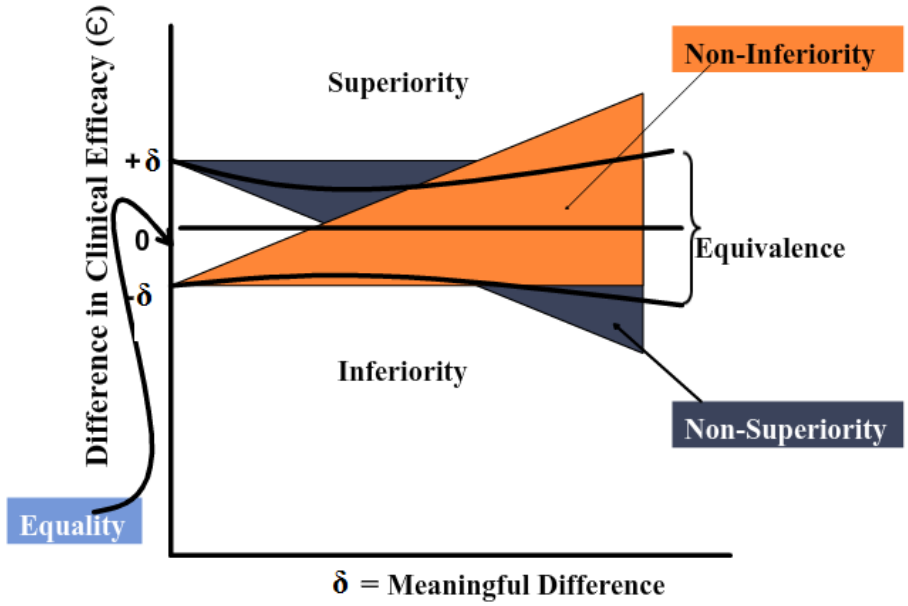


Fig. 1. Pragmatic trial design.

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Outcome assessment and statistical analysis

Other Hypothesis and comparisons



Other Hypothesis and comparisons

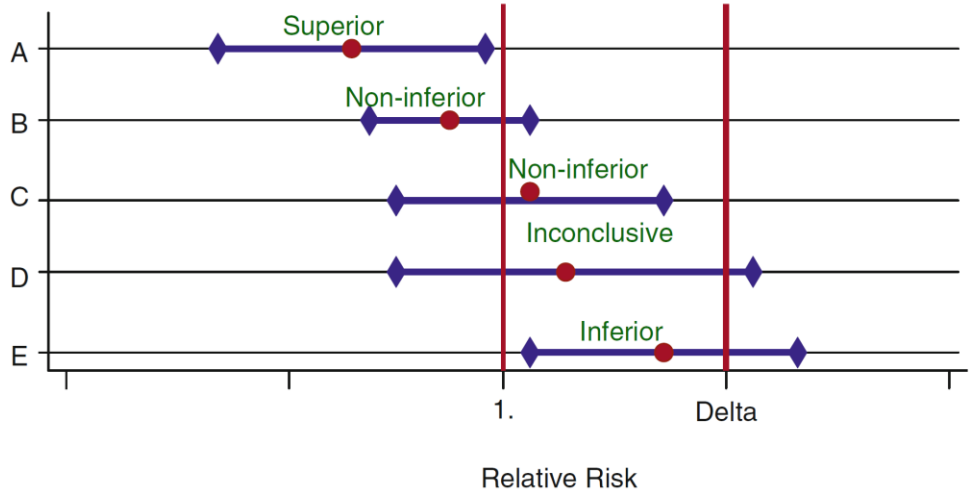
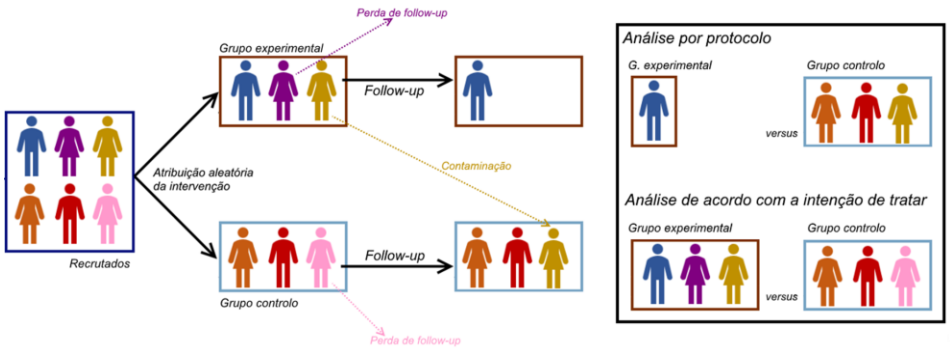


Fig. 18.11 Relative risks and 95% confidence intervals for a series of superiority and non-inferiority trials [181]

Statistical analysis
Intention-to-treat analysis



Statistical analysis

Análise de acordo c/intenção de tratar	Análise por protocolo
Objectivo: identificar o melhor tratamento para a prática clínica • Participantes são analisados de acordo com os resultados da randomização, mesmo que não tenham cumprido com a opção para a qual estavam alocados • Análise que avalia a efectividade da intervenção (dá resposta à questão pragmática) • Deve ser a principal análise reportada num estudo experimental	**Objectivo:** compreender a eficácia da intervenção • Participantes são analisados de acordo com o tratamento/opção que efectivamente completaram • Análise que avalia a eficácia da intervenção (dá resposta à questão explanatória) • Não tem em conta perdas de *follow-up*, contaminação...

Statistical analysis

- **Subgroup analysis** – A difference in treatment efficacy in specific subgroups is only acceptable in the presence of affirmative answers to **all** of the following questions:
  - Does it make sense from the biological and clinical point of view?
  - Is the observed difference clinically and statistically significant?
  - Was the hypothesis described before the beginning of the study? (instead of the results being found by pure chance)
  - Was it simply one of the results of the few assessed subgroups?
  - Are there independent studies confirming the results?

Statistical analysis

- Interim analysis and stopping rules
- Adaptive designs
  - The term "*adaptive design*" is used in many different ways.
  - Used to describe any multi-stage trial where later stages are based, in part, on what happened in earlier stages

HEADS

Statistical analysis

- Adaptive designs

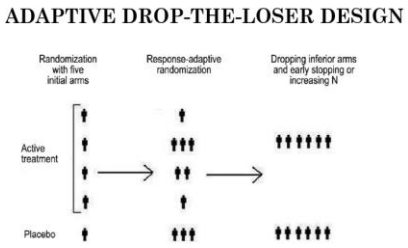


Figure 14. Patient Allocation Adaptive Design

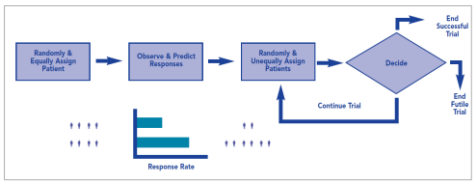
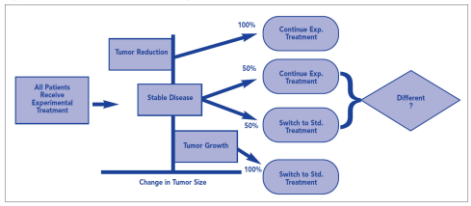


Figure 16. Randomized Discontinuation Design



MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Critical Appraisal of Randomized Clinical Trials

Critical appraisal of clinical trials

- Three basic questions:
  - **Are the results of this study valid**?
  - **Are the (valid) results of this study relevant**?
  - **Are the (valid and relevant) results of this study applicable to my patient**?

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Was there an adequate randomisation?
    - Complete randomized design
    - Random permuted blocks design
    - Randomized cross-over design
    - Randomized stratified (blocked) design
    - Factorial designs
    - Randomized clusters design

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Was the randomisation sequence concealed from the professionals responsible for assigning the interventions?
    - *"Allocation concealment"*
    - *"Concealed randomization list"*
    - Use of sealed, opaque and inviolable envelopes
    - Randomisation center for assigning the intervention by phone contact

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Were the experimental and control groups comparable at baseline regarding other variables that may, eventually, influence the study's results?
    - Baseline characteristics, measurements or variables!
    - Careful analysis of "table 1"

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Were all patients assessed in the groups to which they were initially randomised?
    - Does the analysis respect the initial assignment of interventions?
    - "*Once randomized should be analized!*"
    - *"Intention-to-treat analysis"*

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Were all elements involved in the study blinded or masked on the treatment group?
    - Patients, physicians, researchers responsible for analysing outcome variables, researchers responsible for data analysis (single blind, double blind, triple blind)

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Was the follow-up period sufficiently long and complete?
    - Ideally, the follow-up should be complete
    - Rule of the thumb: at least 80% of follow-up

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Are the outcome variables adequate?
    - Definition of primary and secondary outcomes (end-points)
    - Problems associated with surrogate outcomes (end-points)
    - Problems associated with combined outcomes (end-points)

Speculation on Reasons for Failures of Surrogate End Points*

Disease and Intervention	End Points		Settings in Figure 1†			
	Surrogate	Clinical	A	B	C	D
Cardiologic disorder						
Arrhythmia						
Encainide; flecainide	Ventricular arrhythmias	Survival		+		++
Quinidine; lidocaine	Atrial fibrillation	Survival		+		++
Congestive heart failure						
Milrinone; flosequinan	Cardiac output; ejection fraction	Survival		+		++
Elevated lipid levels						
Fibrates; hormones; diet; lovastatin	Cholesterol levels	Survival		+		++
Elevated blood pressure						
Calcium channel blockers	Blood pressure	Myocardial infarction, survival		+		++
Cancer						
Prevention						
Finasteride	Prostate biopsy	Symptoms; survival	++‡			
Advanced disease						
Fluorouracil plus leucovorin	Tumor shrinkage	Survival		+		++
Other diseases						
HIV infection or AIDS						
Antiretroviral agents	CD4 levels; viral load	AIDS events; survival		+	+	+
Osteoporosis						
Sodium fluoride	Bone mineral density	Bone fractures	+			+
Chronic granulomatous disease						
Interferon-γ	Bacterial killing; superoxide production	Serious infection			++	

* AIDS = acquired immunodeficiency syndrome; HIV = human immunodeficiency virus; + = likely or plausible; ++ = very likely.
† A = surrogate end point not in causal pathway of the disease process; B = of several causal pathways of the disease, the intervention only affects the pathway mediated through the surrogate; C = the surrogate is not in the pathway of the intervention's effect or is insensitive to its effect; D = the intervention has mechanisms of action that are independent of the disease process.
‡ In settings in which only latent disease is prevented.

Fleming, T. R. et. al. Ann Intern Med 1996;125:605-613

Annals of Internal Medicine

Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)

Lancet 1998; **352:** 837–53

The primary endpoint of the trial[w13] was time to first "diabetes-related endpoints" (sudden death, death from hyperglycaemia or hypoglycaemia, fatal or non-fatal myocardial infarction, angina, heart failure, stroke, renal failure, amputation [of at least one digit], vitreous haemorrhage, retinal photocoagulation (endpoint added after the trial onset), blindness in one eye, or cataract extraction); "diabetes-related death" (death from myocardial infarction, stroke, peripheral vascular disease, renal disease, hyperglycaemia or hypoglycaemia, and sudden death); or all-cause mortality.

Critical appraisal of clinical trials

- Are the results of this study valid?
  - Were the comparators adequately chosen?
    - No treatment or placebo
    - Alternative active interventions

Problems in the Design and Reporting of Trials of Antifungal Agents Encountered During Meta-analysis

JAMA. 1999;282:1752-1759

B in patients with cancer complicated by neutropenia. In 3 large trials that comprised 43% of the patients identified for the meta-analysis, results for amphotericin B were combined with results for nystatin in a "polyene" group. Because nystatin is recognized as an ineffective drug in these circumstances, this approach creates a bias in favor of fluconazole. Furthermore, 79% of the patients were randomized to receive oral amphotericin B, which is poorly absorbed and not an established treatment, in contrast to intravenous amphotericin B, which was administered in 4 of 5 placebo-controlled trials, or 86% of patients. It was unclear whether there was overlap among

HEADS

Critical appraisal of clinical trials

- Are the results of this study relevant?
  - What is the **effect size** for the treatment efficacy?
  - What is the **precision of the estimates** for the treatment efficacy?
  - Is the results' presentation adequate?

MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health. Research.

Association or effect measures

- An **association or effect measure quantitatively** assesses the **strength of the (causal) association** between an **intervention or factor** and a certain **outcome**
- In the annex, we provide a list of the effect measures most frequently described in the literature (as well as a brief explanation on confidence intervals)
- There are two types of effect/association measures:
  - **Absolute measures – ARR, NNT**
  - **Relative measures – RR, RRR, OR**

Association or effect measures

Vejamos três interessantes exemplos de ECAs:

REC	REE	RR	OR	RRR	ARR	NNT
50%	25%	0,5	0,333	50%	25%	4
5%	2,50%	0,5	0,487	50%	2,50%	40
0,05%	0,025%	0,5	0,499	50%	0,03%	4000

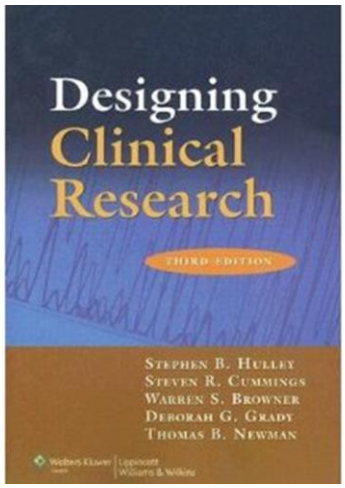
HEADS

Critical appraisal of clinical trials

- Are the valid and relevant results of this study applicable to my patient?
  - Is my patient very different from the study participants?
  - In my setting, is the treatment available and is it applicable in the clinical practice?
  - What are the potential advantages and disadvantages of the treatment in this specific patient?
  - What are the opinions, values and expectations of our patient in relation to the outcome we are trying to avoid and to the treatment we are proposing?

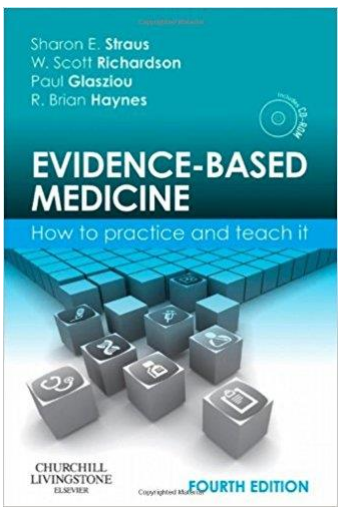
MEDCIDS DEPARTAMENTO DE MEDICINA DA COMUNIDADE, INFORMAÇÃO E DECISÃO EM SAÚDE U.PORTO CINTESIS Health.Research.

Bibliography



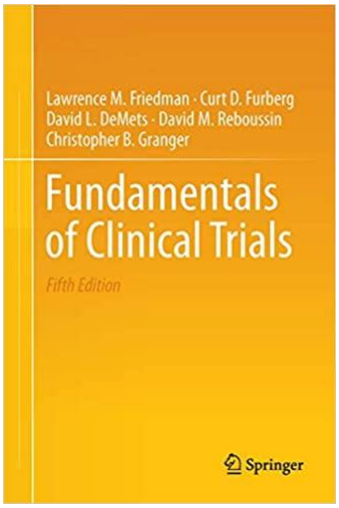
Designing Clinical Research, 4th Ed.
Hulley SB, Cummings SR et al.
Lippincott Williams & Wilkins, 2013

Bibliography



Evidence-Based Medicine: How to Practice and Teach It, 4th Edition by Sharon E. Straus MD, Paul Glasziou MRCGP FRACGP PhD, W. Scott Richardson MD, R. Brian Haynes MD

Bibliography



Fundamentals of clinical trials, 5th Edition by Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., Granger, C. B. Springer (2015).

Questions?

Annex

Measures of effect and association

Association or effect measures

- An **association or effect measure quantitatively** assesses the **strength of the (causal) association** between an **intervention or factor** and a certain **outcome**
- In the annex, we provide a list of the effect measures most frequently described in the literature (as well as a brief explanation on confidence intervals)
- There are two types of effect/association measures:
  - **Absolute measures** – **ARR, NNT**
  - **Relative measures** – **RR, RRR, OR**

Association or effect measures

- As an example, consider a randomised controlled trial (RCT) in patients admitted in an ICU and in which the hospital mortality was assessed in the experimental and in the control groups.

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

Association or effect measures

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Risk of the event in the control group (REC)** – estimate of the probability of the event in the control group (REC= 20/100 = 0,2 = 20%).
- **Risk of the event in the experimental group (REE)** – estimate of the probability of the event in the experimental group (REE = 10/100 = 0,1 = 10%).

Association or effect measures

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Risk ratio (Relative Risk; RR)** – Relative effect size corresponding to the risk in the experimental group dividing by the risk in the control group (**RR = REE/REC = 0,1/0,2 = 0,5**), informing on the relative frequency of the event in the experimental group compared to the control group.

Association or effect measures

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Odds Ratio (OR)** – Relative effect measure, corresponding to the event odds in the experimental group dividing by the event odds in the control group (**OR = (10/90) / (20/80) = 0,11/0,25 = 0,44**), informing on the relative frequency of the event (expressed by odds) in the experimental group compared to the control group.

Association or effect measures

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Absolute risk reduction (ARR)** – Absolute effect measure, estimated by the absolute difference between the risk in the control group and the risk in the experimental group (**RAR = REC-REE = 0,2-0,1 = 0,1 = 10%**). This measure is interpreted as the absolute reduction of the event risk associated with the intervention.

Association or effect measures

Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Relative risk reduction (RRR) –** Measure of potential impact (it is not strictly an effect measure), estimated by dividing the absolute risk reduction by the risk of the event in the control group (**RRR = ARR/REC = 0,1/0,2 = 0,5 = 50%**). The RRR estimates the impact of the intervention in the population and can be interpreted as the risk proportion specifically decreased by the intervention.

Association or effect measures

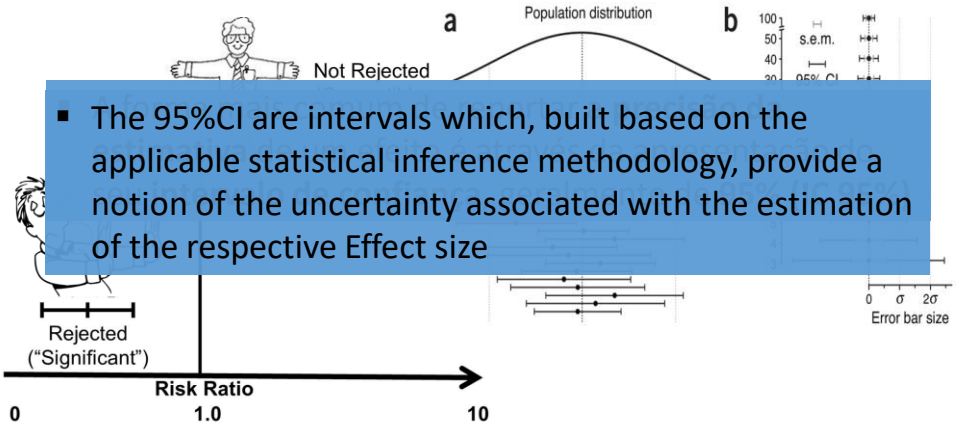
Result/ Group	Hospital deaths	Hospital survivals	Total
Experimental group	10	90	100
Control group	20	80	100
Total	30	170	200

- **Number needed to treat –** Absolute measure, estimated by the inverse of the absolute risk reduction (**NNT = 1/ARR = 1/0,1 = 10**). This measure estimates the number of individuals that would be needed to treat with the intervention in order to prevent a single event (death). In this case, we would need to treat 10 individuals with the intervention to prevent an additional deaths.

Precision of the effect measures

- The **effect measures** provided in a study are always **estimates of the true value of that parameter in the population**
- Naturally, **effect estimates** resulting from **samples of 1000 participants** will be **more precise (less sensitive to random errors related to the sampling process)** than estimates resulting from **samples with 50 participants**
- The **sample size** is an important determinant of the precision. However, the **frequency of the event in the groups**, the **variance of the outcome variable** and the adopted **confidence level** are also relevant factors

Precision of the effect measures



- The 95%CI are intervals which, built based on the applicable statistical inference methodology, provide a notion of the uncertainty associated with the estimation of the respective Effect size

Annex

Statistical analysis of clinical trials

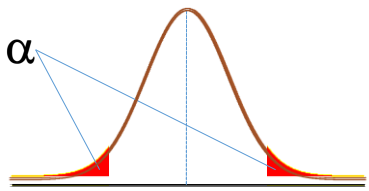
Brief notes on the statistical analysis of clinical trials

- **Statistical Hypothesis Testing:**
  - Hypothesis to be tested – Null hypothesis (H_0)
  - Alternative hypothesis (H_1)
  - Test statistics
  - Probability computed using the test statistics – p-value
  - Decision
  - Type I and type II errors

Statistical Significance (p-value)

- Based on the p-value we can then decide to Accept or Reject H_0
- If the **p-value is small**, we **reject H_0** and say that the result is **statistically significant** (the sample has little compatibility with H_0)
- We should **define beforehand what is the cut-off for "small p-value"**
- **This cut-off is designated as the level of significance (α)**
- Although this choice is arbitrary, in medical research 0.05 is often used
- So, **for α = 0.05, if p<0.05 then we reject H_0**

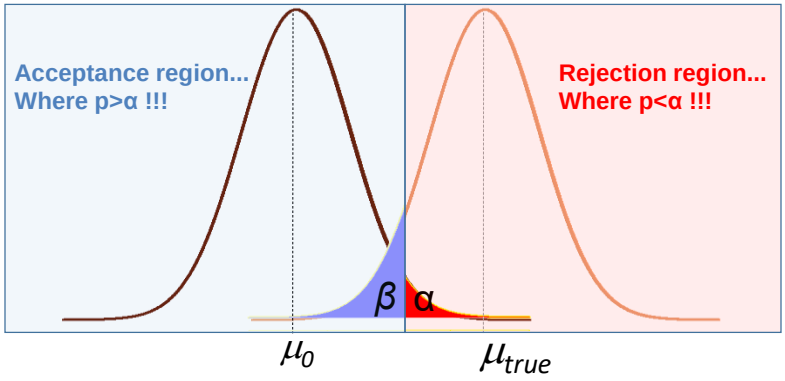
H_0: $\mu_1=\mu_2$
or
H_0: $\mu_1-\mu_2=0$



Errors

$P(\text{Reject } H_0 \mid H_0 \text{ true}) = \alpha$ Type I error

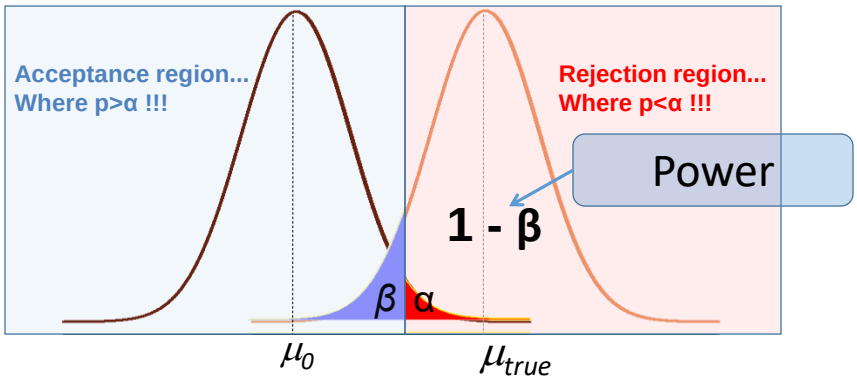
$P(\text{Accept } H_0 \mid H_0 \text{ false}) = \beta$ Type II error



Power

$P(\text{Accept } H_0 \mid H_0 \text{ false}) = \beta$ Type II error

$P(\text{Reject } H_0 \mid H_0 \text{ false})$ Power



POWER is the test ability to detect a difference when in fact it exists