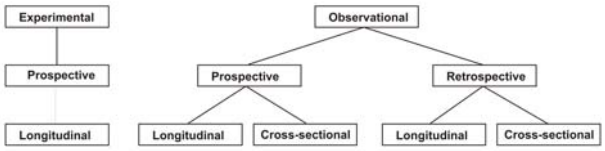
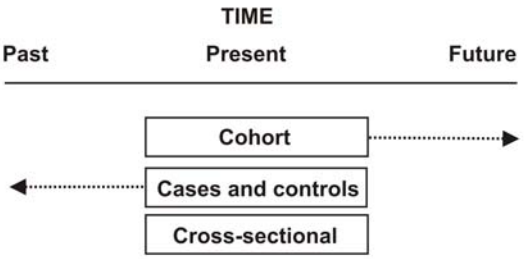
Outline

Types of study design

1. Retrospective studies (of past events) including case-control studies
2. Prospective studies (of past events)
3. Cohort studies or epidemiological design (of ongoing or future events)
4. Clinical trials



Basic structure for different designs



Types of studies

Therapy study
Effectivity of a drug, new surgery or alternative methods
Design: RCT

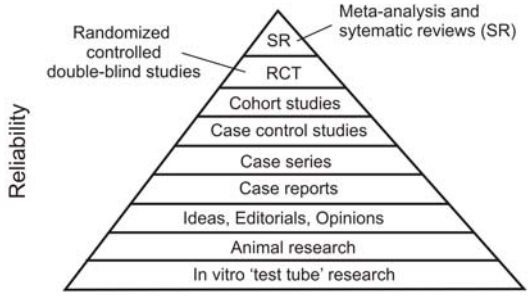
Diagnosis study
Validity and reliability of new diagnostic tests
Design: Cross-sectional

Screening study
Investigation of test results
Design: Cross-sectional

Prognosis study
Progress of an early diagnosed disease
Design: Cohort

Causal study
Association between dangerous substances and a disease
Design: Cohort, Case control

Hierarchy of medical studies



Clinical trials

Clinical studies form a class of all scientific approaches to evaluating medical disease prevention, diagnostic techniques, and treatments. Among this class trials, often called clinical trials, form a subset of those clinical studies that evaluate investigational drugs.

- Phase I trials focus on safety of a new investigational medicine. These are the first human trials after successful animal trials.
- Phase II trials are small trials to evaluate efficacy and focus more on a safety profile.
- Phase III trials are well-controlled trials, the most rigorous demonstration of a drug's efficacy prior to federal regulatory approval.
- Phase IV trials are often conducted after a medicine is marketed to provide additional details about the medicine's efficacy and a mor complete safety profile.

Clinical trials

The goal in a phase I trial is to identify a maximum tolerated dose (MTD), a dose that has reasonable efficacy (i.e. is toxic enough to kill cancer cells) but with tolerable toxicity (i.e. not toxic enough to kill the patient).

Phase I trials are applied to patients from standard treatment failure who are at high risk of death in the short term.

In phase II trials the optimal dose (MTD) is applied to a small group of patients meeting a predefined inclusion criteria (there are also exclusion criteria) and the response rate, the proportion or percentage of patients who respond, is studied.

A second type of phase II trials consist of small comparative trials where we want to establish the efficacy of a new drug against a control or standard regimen.

Clinical trials

Phase III/IV are larger studies and the standard is a randomized double-blind controlled trial ("golden standard").

Controlled: The drug is tested against a control group receiving a placebo or the standard treatment. The size, shape, procedure should be very similar to control psychological and emotional effects.

Randomized: If a patient gets the drug or the placebo is assigned randomly.

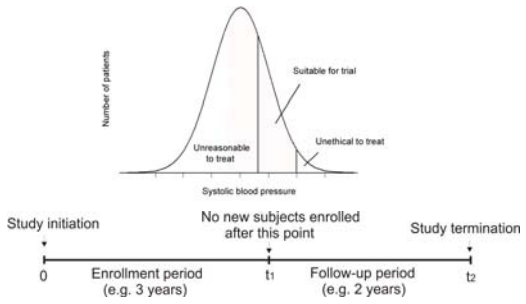
Stratified randomization: If there are expected cofounding variables (e.g. age) patients are stratified and treatment randomly assigned within stratum.

Minimization: A non-random treatment allocation for smaller trials. The allocation is based on the balance of several parameter, so that the $n+1$ treatment is assigned based on the sum of the numbers within the stratified variables (e.g. age$\leq$50 or age$>$50).

Clinical trials

Double-blind: Blind to the patient and blind to the investigator (Triple-blind means that also regulatory officers/statisticians are "blinded").

Selection of subjects: Based on inclusion/exclusion criteria



Alternative designs

- Crossover design
- Within group (paired) comparisons
- Sequential design
- Factorial design
- Adaptive design
- Zelen's design

Sample size

Sample size for Phase II trials and surveys:

$$n = \frac{(z_{1-\alpha})^2 p(1-p)}{d^2} \text{ (response rate)}$$

Sample size for other Phase II trials:

$$n = \frac{(z_{1-\alpha})^2 s^2}{d^2} \text{ (continuous endpoint)}$$

$$n = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{(\frac{1}{2} ln \frac{1+r}{1-r})^2} + 3 \text{ (correlation endpoint)}$$

Sample size

Phase II designs for selection:

$$N = \frac{4(z_{1-\alpha})^2 s^2}{d^2} \text{ (continuous endpoint)}$$

$$N = \frac{4(z_{1-\alpha})^2 \bar{p}(1-\bar{p})}{(p_2 - p_1)^2} \text{ (binary endpoint)}$$

Phase III trials:

$$N = \frac{4(z_{1-\alpha} + z_{1-\beta})^2 \sigma^2}{d^2} \text{ (comparison of 2 means)}$$

$$N = \frac{4(z_{1-\alpha} + z_{1-\beta})^2 \bar{p}(1-\bar{p})}{(p_2 - p_1)^2} \text{ (comparison of 2 proportions)}$$

Number-needed-to-treat

	Outcome		
	+	-	total
Treatment	a	b	a+b
Placebo	c	d	c+d
total	a+c	b+d	N

Experimental event rate: $EER = \frac{a}{a+b}$

Control event rate: $CER = \frac{c}{c+d}$

Relative risk: $RR = \frac{EER}{CER}$

Relative risk reduction: $RRR = \frac{EER - CER}{CER}$

Absolute risk reduction: $ARR = EER - CER$

Number-needed-to-treat: $NNT = \frac{1}{ARR} = \frac{1}{EER - CER}$

Study protocol

- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guidelines for Good Clinical Practise (GCP).
- Formal document outlining the proposed procedures (basically contain any information from patient selection criteria to responsibilities)
- For protocol violations (e.g. patients didn't take their treatments) the only safe way is to keep those in the analysis as intended (intention-to-treat).

Safety

Sponsor

Informing the local site investigators of the true historical safety record of the drug, Monitoring the results of the study (Data Monitoring Commitee (DMC) also known as Data Safety Monitoring Board), Collecting adverse event reports, Write site-specific informed consent

Local site investigator

Conducting the study according to the study protocol, Give truly informed consent (risks, potential benefits)

Institutional review board (IRB) or Ethics Committee

Scrutinize the study for both medical safety and protection to the patients

Regulatory agencies (FDA, EAEM)

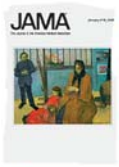
Review all study data before allowing the drug to proceed to the next phase, Audits for the local site investigator

Outline

Medical journals and sites

PubMed www.pubmed.gov

ClinicalTrials.gov



The NEW ENGLAND JOURNAL of MEDICINE

THE LANCET

Volume 371 · Number 9608 · January 19-25, 2008

PLoS MEDICINE

The Cochrane Library

National Institutes of Health
The Nation's Medical Research Agency

How to choose a statistical test

	Type of Data			
Goal	Measurement (from Gaussian Population)	Rank, Score, or Measurement (from Non-Gaussian Population)	Binomial (Two Possible Outcomes)	Survival Time
Describe one group	Mean, SD	Median, interquartile range	Proportion	Kaplan Meier survival curve
Compare one group to a hypothetical value	One-sample t test	Wilcoxon test	Chi-square or Binomial test	
Compare two unpaired groups	Unpaired t test	Mann-Whitney test	Fisher's test (chi-square for large samples)	Log-rank test or Mantel-Haenszel
Compare two paired groups	Paired t test	Wilcoxon test	McNemar's test	Conditional proportional hazards regression
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test	Chi-square test	Cox proportional hazard regression
Compare three or more matched groups	Repeated-measures ANOVA	Friedman test	Cochrane Q	Conditional proportional hazards regression
Quantify association between two variables	Pearson correlation	Spearman correlation	Contingency coefficients	
Predict value from another measured variable	Simple linear regression or Nonlinear regression	Nonparametric regression	Simple logistic regression	Cox proportional hazard regression
Predict value from several measured or binomial variables	Multiple linear regression or Multiple nonlinear regression		Multiple logistic regression	Cox proportional hazard regression

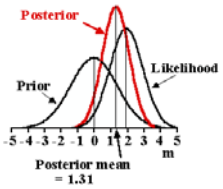
Bayesians vs. Frequentist

Frequentist

The population value is seen as fixed (but unknown) and calculate confidence interval and hypothesis tests. The entire information comes from the data.

Bayesians

The population mean follows a distribution (prior probability). Data can be used to modify the prior probability distribution and gives the posterior probability distribution. Here a 95% credible interval can be constructed, which is narrower than the confidence interval. Difficulties can arise by deciding the prior distribution (prior) and some bayesian methods may lead to intractable computational problems.



Dos and Don'ts

- Don't carry out a significance test, get a large P value, and then interpret this as meaning that there is no difference.
- A confidence interval for the mean difference would be much better than significance tests. A non-significant difference in 10 subjects cannot be interpreted.
- Quote your p values correctly to one significant figure (e.g. $p = 0.007$) (do not use $p < 0.013, p < 0.01, p > 0.05, p = NS$)
- Significant should not be used if you mean important.
- Don't do direct comparison of p-values. It is not correct to compare two groups by testing changes in each one separately. Significance does not depend only on magnitude, but on variability and sample size. A two sample t method should be used to compare the log ratios in the two groups.

Dos and Don'ts

- Always state if you are using SD, SE or CI. Avoid $\pm$
- Do confidence intervals (or SE's) on group means, rather than on comparisons.
- Don't use three-dimensional effects.
- The tests of significance at baseline should not be done. If the subjects are randomized, they come from the same population and the null hypothesis is true. There is no reason to test it.
- Don't analyze the data as if they are all from the same population and ignoring the fact that these 21 groups of subjects are from 9 different trials.
- Don't do Chi-square test analyzes of ordered categorical data.

Guidelines

1. Read the journal's instructions to authors. If they do not cover statistics, use those of one of the major general medical journals.

2. Never, ever, conclude that there is no difference or relationship because it is not significant.

3. Give confidence intervals where you can.

4. Give exact P values where possible, not $P < 0.05$ or $P = NS$, though only one significant figure is necessary.

5. Be clear what your main hypothesis and outcome variable are. Avoid multiple testing.

Guidelines

6. Get the design right, be clear about blinding and randomization, do a sample size calculation if you can.

7. Be clear whether you are quoting standard deviations or standard errors, avoid $\pm$ notation.

8. Avoid bar charts with error bars.

9. Check the assumptions of your statistical methods.

10. Give clear descriptions of your statistical methods.

11. Decide for which baseline characteristics you should adjust in advance, then do it.