Methodological Considerations in Clinical Trials

Ferran Torres
Hospital Clínic Barcelona / Universitat Autònoma Barcelona
EMA: Scientific Advice Working Party (SAWP)
Biostatistics Working Party (BSWP)

LACK OF HARMONISATION
INTERNATIONAL CONFERENCES HARMONISATION
www.ich.org

Data to register in all regions

ICH.org

- E 3 Structure and Content of Clinical Study Reports
- E 6 (R1) Guideline for Good Clinical Practice
- E 8 General Considerations for Clinical Trials
- E 9 Statistical Principles for Clinical Trials
- E 10 Choice of Control Group in Clinical Trials

Globalisation

Regulatory Agencies

07/10/12
Some fundamental books

- Machin D, Davies S. Handbook of Clinical Trials - Phase II.

Phases of Clinical Trials

<table>
<thead>
<tr>
<th>Phase</th>
<th>Purpose</th>
<th>Number of people who take part</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>Small studies of safety and bioavailability, often dose-escalating</td>
<td>Pocock recommends 20 to 80</td>
</tr>
<tr>
<td>Phase II</td>
<td>Small-scale investigation of safety and efficacy, often screens Rx’s to find candidates for Phase III</td>
<td>Pocock: 100-200</td>
</tr>
<tr>
<td>Phase III</td>
<td>Larger trial to establish efficacy and thereby affect clinical practice</td>
<td>Several hundred to several thousand people (++)</td>
</tr>
<tr>
<td>Phase IV</td>
<td>To further assess the long-term safety and effectiveness of the new treatment</td>
<td>Several hundred to several thousand people (****)</td>
</tr>
</tbody>
</table>

Examples

- Pilot study (feasibility): N = 18
- Phase I (toxicity): 20 ≤ N ≤ 40
- Phase II (efficacy): 30 ≤ N ≤ 100
- Phase III (confirmatory): N > 100
- Primary Prevention Trials: N > 10,000
  e.g. BCPT (Tamoxifen): N = 16,000 (13,388)
  PHS (aspirin, β-carotene): N = 22,071

Drug Development Process

1. Preclinical research
2. Compound synthesis & purification
3. Animal testing (ADME)
4. Phase I (first time into humans)
5. Phase II (first data on effectiveness)
6. Phase III (large scale effectiveness & safety)
7. Phase IV (further efficacy, safety, other indications)
Minimise sources of error

Systematic errors (bias)
- "inaccuracy which is different in its size or direction in one of the groups under study than the others"
- Minimise bias by ensuring that the methods used are applied in the same manner to all subjects irrespective of which group they belong to.

Random errors (chance)
- "Inaccuracy which is similar in the different groups of subjects being compared"
- Adequate sample size, accurate methods of measurement

Random vs Systematic error

Random
- ↑ Sample size

Bias
- ↑ Sample size

Forms of Bias
- Subversion Bias
- Technical Bias
- Attrition Bias
- Consent Bias
- Ascertainment Bias
- Dilution Bias
- Recruitment Bias

Bias (cont)
- Resentful demoralisation
- Delay Bias
- Chance Bias
- Hawthorne effect
- Analytical Bias

Type of studies
Internal Validity (validity)
External Validity (generalizability)

Internal Validity (validity)
- Efficacy
- Does receiving treatment work under clean conditions?

External Validity (generalizability)
- Effectiveness
- Does offering treatment help under ordinary circumstances?

Figure 7.3 Efficacy and effectiveness.

METHODOLOGICAL BASIS

Trepanation once was a "scientific cure" for mental illness.

"A clinical trial is a carefully and ethically designed experiment with the aim of answering some precisely framed question"

Sir Austin Bradford Hill (1977)

James Lind
- Born Edinburgh 1716
- On Salisbury in 1747 he allocated 12 men with scurvy
  - Cider
  - Seawater
  - Horseradish, mustard, garlic
  - Nutmeg
  - Elixir Vitriol
  - Oranges and Limes
James Lind 1747, A Treatise of the Scurvy

- “I took 12 patients in the scurvy on board the Salisbury at sea. The cases were as similar as I could have them…they lay together in one place and had one diet common to them all. Two of these were ordered a quart of cider per day…Two others took 25 gutts of elixir vitriol…Two others took two spoonfuls of vinegar…Two were put under a course of sea water…Two others had two oranges and one lemon given them each day…Two others took the bigness of nutmeg. The most sudden and visible good effects were perceived from the use of oranges and lemons, one of those who had taken them being at the end of 6 days fit for duty. The other was appointed nurse to the rest of the sick.”
Terminology: explanatory versus pragmatic

- Explanatory trials
  - estimates efficacy - that is the benefit the treatment produces under ideal conditions
- Pragmatic trials
  - estimates effectiveness - that is the benefit the treatment produces under routine clinical practice


Methodological Tools to diminish Bias in Clinical trials

- Experimental, longitudinal design
- Control Group
- Randomization
- Blinding

JUSTIFICATION OF CONTROL

Why do we need a control group?

- Don’t need a control group if completely predictable results
- But
  - No intervention has 100% efficacy
  - Many diseases recover spontaneously

Regression to the mean

- Occurs when an intervention aimed at a group or characteristic that is very different from average
- Ex:
  - selecting people because of high blood pressure, then measuring them in future will see the BP measurements closer to the mean of the population

Morton and Torgerson BMJ 2003 326:3083-4
Hawthorne Effect

- The presence of the researcher affects the behavior of the subjects (Western Electric Co. Chicago, IL)

- Effect observed due to the study effect rather than a treatment effect

Placebo Effect

- Effect (usually, but not always positive) attributed to the expectation that a therapy will have an effect
- The effect is due to the power of suggestion
- A placebo is an inert medication or procedure

Waber et al. 2008 JAMA Commercial Features of Placebo and Therapeutic Efficacy

http://jama.ama-assn.org/cgi/content/full/299/9/1016

Ferran.Torres@uab.es

EFFECT OF AN INTERVENTION
Diminishing Bias in Clinical trials

- Experimental, longitudinal design
- Control Group
- Randomization
- Blinding

Types of control groups

- Concurrent
- Historical controls
  - Comparison with results obtained in previous similar patients
- No treatment control
  - E.g.: standard practice is observation after surgery
- ‘sham’ treatment control
  - Sham surgery
- Placebo
  - Active groups:
    - A low-dose group
    - Best current therapy

What is randomization

- Scientifically managed chance
- Each participant has the same chance as the other participant in receiving the study treatment

Randomization

- 1st introduced by Fisher in 1926 in agriculture research
- 1st clinical trial used randomization - Amberson et al. (1931)
  - Matching 24 pts with pulmonary tuberculosis into 2 comparable groups of 12 each
  - Flip a coin to decide which group received sanocrysin, a gold compound
- Streptomycin trial by British Medical Research Council (1948)
  - 1st to use random numbers in allocation pt to experimental or control groups
Why Randomise

"To eliminate biases that may lead to systematic differences between treatment groups" – Altman 1991

Randomization

• Randomization tends to produce study groups that are:
  - Comparable with respect to known and unknown risk factors
  - Removes investigator bias in the allocation and treatment of patients
  - Guarantees statistical tests will have valid significance levels.

• Random allocation facilitates blinding

Schulz KF1998, Armitage P1982

Randomisation - Allocation Concealment

• “A technique used to prevent selection bias by concealing the allocation sequence from those assigning participants to intervention groups, until the moment of assignment”

http://www.consort-statement.org/allocationconcealment.htm

Allocation Concealment

• Examples
  - Sequentially numbered opaque sealed envelopes (SNOSE)
  - Sequentially numbered containers
  - Pharmacy controlled
  - Central randomisation

Comparison of concealment

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Effect Size</th>
<th>OR</th>
<th>P &lt; 0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequate</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclear</td>
<td>0.67</td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Inadequate</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Small VS Large Trials

- Small trials tend to give greater effect sizes than large trials
- One explanation is because of poor allocation concealment in small trials
- In ‘secure’ allocation trials the effect reduced effect by 51%.

Randomisation - Sequence Generation

- Simple
- Blocks
- Stratification
- Adaptive randomisation

Randomisation ex.

<table>
<thead>
<tr>
<th>Ex: Simple</th>
<th>Ex: Blocks of fixed size = 4 elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pat Treat</td>
<td>Pat Block Treat</td>
</tr>
<tr>
<td>1 A</td>
<td>1 A</td>
</tr>
<tr>
<td>2 A</td>
<td>1 B</td>
</tr>
<tr>
<td>3 B</td>
<td>2 B</td>
</tr>
<tr>
<td>4 B</td>
<td>3 A</td>
</tr>
<tr>
<td>5 B</td>
<td>4 A</td>
</tr>
<tr>
<td>6 B</td>
<td>5 2 A</td>
</tr>
<tr>
<td>7 A</td>
<td>6 2 A</td>
</tr>
<tr>
<td>8 A</td>
<td>7 2 B</td>
</tr>
<tr>
<td>9 A</td>
<td>8 2 B</td>
</tr>
<tr>
<td>10 B</td>
<td>9 3 A</td>
</tr>
<tr>
<td>11 A</td>
<td>10 3 A</td>
</tr>
<tr>
<td>12 A</td>
<td>11 3 B</td>
</tr>
</tbody>
</table>

Ex.: Simple

Ex.: Blocks of fixed size = 4

Ferran.Torres@uab.es

Blocking

- Ensures close balance of the numbers in each group at all times during trial.
- More importantly when stratified.
- Problem If block size is discovered.
- Remedy: more blinding, varying block size, larger blocks.

Scheme of stratified randomization
Stratified randomization

- When
  - Factors that are believed to significantly influence the outcome

- Small numbers of variables

**MINIMIZATION**

**COMET 1**

- The trial was using MINIMISATION via a computer programme.
- The groups were minimised on age of mother and her ethnicity.
- Programme had a fault.


**COMET 1 - Technical Bias**

<table>
<thead>
<tr>
<th>AGE</th>
<th>Traditional</th>
<th>Combined</th>
<th>Low dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>388</td>
<td>335</td>
<td>331</td>
</tr>
<tr>
<td>&lt;25 years</td>
<td>13</td>
<td>179</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>(3%)</td>
<td>(53%)</td>
<td>(52%)</td>
</tr>
</tbody>
</table>

**Ethics of Randomization**

- Statistician/clinical trialist must sell benefits of randomization
- Ethics ➞ MD should do what he thinks is best for his patient
  - Two MD’s might ethically treat same patient quite differently
  - Whenever there is a reasonable doubt

  1. If MD “knows” best treatment, should not participate in trial
  2. If in doubt, randomization gives each patient equal chance to receive one of therapies (i.e. best)
  3. More ethical way of practicing medicine

- Byar et al. (1976) N Engl
  1. RCT is best method to find out!
  2. Reduces risk of being on inferior treatment
  3. Reduces risk for future patients

**BLINDING**

AGAINST OBSERVATION BIAS...
Blinding

• “The practice of keeping the trial participants, care providers, those collecting data, and sometimes even those analyzing data unaware of which intervention is being administered to which participant”

• Blinding is intended to prevent bias on the part of study personnel

http://www.consort-statement.org/blinding.htm

Blinding

• Open studies are more likely to favor experimental interventions over the controls
  
  Colditz GA 1989

• Appropriate use of blinding may decrease overoptimistic estimates of treatment effect
  
  Juni P et al 2001

Variability & Baseline Imbalances

BLINDING

• LEVELS:
  
  • Open or Unblinded
  
  • Single-blinded
  
  • Double-blinded
  
  • “Triple-blinded”

  External evaluation by a blinded observer

Blinding

• Difficulties:
  
  – Treatments may be different in:
    – Timing
    – Route
    – External appearance

  • Double-dummy

Treatment A

Group A

Placebo

Treatment B

Group B

Placebo

Blinding

• More difficulties (2):

  – Different direct/indirect effects:
    – Laboratory analyses
    – Adverse events:
      » e.g.: bradicardia (beta-blockers)

Ferran.Torres@uab.es

Ferran.Torres@uab.es

Ferran.Torres@uab.es

Ferran.Torres@uab.es
Intention-to-treat analysis (ITT)

Once randomized...

... Always analyzed!

- A strategy for analyzing data in which all participants included regardless they completed the intervention

Attrition Bias

- We can avoid some of the problems with attrition bias by using Intention to Teach (or treat) Analysis, where we keep as many of the patients in the study as possible even if they are no long ‘on treatment’.

Loss to follow-up

- “With as little as 10% loss to follow-up in each arm, the chance of obtaining a false positive result can easily double”

Lachin JM. 2000
In practice, this ideal may be difficult to achieve, for reasons to be described.

The term 'full analysis set' is used to describe the analysis set which is as close as possible to the ITT principle.

Preservation of the initial randomisation
- Important in preventing bias and
- In providing a secure foundation for statistical tests.

In many clinical trials provides a conservative strategy.

Under many circumstances it may also provide estimates of treatment effects which are more likely to mirror those observed in subsequent practice.

Increases external validity.

**Accepted exclusions in ICHE9**

- Including the failure to satisfy major entry criteria (eligibility violations)
- The failure to take at least one dose of trial medication and the
- Lack of any data post randomisation.

**Per Protocol (PP)**

- The 'per protocol' set of subjects:
  - 'valid cases', efficacy sample or the ' evaluable subjects sample'
- A subset of full analysis set who are more compliant with the protocol:
  - (i) the completion of a certain pre-specified minimal exposure to the treatment regimen;
  - (ii) the availability of measurements of the primary variable(s);
  - (iii) the absence of any major protocol violations including the violation of entry criteria.
- The precise reasons for excluding subjects from the per protocol set should be fully defined and documented before breaking the blind.

**Role of the Populations**

- In confirmatory trials it is usually appropriate to plan to conduct both an analysis of the full analysis set and a per protocol analysis.
- However:
  - ITT: the main approach for superiority
  - PP: many times preferred for the non-inferiority testing.
RCT as the Gold Standard

- The randomised controlled trial is widely regarded as the gold standard for evaluating health care technologies because it allows us to be confident that a difference in outcome can be directly attributed to a difference in the treatments and not due to some other factor.

RCT strengths

- Confounding variables minimised
- Only research design which can in principle yield causal relationships
  - can clarify the direction of cause and effect
- Accepted by EBM school
- Don’t have to know everything about the participants

RCT limitations

- Contamination of intervention groups
- Comparable controls
- Problems with blinding
- What to do about attrition?
- Are patients/professionals willing to be in trial different from ‘refusers’? - external validity
- Cost!

Summary

- “Gold standard” of research designs
- Individual patients are randomly allocated to receive the experimental treatment (intervention group) or the standard treatment (control group)
- Maximizes the potential for attribution
- Randomisation guards against selection bias between the two treatment groups
- Standard statistical analysis
- Good internal validity
- May lack generalisability due to highly selected participants
- Can be costly to set up and conduct, ethical issues