Critical appraisal of randomized clinical trial?

Samir Haffar M.D.
Assistant Professor of Gastroenterology
Al-Mouassat University Hospital – Damascus – Syria
Hierarchy of evidence in quantitative studies

Perhaps the first large-scale clinical trial using a properly designed randomized schema
Sir Austin Bradford Hill (1897-1991)

British epidemiologist & statistician
The father of modern RCTs
First RCT in the United States

1951

NIH started a study of adrenocorticotropic hormone (ACTH), cortisone & aspirin in the treatment of rheumatic heart disease*

Number of RCT per year

≈ 20,000 trials published each year
> 500,000 trials in total

Basic structure of a RCT

Parallel trial

Parallel trial is the most frequently used design

Basics of RCT – 1

• **Participants**
  Patients – relatives of pts – healthy volunteers – groups

• **Investigators**
  People who design & carry out study & analyze results

• **Interventions**
  Preventive strategies, screening, & **treatments**

Basics of RCT – 2

Control group should receive one of the following:

1. **Placebo**
   Inert pills that appear identical to trial therapy

2. **Gold standard therapy**
   It may be unethical to treat patient with placebo

3. **New treatment**

Basics of RCT – 3

RCTs are regarded as

- **Quantitative** studies (quantified outcomes)
- Most rigorous method of **hypothesis testing**
- **Experimental** studies versus observational studies
- Gold standard to evaluate **effectiveness of interventions**

Some historical examples of treatments with dramatic effects

- Insulin for diabetes
- Blood transfusion for severe hemorrhagic shock
- Defibrillation for ventricular fibrillation
- Neostigmine for myasthenia gravis
- Tracheotomy for tracheal obstruction
- Drainage for pain associated with abscesses
- Pressure or suturing for arresting hemorrhage

Parachutes reduce risk of injury after gravitational challenge.

Their effectiveness has not been proved with RCTs.
Ethics committee

• **Include:**
  Layman, religious man, lawyers, researchers & clinicians

• **Responsibilities:**
  Protect rights & welfare of research subjects
  Determine if the potential benefits warrant the risks
  Ensure that **informed consent** is obtained
  Prevent unscientific or unethical research
The trial team

- Principal investigator
- Trial coordinator or manager
- Trial programmer
- Data manager or clerks
- Trial statistician
  - Planning phase
  - Interim analyses
  - Final analysis
- Trial secretary
Randomized controlled trial

1. Sample size
2. Randomization
3. Blinding (Masking)
4. Outcomes
5. Intention to treat analysis (ITT)
6. Measurement of treatment effect
7. Applicability of results to your patients

Critical appraisal
Flow diagram for a RCT

Sample size in RCTs

The “Universe” & the “Sample”

Statistical inference

Making statistical inferences about a population from a sample by means of **significance test & CI**

Component of sample size calculation

1. **Type I error (α)**
   - False positive = 0.05

2. **Type II error (β)**
   - False negative = 0.20
   - Power (1 - β)

3. **Event rate in control group**

4. **Event rate in treatment group**

Randomization in RCTs

If the study wasn’t randomized we’d suggest that you stop reading it.
Randomization in RCTs

If the study wasn’t randomized we’d suggest that you stop reading it.

Goal of randomization

Comparable groups to known prognostic factors

Beta-Blocker Heart Attack Trial - Baseline comparisons

<table>
<thead>
<tr>
<th></th>
<th>Propranolol (N-1,916)</th>
<th>Placebo (N-1,921)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Age (yrs)</td>
<td>55.2</td>
<td>55.5</td>
</tr>
<tr>
<td>Male (%)</td>
<td>83.8</td>
<td>85.2</td>
</tr>
<tr>
<td>White (%)</td>
<td>89.3</td>
<td>88.4</td>
</tr>
<tr>
<td>Systolic BP</td>
<td>112.3</td>
<td>111.7</td>
</tr>
<tr>
<td>Diastolic BP</td>
<td>72.6</td>
<td>72.3</td>
</tr>
<tr>
<td>Heart rate</td>
<td>76.2</td>
<td>75.7</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>212.7</td>
<td>213.6</td>
</tr>
<tr>
<td>Current smoker (%)</td>
<td>57.3</td>
<td>56.8</td>
</tr>
</tbody>
</table>

Table comparing baseline characteristics presented in RCT reports
Randomization

- Simple randomization
- Random table
- Block randomization
- Stratified randomization
- Minimization method
- Unequal randomization
- Allocation concealment

Inacceptable

Preferred
## 2 principles of randomization

Regardless of the method of randomization used, investigators should follow two principles:

- **First**: They must define the rules that will govern allocation.
- **Second**: They should follow the same rules strictly throughout the whole study.
Simple randomization

Inacceptable

- Toss of a coin
- Date of birth (even numbers to group A)
- Hospital admission number
- Date seen in clinic

Patients seen this week (group A)
Those seen next week (group B)

Problems arise from openness of allocation system
Allocation concealment

- **Sealed opaque envelope**
  Investigator open several envelopes before allocation
  Allocation seen if envelope held against bright light

- **Remote randomization (preferred)**
  Assignment removed from those making assignments:
  By telephone – Over the internet

Randomization should be distant
& separate from clinicians conducting the trial
RCT of open vs. lap appendectomy

- Trial ran smoothly during the day
- Surgeon’s presence required for lap procedure at night
- Residents at night held semiopaque envelopes up to light & opened first envelope that dictated open procedure
- First eligible patient in the morning allocated to lap group
- If patients seen at night sicker than those seen in the day, this behavior bias results against open procedure

Estimates of treatment effect exaggerated by 40% in trials with unconcealed compared with concealed randomization.
Blinding in RCTs

Blinking /masking in RCTs

Blinding or masking

- Keep one or more of the people involved in the trial unaware of the intervention that is being evaluated
- Purpose: decrease risk of observation bias
- **What matters**
  - Not the number of people blinded during a trial
  - But the number & role of those who are not blinded

Blinding is not always appropriate or possible
Blinding or Masking

Blinding can be implemented in at least 6 levels in RCTs

- Participants
- Investigators who administer interventions
- Investigators taking care of the participants
- Investigators assessing the outcomes
- Data analyst
- Investigators who write results of the trial

Usually the same
Blinding or masking

Depending on blinding extent, RCTs classified as

- Open label (everyone aware)
- Single-blind
- Double-blind
- Triple-blind
- Quadruple-blind & so on
The term ‘double-blind RCT’, so often used to represent the ultimate in design to produce valid results, is confusing.

Why is blinding so important?

- Trials that were not double blinded yielded larger estimates of treatment effects than double blinded trials (OR exaggerated on average by 17%).

- Blinding is weaker than allocation concealment in preventing biases.

A humorous example of blinding/masking

4 Outcomes in RCTs

### Outcomes in RCTs – 1

<table>
<thead>
<tr>
<th>Primary outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>One primary outcome</strong> (usually)</td>
</tr>
<tr>
<td>Most important outcome (stroke in carotid endarterectomy)</td>
</tr>
<tr>
<td><strong>Composite outcomes</strong> (sometimes – can mislead)</td>
</tr>
<tr>
<td>- Drug in MI: death, non fatal MI, hospitalization for ACS</td>
</tr>
<tr>
<td>- Validity depends on similarity in patient importance, treatment effect, &amp; number of events across components</td>
</tr>
<tr>
<td>- Abandoned if large variations exist between components</td>
</tr>
</tbody>
</table>

### Outcomes in RCTs – 2

#### Surrogate outcomes
- Used in case of rare events of clinical importance
- Studies in cytoprotection of *NSAIDs*
- Endoscopic ulcers surrogates of bleeding or perforated PU

#### Secondary outcomes (usually multiple)
- Other variables important to research question (drugs SE)
- Too much emphasis if no change in primary outcome
NSAID-related GI side effects

Relative Severity

- Serious GI Events
- Clinical Ulcers
- Endoscopic Ulcers
- GI Symptoms

Relative Frequency
5 Intention to treat analysis (ITT)

Participants who not complete the study

- Some participants would not complete the study because of **misdiagnosis, non-compliance, or withdrawal**
- When such patients excluded from analysis, we can no longer be sure that important prognostic factors in the 2 groups are similar which lead to potential bias
- To reduce this bias, results should be analyzed on an ‘intention to treat’ basis
Intention to treat analysis

Form of quality control rather than analytic tool

- Strategy in conduct & analysis of RCT ensuring that all patients allocated to treatment or control groups analyzed together as representing that treatment arm whether or not they received prescribed therapy or completed study

- Randomized participants = Analyzed participants

6 Measurement of treatment effect

Measurement of treatment effect in RCTs

- p value (\(p\))
- Relative Risk (\(RR\))
- Odds Ratio (\(OR\))
- Confidence Intervals (\(CIs\))
- Number Needed to Treat (\(NNT\))

Data analyzed as trial proceeds (interim analysis) or at the ends of the trial
Probability value (p Value)

- p value is probability that observed difference between 2 treatment groups might occur by chance.
- Many use p value of 0.05 as cut off for significance.
  - $p < 0.05$: Observed difference between groups is so unlikely to have occurred by chance. Considered as *statistically significant*.
  - $p > 0.05$: Observed difference between groups might have occurred by chance. Considered as *not statistically significant*. 
Probability value (p value)

- \( p > 0.05 \) Statistically insignificant
- \( p < 0.05 \) Statistically significant

Statistically significant \( \rightarrow \) Doesn't mean \( \rightarrow \) Clinically significant
Statistical versus clinical significance

• Pentoxifylline vs placebo in PAD* (1992)
  40 patients randomized to pentoxifylline or placebo
  Maximum pain-free walking distance longer in
  pentoxifylline group than in placebo group \( (p < 0.001) \)
  Conclusion: pentoxiphylline clinically effective

• Close examination of data:
  Difference in maximum walking distance: 3.5 feet
  Doctors & patients consider it not clinically significant

* PAD: Peripheral Arterial Disease
## Risk & Relative Risk (RR)

### Risk

Number of patients fulfill criteria for a given end point divided by total number of patients

i.e.: Diarrhea during tt with antibiotic in 4 of 10 patients

Risk of patients: \( \frac{4}{10} = 0.4 \)

Diarrhea in control group in 1 of 10 persons

Risk of controls: \( \frac{1}{10} = 0.1 \)

### Relative Risk

Risk of patient / risk of control group

**RR**: \( \frac{0.4}{0.1} = 4 \)
# Odds & Odds Ratio (OR)

## Odds

Number of patients fulfill criteria for given endpoint divided by number of patients who do not

i.e.: Diarrhea during tt with antibiotic in 4 of 10 patients

- Odds of patients: 4 / 6 = **0.66**

Diarrhea in control group in 1 of 10 persons

- Odds of controls: 1 / 9 = **0.11**

## Odds Ratio

Odds of patients / odds of control group

**OR** = 0.66 / 0.11 = **6**
Risk & Odds

Risk

\[ \frac{a}{a + b} \]

Odds

\[ \frac{a}{b} \]
Interpretation of RR & OR

RR or OR should be accompanied by their CIs

RR or OR > 1
Increased likelihood of outcome in treatment group

RR or OR < 1
Decreased likelihood of outcome in treatment group

RR or OR = 1
No difference of outcome between tt & control group
Odds ratio or relative risk?

OR will be close to RR if endpoint occurs infrequently (<15%).
If outcome is more common, OR will differ increasingly from RR.

Significance of CI

- When we test a new Crohn’s disease drug on randomly selected sample of patients, the treatment effect we will get will be an estimate of the “true” treatment effect for the whole population of patients with CD in the country.
- *95% CI* of estimate will be range within which we are 95% certain the true population treatment effect will lie.
## Confidence intervals

<table>
<thead>
<tr>
<th>Value</th>
<th>95% CI are commonly used</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90 or 99% CI are sometimes used</td>
</tr>
<tr>
<td><strong>Width of CI</strong></td>
<td>Indicates precision of the estimate</td>
</tr>
<tr>
<td></td>
<td>Wider the interval, less the precision</td>
</tr>
<tr>
<td><strong>CI includes 1</strong></td>
<td>No statistically significant difference</td>
</tr>
<tr>
<td><strong>CI doesn’t include 1</strong></td>
<td>Statistically significant difference</td>
</tr>
</tbody>
</table>
(a) Statistically significant, low precision
(b) Statistically significant, high precision
(c) Not statistically significant, low precision
(d) Not statistically significant, high precision

Influence of sample size on CI precision

Width of CI (precision of the estimate) decreases with increasing sample size

Confidence interval or p value?

• Authors of articles could report both p values & CIs
• CI convey more useful information than p values
• **If only one is to be reported, then it should be the CI**
• p value is less important & can be deduced from CI
Number Needed to Treat (NNT)

- **Relative Risk (RR)**
  Risk in treatment group / risk in control group

- **Relative Risk Reduction (RRR)**
  $1 - RR$

- **Absolute Risk Reduction (ARR)**
  Risk in control group – risk in treatment group

- **NNT** (expressed in clinically relevant way)
  $1 / ARR$
Measurement of treatment effect in RCTs

- p value (p)
- Relative Risk (RR)
- Odds Ratio (OR)
- Confidence Intervals (CIs)
- Number Needed to Treat (NNT)
Subgroup analysis

Post-hoc analysis

- In large trials not demonstrating overall favorable trend, it is common to conduct subgroup analyses to find one or more subgroups in which treatment “really works”
- Literature is replete with unconfirmed subgroup findings
- Post-hoc results should be regarded as inconclusive
- May be of value for hypothesis generation
ISIS-2 trial - Subgroup analysis

- Effects of streptokinase &/or aspirin on short-term mortality in patients admitted with AMI
- Mortality benefits for both active interventions
- In subgroup analyses:
  - Patients born under Zodiac signs of *Gemini & Libra* 5% higher mortality on aspirin vs placebo
  - Patients born under *other Zodiac signs* 30% lower mortality on aspirin vs placebo

ISIS-2 trial

Streptokinase &/or aspirin on AMI mortality

It is very difficult to make a judgment if statistics used in a study are appropriate & applied correctly.

Joe, could you get a significant p-value out of all this?
Basic understanding of medical statistics will enable us to detect the more obvious errors.

7 Applicability of results to your patients

### External validity

**Applicability of results to your patients**

<table>
<thead>
<tr>
<th>Issues needed to consider before deciding to incorporate research evidence into clinical practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Similarity of study population to your population</td>
</tr>
<tr>
<td>• Benefit versus harm</td>
</tr>
<tr>
<td>• Patients preferences</td>
</tr>
<tr>
<td>• Availability</td>
</tr>
<tr>
<td>• Costs</td>
</tr>
</tbody>
</table>

The problem of applying trial results

The Trial

patients

The trial report

The actual patients
Critical appraisal of a RCT

A mother is a perfect expert on her child, but may not be the most objective.
Internal & external validity of a RCT

<table>
<thead>
<tr>
<th>Critical appraisal of a RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>• Internal validity of a trial</strong></td>
</tr>
<tr>
<td>– Randomization</td>
</tr>
<tr>
<td>– Blinding (Masking)</td>
</tr>
<tr>
<td>– Follow-up</td>
</tr>
<tr>
<td>– Outcomes</td>
</tr>
<tr>
<td>– Analysis</td>
</tr>
<tr>
<td>– Biases</td>
</tr>
</tbody>
</table>

**• External validity of a trial (generalizability)**

– Applicability of results to your patients
Bias

• Difference between the study results & the truth
• Of course, we can never know the truth, but we try to come as close as possible by performing & using well-designed & well executed studies

• **Non-systematic bias** (random error or chance)
  Occurs to similar extent in all subjects for both group
  Predictable – Less important than systematic bias

• **Systematic bias** (non-random error)
  Most serious type of bias: under or over-estimation

### Main types of biases in RCTs

<table>
<thead>
<tr>
<th>Biases</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>During planning phase of a RCT</td>
<td>Choice-of-question bias</td>
</tr>
<tr>
<td></td>
<td>Regulation bias</td>
</tr>
<tr>
<td></td>
<td>Wrong design bias</td>
</tr>
<tr>
<td>During course of a RCT</td>
<td>Selection bias</td>
</tr>
<tr>
<td></td>
<td>Observation bias</td>
</tr>
<tr>
<td></td>
<td>Population choice bias</td>
</tr>
<tr>
<td></td>
<td>Intervention choice bias</td>
</tr>
<tr>
<td></td>
<td>Control group bias</td>
</tr>
<tr>
<td></td>
<td>Outcome choice bias</td>
</tr>
<tr>
<td>During reporting of a RCT</td>
<td>Withdrawal bias</td>
</tr>
<tr>
<td></td>
<td>Selective reporting bias</td>
</tr>
<tr>
<td></td>
<td>Fraud bias</td>
</tr>
</tbody>
</table>

Fraud bias

John Darsee (Harvard researcher in cardiology)

- Fabricated data in a study on dogs in 1981
- Fabricated data during his:
  - Undergraduate days [Notre Dame University, (1966-70)]
  - Residency & fellowship [Emory University, (1974-79)]
  - Fellowship [Brigham & Women’s, Harvard, (1979-81)]
- > 100 papers & abstracts most in prestigious journals
- His coauthors had too little contact with the research
  Listed over their objections (had been helpful in the past)
Lessons learned from the Darsee’s affair

1. Little can be done to stop unscrupulous scientist even when he collaborates with knowledgeable colleagues.
2. Inability of peer review to detect the fraud.
4. Focus on responsibilities & contributions of coauthors.
5. Misconduct investigations may need to examine a researcher’s entire work over many years.

One of the lessons learned from Darsee’s case

‘Once a crook, often always a crook’

Darsee was found to have had a long history of faking his results in different projects & in different settings

Existing tools to assess trial quality

• Several components grouped in
  **Scales**   Each item scored numerically
               Overall quality score is generated
  **Checklists** Components evaluated separately
                  No numerical scores

• Systematic search of literature in 1995 identified
  **25 scales** & **9 checklists** for assessing trial quality*

The Jadad scale

Scores: 0 - 5 points – Poor quality if ≤ 2 points

### Appraising a RCT (checklist) – 1

<table>
<thead>
<tr>
<th>Are the results valid?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>At start of trial</strong></td>
</tr>
<tr>
<td>1 Were the patients <strong>randomized</strong>?</td>
</tr>
<tr>
<td>2 Was the randomization <strong>concealed</strong>?</td>
</tr>
<tr>
<td>3 Similar prognostic factors in 2 groups?</td>
</tr>
<tr>
<td><strong>During trial</strong></td>
</tr>
<tr>
<td>4 Was trial <strong>blinded</strong> &amp; to what extent?</td>
</tr>
<tr>
<td><strong>At end of trial</strong></td>
</tr>
<tr>
<td>5 Was <strong>follow-up</strong> complete?</td>
</tr>
<tr>
<td>6 Was <strong>ITT</strong> principle applied?</td>
</tr>
<tr>
<td>7 Was the trial <strong>stopped early</strong>?</td>
</tr>
</tbody>
</table>

# Appraising a RCT (checklist) – 2

## What are the results?

8. How **large** was the treatment effect?
9. How **precise** was estimate of treatment effect (CI)?

## How can I apply the results to patient care?

10. Were the study patients **similar** to my patient?
11. Were all patient-**important outcomes** considered?
12. Are the likely treatment benefits worth **harm & cost**?

Scales or checklists?
No consensus on which is preferable

<table>
<thead>
<tr>
<th>Quality assessment in systematic reviews</th>
<th>Medical journals</th>
<th>CDSR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of SR</td>
<td>78 SR in 204 journals</td>
<td>36 SR</td>
</tr>
<tr>
<td>Checklists</td>
<td>20/78 (26%)</td>
<td>92 %</td>
</tr>
<tr>
<td>Scales</td>
<td>52/78 (67%)</td>
<td>None</td>
</tr>
</tbody>
</table>

*CDSR: Cochrane Database of Systematic Reviews
Moher D et al. Health Technol Assess 1999; 3 (12).*
Improving quality of reports

- RCTs
  - CONSORT*
    - Consolidated Standards of Reporting Trials

- Meta-analysis
  - QUOROM**
    - Quality of Reporting of Meta-analyses

- Diagnostic accuracy study
  - STARD***
    - Standards for Reporting of Diagnostic Accuracy

CONSORT statement

Targeted authors of trial reports rather than readers

- **Experts**  Clinical epidemiologists, journal editors, & biostatisticians published CONSORT statement

- **Aim**  Improve standard of written reports of RCTs

- **Results**  Latest version of CONSORT statement includes

  - **Flow diagram:** Patients progress through a trial
  - **Checklist:** 22 items

---

Flow diagram of a RCT

### CONSORT statement

<table>
<thead>
<tr>
<th>Paper Section &amp; Topic</th>
<th>Item</th>
<th>Descriptor</th>
<th>Reported on Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title &amp; abstract</td>
<td>1</td>
<td>How participants allocated to interventions</td>
<td></td>
</tr>
<tr>
<td>Introduction background</td>
<td>2</td>
<td>Scientific background</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>3</td>
<td>Criteria for participants, settings, locations</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>Details of interventions for each group</td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td>5</td>
<td>Specific objectives &amp; hypotheses</td>
<td></td>
</tr>
<tr>
<td>Objectives</td>
<td>6</td>
<td>Defined <strong>primary &amp; secondary outcomes</strong></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>7</td>
<td>How <strong>sample size</strong> was determined?</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>8-9-10</td>
<td><strong>Allocation concealment</strong>, implementation</td>
<td></td>
</tr>
<tr>
<td>Randomization</td>
<td>11</td>
<td>Whether or not blinding applied</td>
<td></td>
</tr>
<tr>
<td>Blinding (masking)</td>
<td>12</td>
<td><strong>Statistical methods</strong> used</td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>13</td>
<td><strong>Flow diagram</strong> strongly recommended</td>
<td></td>
</tr>
<tr>
<td>Participant flow</td>
<td>14</td>
<td>Periods of recruitment &amp; <strong>follow-up</strong></td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td>15</td>
<td><strong>Baseline characteristics</strong> of each group</td>
<td></td>
</tr>
<tr>
<td>Baseline data</td>
<td>16</td>
<td>No of participants in each group</td>
<td></td>
</tr>
<tr>
<td>Numbers analyzed</td>
<td>17</td>
<td>Summary of results with <strong>95% CI</strong></td>
<td></td>
</tr>
<tr>
<td>Outcomes, estimation</td>
<td>18</td>
<td>Subgroup &amp; adjusted analyses</td>
<td></td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>19</td>
<td><strong>All important adverse events</strong></td>
<td></td>
</tr>
<tr>
<td>Adverse events</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comment</td>
<td>20</td>
<td>Interpretation of the results</td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>21</td>
<td><strong>External validity</strong> of trial findings</td>
<td></td>
</tr>
<tr>
<td>Generalizability</td>
<td>22</td>
<td>General interpretation of results</td>
<td></td>
</tr>
<tr>
<td>Overall evidence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reasons for doing RCTs

- Only study design that can prove causation
- Required by FDA (and others) for **new drugs** and some devices
- Most influential to clinical practice
Disadvantages of RCTs

• Expensive: typically in $ millions
• Time consuming: typically years
• Can only answer a single question
• May not apply to some patients in practice
• May not be practical
• Generally difficult to get funded
• Organizationally complex
Carefully conducted observational studies may provide more evidence than poor RCTs*

Unfortunately, a perfect trial can only exist in our imagination**


References

Understanding Clinical Papers
DAVID BOWERS
ALLAN HOUSE
DAVID OWENS
Second Edition

Randomized Controlled Trials
QUESTIONS, ANSWERS AND MUSINGS
second edition
ALEJANDRO R. JADAD
MURRAY W. ENKIN

Users' Guides to the Medical Literature
ESSENTIALS OF EVIDENCE-BASED CLINICAL PRACTICE
SECOND EDITION
Gordon Guyatt, MD • Drummond Rennie, MD
Maureen O. Meade, MD • Deborah J. Cook, MD

John Wiley & Sons
2006
Blackwell Publishing
2007
Mc Graw Hill
2008
Thank You
Thank You
<table>
<thead>
<tr>
<th>Types of RCTs</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCTs according to how participants are exposed to the interventions</td>
<td>Parallel trials</td>
</tr>
<tr>
<td></td>
<td>Factorial trials</td>
</tr>
<tr>
<td></td>
<td>Cross-over trials</td>
</tr>
<tr>
<td>RCTs exploring different aspects of the interventions they evaluate</td>
<td>Efficacy &amp; effectiveness trials</td>
</tr>
<tr>
<td></td>
<td>Equivalence trials</td>
</tr>
<tr>
<td></td>
<td>Phase III trial</td>
</tr>
<tr>
<td>RCTs by unit of analysis</td>
<td>Body part</td>
</tr>
<tr>
<td></td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td>Group</td>
</tr>
<tr>
<td>RCTs according to the number of participants</td>
<td>Fixed to variable sample size</td>
</tr>
<tr>
<td></td>
<td>N-of-1 trials to mega-trials</td>
</tr>
<tr>
<td>RCTs according to whether investigators know which intervention is being</td>
<td>Open trials</td>
</tr>
<tr>
<td>assessed</td>
<td>Blinded trials</td>
</tr>
<tr>
<td>RCTs that take into account non-randomized individuals &amp; participants’</td>
<td>Zelen’s design</td>
</tr>
<tr>
<td>preferences</td>
<td>Comprehensive cohort design</td>
</tr>
<tr>
<td></td>
<td>Wennberg’s design</td>
</tr>
</tbody>
</table>
References

Randomized Controlled Trials: Questions, Answers and Musings, 2nd Edition
ALEJANDRO R. JADAD, MURRAY W. ENKIN
Blackwell Publishing
2007

Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice, 2nd Edition
GORDON GUYATT, MD, DRUMMOND RENNIE, MD, MAUREEN O. MEADE, MD, DEBORAH J. COOK, MD
McGraw Hill
2008
# Types of RCTs

<table>
<thead>
<tr>
<th>RCTs</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCTs according to how participants are exposed to the interventions</td>
<td>Parallel trials</td>
</tr>
<tr>
<td></td>
<td>Factorial trials</td>
</tr>
<tr>
<td></td>
<td>Cross-over trials</td>
</tr>
<tr>
<td>RCTs exploring different aspects of the interventions they evaluate</td>
<td>Efficacy &amp; effectiveness trials</td>
</tr>
<tr>
<td></td>
<td>Equivalence trials</td>
</tr>
<tr>
<td></td>
<td>Phase III trial</td>
</tr>
<tr>
<td>RCTs by unit of analysis</td>
<td>Body part</td>
</tr>
<tr>
<td></td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td>Group</td>
</tr>
<tr>
<td>RCTs according to the number of participants</td>
<td>Fixed to variable sample size</td>
</tr>
<tr>
<td></td>
<td>N-of-1 trials to mega-trials</td>
</tr>
<tr>
<td>RCTs according to whether investigators know which intervention is</td>
<td>Open trials</td>
</tr>
<tr>
<td>being assessed</td>
<td>Blinded trials</td>
</tr>
<tr>
<td>RCTs that take into account non-randomized individuals &amp; participants’ preferences</td>
<td>Zelen’s design</td>
</tr>
<tr>
<td></td>
<td>Comprehensive cohort design</td>
</tr>
<tr>
<td></td>
<td>Wennberg’s design</td>
</tr>
</tbody>
</table>
The clinical research bridge

The broad range that encompasses the term “clinical research”

### Table C4. 1000 Random Digits

<p>| | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>07048</td>
<td>52841</td>
<td>54969</td>
<td>87057</td>
<td>30570</td>
<td>50494</td>
<td>29936</td>
<td>93967</td>
<td>10641</td>
<td>79871</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09165</td>
<td>56926</td>
<td>17294</td>
<td>03803</td>
<td>31755</td>
<td>11321</td>
<td>33681</td>
<td>12997</td>
<td>17625</td>
<td>25954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35654</td>
<td>69761</td>
<td>83791</td>
<td>63371</td>
<td>28189</td>
<td>19944</td>
<td>04514</td>
<td>56533</td>
<td>89108</td>
<td>27861</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79065</td>
<td>63956</td>
<td>39443</td>
<td>30373</td>
<td>55571</td>
<td>00919</td>
<td>15377</td>
<td>36851</td>
<td>28318</td>
<td>40846</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27969</td>
<td>74368</td>
<td>77782</td>
<td>88616</td>
<td>06368</td>
<td>07345</td>
<td>00725</td>
<td>81221</td>
<td>78417</td>
<td>37992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47528</td>
<td>70548</td>
<td>25078</td>
<td>80729</td>
<td>27806</td>
<td>42877</td>
<td>80287</td>
<td>21759</td>
<td>61980</td>
<td>52447</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65694</td>
<td>95760</td>
<td>64031</td>
<td>24046</td>
<td>77606</td>
<td>91163</td>
<td>51492</td>
<td>20958</td>
<td>18384</td>
<td>49840</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24253</td>
<td>39427</td>
<td>80462</td>
<td>36718</td>
<td>92164</td>
<td>77732</td>
<td>69754</td>
<td>01291</td>
<td>53704</td>
<td>33054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34302</td>
<td>60309</td>
<td>27186</td>
<td>22418</td>
<td>59962</td>
<td>13934</td>
<td>67591</td>
<td>17476</td>
<td>21559</td>
<td>73437</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76809</td>
<td>84341</td>
<td>74012</td>
<td>50947</td>
<td>83214</td>
<td>19967</td>
<td>44219</td>
<td>75929</td>
<td>13182</td>
<td>34858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85183</td>
<td>35958</td>
<td>04301</td>
<td>49628</td>
<td>91493</td>
<td>66103</td>
<td>65699</td>
<td>04241</td>
<td>82441</td>
<td>38112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27541</td>
<td>79187</td>
<td>99777</td>
<td>22894</td>
<td>83283</td>
<td>56218</td>
<td>86183</td>
<td>74497</td>
<td>21070</td>
<td>78935</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74188</td>
<td>09083</td>
<td>54938</td>
<td>79920</td>
<td>27158</td>
<td>24864</td>
<td>31116</td>
<td>33173</td>
<td>43032</td>
<td>52000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13270</td>
<td>57457</td>
<td>30968</td>
<td>65978</td>
<td>67679</td>
<td>91216</td>
<td>47969</td>
<td>39204</td>
<td>46030</td>
<td>93954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89150</td>
<td>53922</td>
<td>40537</td>
<td>23169</td>
<td>46948</td>
<td>05519</td>
<td>72171</td>
<td>85417</td>
<td>31580</td>
<td>98102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49980</td>
<td>44551</td>
<td>99908</td>
<td>46115</td>
<td>92508</td>
<td>77184</td>
<td>44556</td>
<td>69725</td>
<td>42878</td>
<td>60298</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26810</td>
<td>40280</td>
<td>15387</td>
<td>30976</td>
<td>15478</td>
<td>77703</td>
<td>34109</td>
<td>02682</td>
<td>52877</td>
<td>36755</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35056</td>
<td>23942</td>
<td>42645</td>
<td>67063</td>
<td>44118</td>
<td>46433</td>
<td>83172</td>
<td>95689</td>
<td>60923</td>
<td>32769</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09873</td>
<td>65959</td>
<td>77912</td>
<td>70059</td>
<td>07704</td>
<td>16015</td>
<td>57527</td>
<td>09818</td>
<td>84379</td>
<td>35903</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40806</td>
<td>30051</td>
<td>54251</td>
<td>73489</td>
<td>47215</td>
<td>90651</td>
<td>90083</td>
<td>21019</td>
<td>63860</td>
<td>41369</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Random number

Numbers usually have two or more digits

- Select **starting point in the table** (beginning, end, any point of table by a pencil dropped with the eyes closed)
- Select **direction of reading table** (upward - downward)
- Odd numbers: group A – even numbers: group B
  From 01 – 49: group A, from 50 – 99: group B
- Numbers with four digits
  Select position of numbers that determine allocation
  Choose last two digits, or first two, or first & third
Stratified randomization

• First, Identify prognostic factors (or ‘strata’) known to be related to outcome of the study
• Second, Produce separate block randomization lists for different combinations of prognostic factors

It is not practical to stratify on more than one or perhaps two variables
Chemotherapy of breast cancer

• Important prognostic factors:
  Number of metastatic LN: absent, $< 4$, $\geq 4$

• Set of blocks could be generated as follow:
  Breast cancer & no metastatic LN
  Breast cancer & $< 4$ metastatic LN
  Breast cancer & $\geq 4$ metastatic LN

Separate block randomization lists for different combinations of prognostic factors
Minimization Method - 1

3 stratification factors: sex (2), age (3), disease stage (3)

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Female</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 40</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>41 – 60</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td><strong>Disease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage I</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Stage II</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Stage III</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>26</td>
<td>24</td>
</tr>
</tbody>
</table>

50 patients enrolled
the 51st patient is male, age 63, & stage III
Consider lines from the precedent table for that patient's stratification levels only

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
<th>Sign of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>16</td>
<td>14</td>
<td>+</td>
</tr>
<tr>
<td>Age ≥ 60</td>
<td>4</td>
<td>6</td>
<td>−</td>
</tr>
<tr>
<td>Stage III</td>
<td>7</td>
<td>4</td>
<td>+</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>24</td>
<td>2 A, 1 B –</td>
</tr>
</tbody>
</table>
Minimization Method - 3

2 possible criteria

Count only the sign of the difference in each category
Treatment t A is “ahead” in 2 categories out of 3
Assign patient to treatment B

Add the total overall categories (27 As vs 24 Bs)
Treatment A is “ahead,” assign patient to treatment B

Usually agree
Unequal randomization

• Trial comparing a new treatment against a standard one

• Investigator more interested in obtaining information about the new treatment than for the old, where such characteristics are likely to be well known

• **Unbalanced design**
  Allocating larger number of patients to new treatment group
  Power decreases slowly as proportion of new treatment increases

Unequal randomization & power

Reduction in power of a trial as proportion of new treatment increased.

(a) Statistically significant – clinically important
(b) Statistically significant – not clinically important
(c) Not statistically significant – inconclusive
(d) Not statistically significant – true negative
Statistical & clinical significance of CI

(a) Statistically significant, clinically important
(b) Not statistically significant, clinically important
(c) Statistically significant, not clinically important
(d) Not statistically significant, not clinically important
## Main types of biases in RCTs

<table>
<thead>
<tr>
<th>Biases</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>During the planning phase of a RCT</td>
<td>Choice-of-question bias</td>
</tr>
<tr>
<td></td>
<td>Regulation bias</td>
</tr>
<tr>
<td></td>
<td>Wrong design bias</td>
</tr>
<tr>
<td>During the course of a RCT</td>
<td>Selection bias</td>
</tr>
<tr>
<td></td>
<td>Observation bias</td>
</tr>
<tr>
<td></td>
<td>Population choice bias</td>
</tr>
<tr>
<td></td>
<td>Intervention choice bias</td>
</tr>
<tr>
<td></td>
<td>Control group bias</td>
</tr>
<tr>
<td></td>
<td>Outcome choice bias</td>
</tr>
<tr>
<td>During the reporting of a RCT</td>
<td>Withdrawal bias</td>
</tr>
<tr>
<td></td>
<td>Selective reporting bias</td>
</tr>
<tr>
<td></td>
<td>Fraud bias</td>
</tr>
<tr>
<td>During the dissemination of a RCT</td>
<td>Publication bias</td>
</tr>
<tr>
<td></td>
<td>Language bias</td>
</tr>
<tr>
<td></td>
<td>Time lag bias</td>
</tr>
</tbody>
</table>

## Types of RCTs

<table>
<thead>
<tr>
<th>RCTs</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCTs according to how participants are exposed to the interventions</td>
<td>Parallel trials</td>
</tr>
<tr>
<td></td>
<td>Factorial trials</td>
</tr>
<tr>
<td></td>
<td>Cross-over trials</td>
</tr>
<tr>
<td>RCTs exploring different aspects of the interventions they evaluate</td>
<td>Efficacy &amp; effectiveness trials</td>
</tr>
<tr>
<td></td>
<td>Equivalence trials</td>
</tr>
<tr>
<td></td>
<td>Phase III trial</td>
</tr>
<tr>
<td>RCTs by unit of analysis</td>
<td>Body part</td>
</tr>
<tr>
<td></td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td>Group</td>
</tr>
<tr>
<td>RCTs according to the number of participants</td>
<td>Fixed to variable sample size</td>
</tr>
<tr>
<td></td>
<td>N-of-1 trials to mega-trials</td>
</tr>
<tr>
<td>RCTs according to whether investigators know which intervention is</td>
<td>Open trials</td>
</tr>
<tr>
<td>being assessed</td>
<td>Blinded trials</td>
</tr>
<tr>
<td>RCTs that take into account non-randomized individuals &amp; participants’ preferences</td>
<td>Zelen’s design</td>
</tr>
<tr>
<td></td>
<td>Comprehensive cohort design</td>
</tr>
<tr>
<td></td>
<td>Wennberg’s design</td>
</tr>
<tr>
<td>Phase</td>
<td>Objective</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
</tr>
</tbody>
</table>
| I     | Earliest types of studies  
Small numbers of healthy subjects  
Pharmacodynamics, pharmacokinetics & toxicity |
| II    | Carried out in patients  
Find best dose of drug & to investigate safety |
| III   | Major trials aimed at demonstrating efficacy  
Registration of a new product will be based on |
| IV    | Carried out after registration of a product  
Marketing purposes  
Gain broader experience with using the new product |

High quality/relevant data – Pearls

Pearls selected from the rest of lower quality literature

**Ways to reduce bias in studies of therapy**

<table>
<thead>
<tr>
<th>Source of Bias</th>
<th>Strategy to reduce Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Differences at the start of study</strong></td>
<td></td>
</tr>
<tr>
<td>Control &amp; tt group differ in prognosis</td>
<td>Randomization ± stratification</td>
</tr>
<tr>
<td><strong>2 Differences as study proceeds</strong></td>
<td></td>
</tr>
<tr>
<td>Placebo effects</td>
<td>Blinding of patients</td>
</tr>
<tr>
<td>Cointervention</td>
<td>Blinding of caregivers</td>
</tr>
<tr>
<td>Bias in outcome assessment</td>
<td>Blinding of outcome assessors</td>
</tr>
<tr>
<td><strong>3 Differences at completion of study</strong></td>
<td></td>
</tr>
<tr>
<td>Loss to follow-up</td>
<td>Ensure complete follow-up</td>
</tr>
<tr>
<td>Stopping study early (large effect)</td>
<td>Complete study as initially planned</td>
</tr>
<tr>
<td>Patient not receiving assigned tt</td>
<td>Adhere to ITT principle</td>
</tr>
</tbody>
</table>
Some historical examples of treatments with dramatic effects

- Insulin for diabetes
- Blood transfusion for severe hemorrhagic shock
- Defibrillation for ventricular fibrillation
- Neostigmine for myasthenia gravis
- Tracheotomy for tracheal obstruction
- Drainage for pain associated with abscesses
- Pressure or suturing for arresting hemorrhage

Basic Structure of a RCT
Parallel Trial

Most frequently used design

Appraising a RCT (10 questions)

- Did the study ask a clearly focused question?
- Was the study an RCT and was it appropriately so?
- Were participants appropriately allocated to intervention and control groups?
- Were participants, staff, and study personnel blind to participants’ study groups?
- Were all the participants who entered the trial accounted for at its conclusion?
- Were participants in all groups followed up and data collected in the same way?
- Did the study have enough participants to minimise the play of chance?
- How are the results presented and what are the main results?
- How precise are the results?
- Were all important outcomes considered and can the results be applied to your local population?
First RCT in the United States

1951

NIH started a study of adrenocorticotropic hormone (ACTH), cortisone & aspirin in the treatment of rheumatic heart disease*

Ethical principles of research

- All research should be approved by an **ethics committee**
- Study will justify any risk or inconvenience to the subjects
- Researchers are informed of study purpose & must have training to conduct the study with high degree of scientific integrity
- Subjects must be free to withdraw consent at any time & withdrawal must not influence their future treatment
- Subjects must be provided with information on purpose, demands of the protocol prior to their given **informed consent**

Nuremberg Code (1946 – 1947)
Declaration of Helsinki (World Medical Association 1964 → 2002)
Trials in the next 50 years

Much simpler & much larger

- **Large simple RCT**
  Moderate but worthwhile benefits will appear
  Randomize many thousands in breast & intestinal cancer
  Randomize tens of thousands in stroke & heart disease

- Design trials that are extremely simple & flexible
  Simplify entry criteria by use of uncertainty principle
  Simplify treatments
  Simplify enormously data requirements

The Uncertainty Principle

- A patient can be **entered** if, and only if, the responsible clinician is substantially uncertain which of the trial treatments would be most appropriate for that patient.

- A patient **should not be entered** if responsible clinician or patient are, for any medical or non-medical reasons, reasonably certain that one of treatments that might be allocated would be inappropriate for this patient.
Why a RCT?

• Main purpose is to prevent selection bias by distributing characteristics of patients that may influence the outcome randomly between the groups, so that any difference in outcome can be explained only by treatment.

• Thus, there will be balancing of baseline differences between intervention groups that may affect outcome such as:
  - Age
  - Sex
  - Disease activity
  - Duration of disease
Patients not adhered to allocated management

- **Per protocol analysis**
  Excluding participants from analysis
  Those who adhere tend to do better than who do not
  Destroys comparison afforded by randomization

- **Intention-to treat bias**
  If effective treatment & substantial nonadherence
  Underestimates magnitude of treatment effect
  Using protocol ensuring maximal adherence
  **Run-in periods**: exclude nonadherents before R
Blinding

Sometimes called masking

- **Single blind**  Only patients or only investigators are ignorant of assigned treatment
- **Double blind** Patients & investigators are ignorant of assigned treatment
- **Triple blind** Patients, investigators & data evaluators are ignorant of assigned treatment
### History of Streptomycin – 1

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 1943</td>
<td>Isolated by Albert Schatz – PhD student</td>
</tr>
<tr>
<td></td>
<td>Pr Waksman – Rutgers University -NJ</td>
</tr>
<tr>
<td></td>
<td>Developed by the American firm Merck</td>
</tr>
<tr>
<td>1945</td>
<td>Feldman showed effect on TB in guinea pigs</td>
</tr>
<tr>
<td></td>
<td>Merck invested $3.5m in new plant</td>
</tr>
<tr>
<td></td>
<td>10 other firms tried to produce the drug</td>
</tr>
<tr>
<td>July 1946</td>
<td>Feldman visited Britain at instigation of MRC</td>
</tr>
<tr>
<td></td>
<td>Persuasive presentations in Oxford &amp; London</td>
</tr>
<tr>
<td></td>
<td>Ministry of Supply asked MRC to plan CT</td>
</tr>
</tbody>
</table>

### History of Streptomycin – 2

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct 1946</td>
<td>Creation of SPM Clinical Trials Committee&lt;br&gt;Marshall (chairman), Philip Hart (secretary)&lt;br&gt;Bradford Hill (Statistician-Random allocation)</td>
</tr>
<tr>
<td>Nov 1946</td>
<td>50 kg to British government at $ 320,000&lt;br&gt;Only hope to obtain SPM through MRC&lt;br&gt;BBC broadcast many emergency appeals&lt;br&gt;Black market emerged</td>
</tr>
<tr>
<td>1948</td>
<td>BMJ report&lt;br&gt;Pains to defend use of untreated control group</td>
</tr>
</tbody>
</table>

Why is blinding/masking so important

• Vitamin C trial for prevention & treatment of common cold
• Conducted among employees at NIH
• Many of enrollees could not resist temptation to analyze the content of their blinded study medications
• Among participants who did not break the blind, mean duration of colds was similar in the two groups
• Among participants who knew they were taking vit C reported shorter cold durations than those who knew they took placebo

Treatment Allocation by Minimization

Different principle from randomization

• First described by Taves in 1974*

• First participant is allocated at random

For each subsequent participant, we determine which treatment lead to better balance between groups

• Ensure excellent balance between groups for several prognostic factors even in small samples

• Possible by hand or software (minim**, free program)

** http://www-users.york.ac.uk/zmb55/guide/minim.htm
Sources of Bias in RCTs

- External validity
  - 1-Selection
  - Exclusions
- Internal validity
  - 2-Randomisation
  - 3-Follow up
  - Co-intervention
  - Loss to follow up
  - Cross over
  - Contamination
  - 4-Outcomes
  - 5-Results
Wrong or Unreliable Therapeutic Answers

Wrong therapeutic answers are generated by:

- Nonrandomized “outcomes research”
- Small randomized studies
- Small meta-analyses
- Statistically inappropriate analyses
- Large scale randomized evidence

Selective emphasis on particular trials or subgroups
Patients not adhered to allocated management

- **Per protocol analysis**
  Excluding participants from analysis
  Those who adhere tend to do better than who do not
  Destroys comparison afforded by randomization

- **Intention-to treat bias**
  If effective treatment & substantial nonadherence
  Underestimates magnitude of treatment effect
  Using protocol ensuring maximal adherence
  **Run-in periods**: exclude nonadherents before R
Blinding

Sometimes called masking

- **Single blind**  Participants don’t know details of treatment. Researchers do.

- **Double blind**  Both participants & data collectors are ignorant of assigned treatment.

- **Triple blind**  Participants, data collectors, & data evaluators are all blinded.
Production of streptomycin was technically difficult

Porter RW. Chemical Engineering 1946 (Oct).
History of streptomycin

- **Nov 1943**
  Developed by American firm Merck

- **1945**
  10 other firms tried to produce SPM

- **1946**
  50 kg to British government at $ 320,000
  Only hope to obtain SPM through MRC
  BBC broadcast many emergency appeals
  Black market emerged

Porter RW. Chemical Engineering 1946 (Oct).
Treatment Allocation by Minimization

Different principle from randomization

• First described by Taves in 1974*

• First participant is allocated at random
  For each subsequent participant, we determine which treatment lead to better balance between groups

• Ensure excellent balance between groups for several prognostic factors even in small samples

• Possible by hand or software (minim**, free program)

** http://www-users.york.ac.uk/zmb55/guide/minim.htm
Patients not adhered to allocated management

- **Per protocol analysis**
  Excluding participants from analysis
  Those who adhere tend to do better than who do not
  Destroys comparison afforded by randomization

- **Intention-to treat bias**
  If effective treatment & substantial nonadherence
  Underestimates magnitude of treatment effect
  Using protocol ensuring maximal adherence
  **Run-in periods**: exclude nonadherents before randomization
Basic Structure of a RCT

Parallel Trial

Each group exposed only to one of study interventions
Most frequently used design

# Ways to reduce bias in studies of therapy

## Differences at the start of trial

| Difference in prognostic factors | Randomization & stratification |

## Differences as trial proceeds

<table>
<thead>
<tr>
<th>Placebo effects</th>
<th>→</th>
<th>Blinding of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cointervention</td>
<td>→</td>
<td>Blinding of caregivers</td>
</tr>
<tr>
<td>Bias in outcome assessment</td>
<td>→</td>
<td>Blinding of outcome assessors</td>
</tr>
</tbody>
</table>

## Differences at end of the trial

<table>
<thead>
<tr>
<th>Loss to follow-up</th>
<th>→</th>
<th>Ensure complete follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stopping study early</td>
<td>→</td>
<td>Complete study as planned</td>
</tr>
<tr>
<td>Pts not receiving assigned tt</td>
<td>→</td>
<td>ITT principle</td>
</tr>
</tbody>
</table>
Original EBP model

- Practitioner's individual expertise
- Client values and expectations
- Best evidence

EBP
Newer EBP model

Basic elements of clinical decision making

BMC Health Services Research 2002, 2:3
Statistical & clinical significance in CI

(a) Statistically significant – clinically important
(b) Statistically significant – not clinically important
(c) Not statistically significant – inconclusive
(d) Not statistically significant – true negative

The 3 EBP components

Minimal overlap with clinical expertise

Clinician’s guide to evidence-based practices.
The 3 EBP components

- Best Available Research
- Patient Characteristics, Culture, & Preferences
- Clinical Expertise

Minimal overlap with patient preferences & culture

Clinician’s guide to evidence-based practices.
The 3 EBP components

Best Available Research

Patient Characteristics, Culture, & Preferences

Clinical Expertise

Minimal overlap with available research

Clinician’s guide to evidence-based practices.
Sources of medical knowledge

Chin R, Lee BY. Principal & practice of clinical trial medicine. AP
R & D: Research & Development Committee

Development & approval of clinical trials

Research question

Trial development

- Secure a sponsor
- Secure funding
- Protocol development
- Peer review
- Register with EU

Finalized Protocol

Trial approval

- R&D submission
- Ethics submission
- CTA submission

Trial Begins

The 3 EBP components

Major convergence between the 3 components

Rubin A. Practitioner’s guide to using research for EB practice.
John Wiley & Sons, 2007
Clinical epidemiology

Evidence-based medicine

Evidence-based practice

The 3 EBP components

Major convergence between the 3 components

Clinician’s guide to evidence-based practices.
The 3 EBP components

Major convergence between the 3 components

Rubin A. Practitioner’s guide to using research for EB practice.
John Wiley & Sons, 2007
Trial design

- Systematic review
- Meta-analysis
- Randomized controlled trial
- Cohort study
- Case control study
- Cross-sectional study
- Case series & case report

Based on RCTs
Number of randomized trials published*

* Based on Medline search restricted to “Randomized clinical trials”
Annual addition of articles to PubMed

50 years ago: majority of articles published in non-English
Currently: 90% of articles published in English

Sealed opaque envelope
Hierarchy of evidence in quantitative studies

What is bias?

• Tendency of an estimate to deviate in one direction from a true value (underestimation or overestimation)

• More commonly unintentional, & often unrecognized even by researchers themselves

High quality/relevant data – Pearls

Finding high-quality evidence is like searching for ‘rare pearls’

Study types

Observational
- Descriptive
  - Cross-sectional
  - Case series
  - Case report
- Analytic
  - Cohort
  - Case-control

Interventional or experimental
- RCT
Randomization in RCTs

1- Selection

2- Randomization

3- Follow-up

4- Outcomes

5- Analysis

Sir Austin Bradford Hill

- Studied medicine when World War 1 intervened
- Pilot in the World War 1
- Contracted TB: 2 years hospital -2 years convalescence
- Took a degree of Economics by correspondence
- **1922** Attended statistical lectures by Karl Pearson
- **1933** Reader in Epidemiology & Vital Statistics
- **1947** Professor of Medical Statistics
- **1950-52** President of the Royal Statistical Society
High quality/relevant data
Pearls

If it is not valid → It is of no value

If it is not relevant → It is of no value
• If the study wasn’t randomized, we’d suggest that you stop reading it and go on to the next article in your search.
Internal & external validity of a RCT

ISIS-2 trial
Streptokinase &/or aspirin on AMI mortality

Post-hoc analysis

Zodiac signs of Gemini & Libra
5% higher mortality on aspirin compared to placebo

Other Zodiac signs
30% lower mortality on aspirin compared to placebo

Steps of EBM

1 Ask
Steps of EBM

2 Acquire
Steps of EBM

3 Appraise
Critical appraisal of a RCT

Steps of EBM

4 Apply
“All that glisters is not gold”

W. Shakespeare

In “The Merchant of Venice”

Flow chart of evidence based practice

This so-called Hawthorne effect refers to the tendency of people to alter their behavior when they are subject to special attention in a research setting.
Sir Austin Bradford Hill

- Studied medicine when World War 1 intervened
- Pilot in the World War 1
- Contracted TB: 2 years hospital - 2 years convalescence
- Took a degree of economics by correspondence
- 1922 Attended statistical lectures by Karl Pearson
- 1933 Reader in epidemiology & vital statistics
- 1947 Professor of medical statistics
- 1950-52 President of the Royal Statistical Society
Randomization

- Simple randomization
- Random table
- Block randomization
- Stratified randomization
- Minimization method
- Unequal randomization
- Allocation concealment
"EBM is the integration of best research evidence with clinical expertise & patient values”

- David Sackett
Study types

- Observational
  - Cohort study
  - Case-control study
  - Cross-sectional study
  - Case series & case report

- Interventional or experimental
  - RCT
Trial designs

- Systematic review
- Meta-analysis
- Randomized clinical trial
- Cohort study
- Case control study
- Cross-sectional study
- Case series & case report

Secondary research

Primary research
History of randomization

Sir Austin Bradford Hill

- Desirability to use randomization in clinical medicine when he published articles on medical statistics in 1937
- He didn’t recommend randomization of individuals, because he might scare doctors off any use of concurrent controls
- In 1946, when he judged the time was right, he recommended randomization of individual patients & this rapidly gained acceptance among medical scientists
McMaster PLUS project – First level

Critical appraisal filters

1. Valid
2. Ready for clinical attention

50,000 articles/yr from 120 journals

~3,000 articles/yr meet critical appraisal & content criteria (94% noise reduction)

Health Information Research Unit – McMaster University – Canada
Evidence alone does not decide
Combine with other knowledge & values
RCTs as the subject of research

• Important research efforts have used RCTs as the subject rather than the tool of research

• These studies aim to improve the design, reporting, dissemination, & the use of RCTs in health care

Sample size formula for binary outcomes

(\(\alpha = 0.05, \beta = 0.10\), equal number in each group)

\[
N = \frac{10.51 \left[(R + 1) - p_2 \left(R^2 + 1\right)\right]}{p_2 \left(1 - R\right)^2}
\]

<table>
<thead>
<tr>
<th>N</th>
<th>Sample size in each of the groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>Event rate in treatment group (not in formula)</td>
</tr>
<tr>
<td>p2</td>
<td>Event rate in control group</td>
</tr>
<tr>
<td>R</td>
<td>Risk ratio (p1/p2)</td>
</tr>
</tbody>
</table>

If

- \(p_1 = 6\%\)
- \(p_2 = 10\%\)
- \(R = 6\% / 10\% = 0.60\)

\(N = 962\)
Variable in the sample size formula

<table>
<thead>
<tr>
<th>$\alpha$ (Type I error)</th>
<th>Power $(1 - \beta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>0.05</td>
<td>7.58</td>
</tr>
<tr>
<td>0.01</td>
<td>11.68</td>
</tr>
</tbody>
</table>

Being a statistician means never having to say you are certain

Anon

High quality/relevant data
Pearls

If not valid → No value
If not relevant → No value
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1922</td>
<td>Attended statistical lectures by Karl Pearson</td>
</tr>
<tr>
<td>1933</td>
<td>Reader in epidemiology &amp; vital statistics</td>
</tr>
<tr>
<td>1947</td>
<td>Professor of medical statistics</td>
</tr>
<tr>
<td>1950-52</td>
<td>President of the Royal Statistical Society</td>
</tr>
</tbody>
</table>
First properly RCTs

<table>
<thead>
<tr>
<th></th>
<th>Immunisation against whooping cough *</th>
<th>Streptomyacin for pulmonary TB **</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors</strong></td>
<td>MRC</td>
<td>MRC (D’arcy Hart)</td>
</tr>
<tr>
<td><strong>Statistician</strong></td>
<td>Bradford Hill</td>
<td></td>
</tr>
<tr>
<td><strong>Started</strong></td>
<td>Months before Nov 1946</td>
<td>Nov 1946</td>
</tr>
<tr>
<td><strong>Reported</strong></td>
<td>1951</td>
<td>Oct 1948</td>
</tr>
<tr>
<td><strong>Journal</strong></td>
<td>BMJ</td>
<td>BMJ</td>
</tr>
</tbody>
</table>


Hawthorne effect

- Employees of Hawthorne Works of Western Electric Company in Chicago participated in a study to evaluate effect of light levels on work performance.
- Surprisingly, work performance increased, regardless of whether level of light at workplace was increased, kept constant, or decreased.
- Special attention given to workers participated in the study explains improvement in overall performance.