Lecture 7 – The RCT

Mathew J. Reeves BVSc, PhD
Associate Professor, Epidemiology

Objectives

• Understand the central role of randomization, concealment and blinding in RCT
• Understand the major steps in conducting a RCT
• Understand the importance of loss-to-follow-up, non-compliance, and cross-overs
• Understand the reasons for the ITT analysis and why to avoid the PP and AT approaches
• Understand the strengths and weaknesses of RCT's

Experimental (Intervention) Studies

• Investigator completely controls exposure
  – type, amount, duration, and who receives it (randomization)
• Regarded as the most scientifically-vigorous study design. Why?
  – Random assignment reduces confounding bias
  – Concealment reduces selection bias
  – Blinding reduces biased measurement
• Can confidently attribute cause and effect due to the high internal validity of trials
• Trials are not always feasible, appropriate, or ethical
Types of Intervention Studies

- All trials test the efficacy of an intervention and assess safety

- Prophylactic vs Treatment
  - Evaluate efficacy of intervention designed to prevent
disease, e.g., vaccine, vitamin supplement, patient education
  - Evaluate efficacy of controlling drug or intervention or a drug
designed to manage signs and symptoms of a disease (e.g.,
arthritis, hypertension)

- RCT vs Community Trials
  - Individuals, tightly controlled, narrowly focused, highly
selected groups, short or long duration
  - Groups, less rigidly controlled, long duration, usually
primary prevention

---

RCT's – Overview of the Process

- 1. Inclusion Criteria
- 2. Exclusion Criteria
- 3. Baseline Measurements
- 4. Randomization and concealment
- 5. Intervention
- 6. Blinding
- 7. Follow-up (FU) and Compliance
- 8. Measuring Outcomes
- 9. Statistical Analyses

---

RCT: Basic Design

![Diagram of RCT: Basic Design]

- Eligible Sample
- Control (Placebo)
- Treatment
- Population of Interest
- Conceal, Randomize, Blind
Eligibility Criteria

- Explicit inclusion and exclusion criteria
  - Provide guidance for interpretation and generalization of study results
  - Balance between generalizability (external validity) and efficiency

- Criteria should:
  - Capture patients who have potential to benefit
  - Exclude patients that may be harmed, are not likely to benefit, or for whom it is not likely that the outcome variable can be assessed

1. Inclusion Criteria

- Goals – to optimize the following:
  - Rate of primary outcome
  - Expected efficacy of treatment
  - Generalizability of the results
  - Recruitment, follow-up (FU) and compliance

- Identify population in whom intervention is feasible and will produce desired outcome

- ...pick people most likely to benefit without sacrificing generalizability

2. Exclusion Criteria

- Goal: to identify subjects who would "mess up" the study

- Valid reasons for exclusion
  - Unacceptable risk of treatment (or placebo)
  - Treatment unlikely to be effective (not at risk)
    - Disease too severe, too mild, already on meds
  - Unlikely to complete FU or adhere to protocol
  - Other practical reasons e.g., language/cognitive barriers

- Avoid excessive exclusions
  - Decrease recruitment, increased complexity and costs
  - Decreased generalizability (external validity)
3. Baseline Measurements

- What to measure?
  - Tracking info
    - Names, address, phone #, e-mail, SSN
    - Contact info of friends, neighbors, and family
  - Demographics – describe your population
  - Medical History
  - Major prognostic factors for primary outcome
    - Used for subgroup analyses e.g., age, gender, severity

4. Randomization and Concealment

- Results in balance of known and unknown confounders; eliminates bias in Tx assignment, and provides basis for statistical inference

- Randomization process should be described to ensure it is reproducible, unpredictable, and tamper proof

- Simple randomization e.g., coin flip
  - Can result in unequal numbers within treatment groups, especially in small trials

- Blocked randomization
  - Used to ensure equal numbers in each group, randomize within blocks of 4-8

- Stratified blocked randomization
  - Randomize within major subgroups e.g., gender, disease severity

Concealment

- Different from randomization per se

- Unpredictability prevents selection bias
  - Assignment of the next subject should be unknown and unpredictable

- Concealed studies
  - Sealed opaque envelopes, site-specific randomization center

- Unconcealed studies
  - Alternative day assignment, Date of birth

- Conceptually, it is best to view concealment as different from blinding (the latter prevents measurement bias)
What is confounding?

- Bias = systematic error (distortion of the truth)
- Three broad classes of bias:
  - Selection bias
  - Confounding bias
  - Measurement bias
- Confounding:
  - Define: a factor that distorts the true relationship of the study variable of interest by virtue of being related to both the outcome of interest and the study variable.

Example of Confounding:

MASCOTS Stroke Registry – Pre-existing statin use and in-hospital death in women

<table>
<thead>
<tr>
<th>Crude OR= 0.59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statins</td>
</tr>
<tr>
<td>In-hospital death</td>
</tr>
</tbody>
</table>

In this observational study, the risk of death in women hospitalized for acute stroke was 41% lower compared to women not on statins. However, statin users were younger and had fewer comorbidities—both of which are independent risk factors for in-hospital mortality.

<table>
<thead>
<tr>
<th>Adjusted OR= 1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statins</td>
</tr>
<tr>
<td>In-hospital death</td>
</tr>
</tbody>
</table>

+++ Age, comorbidities

5. Intervention

- Balance between efficacy & safety
  - Everyone is exposed to potential side effects but only the few who develop dys&outcome can benefit
  - Hence, usually use "lowest effective dose"
  - Most trials are under-powered to detect side effects—hence Phase IV trials and post-marketing surveillance
- Control group:
  - Placebo or standard treatment
  - Risk of contamination (getting the treatment somewhere else)
6. Blinding

- As important as randomization because it prevents
  - Biased assessment of outcomes post-randomization (i.e., 
    misinterpreted or aspirational bias)
- Blinding also helps reduces non-compliance and contamination
- Blinding is particularly important if have "soft" outcomes
  - e.g., self-report, investigator opinion
- Sometimes blinding is not feasible (i.e., Open trial), if so,
  - Choose a "hard" outcome
  - Standardize treatments as much as possible
- Blinding may be hard to maintain e.g., when obvious side effects are associated with a drug

Single vs. double vs. triple blind?

- Much confusion in use of the terms single, double, and triple blind, hence the study should describe exactly who was blinded.
- Ideally, blinding should occur at all of the following levels:
  - Patients
  - Caregivers
  - Data collectors
  - Adjudicators of outcomes
  - Statisticians

Hakim Z. Bawot/PhD
O Dept. of Epidemiology, 555
7. Loss to follow-up, non-compliance, and contamination

- Needs to be carefully monitored and documented

- If loss-to-follow-up, non-compliance, and contamination are frequent and do not occur at random then results in:
  - major bias
  - decreased power
  - and loss of credibility

- Why lost to follow-up?
  - side effects, moved, died, recovered, got worse, lost interest

7. Loss to follow-up, non-compliance and contamination

- Why poor compliance?
  - side effects, toxic reactions, recovered, got worse, lost interest

- Contamination = cross-overs (esp. in control group if unblinded)

- Maximize FU and compliance by using
  - two screening visits prior to enrollment
  - pre-randomization run-in period using placebo or active drug
  - maintain blinding

Loss to FU, poor compliance, contamination = slow death of a trial

![Diagram of trial flow with contamination and loss to FU]
8. Measuring Outcomes
- Best = Hard clinically relevant endpoints
- e.g., disease rates, death, recovery, complications
- Must be measured with accuracy and precision
- Must be monitored equally in both groups
- Surrogate endpoints
  - used in short-term clinical trials or to reduce size/length of follow-up
  - Must be biologically plausible
  - Association between surrogate and a hard endpoint must have been previously documented
  - e.g., reduced BP in lieu of stroke incidence
- Best to use blinded adjudication, especially for soft outcomes

9. Statistical Analyses
- On the surface analysis is relatively simple because of RCT design:
  - Measure clinical effect with 95% CI using
    - RR, RRR, ARR, NNT
  - For time-dependent outcomes use Kaplan-Meier curves, and/or Cox regression modeling
  - Test for statistical significance (p-value):
    - Categorical outcome = Chi-square test
    - Continuous outcome = t-test
    - Or use non-parametric methods where necessary

RCT's – Basic Analysis
Dichotomous (Disease Yes/No) Outcome

<table>
<thead>
<tr>
<th>Randomize</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp +</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>n1 = a + b</td>
</tr>
<tr>
<td>Exp -</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>d</td>
</tr>
<tr>
<td></td>
<td>n2 = c + d</td>
</tr>
<tr>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

RR = R_a/R_b and ARR = R_c - R_d
Measures of Effect used in RCT's

<table>
<thead>
<tr>
<th></th>
<th>EER = 30/100 = 30%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CER = 40/100 = 40%</td>
</tr>
</tbody>
</table>

EER = Experimental vs. control arms
CER = Control vs. control arms
RR = EER/CER = 30/40 = 75%
ARR = 1 - RR = 25%
NNT = EER - CER = 10 / 10 = 10

9. Statistical Analyses - ITT

- **Intention-to-Treat Analysis**
  - Gold Standard
  - Compare outcomes based on original randomization scheme regardless of missing, non-compliance, dropouts, and lost to follow-up

- **Per Protocol (PP) Analysis**
  - Compare outcomes based on actual treatment received among those who were compliant (analysis drops non-compliance)
  - Ask whether the treatment works among only those that comply

- **As Treated (AT) Analysis**
  - Compare outcomes based on actual treatment received regardless of original assignment
  - Equivalent to analyzing the data as a cohort study
  - Ask whether the treatment works among those that took it

Both PP and AT approaches ignore original randomization and are therefore subject to bias!

ITT vs. Per-Protocol vs. As Treated

- **ITT**
  - All who are randomized, drop out, and are treated
- **Per-Protocol**
  - All who are randomized, drop out, but are treated
- **As Treated**
  - All who are randomized, but not necessarily treated

<table>
<thead>
<tr>
<th></th>
<th>ITT</th>
<th>Per-Protocol</th>
<th>As Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Treatment</td>
<td>YES</td>
<td>DROP</td>
<td>NO</td>
</tr>
<tr>
<td>Dropout</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

* Matthew J. Sydney, MD
  Univ. of Washington, Seattle, WA*

<table>
<thead>
<tr>
<th>Allocated (vs. actual) treatment</th>
<th>Medical (medial)</th>
<th>Medical (surgical)</th>
<th>Surgical (medical)</th>
<th>Surgical (surgical)</th>
<th>ARR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>323</td>
<td>10</td>
<td>308</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Deaths</td>
<td>27</td>
<td>2</td>
<td>15</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Morbidity</td>
<td>8.4% (20/240)</td>
<td>4.0% (1/25)</td>
<td>4.1% (12/294)</td>
<td>33.1% (24/72)</td>
<td></td>
</tr>
<tr>
<td>ITT analysis</td>
<td>7.9% (20/250)</td>
<td>6.2% (16/262)</td>
<td>2.4% (-1.6, 6.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP analysis</td>
<td>8.4% (23/272)</td>
<td>4.1% (13/320)</td>
<td>4.3% (0.7, 8.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT analysis</td>
<td>8.5% (24/280)</td>
<td>4.1% (12/296)</td>
<td>5.4% (1.8, 9.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The problem of the lost-to-follow up?

- LTFU is a common problem that is frequently ignored in the analysis of published RCTs.
- All subjects, LTFU should be included in the ITT analysis, but the problem is that we don’t know their final outcome.
- An ITT analysis done in the face of LTFU is a de facto PP analysis and is therefore biased.
- Some studies impute an outcome measure. Example:
  - Carry forward baseline or worst or last observation
  - Multiple imputation, use a model to predict the outcome
- Better line: Minimal LTFU as much as possible – requires that you follow-up with subjects who were non-compliant, ineligibles, or those dropped out for any reason.

9. Statistical Analyses

- **Subgroup Analysis**
  - Analysis of the primary outcome within sub-groups defined by age, gender, race, disease severity, or any other prognostic variable.
  - Can provide critical information on which sub-groups a treatment works in and which groups it does not.
  - e.g. Low dose ASA was effective in preventing AMI in men but not women.
- Potentially misleading analyses are not pre-planned.
  - Sub-group analysis might be conducted because the primary analysis was non-significant leading to a large number of secondary analyses.
  - This results in the problem of multiple comparisons and increased type I error rates.
10. Summary of RCTs

- **Advantages**
  - High internal validity
  - Able to control selection, confounding and measurement biases
  - True measure of treatment efficacy (cause and effect)

- **Disadvantages**
  - Low external validity (generalizability)
  - Strict enrollment criteria creates a unique, highly selected study population
  - Complicated, expensive, time consuming
  - Ethical and practical limitations