Clinical Trials – Study population

Faiz Ahmad Khan, MD MPH
Respiratory Epidemiology and Clinical Research Unit
Montreal Chest Institute

July 17, 2014
Overview

Study population –
  Eligibility
  Generalizability and recruitment
  Sample size calculations
    Power and clinically meaningful differences

Randomization

Ways to avoid bias
  Blinded vs unblinded studies
    Single/double/triple blinding
  Minimizing LTFU (Withdrawal and drop-outs)

Data collection
Study population

Fundamental point

• The study population should be defined in advance, using unambiguous inclusion (eligibility) criteria. The impact that these criteria will have on ability to generalize, and participant recruitment must be considered.
Study population

Population at large

Population without condition → Population with condition

- With condition but ineligible
- Eligible but not enrolled

Study population

Study sample

Definition of condition → Entry criteria → Enrollment
Aims of Eligibility criteria

Minimize risk / enhance participant safety
Select subjects most likely to benefit from the intervention.
Ineligible patients

• Any trial requires precise definition of eligibility criteria

• Ineligible patients have to be reported

• If the proportion of ineligible patients is too large
  • may reflect poor study organization;
  • or, eligibility criteria were too restrictive.

• Ineligible patients do not affect internal validity.
  • BUT, affect generalizability of results
Internal vs External validity

• Internal validity

  • Judged by whether effect estimated in the trial could differ from the true effect because of systematic (non-random) error
    • selection bias, information bias, confounding
    • Random error is permitted

• External validity or “generalizability”

  • Judged by whether results are valid for patients with the condition that are not in the study population, but treated in similar settings

Generalizability (to all patients with the condition)

- Study subjects are NOT randomly chosen from the study population, because of the restrictions imposed by the eligibility criteria.
- This creates the risk that the findings may not be generalizable.
Generalizability – impact of refusals

Participants must agree to enroll in a study:

Why do some agree to participate while others do not?

How does that affect generalizability? If proportion is low? Or high?
Study population: 6 vs 9 months of TB treatment in PLWH


– Population with condition: PLWH and active TB
– Assessed for eligibility: HIV-infected patients with symptoms and signs suggestive of TB receiving care at TRC clinics in Chennai and Madurai, aged 15 years or above, not moribund and not pregnant
– Entry criteria: PLWH and active TB (positive smear or abnormal CXR not improving with 14d abx), age >14, not moribund, not pregnant, meeting sociological eligibility criteria, Hb>7g/L, WBC >1.1, PLT > 100, ALT<2.5ULN, CR<1.1, gluc<140mg/dL
– Study population: PLWH and active TB meeting the above criteria and receiving care at the clinics where the study was conducted
– Study sample: 334 patients randomized
Study population: 6 vs 9 months of TB treatment in PLWH

Assessed for eligibility: 857

- HIV negative 67
- TB negative 60
- Not fulfilling entry criteria 125
- Absent for assessment 60
- Unwilling to attend as required 183
- Others 28

Randomized: 334

730 patients with condition
25.3% (185) ineligible/could not be assessed for eligibility
211/545 (38.7%) of study population eligible but not enrolled
334/545 (61.2%) of study population were enrolled
Study population: 6 vs 9 months of TB treatment in PLWH

730 patients with condition

- 25.3% (185) ineligible/could not be assessed for eligibility
- 211/545 (38.7%) of study population eligible but not enrolled
- 334/545 (61.2%) of study population were enrolled

Is the study sample representative of the study population?

It would be easier to assess the external validity if we could compare characteristics of eligible-enrolled vs eligible-not enrolled → this information is rarely provided!
Double-blind RCT: high dose INH vs standard INH vs placebo for MDR-TB

• “We recruited consecutive, sputum culture-positive, non-HIV infected patients previously diagnosed with pulmonary TB who developed documented MDR-TB and who reported to the study centre during the time of the study.”

• Excluded: if unwilling to give consent, abnormal renal or hepatic profile, history suggestive of INH hypersensitivity, or were pregnant or lactating were excluded.”

• 134 participants were enrolled.

Katiyar SK et al. IJTL D 2008
Double-blind RCT: high dose INH vs standard INH vs placebo for MDR-TB

- High-dose INH associated with more rapid sputum conversion (HR 2.38, 1.45-3.91), and higher likelihood of being sputum-negative at 6 months (RR 2.37, 1.4-3.84).
- Results internally valid
- No information on number assessed for eligibility, nor number eligible but not enrolled.
- But can they be generalized?

Katiyar SK et al. *IJTLD* 2008
Sample size

Fundamental point

- Clinical trials should have sufficient **statistical power** to detect statistically significant **meaningful** differences between groups.
  - *meaningful* = clinically important/relevant

- Calculation of sample size is an essential part of planning a trial.
Steps in Determining Sample Size

1. Define the response with standard therapy.
   eg. RCT 6 vs 9 months: assumed rate of unfavourable outcomes with 6 months = 20%

2. Decide on a clinically meaningful difference
   - How much is enough to say the new treatment is better (worth it)?
   eg. A reduction from 20% to 10%.
Steps in Determining Sample Size

3. How much **power** do you want
   - Power equals ability to detect a statistically significant difference, if some difference truly exists.
     - often 80% power is used
   - Low power increases risk of falsely concluding no difference, when there IS a difference. (Type 2 error)
     - often a 20% probability of Type 2 error is considered acceptable

4. Costs and Feasibility
   - How much money do you have?
Sample size
Web calculators

Web pages for sample size calculators
http://statpages.org/
http://www.surveysystem.com/sscalc.htm
http://stat.ubc.ca/~rollin/stats/ssize/b2.html

PS Power – can download and use
http://www.mc.vanderbilt.edu/prevmed/ps/index.htm
contains up to date information about the program PS Power. You can download the latest version from there.

[[ Dupont WD and Plummer WD: PS power and sample size program available for free on the Internet. Controlled Clin Trials,1997;18:274 ]]}
Inference for Proportions: Comparing Two Independent Samples

(To use this page, your browser must recognize JavaScript.)

Choose which calculation you desire, enter the relevant population values (as decimal fractions) for \( p_1 \) (proportion in population 1) and \( p_2 \) (proportion in population 2) and, if calculating power, a sample size (assumed the same for each sample). You may also modify alpha and the power, if relevant. After making your entries, hit the calculate button at the bottom.

- Calculate Sample Size (for specified Power)
- Calculate Power (for specified Sample Size)

Enter a value for \( p_1 \):

Enter a value for \( p_2 \):

- 1 Sided Test
- 2 Sided Test

Enter a value for alpha (default is .05):

Enter a value for desired power (default is .80):

The sample size (for each sample) is:

Calculate

http://www.stat.ubc.ca/~rollin/stats/ssize/b2.html
## Estimated sample size required

Estimated sample size required for different cure percentages and detection powers.

(Using [http://stat.ubc.ca/~rollin/stats/ssize/b2.html](http://stat.ubc.ca/~rollin/stats/ssize/b2.html))

<table>
<thead>
<tr>
<th>% Cure expected</th>
<th>Number per group required to detect a difference with power of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx A</td>
<td>Tx B</td>
</tr>
<tr>
<td>65%</td>
<td>90%</td>
</tr>
<tr>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>75%</td>
<td>90%</td>
</tr>
<tr>
<td>65%</td>
<td>85%</td>
</tr>
</tbody>
</table>
## Statistical power: additional considerations

<table>
<thead>
<tr>
<th>Estimated pre-trial</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Control response</td>
<td>60%</td>
</tr>
<tr>
<td>Intervention response</td>
<td>80%</td>
</tr>
<tr>
<td>Expected difference</td>
<td>20%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Total number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>No LTFU &amp; adequ. adh.</td>
<td>182</td>
</tr>
<tr>
<td>10% failure to adhere</td>
<td>226</td>
</tr>
<tr>
<td>10% loss to follow-up</td>
<td>204</td>
</tr>
<tr>
<td>10% LTFU and 10% failure to adhere</td>
<td>250</td>
</tr>
</tbody>
</table>
Azinroimodatn

Tonmaioazndri

Adormnztniaoz... 

Randomization

http://textmechanic.com/Word-Scrambler.html
Randomization in Experimental Studies

Fundamental point

Randomization tends to:

• produce study groups comparable with respect to known and unknown risk factors
• remove investigator bias in the allocation of subjects
• guarantee that statistical tests will have valid significance levels
Randomization: allocation concealment

- **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. Allocation concealment prevents researchers from (unconsciously or otherwise) influencing which participants are assigned to a given intervention group.”

http://www.consort-statement.org/resources/glossary
Randomization: Allocation concealment

- Allocation concealment is necessary to ensure assignment to treatments are truly randomized
  
  - Concealed from whom?

- Goal of the allocation concealment process:
  Investigators and clinicians should not be able to predict the group to which the next enrollee will be assigned (TxA vs TxB, control vs intervention)

- not to be confused with: method used to randomize patients, nor with blinding
Allocation bias

- Selection bias, occurs if the allocation process is predictable **This is what allocation concealment is trying to protect against**
  - eg. investigator knows next enrolee will be assigned to control group (or “suspected” control group). They wait to enrol a patient with worse prognosis.
  - Trials with inadequate concealment “yield up to 40% larger estimates” (Schulz KF & Grimes DA, *The Lancet* 2002)

- Accidental bias, can arise if the randomization procedure does not achieve balance on risk factors or prognostic covariates

Allocation bias may be a more important determinant of outcome than the treatment itself
Adequate randomization

Means that allocation bias minimized:

• **Concealed process** – so investigators do not know in advance. Which of following will be adequate?
  
  – Central randomization
  – Computer generated
  – Random numbers table
  – Draw numbers from a hat
  – Day of the week
  – Toss a coin
Types of randomization

Randomization – participants have the same probability of being assigned to control or intervention arms

Individual randomization
- Simple
- Blocked
- Stratified

Group randomization
Simple randomization

The most elementary form of randomization:

- toss an unbiased coin each time for each consenting subject;
- use a random number producing algorithm (computer generated - more convenient especially for large studies).
- Large trials – should reliably produce groups with equal sizes and distribution of confounders
- Smaller trial – possible to end up with groups of unequal in size or distribution of confounders
## Blocked randomization

Randomization occurs within blocks, to ensure that numbers in each group remain as close to equal as possible, at all times (also called permuted block randomization)

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Avoids imbalance in the number of subjects assigned to each group; Particularly if sample size is small, or numbers within strata/centre are small</td>
<td>• If the study is not blinded, the study staff know the assignment for the last person before randomization of that person</td>
</tr>
</tbody>
</table>
Blocked randomization

Example: Block size = 4. Gives 6 possible combinations of group assignments:

AABB, ABAB, BAAB, BABA, BBAA, and ABBA.

If study unmasked (not blind) then investigator will know which intervention every 4th participant will be assigned (sometimes for 3rd and 4th participant)

- can get around this by randomly varying block sizes
Variable block randomization

Each Block is of different size
Ranges from N=2 to N=16 (or 2-8, or 2-6)
Commonly used in multi-centre studies
Prevents anyone from guessing what the next subject will be randomized to
But – adds complexity
Can result in imbalance:
  If study stopped early, or,
  Many sites and fewer patients at these sites
Stratified randomization
Randomize within sub-groups – defined on basis of most important potential confounders

To improve chances that important baseline characteristics will be similar in the 2 groups.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Reduce variability in group comparison if the stratification is used in the analysis</td>
<td>• Sometime the variables initially thought most important and used for stratified randomization turn out to be unimportant</td>
</tr>
<tr>
<td></td>
<td>• Other factors identified later are more important</td>
</tr>
</tbody>
</table>
Stratified randomization

Example

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>Smoking Hx</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49 yr</td>
<td>Male</td>
<td>1. Current smoker</td>
</tr>
<tr>
<td>50-59 yr</td>
<td>Female</td>
<td>2. Ex-smoker</td>
</tr>
<tr>
<td>60-69 yr</td>
<td></td>
<td>3. Never smoker</td>
</tr>
</tbody>
</table>

In this example, there will be 18 strata...

- Blocked randomization is then performed to ensure an equal number of participants are assigned to each intervention within each strata.
Stratified randomization
By Centre

This should always be done if multi-centre

Balances differences in population

- Differences in patient population

  Illness severity, comorbidities

- Differences in MD practice - referral

- Differences in study staff – refusal rate

- Differences in recruitment rate

Maintains balance if centres drop out
see article by Doig & Simpson *Journal of Critical Care* 2005 – describes step by step how to do stratified and block randomization using SNOSE (sequentially numbered opaque sealed envelopes)
Double-blind RCT: high dose INH vs standard INH vs placebo for MDR-TB

• “We recruited consecutive, sputum culture-positive, non-HIV infected patients previously diagnosed with pulmonary TB who developed documented MDR-TB and who reported to the study centre during the time of the study.”

• Excluded: patients with exposure to SLD for > 30 days and also “patients who were unwilling to give consent, had abnormal renal or hepatic profile, had a history suggestive of INH hypersensitivity or were pregnant or lactating.”

• “Subjects were randomised to three treatment groups by block randomisation to ensure comparable allocation to the trial arms.” www.randomization.com (randomly permuted blocks?)

• “Both study investigators and patients were blinded to the INH dose”

Katiyar SK et al. IJTLD 2008
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>High dose INH</th>
<th>Low dose INH</th>
<th>Placebo</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>38</td>
<td>42</td>
<td>37</td>
<td>0.017</td>
</tr>
<tr>
<td>EMB resistance</td>
<td>52%</td>
<td>55%</td>
<td>76%</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Other baseline characteristics similar: sex, smoking, prior anti-TB treatment, SM-resistance, PZA-resistance, INH MIC

- Differences in age and ethambutol resistance → chance or inadequate allocation concealment?
- Block randomization and double-blinding

Katiyar SK et al. *IJTLD* 2008
Group randomization

• For some interventions (psychosocial, education, etc) randomization by individuals does not work, because there is interaction among subjects (contamination).

• Groups of subjects, clinics or communities are randomized to intervention vs control; The basic sampling units are groups, not individuals.

• This design is not as efficient as individual randomization. A larger sample size required
Appropriately conducted & effective randomization = groups that differ only in terms of the intervention received

→ strongest study design for causal inference

But bias still possible!
Ways to avoid bias during the trial

Confounding (imbalance of known and unknown variables that can affect outcomes)
- Randomization with allocation concealment

Information bias (most importantly differential misclassification of outcome related to exposure)
- blinding
- quality control of data

Selection bias – can occur during allocation and follow-up (differential loss to follow-up related to assigned intervention)
- Randomization with allocation concealment
- Avoid withdrawals, Minimize losses to follow-up
Blinding (masking)

- Bias can occur in many ways in a clinical study
- Caused by investigators and/or patients.
- Caused by conscious factors, subconscious factors, or both

The general solution to the problem of bias is to keep the subject and the investigator blinded, or masked, to the identity of the assigned intervention.
Types of Blinded Studies

Single Blind

- The investigator is aware of the intervention. But the subject is not.

Double Blind

- Neither the subjects nor the investigators responsible for following the subjects know the identity of the intervention assignment.

Triple-Blind

- In addition to subject and investigators, the data analyst, and the committee monitoring the trial, are not told who is getting what.
Importance of Blinding

Example:

*Benefits of the vitamin C in the common cold*

*Lewis et al. Ann NY Acad Sci 1975; 258 : 505-12*

Participants: Medical staff

Evaluation: severity and duration of common cold was self-reported by the participants
Importance of Blinding

Blinding: Many participants could tell (taste) whether they were on active or placebo

Results:
Participants who stated they did NOT know Vitamin C = placebo
BUT, Participants who stated they DID know Vitamin C > placebo
Data collection: blinding of individuals evaluating outcomes

e.g. Canadian cooperative trial of cyclophosphamide and plasma exchange in progressive multiple sclerosis

- participants randomized to cyclo vs. PE vs. placebo
- double-blind
- 2 neurologists:
  - 1 “monitoring” unmasked, 1 “evaluating” masked
  - only masked assessment used.
  - Unmasked: more likely to observe improved outcomes in PE group
  - Masked: no differences in groups

If Blinding not possible

In many studies intervention cannot be blinded

Surgery, different durations of therapy

Solutions:

Objective outcomes
  – Death, days in hospital, lab-confirmed disease (culture/smear/PCR positive)

Blinded independent review of outcomes
  – Serious adverse events, complications
  – X-ray changes
Data gathering during an RCT: Baseline assessment

- Relevant baseline data **should be measured in all study participants** before the intervention starts.

- Useful for:
  - stratification
  - analysis of baseline comparability
    - assess if randomization worked
  - evaluation of change
  - natural history analysis
  - subgrouping
Data collection & quality control

Major types of problems:
- **missing/incomplete data** (one indicator of the quality of the trial)
- **erroneous data** (error will not necessarily be recognized)
- **variability in the observed characteristics** (reduce the opportunity to detect the real changes): random, systematic or combination of both
- **differential misclassification of outcome** – could bias study results.
  - eg. could arise in an open-label trial if placebo group assessed more often/intensely for outcome
Selection bias *during* the trial

- Investigator-initiated withdrawal
- loss-to-follow-up/drop-out of participants related to the outcome and differ by group
Reasons for investigator-initiated withdrawal

• Ineligibility (error made to enrol them)
  • at times inevitable that participants become ineligible after enrolment and allocation to treatment

• Nonadherence or non-compliance

• Poor quality or missing data
  • Especially if drop-out and no idea of outcomes

• Withdrawal of participants after they have been randomized to an intervention group compromises the comparability of the groups provided by randomization
Withdrawal: high risk of bias

- **Ineligibility** (error made to enrol them)
  - problem arises if discovery of ineligibility is not random
  - even if an equal number of subjects are withdrawn due to ineligibility from control and intervention groups, this can alter the results of the trial if their outcomes differed.
  - exception: if difficult to establish eligibility immediately before randomization. Decision to withdraw can be made later by a blinded person, based on data collected at time of randomization.
  - eg. MDR-TB discovered in trial of 6 vs 9 months of treatment for drug-susceptible TB in PLWH
  - eg. Rifapentine+INH vs INH for LTBI: post-allocation ineligibility if source case culture(-), resistant to INH/Rif, or no DST
Withdrawal: high risk of bias

- Non-adherence or non-compliance
  - people who comply with treatment are different than those who do not comply
    - Mortality in trial of lipid-lowering agents:
      Overall: 18% intervention vs 19% placebo
      Comparing compliant vs non-compliant:
        Intervention: 15% vs 25%; placebo: 15% vs 28%
  - non-adherence could be related to intervention or outcome
    - non-adherence in controls could be due to different reasons than non-adherence in intervention group
    - exclusion could create non-comparable groups

Strategies to improve compliance

Factors that maximize compliance:

• Duration of intervention: Shorter = Better

• Simplicity of intervention: Single dose = Better

• Subject selection: Run-in period used to identify OCD
  • can limit external validity of results by excluding non-compliant persons or those who will have more side-effects

• Fully informed consent: Patient really understands
Losses to follow-up

- participants who can no longer be followed up: refuse to participate, move and can’t be contacted

- Different than participants who were able to complete follow-up

- eg. In TBTC 22, open-label RCT loss to follow-up associated with:
  - birth outside USA/CANADA (aOR 2.07, p<0.01)
  - homelessness (aOR 1.94, p<0.01)
  - enrollment at a health department (aOR 2.71, p<0.01)

- losses to follow-up can bias results – particularly when they are differential between intervention & control arms and associated with the outcome

- differential losses could arise because of differences in side-effects or efficacy

Minimizing losses to follow-up

- Employ study personnel responsible for managing & ensuring follow-up
- Call or visit participants that miss appointments
- *Prior to randomization* exclude those unlikely to return for follow-up (likely to move or unwilling to return)
- Obtain lots of contact information
- Follow-up visits done in locations convenient for participants
- Keep follow-up visits short and sweet
- Provide free medical care
- Monetary subsidies or incentives

THANKS!