**Science of Clinical Trial Design: Principals of Study Design**

**Objectives**
- Provide overview of clinical study designs
- Review types and classification of randomized controlled trials (RCT)
- Discuss important concepts related to RCTs
  - Randomization
  - Blinding
  - Endpoints

**Clinical Research: Study Designs**

**Observational**
- Case-control studies
  - Compare those with and without disease to determine effect of one or more exposures
- Cohort studies
  - Compare those with and without exposure and determine effect on one or more diseases

**Interventional**
- Clinical Trial

**Case-Control: Start with Outcome of Interest**

- Exposed to Anti-Lymphocyte Agents
  - PTLD
  - (Cases)
- Not Exposed to Anti-Lymphocyte Agents
  - No PTLD
  - (Controls)

**Cohort: Start with Exposure of Interest**

- Exposed to Anti-Lymphocyte Agents
  - PTLD
  - No PTLD
- Not Exposed to Anti-Lymphocyte Agents
  - PTLD
  - No PTLD

**Clinical Trial**
- Similar to a COHORT study in that patients/participants are enrolled based on EXPOSURE
- Major Difference: Investigators Allocate the Exposure
- Prospective Study comparing the Effect of an Intervention against a Control group
  - Well defined starting point
  - Patients followed forward in time
  - Intervention (e.g., drug, device, procedure etc) is applied in a standard fashion in order to change some aspect of participant
  - Compared to a control group who must be similar enough to the intervention group so that differences in outcome can be attributed to the new therapy
Clinical Trial Phases

Phase I trials
- How does the agent affect the human body?
  - Study dose ranges, safety, tolerability, PK and PD
  - Usually in healthy volunteers but may also be in patients with severe disease (e.g. resistant malignancy, refractory acute rejection)
- Non-randomized design

FKN 056 FOR LIVER, KIDNEY, AND PANCREAS TRANSPLANTATION

THOMAS E. STARZEL1,2
SATORU TOTOYAMA3
ANTHONY J. DODGROTT4
RAMAN VENKATARAMANAN5
ASHOK JAIN6

Departments of Surgery and Pathology, School of Pharmacy, University of Pittsburgh, University of Pittsburgh, and Veterans Administration Medical Center, Pittsburgh, Pennsylvania, USA
THE LANCET, OCTOBER 28, 1989

Clinical Trial Phases

Phase II trials
- Does the agent or intervention have a biologic effect on the disease?
  - Used to set or confirm dose, estimate proportion of responders, and continue to look for toxicity
- Can be non-randomized case series but often randomized

Clinical Trial Phases

Phase III trials
- Is the new agent or intervention (or new use of a treatment) better than the standard?
  - Usually large RCTs

Clinical Trial Phases

Phase IV trials
- Post-approval trials to assess effectiveness and safety in a more generalized population (post-marketing) – different subgroups (e.g. pediatric patients) – different drug combinations

Randomized Controlled Trial (RCT)
- A clinical trial in which subjects are randomized to an intervention or control and followed for occurrence of outcome
- Randomization - process by which each participant has the same chance of being assigned to either treatment group
- Gold Standard for assessing benefit (and major harm) of an intervention
  (Observational studies more suited to determining rare adverse events)
Classification of RCT: Design Types

- Parallel Group (most common)
- Crossover
- Cluster
- Factorial

Parallel Group Trial

- Most common design
- Patient randomized to receive only one treatment
- Patients followed over time to see if outcome of interest occurs

Crossover Trial

- Sequence 1: Run-In – Drug A – Washout Period – Drug B
- Sequence 2: Run-In – Drug B – Washout Period – Drug A
- Outcome of interest measured at the end of each drug treatment
- Subjects receive BOTH treatments - serve as their own control
- Improves precision by removing between patient variability
- No carry-over effects – washout period helps reduce this risk
- Can only evaluate continuous variables (e.g. cholesterol, BP etc)

Cluster Trial

- Eligible “Clusters”
- Treatment A
- Treatment B
- Transplant Program #1
- Transplant Program #2
- Transplant Program #3
- Transplant Program #4

Factorial Trial

- Both Treatment A and B
- Only Treatment B
- Only Treatment A
- Placebo
- Is Treatment B Effective?
- Is Treatment A Effective?

RCT Design Approaches

- Explanatory trials (Efficacy)
  - address whether an intervention works under ideal conditions
- Pragmatic trials (Effectiveness)
  - address whether an intervention works under routine clinical practice
- Type of trial flows directly from the research question being asked (best case scenario vs real-world)
Classification of RCT: By Hypothesis

- **Superiority Trial**
  - Usual way clinicians think
  - “Is new treatment better than our current therapy?”

- **Non-Inferiority Trial**
  - Is new treatment as good as our current therapy but preferable for some other reason?
  - New treatment must offer other advantages in terms of safety, cost, tolerability or convenience

- **Equivalence Trial**
  - Is the new treatment the same as the standard
  - Identical equivalency (confidence interval=0) impossible to determine
  - Effect must lie within predefined interval
  - Usually requires very large sample sizes and rarely conducted

Steps in a Performing a RCT

- Choose a Research Question
- Select Participants
- Measure Baseline Variables
- Randomize
- Administer Treatments
- Follow-up
- Measure Outcomes
- Analyze
- Report Findings

Formulating the Research Question

- The most important step when planning a trial – take your time to get the question right!!
  - Get the opinion of colleagues and other experts
  - Focus on a single primary research question – don’t try to answer too many questions at once
- The research question should be
  - “FINER”: Feasible, Interesting, Novel, Ethical, Relevant

Characteristics of a Good Research Question/Study

- Is the study **Feasible**?
  - Adequate number of subjects
  - Adequate technical expertise
  - Affordable in time and money
  - Manageable in scope

- Is the question **Interesting**?
  - Passion in an idea is essential but not sufficient

- Is the question **Novel** and innovative?
  - Confirms or refutes previous findings
  - Extends previous findings in an important manner
  - Provides new findings

- Is the study **Ethical**?

- Is the question **Relevant**?
  - To scientific knowledge?
  - To clinical and health policy?
  - To future research directions?

Study Population

- Subjects should be at high risk of the outcome (significant disease)
- Likely to benefit from treatment (Highly Responsive)
- Less likely to be harmed by treatment
- Likely to adhere to treatment
- Ease of recruitment

Study Population

- Example: Patients with condition who meet inclusion/exclusion criteria
- Study Sample should ideally reflect the “Target Population”
Symphony Study

Age 18-75
1st Tx or repeat Tx if previous not lost due to rejection
Exclude if:
- PRA >20%
- CIT>30 hrs
- DCD Donor
- GI condition that might interfere with drug absorption
- History of cancer (except skin)

Randomization
- Process by which each participant has the same chance of being assigned to either treatment group
- Eliminates bias in the allocation of treatments
  - In non-randomized designs, the clinician and/or patient influence what intervention is chosen (selection bias)
  - Proper allocation concealment prevents selection bias
- Tends to produce comparable groups
  - Balanced with respect to known and unknown characteristics and prognostic variables
  - Reduces (does not eliminate) the chance that key predictors are unevenly distributed
- Ensures validity of statistical methods of analysis
  - As long as randomization employed, chi-square and t-test can be used without making assumptions on the distributions of baseline variables

Blinding
- Masking of treatment assignment after randomization
- Who can be blinded?
  - Trial participants
  - Investigators and clinical health care team
  - Outcome assessors and data analysts
  - DSMB
- Easiest for drug trials
  - Difficult or impossible for some surgical trials, physical therapies and psychiatric interventions

Blinding - Advantages
- Controls Co-Interventions
  - Patient and/or physician may add other treatments to improve outcomes if they know patient getting placebo
- Controls Contamination
  - Patient and/or physician might seek out or prescribe experimental therapy being studied if they believe it might work and they know patient is on placebo
- Controls Ascertainment Bias
  - “Softer” outcomes such as quality of life or reported side effects may not be sought out or expressed by patient in a similar manner if treatment assignment known
  - Threshold for carrying out other diagnostic tests (e.g. biopsy) might be influenced by knowledge of treatment assignment.
- Prevents Unequal Withdrawals
  - Patient and/or physician might become dissatisfied or impatient if not on new therapy and drop out in disproportionate numbers

Outcome Measures
- Clinical Endpoint
  - Characteristic that reflects how a patient feels, functions or how long they survive (e.g. graft survival, patient survival, quality of life)
- Biomarker
  - Characteristic that is objectively measured as an indicator of normal biologic processes, pathogenic processes or response to therapy (e.g. GFR, proteinuria, serum creatinine, BP, FEV-1)
- Surrogate End-Point
  - Biomarker that is used as a substitute for a clinical endpoint

Surrogate End-Points
Advantages:
- Usually measured earlier in a trial compared to clinical endpoints
  - allows for shorter, cheaper trials to be conducted
  - results in faster decision-making about treatments
- Typically they are continuous variables so all patients in the trial will have an “event”
  - greatly reduces sample size and increases power
Surrogate End-Points

Disadvantages:

- Most biomarkers are NOT valid surrogate endpoints
- Difficult to actually validate
  - Must be prognostic for hard, clinical endpoint
  - Changes in the surrogate endpoint with treatment must predict changes in occurrence of clinical endpoints
  - Full effect of treatment on clinical endpoint should be captured by the surrogate


Renal Transplant Trials – What are the Primary Outcomes?

Systematic Review
All RCTs 1998-2008
N=285

Primary Outcome Clinical Endpoint: 22% Surrogate: 78%

How Well does Acute Rejection at 1-yr Predict All-Cause Graft Loss?

Intention-to-treat Analysis
- Includes all patients that were randomized regardless of whether they received the assigned treatment
- Takes advantage of the control of confounding provided by randomization
- May underestimate true effect of the treatment (conservative)

Per-Protocol Analysis
- Uses only patients who were adherent (i.e. took treatment according to protocol)
- Results should always be interpreted with caution

As-Treated
- Analyzes patients according to what treatment was received
- Results should always be interpreted with caution

Advantages of RCTs

- Prospective design
- Best way to eradicate bias (selection bias)
- Permit efficient evaluation of single or multiple variables in a well-defined patient group
- Allow meta-analysis of similar trials

Intention-to-Treat vs Per-Protocol

Adapted from Guyatt et al, The Principle of Intention to Treat; Users Guide to the Medical Literature
Disadvantages of RCTs

- Expensive and time consuming
  - many trials enroll too few patients
  - many are performed for a short period of time
  - surrogate measures are used instead of important clinical endpoints
- Limited generalizability
  - Performed under ideal circumstances that don’t reflect real world practice (“these are not my patients”)
- Potentially introduce bias
  - selection bias: attributable to imperfect or incomplete randomization
  - exclusion bias: systematic differences in dropouts
  - detection bias: systematic differences in evaluation of outcomes or failure to blind outcome assessors

Points to Remember about RCTs

- More than one RCT may be needed to test efficacy and effectiveness
- In complex environments… RCTs essential means of comparing therapies
- Results from RCTs …essential for the adoption of findings into clinical practice (central element of Evidence-Based Medicine)

 '__Stratified by Age, Gender, Smoking – Blocks of Four__

<table>
<thead>
<tr>
<th>Strain</th>
<th>Age</th>
<th>Sex</th>
<th>Smoking</th>
<th>Group randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40–49</td>
<td>M</td>
<td>Current</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>2</td>
<td>40–49</td>
<td>M</td>
<td>Ex</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>3</td>
<td>40–49</td>
<td>M</td>
<td>Never</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>4</td>
<td>40–49</td>
<td>F</td>
<td>Current</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>5</td>
<td>40–49</td>
<td>F</td>
<td>Ex</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>6</td>
<td>40–49</td>
<td>F</td>
<td>Never</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>7</td>
<td>50–59</td>
<td>M</td>
<td>Current</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>8</td>
<td>50–59</td>
<td>M</td>
<td>Ex</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>9</td>
<td>50–59</td>
<td>M</td>
<td>Never</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>10</td>
<td>50–59</td>
<td>F</td>
<td>Current</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>11</td>
<td>50–59</td>
<td>F</td>
<td>Ex</td>
<td>combo A/B/C</td>
</tr>
<tr>
<td>12</td>
<td>50–59</td>
<td>F</td>
<td>Never</td>
<td>combo A/B/C</td>
</tr>
</tbody>
</table>

Blocking – minimizes imbalance between treatment assignments at all times throughout trial (Treatment A vs Treatment B)
Stratified – ensures equal balance of important predictors between groups at all times in the trial

 '__Attributes and Outcome: Efficacy vs Effectiveness__

| Attribute | Efficacy | Effectiveness | "Pragmatic"
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Question</td>
<td>Can the Rx work under IDEAL circumstances</td>
<td>Does the Rx work under USUAL circumstances</td>
<td></td>
</tr>
<tr>
<td>Eligibility</td>
<td>Strict: Limited to high-risk, highly-responsive, compliant patients</td>
<td>All-comers with the disease or condition</td>
<td></td>
</tr>
<tr>
<td>Rx</td>
<td>Given by “experts” and very closely monitored</td>
<td>Given in routine clinical care</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>Intense frequent follow-up visits</td>
<td>No more than routine clinical care</td>
<td></td>
</tr>
<tr>
<td>Patient Compliance</td>
<td>Closely monitored usually with active interventions to enhance compliance</td>
<td>Usually no interventions to enhance compliance</td>
<td></td>
</tr>
<tr>
<td>Events</td>
<td>Usually restricted to those that answer the biologic question</td>
<td>All harmful events regardless of cause</td>
<td></td>
</tr>
</tbody>
</table>

**What Makes a “Good” Research Study?**

**Good Idea**
- Never been done before
- Common condition/problem
- Major potential impact on morbidity and mortality
- Comparison based upon sound basic science
- Novel approach to care

**Less Interesting Idea**
- Has been done before but not in Canada (me too study)
- No underlying physiological or theoretical framework
- Too complex or convoluted (protocol will never be implemented)
- Methodologically sound but only small increment forward
- Thinking too small

**Why Conduct RCTs?**

- Only study design that can prove causation – observational studies can only show associations
- Historically, magnitude of benefits of interventions was very large …… RCTs unnecessary
  - Blood transfusion for severe post-partum hemorrhage (1818)
  - Insulin for type 1 diabetes (1922)
  - Sulfonamides for puerperal sepsis (1936)
  - Electric shock for ventricular fibrillation (1947)
  - Parachutes for jumping out of planes….

**Why Conduct RCTs?**

- Very few “parachutes” in Medicine these days
- Magnitude of clinical benefit much smaller and as a consequence, unbiased and more powerful methodology essential
- There are multiple diagnostic and therapeutic choices for any given condition/disease and “best” choices not self-evident (just think of the variability in immunosuppressive protocols used at different transplant centers)
  - Does early conversion to sirolimus improve kidney function?
  - Does induction therapy improve outcomes for low risk recipients?
  - Does conversion to belatacept for CNI toxicity improve graft survival?

**Why Conduct RCTs?**

- Very few “parachutes” in Medicine these days
- Magnitude of clinical benefit much smaller and as a consequence, unbiased and more powerful methodology essential
- There are multiple diagnostic and therapeutic choices for any given condition/disease and “best” choices not self-evident
  - Just think of the variability in immunosuppressive protocols used at different transplant centers

---

**When is it Acceptable to Conduct a Trial**

1. **No Preference**
   - “Theoretical Equipoise”
   - Believe that either treatment A or treatment B will be equally beneficial
   - Rarely, if ever occurs in the real world

2. **Uncertainty**
   - “Hunch” that new or different treatment is better
   - Uncertain whether your hunch is correct
   - We believe the new therapy will be better but our uncertainty includes the possibility that it might be no better or even worse

3. **Community Uncertainty**
   - “Clinical Equipoise”

---

Clinical Equipoise

- Sets the moral foundation for the trial
- Clinical equipoise is satisfied when:
  “There is genuine uncertainty within the expert medical community - not necessarily on the part of the individual investigator - about the preferred treatment.”


Clinical Equipoise

- Allows clinical investigators to continue a trial until they have enough statistical evidence to convince other experts of the validity of their results, without a loss of ethical integrity on the part of the investigators.

Implications of Clinical Equipoise

- Obligates investigators to determine that a state of clinical equipoise exists before starting a trial (systematic review of the evidence)

- Requires an organized study and presentation of the evidence to colleagues, ethics boards, funding and regulatory agencies, journal editors and patients

James Lind - HMS Salisbury (1747)

12 Sailors with Scurvy

Divided them into 6 treatment groups (n=2 per group) – all had same diet except for the following
1. Quart of cider daily
2. 25 drops of vitriol (sulfuric acid) daily
3. 6 spoons of vinegar daily
4. ½ pint seawater daily
5. 2 oranges and 1 lemon daily
6. Drink of barley water

Day 6
- ran out of fruit
- one sailor in Group 5 given citrus was back at work and the other was nearly recovered
- only other sailors that showed any improvement were from group 1 (apple cider)

Probably the earliest documented failure of Knowledge Translation as citrus fruit/juice supplement did not become widely used for scurvy until 1794!!

Clinical Research: Study Designs

Observational
- Case-control studies
- Cohort studies

Interventional
- Clinical Trial

Desired Properties of Primary Outcome Measures

- The primary endpoint should be:
  - Appropriate (to the goals of the study)
  - Objective - should require minimal subjective judgment to measure
  - Valid - measures what it is intended to measure
  - Responsive – sensitive to changes due to treatment
  - Clinically available – ideally measured as part of routine clinical care
  - Efficient (i.e. affordable to measure in terms of time and cost)
Blinding

- Masking of treatment assignment after randomization
- Who can be blinded?
  - Trial participants
  - Investigators and clinical health care team
  - Outcome assessors and data analysts
  - DSMB
- Best to avoid confusing term such as single-blind, double-blind, triple-blind etc as they mean different things to different people
- Easiest for drug trials – difficult or impossible for some surgical trials, physical therapies and psychiatric interventions
  - Ensure rigorous follow-up of both groups – blind outcome assessors – specify co-interventions to ensure balance

Advantages of RCTs

- Prospective design
- Best way to eradicate bias (selection bias)
- Permit efficient evaluation of single or multiple variables in a well-defined patient group
- Define optimal care
- Allow meta-analysis of similar trials
- Considered the “gold standard” in evidence-based practice

Disadvantages of RCTs

✓ Expensive and time consuming
  - many enroll too few patients
  - many are performed for a short period of time
  - surrogate measures are used instead of important clinical endpoints

✓ Potentially introduce bias
  - selection bias: attributable to imperfect or incomplete randomization
  - exclusion bias: systematic differences in dropouts
  - detection bias: systematic differences in evaluation of outcomes or failure to blind outcome assessors

✓ Limited generalizability
  - Performed under ideal circumstances that don’t reflect real world practice (“these are not my patients”)

What Should the Ideal RCT Do?

1. Establish if a particular intervention works.
2. Determine the overall benefits and risks for given patients.
3. Minimize the influence of chance, bias, and confounding.
4. Accomplish these objectives with the fewest patients possible (efficiency).

Points to Remember about RCTs

✓ More than one RCT needed to test efficacy and effectiveness

✓ Fundamental methodology used in comparing health interventions

✓ In complex environments... RCTs essential means of comparing therapies

✓ Results from RCTs …essential for the adoption of findings into clinical practice (central element of Evidence-Based Medicine)

Methods of Randomization

Most common methods are:
  - Opaque envelopes
  - On-site computer generated
  - Off-site centralised (telephone, web-based)

✓ All methods ensure that investigators and/or patients do not know treatment assignment until AFTER randomization (Allocation concealment to prevent selection bias)
Components of a Sound Research Question

- "PICOT"
  - Persons
  - Interventions
  - Comparator
  - Outcomes
  - Time

Components of a Sound Research Question

- Persons
  - Type of patient - Severity of condition
  - Source
- Interventions
  - Drug, surgery, procedure, counseling etc that will be studied
  - Given by whom and at what location
- Comparator
  - Usual care – placebo – active comparator
- Outcomes
  - Both benefits and harms
  - How they will be ascertained
- Time
  - When will outcomes be measured

Research Question

Among renal transplant recipients with proteinuria (>200 mg/day), can ramipril (5 mg po BID) administered by nephrologists reduce the risk of transplant failure or death better than placebo over the subsequent 4 years?

RCT Design Approaches

- Explanatory trials (Efficacy)
  - address whether an intervention works under ideal conditions
- Pragmatic trials (Effectiveness)
  - address whether an intervention works under routine clinical practice
- Continuous spectrum – often difficult to lump a trial totally into one category
- Type of trial flows directly from the research question being asked (best case scenario vs real-world)
- Both Efficacy and Effectiveness are important and ideally should be addressed sequentially

Factorial Design

- Test of more than one treatment in the same trial
- Treatments must not interact
  - Difficult to prove, requires large sample
  - e.g. Effect of aspirin on MI is same with and without beta-carotene (Physicians Health Study)
- Best used for unrelated research questions evaluating different outcomes

How Well does eGFR at 1-yr Predict All-Cause Graft Loss?

- Graph showing the relationship between eGFR at 1-yr and all-cause graft loss.