RANDOMIZED CONTROLLED TRIALS

Purpose, design, conduct, analysis and reporting scientifically sound, ethical RCTs

H. Myra Kim, Sc.D.
Odessa Fogarty Workshop
October 2013
Two arm parallel group design

Pool of eligible subjects

R

T

C

stop

time
Randomized Controlled Trials

- RCT utilizes three important steps to avoid potential bias
  - Control Group (to remove placebo effect)
  - Randomization (to avoid selection bias and balance the groups)
  - Blinding (to remove or equalize biases due to patient’s desire to please and investigator’s enthusiasm)
Science vs. Ethics

- The ultimate goal of RCT is to develop the knowledge needed to guide clinical practice and health policy
- Want both scientifically valid and ethically conducted RCTs
- Not separate ideals
- Ethical RCTs begin with scientific validity
- Participant protection and scientifically valid RCTs are not conflicting goals
An example

- The 1954 “field trial” of the Salk poliomyelitis vaccine (a killed-virus vaccine prepared from monkey kidney tissue)

- The largest public health experiment ever
  - the first three grades in the 272 counties in 44 states with the highest incidence of the disease were targeted
  - 1) Experimental areas: 749,236 schoolchildren injected with vaccine or placebo
  - 2) Observed areas: 1,080,680 school children either vaccinated or participated as “observed controls”

http://www.stat.luc.edu/StatisticsfortheSciences/MeierPolio.htm
“The vaccine works. It is safe, effective, and potent.” With these words and in this auditorium on April 12, 1955, Thomas Francis, Jr., Prof of Epidemiology at the UM SPH, announced the results of a poliomyelitis vaccine field trial on 1.8 million children across the nation…” (Plaque in the Rackham Lobby)

- Who does the vaccine belong to?
  “Well, the people, I would say. There is no patent. Could you patent the sun?” (Jonas Salk)

- More recent controversies: For cell-transplant surgical interventions for patients with Parkinson disease, drilling of burr holes into the skull was used as a sham surgery; arthroscopic knee surgery
Rackham Building
CONSORT Statement

- **CONsolidated Standards Of Reporting Trials**
- First published in 1996 to improve the quality of reporting of clinical trials
- Includes a flowchart of progress through the various stages of a trial
- The 22 items and flowchart
- [http://www.consort-statement.org](http://www.consort-statement.org)
CONSORT focuses on reporting

- The quality of a RCT is multidimensional
  - design (including problem selection/formulation)
  - conduct (including monitoring)
  - analysis
  - Reporting
- Will go over CONSORT checklist from both planning from perspectives of the design, conduct, analysis and reporting of scientifically sound, ethical RCTs
Scientific Validity

- **Internal validity** -- extent to which systematic error (bias) is minimized
- **External validity** -- extent to which results of the RCT provide a correct basis for generalization to other circumstances
Internal validity and bias

- **Selection bias** -- biased allocation to treatment groups
- **Performance bias** -- unequal provision of care apart from treatment being evaluated
- **Detection bias** -- biased assessment of outcome(s)
- **Attrition bias** -- biased handling of protocol deviations/dropouts
22 item checklist and a flowchart depicting the passage of participants through the four stages of a RCT
- enrollment
- intervention allocation
- follow-up
- Analysis

Items grouped according to the usual sections of a scientific publication
- Title and abstract (1)
- Introduction (2)
- Methods (3 - 12)
- Results (13 - 19)
- Discussion (20 - 22)
Rationale for CONSORT

- 30% of 67 trials failed to report if outcome assessments were blinded
- 27% of 45 trials did not specify the primary outcome
- 13% of 119 trials excluded patients from the analysis
1. Title and Abstract

- Identify in title how participants were allocated to interventions (e.g., “random allocation,” “randomized,” or “randomly assigned”)
  - Intent is to assure that the paper will be indexed as an RCT
- Structured abstracts recommended with a series of headings regarding
  - Background
  - Objectives
  - Design
  - Methods (Data collection & analysis)
  - Main results
  - Conclusions
In the mid 1990’s Medline searches yielded only about 50% of RCTs relevant to a topic

In 1999, the Cochrane Collaboration had identified 100,000 RCTs that had not been indexed as such by Medline

Cochrane Collaboration

- Meta-analyses for a large number of treatment by disease combinations
- www.cochrane.org
- Contains new and updated entries

Example:
- Exercise therapy for depression
- Melatonin for the prevention and treatment of jet lag
2. Introduction

- Scientific background and rationale for the study
  - Describe the problem that necessitated the work
  - Justify the need for a new trial
    - “Research addressing previously answered questions either denies participants effective treatment or places them at risk for no benefit, or both.”
    - Should ideally include systematic review of previous similar trial or absence of such trials.

- Objectives: questions the trial was designed to answer

- Hypotheses: pre-specified questions being tested to meet the objectives
3. Methods: Design

Example: “This was a multicenter, stratified (6-11 years and 12-17 years of age, with 2:1 imbalanced randomization [2:1]), double-blind, placebo-controlled, parallel-group study conducted in the U.S. (41 sites).”

- Trial type (e.g., parallel group, multi-arm parallel group, cluster randomized, or cross-over)
- Conceptual framework (e.g., superiority, non-inferiority, or equivalence)
- Allocation ratio
- Phase of the trial, if drug trials
- Report important changes to methods (e.g., dropping a center due to poor recruitment, discontinued an arm based on DSMB recommendation)
4. Methods: Participants

- Eligibility criteria for participants
  - Inclusion/exclusion criteria
  - Describe participants

- Settings and locations of recruitment
  - E.g., community or emergency room

- Where the study was carried out
  - E.g., Multi-center trials

- Because they precede randomization, eligibility criteria do not affect the internal validity of the trial, but they do impact generalizability, i.e., to whom the results apply.
5. Methods: Interventions

- Detailed description of the interventions used in the groups and how/when/by-whom administered
- Describe controls/placebos
- Describe “usual care” if control group is to receive “usual care”
6. Methods: Outcomes

- Define primary and secondary outcome measures
- When applicable, describe any methods used to enhance the quality of measurement or reasons for switching the outcome
- Limit the number of primary outcomes
  - More outcomes create problems involving multiplicity of inferences
- Describe how and when the outcomes were assessed
  - Number repeat assessments
  - Times of repeat assessments
6. Methods: Outcomes

Issues with outcomes definition

- >70 outcomes were used in 196 trials of NSAIDS in rheumatoid arthritis
- 640 different instruments used in 2000 trials in schizophrenia
- 149/2,000 trials used unpublished scales - a source of bias

Important:

- Previously validated scales preferred
- Describe number and training of assessors
- Blind assessors to reduce bias
Example: “To detect a reduction in postop hospital stay of 3 days (SD 5 days), which is in agreement with the study of Lobo et al. with a two-sided 5% significance level and 80% power, a sample size of 50 patients per group was necessary, given an anticipated dropout rate of 10%. To recruit this number of patients a 12-month inclusion period was anticipated.”

- Report how sample size was determined
  - the intended size of the clinically important difference
  - N should be “large enough,” but not “too large”
  - Explain why the N is different from the intended, if different

- Interim Analysis
  - If multiple “looks” taken, must be reported
  - Adopt formal stopping rules before the trial and report
8-10. Methods: Randomization

- Random allocation is done to generate groups that are roughly comparable in terms of known and unknown prognostic variables.
- Bias induced by lack of or flawed randomization (e.g., Linus Pauling’s study of Vit C and cold, and Vit C and survival in terminal cancer patients).
  
  “We believe that the ascorbate-treated patients represent a random selection of all the terminal patients in the hospital, even though no formal randomization process was used.”

- Deterministic methods are not random.
- Use random number generators or random number tables.
Example: “Randomization sequence was created using Stata 9.0 (StataCorp, College Station, TX) statistical software and was stratified by center with a 1:1 allocation using random block sizes of 2, 4, and 6.”

Methods
- Random number generators or random number tables
- Stratified random allocations
- Minimization algorithms successfully balance allocations within subsets
- Block randomization block size
Allocation concealment and Implementation of random allocations

- Allocation concealment to prevent selection bias
- **Example:** Assignment using sequentially numbered sealed envelopes, a central telephone, or web site

Important to

- Have sequence generation and assignment of participants to groups by separate people
- Conceal next allocation at the time of enrollment
  - Obtain informed consent before randomization
  - Make the decision to accept or reject a participant in ignorance of the next treatment assignment
11. Methods: Blinding

- Blinding refers to the practice of keeping treatment assignments unknown to:
  - Participants
  - Health care providers
  - Data collectors or those assessing outcomes
  - Analysts

- Seeks to prevent performance and ascertainment bias, but cannot always be implemented
  - E.g., CBT impossible to blind from participants or therapists, but data acquisition should be blinded
  - If it can be done, it eliminates various types of bias
  - “Double blind” refers blinding participants and providers
12. Methods: statistical methods

- Describe methods and the rationale for the method used to compare groups in terms of primary and secondary outcome(s)
  - Use confidence intervals for the estimated effects
  - Use exact p-values (e.g. p=0.002, not p<0.01)
  - Use appropriate analytic methods if data are not independent
  - Use tests of interactions to compare treatment effect between subgroups
13. Results: Participant flow

- Flowchart strongly recommended
- Flowchart information
  - Number evaluated for eligibility
  - Number randomized (N in each group)
  - Number dropped out or lost-to-follow-up
Fig 2. Flow diagram of a multicentre trial of fractional flow reserve vs. angiography for guiding percutaneous coronary intervention (PCI) (adapted from Tonino et al(313)). The diagram includes detailed information on the excluded participants.
13. Results: Participant flow

- Flowchart info has implications for external & internal validity
  - Number evaluated for eligibility
    - N did not meet eligibility criteria
    - N met eligibility criteria, but refused to participate
  - Number randomized (N in each group)
  - Characteristics and reasons for participants not receiving the allocated treatment, lost to follow-up, or being excluded from the analysis after randomization
    - Intent-to-treat analysis: analysis includes all randomized participants in the original groups
    - Per-protocol analysis
  - Documentation of “protocol deviations” important
14. Results: Recruitment

- Specify dates defining the periods of recruitment and follow-up
  - Place the study in historical context
  - Therapies evolve continuously and may affect “routine care”
  - Report rate of recruitment
  - Report min/max/median duration of follow-up if they are not the same for all participants
15. Results: Baseline data

- Present baseline demographic and clinical characteristics by each group
- Random assignment prevents selection bias, but does not guarantee baseline equivalence of the groups
- Comparisons at baseline should be for consideration of the prognostic strength of the variables measured and the size of any imbalances occurred by chance
  - Illogical to test for baseline differences
  - Can adjust for those considered highly associated with the outcomes
16. Results: Numbers analyzed

- Number of subjects in each group included in each analysis
- State the results in absolute numbers when feasible (e.g., 10 of 20, not just 50%)

- Were analyses done by “intention to treat (ITT)?”
  - ITT generally avoids bias associated with non-random loss of participants
  - Was a missing data method considered for outcomes that are missing for some?
- “On treatment” or “per protocol” analyses have excluded drop-outs or non-compliant participants
17. Results: Outcomes and estimation

- For each primary or secondary outcome, summarize the results for each group and report:
  - Means, standard deviations, and confidence intervals for “normal” data; medians and relevant percentiles for skewed data; or counts and percents for discrete outcomes
  - Effect sizes with confidence intervals, i.e., unadjusted and adjusted differences of means for continuous data; odds ratios, relative risks, or risk differences for binary outcomes; hazard ratios for survival analyses
  - Significance tests for group differences
18. Results: Ancillary analyses

- Address multiplicity by reporting any other analyses performed, including subgroup and adjusted analyses
- Indicate analyses pre-specified and exploratory
19. Results: Harms

- Report all important adverse events (AEs) or side effects in each group
  - Define carefully
  - Adverse events – can be related (called side-effects) or not related to the treatment
  - Serious adverse events (SAEs) – reactions that threaten life or function (e.g., reactions with hospitalization); should be reported to promptly to regulators
  - Not all will be due to the intervention
  - RCTs will not detect rare AEs
20. Discussion: Limitation/Interpretation

- Interpret the results, taking into account the study hypotheses and sources of potential bias or imprecision

- Some journals recommend the use of a structured discussion section
  - Brief synopsis of key findings
  - Possible mechanisms/explanations
  - Comparison with other studies
  - Limitations of present study
  - Clinical/research implications
21. Discussion: Generalizability

- Depends on the characteristics of the participants, the trial setting, the treatment regimens tested, and the outcomes assessed
- Interpret findings relative to eligibility criteria, outcome measures, and risk of AEs.
  - Do results apply to groups in other circumstances?
  - How could they be applied to an individual patient?
- % refused in eligible participants may indicate preference or acceptability
- May want to provide a measure helpful in assessing benefit-to-risk balance (e.g., NNT)
- Overall evidence
  - Put current results in context of existing evidence
  - A meta-analysis of existing trials would be good.
Extensions to CONSORT

- **Design Extensions**
  - Cluster Trials (Campbell *et al*, *BMJ*, 328, 2004)
  - Non-inferiority and Equivalence Trials
  - Pragmatic Trials

- **Interventions Extensions**
  - Herbal medical interventions
  - Non-pharmacologic treatment interventions

- **Data Extensions**
  - Patient reported outcomes
  - Harms
The SPIRIT 2013 Statement

- **SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) initiative**
  - A guideline for the minimum content of a clinical trial protocol to facilitate the drafting of high-quality protocols.
  - The 33-item checklist applies to protocols for all clinical trials and focuses on content rather than format.
  - Adherence to SPIRIT would enhance the transparency and completeness of trial protocols for the benefit of investigators, trial participants, patients, sponsors, funders, research ethics committees or institutional review boards, peer reviewers, journals, trial registries, policymakers, regulators, and other key stakeholders.