SCT

SOCIETY FOR CLINICAL TRIALS

Society for Clinical Trials 32nd Annual Meeting

Workshop P1
Essentials of Randomized Clinical Trials

Sunday, May 15, 2011
8:00 AM - 5:00 PM
Plaza A

Part I: Introduction

Yves Rosenberg, M.D, M.P.H.

Program Director, Acting Branch Chief
Atherothrombosis and Coronary Artery Disease Branch
Division of Cardiovascular Sciences
National Heart, Lung and Blood Institute, NIH
6701 Rockledge Dr., Rm. 8148
Bethesda, MD 20892-7956
Tel.: 301-435-1292
Fax: 301-480-3667
E-mail:rosenbey@nih.gov

Essentials of Randomized Clinical Trials
SCT Pre-Conference Workshop – Vancouver, BC
May 15, 2011

I-1

I-2

Introduction to Randomized Clinical Trials

Outline I

- **Historical perspective**
- **Rationale for randomized clinical trials**
  - Rationale for randomization
  - The equipoise issue
  - To blind or not to blind?
- **Key issues in the design of a RCT:**
  - What is the study question? Defining hypothesis, objectives and end-points
  - Defining selection criteria: generalizability vs. homogeneity
  - Selecting the control group: the placebo vs. “usual care” issue

I-3

Introduction to Randomized Clinical Trials

Outline II

- **The different phases of a RCT**
- **Basic RCT Designs**
  - Parallel, cross-over, factorial and cluster designs
  - Large Simple Trials
  - Comparative Effectiveness trials
  - Superiority, Equivalence and Non-Inferiority trials
- **Key elements of a RCT Protocol**
- **Some ethical considerations**
  - Informed Consent Process
  - Patient safety issues

I-4

Historical perspective

Prove thy servants, I beseech thee, ten days; and let them give us pulse to eat, and water to drink. Then let our countenances be looked upon before thee, and the countenance of the children that eat of the portion of the King's meat; and as thou seest, deal with thy servants. So he consented to them in this matter, and proved them ten days. And at the end of ten days their countenances appeared fairer and fatter in flesh than all the children which did eat the portion of the King's meat.

Book of Daniel, Chapter 1, Verses 12 -15

I-5

Historical perspective

I raised myself very early to visit them when beyond my hope I found those to whom I had applied the digestive medicament, feeling but little pain, their wounds neither swollen nor inflamed, and having slept through the night. The others to whom I had applied the boiling oil were feverish with much pain and swelling about their wounds. Then I determined never again to burn thus so cruelly the poor wounded by arquebuses.

Ambroise Paré (1510 – 1590)

I-6

Historical perspective

Lind's Scurvy Study

Nb of Patients: 12

Test Treatments:
- Cyder, 1qt/day
- Elixir vitriol, 25 gutts, 3 times/day
- Vinegar, 2 tsp, 3 times/day
- Bigness of nutmeg 3 times/day
- orange (2) ; lemon (1) /day

Control Treatment
- Sea-water, ½ pt/day

Follow-up: 6 days

Outcome: fit for duty

Lind's Treaty on Scurvy, 1753

I-7

Historical perspective

Key Dates in the History of RCT

- 1747 Lind's Scurvy experiment
- 1800 Waterhouse's smallpox experiments
- 1863 Gull's use of Placebo Treatment
- 1923 Fisher's 1st application of randomization
- 1931 1st use of randomization (and blindness) in a clinical trial
- 1946 Nuremberg Code for Human Experimentation
- 1962 Hill AB Statistical Methods of Clinical and Preventive Medicine
- 1979 Society for Clinical Trials
- 2006 Clinical and Translational Science Awards (CTSAs) program
- 2009: The Recovery Act (ARRA) provides $1.1 billion for Comparative Effectiveness Research.

From Curtis L Meinert. Clinical Trials, Oxford University Press 1986

I-8

Introduction to Randomized Clinical Trials

Outline I

- **Historical perspective**
- **Rationale for randomized clinical trials**
  - Rationale for randomization
  - The equipoise issue
  - To blind or not to blind?
- **Key issues in the design of a RCT:**
  - What is the study question? Defining hypothesis, objectives and end-points
  - Defining selection criteria: generalizability vs. homogeneity
  - Selecting the control group: the placebo vs. "usual care" issue

I-9

Randomized Clinical Trials
Some Terminology

- **Clinical Trial:**
  - An experiment testing medical (e.g. drug, surgical procedure, device or diagnostic test) treatments on human subjects
    - Experiment: a series of observations made under conditions controlled by the scientist
    - Prospective (≠ case-control study)
    - Comparative (≠ cohort study)
    - Involves human subjects
  - A research activity that involves administration of a "test treatment" to some "experimental unit" in order to evaluate that treatment

I-10

Randomized Clinical Trials
Some More Terminology

- **Randomization:** the process of assigning patients to treatment using a random process (such as a table of random numbers)

- **Randomized controlled clinical trial (or randomized clinical trial-RCT):**
  - Clinical trial with at least one control treatment and one test treatment
  - In which the treatment administered are selected by a random process

I-11

Why do we need clinical trials?
Too much Bloodletting!

Treatment choices

(Thanks to Michael Lauer, MD)

I-12

Why do we need clinical trials?
Too much Bloodletting!

A controlled clinical trial in 1809?

In his 1814 Edinburgh MD thesis (see figure), Alexander Hamilton described an experiment that took place during the Peninsular War to assess the effects of bloodletting. Hamilton and two other army surgeons carried out the experiment, which involved 366 sick soldiers in 1809 in the hospital at Elvas, Portugal.

"It had been so arranged, that this number was admitted, alternately, in such a manner that each of us had one third of the whole. The sick were indiscriminately received, and were attended as nearly as possible with the same care and accommodated with the same comforts. One third of the whole were soldiers of the 61st Regiment, the remainder of my own (the 42nd) Regiment. Neither Mr Anderson nor I ever once employed the lancet. He lost two, I four cases; whilst out of the other third [treated with bloodletting by the third surgeon] *thirty five patients died.*"

No corroborative evidence has yet emerged that the trial actually happened, but it is still noteworthy that Hamilton judged that his description of alternation and standard conditions would impress his examiners.[1]

I Milne
Royal College of Physicians, 9 Queen Street, Edinburgh EH2 1JQ, UK

I Chalmers
UK Cochrane Centre, NHS Research and Development Programme, Middle Way, Oxford OX2 7LG, UK

1 **Milne I**, Chalmers I. Tackling bias in assessing the effects of healthcare interventions: early contributions from James Lind, Alexander Lesassier Hamilton and T Graham Balfour. *Proc R Coll Phys Edinb* (in press).

I-13

Why do we need clinical trials?
Too much Bloodletting!

"during the last decades we have certainly bled too little. Pneumonia is one of the diseases in which a timely venesection [bleeding] may save life. To be of service it should be done early. In a full-blooded, healthy man with a high fever and bounding pulse the abstraction of from twenty to thirty ounces of blood is in every way beneficial" (Osler 1892).

I-14

Randomized Clinical Trials
Why Randomize?

"The goal of randomization is to produce comparable groups in terms of general participant characteristics, such as age or gender, and other key factors that affect the probable course the disease would take. In this way, the two groups are as similar as possible at the start of the study. At the end of the study, if one group has a better outcome than the other, the investigators will be able to conclude with some confidence that one intervention is better than the other. "

Friedman et al. Fundamental of Clinical Trials, Mosby Press

I-15

Randomized Clinical Trials
Why Randomize?

- To find out which (if any) of two or more interventions is more effective
- Produce comparable groups
  - Protect against both known and unknown/unmeasured confounders (prognostic factors)
  - Eliminate treatment selection bias
- Best to establish causality
- Can define "Time zero"

I-16

Randomized Clinical Trials
Why Randomize?

- Necessary to detect small but clinically important treatment differences
- Protect against possible time trends in:
  - Patient population and disease characteristics
  - Diagnostic methods and supportive care
- Provides a valid basis for statistical tests of significance

I-17

Randomized Clinical Trials
Why Randomize: The Hormone Replacement Therapy Story

The NEW ENGLAND JOURNAL of MEDICINE

Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the nurses' health study

(N Engl J Med 1991, 325: 756-762)

METHODS. We followed *48,470 postmenopausal women*, 30 to 63 years old… During up to *10 years of follow-up (337,854 person-years)*, we documented *224 strokes, 405 cases of major coronary disease (nonfatal myocardial infarctions or deaths from coronary causes), and 1263 deaths* from all causes.

RESULTS. *After adjustment for age and other risk factors, the overall relative risk of major coronary disease in women currently taking estrogen was 0.56* (95 percent confidence interval, 0.40 to 0.80…

CONCLUSIONS. Current estrogen use is associated with a reduction in the incidence of coronary heart disease as well as in mortality from cardiovascular disease, but it is not associated with any change in the risk of stroke.

I-18

Randomized Clinical Trials

Why Randomize: The Hormone Replacement Therapy Story

June 26, 1995

ESTROGEN FOREVER?
The prevailing medical view is that most should stay on estrogen for the long haul …
At the turn of the century, women died soon after their ovaries quit." Now they live to face heart disease, osteoporosis, increased fractures -- problems that may be prevented in part by taking estrogen...
There may be other risks and other advantages of HRT, but what doctors know is limited by the type of research that has been done. Instead of setting up a group of women on HRT and a carefully matched control group that does not take hormones, studies like the Nurses trial simply look at populations of women who made their own choice whether to take estrogen. "the problem with this.. is that women who take hormones go to doctors more, eat well, exercise and are in better health generally than women who don't take hormones." Thus it is hard to tell whether their lower rates of heart disease or colon cancer or fractures reflect HRT or these other healthy habits.

I-19

Randomized Clinical Trials

Why Randomize: The Hormone Replacement Therapy Story

JAMA

Online article and related content current as of April 13, 2009.

Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Controlled Trial

Writing Group for the Women's Health Initiative Investigators

(JAMA 2002: 288: 321-333)

Design Estrogen plus progestin component of the Women's Health Initiative, a ***randomized controlled primary prevention trial*** (planned duration, 8.5 years) in which *16608 postmenopausal women* aged 50-79 years with an intact uterus at baseline were recruited by 40 US clinical centers in 1993-1998.

Interventions Participants received ***conjugated equine estrogens, 0.625 mg/d, plus medroxyprogesterone acetate, 2.5 mg/d, in 1 tablet (n = 8506) or placebo (n = 8102).***

Main Outcomes Measures The primary outcome was *coronary heart disease (CHD)* (nonfatal myocardial infarction and CHD death), with *invasive breast cancer as the primary adverse outcome*. A global index summarizing the balance of risks and benefits included the 2 primary outcomes plus stroke, pulmonary embolism (PE), endometrial cancer, colorectal cancer, hip fracture, and death due to other causes.

Conclusions ***Overall health risks exceeded benefits*** from use of combined estrogen plus progestin for an average 5.2-year follow-up among healthy postmenopausal US women. All-cause mortality was not affected during the trial. The risk-benefit profile found in this trial is not consistent with the requirements for a viable intervention for primary prevention of chronic diseases, and ***the results indicate that this regimen should not be initiated or continued for primary prevention of CHD.***

I-20

Randomized Clinical Trials

Why Randomize: The Hormone Replacement Therapy Story

July 22, 2002

A large, federally funded clinical trial, part of a group of studies called the Women's Health Initiative (WHI), has definitively shown for the first time that the hormones in question--estrogen and progestin--are not the age-defying wonder drugs everyone thought they were. As if that weren't bad enough, the results, made public last week, proved that taking these hormones together for more than a few years actually increases a woman's risk of developing potentially deadly cardiovascular problems and invasive breast cancer, among other things.

I-21

Randomized Clinical Trials
Why Randomize?

"Mind you, only one doctor out of ten recommends it."

I-22

Randomized Clinical Trials
When Randomize?

- Is there equipoise?
  - *Definition*: A state of genuine uncertainty on the part of the clinical investigators regarding the comparative therapeutic merits of each arm of the trial
  - Trial options must be consistent with *standard of care*: if state of genuine uncertainty exists randomization is an acceptable option
- Clinical equipoise vs. societal equipoise?
- Importance of the informed consent process
  - Accept risk of new treatment
  - Accept concept of randomization
  - Informed about alternative treatment options

I-23

Randomized Clinical Trials
When Randomize?

- Finding "window of opportunity"
  - Too early
    - Not enough "preliminary" evidence :biological plausibility, epidemiologic studies
    - Intervention not "mature" enough (e.g. surgical technique)
  - Too late: intervention already established in clinical practice
- Clinical Equipoise
- Changing Clinical Practice Guidelines

I-24

Randomized Clinical Trials
To Blind or not to Blind?

- **Definition:** concealment (masking) to the patient (single blind), investigator (double) and the monitors (triple) of the identity of the intervention.
  (Opposite = unblinded or open trial)
- **Goal:** avoid bias (systematic error= anything that doe not occur by chance!)

The more subjective the intervention, the more important the blinding!

Bias can occur at any stage of the study: patient assignment, data collection, event ascertainment…

I-25

Randomized Clinical Trials
To Blind or not to Blind?

- Unblinded trial
  - May be the only option: strategies of treatment (drug vs. surgery) behavioral interventions…
  - "True" blinding may be hard: expected biological effect of intervention
  - Easier to carry out and less expensive but…

  Risk of bias generally outweigh benefits!
- Alternative to blinding intervention (if not possible): blind outcome assessment

I-26

Introduction to Randomized Clinical Trials
Outline I

- **Historical perspective**
- **Rationale for randomized clinical trials**
  - Rationale for randomization
  - The equipoise issue
  - To blind or not to blind?
- **Key issues in the design of a RCT:**
  - What is the study question? Defining hypothesis, objectives and end-points
  - Defining selection criteria: generalizability vs. homogeneity
  - Selecting the control group: the placebo vs. "usual care" issue

I-27

Elements of a RCT

What is the Study Question?

From Furberg BD and Furberg CD. Evaluating Clinical Research. Springer Ed.

I-28

Elements of a RCT

What is the Study Question (Who-What-When)?

- Primary question tests the hypothesis
- Hypothesis must include:
  - *Population studied*
  - *Primary outcome of interest*
  - *Intervention studied*
  - *Period of observation*
- Objective: phrase the research question in concise, quantitative terms

I-29

Elements of a RCT

Primary and Secondary Objectives

- Primary objective needs to be defined (determine sample size)
- Secondary objective needs to be:
  - *Defined a priori (avoid post hoc "fishing expedition")*
  - *Chosen parsimoniously (avoid false positive)*
- Primary vs. secondary:
  - *Question of greatest interest/relevance*
  - *Consider feasibility (e.g. mortality vs. morbidity)*

I-30

Elements of a RCT

The Endpoints

- Quantitative measurement required by the objectives (= outcome, response variable)
- Event/condition the trial is designed to ameliorate, delay, prevent…
- Primary endpoint: need to be clearly and rigorously defined (*what is survival?)*
- Endpoints defined by type of measurement used:
  - Discrete, dichotomous *(dead or alive?),* count
  - Continuous *(BP change),* ordered (*toxicity)*

I-31

Elements of a RCT

The Endpoints: what makes a good Primary Endpoint?

- Must answer the primary question (*Co-primary?)*
- Frequency of occurrence must be known in control (determine sample size)
- Must be able to estimate treatment effect: clinical relevance (*minimum desired effect to change practice?)*
- Must be assessed/evaluable in all participants
- Can be measured accurately/reliably/objectively
  - Blinded randomization
  - Blinded assessment (*soft end point?*)
- All patients must be evaluated (no post randomization exclusion/no lost to follow up)

I-32

Elements of a RCT

Other Types of Endpoints

- Intermediate and surrogate
- Combined
- See Part V

I-33

Elements of a RCT

Defining the Study Population

- Subset of population with disease/condition of interest
- Patients enrolled = subset of study population defined by the eligibility criteria
- Inclusion criteria: Define "at risk" population
  - Less inclusive (= more homogeneous population): potential for benefit increase
    - but need to understand mechanism of action of intervention
    - Cannot generalize to other "subgroups"
  - More inclusive (= more heterogeneous population):
    - Increase generalizability
    - But may dilute effect of intervention (increase sample size)
  - Select group more likely to benefit from intervention
    - Higher risk: increase number of events, decrease sample size
    - But: are results applicable to lower risk?

I-34

Elements of a RCT

Defining the Study Population

- **Exclusion criteria:**
  - Insure patient safety (risk/benefit in specific subgroups)
  - Assess competitive risk
  - Assess likelihood of adherence to protocol and intervention

Eligibility criteria will be defined by goal of trial: efficacy vs. effectiveness trial?

I-35

Elements of a RCT

Defining the Study Population: Homogeneity vs. Generalizability

Homogeneity
- **Divergent subgroup of patients (i.e., "atypical" patients) can distort findings for the majority**
- **Restriction of population reduces "noise" and allows study to be done in a smaller sample size**

→ Restrict population to homogenous group

Generalizability
- **At the end of the study, it will be important to apply findings to the broad population of patients with the disease**
- **It is questionable to generalize the findings to those excluded from the study**

→ Have broad inclusion criteria "welcoming" all

From: Virgina Howard

I-36

Elements of a RCT

Defining the Study Population: Efficacy vs. Effectiveness trial

Characteristic	Efficacy Trial	Effectiveness Trial
Goal	Test biological question	Assess "real life" effect of intervention
No participants	< 1,000	> 10,000
Cohort	Homogeneous	Heterogeneous
Data collection	Extensive	Limited
Focus of inference	Internal validity	Generalizability
Eligibility criteria	Strict	Broad

From: Steven Piantadosi. Clinical Trials. A Methodologic Perspective. John Wiley & Sons, Inc 1997

I-37

Elements of a RCT

Choosing an Effectiveness Design

- **Define the *question*: What is the purpose of the trial?**
  - Does the intervention work when applied in usual practice?
- **Define the *setting*: under which conditions will the trial results be applicable?**
  - Ideal setting vs. normal practice?
- **How are *participant*s selected?**
  - Eligibility criteria mostly defined by the condition of interest
- ***Outcome*s of interest?**
  - Direct relevance to practice
  - Will influence clinical decisions and/ health policy decisions

I-38

Randomized Clinical Trials
Nature of "Intervention"

- Drug (or drug regimen)
- Surgical procedure
- Medical device
- Therapeutic modality (radiation, biologic therapy, etc)
- Diet
- Behavioral intervention (education)
- Clinical approach to diagnosis, treatment, symptom management, palliative care, etc. (e.g. strategy)

The common denominator: there is a choice between two alternative approaches; uncertain which is preferable (e.g. equipoise)

I-39

Randomized Clinical Trials
Selecting the Control Group

- Four different types:
  - Placebo
  - No Treatment
  - Different doses or regimens of the treatment
  - Different active treatment (including usual care)
- Control group will be classified based on:
  - Type of treatment used
  - Method of assignment in control group
  - May be more than one control!

I-40

Randomized Clinical Trials
Selecting the Control Group: The Placebo Issue

- **Definitions**
  1. **Clinical:** "A substance having no pharmacological effect but given merely to satisfy a patient who supposes it to be a medicine"

     Goal: to distinguish pharmacological effects from the effects of suggestion
  2. **Research:** "A substance having no pharmacological effect but administered as a control in testing experimentally or clinically the efficacy of a biologically active preparation."

     Goal: to obtain an unbiased assessment of the result of an experiment

I-41

Randomized Clinical Trials
Placebo Control: Scientific Justification

- Minimize subject and investigator bias (when used with randomization and blinding)
- Maximize likelihood of establishing efficacy: encourage optimal conduct of the trial: decrease "incentive" for poor trial conduct (drop-outs, cross-overs, etc)
- Enable distinction between adverse effects of drug/intervention and disease
- Allow for measurement of true effect size: account for the *"placebo effect"*

I-42

Randomized Clinical Trials
The Active Control

- **Positive control**: new therapy compared to known active therapy (randomized, can be blinded)
  - Goal: effectiveness or non-inferiority
  - Based on assumption that previous treatment shown to be effective! (external validation needed)
- **Challenges:**
  - Effect size and safety assessment more difficult
  - Larger sample size
  - Many possible bias: non adherence, concomitant therapies, randomization of inappropriate patients

I-43

Randomized Clinical Trials
Usual Medical Care as Control Group

- State of equipoise: is there a “standard of care”?
- Potential advantages:
  - Increase relevance
  - Increase external validity
  - Increase practicality
- Interpretation of evidence:
  - Is usual care validated by research? Is there a consensus on what is “usual care”?
  - Adherence to guidelines/evidence-based care?

I-44

Introduction to Randomized Clinical Trials
Outline II

- **The different phases of a RCT**
- **Basic RCT Designs**
  - Parallel, factorial, cluster and cross-over designs
  - Large Simple Trials
  - Comparative effectiveness trials
  - Superiority, Equivalence and Non-Inferiority trials
- **Key elements of a RCT Protocol**
- **Some ethical considerations**
  - Informed Consent Process
  - Patient safety issues

I-45

The Different Clinical Trial Phases
Phase I

- First in humans
- Small, uncontrolled
- Healthy volunteers/failed conventional therapy
- Dose-escalation protocols
- Tolerability/toxicity study: Maximum Tolerated Dose (MTD)
- Dose-response models

I-46

The Different Clinical Trial Phases
Phase II

- Test biologic activity/effect
- Estimate rates of adverse events
- Performed in patients with disease/condition of interest
- With or without comparison group
- Strict eligibility criteria

I-47

The Different Clinical Trial Phases
Phase II

- Phase IIa
  - Small scale feasibility studies
  - Intermediate endpoints
- Phase IIb
  - Comparative, randomized
  - Intermediate endpoints

I-48

The Different Clinical Trial Phases
Phase III

- Determine the effectiveness (overall benefit/risk-cost assessment) of new therapies relative to standard therapy
- Large sample size
- Multicenter
- Superiority, equality, equivalence or non inferiority

I-49

The Different Clinical Trial Phases
Descriptive Terminology

- **Early phase/development**:
  - Translational trials (e.g. from lab to clinic)
  - Mechanistic trials
    - Treatment mechanism
    - Dose finding/dose ranging studies
- **Middle development**
  - Safety and activity: probability of benefit?
  - May be randomized (remove selection bias, temporal trends)
  - Intermediate/surrogate outcomes
  - Small sample size

I-50

The Different Clinical Trial Phases
Descriptive Terminology

- **Late development**
  - Comparative studies
    - Treatment efficacy (IIb/III)
    - "Pivotal" trials
    - Large scale/simple trials
    - Superiority or equivalence
  - Late Safety Studies
    - Estimate of incidence of rare serious side effects
    - Very large sample size
    - Causality inference?

I-51

The Different Clinical Trial Phases
Phase IV

- Long term surveillance studies ("post marketing") for safety
- New dosing regimens/indications
- Look for rare side effects
- Often non randomized

I-52

Introduction to Randomized Clinical Trials
Outline II

- **The different phases of a RCT**
- **Basic RCT Designs**
  - **Parallel, cross-over, factorial and cluster designs**
  - **Large Simple Trials**
  - **Comparative Effectiveness Trials**
  - **Superiority, Equivalence and Non-Inferiority trials**
- **Key elements of a RCT Protocol**
- **Some ethical considerations**
  - **Informed Consent Process**
  - **Patient safety issues**

I-53

Basic RCT Designs
Parallel Design

FREEDOM Design

Future REvascularization Evaluation in patients with Diabetes mellitus: Optimal management of Multivessel disease

Eligibility: DM patients with MV-CAD eligible for stent or surgery
Exclude: Patients with acute STEMI, cardiogenic shock

Randomized 1:1

MV-stenting With Drug-eluting stents And abciximab	**CABG With or without CPB**

I-54

Basic RCT Designs
Cross Over Design

- Participant = own control
- Randomize: order of treatment for each patient (e.g. AB vs. BA)
- Advantages
  - Reduce variability → Reduce Sample Size
  - Detect difference in response in individual patient
- Disadvantages
  - Order of treatment should not matter
  - No carry over of effect → test for interaction

I-55

Basic RCT Designs
Factorial Design

Intervention A

Intervention B	Active (A)	Control (A)
Active	a	b
Control	c	d

Cells

a= Active A + Active B
b= Control A + Active B
c= Active A + Control B
d= Control A + Control B

Analysis of a 2 x 2 factorial RCT

Effect of A: ac vs. bd *
Effect of B: ab vs. cd *

*If no treatment interaction

I-56

Basic RCT Designs
Factorial Design

	BP: Intensive (SBP<120)	BP: Standard (SBP<140)	Lipid (statin + fibrate vs statin + placebo): Group A	Lipid: Group B	
Intensive Glycemic Treatment (A1C<6%)	1178	1193	1383	1374	5128*
Standard Glycemic Treatment (A1C 7-7.9%)	1184	1178	1370	1391	5123*
	2362*	2371*	2753*	2765*	10,251

*Primary analyses compare marginals for main effects

ACCORD (Action to Control Cardiovascular Risk in Diabetes)
(ACCORD Study Group, Am J Cardiol 2007;99[suppl]:21i-33i)

I-57

Basic RCT Designs
Factorial Design

- **Advantage:**
  - Two trials for (almost) the price of one
  - Design is best if: two intervention have different mechanisms of actions or different outcomes (e.g. cancer for A and CV disease for B)
- **Disadvantages:**
  - Need to test for possibility of interaction (e.g. A differs based on the presence or absence of B)
  - Test for interaction not very powerful
  - Need to consider gain in cost vs. increased complexity, recruitment and adherence issues and potential for adverse events

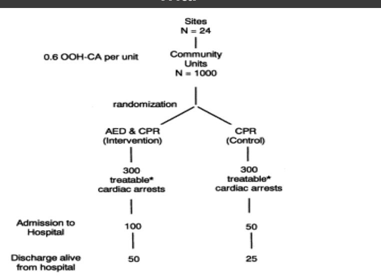
I-58

Basic RCT Designs
Cluster Design

- Cluster design= group randomization
- Group= schools, clinics, villages…
- Sample size: based on number of groups (not individuals)
  - Need to be adjusted by factor N_m (where N= number of cluster each of size m)
  - Need to take into account within-cluster correlation of response (correlation= loss of efficiency)
- Analysis:
  - Cannot use classic statistical methods (correlation)
  - Random effect model
  - Use sensitivity analyses

I-59

Basic RCT Designs
Cluster Design: The Public Access Defibrillation (PAD) Trial



*Treatable arrests are considered to be those witnessed or discovered shortly after collapse.

I-60

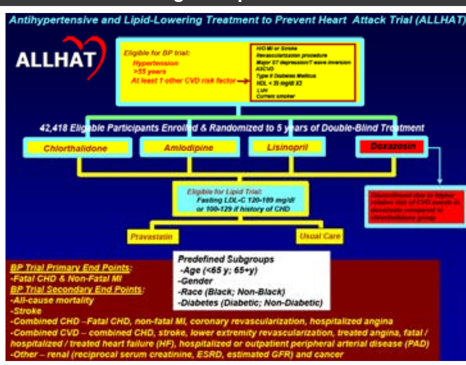
Resuscitation. 2003 Feb;56(2):135-47

Basic RCT Designs
Large Simple Trials

- Provide a more reliable estimate of the effect of intervention
- Needed to uncover smaller treatment effects
  - *That are important in common conditions*
- Increase generalizability
  - *But limit data collection/subgroups and secondary analyses*
- Decrease cost by simplifying design and management
  - *But need strong randomization procedures and reliable outcomes assessment*

I-61

Basic RCT Designs
Large Simple Trials



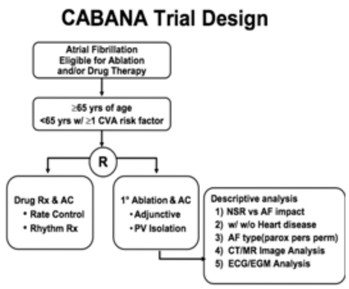
I-62

Basic RCT Designs
Comparative Effectiveness Trials

- A type of health care research that compares the results of one approach for managing a disease to the results of other approaches.
- Comparative effectiveness usually compares two or more types of treatment, such as different drugs, for the same disease. Comparative effectiveness also can compare types of surgery or other kinds of medical procedures and tests.
- Comparative effectiveness research is designed to inform health care decisions by providing evidence on the effectiveness, benefits, and harms of different treatment options. The evidence is generated from research studies that compare drugs, medical devices, tests, surgeries, or ways to deliver health care.

I-63

Basic RCT Designs
Comparative Effectiveness Trials



I-64

Basic RCT Designs
Superiority, Non Inferiority and Equivalence Trials

- **Superiority trial**
  - Is (new) intervention better than no (placebo) intervention or standard intervention?
  - Goal: Demonstrate a difference!
- **Non inferiority trial**
  - Is new intervention *not worse* than standard? (not less effective, but safer, cheaper, etc.)
  - Goal: Demonstrate that new intervention is not worse than the standard by a prespecified Δ (minimum clinically significant difference)
- **Equivalence trial**
  - Are the effects of the two interventions very similar?
  - Goal: Demonstrate that the two interventions are not different by more than the prespecified Δ

I-65

Introduction to Randomized Clinical Trials
Outline II

- **The different phases of a RCT**
- **Basic RCT Designs**
  - **Parallel, cross-over, factorial and cluster designs**
  - **Large Simple Trials**
  - **Comparative Effectiveness Trials**
  - **Superiority, Equivalence and Non-Inferiority trials**
- **Key elements of a RCT Protocol**
- **Some ethical considerations**
  - **Informed Consent Process**
  - **Patient safety issues**

I-66

Key elements of a RCT Protocol
Study Design: Preliminary Considerations

- Demonstrate need for trial
- Establish study objectives
- Choose best approach to problem/question
  - Small vs. large?
  - Less is more!
- Objectives ≠ study goals
  - Objectives: statement about question to answer
  - Goals: what you need to achieve to answer the question

I-67

Key elements of a RCT Protocol
Study Design: Framing the Question

- Toxicity? Efficacy? Effectiveness?
- Feasibility
- Proof of concept
- Pilot study

Why? How? Importance?

→ Outcome

I-68

Key elements of a RCT Protocol
Study Design: Key Steps to Follow

1. Establish study objectives
2. Choose basic study design
3. Determine primary and secondary outcomes
4. Choose type of control
5. Determine need/feasibility of blinding
6. Choose randomization procedure
7. Sample size and power
8. Determine screening, baseline, treatment and follow-up periods
9. Choose patient population
10. Establish treatment modalities

I-69

Elements of a RCT

Protocol: Table of contents (I/IV)

I-70

Elements of a RCT

Protocol: Table of contents (II/IV)

I-71

Elements of a RCT

Protocol: Table of contents (III/IV)

I-72

Elements of a RCT

Protocol: Table of contents (IV/IV)

I-73

Introduction to Randomized Clinical Trials

Outline II

- **The different phases of a RCT**
- **Basic RCT Designs**
  - Parallel, cross-over, factorial and cluster designs
  - Large Simple Trials
  - Comparative Effectiveness Trials
  - Superiority, Equivalence and Non-Inferiority trials
- **Key elements of a RCT Protocol**
- **Some ethical considerations**
  - ***Informed Consent Process***
  - ***Patient safety issues***

I-74

Ethical Issues Specific to Clinical Trials

- Special ethical concerns because treatment is determined by chance
- The arms of the clinical trial must be in clinical equipoise
- Principle of non maleficence, withholding proven treatment from control group
- Periodic analysis of interim data by independent Data and Safety Monitoring Board

I-75

Some Ethical Considerations
Informed Consent Process

- Purpose of the trial
- Nature of the trial
- Procedures of the trial
- Risks and potential benefits and alternatives to participating
- Procedures to maintain confidentiality
- Assurances and contact information

I-76

Some Ethical Considerations
Informed Consent Issues

- Withdrawal
  - Participant is free to withdraw at any time
- New findings
  - Obligation to tell participant of any significant new findings that may affect his/her willingness to continue
- Potential for coercion

I-77

Some Ethical Considerations
Health Information Portability and Accountability Act (HIPAA)

- Research subjects must sign an authorization form that describes the use and disclosure of their protected health information (PHI) for research purposes
- HIPAA authorization wording may be part of informed consent document or a separate form
- Subject must be given signed copy of form with HIPAA disclosure information
- http://privacyruleandresearch.nih.gov/

I-78

Some Ethical Considerations
Where to Go for More Info

- **Human Subjects Research Protection**
  - http://www.hhs.gov/ohrp/
  - Training: http://phrp.nihtraining.com/users/login.php
- **Registry of clinical Trials and Background:**
  - http://clinicaltrials.gov/
- **Regulations and Ethical Guidelines:** http://ohsr.od.nih.gov/guidelines/index.html
  - 45 CFR 46 Protection Of Human Subjects
  - Guidelines for Conduct of Research Involving Human Subjects at NIH (Gray Booklet) (pdf file)
  - The Belmont Report Ethical Principles and Guidelines for the Protection of Human Subjects of Research
  - Nuremberg Code Directives for Human Experimentation
  - World Medical Association Declaration Of Helsinki
- **NIH bioethics Resources:** http://bioethics.od.nih.gov/index.html

I-79

Randomized Clinical Trials
Some key Points

- **Important**
  - in evaluating interventions for the prevention, diagnosis, and treatment of disease
  - Important to obtain unbiased comparisons of interventions
- **Ethical**
  - in the presence of uncertainty (equipoise)
  - present the best choice of therapeutic options to the patients
- **Robust**
  - large trials recommended to increase reliability
- **Applicable to studies of efficacy and of effectiveness**
- **Can answer more than one question at a time (factorial trials and other designs)**
- **In some situations, can randomize entire groups (e.g., communities, medical practices)**

I-80

Randomized Clinical Trials
Some Key References

- Fundamental of Clinical Trials. Lawrence M Friedman, Curt D Furberg, David L DeMets. Springer Verlag Editors
- Clinical Trials: Design, Conduct and analysis. Curtis L Meinert. Oxford University Press
- Successful randomized trials. A Handbook for the 21th Century. Michael Domanski, Sonja McKinlay. Lippincott Williams & Wilkins
- Principles and Practice of Clinical Research. John I Galin. Academic press
- Guide to Clinical Studies and Developing Protocols. Bert Spilker. Raven press
- Clinical Trials. A Methodological Perspective. Steven Piantadosi. John Wiley & Sons, Inc.

I-81

Part II: Project Management in Clinical Trials

SCT Pre-Conference Workshop
Essentials of Randomized Clinical Trials

Dixie Ecklund, RN, MSN, MBA

Associate Director
Clinical Trials Statistical & Data Management Center
Department of Biostatistics
College of Public Health
University of Iowa

II-1

Project Management in Clinical Trials

- Requirements for Clinical Trials vary widely which drives the Project Management model
  - Big Pharma clinical trials
    - Initiated, developed, and managed by Industry Sponsor (e.g. Merck)
    - Data Coordinating Center (DCC)
    - Clinical Coordinating Center (CCC)
    - Statistical Coordinating Center (SCC)
    - Participating Clinical Centers (PCC)
    - Sponsor

II-2

Project Management in Clinical Trials

- Requirements for Clinical Trials vary widely which drives the Project Management model
  - Industry or Federally-funded clinical trials (U01)
    - Initiated and developed by Industry Sponsor or NIH-funded Principal Investigators
    - Managed by Contract Research Organization (CRO) or Academic CRO
    - Data Coordinating Center (DCC)
    - Clinical Coordinating Center (CCC)
    - Statistical Coordinating Center (SCC)
    - Participating Clinical Centers (PCC)
    - Sponsor

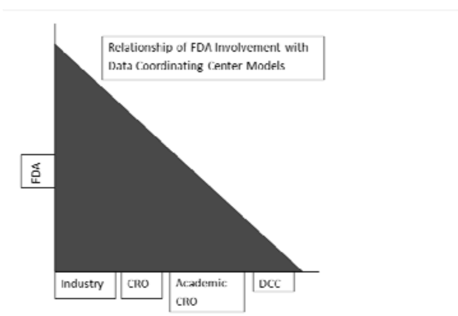
II-3

Project Management in Clinical Trials

- Requirements for Clinical Trials vary widely which drives the Project Management model
  - Federally-funded clinical trials
    - R01 Grants
    - Managed by Academic CRO or Traditional Data Coordinating Center (DCC)
    - Data Coordinating Center
    - Statistical Coordinating Center
    - Interact with CCC and Sponsor
    - Important to define who is doing what

II-4

Project Management in Clinical Trials



II-5

Project Management in Clinical Trials

- **Which model of Data Coordinating Center?**

"PROTHERO TOOK THE CONCEPT OF A FULL SERVICE CRO A BIT TOO FAR."

II-6

DCC Requirements to Successfully Manage Multi-Site Clinical Trials

- Build Good Teams
- Phase I: Grant/Protocol Development
- Phase II: Implementation
- Phase III: Up and Running
  - Study Start-Up Activities
- Phase IV: Ongoing Activities
  - Study Continuation
- Phase V: Study Close-Out

II-7

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Biostatistics***
    - Protocol Development
    - Statistical Analysis Plans
    - Report Generation
    - Interim Analysis
    - Final Analysis

II-8

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Protocol Coordinators***
    - Clinical Coordinators
    - Manage sites
    - Manage and resolve data queries
    - Develop study materials
    - Maintain study supplies

II-9

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Data Managers***
    - Technical Coordinators
    - Develop User's specifications for data entry systems
    - Develop testing plans for data entry systems
    - Validate data entry systems
    - Documentation of validation

II-10

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Information Technology (IT) Developers***
    - Develop Web Applications
    - Develop Data Entry Applications
    - Data Storage
    - Data Back-up and Recovery
    - 21 CFR Part 11 Compliance

II-11

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Regulatory***
    - Responsible for Trial Master File
    - Monitor Site Regulatory Binders
    - IND Safety Reports
    - MedDRA Coding
    - FDA Submissions

II-12

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Fiscal/Administrative***
    - Develop grant budgets
    - Monitor expenditures
    - Human Resource functions
    - Coordinate meeting and travel arrangements

II-13

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Medical Monitors***
    - MedDRA coding
    - Medical writing
    - Aggregate review of Adverse Events
    - Individual review of Serious Adverse Events

II-14

Bring Together a Good Team

- **Data Coordinating Center (DCC) Teams**
  - ***Quality Management***
    - Backbone of all processes
    - Develop and monitor SOPs
    - Standardize training/education
    - Develop center-wide metrics to monitor quality
    - Develop study-specific metrics to monitor quality

II-15

Bring Together a Good Team

- **Clinical Coordinating Center (CCC) Team**
  - Lead Principal Investigator (PI)
  - Lead Study Coordinator
  - Support Staff
- **Participating Clinical Center (PCC) Teams**
  - Site PI
  - Site Study Coordinator
  - Support Staff
- **Sponso**r
  - NIH
  - Foundations
  - Industry

II-16

Glue the teams together

- Written Standard Operating Procedures (SOPs)
- Written Study-Specific Project Work Instructions (PWIs)
- Training and education programs
  - Cross-train whenever possible
- Quality Management initiatives

II-17

The "Reality" in Clinical Trials

II-18

Phase I: Development

- **Protocol Development**
  - PIs for scientific and medical input
  - Biostatisticians for design and analysis input
  - Study Coordinators for practical input
  - Medical Writer to help with readability

II-19

Phase I: Development

- **Study Materials Development**
  - Investigator Brochure (IB) for IND/IDE studies
  - Manual of Procedures (MOP)
  - Laboratory Manuals
  - Informed Consent Templates
  - Source Documents
  - Web and Data System User's Guides
  - Adverse Event System User's Guides
  - Specimen Tracking System User's Guides

II-20

Phase I: Development

- **Develop Safety Monitoring Plan**
  - Identify Medical Monitor(s)
  - Determine level of reporting required
  - Adjudication of events between Medical Monitors

II-21

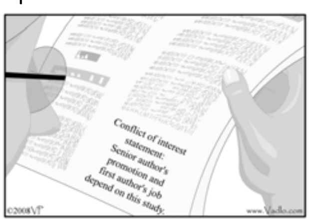
Phase I: Development

- **Develop On-site Monitoring Plan**
  - 100% informed consents
  - 100% inclusion/exclusion criteria
  - Random selection of % of subjects enrolled
  - Site regulatory files

II-22

Phase I: Development

- **Select Qualified Investigators**
  - Search FDA warning letters for debarred investigators
  - Appropriate clinical expertise
  - Adequate staff to perform studies
  - Adequate facilities to perform studies
  - Pool of eligible subjects
  - Conflict of Interest Disclosures



II-23

Phase I: Development

- **Select Qualified Subcontractors**
  - ***On-site Monitoring***
    - Performed by DCC or Contract Research Organization (CRO)
    - Qualified Data Auditors
    - Qualified CRAs
    - Adequate personnel to meet the requirements of the monitoring plan

II-24

Phase I: Development

- **Select Qualified Subcontractors**
  - ***Specimen kit assembly and distribution***
    - Configure specimen kits
    - Supplies on-hand to manufacture kits
    - Ability to meet deadlines to manufacture specimen kits
    - Ability to distribute specimen kits and shipping supplies
    - Ability to collaborate with Specimen Tracking System (if in place)

II-25

Phase I: Development

- **Select Qualified Subcontractors**
  - ***Central laboratories***
    - Appropriate certifications
    - Ability to handle throughput
    - Specialty laboratory specific to research question
    - Able to provide results to Clinical Centers
    - Able to provide results via data transfer to DCC
    - QC processes in place

II-26

Phase I: Development

- **Select Qualified Subcontractors**
  - ***Drug distributor***
    - Receive Investigational Products from Manufacturers
    - Receive approved drugs obtained through Clinical Trial Agreements or purchase
    - Label Investigational Products
    - Appropriate storage facilities
    - Appropriate inventory support
    - Appropriate distribution processes
    - Ability to ship out of country if needed
    - Ability to accept returned products
    - Ability to destroy expired or returned products

II-27

Phase II: Implementation

- **FDA submission for IND/IDE approval**
  - Work with Regulatory Team and Sponsor
  - Assistance from CTSA staff may be available
- **Establish and maintain Trial Master File**
  - May be held by Sponsor
- **Develop Site Regulatory Binders**
  - Prepare tabs and binders

II-28

Phase II: Implementation

- **Collect, QA and Monitor Site Regulatory Documents**
  - 1572s
  - Delegation of Responsibility Log
  - Investigator CVs
  - Investigator Licenses
  - Laboratory Certifications
  - Laboratory Normal Ranges
  - IRB approvals

II-29

Phase II: Implementation

- **Investigator Meeting**
  - Protocol Finalization
  - Procedural discussions
  - Recruitment goals
- **Coordinator Training**
  - May be held in conjunction with Investigator Meeting

II-30

Phase II: Implementation

- **Develop recruitment plan and materials**
  - Identification of Investigators
  - Public website (clinical trials.gov)
  - Call Centers
  - Brochures
  - Google ad campaign
  - Television/radio spots
  - Newspaper advertisements

II-31

Phase II: Implementation

- **Execute subcontracts**
  - This can be a very lengthy process
  - Legal talking to Legal…
- **Prepare for initial DSMB meeting for protocol approval**

II-32

Phase II: Implementation

- **Develop and validate data entry system**
  - Case Report Forms
  - Data Management Plan
  - User's specifications
  - Testing plans
  - Validation documentation

II-33

Phase III: Implementation

- **Develop report shells**
  - Enrollment report
  - Ineligibility report
  - Adverse events/Serious Adverse Events
  - Protocol deviations
  - Missing data
  - Study-specific reports

II-34

Phase III: Up and Running

- **Site Initiation Visits**
  - An opportunity to begin a study on the right path
  - May be done in person or through teleconference or webinars
  - Important agenda items
    - Protocol training
    - Good Clinical Practices
    - Study Coordinator Training on Procedures
    - Data Entry Training and Certification
    - Review of Facilities (if not previously done)

II-35

Phase III: Up and Running

- **Monitor Site IRB approvals**
  - Activate sites when approvals received

II-36

Phase III: Up and Running

- **Develop User Access policies for the Web and Data Entry Systems**
  - Set up user accounts
  - Verify users through Delegation of Responsibilities Log
- **Develop ongoing study training materials**
  - Webinars
  - ppt. presentations
  - Revisions to MOP

II-37

Phase III: Up and Running

- **Develop Statistical Analysis Plan (SAP)**
  - Submit to FDA
- **Distribute study supplies after site activation**
  - Study drug
    - Investigational and approved products
  - Specimen collection kits
    - Kits and shipping supplies
  - Study supplies and equipment
    - Study-specific (e.g. Blood Pressure monitors, EKG machines, Glucometers, etc.)

II-38

Phase IV: Ongoing Activities

- **Protocol amendments**
  - Submit to FDA
  - Submit to IRBs
- **Monitor IRB approvals and renewals**
- **Monitor recruitment, retention and adherence**
  - Site performance tracking tools

II-39

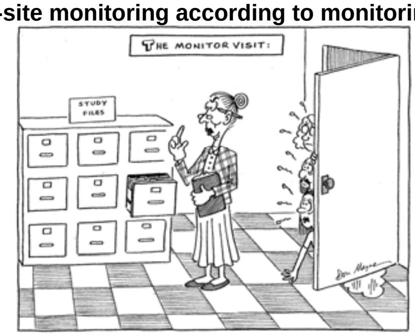
Phase IV: Ongoing Activities

- **Monitor data entry for timeliness**
  - Query resolution
  - Missing data
    - Reports to monitor missing data
    - Work closely with study coordinators to receive all obtainable data
- **Ongoing collection and QA of regulatory documents**

II-40

Phase IV: Ongoing Activities

- **On-site monitoring according to monitoring plan**



II-41

Phase IV: Ongoing Activities

- **On-site monitoring according to monitoring plan**
  - Provide CRA with data listings
  - Source document verification
  - Tools for resolving data discrepancies
  - Monitor drug accountability logs
  - Monitor site regulatory documents
  - Montior Adverse event/Serious adverse event reporting
  - Monitor protocol deviation reporting

II-42

Phase IV: Ongoing Activities

- **Distribute study drugs and supplies**
  - Monitor site utilization
  - Monitor expiration dates
  - Establish trigger points for re-order
- **Monitor drug and supply accountability logs**
  - Internal DCC monitoring to ensure sites don't run out of drug or supplies

II-43

Phase IV: Ongoing Activities

- **Site retraining on protocol and procedures**
  - Study coordinator turnover
  - One on one webinars
  - On-site training
  - Training for cause

Catherine's study team had changed so many times, she'd done more staff inductions than site initiations!

Phase IV: Ongoing Activities

- **Data entry system enhancements**
  - Initial version released at study start-up
  - Don't get caught up in never-ending tweaking of the data entry system
  - Develop enhancements in batches and release preferably no more often than quarterly
  - Hot fixes only for bugs that prohibit data entry
  - Change management software is very useful

II-45

Phase IV: Ongoing Activities

- **Safety Review**

II-46

Phase IV: Ongoing Activities

- **Safety Review**
  - Definitions for AEs/SAEs
    - NCI Common Toxicity Criteria
  - Who determines relatedness?
  - Who determines expectedness?
  - Who determines need to expedite?
  - Who does the expedited safety reporting
  - MedDRA coding
    - Training of coders
    - Agreement of coders
  - Monitor follow-up until resolution
    - Aggregate review of AEs for trends
  - Write safety narratives for Annual Report

II-47

Phase IV: Ongoing Activities

- **Import central laboratory data into database**
  - Establish standardized data sets
  - Establish timeline for data transfers
  - Upload comma-delimited files
  - Email vs. encrypted transfer

- **Interim analyses as described in SAP**
  - Monitor for stopping rules
  - Futility analysis

II-48

Phase IV: Ongoing Activities

- **FDA Annual Report**
  - Determine due date and data lock date
  - Describe protocol activity
  - Describe safety profile
  - Describe new findings

II-49

Phase IV: Ongoing Activities

II-50

Phase IV: Ongoing Activities

- **Lost to Follow-up**
  - Important to make all efforts to obtain endpoint data
  - May need IRB approval to make final contact

II-51

Phase IV: Ongoing Activities

- **Investigator Payments**
  - Based on recruitment activities
  - Based on data entry completion
- **Newsletters**
  - Keep sites informed year-round of study activities
  - Relevant "hot" topics
  - Updates from study PI
  - Updates from DCC
  - Recruitment tips

II-52

Phase IV: Ongoing Activities

- **Update Investigator's Brochure**
  - Review safety profile
  - Submit to FDA
  - Submit to IRBs
- **Annual Investigator/Study coordinator meeting**
  - Opportunity to kick-start recruitment
  - Opportunity for training
    - Protocol
    - Procedures
    - Data entry
  - Collaborate on publications

II-53

Phase IV: Ongoing Activities

- **DSMB Meetings**
  - Follow the charter
  - Teleconference or face-to-face?
  - Establish data lock dates
    - Give coordinators plenty of notice to complete data entry
  - Establish timeline after data lock to:
    - Run reports
    - QA reports
    - Print reports
    - Distribute reports

II-54

Phase IV: Ongoing Activities

- **Steering Committee Meetings**
  - Establish frequency of meetings
  - Teleconference or face-to-face?
  - Steering Committee Chairperson
  - Who are the voting members and what constitutes a quorum
  - Who has responsibility for:
    - Agenda items
    - Minutes
    - Action items

II-55

Phase IV: Ongoing Activities

- **Subcommittee Meetings**
  - Establish frequency of meetings
  - Teleconference or face-to-face?
  - Committee Chairperson
  - Who has responsibility for:
    - Agenda items
    - Minutes
    - Action items

II-56

Phase V: Study Close-Out

II-57

Phase V: Study Close-Out

- **Site Close-Out Visits**
  - Final On-Site Monitoring Visit
  - Reconcile Files
  - Final Drug and Supply Accountability
  - Close-Out Letter

II-58

Phase V: Study Close-Out

- **Resolve all queries and data issues**
- **Data lock**
- **Return all unused study drug (if applicable)**
  - Will there be ongoing study drug treatment?
  - When will subjects/sites be unblinded?
- **How (if) are subjects informed of results?**
  - Keep IRB open at sites if recontact is anticipated
- Analysis programs are developed and debugged

II-59

Phase V: Study Close-Out

- Final analysis done per SAP
- Reports are written, reviewed, and accepted by Steering Committee
- Sponsor and FDA receive final reports
- Publications
  - Lead authors determined through Publication Policy
  - DCC assists with additional analyses as requested
- Submit data sets to clinical trials.gov

II-60

Overall Project Management Tips

- **Develop good teams and working relationships**
  - Identify the Project Champion
  - Provide the teams with the tools and training to successfully accomplish their goals
  - Monitor for meetings that have served their purpose and should be discontinued
  - Monitor for redundancy in meetings
  - Acknowledge and reward exceptional behavior
  - Find the strengths in each team member

II-61

Overall Project Management Tips

- **Ensure that all team members are aware of areas of responsibilities**
  - Never walk out of a meeting without a clear understanding of the deliverables
  - Or who is responsible for the deliverable
  - Or what the expected timeline is for the deliverable

II-62

Overall Project Management Tips

- **Require documentation for all proceedings**
  - Don't rely on memory for previous decisions
  - Distribute minutes and action items after all meetings
  - Post minutes and action items in a shared drive or on the web

II-63

Overall Project Management Tips

- **Be flexible when needed**
  - Good communication will reveal problem areas
  - Must always be willing to re-examine and reprioritize
  - Be willing to look at things from a different viewpoint
  - Solicit input from the staff regularly
  - Disagreement can be healthy if handled well

II-64

Overall Project Management Tips

- **Follow-up on progress**
  - Hold team members accountable for timelines
  - Expect progress reports on regular intervals
  - Look for ways to improve efficiencies
  - Look for ways to maintain staff satisfaction
  - Have some fun along the way!

II-65

Conclusion

- There are many components to juggle in clinical trials research
- Good project management makes clinical trials research more easily accomplished
- Clinical trials work can be very rewarding

II-66

Part III: Data Collection, Reporting, and Quality Control Issues

Laura Lovato, MS

Senior Biostatistician
Department of Biostatistical Sciences
Wake Forest University School of Medicine

SCT Pre-Conference Workshop
Essentials of Randomized Clinical Trials

III-1

Data Collection, Reporting, and Quality Control Issues

Learning Objectives

- GCPs, QC, QA, SOPs
- Primary sources of error in data collection
- Steps in Data Collection
  - Design of data collection forms
  - Standardization of procedures
  - Types of data entry/management systems
- Quality control methods and reporting

III-2

Introduction

"No study is better than the quality of its data."
-Friedman, Furberg and DeMets

"To err is human."

III-3

Introduction: Guidelines for Good Clinical Practice

- Unified standard
- For design, conduct, analyses and reporting of clinical trials that involve human subjects
- To ensure that patients' rights, safety and confidentiality are protected
- To promote scientific validity and data integrity

III-4

Introduction: Specific Principles of GCP Applicable to Data Collection

- Confidentiality of records should be protected
- All clinical trial data should be handled in a way to ensure accurate reporting, interpretation and verification
- An audit trail should be maintained for changes/corrections to forms and electronic data

III-5

Introduction:

Great web sites – GCPs and SOPs

From the U.S. FDA:
http://www.fda.gov/oc/gcp/

From Wake Forest University:
http://www.wakehealth.edu/OR/Research-Monitoring---Oversight--Links-and-Resources.htm

III-6

Introduction: Data Collection and Quality Control

"Any procedure, method, philosophy that is aimed at maintaining or improving the reliability or validity of the data and the associated procedures used to generate them."

- Curtis Meinert

III-7

"Oh look, a data trail to follow."

III-8

Introduction: Quality Control (QC) vs Quality Assurance (QA)

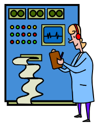
QC involves all process controls and monitoring performed by local staff on a day-to-day basis to maintain data quality

QA involves independent review or auditing of key processes to uncover and remedy problems

III-9

Primary sources of error in data collection process

- Missing data – incomplete or irretrievable
- Incorrect data – more difficult to recognize
- Excess variability – can reduce the opportunity to detect real change



III-10

Steps in Data Collection

- Define key variables
- Standardize & train on procedures (MOP)
- Data Collection
  - Acquisition
  - Recording
  - Entry
  - Study Closeout
- Preparation for analysis

III-11

Steps in Data Collection

Define Key Variables

III-12

Define key variables

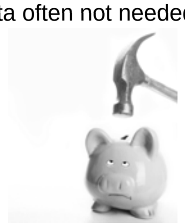
- Depends on trial type and outcomes
- At Baseline: characteristics of enrolled/non-enrolled participants related to major eligibility requirements
- Primary/Secondary outcome measures
- Variables that might confound/mediate/modify association
- Monitoring adherence to the protocol

III-13

Focus on key variables

Only important data should be collected

- As the volume of noncritical data increases, forms become burdensome and complicated leading to confusion
- Clinical care data often not needed as part of trial database



III-14

Steps in Data Collection

Standardization and Training

III-15

Standardization & Training

Pre-trial Quality Control Activities:

- Obtain adequate resources
- Design of case report forms
- Pre-testing
- Design of data management system
- Manual of Procedures (MOP)
- Hiring qualified personnel
- Training and certification

III-16

Standardization & Training

Manual of Procedures
(prior to and during the study)

- Standardized procedures
- Clearly written, detailed instructions
- Timely updates and clarifications
- Accessibility is essential

III-17

II-18

Standardization & Training

Training and Certification

- ❑ Central, regional, or local
- ❑ "Train the trainer" model
- ❑ Use Audio-visuals
- ❑ Certification/recertification to maintain skills

III-19

Standardization & Training

Design of data management system

- ❑ Security features/protection of human subjects' rights (privacy and confidentiality)
- ❑ Controlled Access
- ❑ Identification and authentication

III-20

Standardization & Training

Design of data management system

- Data entry/editing capability
- Desirable features:
  - Ease of screen set up and use
  - Range, field type, skip pattern checks
  - Query system
  - Ability to accomodate double data entry
- Word processing or spreadsheet software not advocated

III-21

Standardization & Training

Design of data management system

- Web-based systems also have administrative functions
  - Communications hub,
  - Information/Resource Center,
  - Coordination of publications process,
  - Management of Adjudication System

III-22

Steps in Data Collection

Data Acquisition

III-23

Design of Case Report Forms

•Purpose:

•To collect complete and accurate data

•To ensure standardization and consistency

•In some cases, to reinforce the protocol

III-24

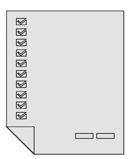
Design of Case Report Forms

- Clean, concise, consistent
- Well-organized with logical flow
- Few “write-in” or “text” answers
- No essay questions!

III-25

Design of Case Report Forms

- Selection of items to be collected
- Timing of visit schedule
- Ordering of Procedures



III-26

Steps in Forms Development

- Examination of Existing Forms (not necessary to “reinvent the wheel”)
- *Data Collection forms in Clinical Trials (Spilker B, Shoenfelder J, Raven Press, New York, 1991)*
- *Talk to someone at your institution/company that has done similar research*
- *Use the web – similar studies may have examples on the public side of their web sites*

III-27

Steps in Forms Development

- •Preparation of initial versions
- •Review by investigators, statisticians, clinic staff, and data management staff
- •Pilot-testing
- •Debriefing and revamping

III-28

Pre-Testing

- •Mock visits/procedures conducted
- •Simulation with practice participants
- •Debriefing is essential to improve procedures
- •Procedures/forms revised accordingly

III-29

Changes to Study Forms

- •Often done early on to improve data collection
- •Can be problematic when done repeatedly throughout the trial
  - •*Results in multiple versions of data sets*
  - •*Can increase risk of errors (clinic, data entry, analysis)*

III-30

Changes to Study Forms

Initial Version

Troponin results
1 At least 5x upper limit of normal
2 At least 2x upper limit of normal but less than 5x
3 Greater than upper limit of normal but less than 2x
4 Within normal limits

III-31

Changes to Study Forms

New Version

Troponin results
1 At least 5x upper limit of normal
2 At least 3x upper limit of normal but less than 5x
3 At least 2x upper limit of normal but less than 3x
4 Greater than upper limit of normal but less than 2x
5 Within normal limits

III-32

Changes to Study Forms

Initial Version

Time to Bed: [ ][ ] __p.m.

Time Arise: [ ][ ] __a.m.

Hours of Sleep: [ ] hours

III-33

Changes to Study Forms

New Version

Time to Bed: [ ][ ] (24 hour clock)

Time Arise: [ ][ ] (24 hour clock)

Hours of Sleep: [ ] hours

III-34

Steps in Data Collection

Data Recording

III-35

Data Recording

- Traditionally, refers to transcribing information onto case report forms (paper -> database)
- Trend toward direct computer entry with no prior hard copy, with no source document
  (e.g., iPad, accelerometers, pedometers Social networks, text messages, smart Phones, video game consoles, IRV)
- Both approaches depend on well-designed forms/data entry screens

III-36

Data Recording

- Direct computer entry:
  - •No source document
  - •Security Risks
    - •Devices could be stolen
    - •Not password protected
    - •Cashe
    - •Used in more public settings
  - •Who pays for device?
  - •Who is actually recording/receiving the information?

III-37

Data Recording: acceptable direct data transmissions

- Aggregate data
- Coded answers that do not describe (or contain metadata that describes) health information
- Health information by itself *without any of the 18 identifiers*
- Behavioral data
  - Food diaries, exercise logs, your 'MII' in WII
- Transmitted raw data without describing meta data
  - Ex. 5.5 is not PHI but HbA1c=5.5 is*
- Outward bound messages (e.g., exercise reminders)

Thanks to Scott Rushing for this slide

II-38

Steps in Data Collection

Data Entry

III-39

Data Entry

Types of traditional data entry systems

- Local
  - Data keyed onsite by clinic personnel
  - Potential for quick resolution of data omissions, errors, and inconsistencies
- Central
  - Forms mailed/faxed to sponsor or data coordinating center
  - Data entered by experienced keyers
  - Forms stored centrally.

III-40

Data Entry

Web-based data entry systems

- Provides flexibility
  - Data entry can be local or mix local/central
  - No specific hardware requirements
  - No specific software requirements for internet browser
- Secure link provided
- Data from multiple sources are consolidated on a central server

III-41

Data Entry

Web-based data entry systems

- Security features/protection of human subjects' rights (privacy and confidentiality)
- Controlled Access
- Identification and authentication
  - Requires valid user id and password
  - Password expire every 90 days
  - Specific access rights based on study function

III-42

Data Entry

Web-based data entry systems

- Audit trail
- Each and every access into the system is documented
- Every page that is accessed is documented
- All versions of any record entered are kept and date/time stamped (with user id)

III-43

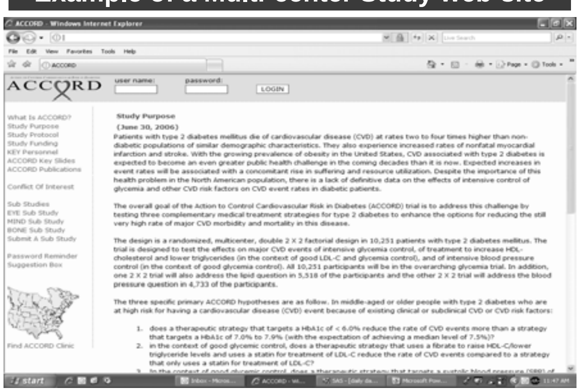
Data Entry

Web-based data entry systems

- Virus protection/scanning strategies to monitor and eliminate security threats
- Database server behind firewall
- Disaster recovery plan
- Regular backup for all data

III-44

Example of a Multi-center Study web-site



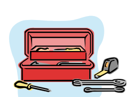
Example of Multi-center Study web-site

Example of Multi-center Study web-site

Web site as a communication hub

Recruitment and Adherence

Chapter 1 - Introduction
Chapter 2 - Recruitment - A. Media Info
Chapter 2 - Recruitment - B. Community Resource
Chapter 2 - Recruitment - C. Med Pro & Institute
Chapter 2 - Recruitment - D. Tools
Chapter 3 - Randomization - Tools & Materials
Chapter 4 - Maintenance - A. Clinic and Participant Tools
Chapter 4 - Maintenance - B. Follow-up Maintenance
Chapter 5 - Retention and Motivational Tools for Staff & Participants
Chapter 5 - Birthday Cards
Chapter 5 - Greeting Cards English
Chapter 5 - Greeting Cards French
Chapter 5 - Greeting Cards Spanish
Chapter 5 - Images
Chapter 5 - Incentive Cards
Chapter 5 - Sympathy Cards



ACCORD Survival Kit

III-48

Web site as a communication hub

Adherence A Short Course 12-02.doc

Participant Retention Letter 4-17-03.doc

Participant Retention Letter 603 Adherence.doc

Participant Retention Letter Cert For Missed Visits.doc

Participant Retention Letter Restart.doc

Red Flag Adherence Worksheet.doc

Visit reminder examples.doc

Search Tips Computer Search.doc

Search Tips Internet Resources

For Finding Lost ACCORD Participants.doc

Study Status Form Q by Q v3 2.pdf

Study Status Form V3 2.pdf

III-49

Steps in Data Collection

Closeout

III-50

Special notes on study closeout

• Continuous monitoring throughout the trial reduces the clean-up job at the end of the study

•Letter to participants (treatment assignment?)

•Lost-to-Follow-up (National Death Index, web-based searches, paid search firm)

III-51

Special notes on study closeout

•"Freezing" data at various points of cleanliness

•Data dictionaries created

•Responsibilities to sponsor (i.e., public use datasets, storing study materials)

III-52

Steps in Data Collection

Preparation for Analysis

III-53

Data Preparation for analysis

- Cleaning/editing
  - Inconsistencies
  - Omissions/discrepancies
- Merging records
- Documenting analysis files
  - Definition of variables/cut points
  - Validation of calculated variables
  - Verification of statistical outliers/distribution of data

III-54

Site Visits

Quality assurance visit of a clinical trial unit (e.g., clinical centers, coordinating center, central lab, etc.) by a team of experts to observe operations and assess performance

III-55

Scientific Misconduct in Clinical Trials

Data Fraud:

- reported in a small number of clinical trials
- refers to:
  - Fabrication (making up data)
  - Falsification (changing or removing data values)

III-56

High Quality Data

- Good Clinical Practice Guidelines
  - Good clinical research practice
  - SOPs
  - Ethical/scientific integrity
- "GIGO"
  - Garbage in, garbage out
  - Inaccurate data are worse than no data

III-57

Quality Control Monitoring Reports

III-58

Basic Monitoring Reports

- Data Monitoring
- Quality Control reports

III-59

Data Monitoring Reports

Examples of the following:

•Recruitment

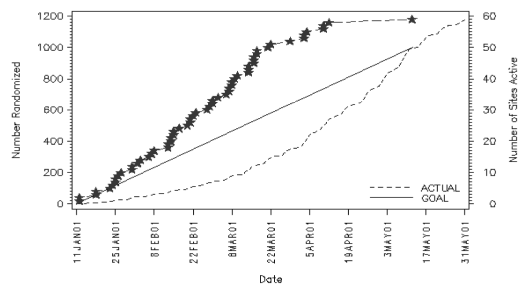
•Baseline and Follow-up data collection (includes lab, ecg, drug distribution, etc.)

•Adherence to protocol (clinicians and participants)

•Lost to follow-up, Refusals

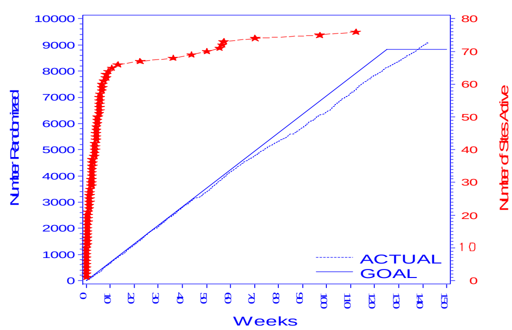
III-60

Recruitment Monitoring Example



III-61

Recruitment Monitoring Example



III-62

Monitoring Baseline Assessments

Are the study groups comparable at the time of randomization?

- Risk or prognostic factors, important demographic characteristics, medical history
- Randomization on average produces balance between groups – no guarantee!
- Correcting an imbalance: adjust in randomization or in analysis

III-63

Monitoring Baseline Assessments

Easiest way: compare each variable by treatment assignment using means, medians, ranges

Note that the groups will never be identical: 5% of the comparisons will show differences at the 0.05 significance level

III-64

Monitoring Follow-up assessments

1. Number of Visits completed as planned: %
2. Completeness of data: missing forms, missing data on forms
3. Quality of data received: data queries on each field (at data entry and/or retrospective data queries)

III-65

Monitoring Adherence

- Come at adherence from many different angles:
  - Participant adherence
  - Clinical site staff adherence to the protocol
- Long-term trials, look at changes over time
- Separate by calendar time, clinic visit, by clinic if a multi-center trial
- Tables and/or graphs

III-66

Monitoring Adherence

ID #	Trial Status	Date of Last Form	Days	Comment (CLICK "Enter a comment" to add your notes)
Ppt#1	**Non-adherent**	**15JAN2008**	**216**	10/20/2007 by Jill Jones (CCN): **Elevated CK > 5X ULN on 2 occasions. Does patient have symptoms?** 12/08/2007 by Joe Smith (CS): **Will reassess for symptoms of myositis at next visit.** 01/15/2008 by Joe Smith (CS): **Yes-pt has symptoms.** 04/17/2008 by Jill Jones (CCN): **Looks like both blinded lipid med and statin were stopped. Last LDL is > 120. Consider checking CK next visit off all lipid meds (looks like he may have some CK elevation even off of lipid meds), then rechallenge with low dose blinded lipid med alone and recheck CK in 6-8 weeks.** 04/28/2008 by Joe Smith (CS): **participant rechallenged on low-dose blinded meds, will check in 6-8 weeks** Enter a comment
Ppt#2	**LTF**	**12DEC2008**	**89**	01/09/2009 by Joe Smith (CS): **This patient has moved to Papua New Guinea for his work and couldn't come for his interval visit in December. Not forwarding address** 01/12/2009 by Jill Jones (CCN): **per our phone conversation, try alternate contacts to see if you can get phone contact info for an events assessment at minimum** Enter a comment

Monitoring Lost to Follow-up, Refused

- Separate groups: Lost to Follow-up versus Participant refusals (withdrawn consent)
- Investigators will want to know why participants are lost (e.g., moved out of range) and refused (e.g., withdrawn consent due to problems with protocol)
- Anticipate participants prone to becoming lost: monitor missed visit patterns and what happened to them
- Second tier: participants not officially LOST or REFUSED, but are no longer coming to the clinic or taking study medications

III-68

Summary

- Learning Objectives
  - GCPs, QC, QA, SOPs
  - Primary sources of error
  - Steps in Data Collection
    - Design of data collection forms
    - Standardization of procedures
    - Types of data entry/management systems
  - Quality control methods and reporting

III-69

Part IV: Treatment Allocation

Michele Melia, Sc.M.

Senior Statistician
Jaeb Center for Health Research
Tampa, FL

SCT Pre-Conference Workshop
Essentials of Randomized Clinical Trials

IV-1

Outline

- What randomization is and why it is used
- Truly random versus not random allocation
- Simple, block, and stratified randomization and when to use them
- Adaptive randomization and some of its pros and cons
- How to administer randomization in a trial

IV-2

What is randomization?

A process by which subjects are randomly assigned to a treatment in a clinical trial

- Neither the participant nor the investigator knows what treatment the participant will receive

IV-3

Why is randomization used?

- Problems arising with treatment assignment in clinical practice:
  - Individuals with certain disease characteristics are generally more likely to receive certain treatments (confounding by indication)
  - Inability to characterize why individuals were assigned to a particular treatment, leading to non-homogeneous groups with different (and unquantifiable) underlying risk
  - Wide variation in outcomes relative to the magnitude of differences due to treatments; treatment differences easily obscured by bias

IV-4

How does randomization work?

- Randomization does:
  - Reduce bias in assigning patients to treatments
  - Ensure valid statistical tests
  - Reduce unwanted variation resulting in improved power for statistical tests (more about this later)
- Randomization does not:
  - *Guarantee* equal distribution of prognostic factors among treatment groups
    - For large studies, the chance of imbalances is small
    - For small studies, the chance of imbalances is larger

IV-5

When is randomization used?

Phase I	Rarely	Not generally necessary to achieve phase I goals of establishing toxicity/maximum tolerated dose/dose response
Phase II	Sometimes	When comparison group is helpful in defining possible biologic and adverse effects, e.g. for highly subjective endpoints. When required by FDA.
Phase III	Almost always	"Gold standard " for reducing bias in assignment of patients to treatment and estimation of treatment effects

Other methods of (non-random) treatment allocation are also sometimes used in CTs:

- Single group with or without historical controls
- Non-random allocation of 2 or more groups

IV-6

Non-random methods of treatment allocation

- Alternating treatments (1st patient gets A, 2nd gets B, 3rd gets A, etc.)
- Alternating assignment by date or day of week (patient gets A if enrolled on even date, B if odd date)
- Using patient initials to determine assignment
  - A-K → treatment 1
  - M-Z → treatment 2

IV-7

Problems with non-random treatment allocation

- Treatment assignment of next patient can be predicted in advance; therefore,
  - – Not truly random
  - – Open to manipulation
  - – Goal of bias reduction can be subverted

IV-8

Basic types of randomization

- Simple
- Block
- Stratified / stratified block

IV-9

Simple Randomization

A sequence from a random number table or generator is used to assign sequential patients to a study treatment using a pre-defined rule. E.g. Even number→A and Odd number→B.

Sequence from random number table	Treatment assignment
7	B
7	B
9	B
2	A
1	B
0	A
6	A

IV-10

Simple Randomization

- Advantages
  - Simple
  - Each new assignment made without regard to previous assignments
- Disadvantages
  - No guarantee of equal or approximately equal sample size in each treatment group at any stage of the trial (including at the end)
    - Imbalance reduces statistical power
    - Estimates of treatment effect are not affected; only precision
  - No protection against long runs of one treatment

IV-11

Block randomization

- Block size that is an integer multiple of the number of treatments is chosen (integer≥2)
- Equal numbers of patients are assigned to each treatment within a block
  - Numbers are proportional rather than equal in the case of unequal allocation
- Overcomes some disadvantages of simple randomization

IV-12

Example: Block Randomization for 2 Treatments

- Possible block sizes are 4, 6, 8, etc.
- For block size of 4, there are 6 treatment-balanced permutations
  - – ABAB, AABB, ABBA, BABA, BBAA, BAAB
- These may be chosen at random with replacement

Sequence from random number table	Treatment assignment
7	--
7	--
9	--
2	AABB
1	ABAB
0	--
6	BAAB

IV-13

Block randomization – cont'd

- Large block size does not protect as well against long runs as small block size
- Small block size makes it easier to guess next treatment
- To make it harder to guess the next allocation when small block sizes are used, block size can be chosen at random from a pre-defined list of block sizes, e.g. 4, 6, 8
- Simple and block randomization do not guarantee balance of treatment groups on important prognostic factors

IV-14

Stratification

- With stratification, a separate, independent randomization sequence is used for each prognostic group (or strata)
- To guarantee treatment balance within strata at all stages of the trial, stratification is combined with blocking
  - Use of simple randomization within strata will not guarantee treatment balance within strata
  - Consequence of imbalance on a prognostic factor is bias in the estimated treatment effect

IV-15

Example – Blocked and stratified randomization

- A randomized trial comparing near versus distance activities while patching for amblyopia (lazy eye) in children 3 to <7 years old
  - Pilot study data suggested that near activities might be less effective in moderate as compared to severe amblyopia
  - Randomization was stratified by amblyopia severity; random block sizes of 4 and 6 also were used

IV-16

Example - continued

- If even, use block size=4; otherwise block size=6
- Use a random shuffle of the block elements

Moderate amblyopia

Random No.	Block size	Random sequence	Treatment assignments
7	6	7 9 2 1 0 6 A A A B B B	0 1 2 6 7 9 B B A B A A
1	6	8 5 1 3 0 7 A A A B B B	0 1 3 5 7 8 B A B A B A

Severe amblyopia

2	4	6 3 1 2 A A B B	1 2 3 6 B B A A
3	6	0 9 5 7 3 4 A A A B B B	0 2 3 5 7 9 A B B A B A

IV-17

Stratified randomization – cont'd

- Chance of imbalance on prognostic factors is small with large sample size
  - Stratification is more important when sample size is small
- As number of stratification factors increases, the number of strata grows very fast, and efficacy with respect to achieving desired balance may decrease
  - Think of case where # strata = sample size
- Be judicious in choice of stratification factors

IV-18

Stratified randomization – cont'd

- If many prognostic factors must be controlled:
  - Consider combining them into an overall index and stratifying on index
  - Consider minimization (more on this in a few moments)
- When analyzing data, it is important to account for stratification
  - If ignored, variability due to the stratification factor is included with error variance
  - If included, variability due to stratification factor is removed from error term, increasing precision

IV-19

Unequal Treatment Allocation

- With unequal treatment allocation, the study is designed to have unequal numbers of patients on the treatments
- Treatment groups of equal size are desirable from a statistical perspective for making treatment group comparisons
  - – Maximizes power for a given sample size
  - – However, loss of power may not be too severe as long as imbalance is not severe, e.g. 2:2:1

IV-20

Unequal Treatment Allocation – cont'd

- Some reasons to consider unequal allocation:
  - – More information is needed on effect of a new treatment (e.g. adverse effects, effect of dose)
  - – Patients may be unwilling to be randomized if probability of assignment to control or placebo is high
  - – To reduce study cost when one treatment is a lot more expensive than the other
- Principles of basic randomization regarding use of blocking and stratification still apply

IV-21

Cluster Randomization

- Clusters of patients are randomized rather than the individual patients
  - Example: In trial of vitamin A supplementation for prevention of mortality in preschool children in Nepal, administrative wards were randomized to supplement or placebo (West KP, Lancet 1991)
- Cluster randomization reduces statistical efficiency (i.e. it requires more patients)
- Usually used when it is not feasible to randomize individual patients

IV-22

Adaptive Allocation (aka Adaptive Randomization)

- Information on previously enrolled patients is used to modify (or adapt) the allocation ratio, i.e. the probability of being assigned to each treatment
- Information used typically is one of:
  - Treatment
  - Covariates (prognostic factors)
  - Response (outcome)
- Other terms:
  - Biased-coin design
  - Urn design
  - Play-the-winner design

IV-23

Treatment Adaptive Randomization

- Allocation ratio is adjusted using the number of patients previously assigned to each treatment
- Basic idea (for trial with 1:1 allocation):
  - If current proportion of patients randomized to A is less than ½, assign current patient to A with probability greater than ½.

IV-24

Treatment Adaptive Randomization

- Advantages
  - Balance on # of patients in each treatment group is achieved at all stages of the trial
  - Harder to guess next assignment than for randomized block design with small block size
- Disadvantages
  - Increased administrative complexity
  - Analysis is more complicated – probability for each assignment is needed

IV-25

Covariate Adaptive Randomization

- Also known as minimization
- Basic idea:
  - If number of previous patients with covariate profile matching the current patient is higher in group A than B, then probability the current patient is randomized to B is increased to greater than ½.

IV-26

Covariate Adaptive Randomization – cont'd

- Advantages
  - Achieves balance among treatments on important covariates
- Disadvantages
  - Intensive administrative effort may be needed (especially if number of covariates is large)
  - Increased risk of breaking masking
  - Unnecessary matching
    - Large sample size alone is likely to result in good balance on covariates
    - Randomization and analysis have been complicated unnecessarily

IV-27

Response Adaptive Randomization

- Also known as 'Play-the-winner' designs
- Basic idea:
  - If current trial results favor treatment A, probability that the patient is randomized to A is increased to greater than ½
- Famous example: ECMO Study (Bartlett, Pediatrics 1985)
  - Start with 2 balls in an urn marked E(cmo) and C(ontrol)
  - If treatment is successful, add a ball marked with that treatment into the urn (along with the selected ball)
  - If not successful, add a ball marked with the opposite treatment (along with the selected ball)

IV-28

Response adaptive allocation - ECMO Study

- Trial ends when 10 balls of 1 type are added with that type declared the winner
- Assuming one treatment has substantially greater chances of survival, this design has high probability of selecting the better treatment as the winner

IV-29

ECMO Study Results

E C	• E(cmo) selected • Patient lives
E C E	• C(ontrol) selected • Patient dies
E C E E	• E selected • Patient lives
E C E E E	• 4^{th}-10^{th} balls: E selected • Patients all live

IV-30

ECMO Study Results

- 10 E balls were added, so ECMO declared the winner
- 2 more patients given E; both lived
- Final counts:
  - 0/1 control patients lived
  - 11/11 ecmo patients lived
- Might be tempted to analyze using Fisher's Exact Test, but cannot, as marginal totals are random variables that contain information about the outcome

IV-31

Response Adaptive Allocation – cont'd

- Advantages
  - Increases chances that patients will get the better treatment
  - Ethically appealing
- Disadvantages
  - Increased administrative complexity
  - Not always possible (e.g. long-term response)
  - Analysis is more complicated; appropriate statistical tests may not exist
  - Ethical difficulties if allocation ratio becomes highly skewed to one treatment

IV-32

Summary – Adaptive Allocation

- Simple randomization or stratified block randomization are generally perfectly adequate when sample size is large
- Consider complex alternatives only if sample size is small

IV-33

Administration of randomization codes

- When the study protocol is finalized, but before the study begins patient enrollment:
  - The randomization schedule is generated (for a non-adaptive randomization scheme)
  - Procedures for obtaining a randomization code for a study patient are defined
  - Procedures for unmasking are defined
  - System for tracking randomizations issued, errors and deviations from schedule, and unmasking is in place

IV-34

Generating the randomization schedule

- A Standard operating procedure (SOP) for generating randomization schedules is desirable. Elements of the SOP should include:
- Who may generate a schedule (preferably this is done by a statistician not involved in day-to-day study operations)
  - Statistician ensures that the schedule adheres to the study design
- Procedures for schedule/code checking

IV-35

Generating the schedule - continued

- Documentation of how the schedule was generated
  - Programs & pseudonumber generator used
  - How to use them
  - Seed(s) used to obtain the schedule in question
- For studies being submitted to FDA, the programs must be validated (and periodically re-validated) and results of validation must be documented

IV-36

Procedures for obtaining a randomization code

- There are many procedures that are commonly used including:

  Centrally administered
  - Telephone call to coordinating center or its surrogate (e.g. answering service)
  - Web-based system

  Locally administered
  - Sequential drug kits
  - Envelope system
  - Computer program installed on local PC

IV-37

Procedures for obtaining a randomization

- Procedures should take into account:
  - Allowable time between request for randomization and issuance of randomization
  - Times of day and days of week that patients will be randomized and attendant staffing needs
    - Coverage for all time zones
  - Ease and convenience for investigators and patients

IV-38

Procedures for obtaining randomization – cont'd

- Procedures should take into account:
  - Vulnerability to manipulation or tampering
    - Centrally-administered systems generally easier to secure
    - Secure local systems are possible with proper safeguards
  - Need for fall back procedure in event that primary procedure isn't working (e.g. web site outage)

IV-39

Procedures for unmasking

- Under what circumstances is unmasking permitted?
- Who may be unmasked?
- How will unmasking be performed?

IV-40

Summary

- Randomization is the primary means for controlling bias in allocation of patients to treatment in a clinical trial
- Randomization helps to generate comparable groups of patients on each treatment
- Randomization enables valid statistical tests for the evaluation of the treatments
- Judicious use of stratification with appropriate analysis can improve statistical power

IV-41

Selected References

- Chow S-C, Liu J-P: Design and Analysis of Clinical Trials, 2nd ed. John Wiley and Sons, 2004; pp 120-153.
- Meinert CLM: Clinical Trials: Design, Conduct, and Analysis. Oxford University Press, 1986; pp 90-112.
- Piantadosi S: Clinical Trials: A Methodologic Perspective. John Wiley and Sons, 2005; pp 331-353.
- Spilker, B: Guide to Clinical Trials. Raven Press, 1991; pp 69-73.
- Controlled Clin Trials 1988; Volume 9, issue 4 has a series of articles on randomization in clinical trials by John Lachin

IV-42

Software

nQuery Advisor can be used to generate randomization lists

For links to randomization software (free) and services (not free) developed and maintained by Martin Bland at University of York see:

http://www-users.york.ac.uk/~mb55/guide/randsery.htm

Disclaimer: endorsement of software and services on this website is not implied

IV-43

Part V: Choice of Endpoints

Susan Halabi, Ph.D.

Associate Professor
Department of Biostatistics and Bioinformatics,
CALGB Statistics and Data Center
Duke University

SCT Pre-Conference Workshop
Essentials of Randomized Clinical Trials

Learning Objectives

By the end of the course, attendees should be able to:

- Identify possible endpoints for their study
- Assess the pros and cons for possible endpoints
- Be able to 'better' choose endpoints that meet study needs

- Missing data
- Intent to treat

V- 2

Outline

- Primary Question
- Primary Endpoints
- Type of Endpoints
- Secondary Endpoints
- Composite Endpoints
- Surrogate Endpoints

V- 3

Choice of the Study Question

"Each clinical trial must have a primary question. The primary question, as well as any secondary or subsidiary questions, should be carefully selected, clearly defined, and stated in advance."

Friedman, Furberg, DeMets

V- 4

Primary Question

- Investigators most interested in answering
- One capable of being adequately answered
- Often framed in the form of a hypothesis
- Can be "superiority" or an "non-inferiority"
- Should be important and clinically relevant

V- 5

Questions vs. Endpoints

- Research question(s) – What we want to show
  - hypothesis
- Endpoint(s) – How to show it
  - single primary outcome
  - limited number of secondary outcomes
- Endpoint(s) are much more specific than question(s)

V- 6

Choice of Primary Endpoint

For a drug or a device to be considered efficacious, it must demonstrate tangible clinical benefit, generally defined as an improvement in survival or improvement in symptoms.

V- 7

Primary Endpoint

- A key decision in designing a trial
- The major determinant of sample size

V- 8

Primary Endpoint

- Consistent with the primary study question
- Clearly defined and specified in advance
- Capable of being ascertained as completely as possible (ideally in every subject)
- Reproducible in research study
- Measured in the same way for all subjects
- Capable of unbiased assessment

V- 9

Examples: Endpoints

- Overall mortality
  - Acute or short-term mortality (e.g., 30-day, 7-day, within the index hospitalization)
  - Long-term mortality (over an extended period of follow-up)
  - Objective, “hard” endpoint
  - Doesn’t require classification/adjudication as to mode of death

V- 10

Examples of Endpoints (cont.)

- Cause-specific mortality
  - e.g., death due to cancer, cardiovascular causes, cardiac causes, sudden death, arrhythmic death
  - Requires classification/adjudication of deaths
  - Cause of death is often difficult to determine

V- 11

Choice of Primary Endpoint- Cancer

- Phase I – proportion of patients who experience a dose limiting toxicity (DLT)
- Phase II- (Non-randomized & Randomized)
  - Tumor shrinkage (Objective response rate)
  - Progression-free survival
- Phase III
  - Overall survival - objective endpoint
  - Time to death due to disease: problems in determining cause of death
  - Progression-free survival

V- 12

Examples of Primary Endpoints- CVD

- Overall mortality
  - 30-day mortality (GUSTO-I)
  - Arrhythmic death/cardiac arrest (MUSTT)
- Incidence of fatal and non-fatal stroke (SHEP)

V- 13

Types of Primary Endpoints

- Binary (e.g., objective response rate, 30-day mortality)
- Ordinal (e.g., toxicity- graded from 0 [none] to grade 5 [death]; two or more seizures)
- Continuous (e.g., quality of life, visual analog scale, CD-4, lymphocyte count)

V- 14

Types of Primary Endpoints

- Time to an event (e.g., overall survival)
- Composite Endpoint (e.g., progression-free survival, fatal/non-fatal mortality)
- Surrogate Endpoint (e.g., PSA decline)

V- 15

Example

Research Question:

Does treating breast cancer women with bisphosphanates increase bone mineral density (BMD)?

Study design: Prospective

Rigorously define:

- By how much ?
  - Any / XX or more
- Since when ?
  - Baseline / last visit
- Relative or absolute difference

V- 16

Example

Research Question:

Does treating breast cancer women with bisphosphanates increase bone mineral density (BMD)?

Study design: Prospective

Rigorously define:

- Bone Mineral Density:
  - -Any / Threshold –XX or more
- How are you measuring bone mineral density?
  - - lumbar

V- 17

Rigorous Assessment Methods

Study protocol should specify….

- Equipment needed (dual energy x-ray absorptiometry (DEXA) scan)
- Time of evaluation (baseline, 12 months)
- Who determines endpoint

V- 18

Reproducible in Research Study

Internal Data
Duplicate measures
Sample / Total study population
Same / different assessors
Same / different methods
Same / different days

External Data
Similar method
Similar personnel
Similar training

V- 19

Assessable in All Groups

- Same methods for all
- Documentation of methods (protocol)
- Same time points for all

V- 20

Composite Endpoints

Composite event

…considered to have occurred if any one of several different outcomes are observed

- e.g. angina pectoris, transient ischemic attack, or myocardial infarction = composite vascular event

V- 21

Composite Endpoints - Advantages

Possible Advantages

- Increases expected event rate
- Increases power
- Reduces sample size
- Shorter study duration
- Combine benefits and risks
- Reduce bias
- Allow multiple important outcomes

V- 22

Composite Endpoints - Disadvantages

Possible Disadvantages

- Confusion in interpreting results
- Additional ‘noise’ may hide differences
- Correlated events - smaller advantage
- Sample size – “minimum clinically important difference”

V- 23

Challenges in the Use of Composite Endpoints

- Complete ascertainment of the component endpoints is required. Missing data may be a problem
- Important nonfatal outcomes need to be adjudicated
- Appropriate design and analysis approaches are required

V- 24

Composite Endpoint- Example

- Skeletal Related events
  - Pathologic bone fracture in the region of cancer involvement
  - Radiation therapy to bone
  - Cancer related surgery to bone
  - Spinal cord or nerve root compression
  - Initiation of bisphosphonate therapy in response to new bone pain symptoms
  - Change of antineoplastic therapy for bone pain due to prostate cancer
  - Death from prostate cancer

V- 25

Primary vs. Secondary Endpoints

Endpoint (outcome)
Determined in each study subject / participant / unit

Primary outcome variable
"… designated or regarded as key in the design or analysis of the results of a trial." – Meinert, CL

Secondary outcome variable
" any other outcome variable used for treatment evaluation" – Meinert, CL

V- 26

Secondary Questions

- Subsidiary questions related to the primary question
- Involve different outcomes than the primary endpoint (e.g., primary endpoint is disease-free survival, secondary endpoint overall survival)
- May relate to sub studies or ancillary studies (e.g., prognostic factors of overall survival)
- May relate to subgroup hypotheses (e.g., stage, responders, non-responders)

V- 27

Secondary Endpoints

- Reasonable to consider several secondary endpoints
  - Primary endpoint: objective response, secondary endpoint: overall survival, toxicity (phase II)
  - Primary endpoint: overall survival, secondary endpoint toxicity, quality of life, etc (phase III)
- If the primary endpoint is a composite, including the individual components as secondary endpoints is desirable
  - Primary endpoint: progression-free survival
  - Secondary endpoint: PSA progression

V- 28

Surrogate Endpoints

- Surrogate endpoints usually are proposed based on biological pathways
- More readily available earlier in the course of the cancer's natural history
- Measurable more frequently, are less costly and thus more "convenient" than the "true" endpoints

V- 29

Surrogate Endpoints - Definition

Surrogate outcome variable

> "A test, measurement, score, or some other similar variable that is used in place of a clinical event in the design of a trial, or in summarizing results from it."

- Believed to be correlated with clinical event
- Perceived utility in yielding detectable treatment difference

– Meinert, CL

V- 30

Criteria for Good Surrogate Endpoints

- Strong statistical association with primary endpt.
- Change in surrogate strongly correlated with change in primary endpoint (but: correlation ≠ causality)
- Surrogate is in the biological pathway of the disease (there may be > 1pathway)
- Short latency (↑surrogate followed by rapid onset of disease)
- Responsive to treatment (effect on surrogate may not equal effect on disease)

V- 31

Prentice Criteria for Surrogate Endpoints

- Prentice developed a formal definition of surrogate endpoint
  - There is a treatment effect with respect to the surrogate endpoint
  - The surrogate endpoint is a prognostic factor of the true endpoint
  - The surrogate endpoint should capture all treatment effects on the true endpoint

V-32

Surrogate Endpoints - Advantages

Possible Advantages

- Smaller sample size
- Endpoint earlier than ideal endpoint
- Easier
- Less costly

V- 33

Surrogate Endpoints - Disadvantages

Possible Disadvantages

- Not well correlated to ideal endpoint
- Mechanism of action unclear
- Less acceptable
- Less clinical relevance
- NO SURROGATE for Safety

V- 34

Surrogate Endpoints- Example 1

Use Prentice's criteria for surrogacy

•50% decrease in PSA over 3 months barely failed one of the surrogate criteria.

•30% decrease satisfied all criteria for 3 and 2 months.

Petrylak et al: JNCI 2006

V-35

Surrogate Endpoints – Example 2

Cardiac Arrhythmia Suppression Trial (CAST)

- Prior evidence of association between arrhythmia and sudden death.
- Wide use of medication to suppress arrhythmia
- Enrolled: patients with asymptomatic or mildly symptomatic ventricular arrhythmia (six or more ventricular premature beats per hour) after myocardial infarction.
- Treatment(s): antiarrhythmic therapy (encainide, flecainide, or moricizine)
- Endpoint(s): death from arrhythmia / initial suppression of their arrhythmia (as assessed by Holter recording)

V- 36

Surrogate Endpoints – Ex. 2 *Continued*

Cardiac Arrhythmia Suppression Trial (CAST)

- March 30, 1989 Results:
  - 75 % had initial suppression of their arrhythmia (surrogate)
  - higher rate of death (primary) from arrhythmia in patients assigned to active drug than the patients assigned to placebo
- "We conclude that neither encainide nor flecainide should be used in the treatment of patients with asymptomatic or minimally symptomatic ventricular arrhythmia after myocardial infarction, even though these drugs may be effective initially in suppressing ventricular arrhythmia."

- Evidence that effect on possible surrogate outcome may differ from effect on clinical outcome

V- 37

Surrogate Endpoints

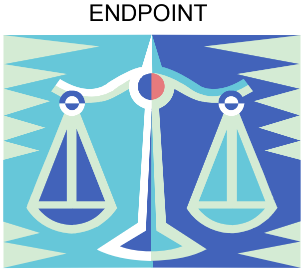
Disease	Definitive Endpoint	Surrogate Endpoint
CVD	MI	Cholesterol level
	CHD	Carotid IMT
	Heart Failure	BNP
	Stroke	Blood pressure
Cancer	Mortality	Tumor size reduction
Prostate Cancer	Overall Survival	PSA
HIV Infection	AIDS/Death	CD4+ count
Glaucoma	Vision Loss	Intraocular pressure

V- 38

Balance and Adjustments

ENDPOINT

Scientific considerations



Practical considerations

V- 39

Endpoint Considerations

Choice of endpoint will affect:

Personnel
Equipment
Facilities
Study duration
Sample size calculations

Resources will affect choice of endpoint

V- 40

Personnel

Who (skill level)
- HS education vs. special training vs. machine

What
- Examination vs. photos vs. lab values

Where
- Local clinic vs. home visit vs. central facility

When and how often
- One point in time vs. repeated measures

Personnel turnover

V- 41

Equipment

Specialized vs. standard

Specific make model vs. approved subset vs. any
- If more than one type – can you switch

Move equipment to people or people to equipment

Technology stable vs. changing/improving

Any information comparing equipment (endpoint)

V- 42

Facilities - Local

Size of room
conduct visit, store files**, forms
measure outcome (distance vision)

Location of room
elderly population – stairs, long walk

Privacy (quality of life, or personal interview)
shared space vs. dedicated trial space

**Know the rules for how long you must keep data forms /specimen
your institution / study sponsor

V- 43

Facilities - Central

Reading Center (photographs, ultrasound, X-rays, etc)
Pathology Center (tissue/ slides)
Radiation Physics Center (dose curves)

- Space – specimens, gradings, storage **
- Ancillary study use of materials
  - Committee to approve use
  - Archiving committee

**Know the rules for how long you must keep data forms /specimen
your institution / study sponsor

V- 44

Study Duration

When is endpoint assessed
- Day 1 vs. 8 weeks vs. all cause mortality
- Length of follow-up beyond primary outcome
- Frequency of assessment

Rate of occurrence
rare event vs. common event

Single vs. multicenter

V- 45

Adjudicated Endpoints

- Subjective Endpoints
- Multiple assessments/assessors – then adjudicate
- Committee – Equal experience or Experts and non-experts
- Where are people located ?
- Adjudicate in person / e-mail
- How often does adjudication happen ?
- What materials does committee need ?
- Grade independently or all together ?

V- 46

Missing Patients (Endpoints)

- Exclusions (never randomized)
  - No bias in randomized comparison
  - Does influence interpretation and generalization
- Withdrawals (deliberately omitted from analysis)
  - Severe bias may arise
  - Withdrawals may be acceptable if based on eligibility criteria determined at baseline and not affected by events subsequent to randomization
- Losses to follow-up (missing outcome data)
  - Bias may arise if the loss is related to the intervention and the outcome

V- 47

Missing Data

- Treatment dropouts do not necessarily have missing outcome data
  - we should design trials (& informed consent processes) so that treatment modifications and/or dropout do not lead to "off-study"
  - such patients should still be followed for outcome
- Patients who need (or want) to modify their therapy may be prognostically different from those who are maintained on the therapy initially assigned (and this may vary by treatment group)

V- 48

Intention to Treat (ITT) Analyses

- Include all individuals randomized
- Include in the group to which they were randomized
- Regardless of what treatment they received or what occurs subsequently
- First analysis of any randomized trial
- Supported by the randomization
- Maintains comparability (expectation)

V- 49

Intention to Treat - 2

- Provides a test of the "policy" ("strategy", "intention")
- Estimate of effectiveness (real world)
  - Efficacy – analyse as treated (ideal world)
- May need to adjust sample size for non-compliance

V- 50

Intention to Treat - 3

What is the goal?

- "pragmatic efficacy" [intent-to-treat]

  or
- "biologic efficacy" [full compliance] *

* may not be attainable (intolerance or toxicity)
danger of false optimism

* may not be straightforward: danger of bias

V- 51

ITT Caveat

In equivalence trials, excessive noncompliance may lead to apparent equivalence which does not reflect reality

- here, intent-to-treat analysis does not have the usual advantage of "conservatism"

V- 52

Incomplete Compliance / Treatment Dropouts

Severe bias may arise if deliberately omitted from analysis

comparing compliers in both groups may be biased;

"as treated" analysis may even be worse

> lose the comparability provided by randomization

V- 53

References

Cardiac Arrhythmia Suppression Trial (CAST) Investigators **Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction** *N Engl J Med*. 1989 Aug 10;321(6):406-12.

Cannon CP. **Clinical Perspectives on the Use of Composite Endpoints** *Controlled Clinical Trials* 1997;18:517-529.

Chen YHJ, DeMets DL, Lan KKG **Monitoring mortality at interim analyses while testing a composite endpoint at the final analysis** *Controlled Clinical Trials* 2003; 24:6–27.

Connolly SJ. **Use and Misuse of Surrogate Outcomes in Arrhythmia Trials** *Circulation* 2006,113;764-766.

D'Agostino RB Jr. Commentary **Debate: The slippery slope of surrogate outcomes.** *Curr Control Trials Cardiovasc Med* 2000, 1:76–78.

DeMets DL. **[Evaluating Therapeutic Interventions: Some Issues and Experiences:] Comment.** *Statistical Science*, 1992, 7, 443-444.

V- 54

References

Freemantle N, Calvert M, Wood J et al **Composite Outcomes in Randomized Trials Greater Precision But With Greater Uncertainty?** *JAMA. 2003;289:2554-2559.*

Friedman LM, Furberg CD, DeMets DL: Fundamentals of Clinical Trials. New York, Springer-Verlag, Fourth edition 2010

Gail MH et al. **On meta-analytic assessment of surrogate outcomes.** *Biostatistics* 2000, 1(3):231-246.

Gluud C et al. **Review : Hepatology may have problems with putative surrogate outcome measures** Journal of Hepatology 46 (2007) 734–742.

Hannah ME et al. **Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial** *Lancet* 2000; 356: 1375–83.

Ioannidis JPA et al. **The Relationship between Study Design, Results, and Reporting of Randomized Clinical Trials of HIV Infection** Controlled *Clin Trials* 1997;18:431-444

V- 55

References

Kelly WK, Halabi S. Oncology Clinical Trials: Successful Design, Conduct and Analysis. Demos Medical Publishing, New York, NY, 2010.

Prentice, RL. **Surrogate endpoints in clinical trials: definition and operational criteria.** *Statistics in Medicine* 1989, 8: 431-440.

Ross S. **Composite outcomes in randomized clinical trials: arguments for and against**. Am J Obstet Gynecol 2007;196:119.e1-119.e6.

Tugwell Pet al. **Powering our way to the elusive side effect: A composite outcome 'basket' of predefined designated endpoints in each organ system should be included in all controlled trials** Journal of Clinical Epidemiology 2005;58: 785–790

The Systolic Hypertension in the Elderly Program (SHEP) Cooperative Research Group. **Rationale and design of a randomized clinical trial on prevention of stroke in isolated systolic hypertension.** J Clini Epid. 1988, 41:1197-1208.

V- 56

Part VI: General Statistical Concepts

Susan Halabi, Ph.D.

Associate Professor
Department of Biostatistics and Bioinformatics,
CALGB Statistics and Data Center
Duke University
SCT Pre-Conference Workshop
Essentials of Randomized Clinical Trials

VI-1

OUTLINE

1) Hypothesis testing
2) Power / Sample Size Calculations
3) Sample size calculations for different outcomes
   - Dichotomous outcomes
   - Continuous outcomes
   - Time-to-event outcomes
4) Software
5) Philosophy of Interim Monitoring

VI-2

TESTING

A primary objective of most clinical trials is to demonstrate the effectiveness and safety of a treatment under investigation.

The purpose of such trials is to:

- Find out which (if any) of the treatments are more effective
- Convince others of the results

VI-3

TESTING

In designing such trials, we need to keep in mind two issues related to participant (patient) heterogeneity:

- The effect of chance
- The effect of bias (whether conscious or unconscious)

These are addressed by:

- ➢ Using randomization for treatment assignment
- ➢ Having adequate numbers of participants in study

VI-4

TESTING

Hypothesis testing involves:

- ➢ Collecting a sample and using the sample to estimate unknown population parameters.
- ➢ Comparing the sample estimate(s) to some hypothesized population value to see if the sample came from the specified population.

VI-5

TESTING

Hypothesis: Statement about a population parameter

Null Hypothesis (H_0): A hypothesis of no difference or status quo; often what we would like to disprove

$$H_0: \mu = 0$$

Alternative Hypothesis (H_A): A statement which contradicts the null hypothesis

$$H_A: \mu \neq 0$$

The goal of hypothesis testing is to collect a sample and determine which hypothesis is 'more likely' to have generated the observed sample.

VI-6

TESTING

Test Statistic: A statistic computed from the sample upon which we will base our decision

Acceptance Region: The range of values for which H_0 is not rejected

Rejection Region: The range of values for which H_0 is rejected

The test statistic must fall into one of these regions.

VI-7

TESTING

The test statistic must fall into one of these regions:

- If the test statistic falls into the rejection region, the test is said to be statistically significant
- If we don't reject H_0, we can't claim to 'accept H_0'
  - Suppose one makes a statement 'all swans are white'
  - To examine this statement, a sample of swans is drawn
  - Two things can happen:
    a) All swans in the sample are white
    b) At least one swan in the sample is not white
  - The event (b) establishes the falsehood of statement
  - However, the event (a) does not prove the statement!

VI-8

TESTING

Type I Error: Rejecting null hypothesis when true (i.e., conclude benefit when none actually exists)

- α = Pr{ Type I error }
  = Pr{ Reject H_0 when true }

Type II Error: Not rejecting null hypothesis when false (i.e., fail to conclude benefit when actual benefit exists)

- β = Pr{ Type II error }
  = Pr{ Fail to reject H_0 when false }

VI-9

TESTING

The decision may be summarized as follows:

CONCLUSION	TRUTH: No Treatment Benefit	TRUTH: Treatment Benefit
Evidence of Treatment Effect	**Type I error (False Positive)**	Correct Result (True Positive)
No Evidence of Treatment Effect	Correct Result (True Negative)	**Type II error (False Negative)**

VI-10

TESTING

Statistical tests quantify the probability of a type I error (false positive result).

For example, an observed difference with $p \leq 0.01$ implies that the probability of obtaining a difference this extreme (or more so) by chance alone is less than or equal to 1%.

VI-11

TESTING

There is a tradeoff between the probability of a type I and a type II error.

Traditionally, type I errors are of greater concern.

Hence, we often fix α at 0.05 and try to take a large enough sample to ensure β is at a reasonable level (<0.20??)

Should this always be the case?

VI-12

TESTING

Example (from Rosner, p. 193-194):

It has been suggested that a certain hospital has lower birth weight babies than the national average.

To see if a special care nursery is needed, a sample of birth weights from the hospital are collected and used to test:

$$H_0: \mu \geq \text{national average}$$

vs.

$$H_A: \mu < \text{national average}$$

VI-13

TESTING

If H_0 is rejected, the hospital will add a special care nursery.

- If a type I error is made, the extra cost of adding a special care nursery will be recommended when it is not needed
- If a type II error is made, a needed special care nursery will not be funded.
  - As a result, some low-birth weight babies may not receive the special attention that they need

VI-14

TESTING

A confidence interval quantifies the uncertainty around the estimated intervention effect.

CI's also indicate the range of values within which we think the true intervention effect lies.

Relationship between CI's and hypothesis tests:

- A (1-α) x 100% confidence interval for μ consists of all values for which H_0 could not be rejected at the α level.

VI-15

TESTING

Jury Trial (criminal law)	Clinical Trial (statistical testing)
• Presume innocent	• Assume the null hypothesis
• Goal: Convict the guilty	• Goal: Detect a true difference (Reject the null hypothesis)
• "Beyond reasonable doubt"	• "Level of significance"
• Requires evidence: Convincing testimony	• Requires evidence: Adequate sample size
• Mistake: Convict an innocent person	• Mistake: False positive (Type I error)

VI-16

Acknowledgement to Susan Hilsenbeck and Sylvan Green

POWER

A primary objective of most clinical trials is to demonstrate the effectiveness and safety of a treatment under investigation.

Hence, sample size calculation plays an important role at the planning stage to ensure sufficient subjects for answering the question of interest.

If sample size is too large, study will waste resources

If sample size is too small, study underpowered and a potentially useful treatment may be discarded.

VI-17

POWER

Sample size calculation is usually performed based on some statistical criteria controlling Type I and/or Type II errors.

With a fixed sample size:

- α increases as β decreases
- α decreases as β increases

The only approach to decrease both α and β is to increase the sample size.

VI-18

POWER

Two common approaches to choosing sample size:

- Precision Analysis: Sample size chosen such that there is a desired precision at a fixed confidence level (i.e., fixed Type I error)
- Power Analysis: Sample size chosen to achieve desired power for detecting clinically/scientifically meaningful difference at a fixed Type I error rate.

In this workshop, we focus on sample size calculation based a power analysis for various situations in clinical trials.

VI-19

POWER

Power of the test is defined as the probability of correctly rejecting the null hypothesis when false.

- Power = 1 - β
  = Pr{ Reject H_0 when false }

VI-20

POWER

Two types of power analysis:

- Sample Size Estimation: Calculation of required sample size for achieving desired power.
- Sample Size Justification: Provide justification for a selected sample size, which is often small due to budget and/or other constraints.

In this workshop, we focus on sample size estimation but the basic principles apply to both approaches.

VI-21

POWER

A *valid* sample size calculation MUST be based on *appropriate tests* for hypotheses that reflect study objectives under a *valid* study design.

Hence, it is important that the following are aligned:

- Study Objective (Hypothesis)
- Study Design
- Statistical Analysis (Test Statistic)
- Sample Size Calculation

Any discrepancies between these items can distort the validity and integrity of the trial.

VI-22

POWER

What must be known to compute sample size?

1. Type of outcome data
2. Type of test
3. Measure of precision or variability
4. The magnitude of treatment difference that the study should be able to detect (δ)
5. Specified Type I error (α)
6. Target Power [or specified Type II (β) error]

VI-23

POWER

Type of Data:

- Dichotomous
  (success or failure; presence or absence)
- Continuous
  (blood pressure; length of hospitalization)
- Time to event
  (time to occurrence of an event of interest)

Sample size estimates for outcomes that do not fall into these categories can usually be approximated by one of them!

VI-24

POWER

Type of Test:

- **Test for Equality**: Show one treatment is more effective than another

$$H_0: \delta = 0 \text{ vs. } H_A: \delta \neq 0$$

- **Test for Superiority**: Show test drug is more effective than an active agent or standard therapy

$$H_0: \delta \leq \varepsilon \text{ vs. } H_A: \delta > \varepsilon$$

where ε is the *superiority* margin.

VI-25

POWER

Type of Test:

- **Test for Non-inferiority**: Show test drug is as effective as an active agent or standard therapy

$$H_0: \delta \leq -\varepsilon \text{ vs. } H_A: \delta > -\varepsilon$$

where ε is the *non-inferiority* margin.

- **Test for Equivalence**: Show no difference of clinical importance between two treatments

$$H_0: |\delta| \geq \varepsilon \text{ vs. } H_A: |\delta| < \varepsilon$$

where ε is the *equivalence* margin.

VI-26

POWER

Type of Test (cont.):

- Important to ensure that the sample size calculation parallels the planned primary analysis.
- The hypothesis of interest should be clearly stated when performing a sample size calculation.
- Each of the above hypotheses has a different sample size requirement in order to achieve a desired power for the corresponding test.
- For this workshop, we will primarily focus on tests of equality between two treatments.

VI-27

POWER

Precision and Variance:

- A more precise method of measurement (i.e. small σ) will permit detection of any given δ with a smaller sample size.
- The importance of precision increases as the desired size of the effect becomes smaller.
- A study with a small sample size will have more uncertainty and will only show statistically significant differences if there is a large difference between the two groups.

VI-28

POWER

Treatment Effect:

- The choice of δ is critical for study planning
- Different choices of δ have major effects on the sample size requirements.
- If δ is small, a large sample size will be required
- Important to ensure the treatment effects have both clinical and statistical meaning
- *Possible* to design study to detect reduction of onset time of local anesthesia from 60 to 59 seconds, but likely not of clinical importance.

VI-29

POWER

Type I Error (Significance level):

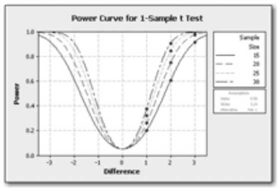
- Pre-set by researchers early in study planning
- Common α values are 0.01, 0.05, and 0.10.
- Often, choose α = .05 more by convention than design
- This implies that we would expect to reject the null hypothesis 5% of the time when it is true (there is no effect).
- May need to adjust for multiple testing

VI-30

POWER

Power:

- Typically set at 80% or 90% for planning purposes
- Power curves are useful since study planning often involves a trade-off between desired sample size, cost, and patient resources



VI-31

POWER

Power (cont.):

- Power curves typically have a sigmoidal shape, with increasing power as n or δ increases.
- Impact of small changes in design parameters depends on shape of power curve.
- If trial design lies near shoulder, small changes in design parameters can seriously affect power.
- Typically, trials designed with 80% power are more susceptible to inaccuracies in design parameters than trials designed with 90% power.

VI-32

POWER

To determine power, we need to specify

- The sample size - *N*
- The significance level - α
- A clinically important difference that we wish to detect - δ
- Any additional nuisance parameters

VI-33

POWER

To determine sample size, we need to specify

- Target power – $P_t = 1-\beta$
- The significance level - α
- A clinically important difference that we wish to detect - δ
- Any additional nuisance parameters

VI-34

POWER

Sample size estimates are approximate:

- Equations often based on approximations to the exact statistical distributions.
- Parameters used in calculations are guesses and have an element of uncertainty

Researchers hope that any errors are small and that the computed sample size is close to the actual number truly needed.

Be conservative (but realistic – always round up!) when estimating sample size!

VI-35

POWER

Small changes in design parameters may yield large changes in the required sample size.

Required sample size increases with:

- Variance of the treatment difference
- Decreasing type I error
- Increasing desired target level of power
- Smaller treatment effects of interest

VI-36

POWER

Note that we cannot separate power from either size of study or magnitude of treatment effect.

Hence, the following statement is ambiguous:

"The trial has 90% power."

All three values must be discussed simultaneously:

"With 500 subjects per group, the trial has 90% power to detect a decrease of 10 mmHg in blood pressure due to the new treatment at the 5% significance level."

VI-37

POWER

Sample size calculation provides the number of *evaluable* subjects required for achieving a desired level of power.

If drop-outs are expected, the sample size should be adjusted upward to ensure a sufficient number of evaluable subjects.

If the response variable can be partially explained by other covariates, the required sample size may be reduced.

VI-38

CONTINUOUS OUTCOME

Suppose that there are two groups of observations:

x_i, i = 1,…,n_1 (treatment)

y_i, i = 1,…,n_2 (control)

Assume that x_i and y_i are *independent* and *normally distributed* with means μ_1 and μ_2, respectively, and a common variance, σ^2.

VI-39

CONTINUOUS OUTCOME

Suppose the hypothesis of interest is:

H_0: $\mu_1 = \mu_2$ vs. H_A: $\mu_1 \neq \mu_2$

Assuming *equal variance* and *equal sample sizes* in the two groups, use the test statistic:

$$Z = \frac{\bar{x} - \bar{y}}{\sigma\sqrt{2/n}}$$

VI-40

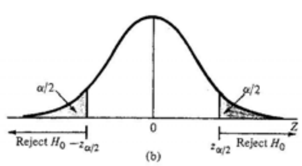
CONTINUOUS OUTCOME

Under the null hypothesis of no treatment effect:

$$Z \sim N(0,1)$$

Hence, we reject the null hypothesis when:

$$|Z| > z_{\alpha/2}$$

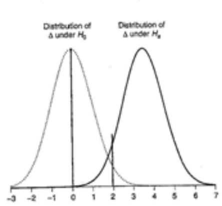


VI-41

CONTINUOUS OUTCOME

Under alternative hypothesis that $\mu_1 = \mu_2 + \delta$ (where δ is a clinically meaningful difference), the distribution is centered away from 0.

Power is the area under the alternative distribution that lies in the rejection region.



VI-42

CONTINUOUS OUTCOME

For given α, β, δ, and σ, the total required sample size is given by:

$$2N = \frac{4\left(Z_{\alpha/2} + Z_{\beta}\right)^2 \sigma^2}{\delta^2}$$

NOTE: This formula is based on a normal (not a t) distribution and assumes either σ is known or *N* is large enough to make this assumption valid.

VI-43

CONTINUOUS OUTCOME

Example:

- In a study of a new diet to reduce cholesterol, a 10 mg/dl difference would be clinically significant.

  δ = 10

- From other data, σ is estimated to be 50 mg/dl.

  σ = 50

- We want a two-sided test with equal sample sizes, α = 0.05, and we desire 90% power.

  $\mathbf{Z_{\alpha/2} = 1.96, Z_{\beta} = 1.28}$

VI-44

CONTINUOUS OUTCOME

Substituting those values into the formula gives:

$$2N = \frac{4\left(1.96 + 1.28\right)^2 (50)^2}{(10)^2} = 1049.8$$

Rounding up yields a required sample size of 2*N* = 1050, or *N* = 525 in each group.

VI-45

CONTINUOUS OUTCOME

How different would the required sample size be if σ were actually 60:

$$2N = \frac{4(1.96+1.28)^2(60)^2}{(10)^2} = 1511.7$$

Rounding up yields a required sample size of 2*N* = 1,512, or *N* = 756 in each group.

This is a big difference in the required sample size considering the relatively small increase in σ.

Be *conservative* in estimates of σ!!

VI-46

DICHOTOMOUS OUTCOME

Compare Drug A (standard) vs. Drug B (new)

p_A = Proportion of failures expected on drug A

p_B = Proportion of failures on drug B which one would want to detect as being different

We want to test

$$H_0: p_A = p_B \text{ vs. } H_A: p_A \neq p_B$$

With significance level α and power = 1−β to detect a difference of $\delta = p_A - p_B$.

VI-47

DICHOTOMOUS OUTCOME

The estimates of p_A and p_B are:

$$\hat{p}_A = r_A / N \quad \text{and} \quad \hat{p}_B = r_B / N$$

With r_A and r_B the number of events in the two groups and *N* the number of subjects in each group.

The usual asymptotic test statistic is:

$$Z = \frac{(\hat{p}_A - \hat{p}_B)}{\sqrt{2\hat{\bar{p}}(1-\hat{\bar{p}})/N}}$$

where $\hat{\bar{p}} = (\hat{p}_A + \hat{p}_B)/2$

VI-48

DICHOTOMOUS OUTCOME

The total sample size required (N in each group) is:

$$N = \frac{\left\{ Z_{\alpha/2}\sqrt{2\overline{p}\left(1-\overline{p}\right)} + Z_{\beta}\sqrt{p_A\left(1-p_A\right)+p_B\left(1-p_B\right)} \right\}^2}{\left(p_A - p_B\right)^2}$$

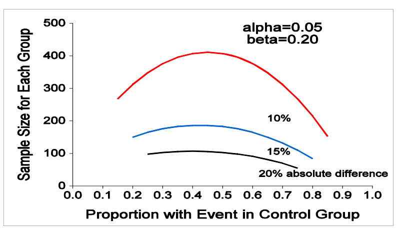
where $\overline{p} = \left(p_A + p_B\right)/2$ and $Z_{\alpha/2}$ and Z_{β} are critical values of the standard normal distribution.

VI-49

DICHOTOMOUS OUTCOME

In general, the variance is largest when $p = 0.5$ and smallest when p is near 0 or 1.

Hence, larger sample sizes are required to detect a change in p_A-p_B when p_A and p_B are near 0.5.



VI-50

DICHOTOMOUS OUTCOME

Example:

- In a clinical trial, the cure rate for the active control agent is assumed to be 65%.

  $p_A = 0.65$

- We want to detect an increase of 20% in cure rate.

  $p_B = 0.85 \rightarrow \delta = (0.85 - 0.65) = 0.20$

- We want a two-sided test with equal sample sizes, α = 0.05, and 80% power.

  $Z_{\alpha/2} = 1.96$, $Z_{\beta} = 0.84$

VI-51

DICHOTOMOUS OUTCOME

Substituting those values into the formula gives:

$$n = \frac{\left[z_{\alpha/2}\sqrt{2\bar{p}(1-\bar{p})} + z_{\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{\delta^2}$$

$$= \frac{\left[1.96\sqrt{2(0.75)(1-0.75)} + 0.84\sqrt{0.65(1-0.65)+0.85(1-0.85)}\right]^2}{(0.85-0.65)^2}$$

$$\approx 73$$

Hence, we require a total sample size of 73 in each group (146 total).

VI-52

TIME TO EVENT OUTCOME

- Elements of the problem
  - Endpoint: time to some event
    - Time to event: survival time
    - Event: failure (deaths, relapses, etc)
  - Required number of failures
  - Total duration of trial: entry (accrual) period and the follow-up period
  - Entry and loss to follow-up rates
  - Hazard rate for each treatment

VI-53

SURVIVOR FUNCTION

- The survivor function, S(t), gives the probability that a person survives longer than some specified time t.
- S(t)=Pr(T>t)
- $S(t_{(0)})=1$, $S(t_{(n)})=0$
- S(4)=0.35 means that 35% of the population survive beyond 4 years.

VI-54

HAZARD FUNCTION

- h(t) =instantaneous potential per unit time for the event to occur, given that the individual survived up to time t.
- $h(t)=\lim_{\Delta t->0} P(t \leq T < t+ \Delta t | T \geq t)/ \Delta t$
- Hazard function is rate and not probability
- $h(t) \geq 0$ and has no upper bound

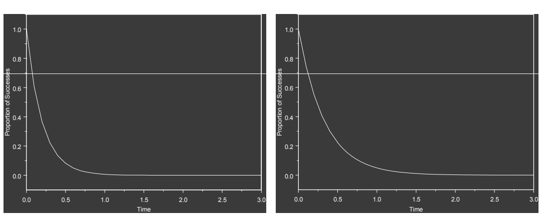
VI-55

Relationship between Hazard and Survivor Functions

- If know one, can determine the other directly
  Exponential Distribution:
  If $h(t)= \lambda$ if and only if $S(t)=e^{-\lambda t}$
- $h(t) = -[(dS(t)/dt)/S(t)]$
- $S(t)= \exp[-\int_0^t h(u)du]$

VI-56

EXPONENTIAL DISTRIBUTION



VI-57

HYPOTHESIS: TIME TO EVENT OUTCOME

- H_0: $\lambda_C = \lambda_E$ vs. H_A: $\lambda_C > \lambda_E$

  Define hazard ratio as:

  $\Delta = \lambda_E/\lambda_C = M_C/M_E$ if we assume exponential failure distribution

- H_0: $\Delta=1$ vs. H_A: $\Delta <1$

VI-58

TIME TO EVENT OUTCOME

In order to compare the groups we need to have a reasonable number of *events*, NOT total observations.

Hence, sample size calculations for comparing two survival curves consists of a two step process:

1) Calculating the Required Number of Events
2) Calculating the Required Number of Patients

Furthermore, the required sample size depends on the accrual and follow-up time for the study.

VI-59

NUMBER of EVENTS

To determine the required number of events, we need to specify:

β_* = Effect (log HR) we wish to detect
α = Significance level used for test
P = Target power
π_1 = Proportion of observations in group 1

$$\text{required \# of events} = \frac{\left(z_{\alpha/2} + z_{\beta}\right)^2}{\pi_1\left(1-\pi_1\right)\beta_*^2}$$

VI-60

NUMBER OF PATIENTS

To calculate the required number of patients to be enrolled, we need to consider the probability of the event over the course of the study.

Once probability of the event has been determined, the required number of subjects can be found from:

$$\text{required sample size} = \frac{\text{required \# of events}}{\Pr\{\text{event}\}}$$

VI-61

NUMBER OF PATIENTS

- A computer program is required to obtain number of patients and the duration of the trial to attain the required Type I and II error rates for given hazard rates

VI-62

NUMBER OF PATIENTS

Accrual period in years assuming different accrual rates and hazard ratios

Δ	60	80	120	160	180
1.20	16.62	12.68	8.73	6.75	6.09
	21.95	16.68	11.40	8.76	7.88
1.25	11.42	8.79	6.14	4.79	4.34
	14.98	11.46	7.93	6.15	5.56
1.30	8.55	6.63	4.68	3.68	3.34
	11.13	8.57	6.00	4.70	4.26
1.35	6.78	5.29	3.77	2.98	2.71
	8.76	6.80	4.81	3.78	3.43

Median in control group=1 year, follow-up is assumed to be 1 year. Upper numbers are based on two-sided type I error rate=0.05, power=80%, whereas lower numbers are based on two-sided type I error rate=0.05, power=90%.

VI-63

PROPORTIONAL HAZARDS MODEL

Two Groups

$$h(t,x_i) = h_0(t)\cdot\exp(\beta x_i)$$

x_i = 1, if new treatment
= 0, if standard treatment

Hazard for person *i* at time *t* is a function of:

- $h_0(t)$: the hazard for those on the standard treatment, *i.e.* $x_i = 0$
- A linear function of group membership (x_i)

VI-64

PROPORTIONAL HAZARDS MODEL

The hazards for subjects in the two treatment groups are:

Standard Treatment ($x_i = 0$): $h(t,0) = h_0(t)$

New Treatment ($x_i = 1$): $h(t,1) = h_0(t)\cdot\exp(\beta)$

Hence, to compare the hazards for an individual on the new treatment vs. one on the standard treatment:

$$HR = \frac{h(t,1)}{h(t,0)} = \frac{h_0(t)\exp\{\beta\}}{h_0(t)} = \exp(\beta)$$

VI-65

PROPORTIONAL HAZARDS MODEL

Hence, a unit increase in *x* multiplies the hazard by an amount that is constant over time:

$$HR = \exp(\beta)$$

Hence, the log-hazard ratio (β) is an unknown coefficient that describes the way survival time is affected by the covariate:

- $\beta = 0$: no effect
- $\beta > 0$: survival is worse with new treatment
- $\beta < 0$: survival is better with new treatment

VI-66

SOFTWARE

Software for power calculations (among many):

- Commercial packages:
  - SAS (PROC POWER)
  - NCSS PASS
  - NQuery
- Free packages:
  - Dr. Russell Lenth's website: http://www.stat.uiowa.edu/~rlenth/Power/index.html
  - PS: Power and sample size calculation http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize

VI-67

Data and Safety Monitoring Boards

DSMBs are often given the responsibility of monitoring the accumulating data.

The DSMB is responsible for assuring that study participants are not exposed to unnecessary or unreasonable risks.

The DSMB is also responsible for assuring that the study is being conducted according to high scientific and ethical standards.

VI-68

Data and Safety Monitoring Boards

Why have DSMBs?

- **Protect safety of trial participants**
- Investigators are in a natural conflict of interest
  - Vested in the study
  - They, and their staff, are paid by the study
- Having the DSMB externally review efficacy and safety data protects:
  - The credibility of the study
  - The validity of study results

VI-69

INTERIM MONITORING

Principle 1 – Composition. The DSMB should have multidisciplinary representation, including topic experts from relevant medical specialties and biostatisticians.

Principle 2 - Conflicts. Individuals with important conflicts of interest (financial, intellectual, professional, or regulatory) should not serve on a DSMB.

Principle 3 – Confidentiality Issues. Trial integrity requires DSMB members not to discuss details of meetings elsewhere.

VI-70

INTERIM MONITORING

DSMB's should periodically review study data.

The study protocol should include a section describing proposed plan for interim data monitoring.

This plan should detail:

- What data will be monitored?
- The timing of all interim analyses?
- The frequency of data reviews.
- Criteria that will guide early termination

VI-71

INTERIM MONITORING

Early DSMB meetings almost exclusively focus on:

- Quality of conduct
  (recruitment, timeliness of data entry, etc.)
- Trial integrity
  (protocol adherence, etc.)

As more data accrue, DSMB meetings will focus on safety issues as well.

Later DSMB meetings may include formal efficacy or futility analyses.

VI-72

INTERIM MONITORING

At end of each meeting, DSMB also summarizes any areas of concern regarding performance and/or patient safety.

Soon thereafter, the DSMB chair will provide a written summary of the board's recommendations.

These letters are extremely important for IRB submissions at each individual site.

VI-73

INTERIM MONITORING

Ethical principles mandate that clinical trials begin with uncertainty as to which treatment is better. (clinical equipoise)

This uncertainty should be maintained during study.

If interim data become sufficiently compelling, ethics would demand that the trial stop and the results made public.

Hence, interim monitoring of safety and efficacy data has become an integral part of modern clinical trials.

VI-74

INTERIM MONITORING

Early termination of a trial should be considered if:

- Interim data indicate intervention is harmful
- Interim data demonstrate a clear benefit
- Significant difference by end of study is probable
- No significant difference by end of study probable
- Severe logistical or data quality problems exist

VI-75

INTERIM MONITORING

The decision to stop a trial early is complex, requiring a combination of statistical and clinical judgment.

Stopping a trial too late means needlessly delaying some participants from receiving the better treatment.

Stopping a trial too early may fail to persuade others to change practice.

Group sequential designs have been developed for interim monitoring of clinical trials to minimize the role of subjective judgment.

VI-76

EFFICACY MONITORING

Consider a clinical trial to compare two normally distributed groups with K interim analyses.

The objective of the trial is to test the null hypothesis of no treatment effect at each interim analysis:

$$H_0: \delta = 0 \text{ vs. } H_A: \delta \neq 0$$

where δ equals difference between treatment means.

At each interim analysis, the null hypothesis is tested using the test statistics $Z_1,\ldots,Z_K$ (Z-statistic for all data observed up to time of k^{th} interim analysis)

VI-77

REPEARTED TESTS of SIGNIFICANCE

Under H_0 (no difference between groups), repeated testing at level α inflates the probability of making at least one type I error.

Even 5-10 tests can lead to serious misinterpretation of trial results.

# of tests	True type I error rate
1	0.05
2	0.08
5	0.14
10	0.19
20	0.25
1000	0.53

VI-78

EFFICACY MONITORING

Solution is to adjust stopping boundaries in such a way to ensure that overall type I error is equal to α:

- Pocock (1977):
  Same critical value at each interim look
- O'Brien & Fleming (1979):
  Nominal significance levels needed to reject H_0 increase as study progresses.
- Haybittle (1971) & Peto et al. (1976):
  Reject H_0 if $|Z_k| \geq 3$ for all interim tests ($k < K$)

VI-79

EFFICACY MONITORING

A comparison of the critical values for the Pocock, O'Brien-Fleming, and Haybittle-Peto methods for *k* = 5 looks and α = 0.05 is given below:

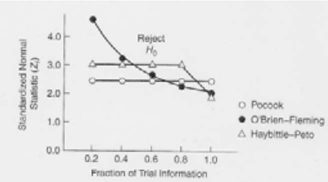


Figure 8.3 Haybittle–Peto, Pocock and O'Brien–Fleming group sequential boundaries for live interim analyses at the one-sided 0.025 significance level. open circle = Pocock; filled circle = O'Brien–Fleming; triangle = Haybittle–Peto, group sequential boundaries.

VI-80

EFFICACY MONITORING

There is a *slight* loss of power with multiple testing.

To account for this, sample size calculations must adjust the sample size upward.

This is accomplished by the following process:

- Compute the required sample size under a fixed sample design.
- Multiply this sample size by an appropriate ratio to account for the multiple testing.

VI-81

EFFICACY MONITORING

The original methodology for group sequential boundaries required that the number and timing of interim analyses be specified in advance.

DSMB's sometimes may require more flexibility as beneficial or harmful trends emerge.

Lan & DeMets (1983, 1989) proposed an 'alpha spending function' which provides more flexible group sequential boundaries.

The approach lends itself well to the accommodation of irregular, unpredictable, and unplanned interim analyses.

VI-82

FUTILITY MONITORING

Power tells whether a clinical trial is likely to have high probability to detect a pre-defined treatment effect of interest.

Very low power implies that a trial is unlikely to reach statistical significance even if there is a true effect.

One should never begin a trial with low power.

However, sometimes low power becomes apparent only after a trial is well under way.

VI-83

FUTILITY MONITORING

Stochastic curtailment uses the concept of conditional power:

$P_k(\theta)$ = Pr{ reject H_0 | θ and observed data so far }

Initially, when k = 0, this is the usual power function.

At the planned termination of the study (stage K), this probability is either 0 or 1.

At interim stage k, conditional power depends on θ.

May want to stop trial for futility if the conditional power drops below some specified level (i.e., 20%).

VI-84

FUTILITY MONITORING

If early results show:

- Intervention better than expected
  → conditional power high
- Intervention worse than expected
  → conditional power low
  (unless sample size increased)

Group sequential methods focus on existing data.

Stochastic curtailment methods consider future data.

VI-85

FUTILITY MONITORING

Clearly, the futility rule is heavily influenced by the assumed value of the treatment difference, θ.

Making an overly optimistic assumption about θ delays decision to terminate the trial.

Several options for the value of θ have been proposed:

- Lan, Simon, & Halperin (1982): Evaluated at value of θ corresponding to alternative hypothesis.
- Evaluated under the null hypothesis.
- Evaluated at the observed treatment effect

VI-86

FUTILITY MONITORING

One limitation of conditional power is that no adjustment is made to account for associated prediction error if observed treatment effect is used.

Interim futility monitoring may also be conducting using other approaches:

- Predictive Power: Mixed Bayesian-Frequentist approach
- Predictive Probability: Bayesian approach

VI-87

SOFTWARE

Software packages for group sequential methods:

- S+SeqTrial (Insightful Corporation)
- EaST (Cytel)
- PEST 4 (University of Reading)
- LanDeM (University of Wisconsin)
- SAS (through the use of Macros)

VI-88

ADAPTIVE DESIGNS

There may be limited information to guide initial choices for study planning.

Since more knowledge will accrue as the study progresses, adaptive designs allow these elements to be reviewed during the trial.

An adaptive design allows for changing or modifying the characteristics of a trial based on cumulative information.

VI-89

ADAPTIVE DESIGNS

Adaptive designs are NOT new.

The broad definition includes topics such as group sequential designs and covariate adaptive randomization techniques.

However, because this is a rapidly expanding area of research, more practical experience is needed.

Both Bayesian and Frequentist approaches should be considered.

VI-90

SUMMARY

The size of a study should be considered early in the planning phase.

Fundamental Principle: Clinical trials should have sufficient statistical power to detect differences between groups considered to be of clinical interest.

Therefore, calculation of sample size with provision for adequate levels of significance and power is an essential part of planning.

VI-91

SUMMARY

There are a variety of approaches for interim monitoring of clinical trial data.

The relationship between clinical trials and practice is very complex, and this complexity is evident in the data monitoring process.

The appropriate monitoring plan depends on the goals of the trial.

VI-92

SUMMARY

Because of the repercussions of stopping a trial early, the decision to stop a trial is complex and requires both statistical and clinical judgment.

Hence, these methods should not be used as a sole basis in the decision to stop or continue a trial.

Other considerations that play an important role in decision making process cannot be fully addressed within the statistical sequential testing framework.

VI-93