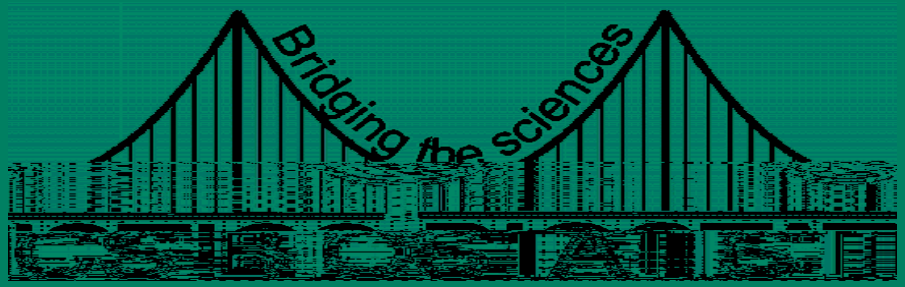
Designing Clinical Trials



Phases of Human Research

- Ideal clinical trial: randomized, double-blinded
- Traditionally four phases of research
  - I=Establish safety, dose-finding, PK studies
  - II=Establish biological activity or potential efficacy
  - III=Randomized comparison of treatment
  - IV=Long-term surveillance in broader population
  - Some hybrid phases (I/II, II/III)

Phase I Trials

Maximum Tolerated Dose:

- ***Highest possible dose with an "acceptable" rate of dose-limiting toxicities***
- ***"acceptable" toxicity rate set by investigators because of potential for benefit to patients***
- ***Example:***
  - ***Dose which produces grade III or worse toxicities in no more than 1 in 3 patients (cancer studies)***

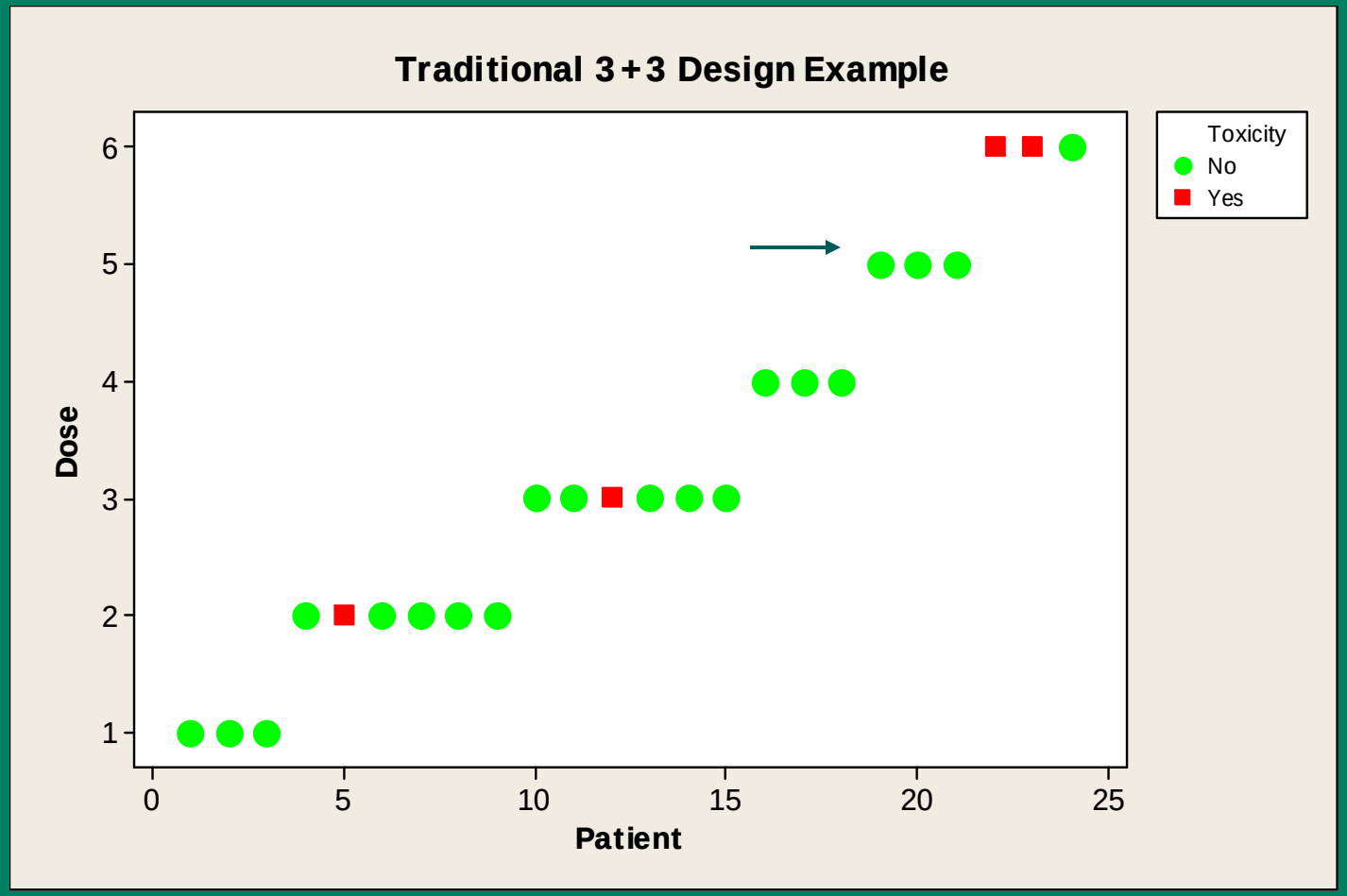
CTSI Biostatistics

Traditional 3+3 Design

Patients enrolled in groups of 3 for each dose level.

- ***None of 3 experience DLT– increase dose level***
- ***1 of 3 experience DLT - enroll additional 3 patients***
- ***>1 of 3 experience DLT – stop; MTD was reached at previous dose***
- ***If <2 out of 6 experience DLT then increase dose level***
- ***If 2 or more out of 6 experience DLT, then MTD was reached at previous dose.***

Traditional 3+3 Design



Traditional 3+3 Design

Drawbacks of 3+3 design

- Biased estimate of MTD (too low)
- MTD defined based on 33% DLT rate
- Moves 64msb03 Tc bw0055 Tw 0.dh Twe

Other Phase I Designs

Continual Reassessment Method (O'Quigley, Biometrics 1990):

- ***After each patient's outcome is known, re-fit the dose-toxicity curve***
- ***Estimate MTD based on the estimated curve***
- ***Enroll next patient at estimated MTD.***
- ***Many modifications to this principle***
- ***More flexible, better statistical properties than 3+3 design***
- ***Requires collaboration with statistician***

Phase II Trials

- Goal: Determination if treatment has any biological activity, estimation of biological activity, estimation of rate of adverse events
- Single well described treatment regimen
- Small sample sizes - <40-50
- Used to select treatment for Phase III trials
- Tighter inclusion/exclusion criteria than Phase III trial – who will benefit most?
- "Quicker" outcome measures if possible
- Often two-stage design

CTSI Biostatistics

CTSI Biostatistics

Two-Stage Designs

- Want to minimize the number of patients treated with an ineffective treatment
- Choose n_1 patients in first stage
  - If r_1 or fewer respond then declare the treatment a failure and stop
  - If $>r_1$ respond then add $n-n_1$ patients (total=n)
- If more than r respond out of n then declare the treatment a success; o/w declare it a failure
- Requires short-term primary endpoint

Two-Stage Design

Example: Suppose we want to compare a new treatment against a historical control with response rate 0.1. Want 80% power when response rate is 0.3, with α=0.05

EN=Expected sample size

PET=Probability of Early Termination

Two-Stage Designs

Simon (Cancer Treatment Reports, 1985)

Calculator for these designs can be obtained online at
http://linus.nci.nih.gov/brb/samplesize/otsd.html

Phase III Trials: Design Features

- Testable hypothesis
  - Outcomes
  - Patient population
- Design
  - Control group
  - Randomization
  - Blinding
  - Data Monitoring

What is the Question?

- ***Includes***
  - ***Treatment Comparison***
  - ***Primary Outcome***
  - ***Patient Population***
- ***Usually one primary question; clearly defined and stated in protocol***
- ***Secondary questions***

Primary Endpoint Example

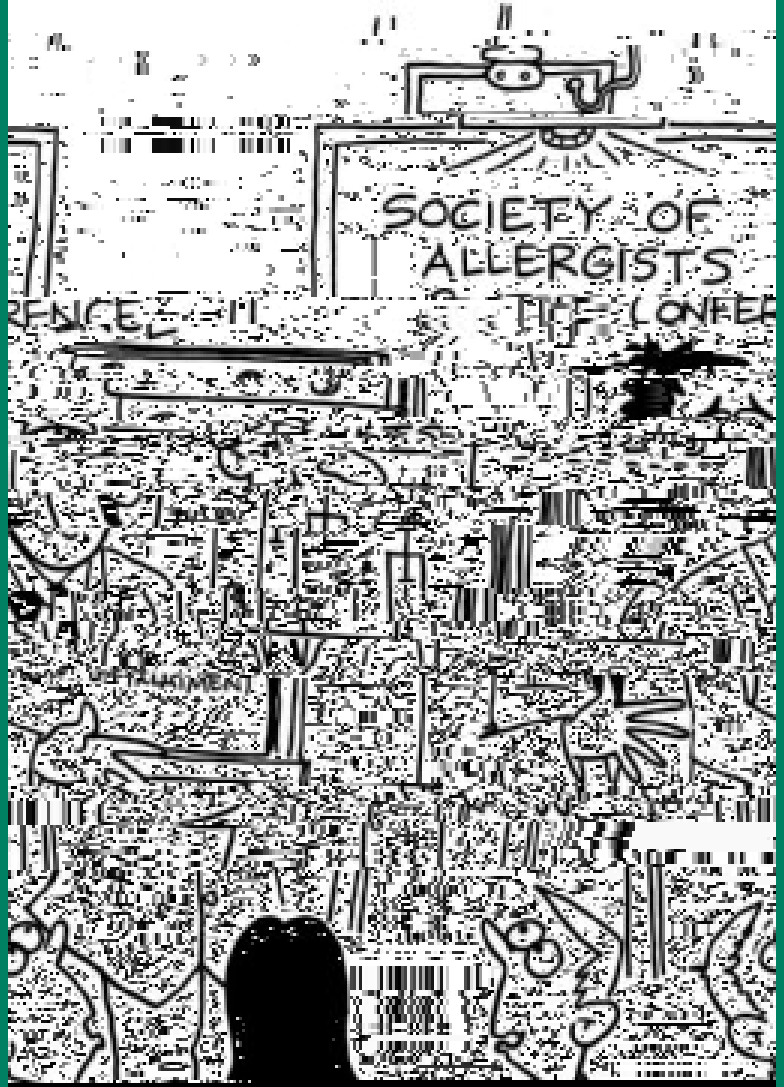
HIV studies: HIV virus destroys immune system, individuals subject to infections which lead to death; now treatments target the virus or specific infections

Possible endpoints:

1. ***Increase in CD4 counts (direct measure of immune function)***
2. ***Viral RNA reduction (measures amount of virus in the body)***
3. ***Time to one or more opportunistic infections***
4. ***Time to death from any cause***
5. ***Time to first infection or death***

CTSI Biostatistics

Sometimes secondary endpoints can be important



Secondary Questions

- Components of composite endpoint
- Other endpoints
- Subgroup analyses
- Toxicity
- Ancillary study questions
- Surrogate outcomes
- Multiple Testing:
  - If enough tests are done, some will be significant by chance alone when there is no intervention effect
- Secondary hypotheses are better at shedding light on results, or on generating new hypotheses; should not be considered definitive results

Subgroup Analysis

- Trial should have reasonable expectations that intervention will be consistent across subgroups; otherwise, do separate studies
- Is effect of treatment in a clinical trial homogeneous across all patients in that trial?
- Example 1: ISIS-2 Trial: 17000 patients with AMI randomized to placebo vs. aspirin (also streptokinase) (Lancet, 1988)
  - Mortality within 1 month: 9% (aspirin) vs. 12% (placebo), $p<0.001$
  - Investigators were urged (by editors) to conduct nearly 40 subgroup analyses
  - Investigators agreed on condition that they could conduct their own subgroup analysis to illustrate unreliability of subgroup findings

Subgroup Analyses

Subgroup defined by astrological sign

# of deaths in 1 month

Astrological sign

Increased variability of results just due to chance when you look at a lot of subgroups.

Excess of type II errors due to multiple comparisons

Primary Question: Example

- Thompson (Blood, 2008) SWOG 9438
- Treatment:
  - Post-transplant therapy with IL-2 versus observation after TBI, etoposide, cyclophosphamide and PBSC autotransplant
- Patient Population:
  - Relapsed NHL
- Outcome: Progression-free survival

Secondary Questions: SWOG NHL Example

- **Secondary Endpoints**
  - OS, Toxicity
- **Subgroup analyses**
  - Test for interaction between chemosensitivity and outcome, or between histology and outcome
  - No significant interactions

Patient Population

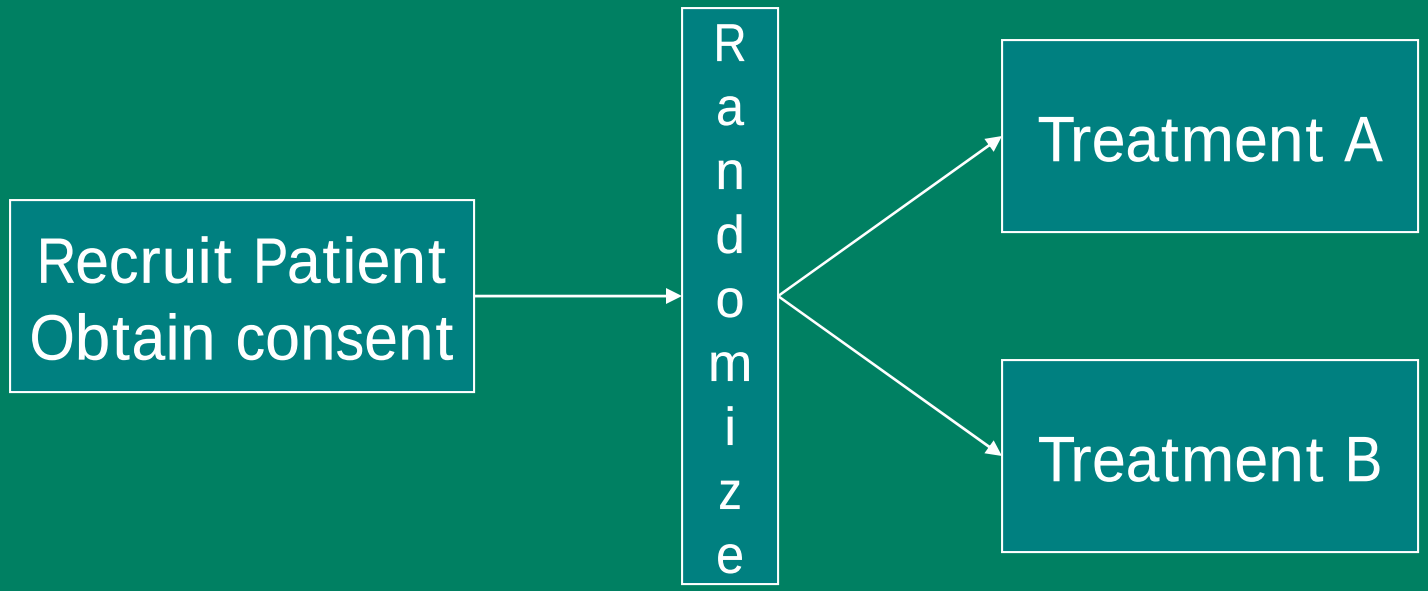
The study population should be defined in advance using clearly defined inclusion/exclusion criteria

Phase II vs. phase III

Patient Population

Randomized Phase III Design Schema

Each patient randomized to one of two treatments:



Randomization

- Process by which all participants are equally likely to be assigned to either the intervention or control groups
- Advantages:
  - Protects against bias in treatment allocation and selection of patients
  - Makes groups comparable or “similar on the average”, balancing out known and unknown prognostic factors
    - An individual study may still be imbalanced

Methods of Randomization

- Simple randomization (Coin flip)
  - Allows for unusual results in treatment assignment
- Random permuted blocks ensure balance over accrual time
  - ABBA | BABA | BBAA | ABAB …
- Stratification ensures balance across important covariates known to be associated with outcome

Randomization Issues

When to randomize: As close to treatment as possible to minimize dropout.

Intention to Treat Principle: All eligible patients are analyzed according to the arm to which they are randomized, even if they deviate from the therapy

Example (cont.)

- Registration prior to tx
- Randomization took place between day 28 and day 80 post tx
- Additional eligibility criteria regarding patient's status at time of randomization (no recurrence/ progression of disease, no serious toxicity)
- 394 patients initially registered, 376 eligible
- 194 were randomized
- 182 eligible but no randomization
    - 46 Refusal
    - 62 grade 4 or 5 toxicity
    - 28 progression

Design Issues-Blinding

- Why patient?
  - May experience a psychological benefit of being on new treatment
  - Attitude toward treatment may affect cooperation in study (compliance with therapy, attendance for evaluation, drop out rates)
  - Attitude toward treatment may affect response

Design Issues - Blinding

- Why Treatment team?
  - Knowledge of treatment may influence:
    - Decisions on dose modification
    - Decisions on intensity of patient examination
    - Decisions on when to discontinue treatment
    - Decisions on ancillary care
  - Enthusiasm for treatment may be expressed to patient

Sample Size

- Patients respond differently to treatment; we are interested in collective effect of treatment
- Larger sample sizes increase precision of estimated treatment effect.
- Calculations based on primary endpoint: Specify likelihood of detecting a significant treatment effect of a particular magnitude (Power)
- Example: Want 80% power to detect if new treatment increases 2 year progression free survival from 40% to 55%

Sample Size

Need to specify:

- ***Type I error***
- ***Power***
- ***Success proportion for control (p_1) and intervention (p_2)***

Formula: N per group is

$$N = 2(z_\alpha + z_\beta)^2 p_c(1 - p_c)/(p_2 - p_1)^2$$

Here

- ***z_α=1.96 for a two-sided 5% level test***
- ***z_β=1.28 and 0.84 for 90% and 80% power***
- ***p_c is the average success proportion***

Sample Size

Tables

- *Hulley et al. (1988)*
- *Example: For P_1=0.35,P_2=0.55,Power=80%, type I error=5%,*
  - *n=96 per arm*

Free software/calculators on Internet

http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize
http://www.math.uiowa.edu/~rlenth/Power/

Other methods for different endpoints

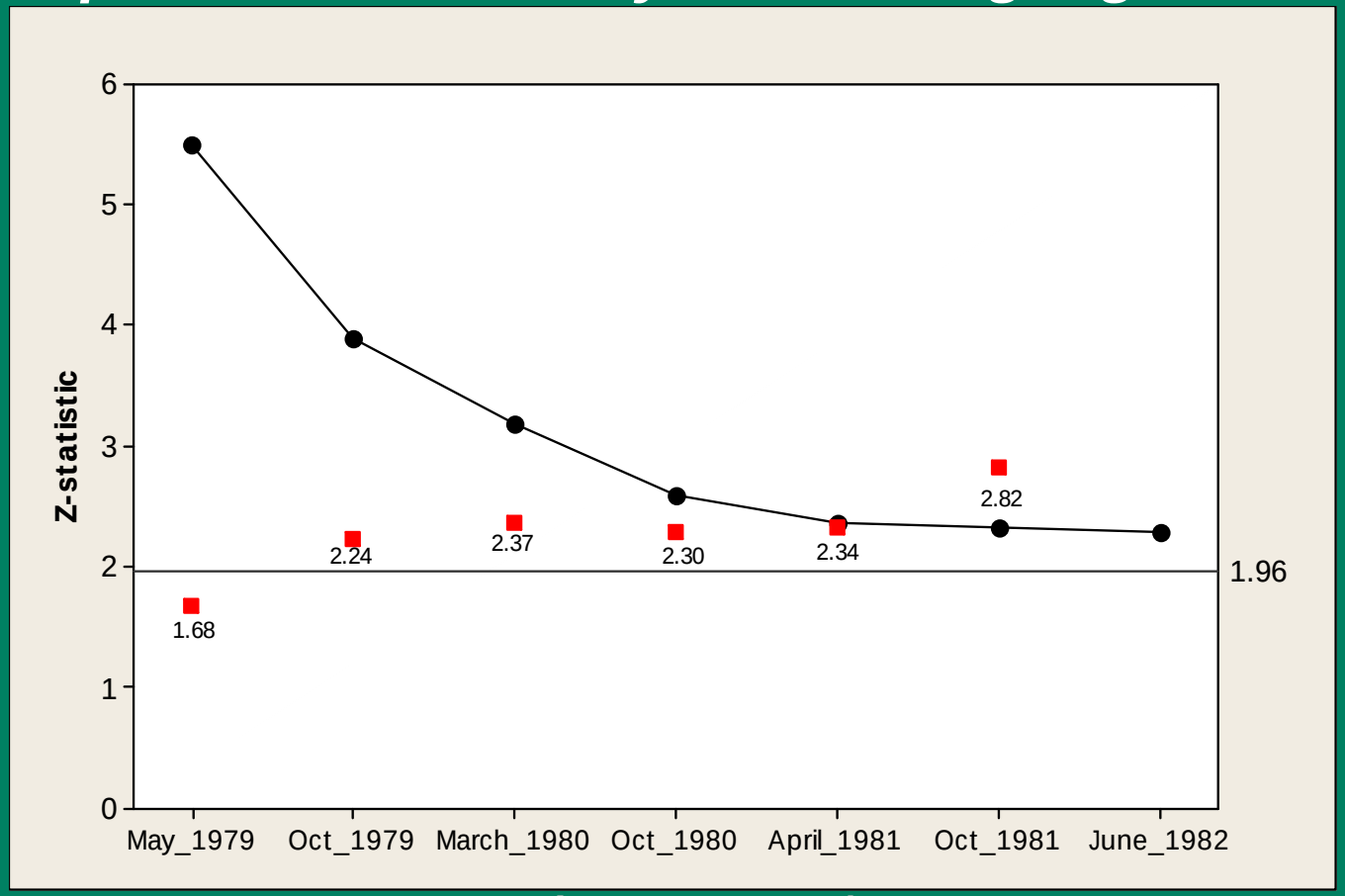
- *Continuous Normal outcomes*

Data Monitoring

Data Safety Monitoring Committee: Responsible for ensuring patient

Beta Blocker Heart attack trial (DeMets CCT 1984)

Comparison of mortality rates using log-rank test



Basic Data Items and Forms

- Detailed in protocol
- Basic variables:
  - Eligibility
  - Evaluability
  - Treatment Summary
  - Outcome Summary
- Forms
  - Study-specific flow sheet
  - Standard off-treatment notice/notice of death
  - Prestudy form (baseline characteristics)

Sample Protocol (Friedman et al. 1998)

1. Background of the Study
2. Objectives
    1. Primary question and response variable
    2. Secondary questions and response variables
    3. Subgroup hypotheses
    4. Adverse effects
3. Design of the Study
    1. Study Population-Inclusion/Exclusion criteria
    2. Sample size assumptions and estimates
    3. Enrollment of participants
        1. Informed consent
        2. Assessment of eligibility
        3. Baseline examination
        4. Intervention allocation (randomization)

Sample Protocol

4. Intervention
    1. Description and Schedule
    2. Measures of compliance
    3. Dose modification and withdrawals
5. Follow-up visit description and schedule
6. Ascertainment of response variables
    1. Data collection and forms
    2. Quality control
7. Data Analysis
    1. Interim Monitoring
    2. Final analysis
8. Termination Policy

4. Administrative organization

Resources

CTSI Biostatistics

Upcoming Lectures