Presented to The Annual DIA Workshop on Statistical Methodology in the Biopharmaceutical Sciences
Philadelphia, PA 03 March 2004

Adaptive Randomization: Promise with Problems
Russell W. Helms, Ph.D.
Rho, Inc.
Chapel Hill, NC
rhelms@RhoWorld.com
www.RhoWorld.com

Before we get started: STAT SIAC

- SIAC: Special Interest Area Community
  - DIA reorganizing: grassroots
  - Newly functioning
  - Needs Assessment
  - Volunteers?
    - Website, newsletter, networking, programming, training
    - Talk to me!
    - Email me: rhelms@RhoWorld.com
    - Email DIA

Adaptive Randomization: Promise with Problems

- Agenda
  - Introduction: What and Why?
  - Types of Randomization
  - Randomization and Inference
  - Advice: What (not) to Do in Practice
  - Conclusions
- Special Thanks
  - Rosenberger and Lachin (2002)
  - McEntegart (2003 DIJ)

Randomization: What & Why?

- What is randomization...?
  - The act of assigning treatments to observational units using (an element of) chance.
  - ...and why do we do it?
  - "...two adequate and well controlled studies..."
  - RCT is the gold standard
  - Control bias, basis for inference
- Related topics: controls, masking

Controls: What & Why?

- Now a gold standard, but...
  - Salk: "I found but one person who rigidly adhered to the idea of a placebo control and he is a bio-statistician who, if he did not adhere to this view, would have had to admit his own purposelessness in life."
  - FDA helps
- Placebo effect, paper challenging it
- Better than what?
  - New information is obtained from a comparison of alternate states
  - Association and causation

Masking: What & Why?

- Masking = Blinding = Veiling
- Masking helps control sources of bias
  - Selection bias: investigators may be able to bias composition of treatment groups by choosing (not) to enroll based on what treatment a particular patient may receive
  - Subconscious preference more likely than deliberate dishonesty
  - e.g., Good prognosis versus bad prognosis
  - e.g., "Likely" to respond to this treatment
  - Worse: investigators (or subjects) may grade outcomes differently
  - Eliminates placebo effect?
Masking: What & Why?

• Masking is not always possible.
  – Emergency unmasking
  – Surgery (Story: Arthroscopic surgery trial for osteoarthritis)
  – Even when planned, patients and investigators are smart, and discern.
  • Side effects, lab tests, response... (Story: Good Friday experiment)
  • Plenty of examples in literature, my experience
  – Implication on randomization and guessing

Randomization: What & Why?

• Why random allocation? 1: Control Bias (throughout)
  – e.g., selection bias
    • Story: Teachers switching treatment groups
  • From Chalmers et al., 1977: 32 Studies of anticoagulants for acute MI. Mortality improvement in:
    – 15/18 studies using historical controls (900 pts), 5/8 studies using nonrandomized concurrent controls (3,000 pts).
      » Pooled: 50% reduction in mortality
    – 1/6 RCT (3,800 pts).
      » Pooled: 20% reduction in mortality
    – Bias > effect?
  • Large effect versus small effect: Other publications show examples of inappropriately favoring new interventions

Randomization: What & Why?

• Why random allocation? 2: Basis for inference (section later)

Randomization: What & Why?

• Issues specific to RCT in Drug Development...
  – Sequential appearance of potential subjects
  – Inclusion/Exclusion criteria
  – Multicenter trials
  – Phases with different goals
  – Implemented by very intelligent people
  – Implemented by people with (sometimes poor) judgment
  • Story: Investigator "too busy" to call CIRS
  • Rescue Story: Investigators "forgot" => multiple randomizations
  – Studying people with independent will: plan for some chaos

Four Types of Randomization

• Complete randomization
  – Simple coin tossing
• Restricted randomization (e.g., permuted blocks)
  – Complex coin tossing: add restrictions on the coin and how it is used to ensure balance of treatment arms
• Covariate-adaptive randomization (e.g., Pocock and Simon)
  – Used to ensure balance of treatment arms with respect to certain known covariates
• Response-adaptive randomization
  – Treatment assignments depend upon previous patient responses
Types of Randomization: Complete Randomization

- Complete randomization
  - Simple coin tossing
  - Not adaptive
  - No selection bias: impossible to predict
  - Chance of (possibly severe) treatment imbalance: one treatment may be assigned more frequently than another at any point in the trial, including the end.
  - Why want equal (balanced) allocation?
    * Max power: generally not much of an issue, but extreme imbalances, while rare, can be a serious problem. (Seek balance for insurance.)
    * Equipoise and ethics
    * Facilitates CTM supply management
    * Chronological bias: patients differ over time of enrollment
      - e.g., epidemics in vaccine or anti-infective trials

Types of Randomization: Restricted Randomization

- Restricted randomization
  - Complex coin tossing: add restrictions on the coin and how it is used to ensure balanced allocation of treatments. Very common.
  - Different assignments are somewhat dependent:
    * Adaptive at simplest level, but not by strict definition
    * Examples:
      - Random Allocation Rule:
        - Urn drawing without replacement. Avoid because:
          - Deterministic in late stages: selection bias, chronological bias, inference
      - Truncated Binomial design
        - Binomial until one treatment group “full”. Avoid because:
          - Deterministic in late stages: selection bias, chronological bias, inference
          - Possibility of severe covariate imbalances and resulting “accidental bias
      - Permuted blocks (more on next page)

Types of Randomization: Restricted Randomization

- Permuted Block design: most common in our industry
  - Promise: Balance at end and throughout: every n-th subject.
    [BABA] [AABB] [ABAB] … Important when:
    - time-heterogenous covariate related to treatment outcome: avoid chronological bias
    - interim analyses are planned
    - The trial might be stopped early.
  - Problems:
    - Overall imbalance possible with incomplete blocks
      - inference and “wasted” data (more later)
    - Large element of determinism (every block)
      - selection bias: secret and/or variable block size (worse?)
      - determinism and inference (potentially small reference set)

Types of Randomization: Restricted Randomization

- (Generalized) Biased Coin Designs
  - Examples
    * Effron’s Biased Coin design
      - Vary p from ½ when balance gets “bad”
      - Only two values of p (1/2 or other)
      - Some more deterministic modifications aim to prevent extreme imbalance
    * Wei’s Urn design
      - Adaptive biased coin design: p varies by degree of imbalance
        - Urn with replacement plus a few
      - Promise: balance, difficult to predict (selection bias)
      - Problems: analysis complications, chronological bias

Maximal Procedure (Berger, Ivanova, Knoll, 2003)

- Relatively new, relatively untried
- Maximal reference set subject to maximum tolerated imbalance condition
  - Focus on terminal balance, maximum tolerated imbalance, and ignore forced returns to exact balance (as in permuted blocks)
  - Can implement with complete randomization and replacement randomization (discarding sequences that don’t fit constraints)
- Promise: controls balance (chronological bias), difficult to predict (selection bias), not necessarily deterministic
- Problems: untested, explaining to investigators, IRBs, etc.
Types of Randomization: Covariate-Adaptive Randomization

- Often covariates or prognostic factors of interest besides treatment affect outcome.
  - e.g., site in a multi-center trial is often the greatest source of variability
  - e.g., infections: nosocomial versus community-acquired, resistant versus not
  - e.g., baseline severity of disease
  - Arthritis examples:
    - Duration of illness: 2-3 years versus 10+ years (potentially continuous)
  - Prior treatment history
- Accidental bias may result from imbalances on such factors.
  - Especially true for permuted block randomization, as the asymptotic decline of accidental bias is slow for that design (Rosenberger and Lachin)

Types of Randomization: Covariate-Adaptive Randomization

- Covariates besides treatment affect outcome:
- Solution 1: Stratification = “Separate” randomization for each level of factor, with any randomization method so far discussed.
  - Promise: stratification can protect from accidental bias (known covariates + correlates) and also increase power.
  - Note: stratification on randomization requires stratification in analysis; stratification in analysis is possible without stratification in randomization, similar power gains
  - Problems:
    - treatment balance within strata may reduce overall balance
    - Number of strata increases rapidly and can make overall balance impossible

Types of Randomization: Covariate-Adaptive Randomization

- Solution 2: Covariate-adaptive randomization
  - Randomization of subject # (k+1) depends on balance among first k subjects, assign k+1 the treatment that “minimizes” the imbalance
  - Zelen’s Rule: Use schedule but swap when balance “bad”
  - Wei’s marginal Urn design: similar to urn design with urns for margins of covariates, instead of an urn for each stratum
  - "Optimal" Design: Based on linear model, minimize variance of estimated treatment effect in presence of covariates.
    - Note different than balancing over covariates to mitigate bias.
    - Difficult to implement and to explain to nonstatisticians.

Types of Randomization: Covariate-Adaptive Randomization

- Pocock and Simon (sometimes called "minimization")
  - Independent generalization of Taves; incorporates stochastic element.
  - Measures imbalance on each factor then combines to get an overall measure of imbalance associated with each treatment. “Bad” imbalance implies a high probability (e.g., biased coin) of assigning a treatment.
  - Tip: Unlike stratification, can account for many levels of “strata”.
  - Metrics are arbitrary, but intuitive. Little research on properties, but used with some frequency.
    - Tip: simulation to discern properties for each implementation
    - Tip: Balances margins, not cells.
    - Metrics: Begg & Iglewicz, Atkinson, Frane (lead to “Optimal”)
    - Tip: Imbalance measure based on variance better than based on range: better avoid extreme imbalances
    - Tip: Include factors in analysis

Types of Randomization: Covariate-Adaptive Randomization

- Promise:
  - Insurance: extreme covariate imbalance; while unlikely, may make a trial difficult/impossible to interpret.
  - Enhances credibility for less statistically sophisticated reviewers who may misunderstand or mistrust model-based adjustments for imbalance
  - Increases precision and interpretability for interim analyses and early-stopped trials
  - Avoids analysis problems of empty cells, and
  - Handles very many factor levels

Types of Randomization: Covariate-Adaptive Randomization

- Problems:
  - Implementation, analysis can be complicated (inference)
  - Possibility of selection bias increases as p→1
  - Intuitive method, properties not (yet) fully explored
  - Doesn’t guarantee balance of unmeasured covariates (accidental bias), though simulation studies suggest it may help (Rosenberger & Lachin)
  - Balance may be overrated
Types of Randomization: Covariate-Adaptive Randomization

• In general,
  – The more stochastic, the better
    • ICH E9 Guidance: Use random element in all assignments.
    • Forcing treatment balance among covariates may be
      overrated: be quite sure it’s necessary before stratifying or
      implementing covariate-adaptive procedure.
    • EXCEPTION: Site
  – Many cells: stick with modified P&S: stochastic (e.g., p=0.8)
    • great balance and difficult to predict,
    • but unknown properties
  – Comparison to permuted block design:
    • P&S less chance of selection bias as long as p < 1
    • P&S less deterministic as long as p < 1
    • P&S may be better at preventing accidental, chronological bias

Types of Randomization: Response-Adaptive Randomization

• Treatment assignments depend upon previous patient
  responses, e.g.,
  – “Play-the-winner”: highly deterministic
  – “Two-armed bandit”: more like generalized biased coin
• Promise:
  – Ethics: maximize number of patients on superior treatment
• Problems:
  – Requires rapid ascertainment of patient outcomes
  – Lower power (larger sample size offsets ethical advantage)
  – Chronological bias. Story: ECMO example (1 F, 9 S, Stop)
  – Analysis complicated: inference next slide!
• Comment: Other CT designs (NonStop designs, interim
  analyses) can accomplish similar goals

Randomization and Inference

• A goal of randomization is to aid in inference from the
  trial. Two methods:
  – Population-Model Tests
    • Editorial: Practical, simulation-supported hand-waving
  – Permutation Tests (Randomization model)
    • Editorial: Mathematically rigorous complications

Randomization and Inference:

Population Model

Population A:

\[ Y \sim G(y | \theta_A) \]

Sample at Random

\[ n_A \text{ Subjects: } Y_{A1} \sim G(y | \theta_A) \]

H_0: \theta_A = \theta_B

Note: Adapted from (Lachin, 1988)

Population versus Permutation

• Approaches address different questions (null
  hypotheses).
  – Permutation: Probability that observed effect (or more) in n
    patients actually randomized into the trial could have
    occurred by chance
    • Allows conclusions about effects of treatment in patients actually
      studied.
  – Population: Probability that at the observed effect (or more)
    in n patients studied could have been observed in samples of n_A and n_B
    drawn at random from respective populations.
    • Allows conclusions about effects of treatment in hypothesized
      general population.
  – Thus generalization may be different, but this is really a
    bigger issue of trial design (inclusion/exclusion criteria, etc.)
### Randomization and Inference: Population versus Permutation

- Approaches address design differently:
  - Permutation tests: assumes outcomes fixed, treatment assignment is a random variable
  - Population model: treatment (populations) fixed, outcome is a random variable
- Approaches have different assumptions:
  - Permutation tests require virtually no assumptions.
  - Population model requires faith in untestable assumptions.
  - Population model based on Neyman-Pearson.
- Type of randomization matters for permutation tests, but is ignorable for population
  - This bothers some in the strict “analyze as randomize” crowd.
- Power and sample size must be based on population
  - Sometimes permutation test wastes data
    - e.g., unfilled blocks

### Recommendations: What (not) to Do

- Rosenberger’s and Lachin’s recommendation: two stage inference:
  - Tests are permutation tests
  - Estimates based on population model
- My recommendation:
  - Use population model through Phase-II.
  - Before Phase-III, ask FDA.
- If permutation test is possible, be wary of the size of the reference set: don’t use a design whose minimum p-value is > 0.05!

### Conclusions

- Promise? Bias control and inference.
- Problems? Bias control and inference.
- In practice, nearly all randomization designs are adaptive in the simplest sense: one treatment assignment is at least partially dependent on another.
  - Permutated block designs, while common, suffer from many of the same problems as “more adaptive” designs.
- Adopt a method of inference faithfully regardless of randomization details. Have faith? Either you believe the population model or you don’t.
- Population Model Analysis is independent of randomization details; Permutation Test Analysis depends on them absolutely.
  - Permutation tests can be complicated, especially with chaos.
  - Controlling selection bias is paramount.
Conclusions

- The Powerpoint presentation will be available on the internet:

- Email questions, comments, etc. to:
  [RHelms@RhoWorld.com](mailto:RHelms@RhoWorld.com)

- UNIX types: you don’t have to use capital letters…

- Thank you!