Biost 524: Design of Medical Studies

Lecture 5:
Methods of Randomization

Scott S. Emerson, M.D., Ph.D.
Professor of Biostatistics
University of Washington

April 21, 2010

1

Lecture Outline

- Randomization
- Analytic Models
- Completely Randomized Designs
- Blocked Randomization
- Stratified Randomization
- Covariate Adaptive Randomization

2

Nonadaptive Randomization

Randomization

Where am I going?

- The goal of a RCT is to assess cause and effect
- Randomization is the tool that allows this, but only for the scientific and statistical hypotheses that are based on randomization

3

Treatment of Variables

- Measure and compare distribution across groups
  - Response variable in regression
- Vary systematically (intervention)
- Control at a single level (fixed effects)
- Control at multiple levels (fixed or random effects)
  - Stratified (blocked) randomization
- Measure and adjust (fixed or random effects)
- Treat as “error”

4

Predictor of Interest

- The predictor of interest is varied systematically
  - *r* subjects on experimental treatment : 1 control

5

Cause and Effect

- Necessary conditions for establishing cause and effect of a treatment
  - The treatment should precede the effect
    - Beware protopathic signs
      - Marijuana and risk of MI within 3 hours
  - When comparing groups differing in their treatment, the groups should be comparable in every other way (at baseline)

6

Major Scientific Tool

- Randomization is the major way in which cause and effect is established
  - Ensures comparability of populations
    - Each treatment group drawn from same population
    - Differences in other prognostic factors will only differ by random sampling
      - Provides balance on the total effect of all other prognostic factors
      - May not provide balance on each individual factor
- NB: Sequential allocation of patients is not randomization
  - Possible time trends in recruitment, treatments, etc.

7

Nonadaptive Randomization

Analytic Models

Where am I going?

- Randomization serves as the basis for ascribing cause and effect
- However, to realize this we must consider the statistical foundations for inference, which include
  - Population model
  - Randomization model

8

Analytic Randomization Models

- Population model
  - Ensures treatment arms drawn from same population initially
  - Test weak null hypothesis of no treatment effect on summary measure of interest
    - E.g., test of equal mean outcome
    - Can allow for treatment differences between arms on other aspects of outcome distribution

9

Analytic Randomization Models

- Randomization model
  - Conditions on the sample obtained
    - E.g., permutation tests
    - Pretends that all outcomes were pre-ordained absent a treatment effect
  - Tests strong null hypothesis of no treatment effect whatsoever
    - Under the null hypothesis, any difference in outcome must have been randomization imbalance

10

Comments: Strong vs Weak Null

- Logical implications
  - Strong Null ➔ Weak Null
  - Rejection of Weak Null ➔ Rejection of Strong Null

- Advantages / Disadvantages of Strong Null
  - Can always test strong null via permutation tests
  - Assumption of strong null not in keeping with scientific method
    - Assumptions are more detailed than primary question
      - Primary question usually about first moment
      - Semiparametric assumptions are about all moments
    - Consider bone marrow transplantation

11

Comments: Choice of Analytic Models

- First choice: Population model
  - Randomization model does not typically allow testing of nonzero null hypotheses (e.g, noninferiority)
  - Randomization model does not allow distribution-free estimation of confidence intervals
    - For CI, we must know distribution under alternatives

- But the randomization model is an important fall back position
  - I generally feel uncomfortable in settings where a population model rejected a weak null but a randomization model could never reject the strong null
  - (cf: Deterministic minimization methods)

12

Points Meriting Repeated Emphasis

- Randomization is our friend…
  - If we randomize, we do not (on average) need to worry about differences between the treatment groups with respect to factors present at time of randomization
    - Any difference in outcomes can be attributed to treatment
      - Again, recognize that treatment can lead to differential use of other ancillary treatments, however
- But like all friends, we must treat it with respect.
  - We must analyze our data in groups defined at the time of randomization
    - Discarding or missing data on randomized subjects may lead to bias
      - It certainly leads to diminished scientific credibility

13

Impact on Data Analysis

- In presence of randomized treatment assignment
  - Intent to treat analysis (ITT)
  - Based on randomization
    - "Modified ITT" acceptable for efficacy?
      - Efficacy within strata identified pre-randomization
      - Safety in all subjects
  - Science: Population model (not randomization model)
    - My view: "Permutation Tests Considered Harmful"

14

Points for Further Elucidation

- Confounding not an issue (on average)
  - P value measures probability of observed effects occurring due only to randomization imbalance
- Gain precision if
  - Control important predictors, or
  - Adjust for stratification variables
- Subgroup analyses
  - If effect modification is concern

15

Nonadaptive Randomization

Complete Randomization

Where am I going?

- The simplest form of randomization is independent randomization of each individual
- Within the context of a completely randomized design, we can explore its performance with respect to
  - Bias,
  - Face validity, and
  - Precision.

16

Randomization Strategies

- Complete randomization (CRD)
- Blocked randomization
  - Ensure balance after every *k* patients
  - Ensure closer adherence to randomization ratio
  - Undisclosed block sizes to prevent bias
- Stratified randomization
  - Separately within strata defined by strong risk factors
    - Lessens chance of randomization imbalance
  - Need to consider how many variables can be used
- Dynamic randomization
  - Adaptive randomization to achieve best balance on marginal distribution of covariates

17

Complete Randomization (CRD)

- With each accrued subject a (possibly biased) coin is tossed to determine which arm
  - Probability of treatment arm = *r / (r + 1)*
  - Independence of successive randomizations
- Issues
  - Bias
  - Face validity
  - Precision

18

CRD: Unbiased

- On average (across repeated experiments)
  - No correlation between treatment variable and other covariates
  - Individual type I errors come from samples in which other covariates are imbalanced

$$\beta_1 = \gamma_1 + \rho_{XW} \frac{\sigma_W}{\sigma_X} \gamma_2$$

19

Face Validity: Table 1

	Methotrexate Arm		Placebo Arm	
	n	Mean (SD; Min – Max)	n	Mean (SD; Min – Max)
Age (yrs)	132	50.4 (8.5; 32 - 69)	133	52.2 (8.5; 26 - 67)
Female	132	92.4%	133	92.5%
Pruritus score	116	7.7 (3.8; 4 - 16)	124	6.9 (3.8; 4 - 20)
Splenomegaly	131	8.4%	133	10.5%
Telangiectasia	132	4.6%	133	11.3%
Edema	132	6.1%	133	3.0%
Alkaline phosphatase	132	242.6 (145.9; 53 - 933)	133	245.0 (187.6; 66 - 1130)
ALT	131	54.5 (41.7; 12 - 202)	132	50.6 (41.4; 12 - 311)
Total bilirubin	132	0.7 (0.4; 0.1 - 2.7)	133	0.7 (0.4; 0.1 - 2.4)
Albumin	132	4.0 (0.3; 3.1 - 6.0)	133	4.0 (0.3; 3.0 - 4.8)
Prothrombin time INR	124	1.0 (0.1; 0.7 - 1.3)	132	1.0 (0.1; 0.7 - 1.3)
Mayo score	128	3.8 (0.8; 1.6 - 6.3)	133	3.9 (0.8; 1.6 - 6.1)
Avg stage	128	2.2 (0.9; 1.0 - 4.0)	128	2.3 (0.9; 1.0 - 4.0)
Avg fibrosis	128	1.2 (0.8; 0.0 - 3.0)	128	1.3 (0.9; 0.0 - 3.0)

20

CRD: Face Validity

- Table 1: Potential for imbalance in covariates
  - Depends on number of covariates and correlations among them
  - Probability of at least one "significant" imbalance

Number Displayed	Worst Case	Correlation 0.00	0.30	0.50	0.75	0.90
1	.050	.050	.050	.050	.050	.050
2	.100	.098	.095	.090	.081	.070
3	.150	.143	.137	.126	.104	.084
5	.250	.226	.208	.184	.138	.101
10	.500	.401	.353	.284	.193	.127
20	1.000	.642	.540	.420	.258	.154
50	1.000	.923	.806	.624	.353	.193

21

CRD: Face Validity

- Of course, statistical significance is not the issue
  - "Conditional confounding"
    - How does unadjusted estimate compare to adjusted estimate?
    - Product of sample correlation between *X* and *W* and adjusted association between *Y* and *W*

$$\beta_1 = \gamma_1 + r_{XW}\frac{\sigma_W}{\sigma_X}\gamma_2$$

22

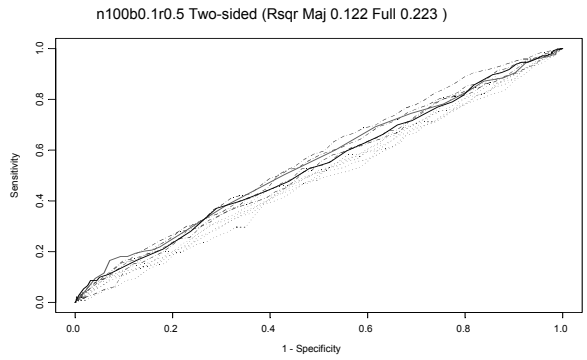
Demonstration of the Problem

- Consider a CRD in presence of
  - 4 highly correlated predictors with larger importance
  - 6 independent predictors with smaller importance
  - No treatment effect
- Questions about unadjusted analysis
  - What is type I error? ➔ 0.025
  - What does imbalance in predictors tell us about type I error?
    - Sensitivity, specificity of imbalance in predictors under null hypothesis
    - Dependence on R^2 of measured covariates

23

Low Association: 2-sided

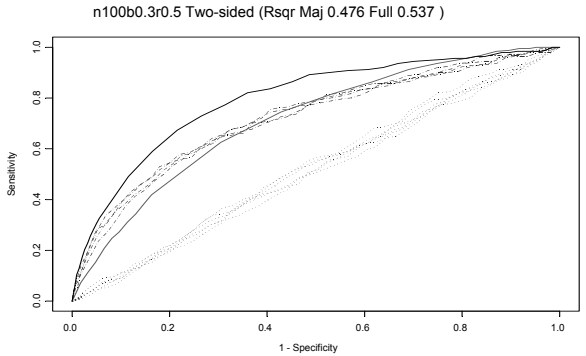
- ROC curve for covariate imbalance "explaining" statistical significance under the null



24

High Association: 2-sided

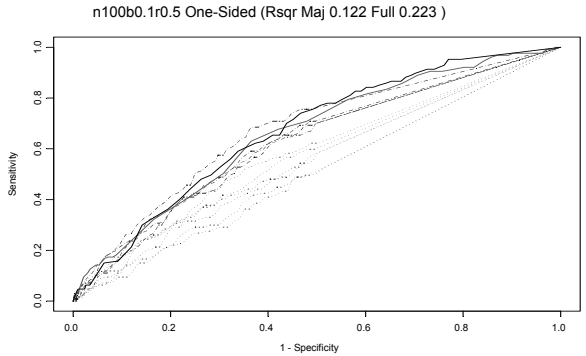
- ROC curve for covariate imbalance “explaining” statistical significance under the null



25

Low Association: 1-sided

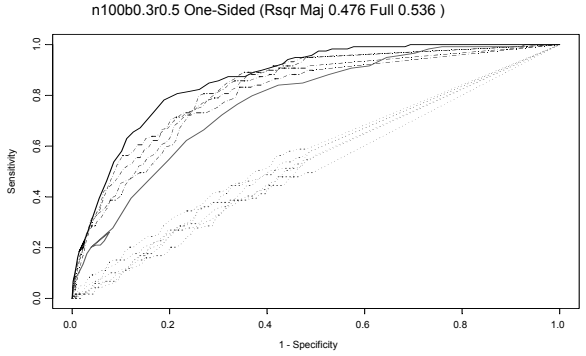
- ROC curve for covariate imbalance “explaining” statistical significance under the null



26

High Association: 1-sided

- ROC curve for covariate imbalance “explaining” statistical significance under the null



27

Face Validity

- Spurious results due to covariate imbalance
  - Unconditionally: Unbiased so no problem
  - Conditional on obtained randomization:
    - IF covariates are strongly predictive of outcome, then covariate imbalance is predictive of type I error
    - But need to consider that combined effect of other measured and unmeasured covariates may provide balance
- Ultimately, however, we need to have credible results
  - We do not always get to choose what others believe

28

Precision

- Impact of completely randomized design on precision of inference
  - Impact of imbalance in sample sizes
    - The number accrued to each arm is random
  - Impact of imbalance in covariates
    - "One statistician's mean is another statistician's variance"

29

Randomization Ratio

- Most efficient
  - When test statistics involve a sum, choose ratio equal to ratio of standard deviations
- Most ethical for patients on study
  - Assign more patients to best treatment
    - Many sponsors / patients presume new treatment
    - (Adaptive randomization: Play the winner)
- Most ethical for general patient population
  - Whatever is most efficient (generally not adaptive)
- Other goals
  - Attaining sufficient patients exposed to new treatment
  - Maintaining DSMB blind

30

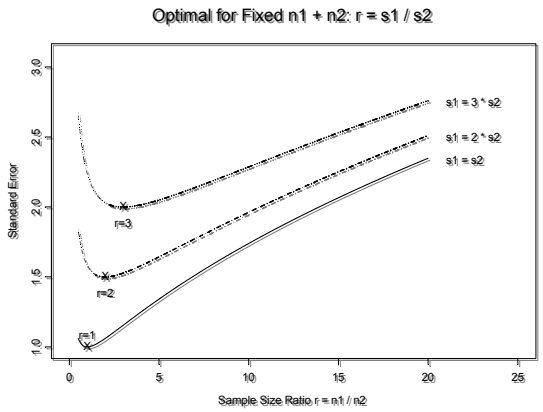
Comment: Optimal *r* (Fixed *n*)

- Suppose we are constrained by maximal sample size $n = n_1 + n_2$
  - Smallest standard error when

$$r = \frac{n_1}{n_2} = \frac{\sigma_1}{\sigma_2}$$

31

Comment: Optimal *r* (Fixed *n*)



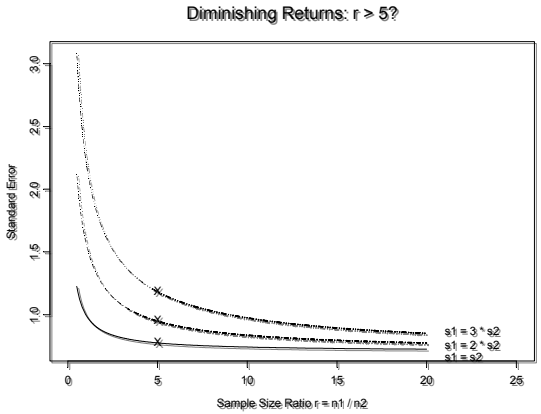
32

Comment: Diminishing Returns

- When we are unconstrained by maximal sample size we still hit a point of diminishing returns
  - Often quoted: *r = 5*

33

Comment: Diminishing Returns



34

CRD: Efficiency Loss from Wrong Ratio

- CRD may not attain optimal ratio
  - Following table explores practical inefficiency
  - (True inefficiency is infinite due to possibility of no subjects randomized to one group)

N	r= 1	r= 2	r= 3	r= 5	r=10
20	1.0599	1.0652	1.0694	***	***
50	1.0213	1.0219	1.0229	1.0258	1.0282
100	1.0103	1.0104	1.0106	1.0111	1.0130
200	1.0051	1.0051	1.0051	1.0053	1.0056
500	1.0020	1.0020	1.0020	1.0020	1.0021
1000	1.0010	1.0010	1.0010	1.0010	1.0010

35

CRD: Probability 0 or 1 on an Arm

- CRD may not randomize at least two subjects per arm

N	r= 1	r= 2	r= 3	r= 5	r=10
20	0.0000	0.0033	0.0243	0.1304	0.4459
50	0.0000	0.0000	0.0000	0.0012	0.0511
100	0.0000	0.0000	0.0000	0.0000	0.0008
200	0.0000	0.0000	0.0000	0.0000	0.0000
500	0.0000	0.0000	0.0000	0.0000	0.0000
1000	0.0000	0.0000	0.0000	0.0000	0.0000

36

CRD: Efficiency Loss from Imbalance

- Covariates may be imbalanced across arms
  - Variability across replicated experiments increased if important predictor not controlled
    - Increased within group variance

Unadjusted Model $\left[se\left(\hat{\beta}_1\right)\right]^2 = \frac{Var\left(Y \mid X\right)}{nVar\left(X\right)}$

$$Var\left(Y \mid X\right) = \gamma_2^2 Var\left(W \mid X\right) + Var\left(Y \mid X, W\right)$$

37

CRD: Improved Performance

- If we adjust for important covariates, we will often gain precision
  - Face validity in Table 1 if readers recognize that adjustment accounts for any observed imbalance
- Caveats:
  - If covariate imbalance by arm, model misspecification can be an issue re conditional bias
  - If covariate imbalance by arm, lack of effect can be an issue re variance inflation
  - If adjustment not TOTALLY prespecified, "intent to cheat" analysis can be an issue
    - Not too much loss of precision from imperfect model

38

CRD: Continuous vs Dichotomized

	CRD – Continuous Adjust				CRD – Dichotomized Adjust			
Tx Eff	SE Slope		Power		SE Slope		Power	
	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj
0.0	.281	.211	.026	.024	.284	.231	.023	.026
0.1	.278	.209	.053	.062	.284	.229	.045	.062
0.3	.279	.209	.178	.285	.287	.231	.184	.243
0.5	.281	.209	.423	.655	.279	.225	.409	.581
0.7	.279	.209	.696	.909	.281	.229	.699	.858

39

Nonadaptive Randomization

Blocked Randomization

Where am I going?

- Blocking is sometimes used to ensure
  - Proper ratio of sample sizes across groups, and
  - Balance across arms over time

40

Randomization Strategies

- Complete randomization
- Blocked randomization
  - Ensure balance after every *k* patients
  - Ensure closer adherence to randomization ratio
  - Undisclosed block sizes to prevent bias
- Stratified randomization
  - Separately within strata defined by strong risk factors
    - Lessens chance of randomization imbalance
  - Need to consider how many variables can be used
- Dynamic randomization
  - Adaptive randomization to achieve best balance on marginal distribution of covariates

41

Issues with CRD

- Imbalance across arms in sample sizes
  - Not much of an issue with large sample sizes
  - Could be problematic with sequential sampling
    - Interim analyses of data early in the study
- Imbalance across arms in time trends
  - Outcome may be associated with time of accrual

42

Mechanisms Leading to Time Trends

- Patients accrued early may differ from those accrued later, because
  - Backlog of eligible patients
  - Startup of new clinical sites
  - Pressure to increase accrual
  - Secular trends in beliefs about intervention
    - (Made much worse if any interim results leak out)
  - Secular trends in diagnostic tools used for eligibility
  - Secular trends in ancillary treatments

43

Blocked Randomization

- Within every sequence of *k* patients, the ratio of treatment to control is exactly *r : 1*
  - Within each “block” ordering of treatments is random
- Important caveats:
  - Investigators must not know block size
    - Otherwise, decisions to enroll patients might be affected by knowledge of next assignment
  - Hence, often use “concealed blocks of varying sizes”

44

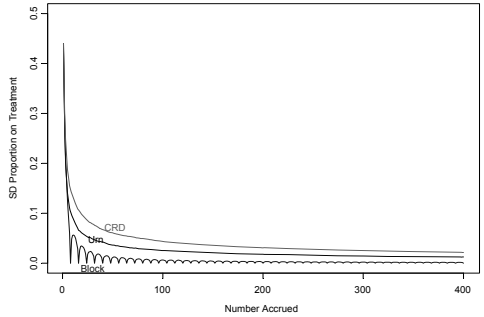
Alternative Strategy: Urn Model

- Begin with *k* white balls and *rk* black balls in an urn
- Upon accrual of a patient draw a ball from urn
  - White ➔ control; black ➔ treatment
  - After every white ball withdrawn, return 1 white ball and *rm* black balls
  - After every *r*-th black ball withdrawn, return *r* black balls and *m* white balls
- Such a strategy tends to behave like small blocks early and complete randomization later, depending on *k* and *m*

45

Comparison of Blocking Strategies

- SD proportion on treatment for 3:1 randomization
  - Urn (k=1, m=1) vs Blocking (size = 8) vs CRD



46

Statistical Inference

- Impact on statistical inference relative to CRD
  - Bias properties unchanged
  - Face validity largely unchanged
    - We rarely report accrual patterns over time
  - Precision slightly improved due to achieving closer to desired randomization ratio
  - Precision could be improved if adjust for blocks as a random effect in analysis
    - This is rarely done, except in re-randomization test
      - Large number of small blocks, often with small variance of the random effects

47

Nonadaptive Randomization

Stratified Randomization

Where am I going?

- Stratified randomization is sometimes used to ensure proper ratio of sample sizes across subgroups defined by important covariates, thereby
  - Decreasing conditional bias,
  - Improving face validity, and
  - Possibly improving precision
- Major improvements in precision are gained only with adjustment for important stratification variables

48

Randomization Strategies

- Complete randomization
- Blocked randomization
  - Ensure balance after every *k* patients
  - Ensure closer adherence to randomization ratio
  - Undisclosed block sizes to prevent bias
- Stratified randomization
  - Separately within strata defined by strong risk factors
    - Lessens chance of randomization imbalance
  - Need to consider how many variables can be used
- Dynamic randomization
  - Adaptive randomization to achieve best balance on marginal distribution of covariates

49

Issues with CRD

- Imbalance across arms in covariate distribution
  - Loss of face validity
  - Conditional bias
  - Not much of an issue with large sample sizes
  - Could be problematic with sequential sampling
    - Interim analyses of data early in the study
  - Could be problematic with subgroup analyses
    - Possibility of very inefficient randomization ratio in small subgroups

50

Stratified Randomization

- Strata are defined based on values of important covariates
  - E.g., sex, age, disease severity, clinical site
- Within each stratum defined by a unique combination of stratification variables, CRD or blocked randomization
- Important caveats:
  - Number of strata is exponential in number of stratification variables
    - E.g., 4 two level stratification variables ➔ 16 strata

51

Statistical Inference

- Impact on statistical inference relative to CRD
  - Bias properties unchanged
  - Face validity improved for most important variables
  - Precision improved due to achieving closer to desired randomization ratio
  - Precision could be further improved if adjust for stratification variables in analysis
    - This should be done
      - Without adjustment for strata, may even lose power for some alternatives
    - Requires pre-specification of analysis model to avoid “intent to cheat” analysis

52

CRD vs Orthogonal Randomization

	CRD – Continuous Adjust				Orthogonal Randomization			
Tx Eff	SE Slope		Power		SE Slope		Power	
	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj
0.0	.281	.211	.026	.024	.206	.206	.005	.026
0.1	.278	.209	.053	.062	.208	.208	.013	.069
0.3	.279	.209	.178	.285	.205	.205	.115	.313
0.5	.281	.209	.423	.655	.205	.205	.403	.684
0.7	.279	.209	.696	.909	.205	.205	.759	.924

Advantages

- Additional advantages of stratification
  - Balance within clinical center
    - Especially if quality control issues
  - Balance for interim analyses
  - Balance for subgroup analyses

Adaptive Randomization

Covariate Adaptive Randomization

Where am I going?

- Stratified randomizations has drawbacks in the presence of sparse data
- Some authors have described dynamic randomization processes that will allow balancing on more covariates

Issues with Stratified Analyses

- The need to stratify on all combinations of variables
  - Good news:
    - Balances on interactions as well as main effects
  - Bad news:
    - Effect of interactions might be quite small
    - Really only need to adjust on “counterfactual” outcome based on linear combination of all covariates

Dynamic Randomization

- Subjects are assigned to the treatment arm that will achieve best balance
  - "Minimization": minimize the difference between the distribution of covariate effects between arms
    - Define a "distance" between arms for covariate vectors
    - Probability of assignment depends upon arm that would provide smallest difference

57

Distance Between Arms

- Two arms are "distant" based on one of:
  - Randomization ratio very different from *r : 1* in some stratum
  - Summary measure of distribution of $(W_{i1},\ldots,W_{ip})$ differs
    - Mean, median, variance, …
  - Distribution of $(W_{i1},\ldots,W_{ip})$ differs
  - Contribution of covariates to the outcome differs

58

Conditional Confounding

Unadjusted : $g\left[\theta \,|\, X_i\right] = \beta_0 + \beta_1 \times X_i$

Adjusted : $g\left[\theta \,|\, X_i, \vec{W}_i\right] = \gamma_0 + \gamma_1 \times X_i + \vec{W}_i^T \vec{\delta}$

Unadjusted : $g\left[\theta \,|\, \mathbf{X}\right] = \mathbf{X}\vec{\beta}$

Adjusted : $g\left[\theta \,|\, \mathbf{X}, \mathbf{W}\right] = \mathbf{X}\vec{\gamma} + \mathbf{W}\vec{\delta}$

$$\vec{\beta} = \vec{\gamma} + \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}\vec{\delta}$$

59

Conditional Confounding

$g\left[\theta \,|\, \mathbf{X}\right] = \mathbf{X}\vec{\beta}$ $\qquad$ $g\left[\theta \,|\, \mathbf{X}, \mathbf{W}\right] = \mathbf{X}\vec{\gamma} + \mathbf{W}\vec{\delta}$

$$\vec{\beta} = \vec{\gamma} + \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}\vec{\delta}$$

$$\beta_1 = \gamma_1 + \sum_{j=1}^{p}\left(\overline{W}_{1j\bullet} - \overline{W}_{0j\bullet}\right)\delta_j$$

$$\overline{W}_{kj\bullet} = \frac{1}{n_k}\sum_{i=1}^{n} W_{ij}1_{[X_i=k]}$$

60

Distance Metrics

Based on contribution to confounding:

$$d\left(\vec{X}, \mathbf{W}\right) = \left| \sum_{j=1}^{p} \left(\overline{W}_{1j\bullet} - \overline{W}_{0j\bullet}\right) \delta_j \right|$$

Weighted distance between standardized means:

$$d\left(\vec{X}, \mathbf{W}\right) = \sum_{j=1}^{p} c_j \left| \frac{\overline{W}_{1j\bullet} - \overline{W}_{0j\bullet}}{SD\left(W_j\right)} \right|^{\lambda}$$

Weighted imbalance in n across strata $\Omega_1, \ldots, \Omega_s$:

$$d\left(\vec{X}, \mathbf{W}\right) = \sum_{s=1}^{S} c_s \left| \sum_{i=1}^{n} 1_{[X_i=1]} 1_{[\vec{W}_i \in \Omega_s]} - \sum_{i=1}^{n} 1_{[X_i=0]} 1_{[\vec{W}_i \in \Omega_s]} \right|^{\lambda}$$

61

Implication

- Spurious associations will be minimized if means of important predictors are balanced across treatment arms
  - The greater the value of δ_j the more important it is for the means of the j-th covariate to be equal
    - (Presumes linear model reasonable approximation)
  - We could use estimates of the of δ_j's to define the distance between the arms (or just balance means)
- Balancing group sizes across covariates will tend to have means balanced by randomization
  - Group sizes within strata may matter for subgroup analyses

62

Probability of Assignment

- Subjects are assigned to the treatment arm that will achieve best balance
  - When i-th patient accrued, compute a randomization probability

$$\Delta_i = d\left(\vec{X}, \mathbf{W} \mid X_i = 1\right) - d\left(\vec{X}, \mathbf{W} \mid X_i = 0\right)$$

$$\pi_i = \Pr\left(X_i = 1\right) = f\left(\Delta_i\right)$$

  - $0 \leq \pi_i \leq 1$
  - Larger values of Δ_i ➔ smaller values of π_i
  - Probably best to avoid $\pi_i = 0$ and $\pi_i = 1$

63

Inference: Population Model

- Impact on statistical inference relative to CRD
  - Bias properties unchanged
  - Face validity improved for most important variables
  - Precision improved due to achieving closer to desired randomization ratio
  - Precision could be further improved if adjust for stratification variables in analysis for population model
    - This should be done
    - Requires pre-specification of analysis model to avoid "intent to cheat" analysis

64

Inference: Randomization Model

- Proschan, Brittain, Kammerman, 2010: Precision could be greatly hampered if you analyze under randomization hypothesis
- Alternative randomization schemes may be quite restrictive, especially under unequal randomization
  - Suppose sequential allocation
    - Randomization P value is identically 1 (or 0.5?)
  - If dynamic randomization has $\pi_i = 0$ or $\pi_i = 1$ too often, range of randomization P values is greatly restricted
- Also: Statistical analysis can be quite involved

65

Advantages / Disadvantages

- Advantages
  - Typically improved face validity
  - Can handle an arbitrary number of covariates
    - Depending on distance metric
- Disadvantages
  - Logistically more involved
  - Decreased credibility if too deterministic
    - Approaches sequential allocation
  - Some analytic strategies more complex
  - Does not necessarily facilitate subgroup analyses
    - Unless distance metric chosen carefully

66

Adaptive Randomization

Response Adaptive Randomization

Where am I going?

- Some authors have described dynamic randomization processes that attempt to minimize exposure of patients to harmful treatments

67

Ethics

- Clinical trials are experiments in human volunteers
- Individual ethics
  - Patients on trial: Avoid continued administration of inferior treatment
  - Patients not yet on trial: Avoid starting inferior treatment
- Group ethics
  - Facilitate rapid adoption of new beneficial treatments
  - Avoid prolonging study of ineffective treatments

68

Solutions

- Most commonly used
  - Sequential sampling
    - Interim analyses of data
    - Terminate trials when credible decisions can be made
- Also proposed
  - Response adaptive randomization
    - Change randomization probabilities as evidence accumulates that one treatment might be best
    - "Play the winner"

69

Play the Winner: Urn Model

- Begin with *k* white balls and *k* black balls in an urn
- Upon accrual of a patient draw a ball from urn
  - White ➔ control; black ➔ treatment
- Observe outcome
  - If outcome is good, return *m+1* balls of same color as withdrawn
  - If outcome is bad, return 1 ball of same color as withdrawn and *m* balls of opposite color

70

Bayesian Methods

- An explicit Bayesian approach could to dynamic randomization could base the randomization ration on the current posterior probability that one treatment is superior
  - Ultimately, that posterior probability is based on the number of good outcomes on each treatment
- Advantage of using Bayesian posterior probability
  - Can easily handle continuous outcomes
  - Can easily handle continuous randomization probabilities

71

Analytic Issues

- Treatment of successive patients is not independent of previous patients treatment and results
  - Possible bias in accrual of future patients
- Conditionally biased estimates of treatment effect in arm with lower sample sizes
  - Bad early results tend to preclude regression to mean
- Randomization hypothesis can lead to quite unconvincing results

72

Example: ECMO Study

- Randomized clinical trial of extracorporeal membrane oxygenation in newborns
  - Randomized PTW design with k=1
- Data:
  - First patient on ECMO survived
  - Next patient on control died
  - Next 9 patients on ECMO survived
- Inference (Begg, 1990)
  - P value of 0.001, 0.051, 0.083, 0.28, 0.62?

73

Comments

- This experience has tempered enthusiasm for randomized PTW
  - Interestingly, follow-up studies had 67% survival on conventional therapy
- I believe there can be times that this will work, but
  - There needs to be a clear dilemma re individual ethics
  - There will tend to be decreased group ethics
  - It takes a lot of planning in order to obtain results that will be sufficiently credible
    - Assuming your conclusion will not cut it

74