Sequential Rematching Randomization and Adaptive Monitoring with the Second-Generation p-value to increase the efficiency and efficacy of Randomized Clinical Trials

Jonathan Chipman, MS

Department of Biostatistics
Vanderbilt University

May 10, 2019

RCT Costs and Benefits

2015-2016 Pivotal Trials accepted by FDA **?**

- 138 Trials to approve 59 novel therapeutic drugs
- Enrolled on median 488 patients (IQR: 230, 740)
- Cost per patient on median $41K (IQR: $32K - $82K)
- 53 Used surrogate marker as primary endpoint

NINDS Reviewed costs of trials between 1977 and 2002

- Overall quality of life benefits >> Overall costs

"Reducing the costs of trials is absolutely crucial for the public good" – Dr. Claiborne Johnston (UCSF) **?**

Outline

Three methods paper to increase RCT efficiency and efficacy.

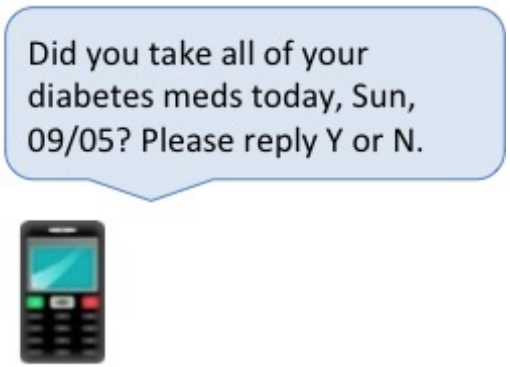
- Motivating Example: REACH Trial
- Paper 1: Sequential Rematched Randomization
- Paper 2: Adaptive Monitoring Using the Second Generation p-value
- Paper 3: sgpvAM R package and methods to controll adaptive monitoring errors

REACH Trial

Rapid **E**ducation/Encouragement **A**nd **C**ommunications for **H**ealth

Population:	Adults with Type 2 Diabetes (DM)
Purpose:	Increase glycemic control and adherence to medications
Main Intervention:	Text message-delivered diabetes support for 12 months
Outcome:	Glycemic control (A1c) compared to control
Multi-site enrollment:	512 patients from Vanderbilt and Non-Vanderbilt Clinics
Clinical Relevance:	Change clinical practice – Decrease in A1c of 0.5 or more

REACH



Key Baseline Covariates

Biological Factors

- Baseline A1c*
- Age at baseline
- Time since DM dx*
- DM type*
- Race / Ethnicity

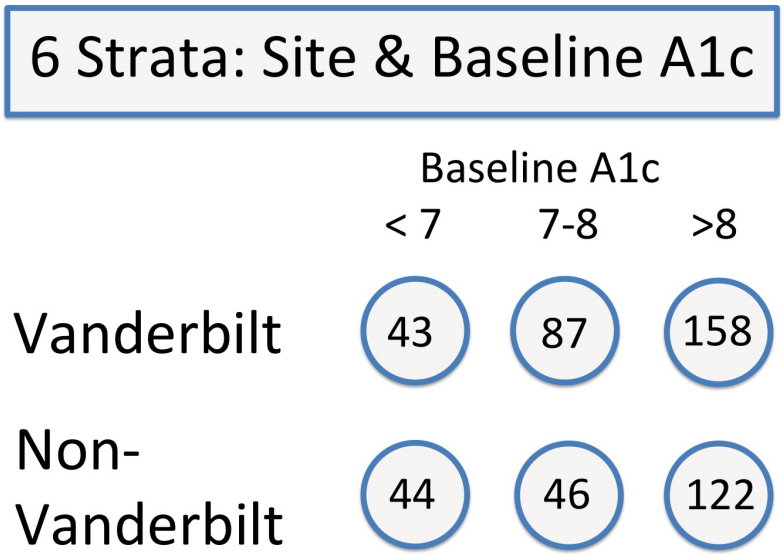
Socio-economic Factors

- Yrs of education
- Income level
- Insurance type

* Greater association with outcome

Stratified Block Randomization

- Randomize within categories of similar patients (ex: Site)
- Continuous covariates must be categorized
- REACH Example: Randomize within site and Baseline A1c (6 strata)



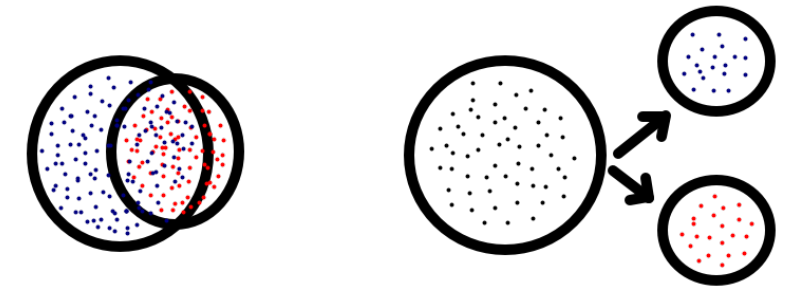
Stratified Block Randomization

- Limited in number of covariates creating strata
- 4000+ Strata - Stratifying on all covariates*
- 2 * 3 * 3 * 2 = 36 Strata - Stratifying on:
  - 2 Sites
  - 3 Baseline A1c levels
  - 3 DM Types
  - 2 Time since DM Dx levels

*Continuous baseline covariates categorized as:

Baseline A1c:	< 7	7-8	≥ 8
Age:	<60	≥ 60	
Yrs of Education:	<12	≥ 12	
Time since DM Dx:	<10	≥ 10	

Matched Randomization [?]



Observational Study: Bipartite (2 group) matching

Randomization Trial: Non-bipartite matching

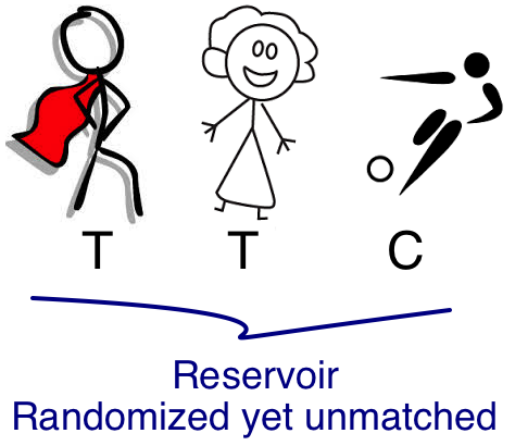
Refined stratification into pairs of patients using distance matrix

- Pairs selected for smallest total distance
- Allows for any number of continuous and categorical covariates
- All participants known before randomization

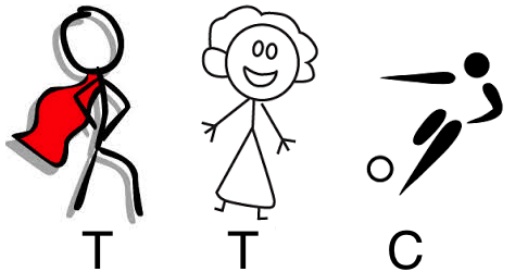
Sequential Matching [?]

Sequential Matching [?]

Sequential Matching [?]

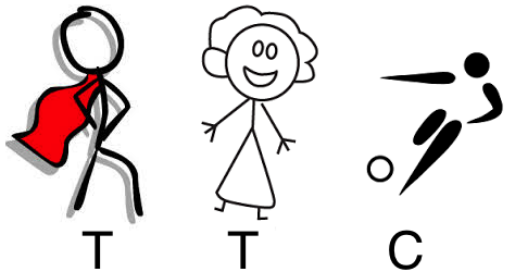


Sequential Matching [?]



If $\text{Distance}_2 < \text{Threshold}$

Sequential Matching [?]



If $\text{Distance}_2 > \text{Threshold}$

Matching OTF, some additional notes [?]

Pre-specification:

- Initial reservoir size
- Threshold to denote degree of similarity

Fixed Threshold: Match is better than [20%] of random matches

- Mahalanobis Distance of random pairs scales to $F_{(p,n-p)}$
- Assumes normally distributed baseline covariates

Reservoir:

- Not required to deplete
- May result in unequal treatment group sizes

Extension 1: Dynamic Threshold

Dynamic percentile of random match distances

1. Chance of matching now versus later
2. Random match distances (from data, ie not assumed normal)
3. Remove threshold at end to ensure all patients match

Patients enroll in b enrollment blocks
$||U_b||$: Number of unmatched patients at bth enrollment
$||R_b||$: Number remaining patients after bth enrollment

$$Q_b = \frac{||U_b|| - 1}{||U_b|| + ||R_b|| - 1}$$

$$Threshold_b = \begin{cases} F_b^{-1}(Q_b) & ||U_b|| < ||R_b|| \\ \text{best match(es)} & ||U_b|| \geq ||R_b|| \end{cases}$$

F_b is estimated by $\hat{F}_b$ by bootstrap sampling random matches from distance matrix.

Extension 2: Sequential Re-Matching

Sequential Matching

- Formed matches remain through study

Sequential Re-Matching

- Matches are allowed to break if a better match enters
- However, participants keep original treatment assignment
  - New participant: Match to anybody
  - Former enrolled participant: Match to anyone of opposite treatment or new participant

Case study using REACH Data to comparing randomization schemes

Balance of baseline covariates

Efficiency estimating treatment effect

Results will vary based on baseline covariates

- Number of patients
- Number of covariates
- Distribution of covariates
- Adjusted R^2 of covariates

Simulations to evaluate Sequential Rematching Randomization

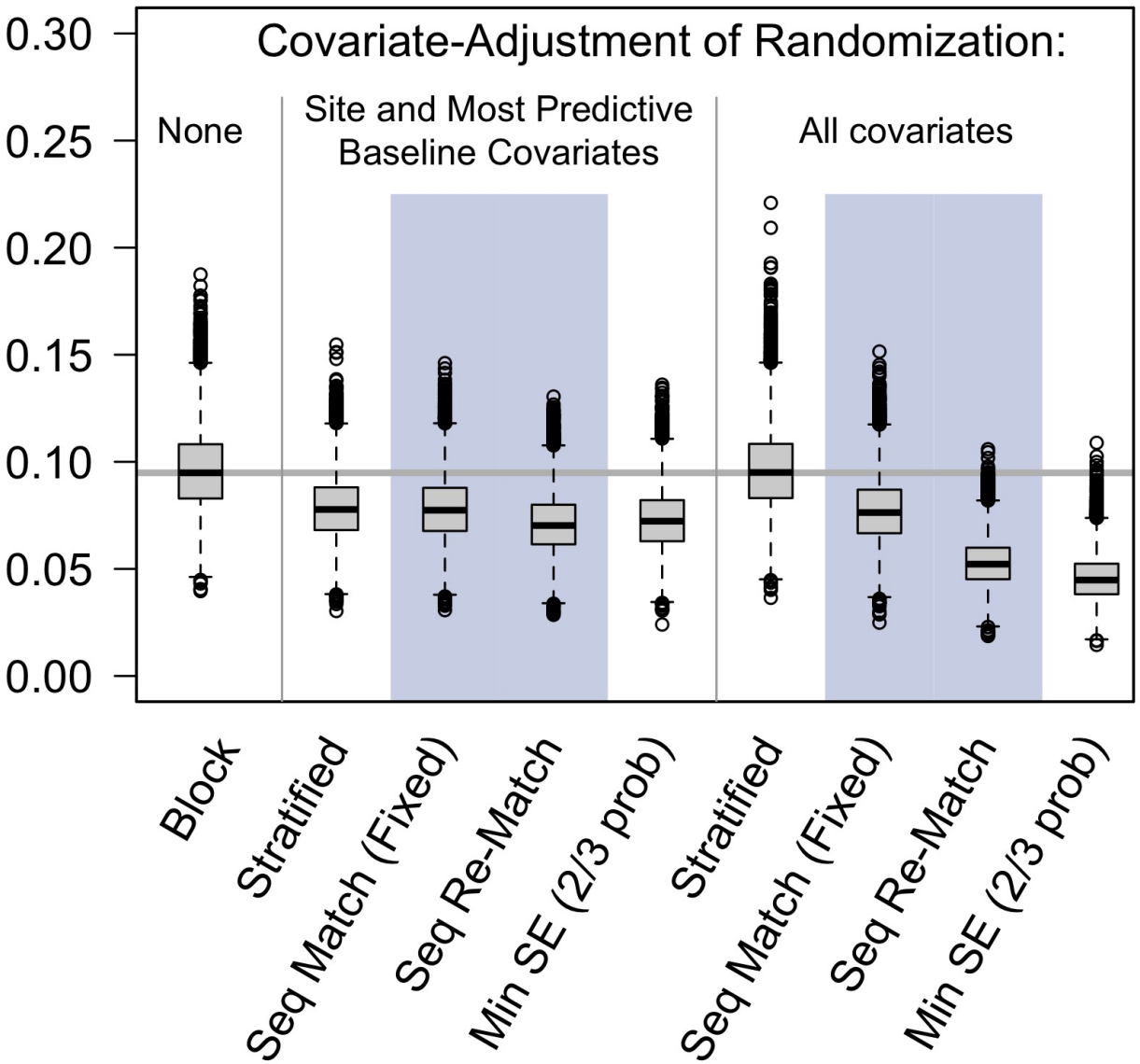
Generate 20K MCMC replicates under potential outcomes framework:

- Bootstrap sample 512 observations from REACH trial
- Outcomes
  - Y(0) = Predicted three month A1c + random residual
  - Y(1) = Y(0) - 0.5
- Generate randomization scheme under
  - Block Randomization
  - Stratified Block Randomization
  - Matched Randomization
- Observe
  - Balance: Average Standardized Mean Difference among all covariates
  - Efficiency: CI Width of 1) fully-adjusted linear model and 2) Permutation distribution

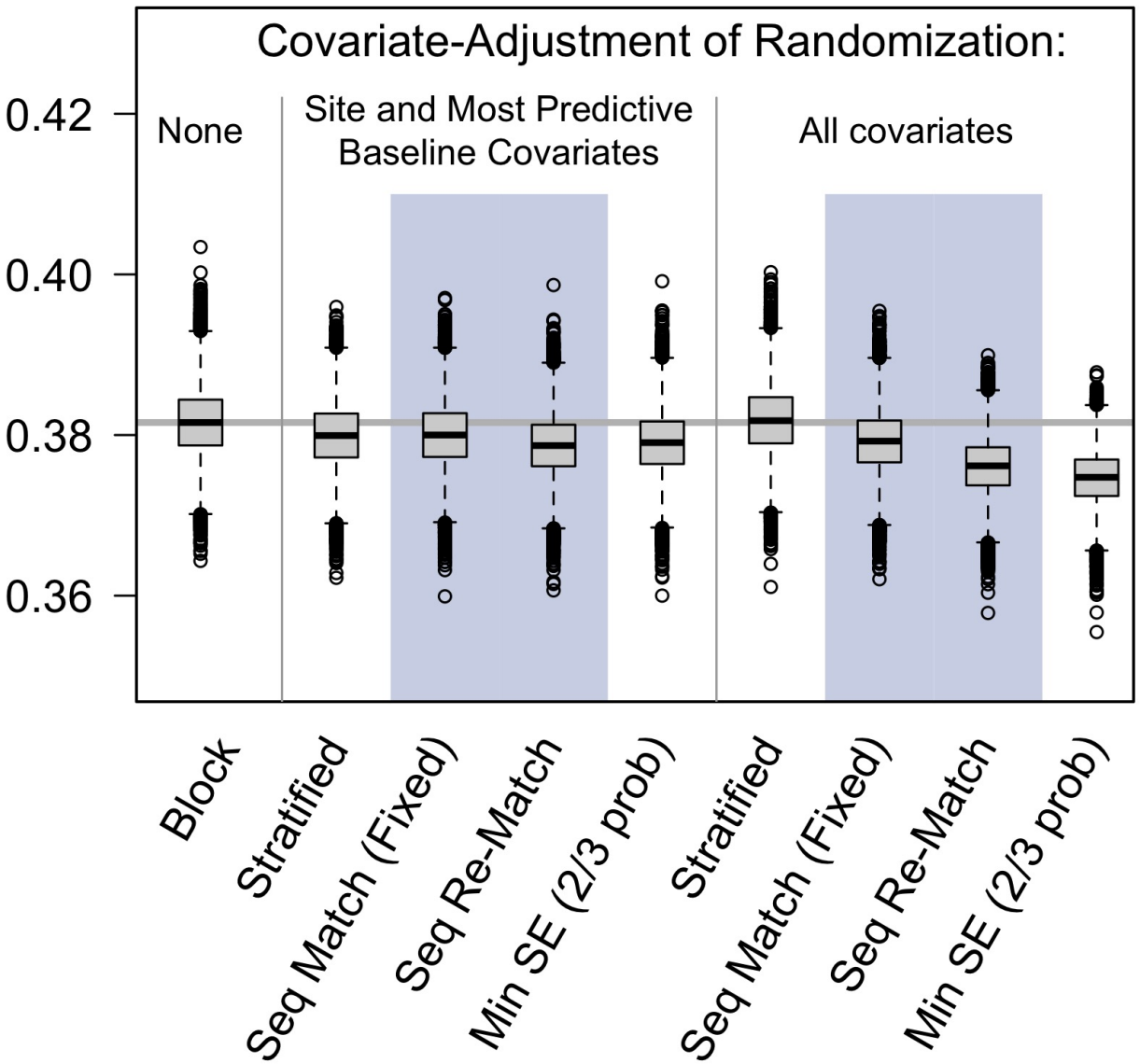
Where Stratified and Matched randomization adjusts for

- Site, Baseline A1c, DM Type, and Time since DM Dx (Adj)
- All baseline covariates (3m A1c Adj $R^2 = 0.46$)

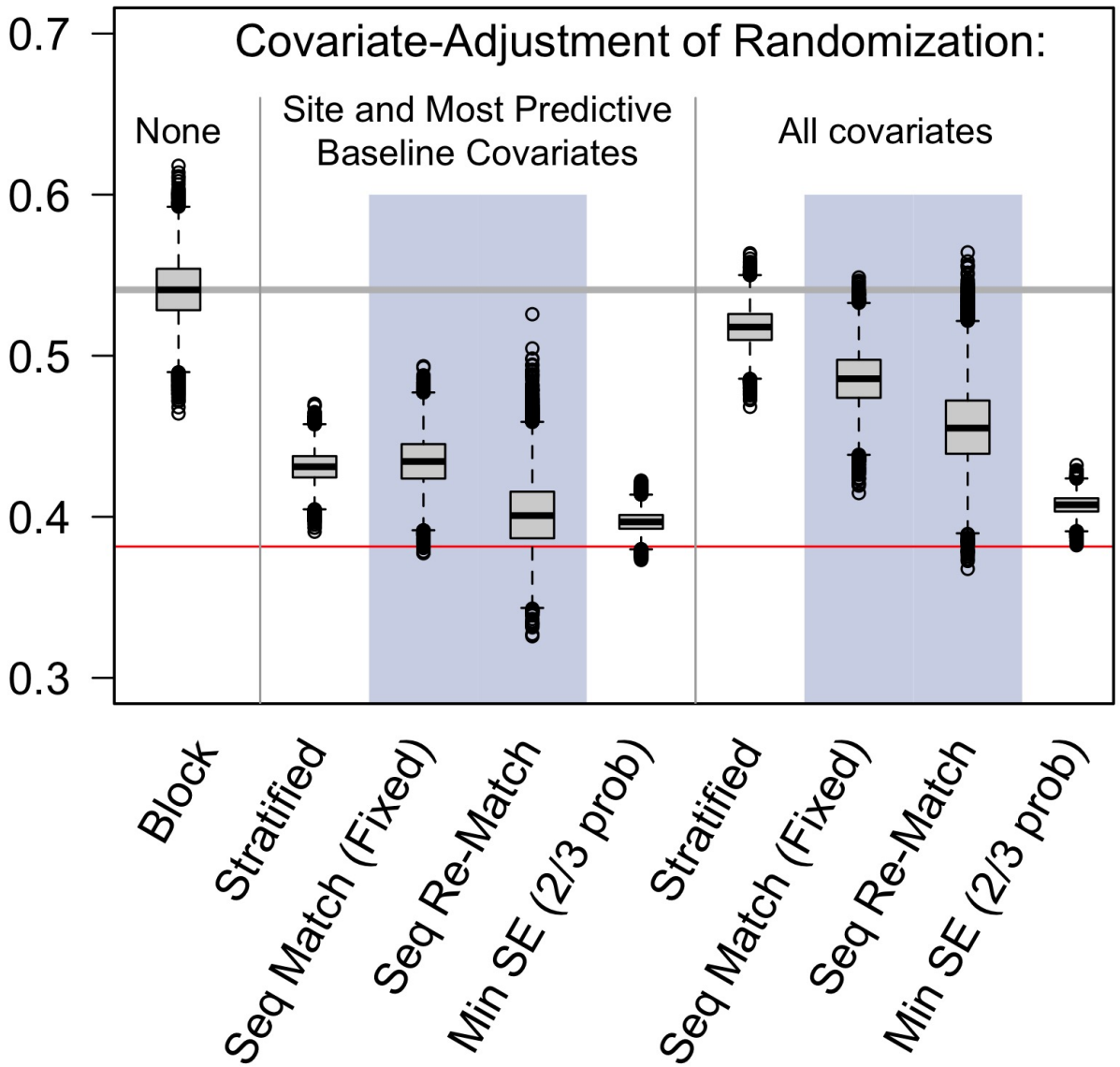
Average Standardized Mean Difference (all baseline covariates)



Covariate-Adjusted OLS Confidence Interval Width of estimated treatment effect



Permutation Confidence Interval Width of estimated treatment effect



Recommendations regarding balance and efficiency

Conclusions

- Greatest overall balance by adjusting for all baseline covariates
- Greatest linear gains in efficiency
  - Model efficiency: adjusting randomization to all baseline covariates
  - Permutation efficiency: adjusting randomization to priority covariates
- Permutation efficiency achieves nearly same efficiency as fully-adjusted model with Block Randomization

Practical Recommendations

- Recommend adjusting to all baseline covariates
- Adaptively monitor until reaching a clear clinical conclusion

References