Measuring The Impact of Patient Preference in Clinical Trials

Denise Esserman, PhD

Professor of Biostatistics, Yale School of Public Health
Co-director, Yale Center for Analytical Sciences
Director, Biostatistics Core, Yale Program on Aging
denise.esserman@yale.edu

May 22, 2024

Funding Disclosure

This work was supported through the Patient Centered Outcomes Research Institute (PCORI) and Yale's CTSA Award.

Disclaimer: All statements in this presentation, including its findings and conclusions, are solely those of the author and do not necessarily represent the views of PCORI or its Board of Governors, or the Methodology Committee

No other relevant disclosures.

Overview

1. Introduction
2. Two-stage Clinical Trial Designs
3. Partially Randomized Preference Design
4. Additional Work and Considerations for the Future

Patient Preference

According to PCORI: *Deciding between alternative options in health care requires an understanding of how to balance the benefits and risk of each treatment option and an understanding of how each option may apply differently to different patients, given their unique personal characteristics.*[1]

As stated by McPherson and Britton: *Until the therapeutic effects of preference itself are more clearly understood, understanding the true therapeutic effects will be compromised, at least in principle.*[2]

Patient Preference's Impact on Clinical Trials

- Patient preference plays a role in clinical practice, and is at the heart of patient-centered outcomes research; thus, ignoring the impact on outcomes could results in missing important determinants of study outcomes.
- Need to understand how a patient's preference for a certain treatment over another affects his/her outcome response.
- A patient may have different psychological response to a treatment he/she deems favorable.
- In fact, it could be that the preference effects are larger than the direct effect of the treatment.

Areas of Substantial Impact in Clinical Trials

- Unblinded Trials
- Behavioral Interventions
- Trials where patients must sustain a demanding role
- Trials of equivalence or non-inferiority

Unblinded Trials

- While blinding is the gold standard, may not be possible to blind participants.
- Patient preference for treatment may be different and could bias results of the trial through systematic bias introduced by patient beliefs.[3–6]
    - Example: Several well-designed studies of airway clearance techniques were prematurely terminated,[7,8] with others compromised,[9,10] due to the patterns of dropouts attributable to patient preferences.
- Patients may believe newer treatment is better.

Behavioral Interventions/Demanding Role

- The motivation of a patient to follow a treatment may be influenced by a preference a patient has before beginning any course of action.[11]
- Allowing the patient to choose the intervention that is most suitable, in terms of level of commitment and willingness to participate in the task, may lead to enhanced motivation and thus efficacy of the intervention.[11,12]
- It is important to target treatments toward the patients who understand why they are being prescribed and are motivated to adhere to that treatment.

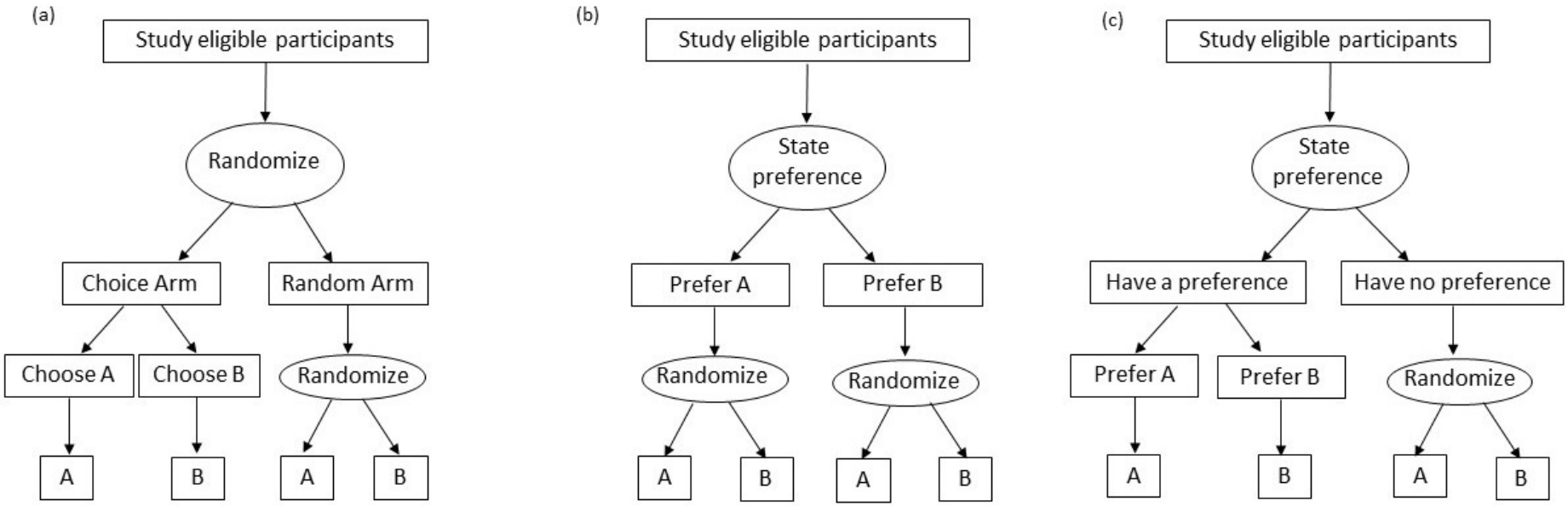
Equivalence/Non-inferiority

- In cases where equivalence has already been demonstrated among multiple treatments, accurate estimation of the impact of patient preference may help clinicians to better treat patients.

- This is especially important with participatory treatments, where patients are educated and motivated to manage their conditions without the assumption that there is only one right way.[11]
    - Example: Group versus individualized therapy for the treatment of psychiatric illnesses such as depression and anxiety.

- Shared decision-making models promote better outcomes and are especially important for patients for which multiple reasonable treatment options exist, with none clearly outperforming the others.[13]

Examples of Trial Designs Incorporating Patient Preference



Designs: (a)Two-stage randomized clinical trial[15]; (b) Fully randomized design; (c) Partially randomized preference design.

Simple linear model

Response y_{ijk} for patient k receiving treatment i and prefers treatment j is modeled as:

$$y_{ijk} = \mu + \tau_i + \nu_j + \pi_{ij} + \epsilon_{ijk}$$

μ is the overall response
τ_i treatment effect of treatment i
ν_j selection effect of preferred treatment j
π_{ij} preference effect for the received treatment i and preferred treatment j
ϵ_{ijk} random error term

We assume that all patients have a treatment preference (no undecided/indifferent individuals) and follow some constraints (i.e., $\sum \tau_i = 0$).

Effect Estimates for two-treatments[15]

Response y_{ijk} for patient k receiving treatment i=1,2 and prefers treatment j=1,2 is modeled as:

$$y_{ijk} = \mu + \tau_i + \nu_j + \pi_{ij} + \epsilon_{ijk}$$

- *Treatment Effect*: The average effect a particular treatment will have in specified population, $\Delta_\tau = \tau_1 - \tau_2$.
- *Selection Effect*: The difference in treatment effect influenced by self-selection of a given treatment by patients, $\Delta_\nu = \nu_1 - \nu_2$.
- *Preference Effect*: The change in outcome from the interaction between a patient's preferred treatment and the treatment actually received, $\Delta_\pi = (\pi_{11} + \pi_{22}) - (\pi_{12} - \pi_{21}))/2$.

Estimable effects for different designs*

Effect	Design			
	Parallel Group	Two-stage	Fully randomized preference	Partially Randomized preference
Treatment Effect	Yes	Yes	Yes[a]	Yes[a]
Selection Effect	No	Yes	Yes[b]	Yes[b]
Preference Effect	No	Yes	Yes	No

[a]Valid in preference subgroups

[b]Potentially biased

*Reproduced from Walter et al.[15]

Estimable effects for different designs*

Effect	Design			
	Parallel Group	Two-stage	Fully randomized preference	Partially Randomized preference
Treatment Effect	Yes	Yes	Yes[a]	Yes[a]
Selection Effect	No	Yes	Yes[b]	Yes[b]
Preference Effect	No	Yes	Yes	No

[a]Valid in preference subgroups

[b]Potentially biased

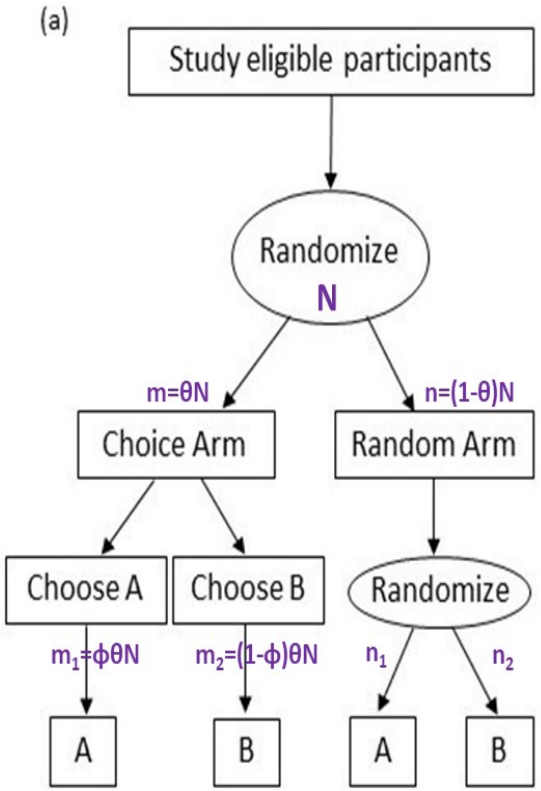
*Reproduced from Walter et al.[15]

Two-stage clinical trial design

- In a completely randomized trial, it is difficult to distinguish between a treatment that has failed because of its lack of efficacy and one that has failed because of patient preferences.
- In order to address this concern, the two-stage design could be become the new standard for testing behavioral interventions.
- While the two-stage design may be more complex than the standard RCT, and require a larger number of participants, it allows researchers the opportunity to assess the impact of these preferences on treatment effects.
- Therefore, the costs of increasing the sample size in a fully randomized design to detect preference effects needs to be weighed against the probability that preference effects could substantially bias the outcomes.[3]

Assumptions of the Two-Stage Design proposed by Rucker[16]

- Randomization in the first stage has no influence on a patient's response.
- The rate of preference for a particular treatment will be the same in the choice arm as in the random arm.
- The test statistics follow an approximate normal distribution (based on method of moments approach).



Observable Data

Actual Treatment		Choice arm (total sample size=m)		Random Arm (total sample size=n)
		Choose A	Choose B	
1	Sample size	m_1		n_1
	Observations	$x_{11}, \cdots, x_{1m_1}$		$y_{11}, \cdots, y_{1n_1}$
	Mean	μ_{11}	(μ_{12})	μ_1
2	Sample size		m_2	n_2
	Observations		$x_{21}, \cdots, x_{2m_2}$	$y_{21}, \cdots, y_{2n_2}$
	Mean	(μ_{12})	μ_{22}	μ_2

Note: m_1 is a random variable following a binomial distribution with probability ϕ, proportion of participants preferring treatment A.

Premised on principles of randomization, we can determine:

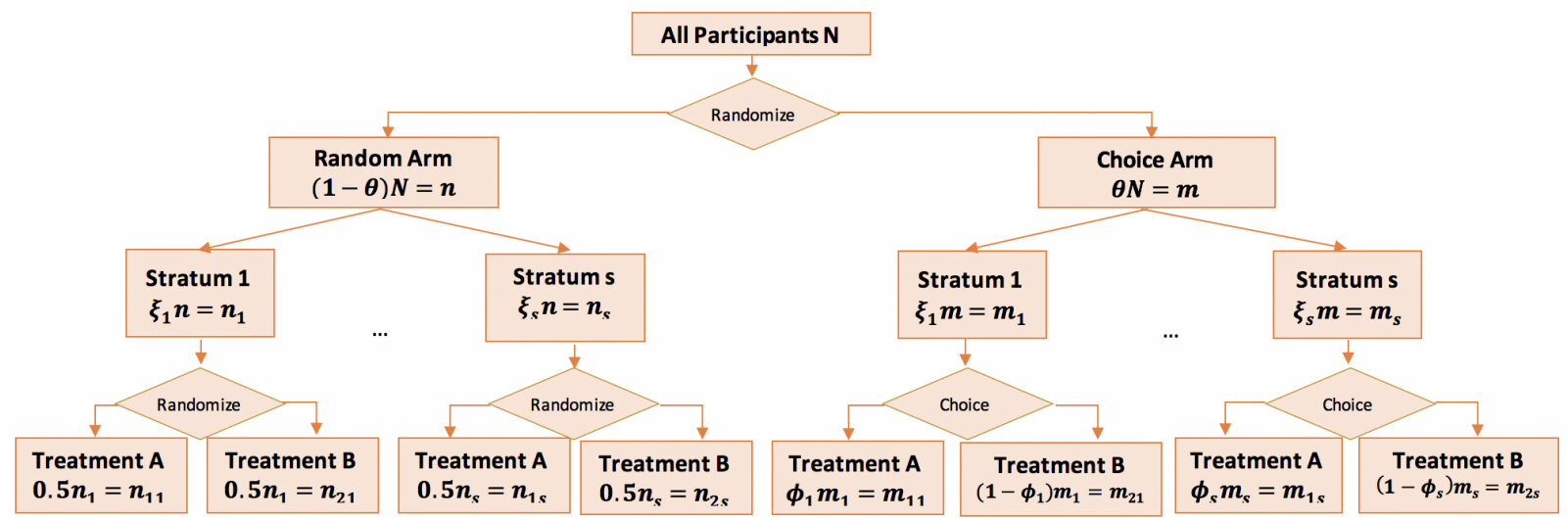
$$\mu_{12} = \frac{\mu_1 - \phi\mu_{11}}{1-\phi} \text{ and } \mu_{21} = \frac{\mu_2 - (1-\phi)\mu_{22}}{\phi}$$

Expansions in Two-Stage Design Methodology

- Sample Size Estimation[17]
- Optimal allocation of participants[18]
- Include Stratification[19]
- Binary Outcomes[20]
- Count Outcomes[21]
- Undecided Individuals[22]
- Survival[23]
- Group Sequential Monitoring[24]

Stratification

- Concern that certain subgroups may have differential preference rates which can be addressed in the design of the study.
  - Ex: Older individuals may prefer a medical intervention, while younger individuals may prefer a surgical intervention.
- Define a new parameter ξ_l, the proportion of individuals in stratum l.



- Construct an overall test statistic that is calculated through a weighted sum of the stratum-specific test statistics.

Binary/Count Outcomes

- Needed to expand to include alternative outcomes
  - Binary: Received treatment or attained "cure" for Hepatitis C
  - Count: Number of infections

- For large sample, can assume the test statistic follows normal distribution. Normal approximation tends to work well for sample sizes $N > 50$.

- For small sample, derived a conditional exact distribution for both binary[20] and count data[21].

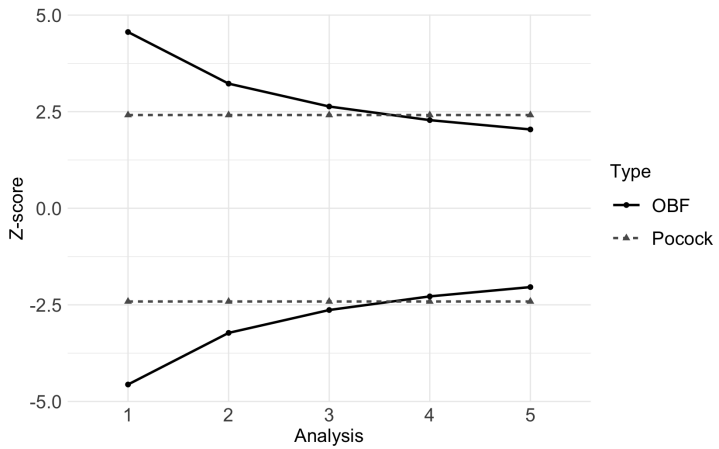
- Developed closed form sample size under normal approximation.

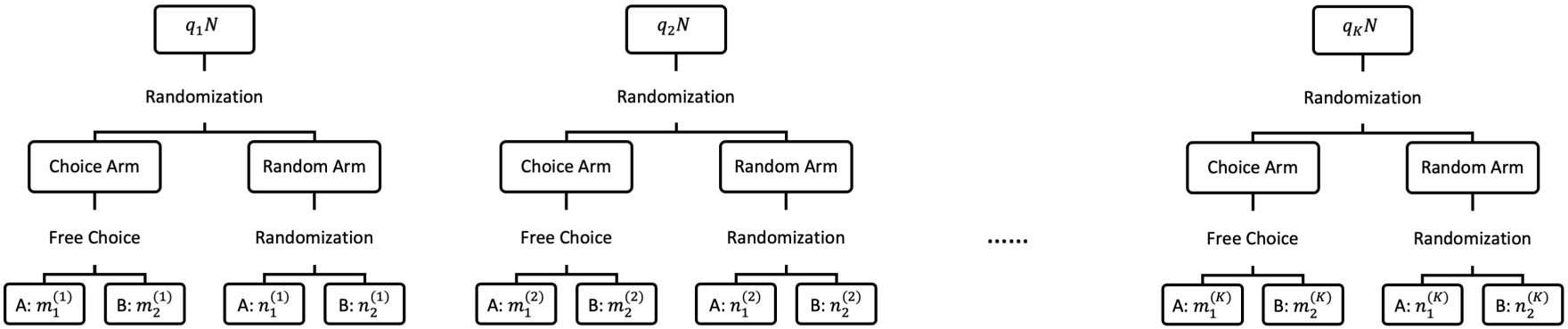
- For count, found results were robust: no observable difference when considered over-dispersion and zero-inflation.

- Expansion includes methods for both unstratified and stratified models.

Group Sequential Monitoring

- One of the limitations of employing of the two-stage design may be that the sample size required is almost always larger than a traditional RCT.
- Developed a sequential monitoring technique to reduce the expected sample size for a two-stage trial.
- Allows early stopping of the trial if sufficient evidence of the primary objective is obtained during a planned interim analysis.
- We considered application of sequential monitoring methods to two-stage design for continuous and binary measures[24] using both O'Brien-Fleming[25] and Pocock[26] stopping boundaries.





First Analysis:

Participants Accrued in Each Arm	
Choice Arm A	$m_1^{(1*)} = m_1^{(1)}$
Choice Arm B	$m_2^{(1*)} = m_2^{(1)}$
Random Arm A	$n_1^{(1*)} = n_1^{(1)}$
Random Arm B	$n_2^{(1*)} = n_2^{(1)}$

Second Analysis:

Participants Accrued in Each Arm	
Choice Arm A	$m_1^{(2*)} = m_1^{(1)} + m_1^{(2)}$
Choice Arm B	$m_2^{(2*)} = m_2^{(1)} + m_2^{(2)}$
Random Arm A	$n_1^{(2*)} = n_1^{(1)} + n_1^{(2)}$
Random Arm B	$n_2^{(2*)} = n_2^{(1)} + n_2^{(2)}$

.....

K^{th} Analysis:

Participants Accrued in Each Arm	
Choice Arm A	$m_1^{(K*)} = m_1^{(1)} + \cdots + m_1^{(K)}$
Choice Arm B	$m_2^{(K*)} = m_2^{(1)} + \cdots + m_2^{(K)}$
Random Arm A	$n_1^{(K*)} = n_1^{(1)} + \cdots + n_1^{(K)}$
Random Arm B	$n_2^{(K*)} = n_2^{(1)} + \cdots + n_2^{(K)}$

Illustrative Example Results - HCV Treatment

			Average GS-TSPD		Maximum GS-TSPD	
Outcome	Effect	Fixed Design	Pocock	OBF	Pocock	OBF
HRQoL	N_τ	337	244	263	388	343
	N_ν	223	185	195	256	227
	N_π	50	45	46	58	51
Cure Rate	N_τ	764	543	612	880	777
	N_ν	863	628	688	993	877
	N_π	322	234	260	371	328

Required sample sizes for detecting treatment (N_τ), selection (N_ν), and preference effects (N_π) for health related quality of life (HRQoL) and Hepatitis C virus (HCV) cure using fixed sample two-stage preference design (TSPD) and group sequential (GS)-TSPD for HCV patients receiving care at either a specialty clinic or mobile medical clinic.

Note: HRQoL and ultimate cure rate are different endpoints under which the required sample size calculations are subject to different design parameters, including true effect size. For HRQoL,we assume the following: treatment effect $\Delta_\tau = 2$, selection effect $\Delta_\nu = 2.1$, and preference effect $\Delta_\pi = 3.9$. For HCV cure rate, we assume: $\Delta_\tau = 0.15$,$\Delta_\nu = 0.185$, and$\Delta_\pi = 0.297$. There are at most L = 3 analyses. We set the type I error rate to 0.05 and the power to 90%.

Key Takeaways - Two-stage design

- The two-stage randomized trial design is a useful tool for researchers seeking to understand the importance of patient treatment preferences in determining health outcome.
- There is evidence that preference and selection effects may be more important to consider when the treatment effect is small or nonexistent.[27]
- There has been a lot of work to extend and improve methods for two-stage clinical trials to measure the impact of patient preference on trial outcomes.

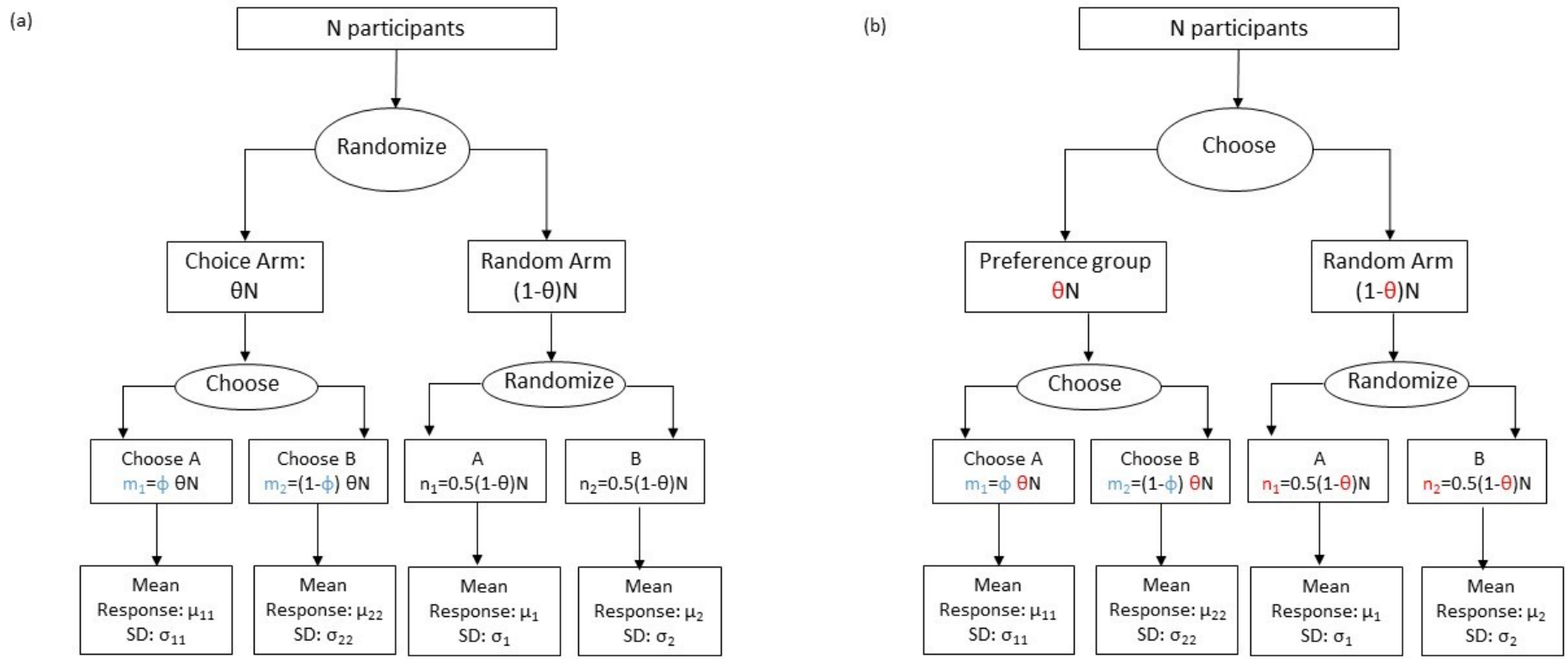
Partially Randomized Preference Design (PRPD)

- Lack of equipoise can result when individuals having strong preferences for one of the treatments, such as pharmacological vs non pharmacological therapy for pain management, which can lead to declining enrollment resulting in a less pragmatic and generalizable trial.[28,29]

- With PRPD, patients with a treatment preference are given their preferred treatment.

- May be a more pragmatic alternative to the two-stage design, but the preference effect is not directly estimable and the selection effect is potentially biased.[15]
  - Patient may not understand the two stages of randomization.
  - Patients may be willing to contribute data but not willing to be randomized.

Partially Randomized Design



Comparison of the two-stage (panel a) and partially randomized (panel b) preference designs.

Estimable effects for different designs*

Effect	Design			
	Parallel Group	Two-stage	Fully randomized preference	Partially Randomized preference
Treatment Effect	Yes	Yes	Yes[a]	Yes[a]
Selection Effect	No	Yes	Yes[b]	Yes[b]
Preference Effect	No	Yes	Yes	No

[a]Valid in preference subgroups

[b]Potentially biased

*Reproduced from Walter et al.[15]

Methodology Development PRPD[30]

- In removing the first stage of randomization, we remove the mathematical properties we gain with randomization, and the ability to estimate unbiased treatment preference effects.
- We proposed using a propensity score model (stratification) to mimic the first stage of randomization and allow for estimation of the overall treatment, selection and preference effects with minimal bias.
- We derived a set of closed form sample size formulas for detecting all three effects through adaption of stratified methodology we developed.[19]
- We found we needed 5 to 10 propensity score strata to correct for biases, with the exact number potentially depending on treatment effect heterogeneity.

Limitations of Propensity Score PRPD

- Assume that all patients have an underlying treatment preference (no undecided patients), but there is a subset who are unwilling to state preference and/or willing to be randomized.
  - This is a strong assumption - and possible deviation can lead to bias under our current proposed methods.

- Not unique to our model, validity is dependent on conditional exchangeablity assumption (no unmeasured confounders) and correct specification of the propensity score model.

- We can measure a set of clinically important covariates that are associated with first stage allocation.

- Only considered stratification - which may not perform well for small sample sizes and could lead to residual bias in presenece of strong confounding.

Other Areas to Consider for Future Work

- Undecided patients for all outcome distributions
- Impact of differential dropout
- Longitudinal outcomes
- Clustering
- More pragmatic designs - extensions for the partially randomized preference design (i.e., propensity score weighting or matching; undecided participants)
- Adjustment for covariates beyond accounting for stratification[31]

Ackowledgements

This work would not be possible without contributions of trainees and key collaborators.

Trainees

- Dr. Briana Cameron
- Yu Shi
- Yumin Wang
- Can Meng
- Ruyi Liu
- Xian Gu

Key Collaborators

- Dr. Fan Li
- Dr. Peter Peduzzi
- Dr. Michael Kane
- Dr. Ondrej Blaha
- Dr. Mary Ryan
- Dr. Stephen Walter

References

1. http://www.pcori.org/sites/default/files/PCORI-PFA-2015-Cycle3-Methods.pdf
2. McPherson K, Britton A. Preferences and understanding their effects on health. *Quality in Health Care. 2001; 10(Suppl I):i61-i66.*
3. *Halpern SD. Evaluating preference effects in partially unblinded, randomized clinical trials. Journal of Clinical Epidemiology. 2003;56:109-115.*
4. *McPherson K, Britton AR, Wennberg JE. Are randomized controlled trials controlled? Patient preferences and unblind trials. Journal of the Royal Society of Medicine. 1997;90:652-656.*
5. *Mu W, Shang H. Understanding patient values and the manifestations in clinical research with traditional chinese medicine - with practical suggestion for trial design and implementation. Evidence-Based Complementary and Alternative Medicine. 2013.*
6. *McPherson K. Do patients' preferences matter? British Medical Journal. 2008;337:a2034.*
7. *McIlwaine M, Wong L, Chilvers M, Davidson G. Long-term comparative trial of two different physiotherapy techniques: postural drainage with percussion an autogenic drainage, in the treatment of cystic fibrosis. Pediatric Pulmonology. 2010;45:1064-1069.*
8. *Sontag MK, Quittner AL, Modi AC, et al. Lessons learned from a randomized trial of airway secretion clearance techniques in cystic fibrosis. Pediatric Pulmonology. 2010;45(3):291-300.*
9. *McIlwaine M, Alarie N, Davidson GF, et al. Long-term multicentre randomised controlled study of high frequency chest wall oscillation versus positive expiratory pressure mask in cystic fibrosis. Thorax. 2013;66(8):746-751.*
10. *Pryor JA, Tannenbaum E, Scott SF, et al. Beyond postural drainage and percussion: Airway clearance in people with cystic fibrosis. Journal of Cystic Fibrosis. 2010;9(3):187-192.*
11. *Brewin CR, C B. Patient preferences and randomised clinical trials. British Medical Journal. 1989;299:313-315.*
12. *Clark NM, Janz NK, Dodge JA, et al. The effect of patient choice of intervention on health outcomes. Contemporary Clinical Trials. 2008;29:679-686.*
13. *Barclay L. Multiple Sclerosis Discovery Forum. MGH and ACP; 2014.*
14. https://www.cdc.gov/hepatitis/hcv/hcvfaq.htm#section1

References (cont)

15. Walter SD, Turner R, Macaskill P, McCaffery KJ, Irwig L. Beyond the treatment effect: Evaulating the effects of patient preferences in randomised trials. *Statistical Methods in Medical Research. 2017;26(1): 489-507.*
16. *Rucker G. A two-stage trial design for testing treatment, self-selection and treatment preference effects. Statistics in Medicine. 1989; 8:477-485.*
17. *Turner RM, Walter SD, Macaskill P. McCaffery KJ, Irwig L. Sample size and power when designing a randomized trial for the estimation of treatment, selection, and preference effects. Medical Decision Making. 2014 Aug;34(6):711-9.*
18. *Walter SD, Turner RM, Macaskill P, McCaffery KJ, Irwig I. Optimal allocation of participants for the estimation of selection, preference and treatment effects in two-stage randomised trial design. Statistics in Medicine. 2012;31(13): 1307-1322.*
19. *Cameron B, Esserman DA. Sample size and power for stratified doubly randomized preference design. Stat Methods Med Res. 2016.*
20. *Cameron B, Peduzzi P, Esserman DA. Extensions to the two-stage randomized trial design for testing treatment, self-selection, and treatment preference effects to binary outcomes. Statistic in Medicine. 2018; 37:3147-3178.*
21. *Shi Y, Cameron B, Gu X, Kane M, Peduzzi P, Esserman DA. Two-stage randomized trial design for testing treatment, preference and self selection effects for count outcomes. Statistics in Medicine. 2020; 39:3653-3683.*
22. *Walter SD, Turner RM, Macaskill P, McCaffery KJ, Irwig L. Estimation of treatment preference effects in clinical trials when some participants are indifferent to treatment choice. BMC Med Res Methodol. 2017; 17(1):29.*
23. *Chahine RA, Aban I. Analyis of survival outcomes using likelihood ratio test in trials incorporating patient's treatment choice. Journal of Applied Statistics. 2023. DOI: 10.1080/02664763.2023.2199177*
24. *Liu R, Li F, Esserman D, Ryan M. Group sequential two-stage preference designs. Statistics in Medicine. 2024; 43:315-341.*

References (cont)

25. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics. 1979;35(3):549-556.
26. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. Biometrika. 1977;64(2):191-199.
27. McCaffery KJ, Turner R, Mccaskill P, Walter SD, Chan SF, Irwig L. Determining the impact of informed choice: separating treatment effects from the effects of choice and selection in randomized trials. *Med Decis Making. 2011; 31(2): 229-236.*
28. *Peduzzi P, Kyriakides T, O'Connor TZ, Guarino P, Warren SR, Huang GD. Methodological issues in comparative effectiveness research: clinical trials. Am J Med. 2010;123(12 Suppl 1):e8-15.*
29. *Sidani S, Fox M, Epstein D. Conducting a two-stage preference trial: utility and challenges. Int J Nurs Stud. 2015;52(5):1017-1024.*
30. *Wang Y, Li F, Blaha O, Meng C, Esserman D. Design and analysis of partially randomized preference trials with propensity score stratification. Statistical Methods in Medical Research. 2022; 31(8): 1515-1537.*
31. *Long Q, Little RJ, Lin X. Causal inference in hybrid intervention trials involving treatment choice. J AmStat Assoc. 2008;103(482):474-484.*

Example: Sample Size Estimation for Binary Outcome

Recall: We have three effects that we can estimate in the two-stage design.

- *Treatment Effect*: $\Delta_\tau = \tau_1 - \tau_2$.
- *Selection Effect*: $\Delta_\nu = \nu_1 - \nu_2$.
- *Preference Effect*: $\Delta_\pi = ((\pi_{11} + \pi_{22}) - (\pi_{12} + \pi_{21}))/2$

Closed Form Sample size Calculations[17]

1. Derive the variance
2. Assuming approximate normality, we can use the standard sample size formula:

$$N = \frac{2\text{Var}(\delta)(Z_{1-\alpha/2} + Z_{1-\beta})^2}{\delta^2}$$

3. Construct formula through substitution

Sample Size Formulas for Binomial Distribution[20]

$$N_\tau = 2 * \frac{p_1(1-p_1)+p_2(1-p_2)}{(1-\theta)\Delta\tau^2}\left(Z_{1-\alpha/2}+Z_{1-\beta}\right)^2$$

$$N_\nu = \frac{\left(Z_{1-\alpha/2}+Z_{1-\beta}\right)^2}{4\theta\Delta\nu^2\phi^2(1-\phi)^2}\left[\phi p_{11}(1-p_{11})+(1-\phi)p_{22}(1-p_{22})+\frac{\left(\phi^2 d_1+(1-\phi)^2 d_2\right)^2}{\phi(1-\phi)}\right.$$
$$\left.+2\left(\frac{\theta}{1-\theta}\right)\left(\phi^2 p_1(1-p_1)+(1-\phi)^2 p_2(1-p_2)\right)\right]$$

$$N_\pi = \frac{\left(Z_{1-\alpha/2}+Z_{1-\beta}\right)^2}{4\theta\Delta\pi^2\phi^2(1-\phi)^2}\left[\phi p_{11}(1-p_{11})+(1-\phi)p_{22}(1-p_{22})+\frac{\left(\phi^2 d_1-(1-\phi)^2 d_2\right)^2}{\phi(1-\phi)}\right.$$
$$\left.+2\left(\frac{\theta}{1-\theta}\right)\left(\phi^2 p_1(1-p_1)+(1-\phi)^2 p_2(1-p_2)\right)\right]$$

Motivating Example: Hepatitis C Treatment

- Approximately 3.5 million people in the United States are currently estimated to be infected with Hepatitis C virus, a leading cause of death and end-stage liver disease.[14]
- Vulnerable population, which includes significant number of alcohol and drug users.
- Many providers used to require patients to have abstained from drug or alcohol use for a minimum of 6 months - reducing likelihood of seeking treatment.

Motivating Example (cont)

- Mobile medical clinics (MMC) are an attempt to remove many barriers associated with the traditional healthcare setting or seeking treatment at a specialty clinic.
- Believed that alcoholics and drug users would have different preference for receiving care in the specialty clinic versus the MMC and that this would influence outcomes.
- Outcome(s) of interest: the proportion of patients initiating treatment (or those achieving sustained Virologic response).

Motivating Example (cont)

Hypothesis: Allowing patients with Hepatitis C to choose where they receive treatment - mobile medical clinic or specialty clinic - will result in improved health outcomes.

Primary outcome: Hepatitis C sustained virologic response ("cure").

Stratification variable: Alcohol and Drug User Status.

Assumptions

Proportion cured of Hepatitis C for stratum of drug and alcohol users					
	Random	Choice			
Actual treatment	Response proportion	Specialty	MMC	Preferred Specialty	Preferred MMC
Specialty Clinic	0.65	0.75	-	0.75	$\frac{0.65-0.3x0.75}{0.7} = 0.61$
MMC	0.5	-	0.7	$\frac{0.5-0.7x0.7}{0.3} = 0.03$	0.7

Proportion cured of Hepatitis C for stratum of nondrug and alcohol users					
	Random	Choice			
Actual treatment	Response proportion	Specialty	MMC	Preferred Specialty	Preferred MMC
Specialty Clinic	0.85	0.9	-	0.9	$\frac{0.85-0.5x0.9}{0.5} = 0.8$
MMC	0.7	-	0.9	$\frac{0.7-0.5x0.9}{0.5} = 0.5$	0.9

MMC = Mobile Medical Clinics
30% of individuals will be drug users and 70% will not.
70% of drug users will prefer the MMC and 50% of non-users will prefer the MMC.

$$N_\nu = \frac{\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{4\theta\Delta\nu^2} \sum_{l=1}^{s} \frac{\xi_l}{\phi_l^2(1-\phi_l)^2} \left[\phi_l p_{11l}(1-p_{11l}) + (1-\phi_l)p_{22l}(1-p_{22l}) + \frac{\left(\phi_l^2 d_{1l} + (1-\phi_l)^2 d_{2l}\right)^2}{\phi_l(1-\phi_l)} + 2\left(\frac{\theta}{1-\theta}\right)\left(\phi_l^2 p_{1l}(1-p_{1l}) + (1-\phi_l)^2 p_{2l}(1-p_{2l})\right)\right]$$

where $\Delta\nu = \sum_{l=1}^{s} \xi_l \Delta\nu_l = \sum_{l=1}^{s} \xi_l \frac{\phi_l d_{1l} - (1-\phi_l)d_{2l}}{2\phi_l(1-\phi_l)}$ is the weighted average of the stratum-specific selection effects.

Calculations of the Effects

Preference and selection effect sizes for drug and alcohol users

Preference Effect		**Selection Effect**	
Average response proportion: received preferred treatment=0.725	Average response proportion: did not receive preferred treatment=0.32	Average response proportion: prefer specialty=0.39	Average response proportion: prefer MMC=0.655
Preference effect=0.725-0.32=0.405		Selection effect=0.39-0.655=-0.265	

Preference and selection effect sizes for non-drug and alcohol users

Preference Effect		**Selection Effect**	
Average response proportion: received preferred treatment=0.9	Average response proportion: did not receive preferred treatment=0.65	Average response proportion: prefer specialty=0.7	Average response proportion: prefer MMC=0.85
Preference effect=0.9-0.65=0.25		Selection effect=0.7-0.85=-0.15	

Sample Size Estimates

	80% Power		**90% Power**	
	Stratified	Unstratified	Stratified	Unstratified
N_τ	265	277	355	371
N_π	303	288	405	385
N_ν	789	817	1056	1093

Example for Count Outcome

- Antimicrobial use at the end of life.
  - Fever and infections common acute complications experienced by terminally ill patients.
  - Antimicrobials often prescribed with non-trivial risks.
  - Caregivers may forgo extended survival for greater quality of life.
  - No trials to determine whether receiving versus not receiving antimicrobials results in prolonged survival and symptom relief in terminally ill.

- Two-stage design would be ideal to answer whether prescribing antimicrobials increases the number of symptoms experienced in patients in the end stages of life (treatment effect) and whether patients/caregivers who prefer to receive antimicrobials over not receiving antimicrobials experience less symptoms (selection effect).

Assumptions and Sample Size Estimates

Parameter	Value	Description
ϕ	0.4	40% of patients/caregivers prefer antimicrobials
λ_1	6	Mean number of symptoms in those randomized to receive antimicrobials
λ_2	4	Mean number of symptoms in those randomized to not receive antimicrobials
λ_{11}	5	Mean number of symptoms in those receiving antimicrobials and choosing to receive antimicrobials
λ_{22}	5	Mean number of symptoms in those not receiving antimicrobials and choosing not to receive

$\Delta_\tau = \lambda_1 - \lambda_2 = 2$ $\quad \Delta_\nu = \frac{\phi(\lambda_{11}-\lambda_1)-(1-\phi)(\lambda_{22}-\lambda_2)}{2\phi(1-\phi)} = -2.08$ $\quad \Delta_\pi = \frac{\phi(\lambda_{11}-\lambda_1)+(1-\phi)(\lambda_{22}-\lambda_2}{2\phi(1-\phi)} = 0.85$

$$N_\nu = \frac{(Z_\alpha + Z_\beta)^2}{4\theta\Delta\nu^2\phi^2(1-\phi)^2}\left[\phi\lambda_{11} + (1-\phi)\lambda_{22} + \frac{\left(\phi^2 d_1 + (1-\phi)^2 d_2\right)^2}{\phi(1-\phi)} + 2\left(\frac{\theta}{1-\theta}\right)\left(\phi^2\lambda_1 + (1-\phi)^2\lambda_2\right)\right]$$

	Type I Error=5%		Type I Error=2.5%	
	80% Power	90% Power	80% Power	90% Power
N_τ	79	106	96	125
N_ν	157	210	190	248
N_π	4289	5741	5193	6781

Analysis of Survival Outcomes Using Likelihood Ratio Test in Trials Incorporating Patient's Treatment Choice

Rouba A. Chahine, PhD [1] Inmaculada Aban, PhD [2]

[1] RTI International

[2] University of Alabama At Birmingham

May 22, 2024

Overview

- Introduction
- Likelihood based approach, LRT
- Evaluate the LRT using simulations
- Design a study
- Conclusions & Limitations

RTI INTERNATIONAL

Introduction

Currently Available Methods of analysis

- ANOVA approach
  - Normal outcomes, Rücker (1989) & Turner et al. (2014)
  - Case where participants are indifferent to treatment choice, Walter et al. (2017)
  - Binary outcomes, Cameron et al. (2018)
  - Count outcomes, Shi et al. (2020)
  - Group Sequential, Liu et al. (2024)
- likelihood approach
  - Causal inference in hybrid intervention trials involving treatment choice, Long et al. (2008)

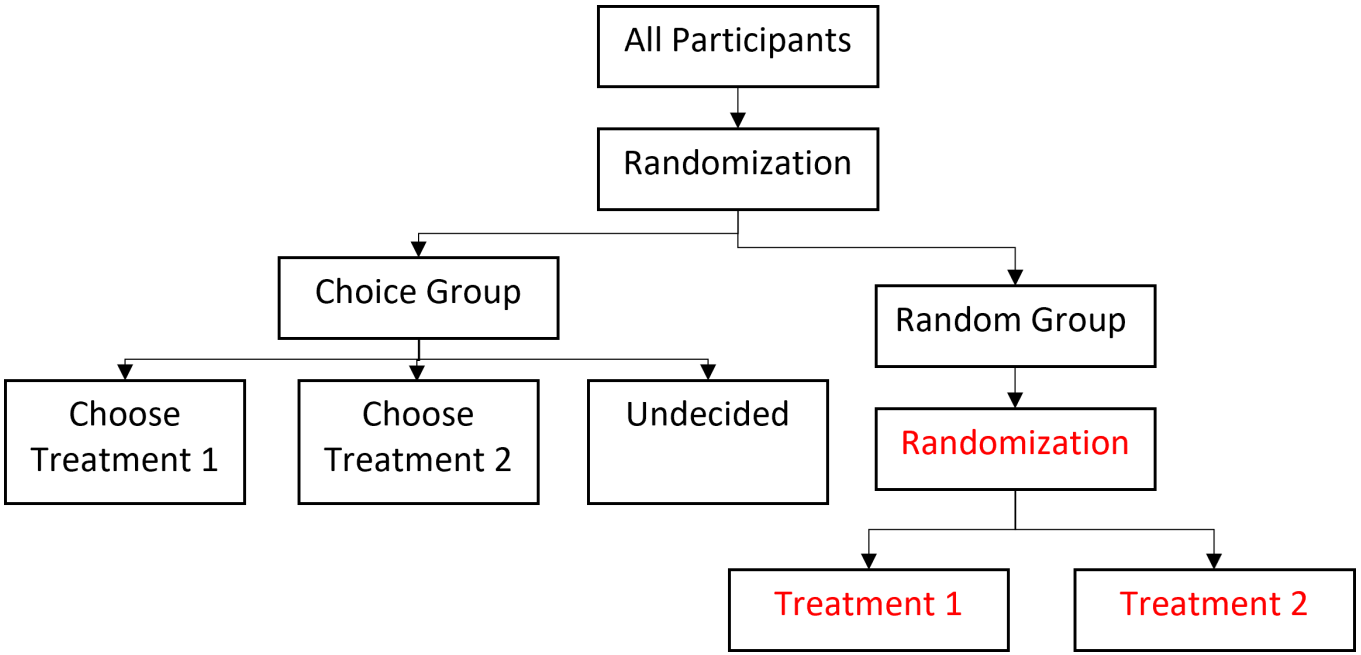
RTI INTERNATIONAL

Interest

- Time to treatment or diagnosis to an important clinical event
- Challenge:
  - Participants starts at enrollment but terminates before the end of the study without observing the event of interest
  - Not all participants join the study at the same time and will consequently have different follow-up times
  - Data collected from individuals who are lost to follow-up or never experienced the event of interest cannot be ignored as we know their survival time is at least as long as the period documented.
- Can we incorporate censoring in the model? Chahine et al. (2023)
- Can we incorporate covariates in the model? Chahine et al. (2024)

RTI INTERNATIONAL

Design Flow Chart



Notes:

- Red: Traditional clinical trials
- ”Undecided” is not considered here

Method Based on Likelihood Approach

Assumptions of the proposed method:

- Participants in the random arm are not asked about the treatment of their choice
- Participants in the choice arm will make a choice and will receive the treatment of their choice
- Non-informative censoring, i.e., the censoring mechanism does not contain information about the parameters of the event time distribution

RTI INTERNATIONAL

Method Based on Likelihood Approach

Notations

- T, treatment indicator
- C, choice indicator
- R, group randomization indicator
- S, censoring indicator
- Y, outcome of interest
- β & α, model parameters

RTI INTERNATIONAL

The Likelihood

- Conditional expected value of Y for choice and random group,

$$\mu_C = E(Y_i | T_i, \boldsymbol{X_i}, R_i = 0)$$
$$\mu_R = E(Y_i | Y_i, \boldsymbol{X_i}, R_i = 1)$$

with $g(.)$ monotonic link function,

$$g(\mu_C) = \beta_{C0} + \beta_{CT} T_i + \boldsymbol{\alpha X_i}$$
$$g(\mu_R) = \beta_{R0} + \beta_{RT} T_i + \boldsymbol{\alpha X_i}$$

- Conditional likelihood for N participants,

$$L = f(Y|T, R, \beta, \alpha) = \prod_{i=1}^{N} (f_C(y_i|t_i, \boldsymbol{\beta_C}, \alpha))^{(1-r_i)} (f_R(y_i|t_i, \boldsymbol{\beta_R}, \alpha))^{r_i}$$

The Likelihood

- To account for right-censoring, the likelihood is written as,

$$f(Y|T,R,\beta,\alpha) =$$
$$\prod_{i=1}^{N}\left[(f_C(y_i|t_i,\beta_{\boldsymbol{C}},\alpha))^{(1-s_i)}(S_C(y_i|t_i,\beta_{\boldsymbol{C}},\alpha))^{s_i}\right]^{(1-r_i)}$$
$$\times\left[(f_R(y_j|t_j,\beta_{\boldsymbol{R}},\alpha))^{(1-s_i)}(S_R(y_j|t_j,\beta_{\boldsymbol{R}},\alpha))^{s_i}\right]^{r_i}$$

RTI INTERNATIONAL

Test Statistic

- Likelihood Ratio Test,

$$\lambda(y) = \frac{L(\hat{\beta}_{H_0}, \hat{\alpha}_{H_0} | y, t, r)}{L(\hat{\beta}, \hat{\alpha} | y, t, r)}$$

- The null hypothesis, H_0, will be rejected when

$$-2\log(\lambda(y)) \leq \chi^2_{d,\alpha}$$

d degree of freedom calculated as the difference between the number of free parameters specified under the restricted model anf those under the null hypothesis

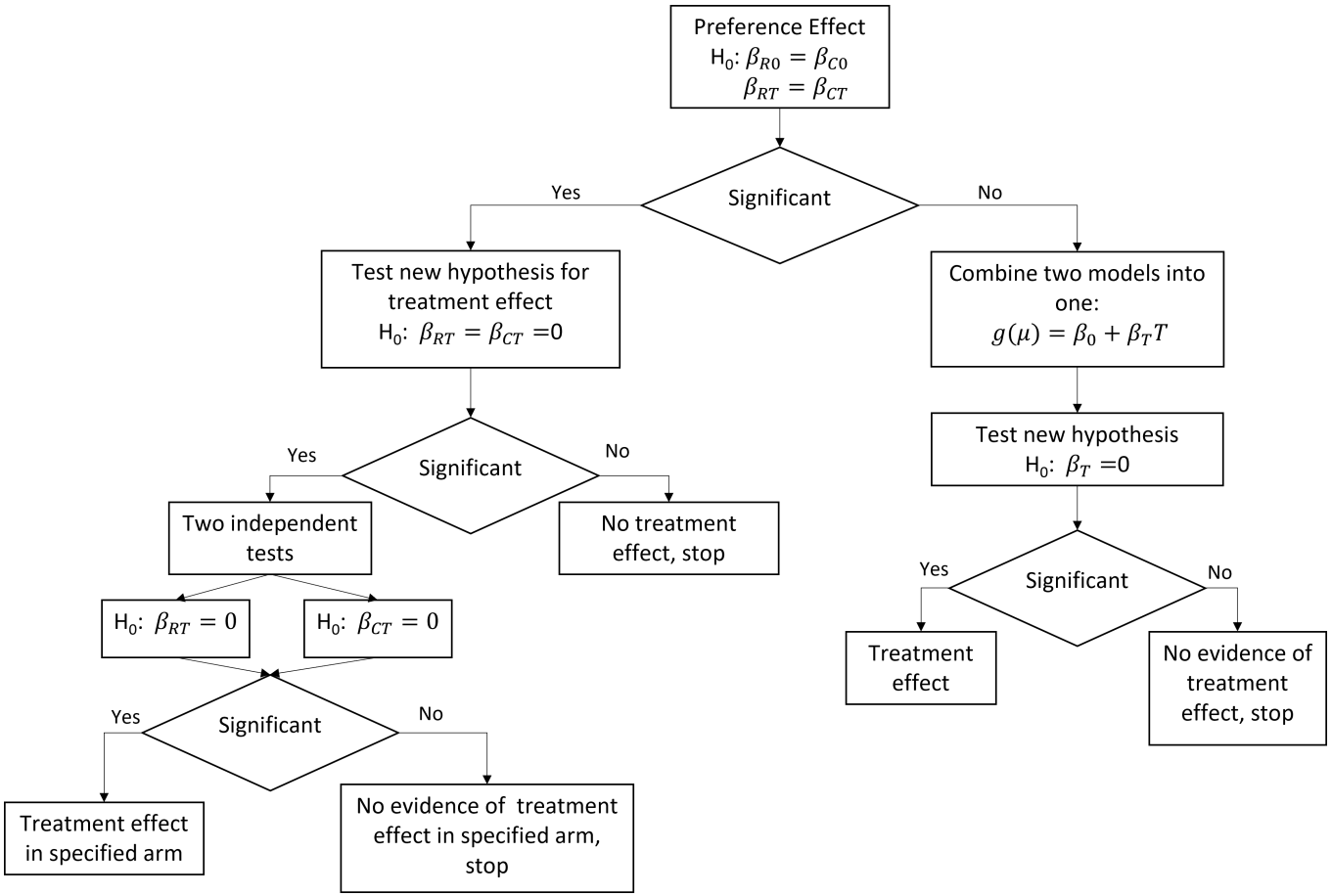
Test Statistic

Computational considerations:

- MLEs do not have closed forms
- OPTIM in R using a quasi-Newton algorithm (L-BFGS-B)

RTI INTERNATIONAL

Flow chart detailing suggested steps/tests to be performed



To account for multiple testing of nested hypotheses of the left side of the flow chart, we propose using approach similar to Holm's.

Simulations

Evaluating Likelihood Ratio Test Methods

- Estimates
- Type I Error
- Power

Simulations

For the survival simulations, we focus on Weibull distribution.

- Commonly used for failure time
- Weibull shape parameter "a" plays a role in failure rate
  - $a < 1$, decreased hazard
  - $a = 1$, constant hazard
  - $a > 1$, increased hazard

RTI INTERNATIONAL

Simulations, Examine the Estimates

1. 4,000 simulated trials with N=400
2. 50-50 between the choice and random arms
3. 50-50 randomization in the random arm
4. 40% proportion choosing Treatment 1
5. 3 cases of hazard: $a = 0.5, a = 1, a = 1.5$
6. 30% censoring

RTI INTERNATIONAL

Simulations, Examine the Estimates

Parameters	True value	Sample Mean (Sample Standard Error)		
		a=0.5	a=1	a=1.5
β_{R0}	2.5	2.492 (0.276)	2.497 (0.134)	2.498 (0.087)
β_{RT}	1	0.999 (0.373)	0.998 (0.182)	0.998 (0.123)
β_{C0}	2.5	2.485 (0.258)	2.493 (0.121)	2.497 (0.079)
β_{CT}	1.5	1.504 (0.375)	1.500 (0.189)	1.502 (0.126)
a		0.504 (0.025)	1.009 (0.052)	1.511 (0.076)

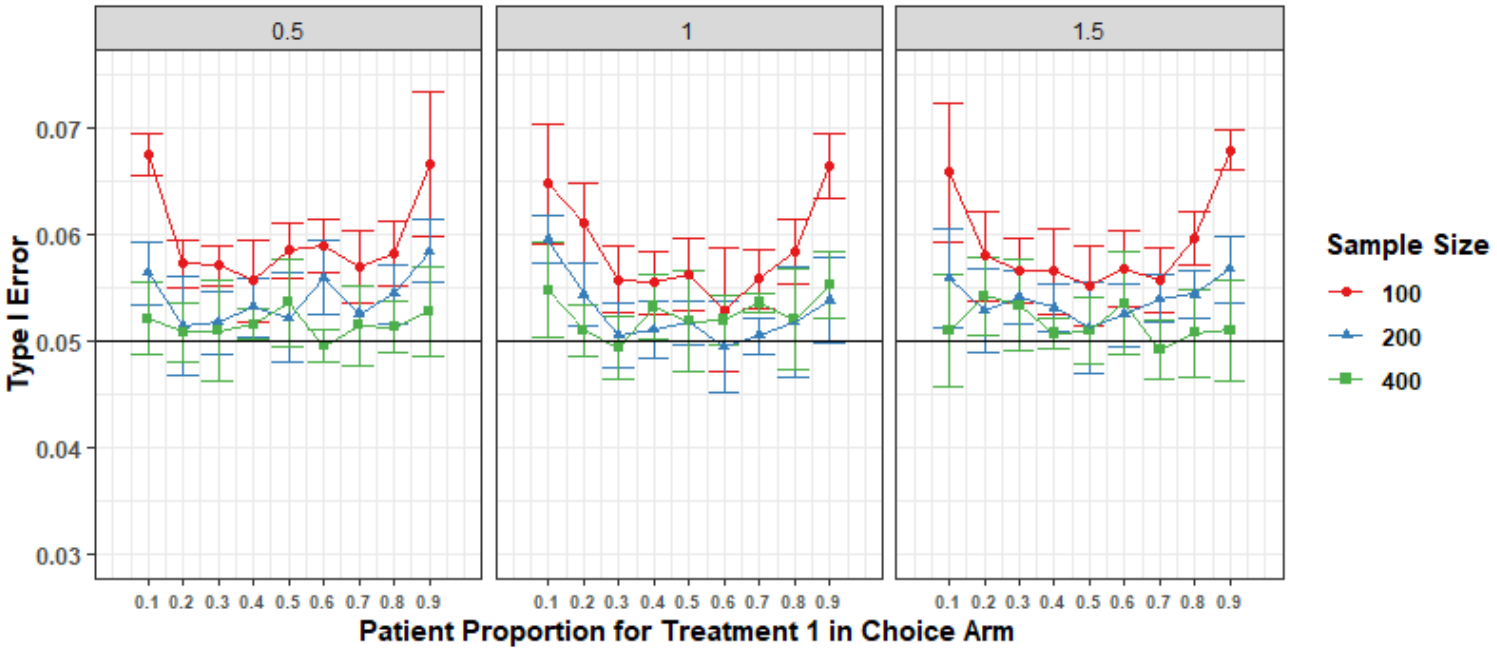
RTI INTERNATIONAL

Simulations, Type I Error

1. 4,000 simulated trials with N=100, 200 & 400
2. Re-run 10 times to account for simulation variation
3. 50-50 between the choice and random arms
4. Proportion choosing Treatment 1 vs Treatment 2 varied from 10% to 90%
5. 30% censoring

RTI INTERNATIONAL

Simulations, Type I Error for the Preference Effect

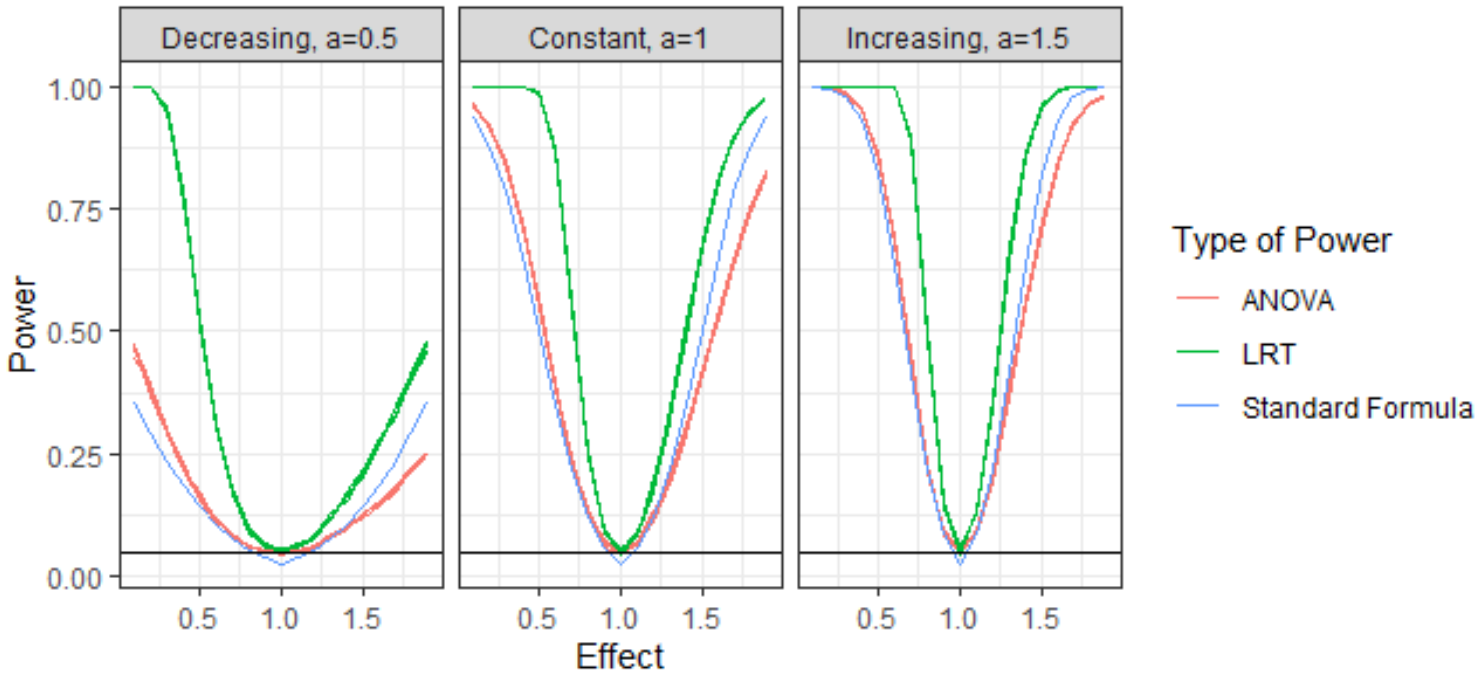


Simulations, Power

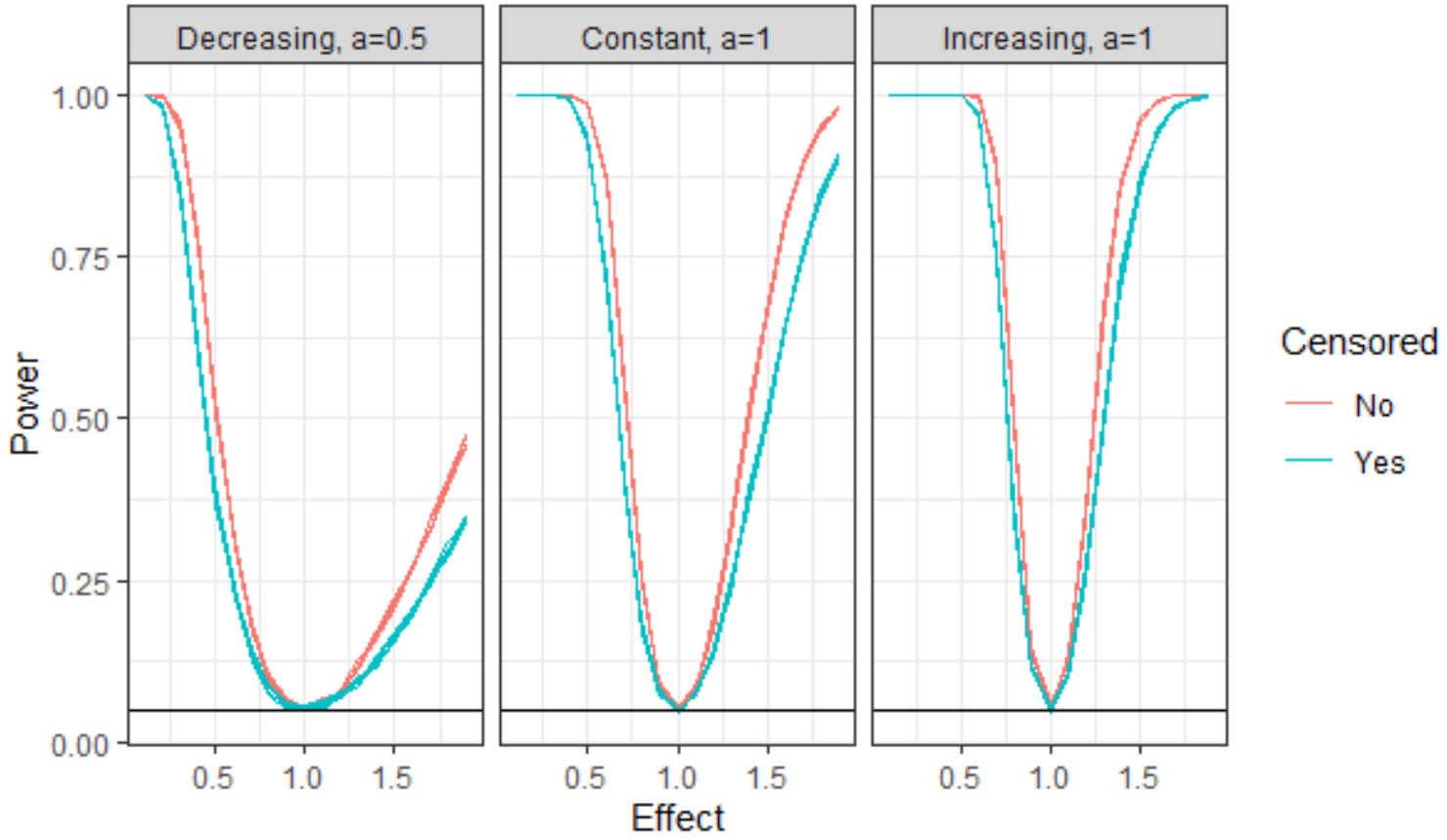
1. 4,000 simulated trials with N=400
2. Re-run 10 times to account for simulation variation
3. 50-50 between the choice and random arms
4. 40% Proportion choosing Treatment A
5. No covariates, nor censoring to enable comparison with ANOVA and standard sample size formula, Turner et al. (2014)

RTI INTERNATIONAL

Simulations, Power



Simulations, Power



Illustration

Hypothetical scenario based on a real study, STEP for MS, Molt et al. (2019)

What is STEP for MS?

- Looks at effectiveness of an exercise training program for participants with MS to improve walking performance, Timed-25-Feet Walk
- Compares between in-person facility based (FB) and telerehabilitation home based (TELE)
- Incorporate an choice arm where participants are allowed to choose mode of deliverance
- Non-inferiority study

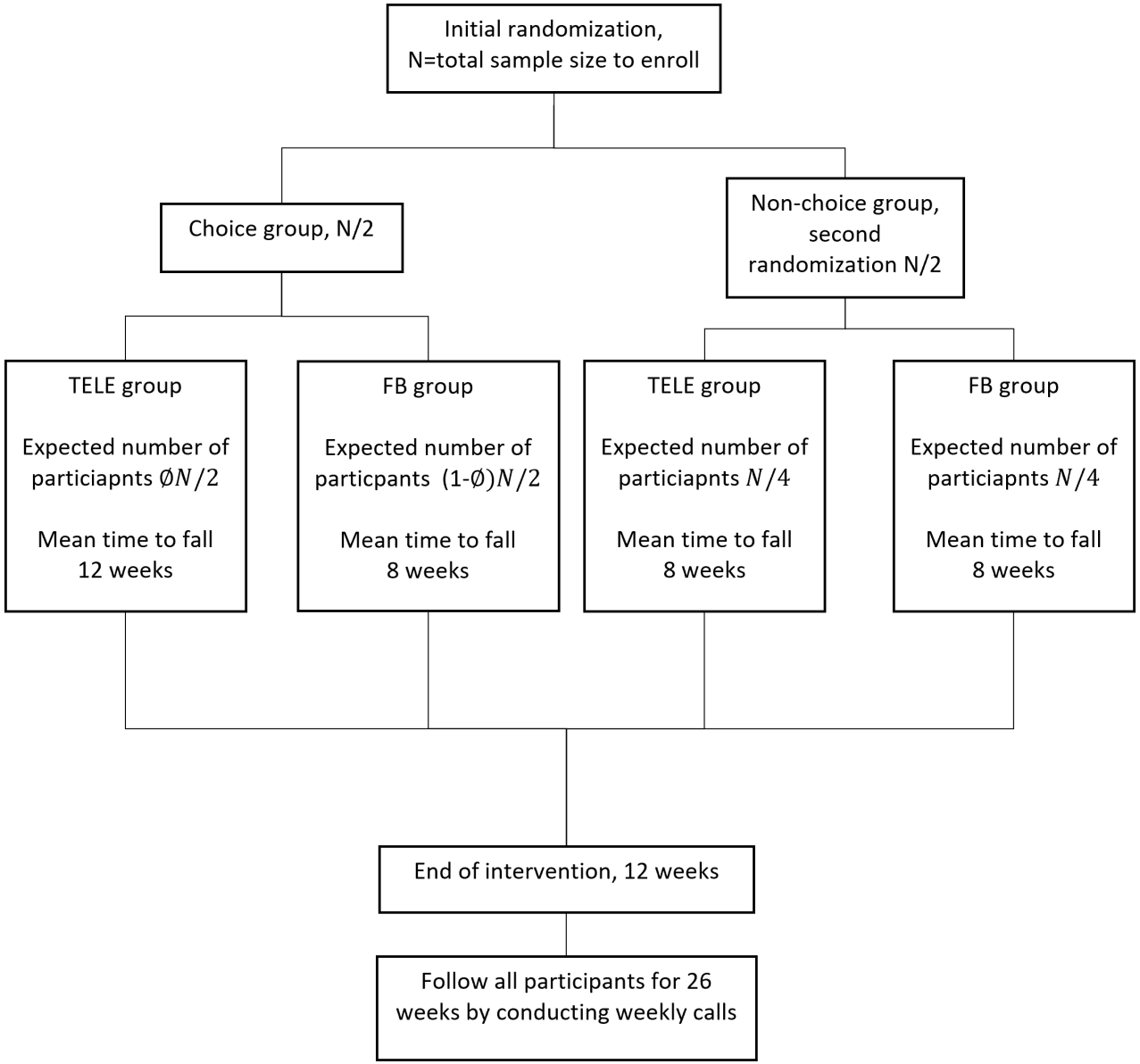
RTI INTERNATIONAL

Illustration

Our Scenario

- Time to first fall after randomization
- No treatment effect
- 50% increase in mean time until first fall for TELE choice group
- 10% drop out
- 30% censoring
- Three proportions of choosing TELE: 20, 40, and 80%

Illustration



Illustration

Power for preference effect for a hypothetical rehabilitation study in both cases of constant ($a = 1$) and decreased hazards ($a = 0.7$) of falling

ϕ	N	Choice Group		Random Group		Power	
		TELE	FB	TELE	FB	Decreasing hazard	Constant hazard
20%	400	40	160	100	100	20%	38%
	600	60	240	150	150	27%	52%
	800	80	320	200	200	36%	64%
	1000	100	400	250	250	42%	74%
	1200	120	480	300	300	50%	81%
40%	400	80	120	100	100	28%	52%
	600	120	180	150	150	40%	70%
	800	160	240	200	200	50%	82%
	1000	200	300	250	250	61%	91%
	1200	240	360	300	300	67%	95%
80%	400	160	40	100	100	36%	64%
	600	240	60	150	150	51%	83%
	800	320	80	200	200	65%	92%
	1000	400	100	250	250	74%	96%
	1200	480	120	300	300	82%	98%

RTI INTERNATIONAL

Conclusions & Limitations

- Proposed a likelihood based method as an alternative to ANOVA mehtod
  - Allows for the inclusion of censoring & covariates in the model
- Survival simulation results
  - Estimates close to true parameters
  - Type I error performed well for N $\geq$ 200
- More work has been done and published for Normal, Binary, and count outcomes with covariates
- Limitation
  - Sensitivity to distribution misspecification
  - Proportion of participants choosing one treatment vs another, pilot study

RTI INTERNATIONAL

Thank you

RTI INTERNATIONAL


R. Chahine & I. Aban (2024). Likelihood Based Inferences for Trials Incorporating Participant?s Treatment Choice. In *Contemporary Clinical Trials Communicatins*, DOI: 10.1016/j.conctc.2024.101306


R. Chahine & I. Aban (2023). Analysis of survival outcomes using likelihood ratio test in trials incorporating patient's treatment choice. In *Applied Statistics*, DOI: 10.1080/02664763.2023.2199177


B. Cameron, M. J. Kane, & D. Esserman (2020). Preference: an r package for two-stage clinical trial design accounting for patient preference. In: *Journal of Statistical Software* 94, pp. 1?16. DOI: 10.18637/jss.v094.c02.


B. Cameron, P. Peduzzi, & D. Esserman (2018). Extensions to the two-stage randomized trial design for testing treatment, self-selection, and treatment preference effects to binary outcomes. In: *Statistics in Medicine* 37.22, pp. 3147?3178. DOI: 10.1002/sim.7830.


G. Casella & R. Berger (2002). *Statistical Inference*. Second Edi. Belmont, CA: Brooks/Cole Cengage Learning. ISBN: 978-0534243128.


M. R. Janevic, N. K. Janz, J. A. Dodge, X. Lin, W. Pan, B. R. Sinco, & N. M. Clark (2003). The role of choice in health education intervention trials: A review and case study. In: *Social Science and Medicine* 56.7, pp. 1581?1594. DOI: 10.1016/S0277-9536(02)00158-2.


Q. Long, R. J. Little, & X. Lin (2008). Causal inference in hybrid intervention trials involving treatment choice. In: *Journal of the American Statistical Association* 103.482, pp. 474?484. ISSN: 01621459. DOI: 10.1198/016214507000000662.


R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL:


G. Rücker (1989). A two stage trial design for testing treatment, self-selection and treatment reference effects. In: *Statistics in Medicine* 8.4, pp. 477?485.


Y. Shi, B. Cameron, X. Gu, M. Kane, P. Peduzzi, & D. A. Esserman (2020). Two-stage randomized trial design for testing treatment, preference, and self-selection effects for count outcomes. In: *Statistics in Medicine* 39.25, pp. 3653?3683. DOI: 10.1002/sim.8686.


R. M. Turner, S. D. Walter, P. Macaskill, K. J. McCaffery, & L. Irwig (2014). Sample size and power when designing a randomized trial for the estimation of treatment, selection, and preference effects. In: *Medical Decision Making* 34.6, pp. 711?719. ISSN: 1552681X. DOI: 10.1177/0272989X14525264.


S. D. Walter, R. M. Turner, P. Macaskill, K. J. McCaffery, & L. Irwig (2017). Estimation of treatment preference effects in clinical trials when some participants are indifferent to treatment choice. In: *BMC Medical Research Methodology* 17.1, pp. 1?10. DOI: 10.1186/s12874-017-0304-x.


M. Zelen (1979). A New Design for Randomized Clinical Trials. In: 300.22, pp. 1242?1245. DOI: 10.1056/nejm197905313002203.

RTI INTERNATIONAL

PRPP-SMART Design and Analyses

Kelley M Kidwell, PhD
University of Michigan
SCT
May 22, 2024

DISCLOSURES

This work was supported by a Patient Centered Outcomes Research Initiative Improving Methods contract.

Disclaimer: All statements in this presentation, including its findings and conclusions, are solely those of the author and do not necessarily represent the views of PCORI or its Board of Governors, or the Methodology Committee.

I am a part-owner of www.smart-workshops.com

No other relevant disclosures.

PRPP-SMART MOTIVATION

- RCTs are the gold standard for treatment comparisons but they **exclude or ignore participants with treatment preference**
- When testing **different treatment modalities** (e.g. behavioral, medication, surgery) in one trial many participants will have treatment preferences
- Trial designs that do not allow for these preferences can **suffer in accrual, adherence, retention, and external validity of results**
- Many diseases/disorders require **sequential, tailored treatment**
- Motivated by BACPAC trial: a SMART investigating treatment for chronic lower back pain

HOW PROVIDERS TREAT

- Ongoing care and follow up
- Therapies are not set in stone
- Therapies can be changed, intensified, discontinued
- Treatment decisions can be based on health progress, treatment adherence, side effects, and patient choice
- Follow-up therapy based on experience, guidelines, clinical trials

- **Dynamic Treatment Regimen (DTR)** aka adaptive intervention
  - Sequence of **individually tailored decision rules** that specify whether, how and/or when to alter the intensity, type, dose or delivery of treatment at critical decision points in the course of care
  - Guide/Formula for treatment
  - Evidence-based

WHEN AND HOW TO CONSTRUCT DTRS

DTRs are useful when

- Waxing and waning of disease/disorder
- No widely effective treatment
- Treatments may be costly or burdensome
- Adherence problems
- Within and between person heterogeneity

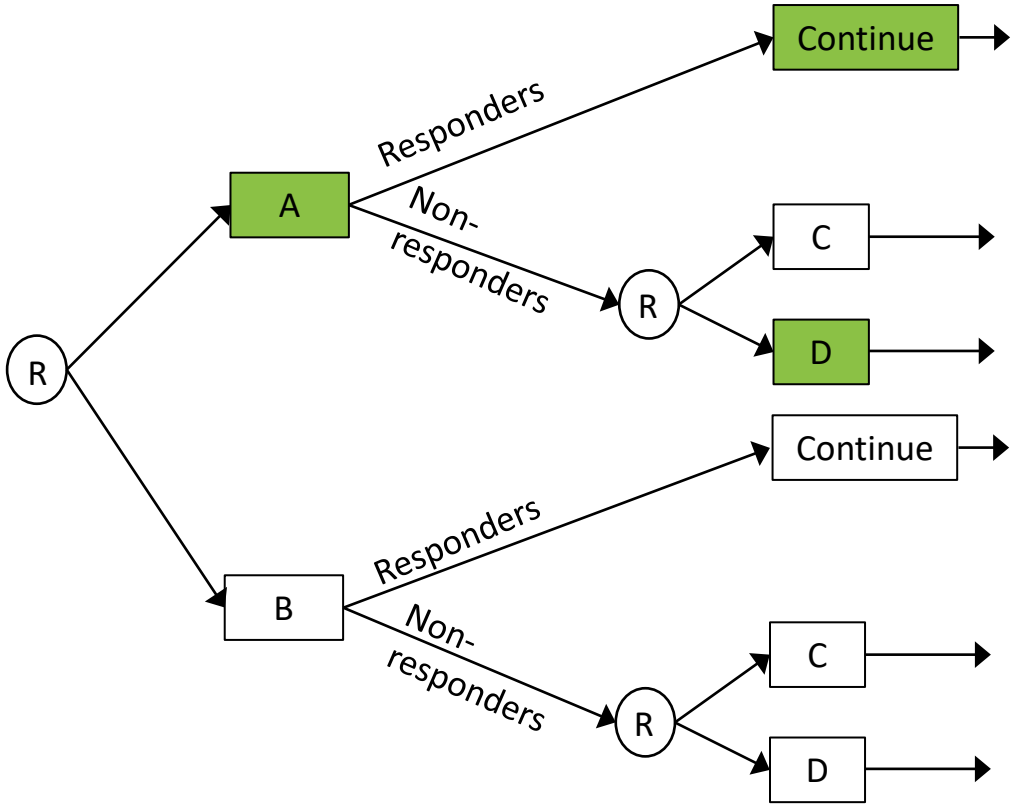
Questions to develop DTR

1. What is the best **first-line** intervention?
2. What is the best **measure of response** to see if the intervention is successful?
3. When is the best **time to measure response** to the initial intervention?
4. What is the best **subsequent** treatment among **non-responders**?
5. What is the best **subsequent** treatment among **responders**?

SMART

- Sequential, multiple assignment, randomized trial
- A type of **multi-stage randomized design**
- Trial participants are randomized to a set of treatment options at critical decision points over the course of treatment
- **All individuals** participate in all stages of the trial
- Subsequent randomization is based on information leading up to that point
- DTRs embedded in design

(A,Continue,C) (B,Continue,C)
(A,Continue,D) (B,Continue,D)



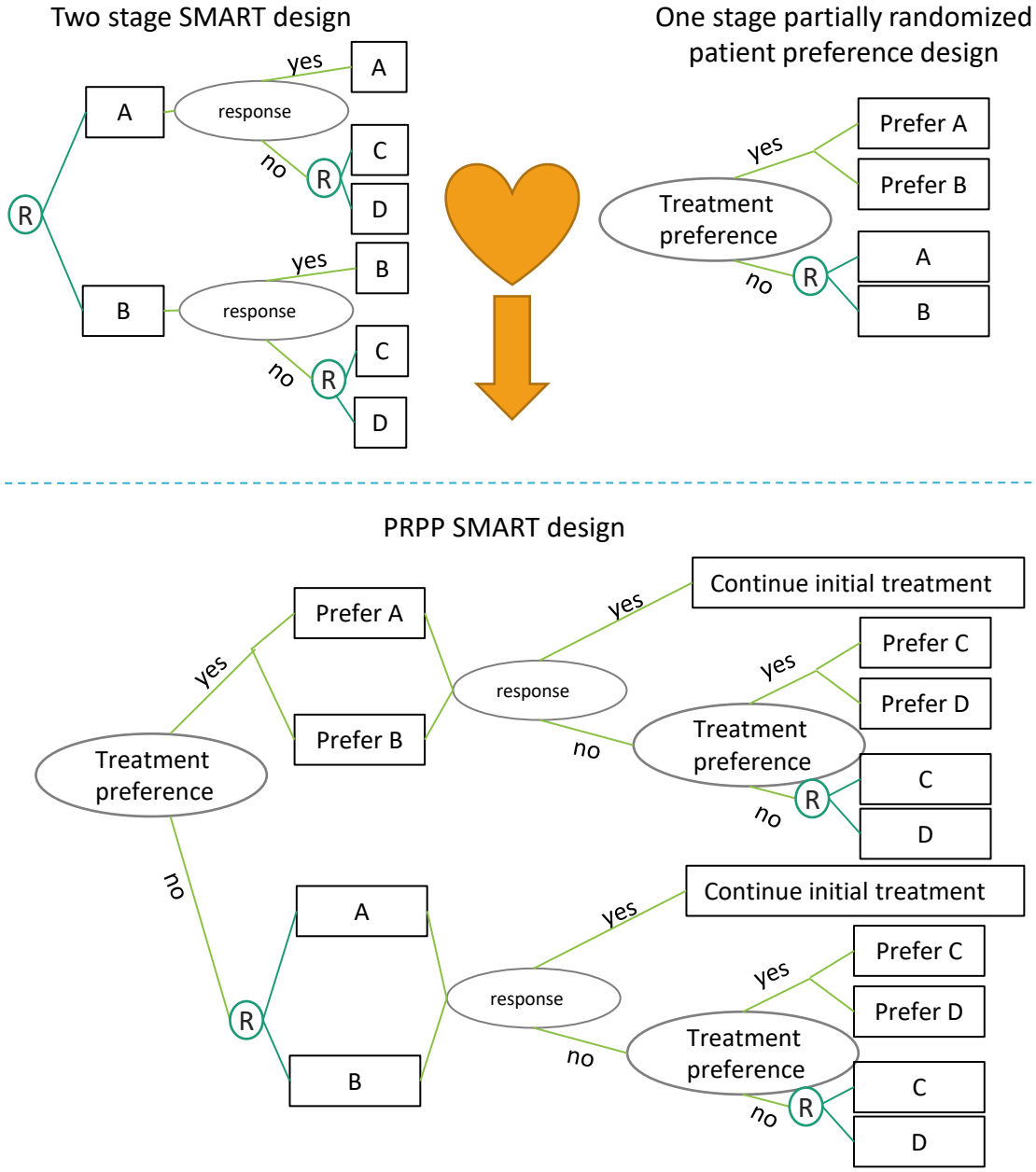
SMARTS IN THE FIELD

- Oncology
- Drug abuse
- ADHD
- Alcoholism
- Obesity
- OCD
- Pain
- Autism
- Depression
- Insomnia
- Bipolar
- Conduct problems
- Smoking cessation
- Suicide prevention

Many settings with sequential treatment of different modalities, where patients are likely to express treatment preferences

DEVELOPMENT OF PRPP-SMART DESIGN

- PCORI Funded Methods Contract
- Design and Methodological Development for a Patient Preference SMART, 2021-2024
  - Summary
  - Mari Wank, Sarah Medley & Tom Braun, Roy Tamura



Advisory Board: Dan Clauw, Dave Williams, Lu Wang, Denise Esserman, Jane Perlmutter, Chris Veasley, Andrea Anderson

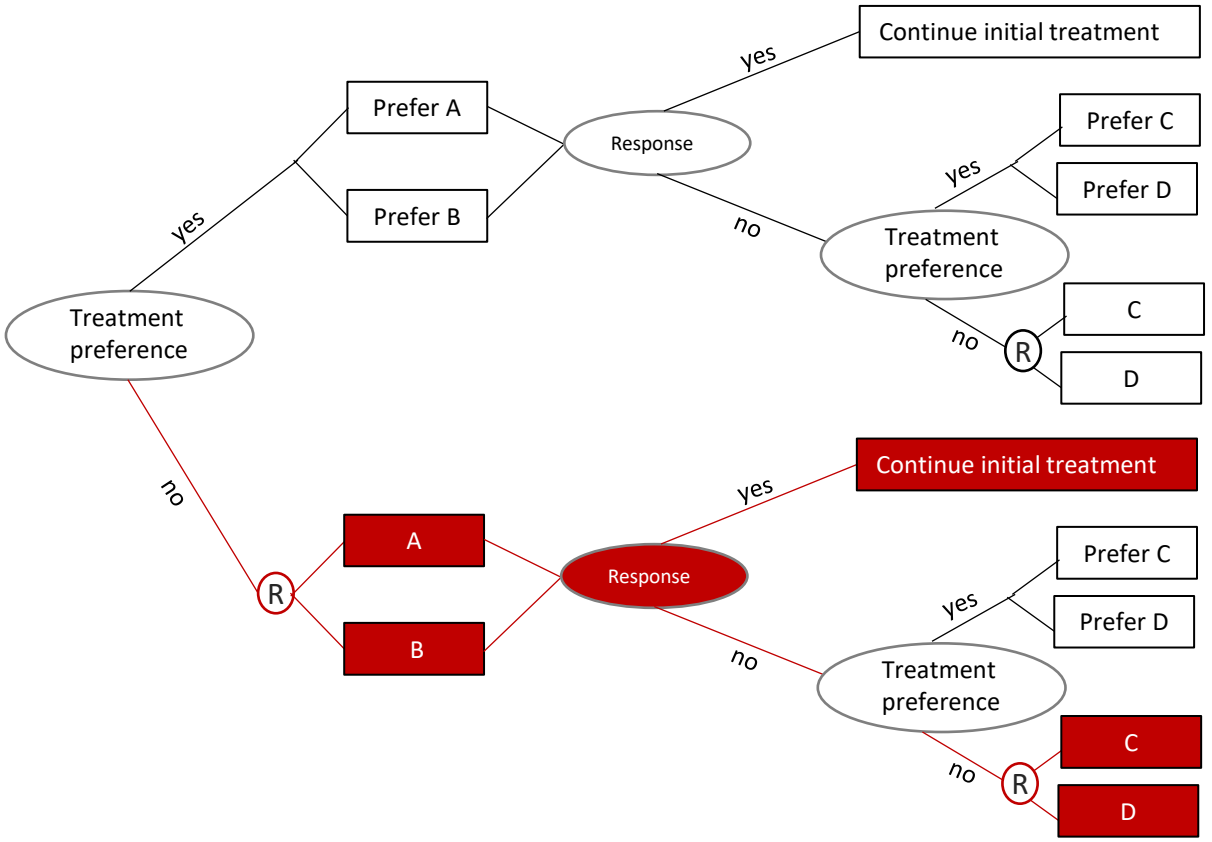
4 EMBEDDED TRIALS

1. It is intuitive to think about a PRPP-SMART as consisting of **four sub-trials**, each with different subgroups of participants
   - 1 - Participants who are always randomized
   - 2 & 3 - Participants who are randomized at one stage
   - 4 - Participants who always have a preference and are never randomized

2. Our methods focus on **combining the data from these four trials together** to produce **efficient** treatment effect estimates with **minimal bias**
   - Frequentist & Bayesian methods
     - Extension of common **SMART weighted & replicated regression**
     - Extension of our rare disease **snSMART Bayesian joint stage models**

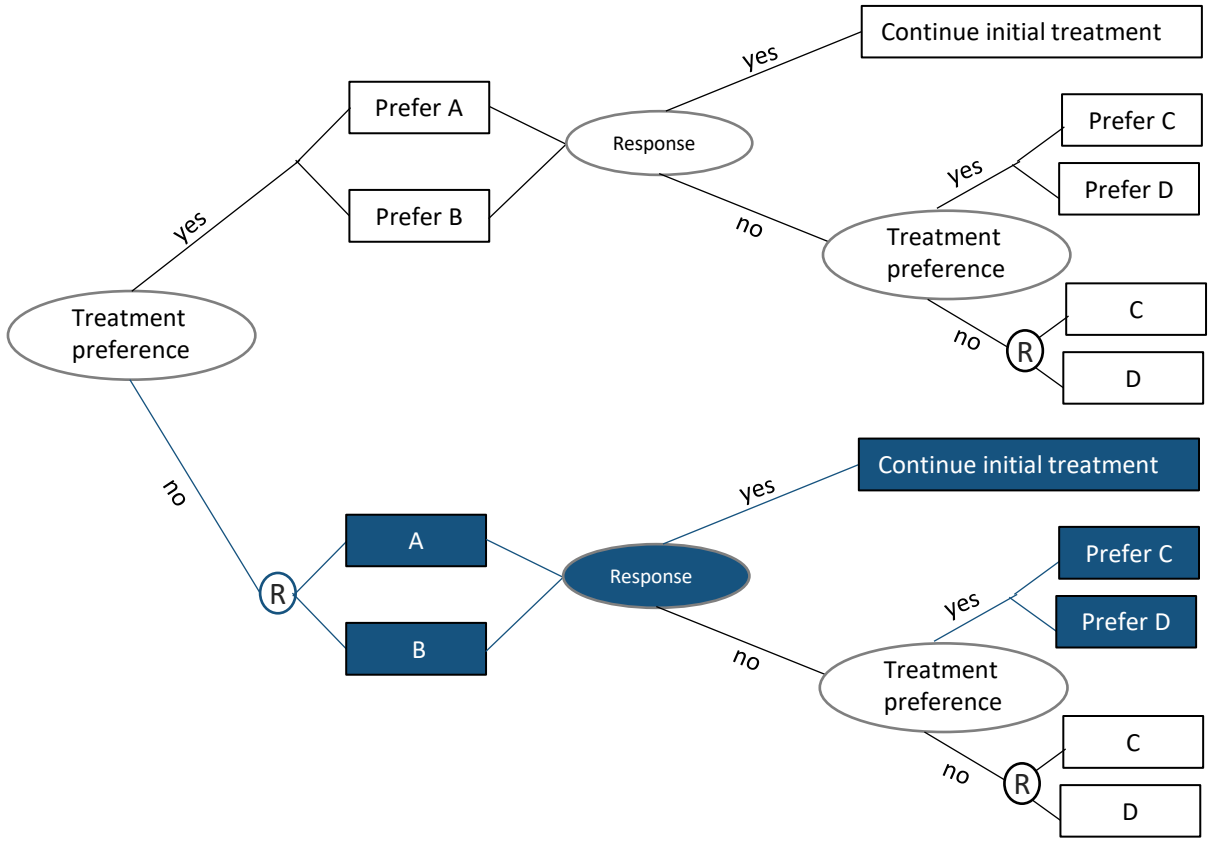
4 EMBEDDED TRIALS: TRIAL 1

Traditional Two-Stage SMART (Randomize-Randomize Trial)



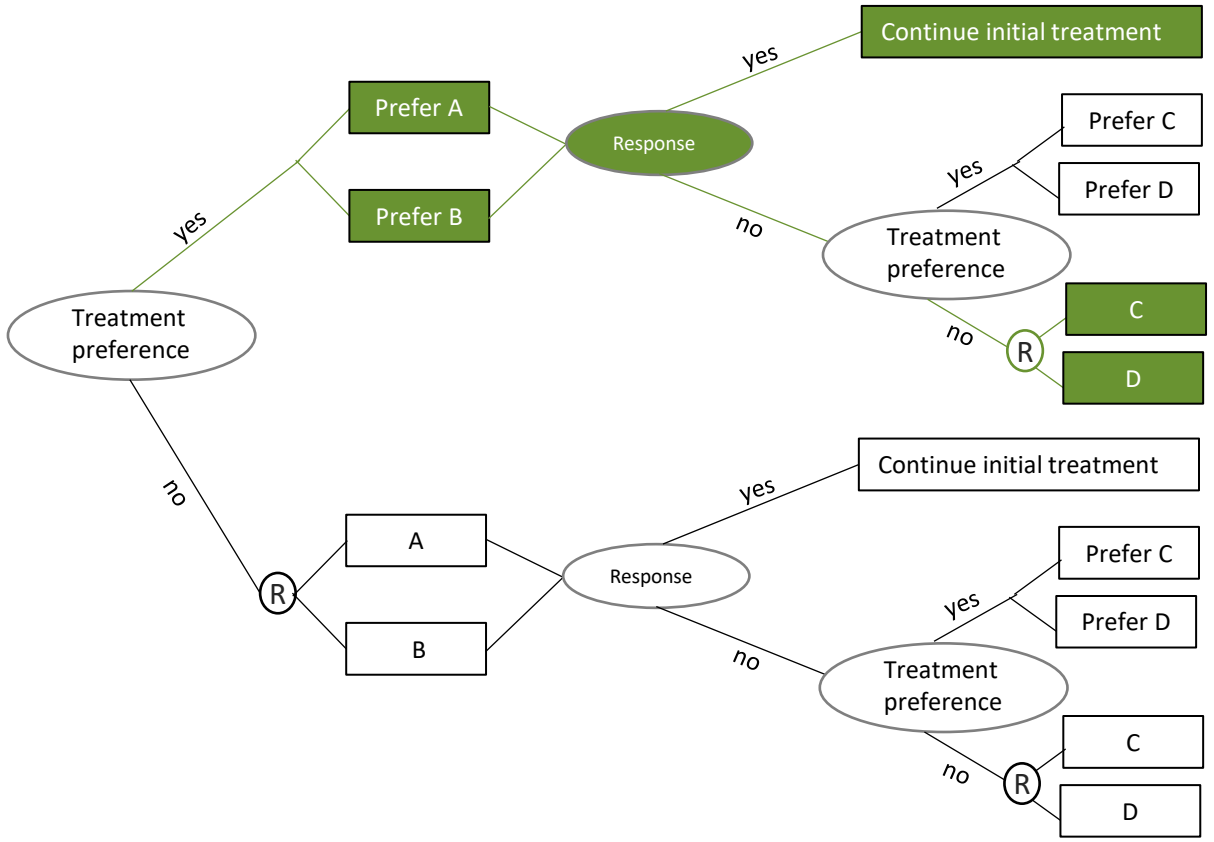
4 EMBEDDED TRIALS: TRIAL 2

Randomized-Preference Trial



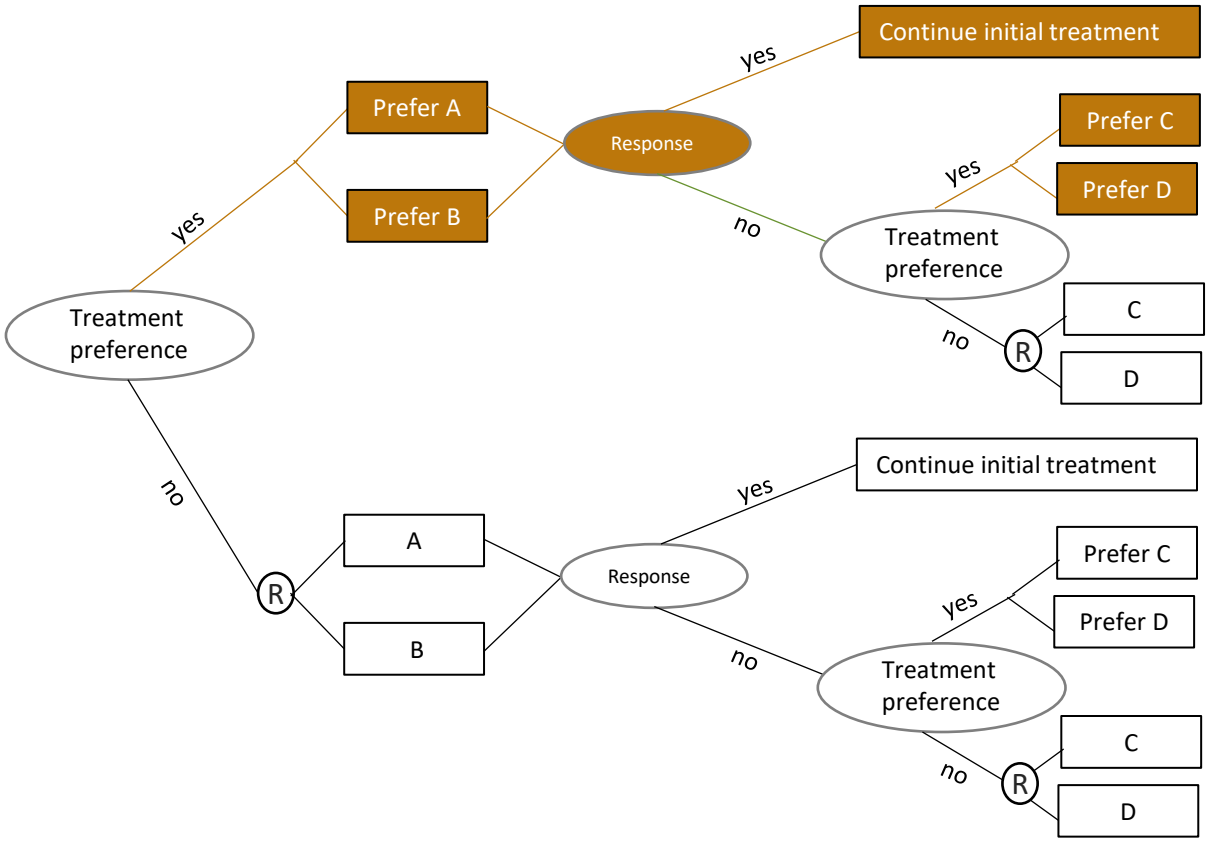
4 EMBEDDED TRIALS: TRIAL 3

Preference-Randomize Trial



4 EMBEDDED TRIALS: TRIAL 4

Preference-Preference Trial



PRPP-SMART DTRS

- DTRs from a PRPP-SMART additionally depend on preference
- Denoted as $[\mathrm{T_1}\,\mathrm{T_{2,R}}\,\mathrm{T_{2,NR}}]_{\mathrm{P_1P_2}}$
  - T_1 and $T_{2,R} \in \{A,B\}$ and $T_2 \in \{C,D\}$ are the first and second stage treatment assignments
  - $P_1 \in \{0,1\}$ and $P_2 \in \{0,1\}$ designate treatment received based on treatment indifference (0) or preference (1)
- AAC_{01} represents the DTR in which participants are indifferent to treatment and receive A in stage 1, with responders continuing to receive A and non-responders who expressed a preference being given their preferred treatment C in stage 2
- In a two-stage PRPP-SMART we have 16 embedded DTRs
  - 4 "standard" SMART DTRs: **Indifference DTRs**: AAC_{00}, AAD_{00}, BBC_{00}, BBD_{00}
  - 12 DTRs involving preference

WEIGHTED AND REPLICATED REGRESSION MODEL

- Weighted and replicated regression model (WRRM) is the most frequently used method to estimate DTRs in a SMART
    - Historically been implemented in a frequentist framework
- WRRM
    - Prevents bias in the estimation of DTRs
    - Allows all embedded DTRs to be estimated simultaneously
- Consists of three components
    - Data replication
    - Data weighting
    - Regression model

PRPP-SMART WRRM

- Data replication
  - Trick to **simultaneously estimate all embedded DTRs** in a PRPP-SMART using standard statistical software
  - **Data for subjects who are consistent with more than one DTR need to be replicated**
  - Responders are consistent with 4 DTRs
    - Ex: Responders to randomized treatment A in stage 1 are consistent with DTRs AAC_{00}, AAC_{01}, AAD_{00}, AAD_{01}
    - We replicate each responder in our data 4 times
    - Fill in the stage 2 (T_2) treatment assignment and stage 2 indicator of whether treatment was assigned via preference or randomization (P_2) so that each responder corresponds to one of the four DTRs it is consistent with

- Data weighting
  - **Imbalance of observations resulting from the PRPP-SMART design**
    - restricted randomization
  - Weight subjects according to their **inverse probability of receiving their own treatment path**

REGRESSION MODEL: BINARY OUTCOME

- We develop Bayesian and frequentist weighted and replicated regression models (WRRMs) which allow for simultaneous estimation of all embedded DTRs within the two-stage PRPP-SMART
- PRPP-SMART regression model (fit on the weighted and replicated data):

$$\text{logit}(\pi^{P_1P_2}_{T_1T_{2,R}T_{2,NR}}) = \alpha_{P_1P_2} + \beta_{P_1}T_1 + \theta_{P_2}T_2 + \gamma T_1T_2$$

Where,

$$\pi^{P_1P_2}_{T_1T_{2,R}T_{2,NR}} = Pr(Y = 1 | T_1T_{2,R}T_{2,NR}P_1P_2)$$

is the probability of response after stage 2 in patients receiving regime $[\mathrm{T_1T_{2,R}T_{2,NR}}]_{\mathrm{P_1P_2}}$

Note: In this model we use **ALL** subjects enrolled in the PRPP-SMART to estimate the probabilities of response for the embedded DTRs

METHOD CONT.

- The previous WRRM implies 4 regression models to estimate the probabilities of response for our 16 DTRs
- These equations represent the regression models that would be used to estimate the DTRs from each sub-trial of a PRPP-SMART
- The first equation is of primary interest as it will estimate the indifference DTRs that would be produced from a SMART

$$\text{logit}(\pi^{00}_{T_1 T_{2,R} T_{2,NR}}) = \alpha_{00} + \beta_0 T_1 + \theta_0 T_2 + \gamma T_1 T_2$$

$$\text{logit}(\pi^{01}_{T_1 T_{2,R} T_{2,NR}}) = \alpha_{01} + \beta_0 T_1 + \theta_1 T_2 + \gamma T_1 T_2$$

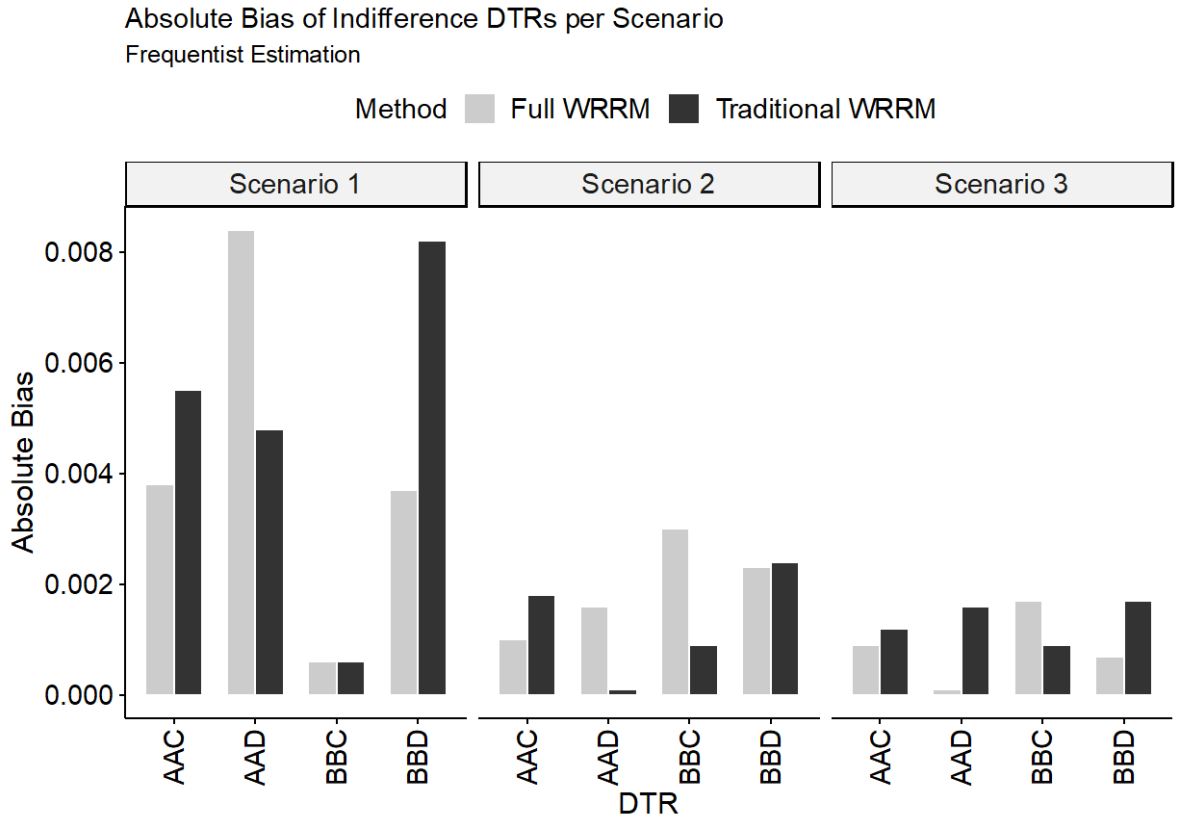
$$\text{logit}(\pi^{10}_{T_1 T_{2,R} T_{2,NR}}) = \alpha_{10} + \beta_1 T_1 + \theta_0 T_2 + \gamma T_1 T_2$$

$$\text{logit}(\pi^{11}_{T_1 T_{2,R} T_{2,NR}}) = \alpha_{11} + \beta_1 T_1 + \theta_1 T_2 + \gamma T_1 T_2$$

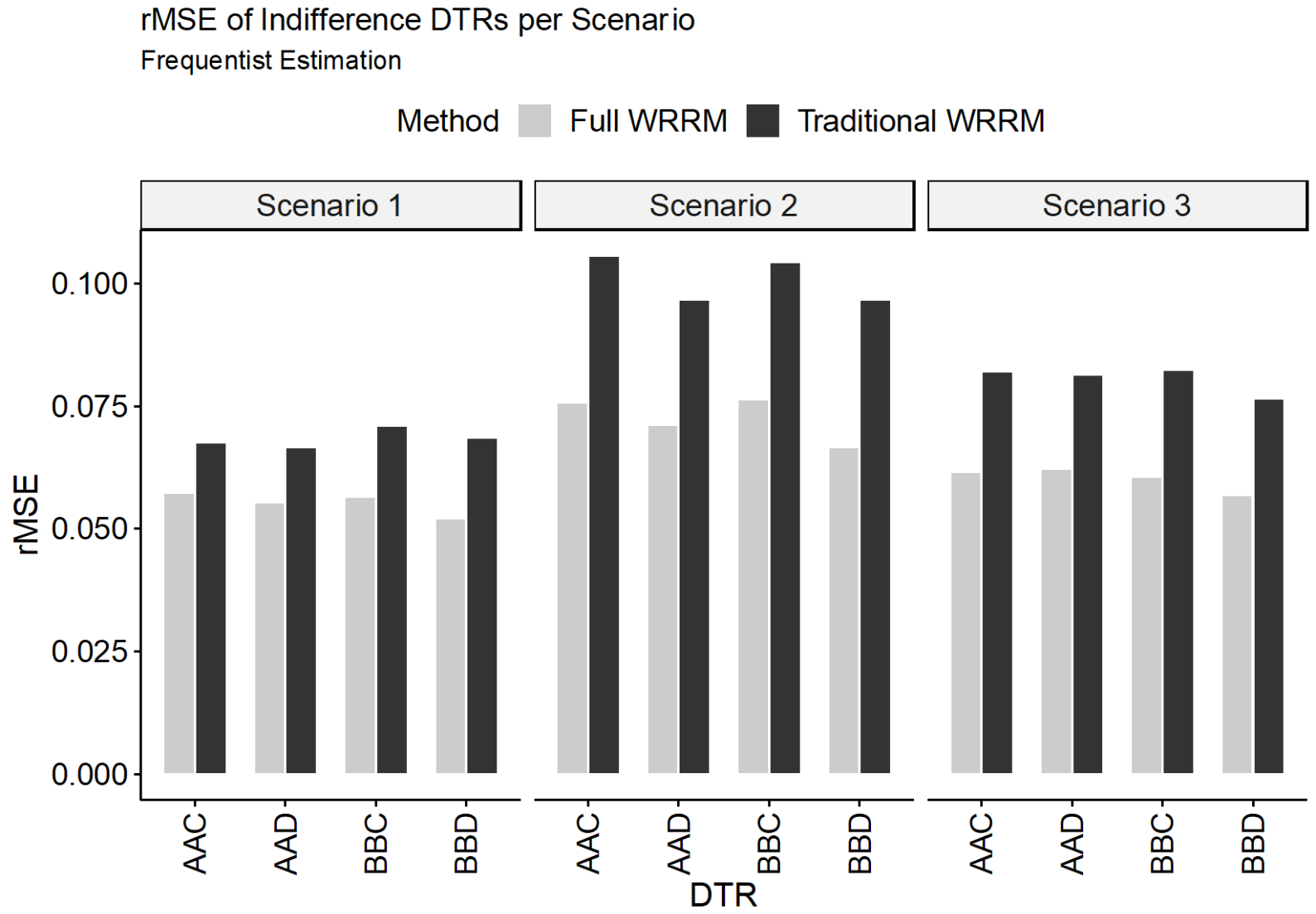
RESULTS: BIAS

- We present results for our frequentist WRRM; Bayesian WRRM produces similar results
- Compare to more traditional PRPP analysis only considering patients randomized to treatment

- "Full WRRM": uses both randomized and non-randomized patients
- "Traditional WRRM": uses only patients randomized to treatment
- Scenario
  - 1: Even split of patients expressing no preference/preference in both stages
  - 2: Majority of patients expressing preference at both stages, and this is equal across stages
  - 3: 50% expressing no preference in stage 1 and 1/3 expressing no preference in stage 2



RESULTS: RMSE



BINARY OUTCOME WRRM CONCLUSIONS

- Our Bayesian and Frequentist WRRMs uses data from **both** randomized and non-randomized subjects for efficient estimation of the DTR effects

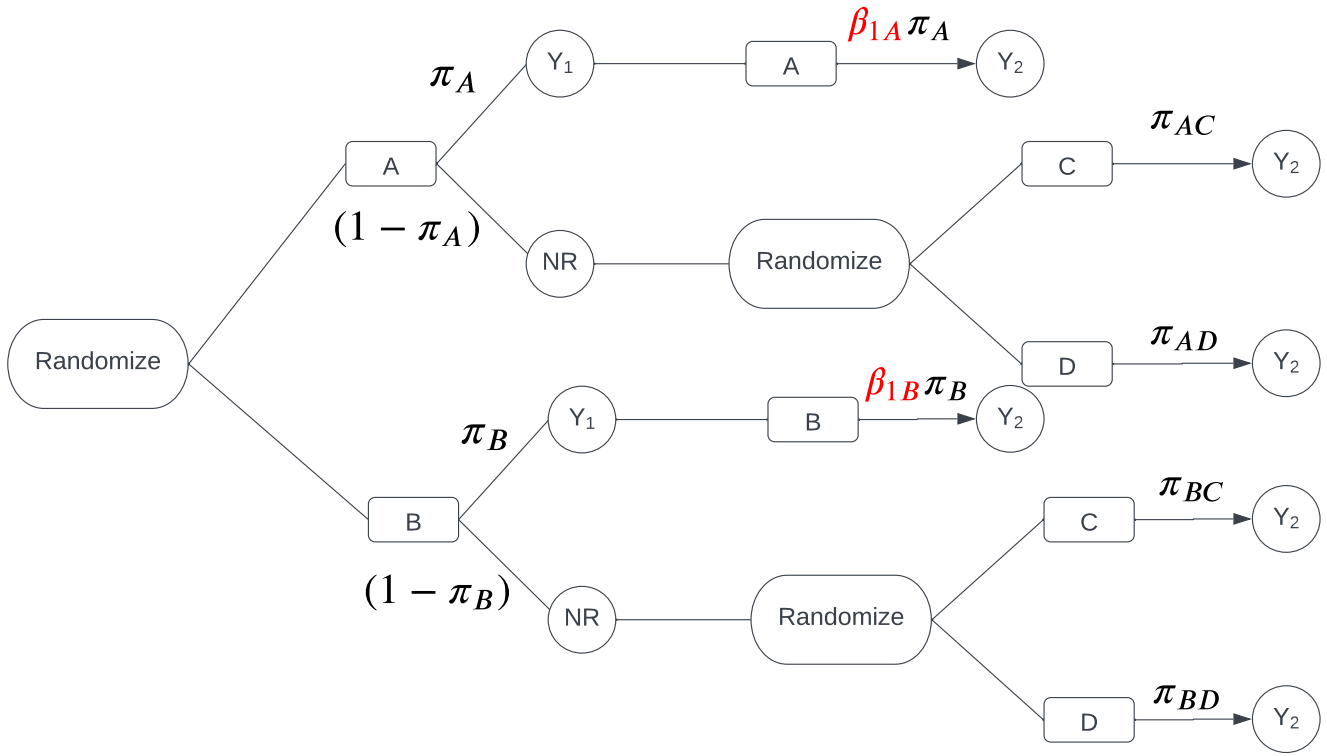
- Our WRRMs…
  - Produces indifference DTR estimates with **minimal bias** despite the inclusion of non-randomized patients in the analysis
  - Results in **more efficient indifference DTR estimates** compared to a traditional analysis
  - Produces indifference DTR estimates with **nominal coverage**

BJSM INTRO

- We propose a Bayesian Joint Stage Model (BJSM) to estimate DTRs and stage-specific main treatment effects from a PRPP-SMART
- In a BJSM, we estimate the DTRs and stage-specific main treatment effects by linking the response rates of our binary outcome variables Y_{is} for randomized and non-randomized subjects via linkage parameters
- Notation:
  - $T_1 \in \{j = A, B\}$ and $T_2 \in \{k = C, D\}$ first and second stage treatment assignments
  - $P_1 \in \{0, 1\}$ and $P_2 \in \{0, 1\}$ designate treatment assignment through preference (1) or randomization (0) in the first and second stages
  - $s = 1, 2$ denote the stage of the trial
  - For subject $i = 1, \ldots, N$, let Y_{is} be a binary outcome collected at the end of stage s
    - Y_{i1}: binary outcome collected at end of stage 1
    - Y_{i2}: binary outcome collected at end of stage 2

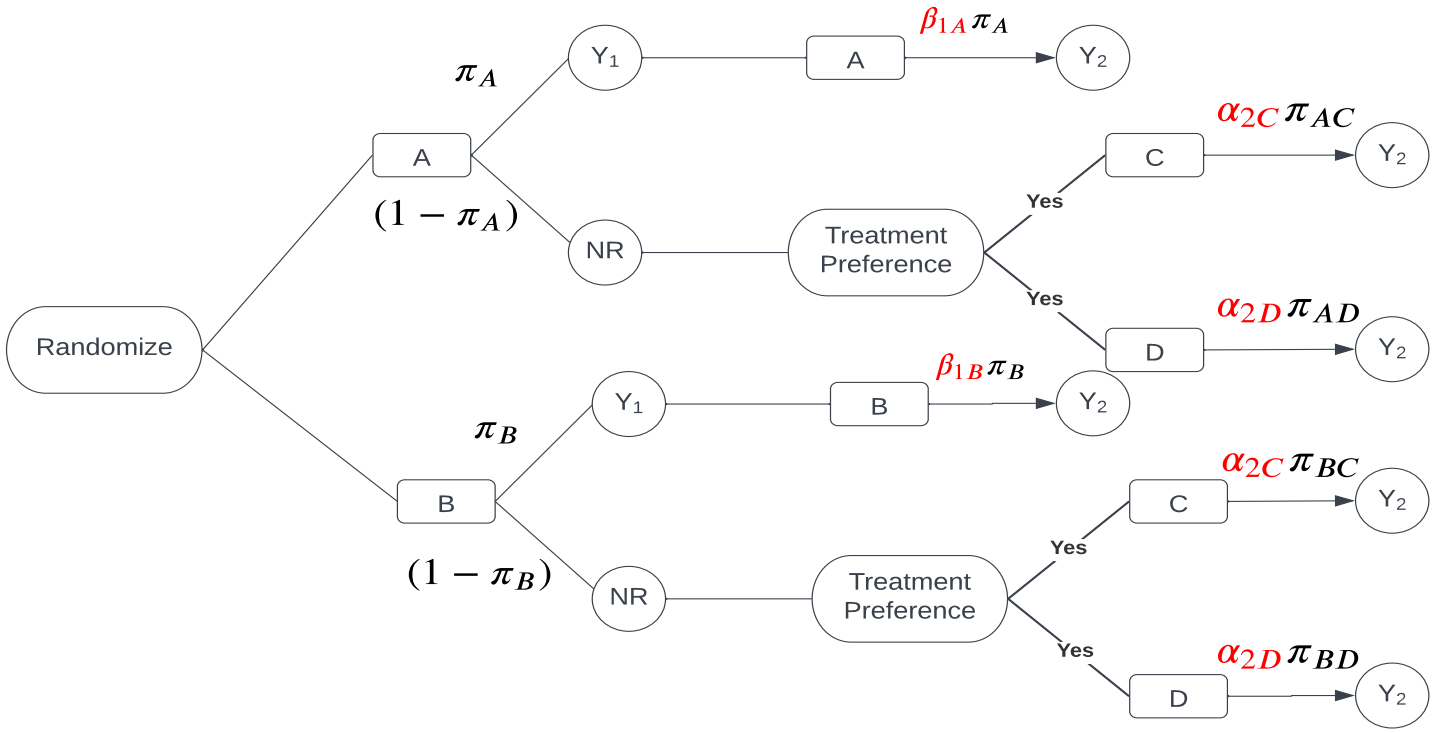
BJSM RESPONSE RATES

- It is useful to think about a BJSM in terms of the four sub-trials of a PRPP-SMART
- Red parameters denote linkage parameters
- Traditional two-stage SMART



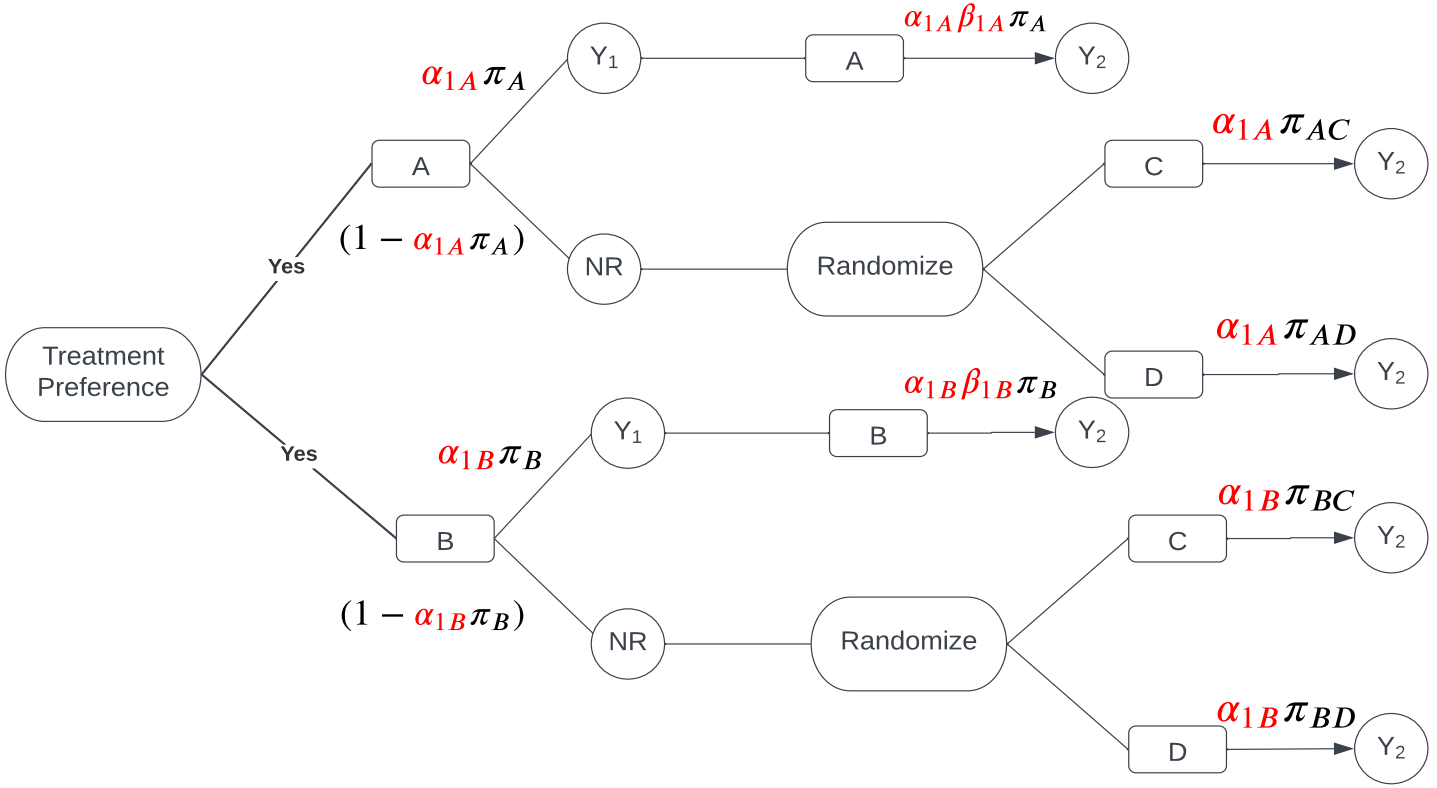
BJSM RESPONSE RATES

- It is useful to think about a BJSM in terms of the four sub-trials of a PRPP-SMART
- Red parameters denote linkage parameters
- **Randomize-Preference Trial**



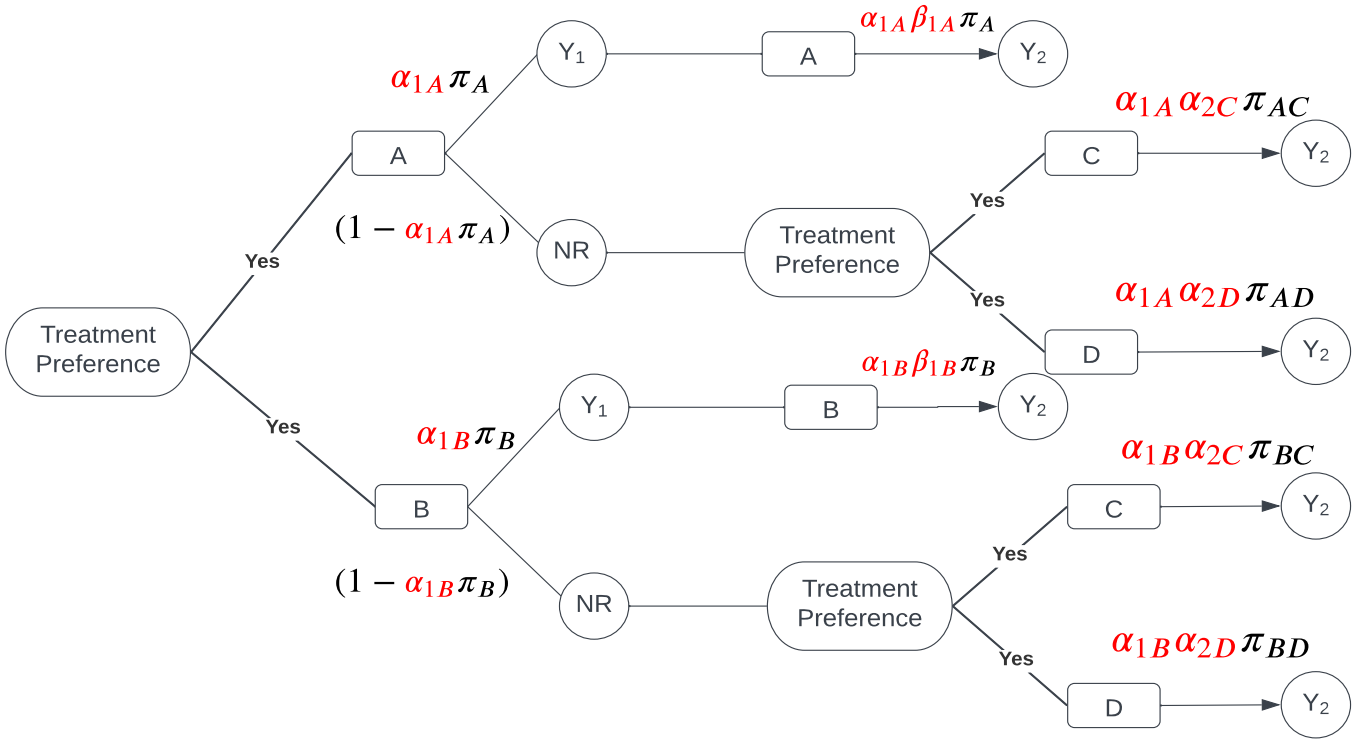
BJSM RESPONSE RATES

- It is useful to think about a BJSM in terms of the four sub-trials of a PRPP-SMART
- Red parameters denote linkage parameters
- **Preference-Randomize Trial**



BJSM RESPONSE RATES

- It is useful to think about a BJSM in terms of the four sub-trials of a PRPP-SMART
- Red parameters denote linkage parameters
- **Preference-Preference Trial**



BJSM

- Define the BJSM as follows:
  - First and second stage response distributions

$$Y_{i1}|\pi_j, \alpha_{1j}, P_{1i} \sim \text{Bernoulli}\Big((\pi_j\alpha_{1j})^{P_{1i}}(\pi_j)^{1-P_{1i}}\Big)$$

$$\begin{aligned} Y_{i2}|Y_{i1}, \pi_j, \pi_{jk}, \beta_{1j}, \alpha_{1j}, \alpha_{2k}, P_{1i}, P_{2i} \sim \text{Bernoulli}\Bigg(& \Big[(\beta_{1j}\pi_j)^{Y_{1i}}(\pi_{jk})^{1-Y_{1i}}\Big]^{(1-P_{1i})(1-P_{2i})} \\ & \Big[(\beta_{1j}\pi_j)^{Y_{1i}}(\alpha_{2k}\pi_{jk})^{1-Y_{1i}}\Big]^{(1-P_{1i})P_{2i}} \\ & \Big[(\alpha_{1j}\beta_{1j}\pi_j)^{Y_{1i}}(\alpha_{1j}\pi_{jk})^{1-Y_{1i}}\Big]^{P_{1i}(1-P_{2i})} \\ & \Big[(\alpha_{1j}\beta_{1j}\pi_j)^{Y_{1i}}(\alpha_{1j}\alpha_{2k}\pi_{jk})^{1-Y_{1i}}\Big]^{P_{1i}P_{2i}}\Bigg) \end{aligned}$$

BJSM PRIORS

- Specify the following priors
- Prior hyperparameters
  - $\theta_1 = \delta_1 = 1/3$
    - Gives a prior mean of 0.5
  - $\theta_2 = \delta_2 = 1/2$
    - Gives a prior mean of 1
    - First stage responders who receive the same treatment in the second stage are assumed, on average, to have the same response rate in the first and second stages
  - $\theta_3 = \delta_3 = 1/2$
    - Gives a prior mean of 1
    - Consider initially that there is no effect of preference but add substantial variance to this prior to allow for the possibility of preference effects
    - Believe there will not be extreme preference effects in a PRPP-SMART

$$\pi_j \sim \text{Beta}(\theta_1, \delta_1)$$

$$\pi_{jk} \sim \text{Beta}(\theta_1, \delta_1)$$

$$\beta_{1j} \sim \text{Gamma}(\theta_2, \delta_2)$$

$$\alpha_{1j} \sim \text{Gamma}(\theta_3, \delta_3)$$

$$\alpha_{2k} \sim \text{Gamma}(\theta_3, \delta_3)$$

BJSM DTRS

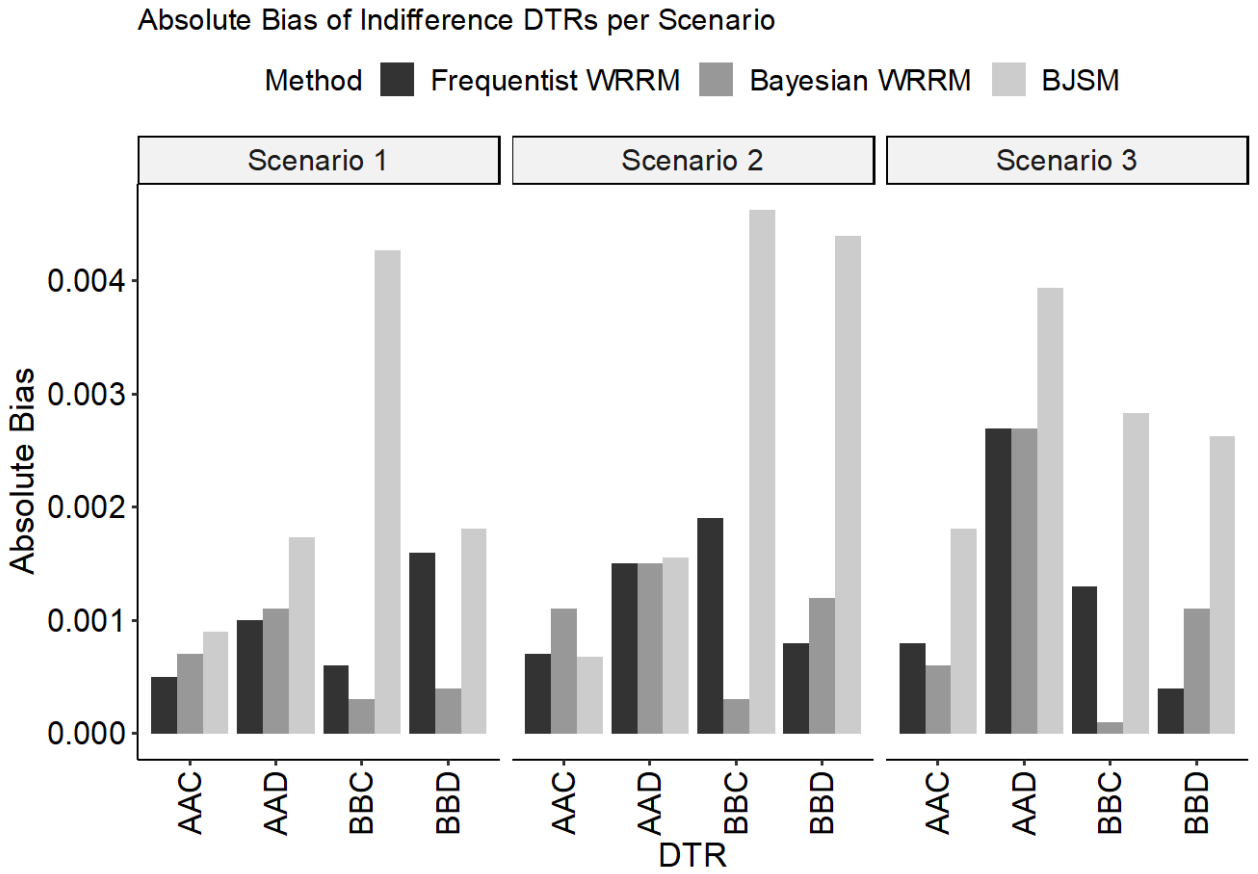
- Want to estimate the second stage outcome or response rates of the 16 embedded DTRs
  - $\pi^{P_1P_2}_{T_1T_{2,R}T_{2,NR}} = \Pr(Y_2 = 1 | T_1T_{2,R}T_{2,NR}P_1P_2)$
- First compute the posterior draws of the BJSM parameters using MCMC: π_j, π_{jk}, β_{1j}, α_{1j}, α_{2k}
- Then compute the posterior draws for each DTR $\pi^{P_1P_2}_{T_1T_{2,R}T_{2,NR}}$ via the following equation
  - Estimate the DTR effects via posterior means
  - Credible intervals are obtained by looking at the quantiles of the posterior draws ($q_{0.25}$, $q_{97.5}$)

$$\pi^{P_1P_2}_{T_1T_{2,R}T_{2,NR}} = \Big(\Big[\pi_j(\beta_{1j}\pi_j) + (1-\pi_j)(\pi_{jk}) \Big]^{(1-P_{1i})(1-P_{2i})} \Big[\pi_j(\beta_{1j}\pi_j) + (1-\pi_j)(\alpha_{2k}\pi_{jk}) \Big]^{(1-P_{1i})P_{2i}} \Big[\alpha_{1j}\pi_j(\alpha_{1j}\beta_{1j}\pi_j) + (1-\alpha_{1j}\pi_j)(\alpha_{1j}\pi_{jk}) \Big]^{P_{1i}(1-P_{2i})} \Big[\alpha_{1j}\pi_j(\alpha_{1j}\beta_{1j}\pi_j) + (1-\alpha_{1j}\pi_j)(\alpha_{1j}\alpha_{2k}\pi_{jk}) \Big]^{P_{1i}P_{2i}} \Big)$$

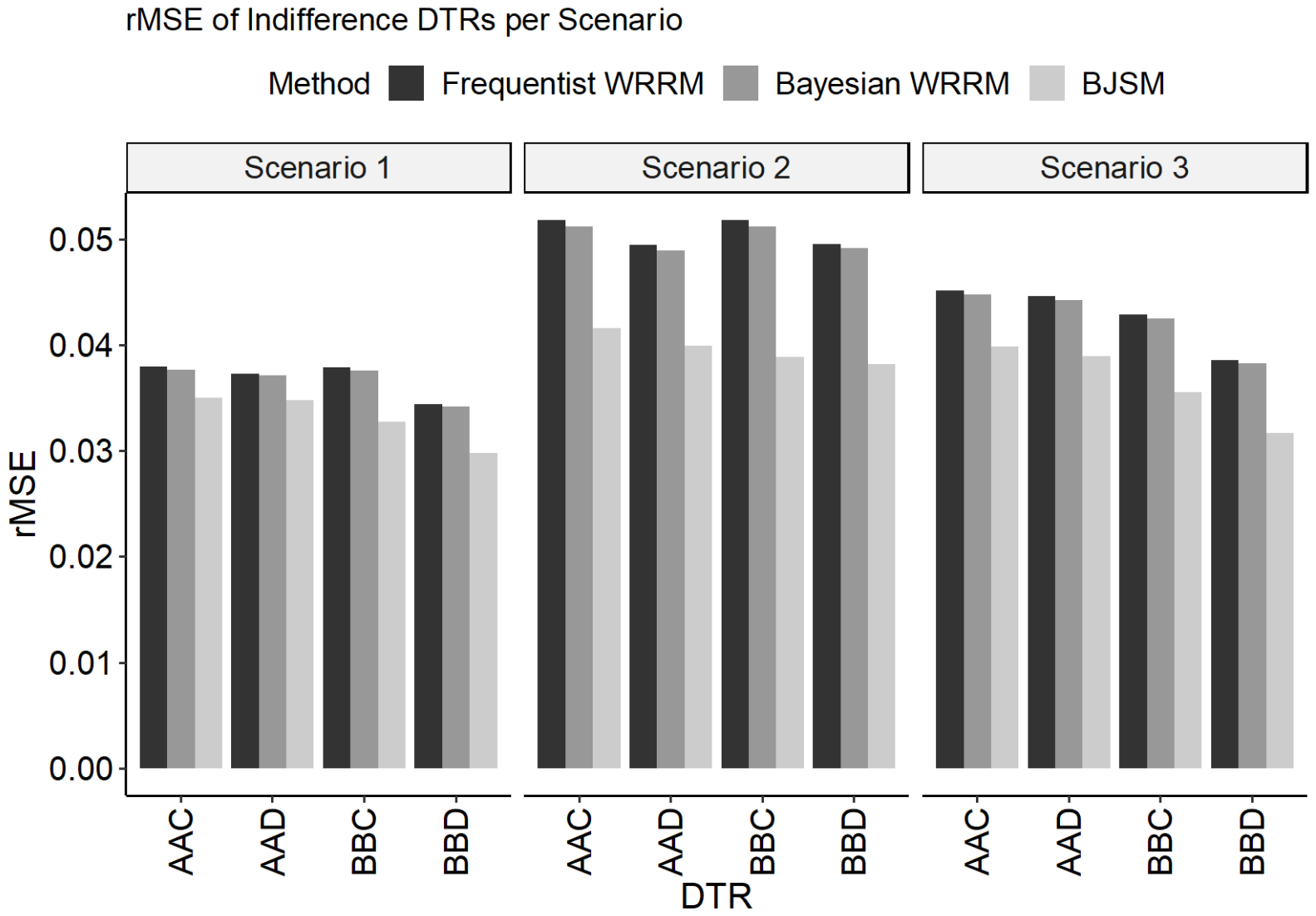
RESULTS: BIAS

- We compare BJSM to the frequentist and Bayesian WRRMs

- Scenario
  - 1: Even split of participants expressing indifference/preference in both stages
  - 2: Majority of participants expressing preference at both stages, and this is equal across stages
  - 3: 50% expressing indifference in stage 1 and 1/3 expressing indifference in stage 2



RESULTS: RMSE



BJSM CONCLUSIONS

- BJSM produces indifference DTR estimates with **minimal bias**
- BJSM produces **more efficient DTR estimates** compared to a WRRM
  - WRRM solely utilizes second-stage outcomes to estimate DTR response rates
  - **BJSM uses both first and second-stage outcomes**
    - Capitalizes on the correlation between repeated outcomes to enhance the efficiency of our estimates
- **Nominal coverage** is achieved for all indifference DTRs

CURRENT/FUTURE WORK

1. Sample size calculations for PRPP-SMART
2. A Bayesian dynamic borrowing analysis method for continuous outcomes
3. R Software and applets to facilitate dissemination to statisticians
4. Greater dissemination to clinical/scientific investigators

PRPP-SMART SUMMARY

- PRPP-SMART design may be considered for more **pragmatic, effectiveness trials**.
  - where treatments are of **different modalities**
  - And treatments are given **sequentially, tailored** to response
- A PRPP-SMART will require more participants than a SMART, but may be the only way to answer specific treatment questions
  - **If preference is not included the trial could suffer from accrual and retention issues**
- Most people have treatment preference, but all may not express it or be indifferent to some options
  - Inherent **patient phenotypes**
  - Interest may span beyond indifference DTR effects

Thank you!

kidwell@umich.edu

Learn more about SMART design:
https://smart-workshops.com/smart-design-info

EXTRA SLIDES

DTR EXAMPLES

#1: Weight Loss

- For an overweight or obese individual, begin with individual behavioral weight loss treatment (IBT) for 5 weekly sessions.
- If at the end of the 5th session, the individual has lost 5 or more pounds, then continue IBT for 5 more sessions.
- If at the end of the 5th session, the individual has lost <5 pounds, add meal replacements to IBT for 5 sessions

#2: Alcohol Abuse

- First take naltrexone and receive in person medical management
- After 2 weeks, but before 8 weeks, if the individual has 2 or more heavy drinking days then add cognitive behavioral therapy
- After 2 weeks, but before 8 weeks, if the individual has less than 2 heavy drinking days, replace in person medical management with telephone disease management

DYNAMIC TREATMENT REGIMENS

- A DTR is a sequence of **decision rules that outlines how to adapt treatments** for an individual patient over time
- Think of as if-then statements:
  - Start with treatment A
    - IF participant responds to treatment
      - THEN continue A
    - IF participant does not respond to treatment
      - THEN administer C
- $[T_1 T_{2,R} T_{2,NR}]$: T_1 and $T_{2,R} \in \{A,B\}$ and $T_2 \in \{C,D\}$ are the first and second stage treatment assignments
  - AAC represents the DTR where participants receive A in stage 1, with responders continuing to receive A, and non-responders receiving C in stage 2

BENEFITS OF SMART DESIGN

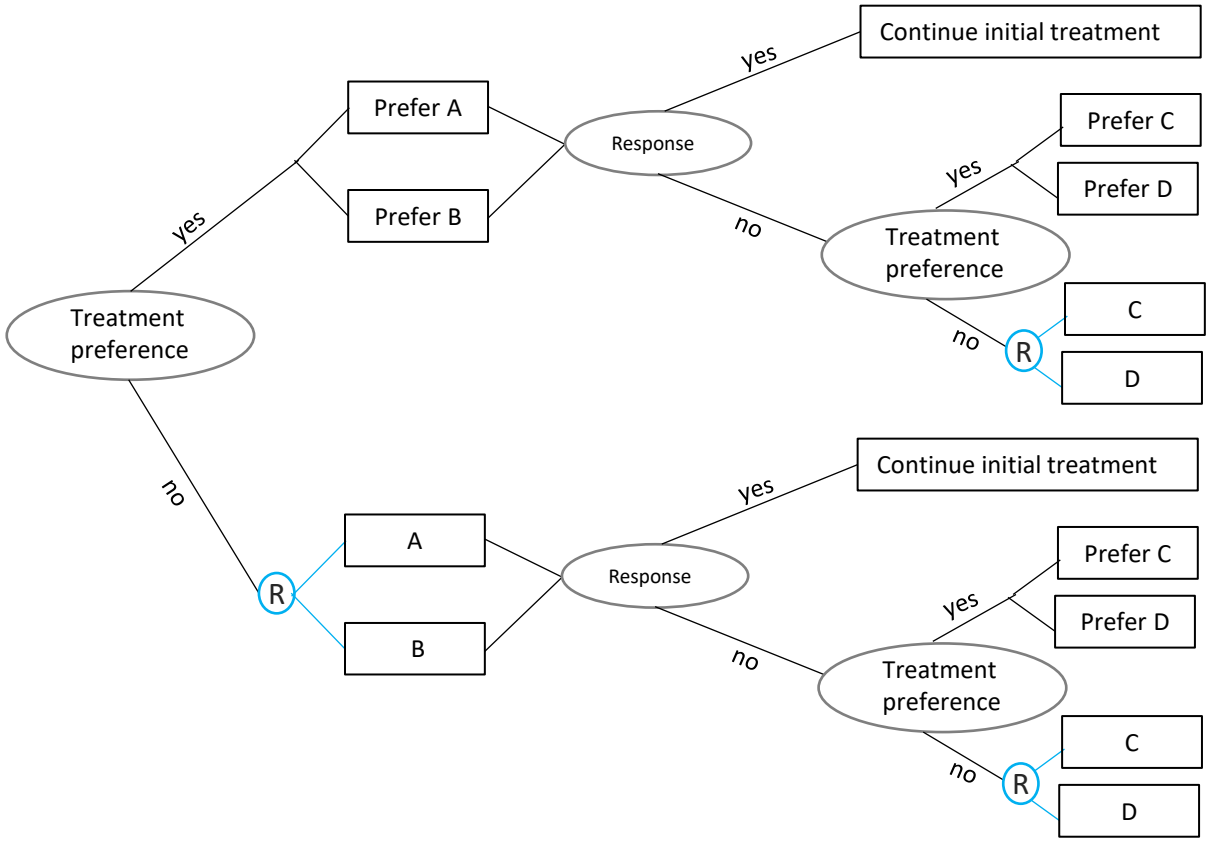
- **Delayed Effects**
  - treatment synergies or antagonisms

- **Prescriptive Effects**
  - initial treatment may elicit symptoms to better match individual to subsequent treatment

- **Sample Selection Effects**
  - individuals who enroll in, remain in or are adherent in a SMART may be different from those in other designs

TWO-STAGE PRPP-SMART DESIGN

A multi-stage trial where, at each stage, patients either receive their preferred treatment, or if they have no preference, they are randomized



Key

- R : Randomization
- A, B : 1st stage treatment options
- C, D: 2nd stage treatment options

REGRESSION MODEL: CONTINUOUS OUTCOME

- We extend the same WRRM to continuous outcomes
- The model can equivalently be represented as

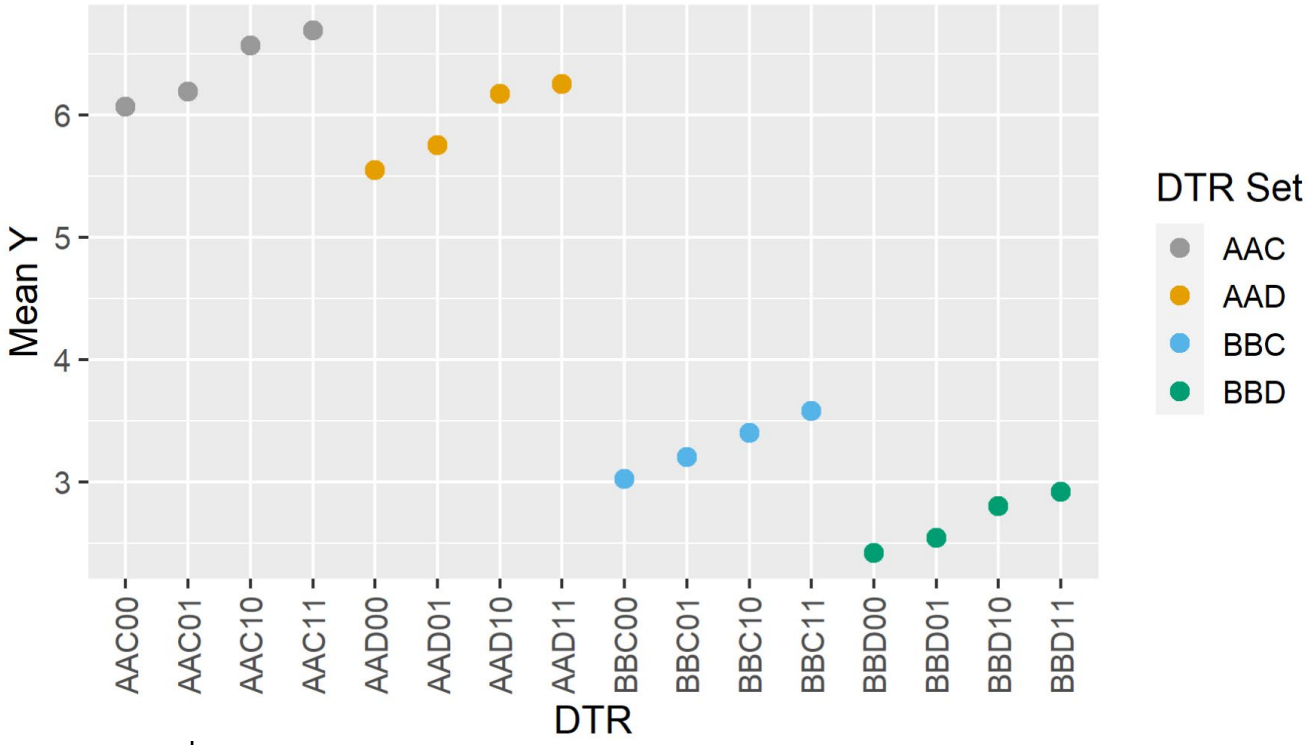
$$E[Y \mid T_1, T_2, P_1, P_2] = \alpha + \beta T_1 + \theta T_2 + \gamma T_1 T_2 + \beta_B P_1 + \beta_A P_1 T_1 + \theta_D P_2 + \theta_C P_2 T_2 + \gamma_P P_1 P_2$$

- We also consider a model with fewer parameters to gain efficiency in DTR estimates (reduced model)

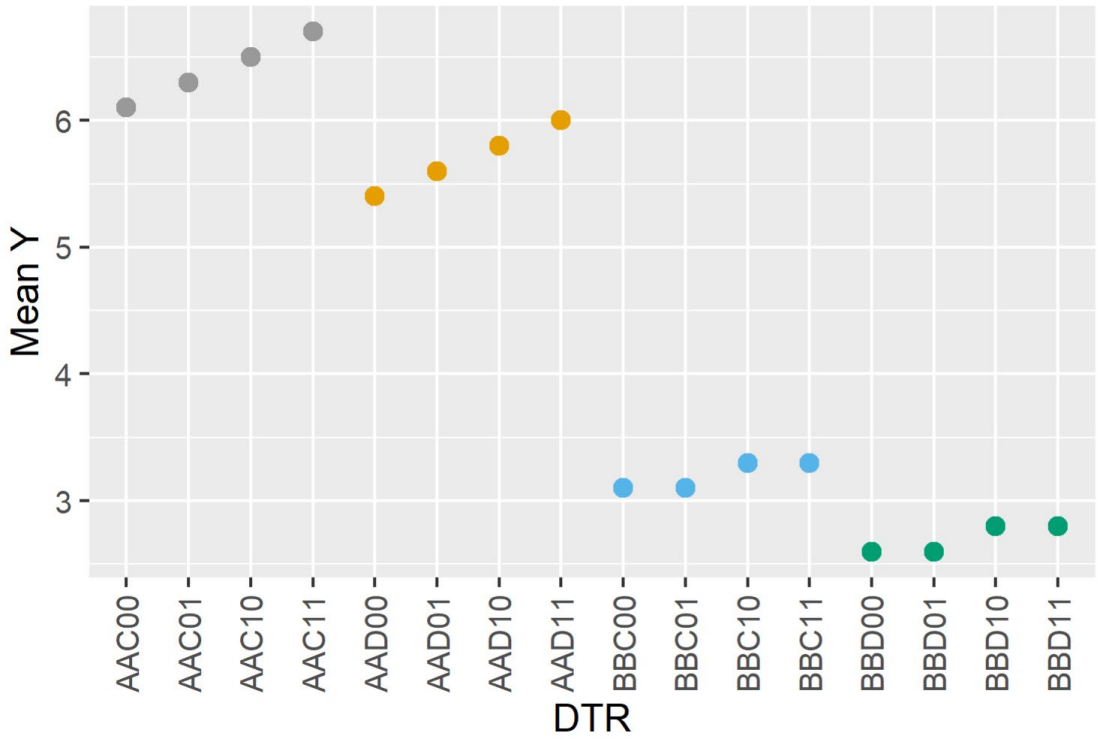
$$E[Y \mid T_1, T_2, P_1, P_2] = \alpha + \beta T_1 + \theta T_2 + \gamma T_1 T_2 + \beta_P P_1 + \theta_P P_2$$

SUB-TRIAL EFFECT TYPES

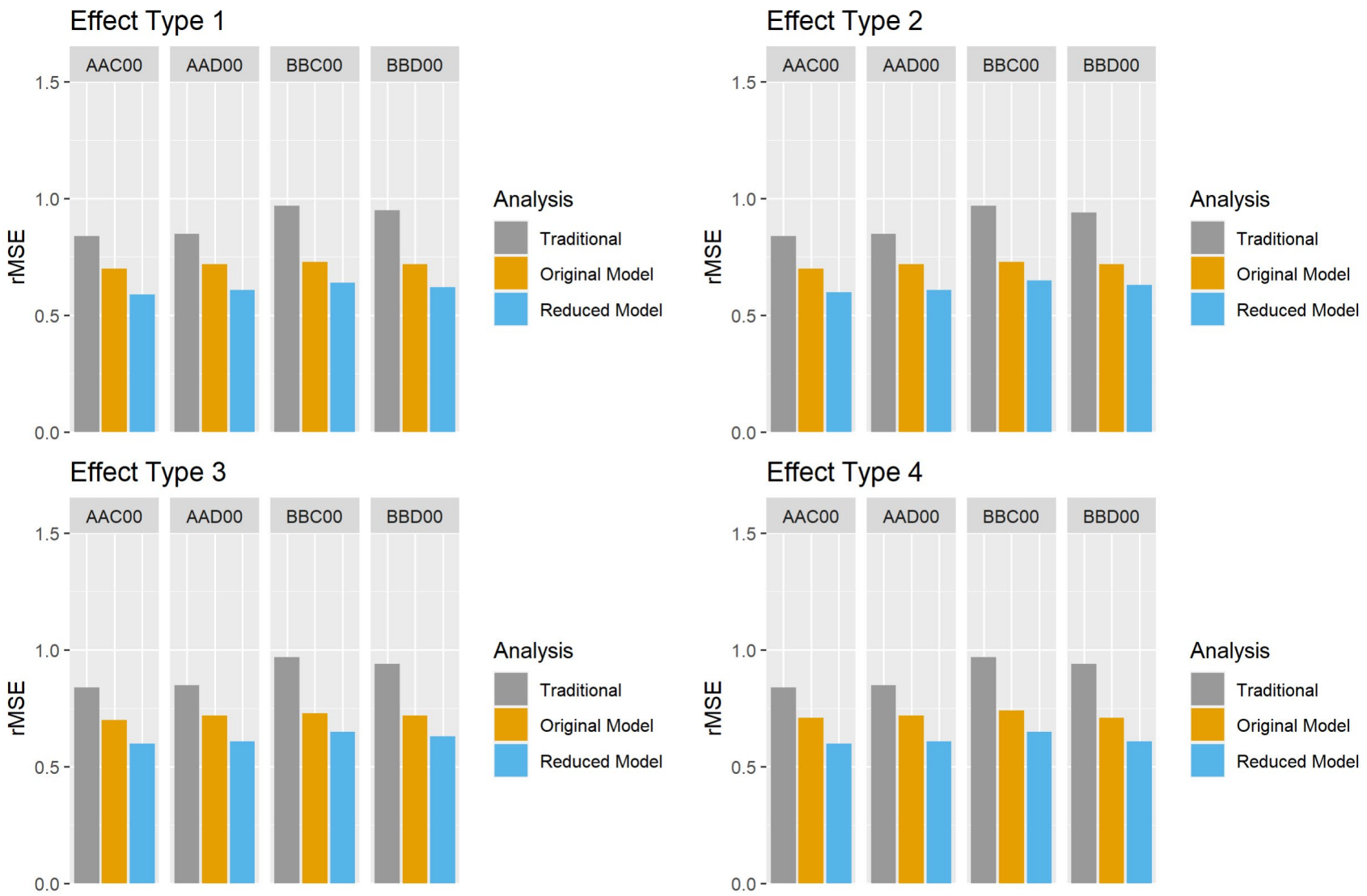
Type 1: Similar relationship between sub-trials in all DTRs



Type 2: Higher outcomes for individuals that preferred A in stage 1



RESULTS: SMALL SUB-TRIAL EFFECTS



RESULTS: SMALL SUB-TRIAL EFFECTS

- When the largest standardized effect size between sub-trials is 0.1 or less, both models have minimal bias with N = 500
- Results are consistent across different sub-trial effect types
- The reduced model has similar levels of bias as the original model
- Standard errors are at least 15% smaller for each DTR in the reduced vs. original model.
  - The rMSE for the reduced model is always smaller

If outcomes are similar across sub-trials, use the reduced model to gain efficiency

RESULTS: LARGE SUB-TRIAL EFFECTS

- When sub-trial effect sizes are moderate, the results differ between sub-trial effect types
- The model fit matters & sub-trial effects cannot be too complex
  - The reduced model is sufficient for sub-trial effect type 1, but does not fit well in the other types
- We observe minimal bias with N = 1,000

If outcomes are different for preference sub-trials, need to correctly model the treatment & preference relationship to avoid biased DTR estimates

CONTINUOUS OUTCOME WRRM CONCLUSIONS

- When sub-trial effects do not exceed treatment effect sizes, our **Frequentist WRRM produces DTR estimates with minimal bias and nominal coverage**
  - PRPP-SMART sample size depends on treatment and sub-trial effect sizes (future work)
  - Method is particularly **robust when sub-trial effects are small**
  - Can **gain efficiency with reduced model** is sub-trial effects are small or the relationship between sub-trials is simple (as in effect type 1)

- Simulations considering various subtrial effects provided a better understanding of the **limitations of our methods**
  - We expect our method is appropriate under common scenarios

BJSM DTRS

- We can break down the previous equation into distinct equations each representing how to estimate the DTRs from the four sub-trials which constitute a PRPP-SMART
- The first equation is of primary interest as it will estimate the indifference DTRs that would be produced from a SMART

$$\pi^{00}_{T_1 T_{2,R} T_{2,NR}} = \pi_j(\beta_{1j}\pi_j) + (1-\pi_j)(\pi_{jk})$$

$$\pi^{01}_{T_1 T_{2,R} T_{2,NR}} = \pi_j(\beta_{1j}\pi_j) + (1-\pi_j)(\alpha_{2k}\pi_{jk})$$

$$\pi^{10}_{T_1 T_{2,R} T_{2,NR}} = \alpha_{1j}\pi_j(\alpha_{1j}\beta_{1j}\pi_j) + (1-\alpha_{1j}\pi_j)(\alpha_{1j}\pi_{jk})$$

$$\pi^{11}_{T_1 T_{2,R} T_{2,NR}} = \alpha_{1j}\pi_j(\alpha_{1j}\beta_{1j}\pi_j) + (1-\alpha_{1j}\pi_j)(\alpha_{1j}\alpha_{2k}\pi_{jk})$$