Some aspects of SMART design: Methodological Developments and an Application in Non-inferiority Trial

Samiran Ghosh [1]

Erina Paul [1], Bibhas Chakraborty [2], Alla Sikorski [3]

[1]Wayne State University, School of Medicine
[2]National University of Singapore
[3]Michigan State University

Department of Family Medicine and Public Health Sciences

CENTER FOR MOLECULAR MEDICINE AND GENETICS

WAYNE STATE UNIVERSITY

Funding Sources

- ME-1409-21410 a PCORI contract as a PI
- P30-ES020957 a NIEHS Center grant (Co-I)
- HD-089875-01 a U19 NIH grant on HIV Prevention (Co-I)
- R01 HD089000-NICHD grant (Co-I)
- R01 NS079429 NINDS grant (Co-I)
- R21 HS024750 an AHRQ grant (Co-I)

Outline

Efficacy Seeking Experimental Design

- Efficacy (de-jure) is the capacity for beneficial change (or therapeutic effect) of a given intervention under Idealized condition (or as directed).
- To determine whether a drug or an intervention is efficacious we often proceed via Randomized Controlled Trial (RCT).
- RCT is based on three basic principals of Design of Experiment, 1. Randomize, 2. Replicate and 3. Blocking.
- Note, RCT involves human subjects.
- Prospective, controlled, experiment under strict inclusion-exclusion criteria (hence the term **Idealized**).
- Note also, to declare something as efficacious we need a "Control" group.

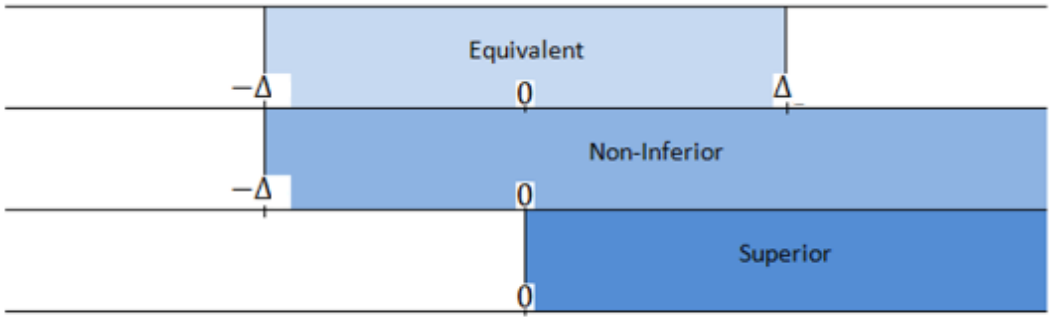
RCT

- Randomized controlled clinical trials (RCTs) are an indispensable source of information about efficacy of treatments in almost any disease area.
- RCTs place a strong emphasis on internal validly with randomization, replication, double-blinding, and control or comparison group/s.
- The goal is to determine whether certain intervention is efficacious compared to a control group.
- In the absence of an effective treatment the usefulness of placebo controlled RCTs are uncontroversial.
- However, in the presence of an established effective regime, placebo controlled RCTs are non-ethical.
- A big question is then what that *Control* group should be!!!

Why Non-Inferiority?

- Implication of placing an Active Comparator arm in RCT is huge.
- This gives rise to Superiority and Non-Inferiority (NI) trial.
- When Superiority of an Intervention is questionable yet the intervention has some desirable features, Non-Inferiority could be an option.
- Such as, it may be less toxic, less invasive, less costly and/or less debilitating, and hence preferable to a sub-group of patients.
- Albeit intervention may be *slightly less* efficacious or inferior within an *acceptable* range.
- NI trials are typically active-control trial (i.e. no placebo).

- It requires a clinically acceptable margin ($\Delta > 0$).



Illustrates the Equivalence, Non-Inferiority and Superiority intervals.

- Choice of this margin is a non-trivial issue.
- ICH, CPMP and FDA regulatory documents provide some general guideline.
- In this talk, we are focusing on **Non-inferiority (NI)** trial for a special type of RCT, **Sequential, Multiple Assignment, Randomized Trial (SMART)**.
- First, let's talk about **NI** in general setting.

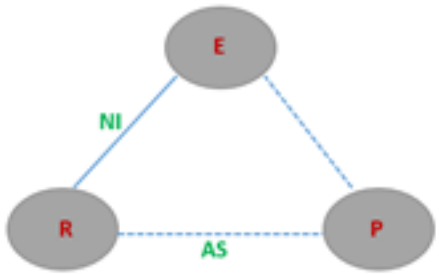
Non-Inferiority Trial

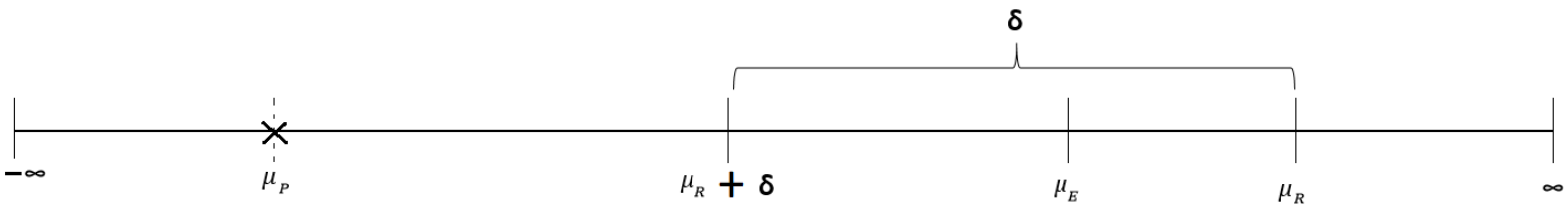
- In a classical Non-inferiority trial placebo arm is absent.
- This makes ethical sense but may lead to serious inferential consequence.
- An Experimental Intervention (E) is compared with an active reference/control (R) (e.g. established regime).
- WLOG, we assume larger mean implies better efficacy.
- Done mostly for ethical reason when it is established that R is clinically preferable to a placebo/standard-of-care (P) in a population/sub-population of interest.
- It requires a clinically acceptable margin ($\delta = -\Delta < 0$).

Hypothesis

For the two-arm NI, we are interested in testing (for a given/chosen $\delta = -\Delta < 0$: NI margin)

$$H_0 : \mu_E - \mu_R \leq \delta \text{ vs. } H_1 : \mu_E - \mu_R > \delta.$$



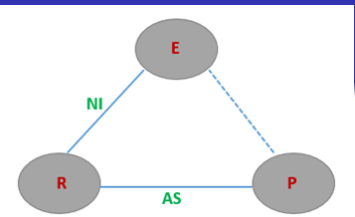


Assumptions

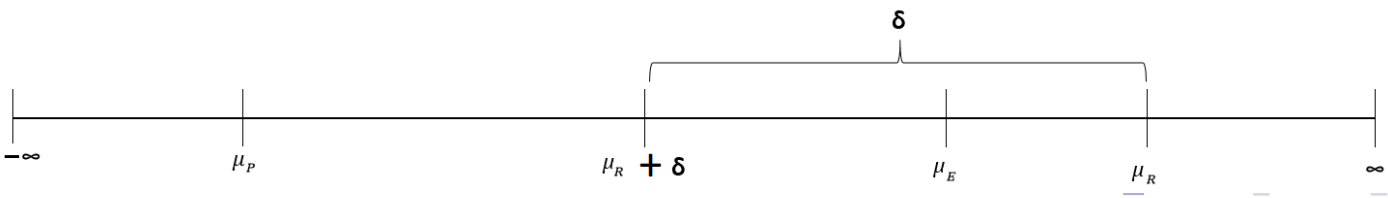
Since in a two-arm NI-trial placebo arm is absent, we are making implicit assumption of constancy and assay sensitivity.

- **Constancy**: The historic difference between the R and P are still valid in the current trial. i.e. we still reject the null hypothesis in the current setup.
- **Assay Sensitivity (AS)**: The ability of current trial to distinguish an effective treatment from a less effective or ineffective intervention (e.g. placebo).
- In a 2-arm trial NI margin needs to be chosen based on historical data. This will require **External Validation**.

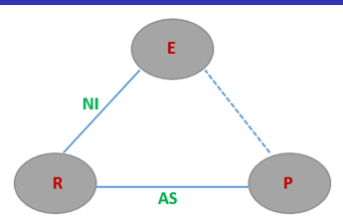
Gold Standard Design



- To avoid some of these assumptions, if it is ethically OK a placebo arm is added for internal validation.
- Resulting design is a 3-arm "Gold Standard design" which does not require External Validation.
- Note: We still need to worry about the NI margin δ.
- ICH (E9) gives some general guideline but still the solution is elusive and far from consensus even with a placebo arm.
- For 3-arm trial Pigeot et al. (2003) proposed a method of constructing "δ" as a fraction of difference between R and P in the classical setup under the assumption of homogeneity and normality.



Fraction Margin Approach



- Fraction margin approach (Pigeot et al 2003): In a three-arm trial, the construction of δ

$$\delta = (\mu_R - \mu_P)\tau,$$

with $\tau < 0$, assuming the gatekeeping AS condition ($\mu_R > \mu_P$).

- Three arm non-inferiority test can be rewritten as:

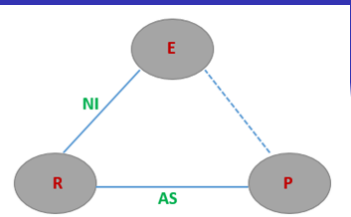
$$H_0 : \mu_E - \mu_R \leq (\mu_R - \mu_P)\tau \text{ vs. } H_1 : \mu_E - \mu_R > (\mu_R - \mu_P)\tau.$$

- After some algebra and putting $\theta = 1 + \tau$, where $\theta \in [0.5, 1)$

$$H_0 : \mu_E - \theta\mu_R - (1-\theta)\,\mu_P \leq 0 \quad \text{vs.} \quad H_1 : \mu_E - \theta\mu_R - (1-\theta)\,\mu_P > 0.$$

- To construct margin this approach makes the assumption $\mu_R - \mu_P > 0$.

Fraction Margin Approach (Continued . . .)



- Pigeot et al. (2003) argued that one first must reject the AS null hypothesis, AS $\Longrightarrow$ $K_0 : \mu_R \leq \mu_P$ vs. $K_1 : \mu_R > \mu_P$.
- However, due the hierarchical ordering of K_0 and H_0 no α adjustment is needed for NI testing.
- They provided power as a function of $\tau[\delta = \tau(\mu_R - \mu_P)]$. Note, δ is not directly specified, but % of effect E must retain over R is specified (i.e., effect retention).
- Koch and Tangen (1999) suggested the difference between R and E is expected to be much smaller than the difference of both treatments compared to placebo.
- It is also not very ethical to put large sample on placebo.
- Equal allocation may not provide optimal power in NI.

Shall We Test for NI Only?

- Hida and Tango, 2009 argued "The NI with AS is only established" when H_0 and K_0 are jointly rejected.
- This approach of is known as **Fixed Margin approach** as δ cannot be constructed rather has to be pre-chosen.
- It is basically a continuation of two-arm NI where δ is fixed or pre-chosen (i.e., from historical trial).
- Usual t-test is used and since joint rejection of H_0 and K_0 are achieved by the Intersection-Union test (IUT), no α adjustment is needed.
- IUT principle preserves α but may produce biased test.
- In this talk, we are focusing on **Fraction Margin Approach** for the 3-arm **SMART** design.
- Let's now talk about the standard **SMART** design.

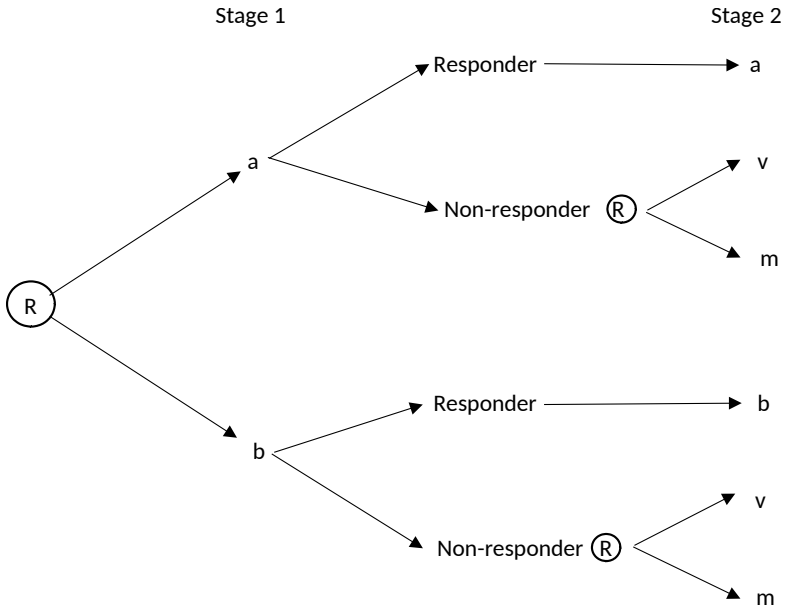
Adaptive Intervention (AI):

- A sequence of individualized intervention options that specify for whom and under what conditions different intervention alternatives should be provided, in order to address the changing needs of individuals as they proceed over time.
- A.K.A. adaptive treatment strategies, dynamic treatment regimes, stepped care, treatment policies.

SMART:

- A novel trial design that was developed to inform the development of adaptive interventions.
- Multi-stage trials; same participants throughout the study.
- Each stage corresponds to a decision rule.
- At each stage, subjects are randomized to set of treatment options.

- Inform the development of AIs while more closely mimicking treatment process.
- Develop a proposal for an AI, which could then be tested in a 2-arm randomized trial against an appropriate alternative.
- Evaluate the timing, sequencing and tailored selection of treatments through randomization.
- Collect information on other variables (besides tailoring variable of interest) and use observed data to estimate more personalized decision rules.

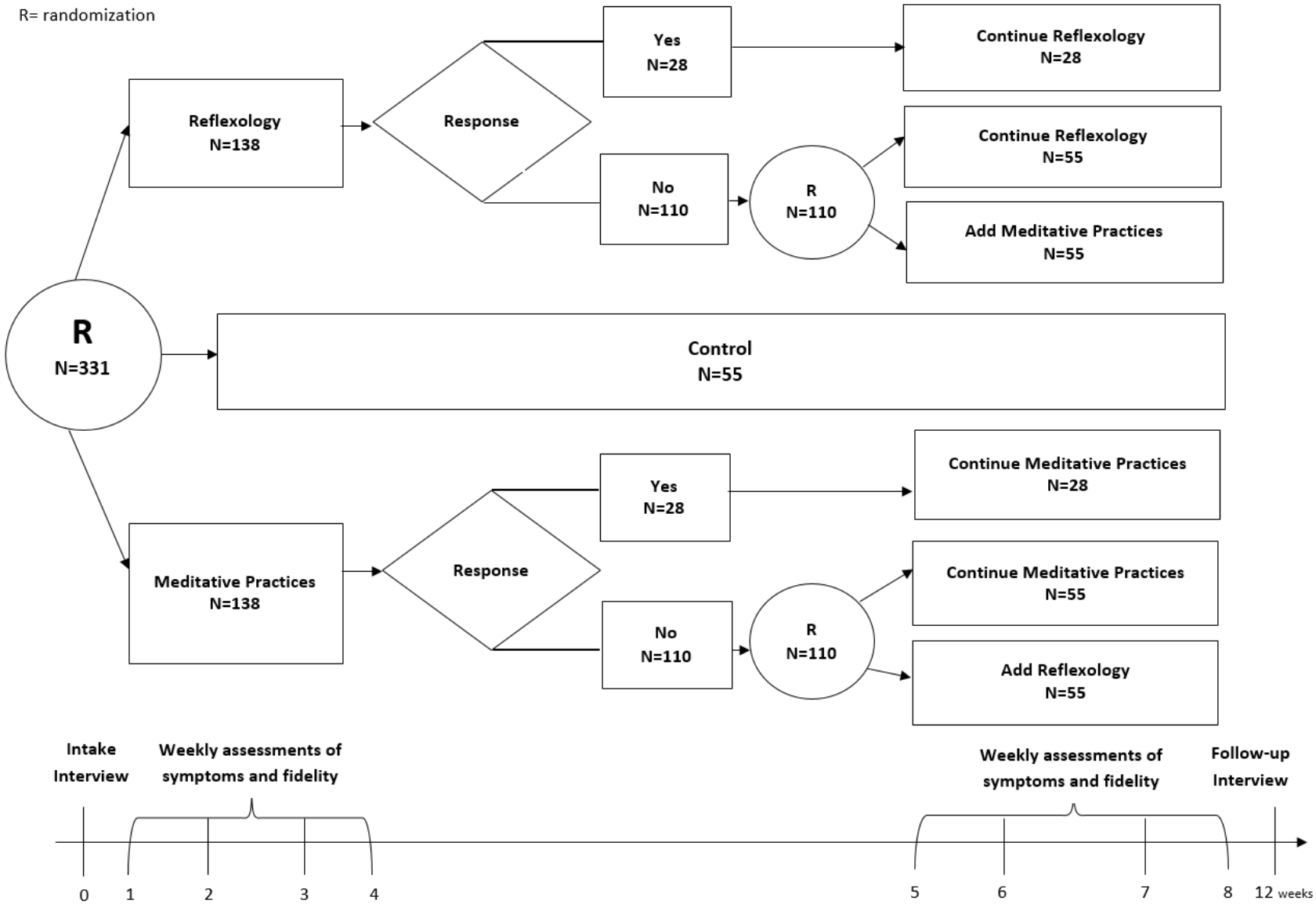


- (R): Randomization.
- First stage interventions: a, b.
- In standard SMART, the placebo arm is absent.
- This makes ethical sense because everybody is getting some treatment.

- Based on the individual's progress during the first intervention stage, the participant can be a responder or non-responder.
- Definition of responder is fixed at priori.
- At the second stage, responders are allowed to continue with the same initial intervention option, whereas non-responders are re-randomized to one of two second-stage options: *v* or *m*.

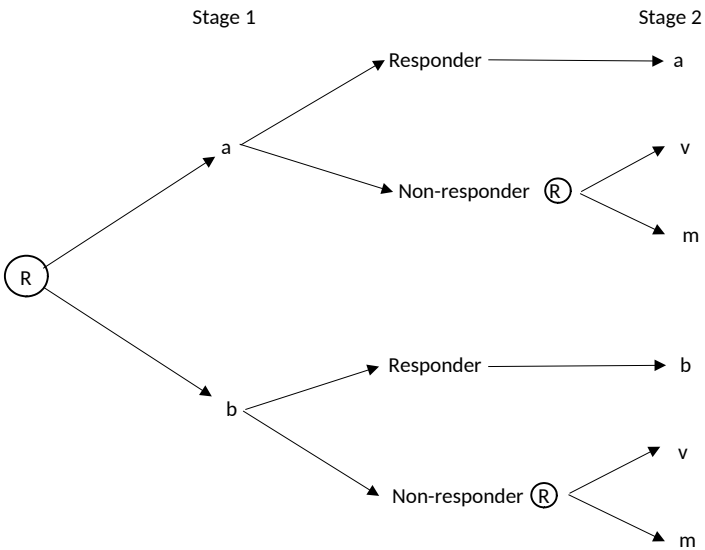
- Treatment Alternatives: must be used or could be used in practice - one may ask a clinician or community organizer how he/she would treat a person over time to develop the possible embedded AIs.
- Tailoring Variable: denotes early signs of non-response; use low dimensional summary to restrict second-stage treatments; must be agreed upon by experts in the field and clinically feasible.
- Ability to collect research evaluations on schedule necessary for critical decisions.
- Organized treatment and research team with systematic data collection.
- SMART design may differ depending on ethical, feasibility or strong scientific concerns.
- Goal is to keep the SMART design simple and the AIs realistic.
- Let's see an example (details will be discussed later) ....

SMART-MR Example: ("MR" for meditative practices/reflexology) for symptomp mangement during cancer treatment

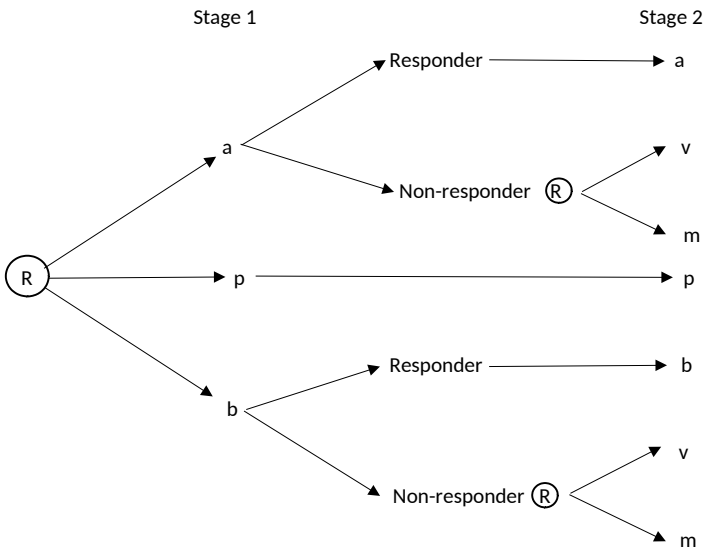


General Structure

- Control arm is to compare it with the intervention sequences.
- Control arm is recommended by NIH in the this SMART design.



(a) Standard SMART



(b) SMART with Placebo/Control

Adaptive Interventions

- Five AIs: With first stage intervention
  - a: $d_1 = (a, a^R v^{1-R})$, $d_2 = (a, a^R m^{1-R})$,
  - b: $d_3 = (b, b^R v^{1-R})$, $d_4 = (b, b^R m^{1-R})$,
  - p: $d_5 = (p, p)$.
- **Distinct path (DP):** Starting with different initial interventions
  - $\{d_1, d_5, d_3\}$,
  - $\{d_1, d_5, d_4\}$,
  - $\{d_2, d_5, d_3\}$,
  - $\{d_2, d_5, d_4\}$.
- **Shared path (SP):** Starting with the same initial intervention including placebo
  - $\{d_1, d_5, d_2\}$,
  - $\{d_3, d_5, d_4\}$.

Outcome

- T_1 and T_2: Intervention options at stages 1 and 2.
- R: response indicator ($R = 1$ for responders and $R = 0$ for non-responders).
- Observed data trajectory for the i^{th} individual in a SMART is given by $(T_{1i}, R_i, T_{2i}, Y_i)$, $i = 1, \ldots, N$, where N is the total number of individuals in the trial.
- An individual's potential outcome (Robins, 1997): Y_{T_1,T_2} where $T_1 \in \{a, p, b\}$; $T_2 \in \{a, v, m\}$ if $T_1 = a$; $T_2 \in \{b, v, m\}$ if $T_1 = b$; and $T_2 \in \{p\}$ if $T_1 = p$.
- Assume $E(Y_{T_1,T_2}) = \mu_{T_1,T_2}$ and $\text{Var}(Y_{T_1,T_2}) = \sigma^2$.

Probabilities

- π_{T_1}: First-stage randomization probability in favor of intervention option T_1.
- π_{T_1,T_2}: Second-stage randomization probability for those who started with the first-stage option T_1, in favor of intervention option T_2.
- $\pi_p = 1 - \pi_a - \pi_b$, $\pi_{a,m} = 1 - \pi_{a,v}$, $\pi_{b,m} = 1 - \pi_{b,v}$.
- A responder is assigned to the treatment sequence (T_1, T_1) with probability π_{T_1}.
- A non-responder is assigned to a treatment sequence (T_1, T_2) with probability $\pi_{T_1} \times \pi_{T_1,T_2}$.
- Using the principles of inverse probability weighting, the weight used for the i^{th} individual in any embedded AI d is $1/(\pi_{T_{1i}} \pi_{T_{1i}T_{2i}}^{1-R_i})$.

Mean of observed mean

- Define the observed mean:

$$\bar{Y}_d = \frac{1}{N}\sum_{i=1}^{N} W_i^d Y_i, \quad W_i^d = \frac{I\{T_{1i}=T_{1i}, T_{2i}=T_{1i}^{R_i} T_{2i}^{1-R_i}\}}{\pi_{T_{1i}} \pi_{T_{1i},T_{2i}}^{R_i}}.$$

- Define $\gamma_{T_1} \in \{\gamma_a, \gamma_p, \gamma_b\}$ as the response rate to the initial intervention options T_1, with respect to some pre-specified definition of response. In this case, $\gamma_p = 1$.
- Population mean:

$$\mu_d = E(\bar{Y}_d) = \gamma_{T_1}\mu_{T_1,T_1} + (1-\gamma_{T_1})\mu_{T_1,T_2}.$$

Variance and covariance of observed mean

- Population variance:

$$\begin{aligned}\mathrm{Var}(\bar{Y}_d) &= \frac{1}{N}\left\{\frac{1-\gamma_{T_1}+\gamma_{T_1}\pi_{T_1,T_2}}{\pi_{T_1}\pi_{T_1,T_2}}\sigma^2+\frac{\gamma_{T_1}(1-\gamma_{T_1}\pi_{T_1})}{\pi_{T_1}}\mu^2_{T_1,T_1}\right.\\ &\quad+\frac{(1-\gamma_{T_1})(1-(1-\gamma_{T_1})\pi_{T_1}\pi_{T_1,T_2})}{\pi_{T_1}\pi_{T_1,T_2}}\mu^2_{T_1,T_2}\\ &\quad\left.-2\gamma_{T_1}(1-\gamma_{T_1})\mu_{T_1,T_1}\mu_{T_1,T_2}\right\}=\frac{\sigma^2_d}{N}.\end{aligned}$$

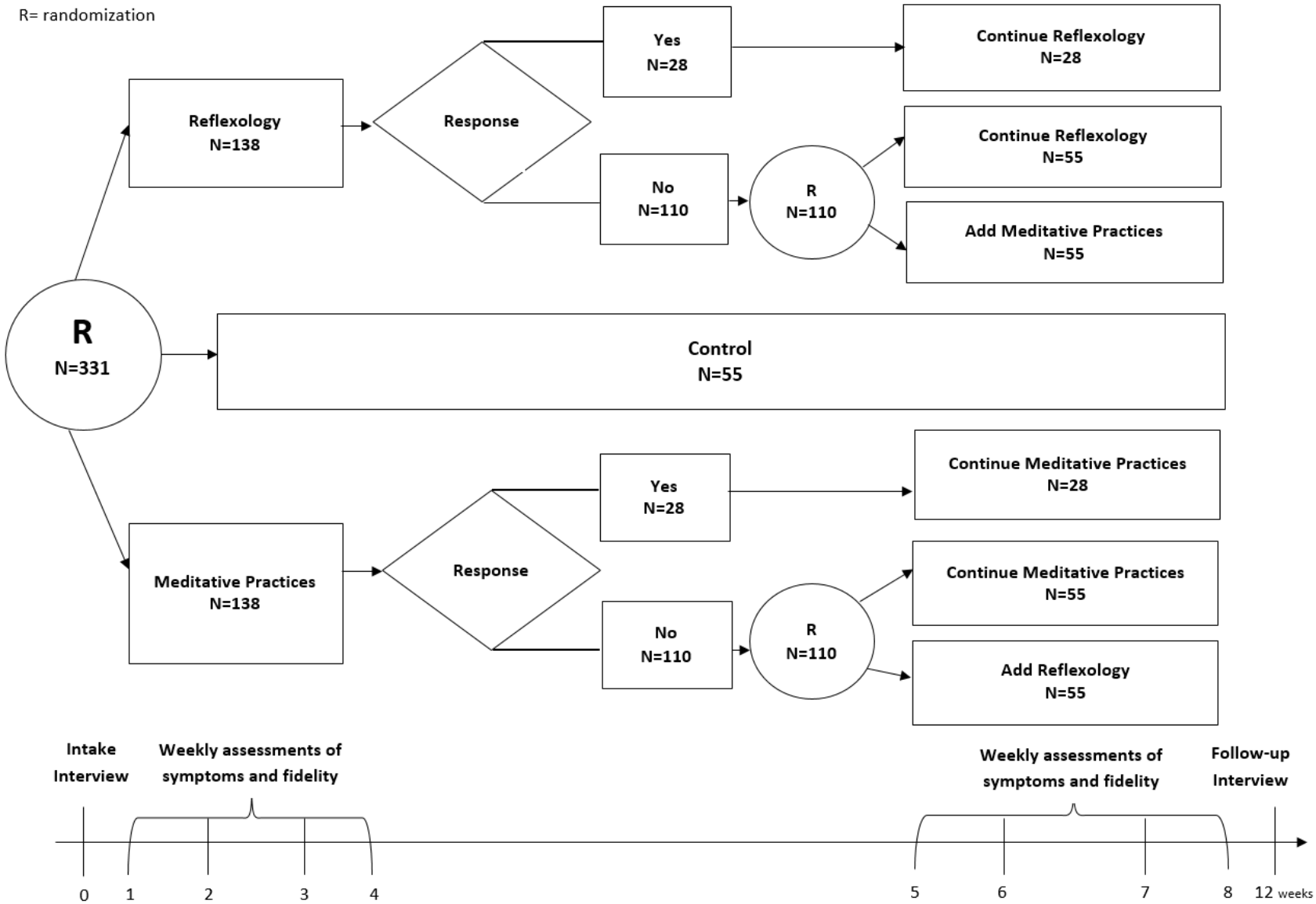
- For the SP, the covariance between $\bar{Y}_{d_1}$ and $\bar{Y}_{d_2}$:

$$\begin{aligned}\mathrm{Cov}(\bar{Y}_{d_1},\bar{Y}_{d_2}) &= \frac{1}{N}\left\{\frac{\gamma_a(\sigma^2+\mu^2_{a,a})}{\pi_a}-(\gamma_a\mu_{a,a}+(1-\gamma_a)\mu_{av})(\gamma_a\mu_{a,a}\right.\\ &\quad\left.+(1-\gamma_a)\mu_{a,m})\right\}=\frac{\sigma_{d_1d_2}}{N}.\end{aligned}$$

- Let's talk about the example.....

- Existing methods for using SMART study is to compare AIs and associated power planning resources, are suitable for traditional superiority testing, where the goal is to investigate whether one AI is more efficacious than the other.
- In NI testing, the goal is to establish that a new intervention yields favorable outcomes, which when compared to another reference intervention, are not below some pre-specified margin.
- The NI margin captures how close the new intervention must be to the established one in terms of the expected outcome in order for the new intervention to be considered NI to the active reference.
- In the 3-arm SMART study, the goal is to test the non-inferiority of an AI with respect to another AI in the presence of Placebo/Standard-of-care arm.
- Before combining the idea of **Non-inferiority** testing with **SMART**, let's see the example we have mentioned earlier.

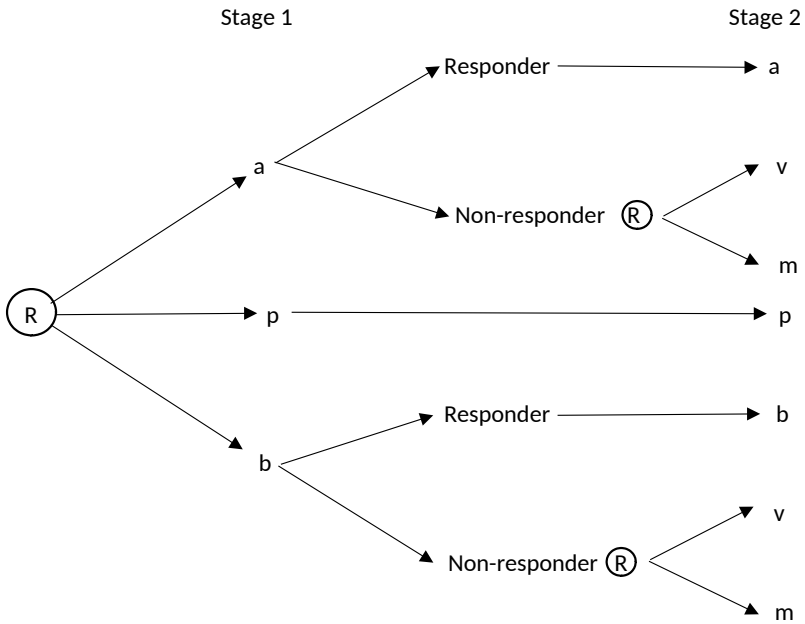
Example: SMART-MR ("MR" for meditative practices/reflexology) during cancer treatment



Example: SMART-MR ("MR" for meditative practices/reflexology) during cancer treatment

- The participant flow through SMART-MR ("MR" for meditative practices/reflexology) for symptom management during cancer treatment.
- Counts are what we projected before the study started.
- Interventions: Reflexology, Meditative practices, placebo/control.
- In SMART-MR, we have
  - a sequence that starts with reflexology, then meditative practices intervention is added.
  - a sequence that starts with meditative practices, then reflexology is added.
  - a control arm is to compare intervention sequences to control. While each intervention has been previously compared to control, adding two interventions sequentially is not guaranteed to produce a better outcome compared to one intervention.
- Also, in SMARTs of self-management interventions, one cannot switch to another intervention.
- This example provides motivation for **NI** testing in **SMART**.

- Three-arms:

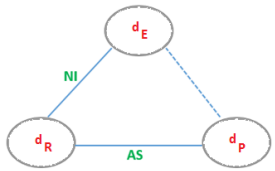


  - Reference: d_R (with first stage intervention as a or b),
  - Experimental: d_E (with first stage intervention as a or b),
  - Placebo: d_P (with both stage interventions as p).
- Corresponding means: μ_{d_R}, μ_{d_E}, and μ_{d_P}.

- In terms of SMART design, the two arm NI ($\delta = -\Delta < 0$: NI margin) can be written as

$$H_0 : \mu_{d_E} - \mu_{d_R} \leq \delta \text{ vs. } H_1 : \mu_{d_E} - \mu_{d_R} > \delta.$$

- As the placebo arm is absent, μ_{d_R} vs. μ_{d_P} (AS) can't be tested.
- Followed by the Fraction margin approach (Pigeot et al 2003), the construction of δ in the three-arm SMART can be modified as $\delta = (\mu_{d_R} - \mu_{d_P})\tau$, where $\tau < 0$ assuming the AS condition ($\mu_{d_R} > \mu_{d_P}$) holds.

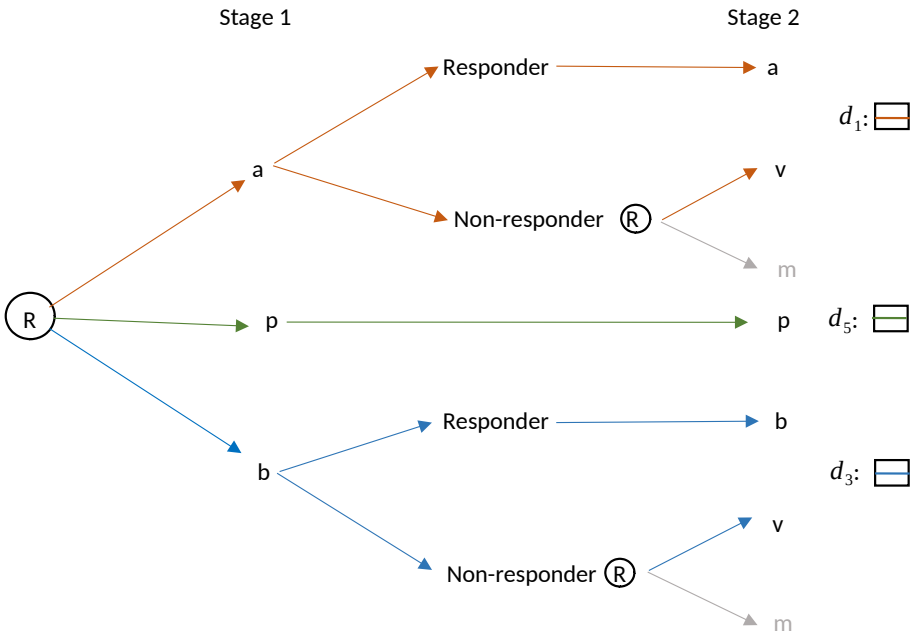


- After some algebra and putting $\theta = 1 + \tau$, three arm **NI-SMART** can be reduced to

$$H_0 : \mu_{d_E} - \theta\mu_{d_R} - (1-\theta)\,\mu_{d_P} \leq 0 \quad \text{vs.} \quad H_1 : \mu_{d_E} - \theta\mu_{d_R} - (1-\theta)\,\mu_{d_P} > 0.$$

- Now, we illustrate the **NI** testing for **Distinct** and **Shared** paths.

- As we have mentioned earlier, **Distinct path** is starting with different initial interventions.
- For example, Experimental path: $d_E = d_1 = (a, a^R v^{1-R})$, Reference path: as $d_R = d_3 = (b, b^R v^{1-R})$, and Control path: $d_P = d_5 = (p, p)$.



- Unscaled test statistic of the linear contrast $\mu^{DP}_{d_1 d_5 d_3} = \mu_{d_1} - \theta\mu_{d_3} - (1-\theta)\,\mu_{d_5}$, is given by

$$T^{DP}_{d_1 d_5 d_3} = \bar{Y}_{d_1} - \theta\bar{Y}_{d_3} - (1-\theta)\,\bar{Y}_{d_5}$$

with variance $\sigma^2_{d_1 d_5 d_3} = (\sigma^2_{d_1} + \theta^2\sigma^2_{d_3} + (1-\theta)^2\sigma^2_{d_5})/N$, where

$$\begin{aligned}
\sigma^2_{d_1} &= \frac{1-\gamma_a+\gamma_a\pi_{a,v}}{\pi_a\pi_{a,v}}\sigma^2 + \frac{\gamma_a(1-\gamma_a\pi_a)}{\pi_a}\mu^2_{a,a} + \frac{(1-\gamma_a)(1-(1-\gamma_a)\pi_a\pi_{a,v})}{\pi_a\pi_{a,v}}\mu^2_{a,v} \\
&\quad -2\gamma_a(1-\gamma_a)\mu_{a,a}\mu_{a,v}, \\
\sigma^2_{d_3} &= \frac{1-\gamma_b+\gamma_b\pi_{b,v}}{\pi_b\pi_{b,v}}\sigma^2 + \frac{\gamma_b(1-\gamma_b\pi_b)}{\pi_b}\mu^2_{b,b} + \frac{(1-\gamma_b)(1-(1-\gamma_b)\pi_b\pi_{b,v})}{\pi_b\pi_{b,v}}\mu^2_{b,v} \\
&\quad -2\gamma_b(1-\gamma_b)\mu_{b,b}\mu_{b,v}, \\
\sigma^2_{d_5} &= \frac{1}{\pi_p}(\sigma^2 + \mu^2_{p,p}(1-\pi_p)).
\end{aligned}$$

- The large sample distribution of the standardized statistic:

$$Z^{\text{DP}}_{d_1 d_5 d_3} = \frac{T^{\text{DP}}_{d_1 d_5 d_3} - \mu^{\text{DP}}_{d_1 d_5 d_3}}{\sqrt{\sigma^2_{d_1 d_5 d_3}}} \overset{H_0}{\sim} \text{Normal}(0,1).$$

- Reject H_0 and conclude non-inferiority if

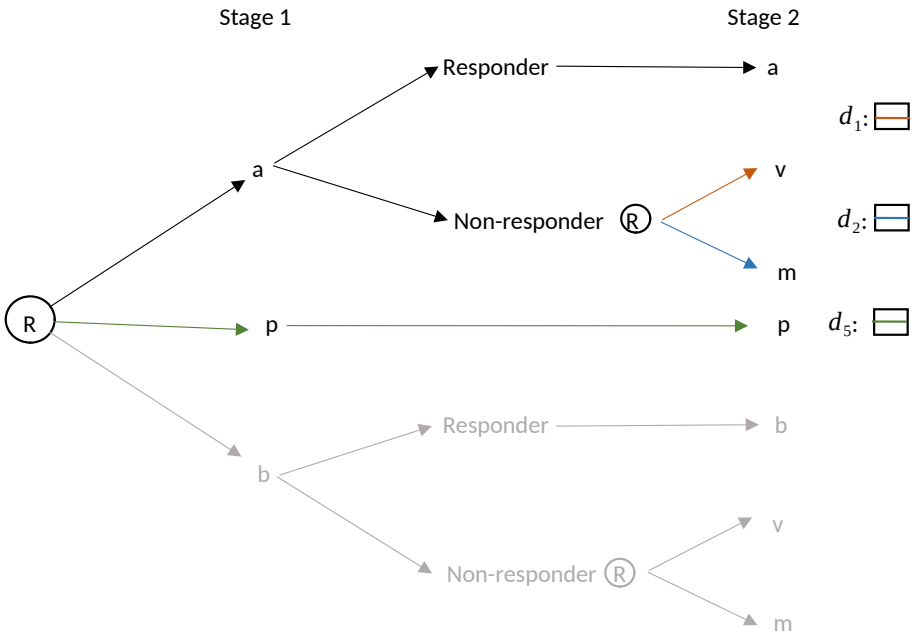
$$Z^{\text{DP}}_{d_1 d_5 d_3} > z_\alpha (\text{one-sided test}).$$

- Under a specified effect size ψ^{DP} such that $\mu_{d_1} - \theta\mu_{d_3} - (1-\theta)\,\mu_{d_5} = \psi^{\text{DP}}$, the required sample size (assuming the equal variance) is

$$N^{DP} = \frac{(z_\alpha + z_\beta)^2}{\eta^{\text{DP}^2}},$$

where $\eta^{\text{DP}} = \psi^{\text{DP}} / \sqrt{\sigma^2_{d_1 d_5 d_3}}$: standardized effect size.

- As we have mentioned earlier, **Shared path** is starting with the same initial intervention including placebo.
- For example, Experimental path: $d_E = d_1 = (a, a^R v^{1-R})$, Reference path: as $d_R = d_2 = (a, a^R m^{1-R})$, and Control path: $d_P = d_5 = (p, p)$.



- Unscaled test statistic of the linear contrast $\mu^{SP}_{d_1 d_5 d_2} = \mu_{d_1} - \theta\mu_{d_2} - (1-\theta)\,\mu_{d_5}$, is given by

$$T^{SP}_{d_1 d_5 d_2} = \bar{Y}_{d_1} - \theta\,\bar{Y}_{d_2} - (1-\theta)\,\bar{Y}_{d_5}$$

with variance $\sigma^2_{d_1 d_5 d_2} = (\sigma^2_{d_1} + \theta^2\sigma^2_{d_2} + (1-\theta)^2\sigma^2_{d_5})/N$, where

$$\begin{aligned}\sigma^2_{d_2} \quad = \quad & \frac{1-\gamma_a+\gamma_a\pi_{a,m}}{\pi_a\pi_{a,m}}\sigma^2 + \frac{\gamma_a(1-\gamma_a\pi_a)}{\pi_a}\mu^2_{a,a} + \frac{(1-\gamma_a)(1-(1-\gamma_a)\pi_a\pi_{a,m})}{\pi_a\pi_{a,m}}\mu^2_{a,m} \\ & -2\gamma_a(1-\gamma_a)\mu_{a,a}\mu_{a,m}.\end{aligned}$$

- The large sample distribution of the standardized statistic:

$$Z^{SP}_{d_1 d_5 d_2} = \frac{T^{SP}_{d_1 d_5 d_2} - \mu^{SP}_{d_1 d_5 d_2}}{\sqrt{\sigma^2_{d_1 d_5 d_2}}} \overset{H_0}{\sim} \text{Normal}(0,1).$$

- Reject H_0 and conclude non-inferiority if

$$Z^{\text{SP}}_{d_1 d_5 d_2} > z_\alpha (\text{one-sided test}).$$

- Under a specified effect size ψ^{DP} such that $\mu_{d_1} - \theta\mu_{d_2} - (1-\theta)\,\mu_{d_5} = \psi^{\text{SP}}$, the required sample size (assuming the equal variance) is

$$N^{SP} = \frac{(z_\alpha + z_\beta)^2}{\eta^{\text{SP}^2}},$$

where $\eta^{\text{SP}} = \psi^{\text{SP}} / \sqrt{\sigma^2_{d_1 d_5 d_2}}$: standardized effect size.

- Let's see some simulation studies for three arm **NI-SMART** to explore power/sample size requirement.

1. Specify the model parameters and set COUNT $= 0$.
2. For the power curves, fix the value of N. For the sample size and power calculation, calculate N for the fixed the type-I error rate of 5%, the desired power at 80%, and η at a fixed value.
3. Generate M many datasets for DP or SP.
4. For m^{th} dataset,
   1. Compute $Z_{\text{obs}}^{(m)}$, i.e., $Z_{d_1 d_5 d_3}^{\text{DP}}$ for DP or $Z_{d_1 d_5 d_2}^{\text{SP}}$ for SP.
   2. Calculate the p-value: $P_{H_0}(Z_{\text{obs}}^{(m)} > 0)$.
   3. If p-value $< \alpha$, we reject the null hypothesis and increase the COUNT$^{(m)}$ by 1, otherwise 0.
5. Calculate the empirical power as follows:

$$\hat{\phi} = \frac{1}{M}\sum_{m=1}^{M} \text{COUNT}^{(m)} = \frac{1}{M}\sum_{m=1}^{M} I\left(P_{H_0}(Z_{\text{obs}}^{(m)} > 0) < \alpha\right),$$

where $I(\cdot)$: indicator variable and COUNT$^{(m)}$: COUNT at m^{th} simulation.

Table: Standardized effect size (η^{DP}), γ response rate, Sample size (N^{DP}) and empirical power ($\hat{\phi}$) for distinct path

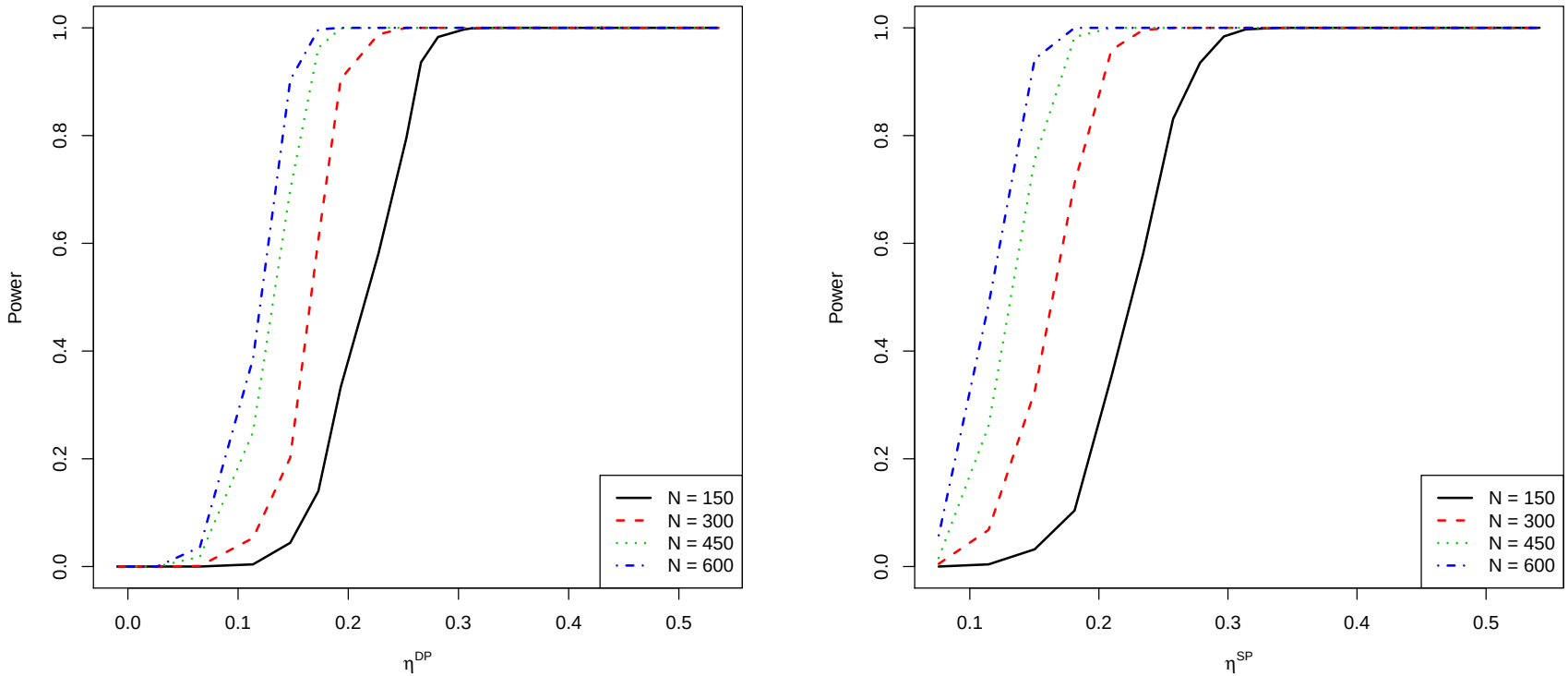
$\xi_{2a,v}$	γ_b	$\theta = 0.5$			$\theta = 0.6$			$\theta = 0.7$		
		η^{DP}	N^{DP}	$\hat{\phi}$	η^{DP}	N^{DP}	$\hat{\phi}$	η^{DP}	N^{DP}	$\hat{\phi}$
3.5	0.10	0.163	234	0.838	0.155	258	0.829	0.147	288	0.825
	0.15	0.161	240	0.841	0.154	264	0.834	0.145	297	0.820
	0.20	0.160	243	0.841	0.152	270	0.829	0.143	306	0.832
	0.25	0.159	246	0.836	0.150	276	0.831	0.141	312	0.832
	0.30	0.157	252	0.851	0.149	282	0.855	0.139	321	0.838
	0.35	0.156	255	0.828	0.147	288	0.833	0.137	330	0.832
4.0	0.10	0.178	198	0.847	0.170	216	0.842	0.161	240	0.832
	0.15	0.176	201	0.859	0.168	219	0.832	0.159	246	0.838
	0.20	0.175	204	0.853	0.167	225	0.846	0.158	249	0.832
	0.25	0.174	207	0.846	0.165	228	0.836	0.156	255	0.844
	0.30	0.172	210	0.839	0.164	231	0.854	0.154	261	0.848
	0.35	0.171	213	0.852	0.162	237	0.842	0.153	267	0.836
4.5	0.10	0.192	171	0.881	0.184	183	0.861	0.175	204	0.859
	0.15	0.190	171	0.856	0.182	189	0.860	0.173	207	0.836
	0.20	0.189	174	0.862	0.181	192	0.861	0.171	213	0.849
	0.25	0.188	177	0.850	0.179	195	0.848	0.170	216	0.856
	0.30	0.186	180	0.845	0.178	198	0.855	0.168	219	0.845
	0.35	0.185	183	0.867	0.176	201	0.863	0.167	225	0.849

Hypothesis: $H_0 : \mu_{d_1} - \theta\mu_{d_3} - (1-\theta)\,\mu_{d_5} \leq 0$ vs. $H_1 : \mu_{d_1} - \theta\mu_{d_3} - (1-\theta)\,\mu_{d_5} > 0$.

Table: Standardized effect size (η^{SP}), γ response rate, Sample size (N^{SP}) and empirical power ($\hat{\phi}$) for shared path

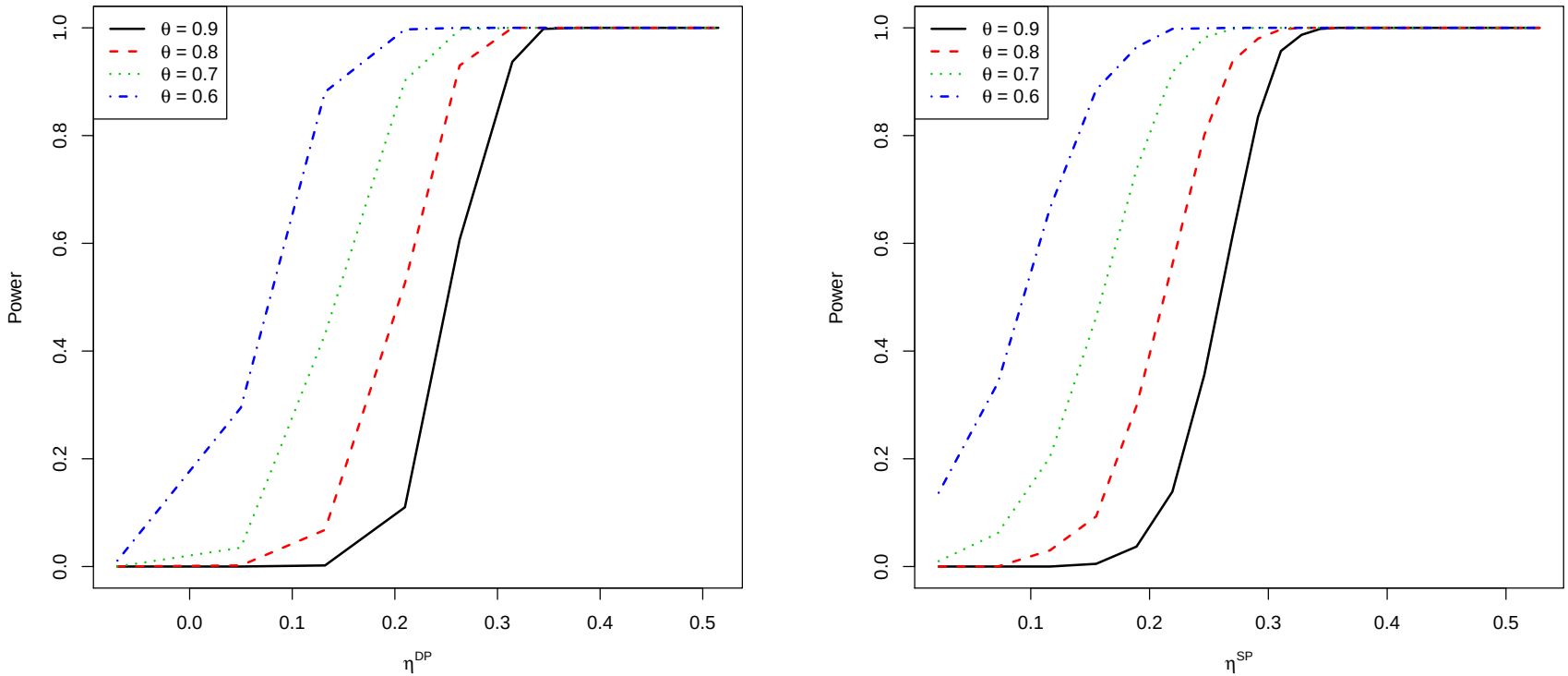
$\xi_{2a,v}$	γ_a	$\theta = 0.5$			$\theta = 0.6$			$\theta = 0.7$		
		η^{SP}	N^{SP}	$\hat{\phi}$	η^{SP}	N^{SP}	$\hat{\phi}$	η^{SP}	N^{SP}	$\hat{\phi}$
3.5	0.10	0.201	153	0.834	0.197	162	0.857	0.192	171	0.845
	0.15	0.197	162	0.840	0.192	168	0.840	0.187	180	0.848
	0.20	0.192	168	0.851	0.188	177	0.847	0.182	189	0.862
	0.25	0.188	177	0.856	0.183	186	0.864	0.177	198	0.832
	0.30	0.184	183	0.848	0.179	195	0.835	0.172	210	0.844
	0.35	0.180	192	0.836	0.174	204	0.845	0.167	222	0.857
4.0	0.10	0.222	126	0.849	0.218	132	0.866	0.212	138	0.845
	0.15	0.216	132	0.861	0.212	138	0.851	0.206	147	0.847
	0.20	0.211	141	0.845	0.207	147	0.866	0.201	156	0.853
	0.25	0.206	147	0.864	0.201	153	0.869	0.195	165	0.859
	0.30	0.201	153	0.838	0.196	162	0.847	0.189	174	0.857
	0.35	0.196	162	0.847	0.191	171	0.863	0.184	186	0.872
4.5	0.10	0.240	108	0.890	0.236	111	0.863	0.230	117	0.856
	0.15	0.234	114	0.885	0.230	117	0.870	0.224	126	0.861
	0.20	0.228	120	0.840	0.224	126	0.864	0.218	132	0.866
	0.25	0.223	126	0.858	0.218	132	0.878	0.211	141	0.871
	0.30	0.217	132	0.872	0.212	138	0.878	0.205	150	0.878
	0.35	0.211	141	0.886	0.206	147	0.850	0.199	159	0.874

Hypothesis: $H_0 : \mu_{d_1} - \theta\mu_{d_2} - (1-\theta)\,\mu_{d_5} \leq 0$ vs. $H_1 : \mu_{d_1} - \theta\mu_{d_2} - (1-\theta)\,\mu_{d_5} > 0$.



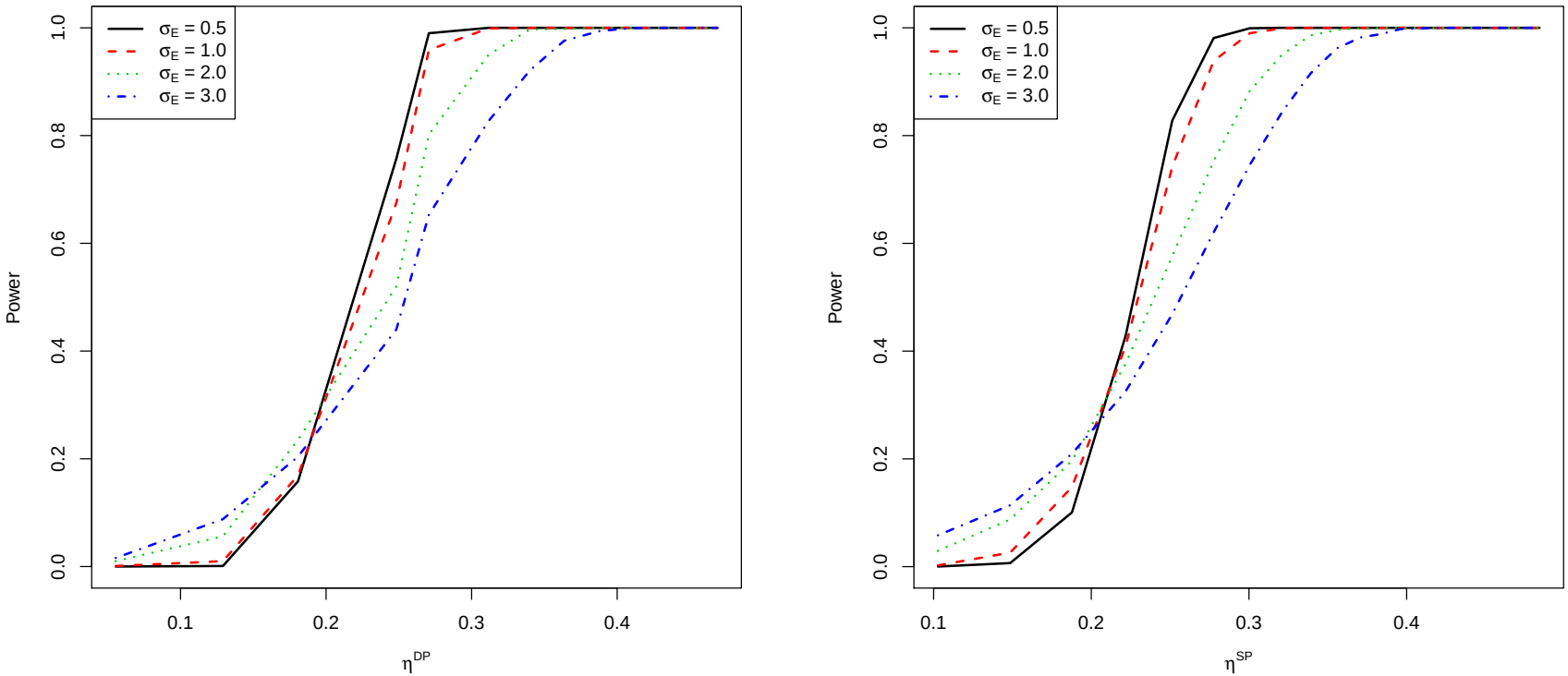
(a) (b)

Figure: Power curves for different N and $\theta = 0.8$, (a) Distinct path and (b) Shared path



(a) (b)

Figure: Power curves for different θ and $N = 120$, (a) Distinct path and (b) Shared path



(a) (b)

Figure: Power curves for different σ_E^2 with $\theta = 0.8$ and $N = 150$, (a) Distinct path and (b) Shared path

Future work:

- Joint testing procedure of Hida-Tango can be done for 3-arm NI-SMART.
- Sequential testing (AS $->$ NI) is another extension in this setting.
- In this study, we explored the continuous outcome.
- Binary, count, time-to-event outcomes can also be explored.
- For non-continuous outcome, defining NI margin via Fraction-margin approach is not unique. (Chowdhury et al., 2018)
- For NI testing Bayesian methods are well established as substantial historical information is available for active reference (and possibly for placebo/standard-of-care).
- Similar methods can be developed in the NI-SMART framework.

Thank You

Any question or suggestion?

Selected references

- Nahum-Shani, I., Qian, M., Almiral, D., Pelham, W., Gnagy, B., Fabiano, G., . . . , Murphy, S.A. (2012). Experimental design and primary data analysis methods for comparing adaptive interventions. Psychological Methods, 17, 457-477.
- Pigeot, I., Schäfer, J., Röhmel, J., and Hauschke, D. (2003). Assessing non-inferiority of a new treatment in a three-arm clinical trial including a placebo. Statistics in Medicine, 22(6):883-899.
- Robins, J. M. (1997). Causal inference from complex longitudinal data. In M. Berkane (Ed.), Latent variable modeling and applications to causality: Lecture notes in statistics (pp. 69 -– 117). New York, NY, Springer-Verlag.
- Ghosh, S., Ghosh, S., and Tiwari, R. C. (2016). Bayesian approach for assessing non-inferiority in a three-arm trial with pre-specified margin. Statistics in medicine, 35(5), 695-708.
- Chowdhury, S., Tiwari, R. C., and Ghosh, S. (2018). Non-inferiority testing for risk ratio, odds ratio and number needed to treat in three-arm trial. Computational Statistics & Data Analysis.