The Use of Decreasingly Informative Priors in Adaptive Clinical Trial Designs

Roy T. Sabo

Department of Biostatistics
School of Public Health
Virginia Commonwealth University

Monday, November 4, 2024

Response-Adaptive Allocation (Rosenberger et al, 2001)

- Two Sample Optimal Design:
  - (Binary) Reduces treatment failures $\left(\sum_{j=1}^{K} n_j(1-p_j)\right)$
  - (Cont) Minimize/Maximize Response $\left(\sum_{j=1}^{K}\sum_{i=1}^{n_j} y_{ij}\right)$
  - Allocation weights:

$$w_A = \frac{\sqrt{p_A}}{\sqrt{p_A}+\sqrt{p_A}}, w_B = 1 - w_A$$

$$p_j \quad \text{estimated with sample proportion } \hat{p}_j$$

- Problems with RA design:
  - Undefined (when $y_j = 0$ or $y_j = n_j$) or unstable (highly variable) allocation weights early in trial (Thall 2007).
  - Ethical concerns up for debate

- Possible Solutions:
  - *Hard Lead-In*: equal allocation until pre-determined accrual target
  - *Start-Up Design*: Equal allocation until all parameters in all groups are estimable (Haines and Sadiq, 2015)
  - *Natural lead-in*:
    - Restricts allocation proportions in early phases of trial
    - Ability to adapt increases as trial continues
    - Often use Bayes estimators

- Two-Group Natural Lead-In (Thall and Wathen, 2007):
  - Allocation weights:

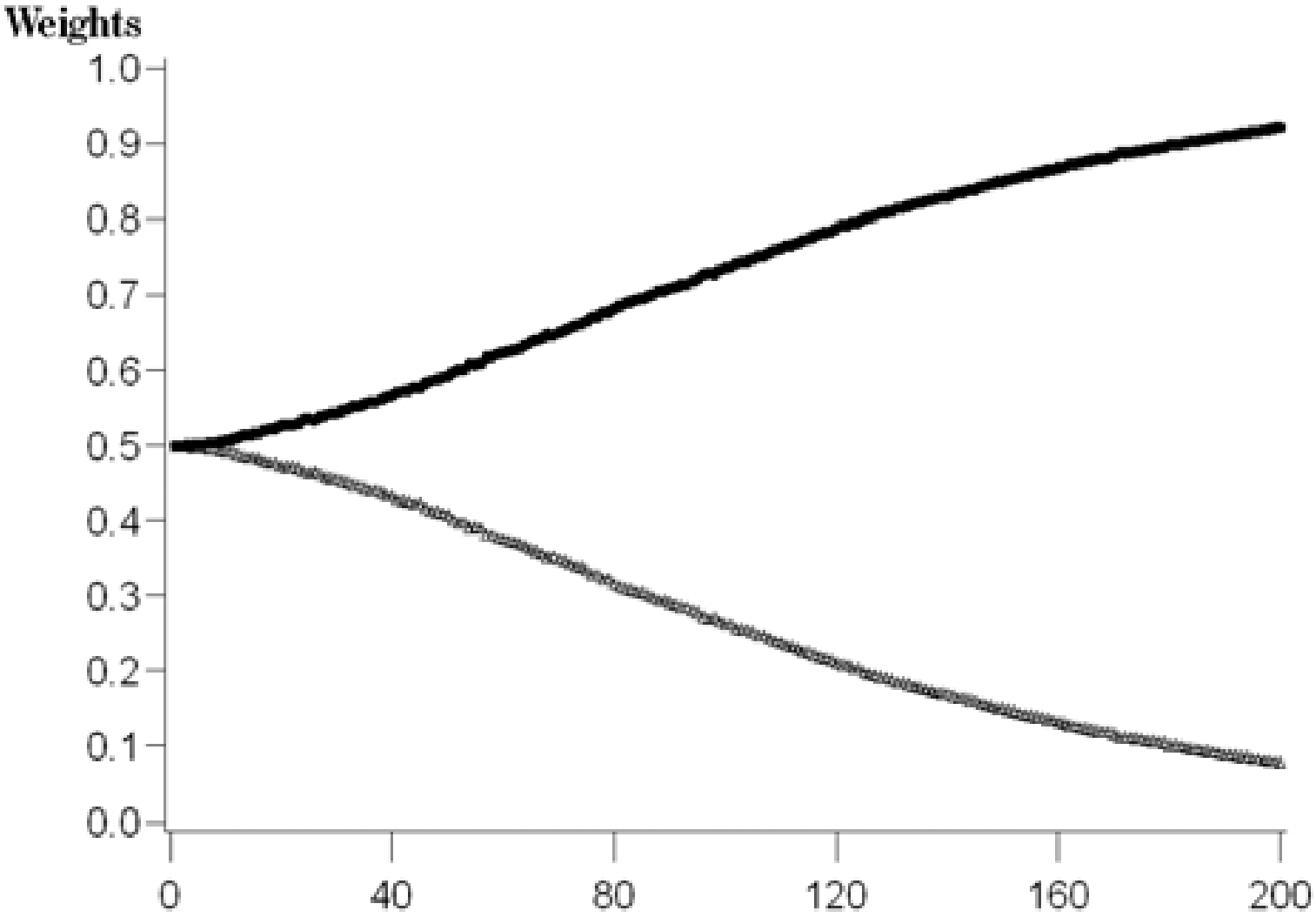
$$w_A = \frac{p_A^{n/2N}}{p_A^{n/2N} + p_B^{n/2N}}, \; w_B = 1 - w_A$$

  - Three-Group Natural Lead-In (Bello and Sabo, 2016):

$$w_j = \frac{\left(w_j^*\right)^{n/2N}}{\left(w_1^*\right)^{n/2N} + \left(w_2^*\right)^{n/2N} + \left(w_3^*\right)^{n/2N}}, \; j = 1, 2, 3$$

  - For both methods
    - Weights begin at equal allocation
    - Weights converge toward "Full RA" weights as $n \to \infty$
    - Inestimable weights still a problem early

- Natural Lead-In Weight Behavior
- True Efficacy: $p_1 = 0.5$ $p_2 = 0.3$, $N = 200$



- Problem: Natural Lead-In is *ad hoc*
- Can we design approach that
  - Is Based on optimal design
  - Achieves same gradual adaptation

Decreasingly Informative Priors (Sabo ,2014)

- Functions of observed (n) and planned (N) sample sizes
- Highly informative:
  - When n is small, more information in prior than likelihood
  - As $n \uparrow$: information incrementally transferred to likelihood
  - Initially low variance that increases as $n \uparrow$
- Skeptical:
  - Mean/Mode centered at null-hypothesized value
  - If there are two or more groups: assigned the same prior
- Main Idea: As n increases, information incrementally transfers from prior to likelihood

- General Set-Up:

$$\begin{aligned} P(\theta|y) &\propto f(y|\theta)\pi(\theta|\theta_0, n, N)g(\theta_0|\lambda) \\ y \sim f(y|\theta) &\rightarrow \text{Observed outcome and likelihood} \\ \theta \sim \pi(\theta|\theta_0, n, N) &\rightarrow \text{Parameter of interest and DIP} \\ \theta_0 \sim g(\theta_0|\lambda) &\rightarrow \text{Hyperprior on } \theta_0 \text{ with hyperparameter } \lambda \end{aligned}$$

- Note
  - $g(\theta_0|\lambda)$ is often a point mass at θ_0
  - θ_0 is determined by the null hypothesis

How Do We Parameterize DIP

- DIP should be skeptical
- As n increases $\rightarrow$ information incrementally transfers from prior to likelihood

- How much weight should the prior get?
  - Influence of prior should be proportional to ratio of unobserved to observed sample size
  - Thus
    - Equate effective sample size to unobserved sample size
    - Equate prior mode to null hypothesized value θ_0
    - Solve the two equations for prior parameters

ESS: Neuenschwander et.al (2020)

- Expected Local-Information-Ratio (ELIR)
  - Ratio of the prior information to the Fisher information

$$ESS_{ELIR} = E_\theta\{r(\theta)\} = E\left\{\frac{i(p(\theta))}{i_F(\theta)}\right\}$$

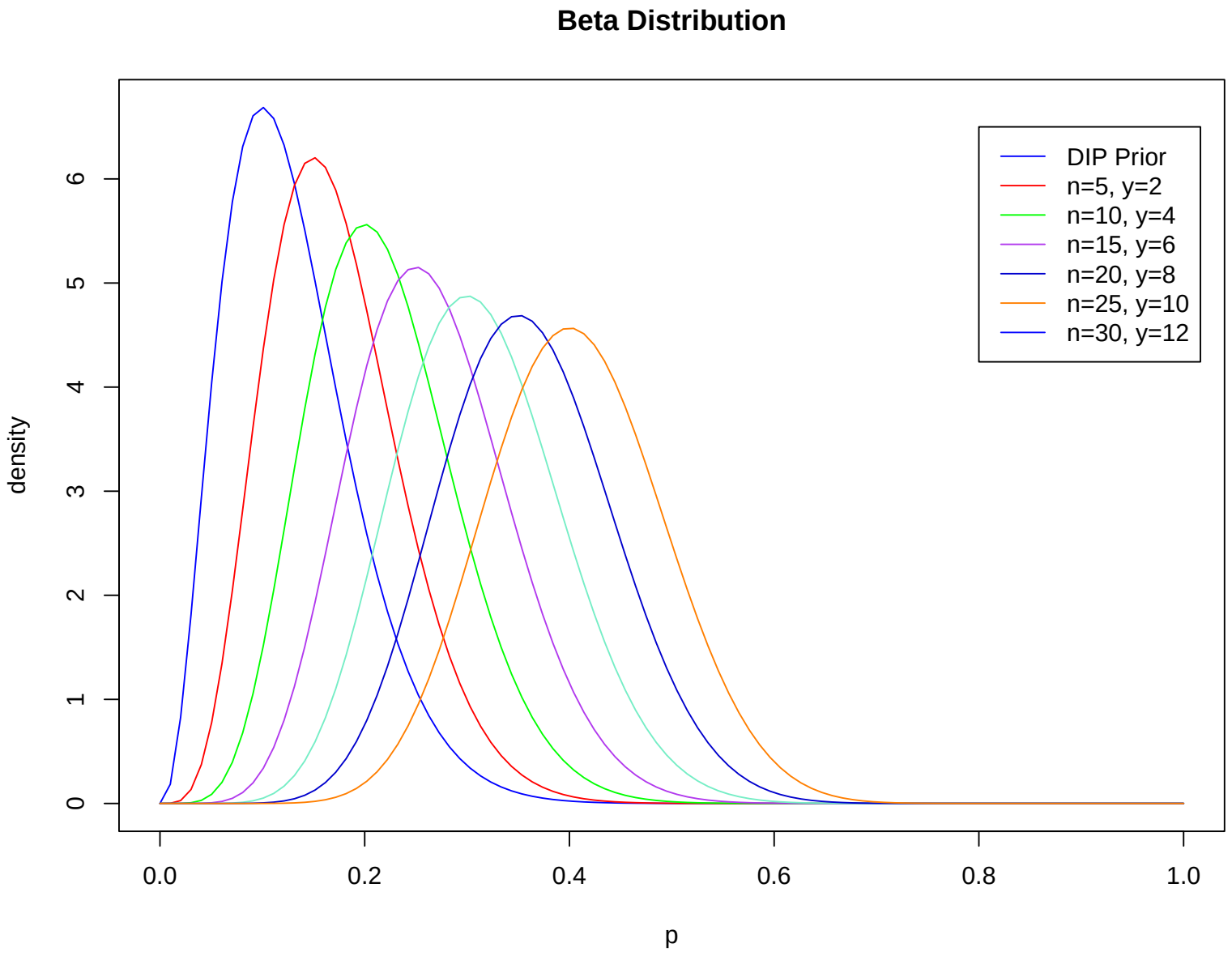
- Beta Distribution ($p \sim \text{beta}[a, b]$)

$$\begin{aligned} \text{Mean} \quad & \frac{a-1}{a+b-2} = p_0 \\ \text{ESS} \quad & E_\theta\{r(\theta)\} = a + b = N - n \\ \therefore \quad & a = 1 + p_0(N-n) \\ \& \quad & b = 1 + (1-p_0)(N-n) \end{aligned}$$

- Say we have binary outcomes in K groups.
- Model: beta-binomial conjugate pair.
- Estimate Posterior Mean from the either of the following

- Point Mass DIP centered at p_0.
  - $f(.) \rightarrow y_k \sim \text{binomial}(n_k, p_k)$
  - DIP: $\text{p}_k \sim \text{beta}\left[1 + p_0(N-n), 1 + (1-p_0)(N-n)\right]$
  - $p_0 = p_0$ with probability 1
  - $p_k \sim \text{beta}\left[1 + y_k + p_0(N-n), 1 + (n_k - y_k) + (1-p_0)(N-n)\right]$

- DIP centered at p_0 with hyperprior.
  - $f(.) \rightarrow y_k \sim \text{binomial}(n_k, p_k)$
  - DIP: $p_k \sim \text{beta}\left[1 + p_0(N-n), 1 + (1-p_0)(N-n)\right]$
  - $p_0 \sim \text{beta}\left[1 + \delta_1, 1 + \delta_2\right]$
  - $p_k \sim f(p_k | y_k, n, N, p_0, \delta_1, \delta_2)$


Beta Distribution
density
p
0
1
2
3
4
5
6
0.0
0.2
0.4
0.6
0.8
1.0
DIP Prior
n=5, y=2
n=10, y=4
n=15, y=6
n=20, y=8
n=25, y=10
n=30, y=12

Donahue and Sabo (2021)

- Continuous Outcomes: Normal-Normal Conjugate Pair

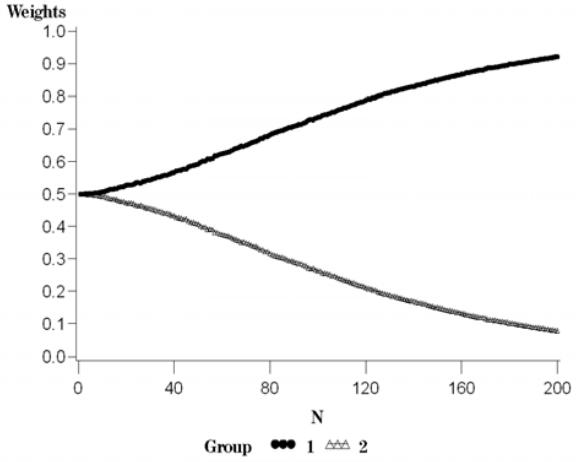
$$
\begin{aligned}
y_k &\sim N(\mu_k, \sigma_k^2) \\
\text{DIP: } \mu_k|\sigma_k &\sim N(\mu_k, \sigma_k^2/(N-n)) \\
\text{DIP: } \sigma_k^2 &\sim IG(N-n, \sigma_o^2) \\
\mu_0 &\sim N(0, 1000)
\end{aligned}
$$

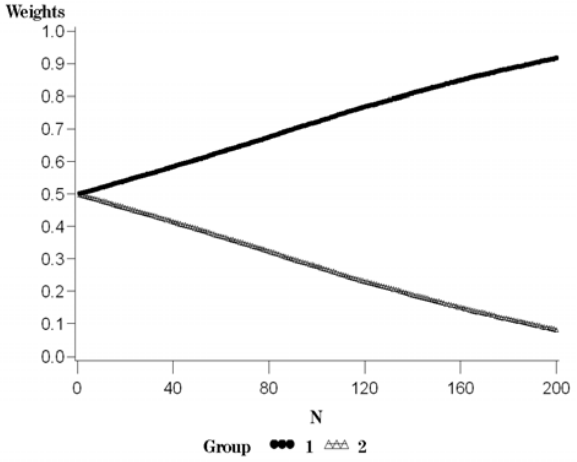
Sabo (2014); Donahue and Sabo (2021)

- Can plug Posterior DIP estimate directly into Rosenberger (2001) optimal designs
- Allocation weights:

$$w_A = \frac{\sqrt{p_A}}{\sqrt{p_A} + \sqrt{p_A}}, w_B = 1 - w_A$$

p_j estimated with poster mean from DIP model





- Two Group Case (DIP with Point Mass)
- True Efficacy: $p_1 = 0.5$ and $p_2 = 0.3$, $N = 200$

	$p_1 = 0.25, p_2 = 0.1, R = 1.58$ $N = 200$			$p_1 = 0.55, p_2 = 0.4, R = 1.17$ $N = 352$		
	Bal.	NIB	DIP	Bal.	NIB	DIP
Exp. Succ.	35.1	36.5	36.1	167.2	167.9	167.8
	(3.85)	(4.28)	(4.00)	(7.86)	(8.92)	(8.57)
$\hat{n}_1$	100.2	110.6	105.3	175.7	183.1	181.4
	(7.13)	(7.81)	(6.89)	(9.47)	(10.23)	(9.65)
$\hat{n}_2$	99.8	89.4	94.7	176.3	168.9	170.6
	(7.13)	(7.81)	(6.89)	(9.47)	(10.23)	(9.65)
Power	80.0%	81.2%	80.3%	80.0%	81.3%	80.1%
R_{50}	–	1.24 (0.15)	1.06 (0.04)	–	1.08 (0.04)	1.05 (0.03)
R_{75}	–	1.40 (0.22)	1.14 (0.07)	–	1.13 (0.06)	1.10 (0.04)
R_{100}	–	1.58 (0.30)	1.53 (0.25)	–	1.17 (0.07)	1.17 (0.07)

Thall and Simon (1994AB)

- Bayesian Early Termination in Phase II Trials
  - One-Sample

$$H_0: \quad \theta = \theta_0$$
$$H_A: \quad \theta = \theta_0 + \delta$$

  - Two-Sample

$$H_0: \quad \theta_1 = \theta_2$$
$$H_A: \quad \theta_1 = \theta_2 + \delta$$

  - Decision Rule: Stop study/trial for

Superiority: $P(\theta > \theta_0 + \delta | y, \pi(\theta)) > p_s$
$P(\theta_1 > \theta_2 + \delta | y, \pi(\theta)) > p_s$

Futility: $P(\theta > \theta_0 + \delta | y, \pi(\theta)) < p_f$
$P(\theta_1 > \theta_2 + \delta | y, \pi(\theta)) < p_f$

Wang et al (2022)

- Altered Model for Decision Rule: Stop study/trial for

$$\text{Superiority:} \quad P(\theta > \theta_0 + \delta | y, \pi(\theta|\theta_0, N, n)) > p_s$$
$$P(\theta_1 > \theta_2 + \delta | y, \pi(\theta|\theta_0, N, n)) > p_s$$

$$\text{Futility:} \quad P(\theta > \theta_0 + \delta | y, \pi(\theta|\theta_0, N, n)) < p_f$$
$$P(\theta_1 > \theta_2 + \delta | y, \pi(\theta|\theta_0, N, n)) < p_f$$

- Where $\pi(\theta|\theta_0, N, n)$ is now a DIP
  - Forces posterior probabilities close to 0.5 at start of trial
  - Gradually allows probabilities to adapt as trial continues

Wang et al (2022)

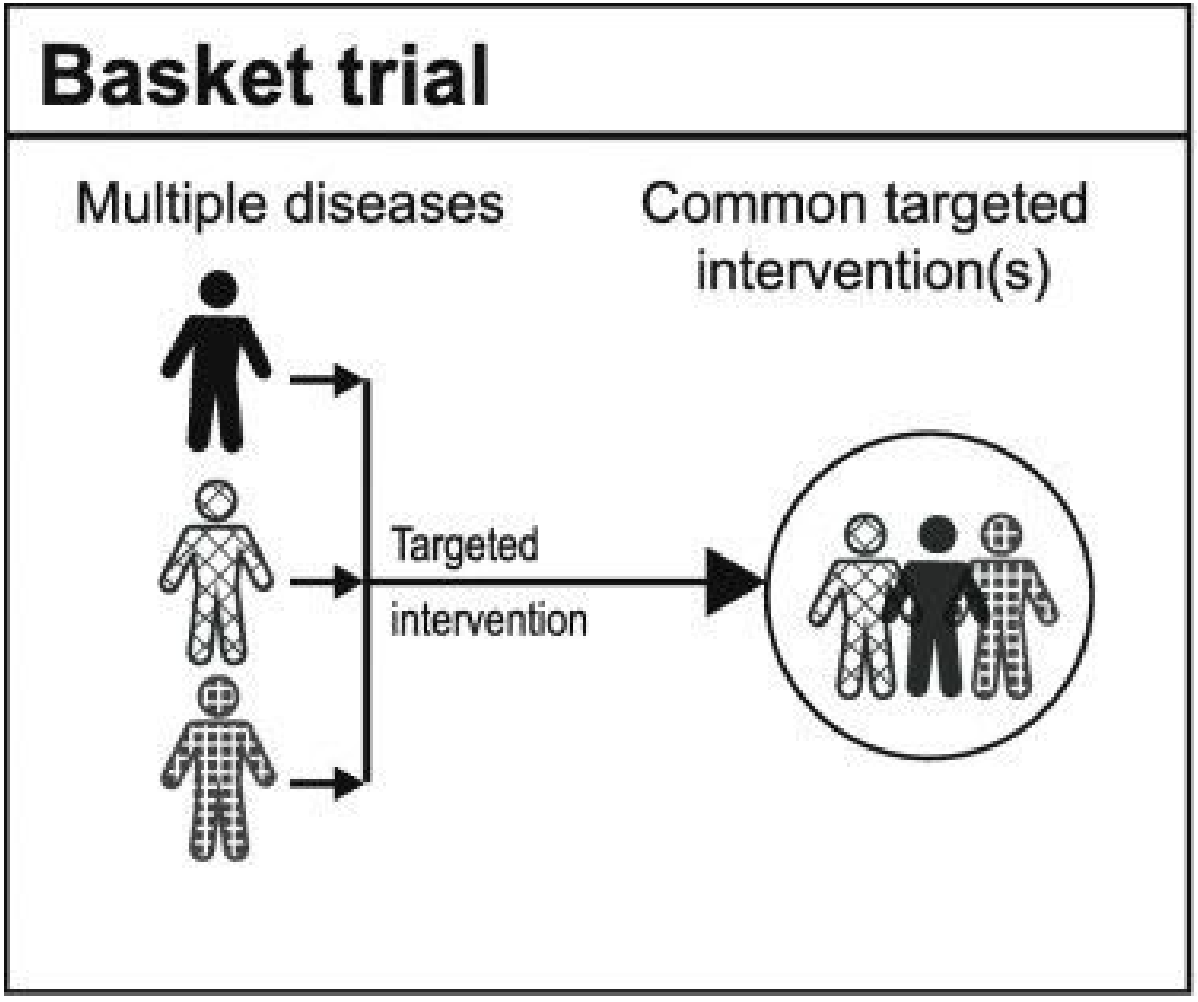
- Goal: Search for minimum viable sample size for each design
  - Power $\geq 80\%$
  - Type I Error Rate $\leq 5\%$ under H_0
  - Searchable Parameters: N, p_s, p_f
- Prior Information for Bayesian Approach
  - Non-Informative $(a, b = 1)$, $a + b = 2, 6, 10$
  - Hyperprior: $p_0 \sim \text{beta}(1, 1)$
- Trial Simulation
  - Subjects recruited until trial is stopped (futility or efficacy) or planned sample size N is reached
  - Upper (efficacy) and lower (futility) decision boundaries set to $p_s \in (0.80, 0.99)$ and $p_f \in (0.01, 0.10)$ respectively
  - Total sample size: $N \in (10, 100)$ and $N_j \in (10, 100)$

One-Sample Results

Model	p_0	p_1	Sample Size	Efficacy Rate	Type I Error
DIP	0.1	0.20	76	0.98	0.050
Bayes $(Beta(1,1))$	0.1	0.20	88	0.99	0.076
Bayes $(a+b=2)$	0.1	0.20	85	0.99	0.050
Bayes $(a+b=6)$	0.1	0.20	90	0.99	0.051
Bayes $(a+b=10)$	0.1	0.20	86	0.98	0.050
DIP	0.1	0.30	22	0.98	0.050
Bayes $(Beta(1,1))$	0.1	0.30	24	0.99	0.052
Bayes $(a+b=2)$	0.1	0.30	25	0.97	0.050
Bayes $(a+b=6)$	0.1	0.30	25	0.96	0.050
Bayes $(a+b=10)$	0.1	0.30	25	0.95	0.050

Two-Sample Results

Model	p_1	p_2	Sample Size	Efficacy Rate	Type I Error
DIP	0.45	0.3	196	0.93	0.111
Bayes ($Beta(1,1)$)	0.45	0.3	199	0.97	0.146
Bayes ($a+b=2$)	0.45	0.3	190	0.97	0.162
Bayes ($a+b=6$)	0.45	0.3	185	0.96	0.136
Bayes ($a+b=10$)	0.45	0.3	198	0.96	0.125
DIP	0.50	0.3	168	0.97	0.050
Bayes ($Beta(1,1)$)	0.50	0.3	192	0.99	0.061
Bayes ($a+b=2$)	0.50	0.3	197	0.99	0.053
Bayes ($a+b=6$)	0.50	0.3	178	0.99	0.050
Bayes ($a+b=10$)	0.50	0.3	198	0.99	0.052



- Can "borrow" information across sub-groups to gain power
- Can also assess sub-groups separately
- Early Termination for superiority or futility

Analytic Framework: Simon et al (2016)

- Stratum K Efficacy (Assuming Heterogeneity)

$$
\begin{aligned}
P_{k,ind} &= \frac{\gamma \times \left(p_A^{y_k}\left(1-p_A\right)^{n_k-y_k}\right)}{\gamma \times \left(p_A^{y_k}\left(1-p_A\right)^{n_k-y_k}\right) + (1-\gamma) \times \left(p_0^{y_k}\left(1-p_0\right)^{n_k-y_k}\right)} \\
K &= \text{The number of groups} \\
(p_A,\, p_0) &= \text{Success rate under alternative and null} \\
(y_k,\, n_k) &= \text{Number of success and total participants in group K} \\
\gamma &= \text{Pre-specified overall probability of effectiveness}
\end{aligned}
$$

- Overall Efficacy (Pooled Data)

$$
P_0 = \frac{\gamma \times \Pi_{k=1}^{K} p_A^{y_k}\left(1-p_A\right)^{n_k-y_k}}{\gamma \times \Pi_{k=1}^{K} p_A^{y_k}\left(1-p_A\right)^{n_k-y_k} + (1-\gamma) \times \Pi_{k=1}^{K} p_0^{y_k}\left(1-p_0\right)^{n_k-y_k}}
$$

Analytic Framework: Simon et al (2016)

- Posterior Probability that therapy is active in stratum K

$$\begin{aligned} P_k &= \pi \times P_0 + (1-\pi) P_{k,ind} \\ \pi &= \text{Estimate of degree of homogeneity and function of } \lambda \\ \lambda &= \text{Pre-specified degree of correlation between strata} \end{aligned}$$

- By this framework
  - As $\pi \to 1$: evidence for homogeneity between strata is high, then estimate pools information across strata.
  - As $\pi \to 0$: evidence for homogeneity between strata is low, then estimate does not pool information across strata.

Our DIP-Based Approach

- Estimate γ instead of using pre-specification

$$\begin{aligned} \gamma_k &= P(p_k \geq p_A | n_k, y_k, n, N, p_0) \\ \text{DIP:} & \\ \pi(p_k | p_0, n, N) &\sim \text{Beta}(1 + (N-n)p_0, 1 + (N-n)(1-p_0)) \end{aligned}$$

- By this set-up
  - Prior Mode is $p_0 \Rightarrow$ Assumes therapy does not work
  - γ_k will be low early in trial
  - Reduce False Positive Rate & Early Stopping Rate

Model Comparisons

True Pos.	False Neg	False Pos	True Neg	Sample Size	Sens	Spec
			Original Approach			
0	0	0.29	2.71	17.6	na	0.90
0.69	0.31	0.27	1.73	19.5	0.69	0.86
1.55	0.45	0.20	0.80	18.8	0.77	0.80
2.58	0.42	0	0	14.8	0.86	na
			DIP Approach			
0	0	0.06	2.82	35.4	na	0.98
0.74	0.17	0.04	1.86	38.7	0.81	0.98
1.52	0.25	0.06	0.92	38.5	0.86	0.94
2.46	0.36	0	0	33.2	0.87	na

$$n = 20 \text{ Per Group}, \; P_0 = 0.05, \; P_A = 0.25$$

Model Comparisons

True Pos.	False Neg	False Pos	True Neg	Sample Size	Sens	Spec
			Original Approach			
0	0	0.21	1.39	54.9	na	0.87
0.32	0.16	0.21	0.73	56.9	0.67	0.78
0.79	0.21	0.16	0.28	56.6	0.79	0.65
1.46	0.21	0	0	53.8	0.88	na
			DIP Approach			
0	0	0.11	2.77	29.6	na	0.96
0.54	0.34	0.09	1.78	35.0	0.62	0.95
1.13	0.55	0.08	0.85	38.2	0.67	0.92
1.96	0.68	0	0	37.9	0.75	na

$n = 20$ Per Group, $P_0 = 0.2$, $P_A = 0.4$

Summary

- The DIP Approach:
  - Create skeptical and informative prior(s) based on null hypothesis
  - Prior(s) decrease in information as trial continues
  - "Tempers" extreme outcomes in early stages of trials
- In Outcome-Adaptive Allocation:
  - Modest improvement in successes vs. balanced case
  - More stable variability than other approaches
- Early Trial Termination/Basket Trials:
  - Can improve error rates and sample sizes in some cases
  - Performance dependent upon means

Acknowledgements

- Colleagues
  - Ghalib Bello, Ph.D., Ichan School of Medicine at Mount Sinai
  - Erin Donahue, Ph.D., Levine Cancer Institute
  - Chen Wang, Ph.D., Medpace
  - Darya Shokouhi Niaki, Doctoral Trainee

- This work was supported in part by:
  - Biostatistics Shared Resource of the VCU Massey Comprehensive Cancer Center; NIH-NCI Cancer Center Support Grant P30 CA016059.
  - Biostatistics, Epidemiology and Research Design (BERD) Core; C. Kenneth and Dianne Wright Regional Center for Clinical and Translational Science (CTSA Award #*UM*1*TR*004360

References

- Rosenberger WF, Stallard N, Ivanova A, Harper CN, Ricks ML (2001). Optimal adaptive designs for binary response trials. *Biometrics* 57: 909–913
- Thall PF, Wathen JK (2007). Practical Bayesian adaptive randomization in clinical trials. *European Journal of Cancer* 43(5): 859â866
- Haines LM, Sadiq H (2015). Start-up designs for response-adaptive randomization procedures with sequential estimation. *Statistics in Medicine* 34(21): 2958-â2970
- Bello G, Sabo R (2016). Outcome-adaptive allocation with natural lead-in for three-group trials with binary outcomes. *Journal of Statistical Computation and Simulation*. 86(12): 2441-â2449
- Sabo RT (2014). Adaptive allocation for binary outcomes using decreasingly informative priors. *Journal of Biopharmaceutical Statistics*. 24(3): 569-â578
- Neuenschwander B, Weber S, Schmidli H, OâHagan A (2020). Predictively consistent prior effective sample sizes. *Biometrics* 76: 578-â587
- Donahue E, Sabo RT (2021). A natural lead-in approach to response-adaptive allocation for continuous outcomes. *Pharmaceutical Statistics*. 20(3): 563-572

References

- Thall PF, Simon R (1994A). Practical bayesian guidelines for phase iib clinical trials. *Biometrics*, 50(2):337â-349
- Thall PF, Simon Rb(1994B). A bayesian approach to establishing sample size and monitoring criteria for phase II clinical trials. *Controlled Clinical Trials*, 15(6):463â-481
- Wang C, Sabo RT, Mukhopadhyay N, Perera R (2022). Early termination in single-parameter model phase II clinical trial designs using decreasingly informative priors. *International Journal of Clinical Trials*. 9(2):107–117
- Simon R, Geyer S, Subramanian J, Roychowdhury S (2016). The Bayesian basket design for genomic variant-driven phase II trials. *Seminars in Oncology*. 43: 13–18

Thank You

Questions?