Advanced Strategies in Clinical Trials: Optimal Design for Multiple Dose and Adaptive Information Borrowing with SAM

Presenter: Peng Yang

Sep 26, 2024

Contents

Background – Oncology trial design

Background – Oncology trial design

Traditional Paradigm:

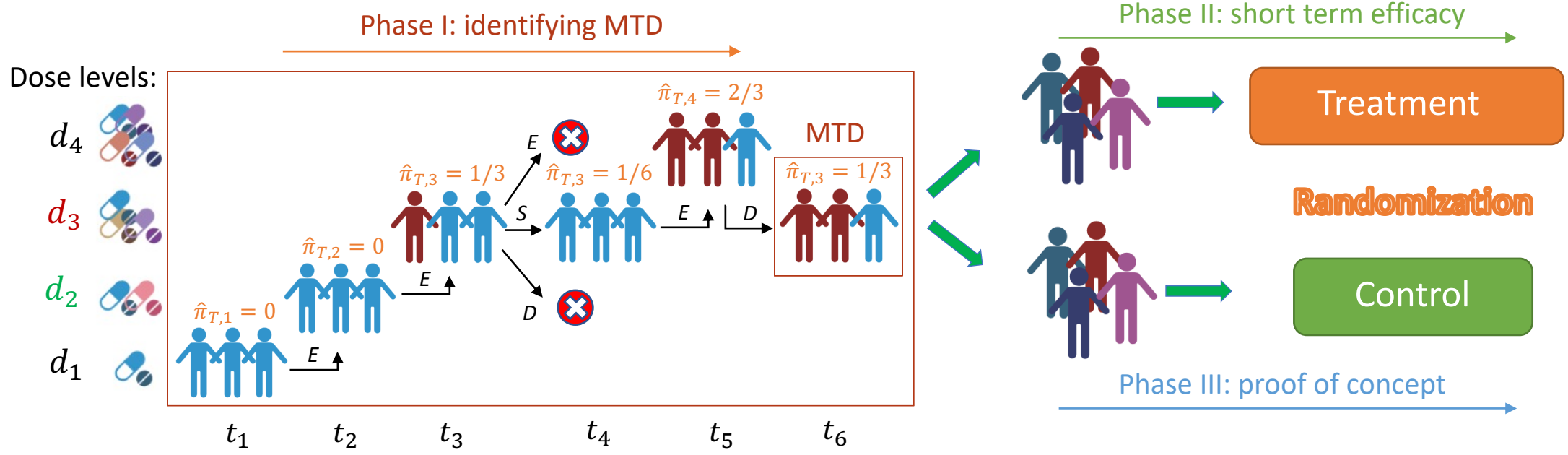
Phase I: **Safety** MTD -> RP2D → Phase II: **Short-term Efficacy** ORR → Phase III: **Long-term Efficacy** PFS/OS

MTD - Maximum tolerated dose

ORR – Objective response rate

RP2D – Recommended phase II dose

Endpoint: Binary | Binary/continuous | Survival

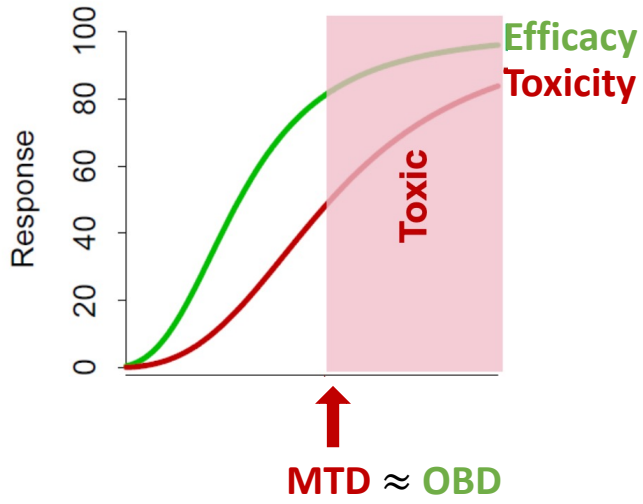


Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

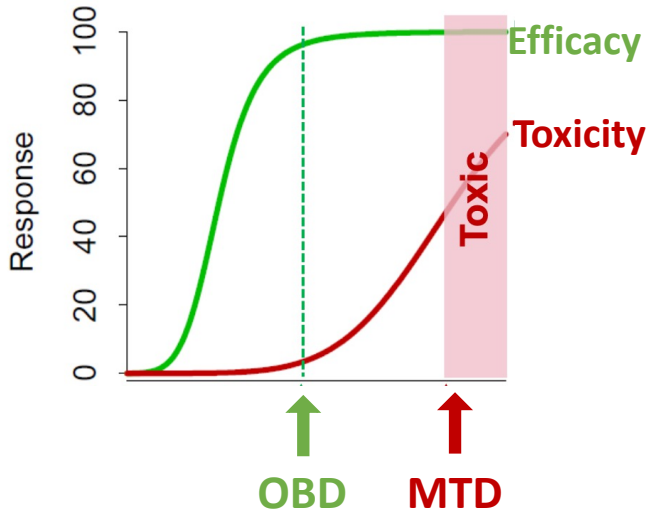
Background and Literature Survey

- In 2022, FDA initiated *Project Optimus* "to reform the dose optimization and dose selection paradigm in oncology drug development."
- Paradigm shifting from maximum tolerated dose (MTD) to optimal biological dose (OBD).

Conventional Chemotherapy



Targeted Therapies



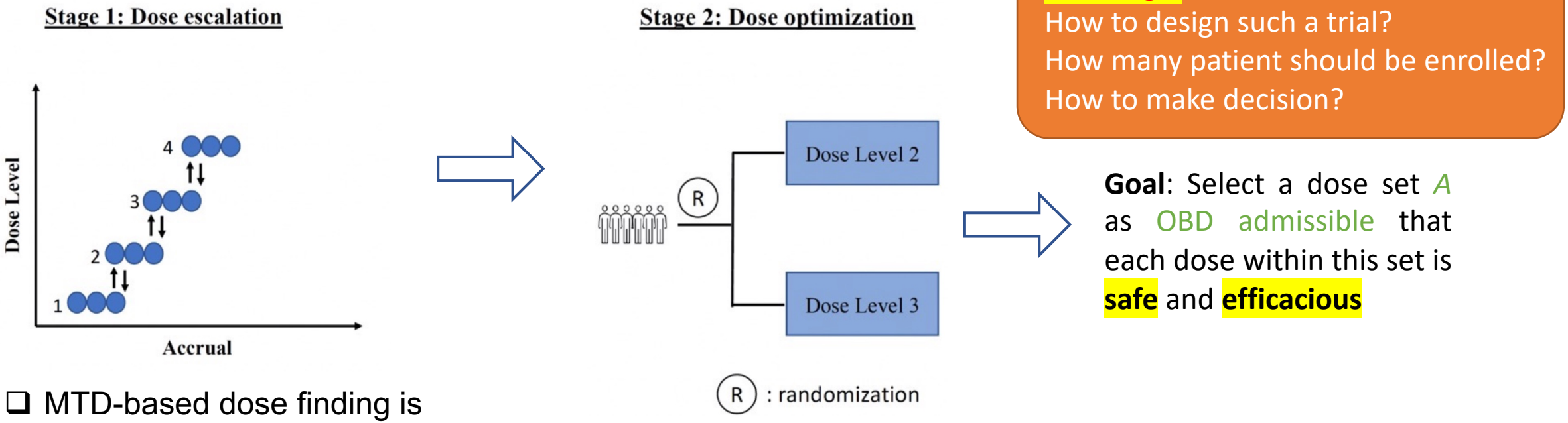
- MTD-based dose finding is often appropriate to inform RP2D
- Safety alone is not sufficient to inform optimal RP2D

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Background – Design strategies to find OBD

- Trial designs to compare multiple dosages (FDA, 2023)
  - The trial should be sized to allow for sufficient assessment of activity, safety, and tolerability for each dosage.
  - The trial does not need to be powered to demonstrate statistical superiority of a dosage or statistical non-inferiority among the dosages.

Challenge:
How to design such a trial?
How many patient should be enrolled?
How to make decision?



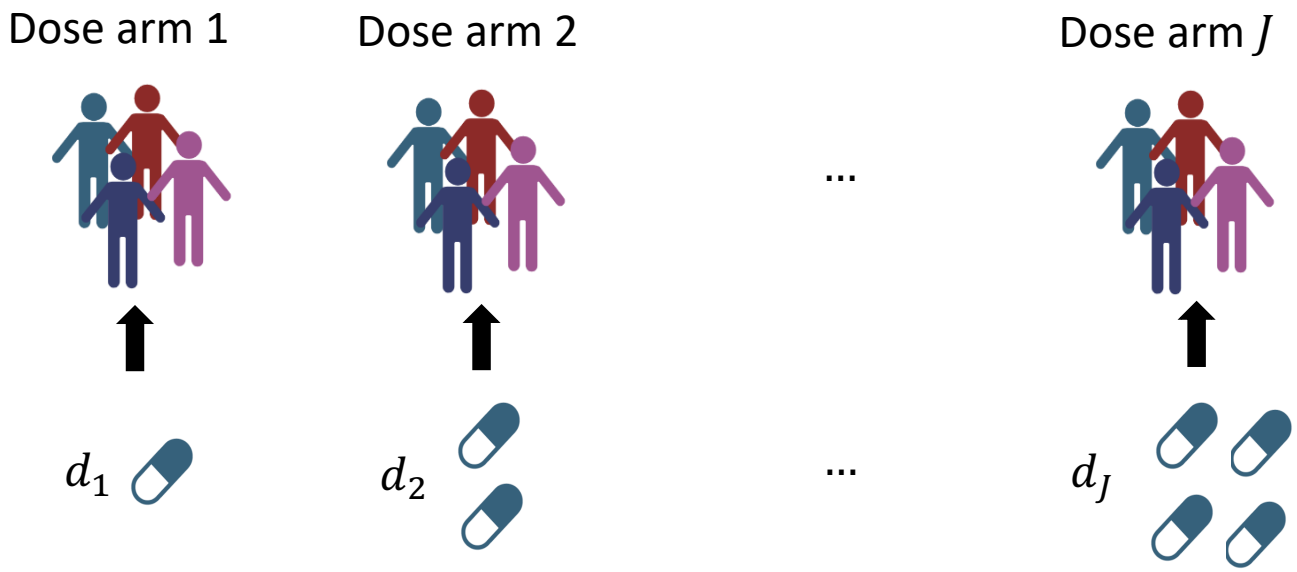
Goal: Select a dose set *A* as OBD admissible that each dose within this set is **safe** and **efficacious**

- MTD-based dose finding is often appropriate

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Set ups

- Consider a multiple-dose randomized trial, where a total of $J \times n$ patients are equally randomized to J doses, with with $d_1 < d_2 < \cdots < d_J$.
  - The In most applications, J = 2 or 3, and d_J is often the MTD or maximum administered dose.



- Y_T and Y_E denote the binary toxicity and efficacy endpoints, respectively.
  - Example of Y_T: dose-limiting toxicity, dichotomized total toxicity burden.
  - Example of Y_E: objective responses, efficacy surrogate endpoints.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Assumptions

- Let $\pi_{T,j} = \Pr(Y_T = 1|d_j)$ and $\pi_{E,j} = \Pr(Y_E = 1|d_j)$ denotes the probability of the occurrence of toxicity and efficacy events.
- Assuming that $\pi_{T,j}$ and $\pi_{E,j}$ are **non-decreasing** with respect to the increasing of dose levels.
  - Randomized dose optimization trials with same drug but with **ordered doses.**
- For toxicity endpoint, we assume
  - $\phi_{T,0}$: the **null toxicity** rate that is high and deemed **unacceptable**;
  - $\phi_{T,1}$: the **alternative toxicity** rate that is low and deemed **acceptable**;
- For efficacy endpoint, we assume
  - $\phi_{E,0}$: the **null efficacy** rate that is low and deemed **unacceptable**;
  - $\phi_{E,1}$: the **alternative efficacy** rate that is high and deemed **acceptable**

Challenge:
How to modeling joint toxicity and efficacy data?
How to account for the fact that doses are ordered

OBD admissible

$\pi_{T,j}$ / $\pi_{E,j}$	$\phi_{T,0}$	$\phi_{T,1}$
$\phi_{E,0}$	✗	✗
$\phi_{E,1}$	✗	✓

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Global type I error

- H_0: None of the doses is the OBD. **Challenge**: consists of multiple hypotheses!

$$H_0(s,k): \underbrace{\pi_{T,1}=\pi_{T,2}=\cdots=\pi_{T,s}=\phi_{T,1}}_{\text{acceptable toxicity}} < \underbrace{\pi_{T,s+1}=\cdots=\pi_{T,k}=\pi_{T,k+1}=\cdots=\pi_{T,J}=\phi_{T,0}}_{\text{unacceptable toxicity}};$$

$$\underbrace{\pi_{E,1}=\pi_{E,2}=\cdots=\pi_{E,s}=\pi_{E,s+1}=\cdots=\pi_{E,k}=\phi_{E,0}}_{\text{unacceptable efficacy}} < \underbrace{\pi_{E,k+1}=\cdots=\pi_{E,J}=\phi_{E,1}}_{\text{acceptable efficacy}},$$

Safe but futile | Futile and overly toxic | Efficacious but overly toxic

where $s, k \in \{0,1,\ldots,J\}$ with $s \le k$.

		d_1	d_2			d_1	d_2
$H_0(0,0)$	Tox	$\phi_{T,0}$	$\phi_{T,0}$	$H_0(0,0)$	Tox	$\phi_{T,1}$	$\phi_{T,0}$
	Eff	$\phi_{E,1}$	$\phi_{E,1}$		Eff	$\phi_{E,0}$	$\phi_{E,1}$
$H_0(0,1)$	Tox	$\phi_{T,0}$	$\phi_{T,0}$	$H_0(0,1)$	Tox	$\phi_{T,1}$	$\phi_{T,0}$
	Eff	$\phi_{E,0}$	$\phi_{E,1}$		Eff	$\phi_{E,0}$	$\phi_{E,0}$
$H_0(0,2)$	Tox	$\phi_{T,0}$	$\phi_{T,0}$	$H_0(0,2)$	Tox	$\phi_{T,1}$	$\phi_{T,1}$
	Eff	$\phi_{E,0}$	$\phi_{E,0}$		Eff	$\phi_{E,0}$	$\phi_{E,0}$

OBD admissible

$\pi_{E,j}$ \ $\pi_{T,j}$	$\phi_{T,0}$	$\phi_{T,1}$
$\phi_{E,0}$	✗	✗
$\phi_{E,1}$	✗	✓

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Global type I error

- H_0: None of the doses is the OBD. **Challenge**: consists of multiple hypotheses!

$$H_0(s,k): \underbrace{\pi_{T,1}=\pi_{T,2}=\cdots=\pi_{T,s}=\phi_{T,1}}_{\text{acceptable toxicity}} < \underbrace{\pi_{T,s+1}=\cdots=\pi_{T,k}=\pi_{T,k+1}=\cdots=\pi_{T,J}=\phi_{T,0}}_{\text{unacceptable toxicity}};$$

$$\underbrace{\pi_{E,1}=\pi_{E,2}=\cdots=\pi_{E,s}=\pi_{E,s+1}=\cdots=\pi_{E,k}=\phi_{E,0}}_{\text{unacceptable efficacy}} < \underbrace{\pi_{E,k+1}=\cdots=\pi_{E,J}=\phi_{E,1}}_{\text{acceptable efficacy}},$$

Safe but futile | Futile and overly toxic | Efficacious but overly toxic

where $s, k \in \{0,1,\ldots,J\}$ with $s \le k$, and

- Define *global type I error* that encompasses all $H_0(s,k)$ as follows:

$$\alpha = \Pr(\text{reject } H_0 | H_0)$$

$$= \max_{0\le s\le J,\ s\le k\le J} \{\alpha(s,k)\}.$$

where $\alpha(s,k) = \Pr(\text{reject } H_0(s,k) | H_0(s,k))$.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Generalized powers

- Consider H_1: At lease one does is the OBD admissible
  - **challenge**: Consists of multiple hypotheses

$$H_1(u,v): \underbrace{\pi_{T,1}=\pi_{T,2}=\cdots=\pi_{T,u}=\pi_{T,u+1}=\cdots=\pi_{T,v}=\phi_{T,1}}_{\text{acceptable toxicity}} < \underbrace{\pi_{T,v+1}=\cdots=\pi_{T,J}=\phi_{T,0}}_{\text{unacceptable toxicity}};$$

$$\underbrace{\pi_{E,1}=\pi_{E,2}=\cdots=\pi_{E,u}=\phi_{E,0}}_{\text{unacceptable efficacy}} < \underbrace{\pi_{E,u+1}=\cdots=\pi_{E,v}=\pi_{E,v+1}=\cdots=\pi_{E,J}=\phi_{E,1}}_{\text{acceptable efficacy}},$$

Safe but futile | Safe and efficacious | Efficacious but overly toxic

where $u, v \in \{0,1,2,\ldots,J\}$ with $u < v$.

: OBD admissible

		d_1	d_2
$H_1(0,1)$	Tox	$\phi_{T,1}$	$\phi_{T,0}$
	Eff	$\phi_{E,1}$	$\phi_{E,1}$
$H_1(0,2)$	Tox	$\phi_{T,1}$	$\phi_{T,1}$
	Eff	$\phi_{E,1}$	$\phi_{E,1}$
$H_1(1,2)$	Tox	$\phi_{T,1}$	$\phi_{T,1}$
	Eff	$\phi_{E,0}$	$\phi_{E,1}$

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Generalized powers

- Consider H_1: At lease one does is the OBD admissible
  - **challenge**: the standard definition of power, which reject the H_1, is not sufficient to account for the characteristics of dose optimization

$$H_1(u,v): \underbrace{\pi_{T,1}=\pi_{T,2}=\cdots=\pi_{T,u}=\pi_{T,u+1}=\cdots=\pi_{T,v}=\phi_{T,1}}_{\text{acceptable toxicity}} < \underbrace{\pi_{T,v+1}=\cdots=\pi_{T,J}=\phi_{T,0}}_{\text{unacceptable toxicity}};$$

$$\underbrace{\pi_{E,1}=\pi_{E,2}=\cdots=\pi_{E,u}=\phi_{E,0}}_{\text{unacceptable efficacy}} < \underbrace{\pi_{E,u+1}=\cdots=\pi_{E,v}=\pi_{E,v+1}=\cdots=\pi_{E,J}=\phi_{E,1}}_{\text{acceptable efficacy}},$$

Safe but futile | Safe and efficacious | Efficacious but overly toxic

where $u, v \in \{0,1,2,\ldots,J\}$ with $u < v$.

: OBD admissible

		d_1	d_2
$H_1(0,1)$	Tox	$\phi_{T,1}$	$\phi_{T,0}$
	Eff	$\phi_{E,1}$	$\phi_{E,1}$
Selection		✓	
			✓
		✓	✓

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Generalized powers

- *Generalized power I*:

$$\beta_1(u,v) = \Pr\left(\text{reject } H_0 \text{ and all does in } A \text{ are truly safe and efficacious} \;\middle|\; H_1(u,v)\right).$$

where *A* denotes the admissible dose set selected by the design.

- *Generalized power II*:

$$\beta_2(u,v) = \Pr\left(\text{reject } H_0 \text{ and at leaset one doe in } A \text{ are truly safe and efficacious} \;\middle|\; H_1(u,v)\right).$$

	β_1	β_2
Tolerate false positive	✗	✓

		d_1	d_2
$H_1(0,1)$	Tox	$\phi_{T,1}$	$\phi_{T,0}$
	Eff	$\phi_{E,1}$	$\phi_{E,1}$
Selection		✓	
			✓
		✓	✓

- Both generalized powers are stricter than the standard power.
- The choice of power depends on the characteristics of the trial and user's tolerability of false positive.
- Power II is a good option when reducing the sample size is of top priority.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Research Plans and Methodology: Generalized powers

Theorem 1 *Define the least favorable set* $\widetilde{H}_1 = \{H_1(j), j = 1, \cdots, J\}$, *where*

$$H_1(j) = \left(\begin{array}{ccc} \underbrace{\pi_{T,1} = \cdots = \pi_{T,j-1} = \phi_{T,1}}_{\textit{safe but futile}} & \underbrace{\pi_{T,j} = \phi_{T,1}}_{\textit{safe and efficacious}} & \underbrace{\pi_{T,j+1} = \cdots = \pi_{T,J} = \phi_{T,0}}_{\textit{toxic and efficacious}} \\ \underbrace{\pi_{E,1} = \cdots = \pi_{E,j-1} = \phi_{E,0}} & \underbrace{\pi_{E,j} = \phi_{E,1}} & \underbrace{\pi_{E,j+1} = \cdots = \pi_{E,J} = \phi_{E,1}} \end{array} \right).$$

For any $H_1(u, v)$, *with* $u, v \in \{0, 1, 2, \ldots, J\}$ *and* $u < v$, *there exists an* $H_1(j)$ *such that* $\beta_i(j) \leq \beta_i(u, v)$, $i = 1, 2$, *where* $\beta_1(j)$ *and* $\beta_2(j)$ *denote the generalized power I and II under* $H_1(j)$, *respectively.*

Global power I and II can be minimized as follows:

$$\beta_i = \min_{j \in \{1,\ldots,J\}} \beta_i(j) \text{ for } i = 1, 2.$$

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: The MERIT design

- MERIT (Multiple-dosE RandomIzed Phase-II Trial)
  - Specify target global type I error and power α^* and β^*;
  - Randomize $J \times n$ patients equally to J doses;
  - In any dose arm d_j, if $n_{E,j} \geq m_E$ and $n_{T,j} \leq m_T$, we reject H_0 and claim that d_j is OBD admissible, where m_T and m_E are decision boundaries for toxicity and efficacy, respectively.

*$n_{E,j}$ and $n_{T,j}$ are the total number of patients who experience efficacy and toxicity events in dose arm d_j.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Methodology: Determine design parameters

- Given the pre-specified global type I error α^* and global power β_1^* or β_2^*, MERIT design can obtain the optimal design parameters (n, m_T, m_E) through numerical search using the following algorithm:

1. Set $n = 1, \ldots, N$, where N is a large number;
2. Given a value of n, enumerate all possible values of $m_E, m_T \in (0, 1, \ldots, n)$. Given a set of (n, m_T, m_E), calculate the type I error and powers β_1 or β_2;
3. Repeat steps 1 and 2 untie we find the smallest n and corresponding m_E and m_T, such that $\alpha \leq \alpha^*$ and $\beta_1 \geq \beta_1^*$ or $\beta_2 \geq \beta_2^*$.

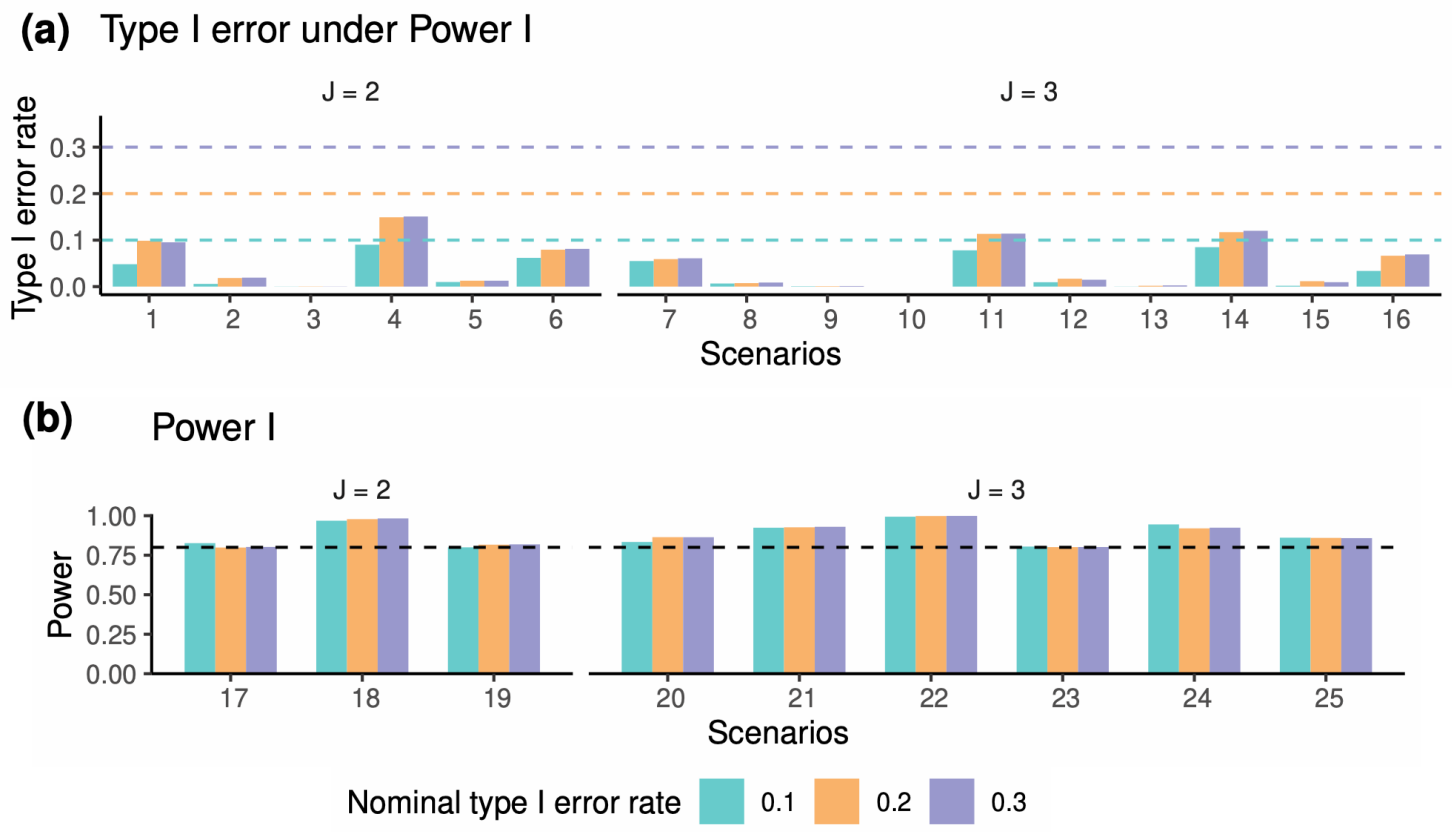
	β_1						β_2					
β^*	$\alpha^* = 0.1$			$\alpha^* = 0.2$			$\alpha^* = 0.1$			$\alpha^* = 0.2$		
	n	m_T	m_E	n	m_T	m_E	n	m_T	m_E	n	m_T	m_E
0.6	26	7	6	23	6	26	25	6	5	18	5	4
0.7	33	9	7	30	8	33	33	8	6	24	7	5
0.8	44	12	8	39	11	44	39	11	8	30	8	5

Optimal design parameters (n, m_T, m_E) of MERIT design, when $J = 2, (\phi_{T,0}, \phi_{T,1}) = (0.4, 0.2)$, and $(\phi_{E,0}, \phi_{E,1}) = (0.2, 0.4)$.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Operating Characteristics: under power I

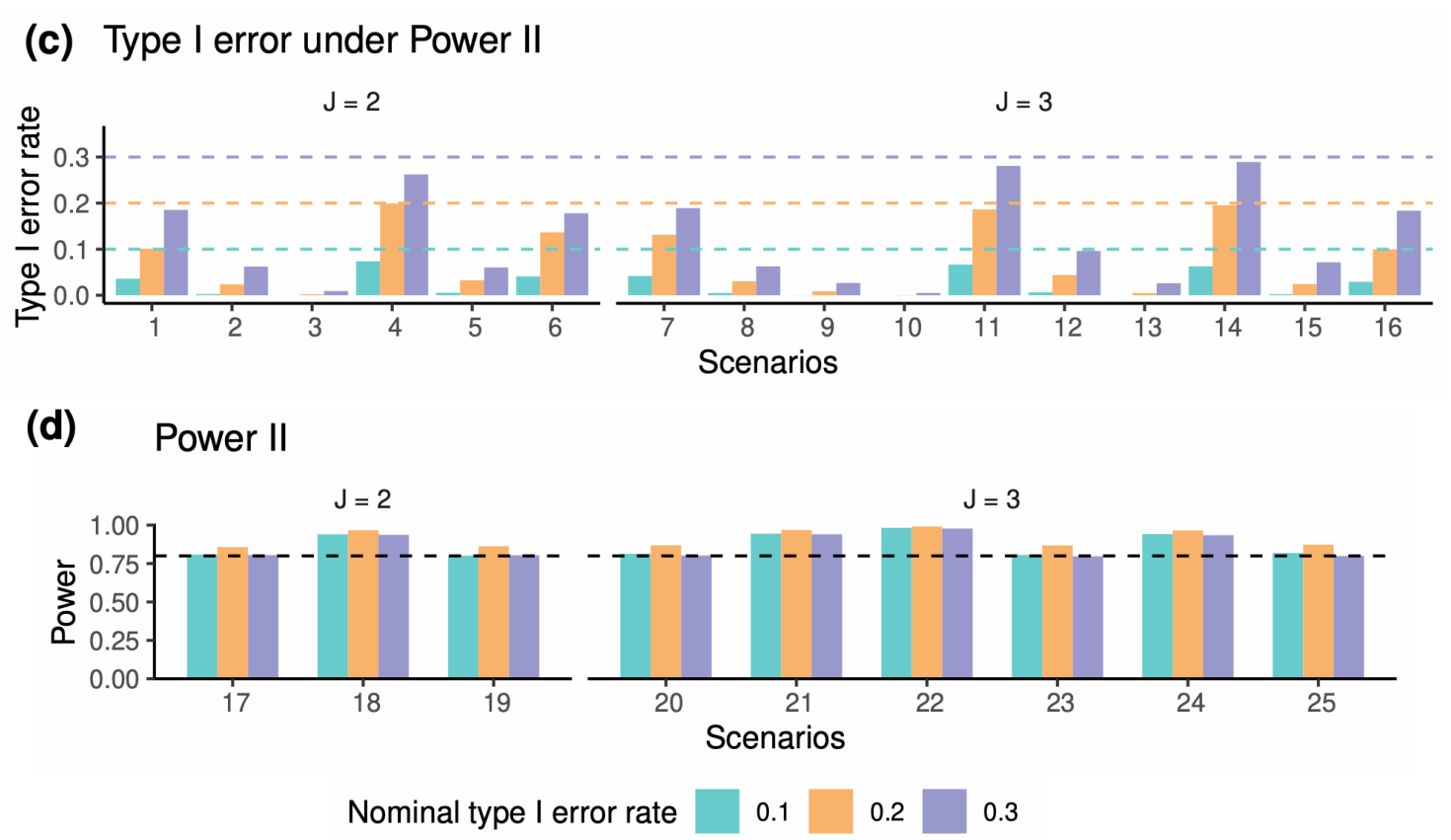
- Type I error and power of MERIT design when $(\phi_{T,0}, \phi_{T,1}) = (0.4, 0.2)$, and $(\phi_{E,0}, \phi_{E,1}) = (0.2, 0.4)$.



Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

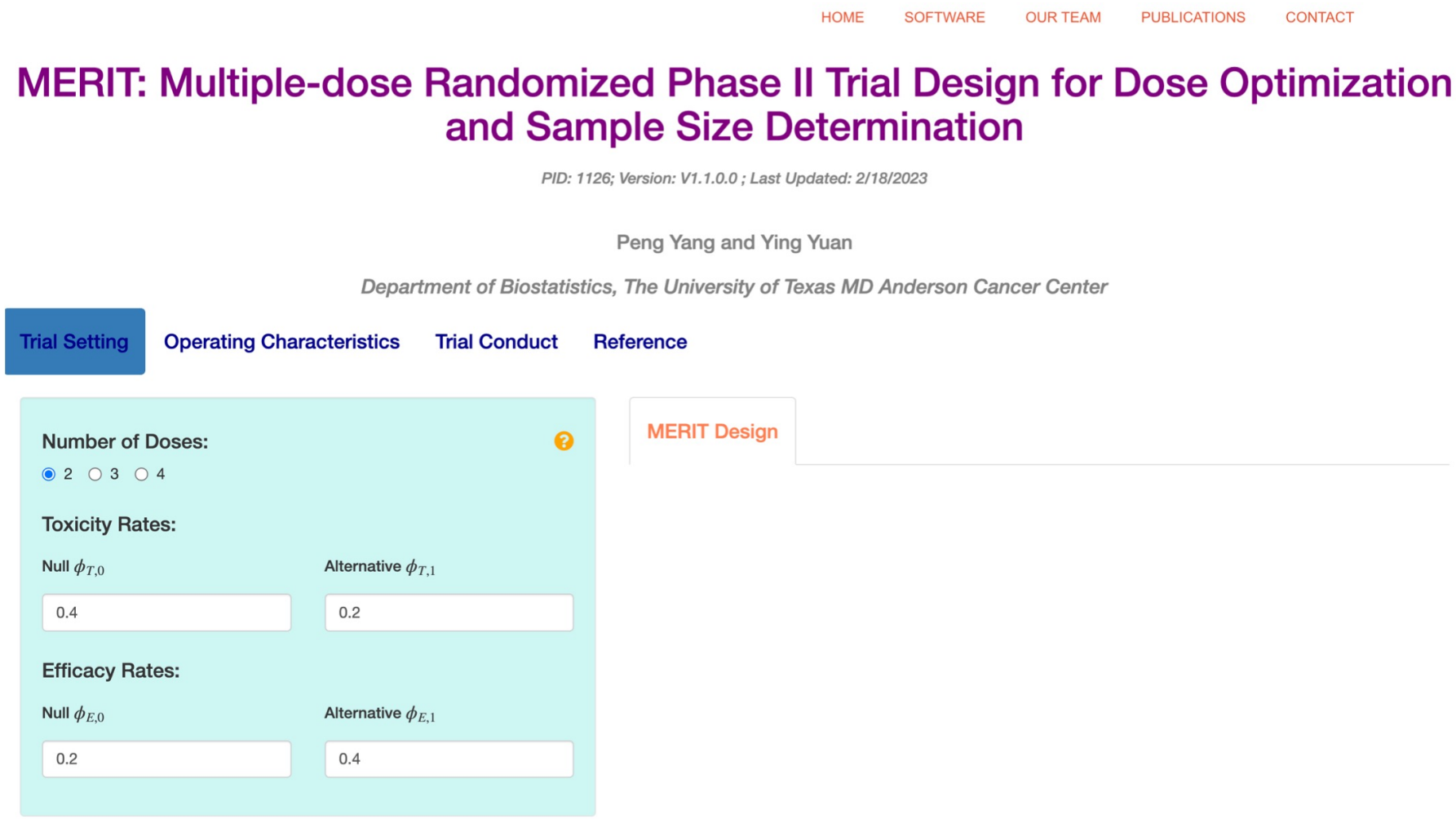
Operating Characteristics: under power II

- Type I error and power of MERIT design when $(\phi_{T,0}, \phi_{T,1}) = (0.4, 0.2)$, and $(\phi_{E,0}, \phi_{E,1}) = (0.2, 0.4)$.



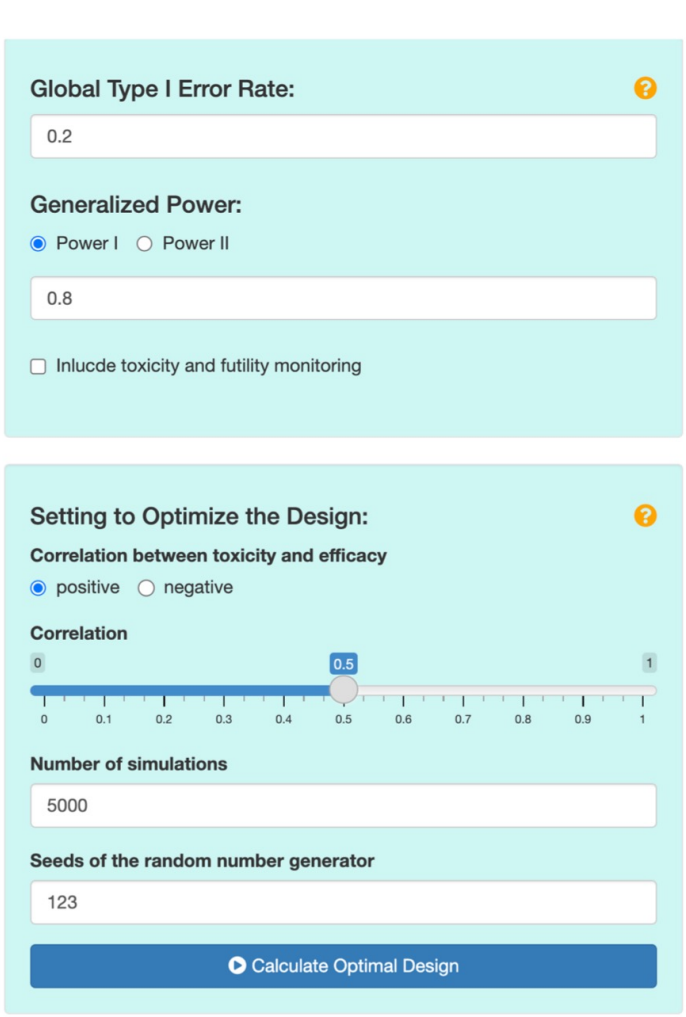
Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

↳ Software: shiny app on www.trialdesign.org



Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Software: shiny app on www.trialdesign.org



Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

↳ Software: shiny app on www.trialdesign.org

MERIT: Multiple-dose Randomized Phase II Trial Design for Dose Optimization and Sample Size Determination

PID: 1126; Version: V1.1.0.0 ; Last Updated: 2/18/2023

Peng Yang and Ying Yuan

Department of Biostatistics, The University of Texas MD Anderson Cancer Center

Trial Setting | Operating Characteristics | Trial Conduct | Reference

Number of Doses:

2 3 4

Toxicity Rates:

Null $\phi_{T,0}$: 0.4 Alternative $\phi_{T,1}$: 0.2

Efficacy Rates:

Null $\phi_{E,0}$: 0.2 Alternative $\phi_{E,1}$: 0.4

MERIT Design

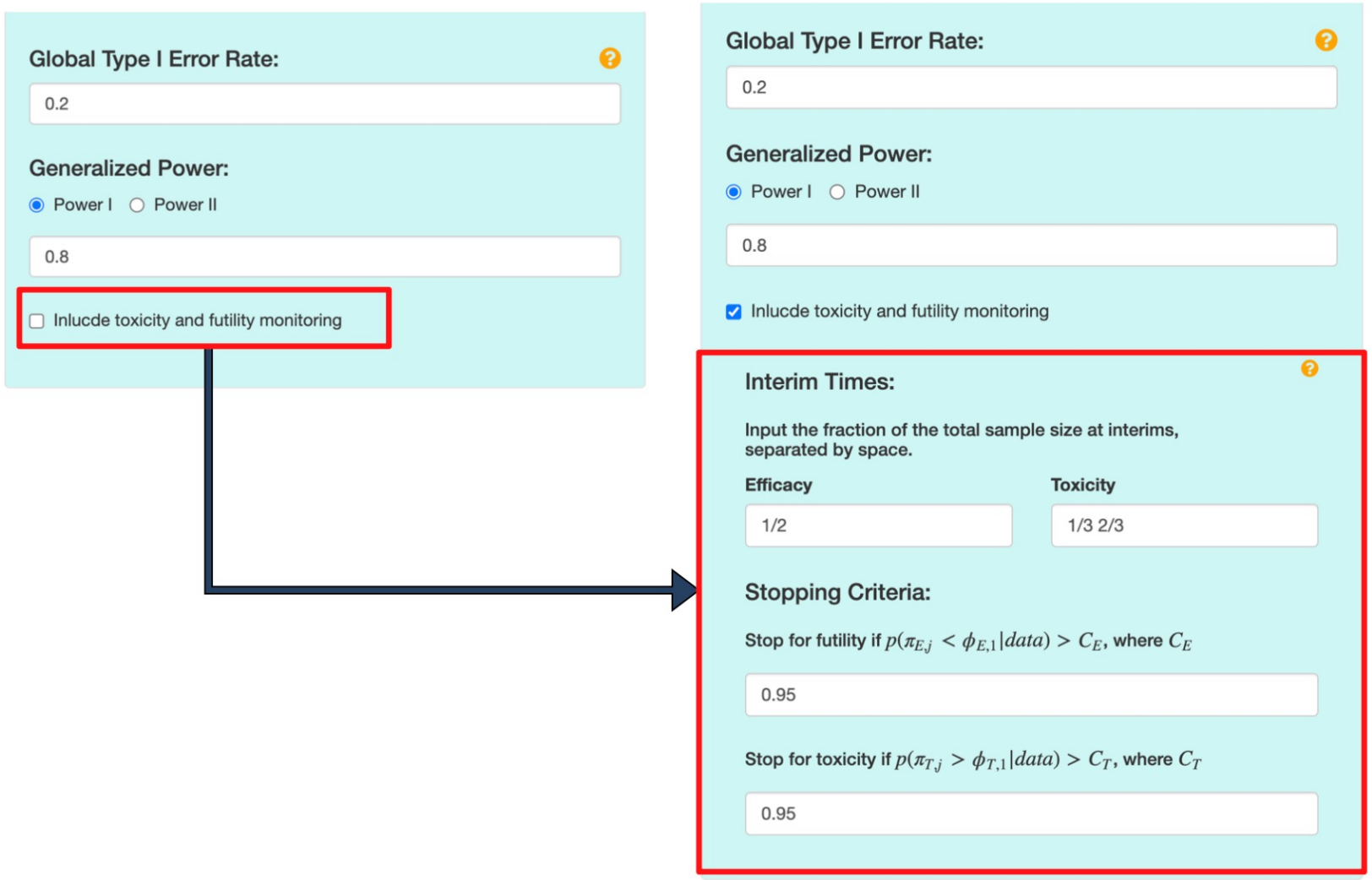
Download MERIT Design

Design Description

In this trial, the toxicity rates of 0.4 and 0.2 are considered unacceptable and acceptable, respectively, while the efficacy rates of 0.2 and 0.4 are considered unacceptable and acceptable, respectively. In order to control the global Type I error rate at 0.2 and achieve a global generalized power I of 0.8, a minimum sample size of 44 per arm is required. The generalized power I is defined as the probability of rejecting the null hypothesis that none of the doses are considered optimal biological doses (OBD) admissible and all doses identified as OBD admissible are truly safe and efficacious given the alternative hypothesis that at least one dose is OBD admissible. At the end of the trial, perform isotonic regression on toxicity and efficacy data across all doses. A dose will be considered OBD admissible if the isotonically transformed number of toxicity <= 13 and the isotonically transformed number of efficacy >= 13.

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

↳ Software: shiny app on www.trialdesign.org



Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Software: shiny app on www.trialdesign.org

Number of Doses:

(•) 2 ( ) 3 ( ) 4

Toxicity Rates:

Null $\phi_{T,0}$: 0.4 Alternative $\phi_{T,1}$: 0.2

Efficacy Rates:

Null $\phi_{E,0}$: 0.2 Alternative $\phi_{E,1}$: 0.4

Global Type I Error Rate: 0.2

Generalized Power:

(•) Power I ( ) Power II

0.8

[x] Inlucde toxicity and futility monitoring

Interim Times:

MERIT Design

Download MERIT Design

Design Description

In this trial, the toxicity rates of 0.4 and 0.2 are considered unacceptable and acceptable, respectively, while the efficacy rates of 0.2 and 0.4 are considered unacceptable and acceptable, respectively. In order to control the global Type I error rate at 0.2 and achieve a global generalized power I of 0.8, a minimum sample size of 45 per arm is required. The generalized power I is defined as the probability of rejecting the null hypothesis that none of the doses are considered optimal biological doses (OBD) admissible and all doses identified as OBD admissible are truly safe and efficacious given the alternative hypothesis that at least one dose is OBD admissible. At the end of the trial, perform isotonic regression on toxicity and efficacy data across all doses. A dose will be considered OBD admissible if the isotonically transformed number of toxicity <= 13 and the isotonically transformed number of efficacy >= 13.

During the trial, the toxicity and efficacy of each dose arm will be monitored independently using the stopping criteria outlined in Table 1. If the isotonically transformed toxicity and efficacy acrosss topping boundaries, enrollment in that particular dose arm will be suspended.

Table 1. Stopping boundaries for toxicity and efficacy.

CSV Excel PDF Print Search:

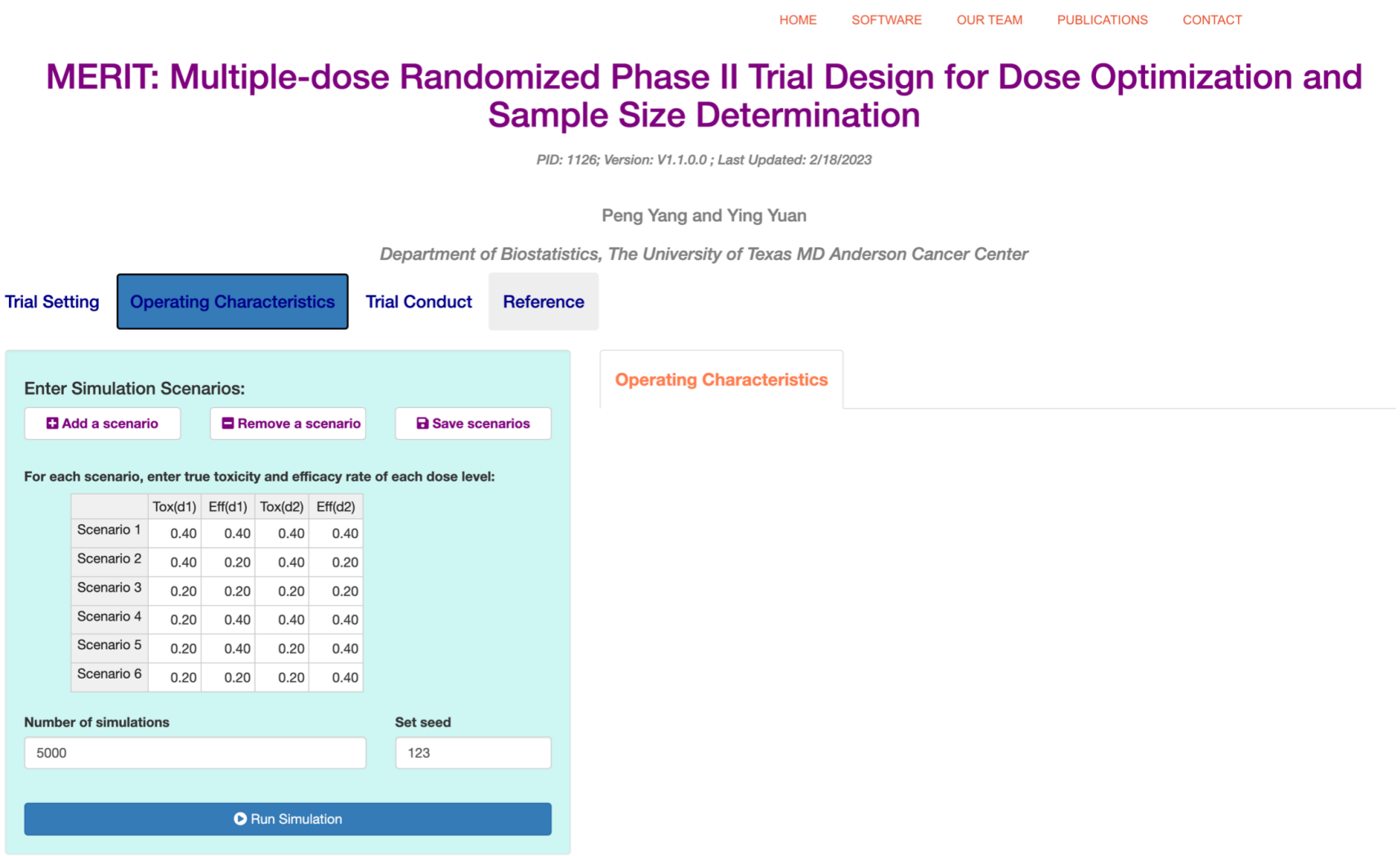
# of patients treated	Stop if # toxicity >=	Stop if # efficacy <=
15	6	NA
23	NA	5
30	10	NA

Showing 1 to 3 of 3 entries Previous 1 Next

Note: 'NA' means that this endpoint will not be used to make go/no-go decision at the interim

Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

↳ Software: shiny app on www.trialdesign.org



Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Software: shiny app on www.trialdesign.org

Operating Characteristics

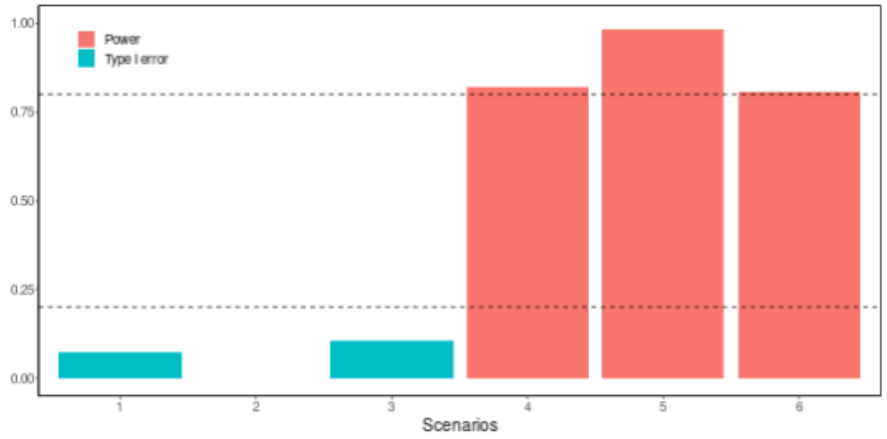
Copy CSV Excel Print Search:

Scenarios	Metrics	Values	Average sample size
1	Type I error	0.075	45
2	Type I error	0.000	45
3	Type I error	0.104	45
4	Power	0.821	45
5	Power	0.982	45
6	Power	0.805	45

Showing 1 to 6 of 6 entries Previous 1 Next

Download Figure 1

Figure 1. Type I error and power of MERIT design when unacceptable and acceptable toxicity rates are 0.4 and 0.2, and unacceptable and acceptable efficacy rates are 0.2 and 0.4. The horizontal dashed lines represent the nominal values of type I error (0.2) and power I (0.8).



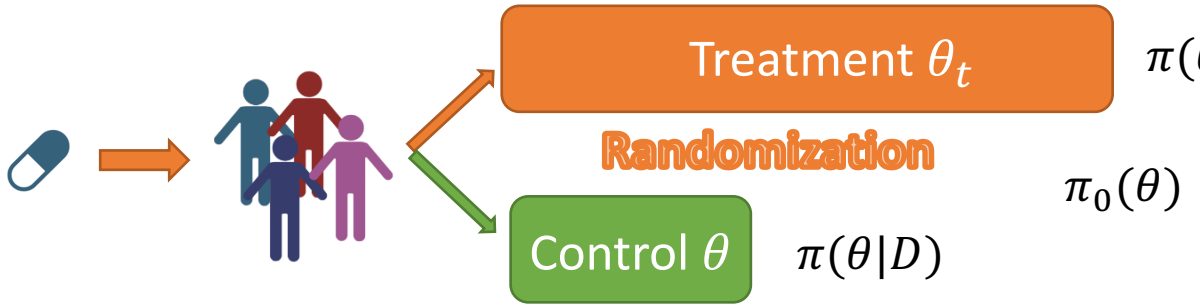
Part 1: Design and Sample Size Determination for Multi-Dose Randomized Trial

Conclusion

- In this project, we proposed MERIT design for a multiple dose randomized clinical trial by considering both toxicity and efficacy data.
- MERIT design provides a rigorous statistical framework for sample size determination and optimal dose selection.
  - This design extends beyond the traditional hypothesis testing framework, introducing structure null and alternative hypothesis to account for the ordered nature of doses across arms.
  - The sample size is determined by rigorously defining a generalized type I error and power, showing a sample size 20 to 40 per dosage arm often offers reasonable type I error and power.
- This wok has been published in *Statistics in Medicine.*
- MERIT design is available on www.trialdesign.org as a shiny app.

Part 2: Self-adapting Mixture Prior to Dynamically Borrow Information from Historical Data

Background and literature Survey



Bayes' Theorem: $\pi(\theta|D) = \dfrac{\pi(D|\theta)\pi_0(\theta)}{\pi(D)}$

$$\pi(\theta|D, D_h) = \frac{\pi(D|\theta)\pi(\theta|D_h)}{\pi(D, D_h)}$$

Hypothesis Testing:

H_0: $\theta_t \leq \theta$ Treatment is not superior than control;

H_1: $\theta_t > \theta$ Treatment is superior than control.

$$\Pr(\theta_t > \theta \,|D, D_t) > C$$

where *C* is probability cutoff to maintain desirable type I error rate.

Decision Making:

$$\Pr(\theta_t > \theta \,|D, D_t, D_h) > C$$

Background and literature Survey

- Borrowing information from historical or real-world data has great potential to improve the efficiency and feasibility of clinical trials.
- In the literature $\pi(\theta|D_h) \propto L(\theta|D_h)^{\alpha}\, \pi_0(\theta)$
  - *Ibrahim et al. (2000)* proposed power priors, which use a power parameter to acknowledge the possibility of prior-data conflict and discount the historical data for information borrowing;
  - *Hobbs et al. (2011)* proposed commensurate priors that control information borrowing based on the commensurability between historical data and current data; $\pi(\theta|D_h, \tau) \propto L(\theta|D_h)\pi(\theta|\theta_h, \tau)\pi_0(\theta)$
  - *Schmidli et al. (2014*) proposed robust meta-analytic predictive (MAP) prior, which mixes a MAP prior with a vague prior. $\pi(\theta|D_h) = w\, L(\theta|D_h)\pi_0(\theta) + (1-w)\pi_0(\theta)$
- It is important to acknowledge the prior data conflicts, and the neglect of this may cause a loss of power and inflate the type I error.

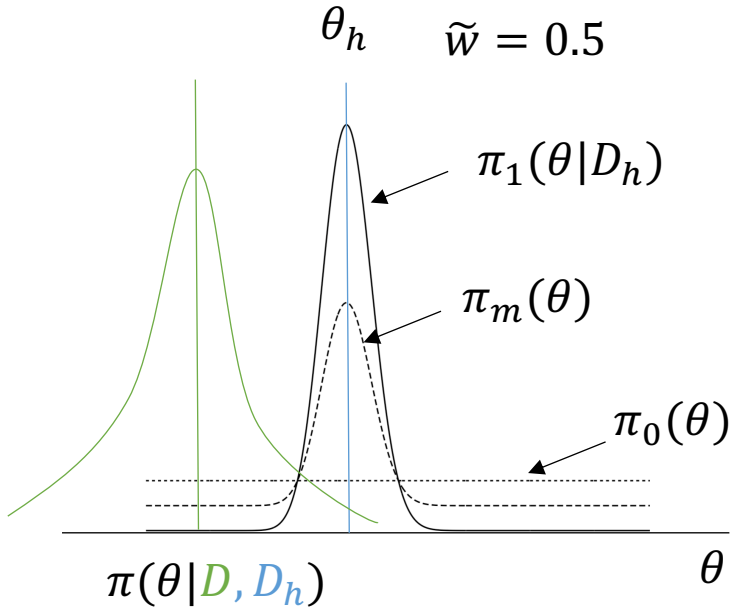
Methodology: Mixture prior

- To acknowledge the possibility of prior-data conflict and improve the robustness of the inference, Schmidli et al. (2014) proposed mixture priors:

$$\pi_m(\theta) = \tilde{w}\pi_1(\theta) + (1-\tilde{w})\pi_0(\theta),$$

where $\tilde{w}$ is a pre-specified fixed mixing weight that controls the degree of information borrowing from D_h.

- $\boldsymbol{\pi_1(\theta)}$ **is an informative prior** $\boldsymbol{\pi_1(\theta|D_h)}$ that has been constructed based on historical D_h using a certain methodology.
- $\boldsymbol{\pi_0(\theta)}$ **denotes a non-informative** or vague prior.
- Ideally, the mixing weight $\tilde{w}$ should reflect the degree of relevance of the historical data to the new trial
- Unfortunately, such information is rarely known as *a priori*.

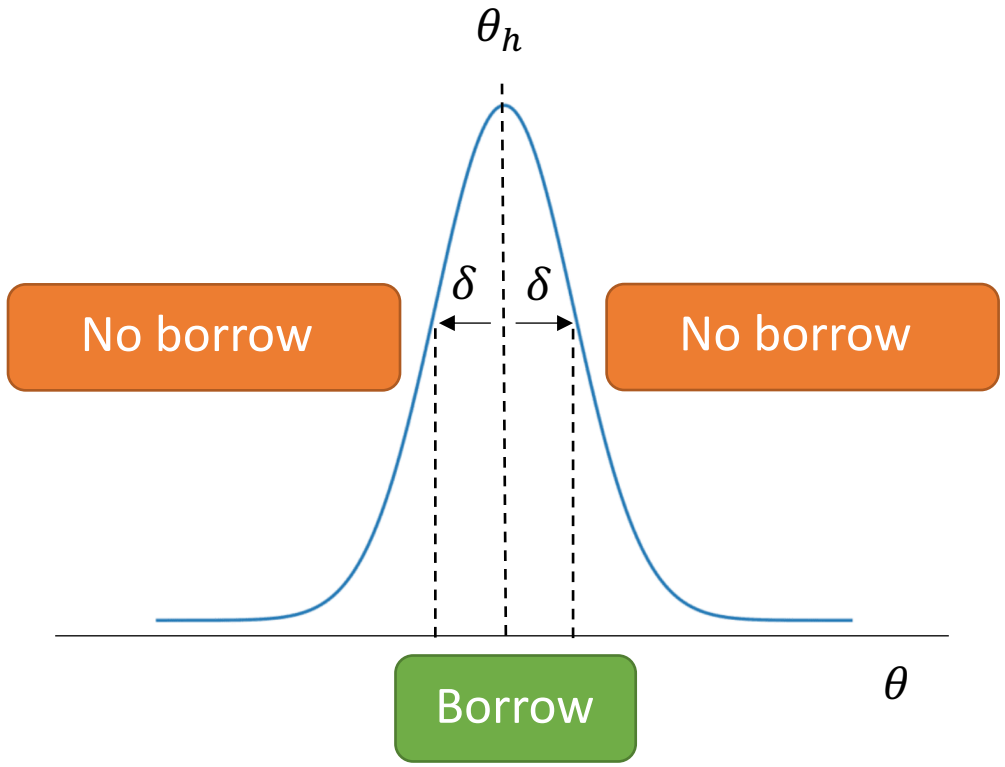


Methodology: Self-adapting mixture prior

- To account for potential prior-data conflict, we propose an empirical way of pre-determine w that takes both historical data and current trial data into consideration:

- Let θ_h denote the treatment effect associated with D_h, which could be the same as or significantly different from θ;

- Let δ denote a clinically significant difference (CSD) in the treatment effect such that if $|\theta_h - \theta| \geq \delta$, θ_h and thus it is not clinically sound to borrow any information from D_h.



Methodology: Self-adapting mixture prior

- To proceed, we define two models (or hypotheses), denoted by H_0 and H_1,

$$H_0 : \theta = \theta_h, \qquad H_1 : \theta = \theta_h + \delta \text{ or } \theta = \theta_h - \delta.$$

  - Under H_0, $\pi_1(\theta)$ and D are consistent, thus it is appropriate to use $\pi_1(\theta)$ to borrow information from D_h;
  - Under H_1, the treatment effect of D and D_h are different to the degree that no information should be borrowed, thus, $\pi_0(\theta)$ should be used for the posterior inference of θ.

- We propose to use the likelihood ratio as the evidence of favoring H_1 versus H_0 in a data-driven way,

$$R = \frac{p(D|H_0,\theta)}{p(D|H_1,\theta)} = \frac{p(D|\theta=\theta_h)}{\max\{p(D|\theta=\theta_h+\delta), p(D|\theta=\theta_h-\delta)\}},$$

where R is the likelihood ratio statistics.

Methodology: Self-adapting mixture prior

- The self-adjusting mixture prior is formed as

$$\pi_{sam}(\theta) = w\pi_1(\theta) + (1-w)\pi_0(\theta),$$

where $w = \frac{R}{1+R}$ is the *SAM weight*.

- It is important to note that, unlike the fixed-weight mixture prior, where its mixing weight $\tilde{w}$ is a constant, w is a function of D and D_h (i.e., data-dependent).

Theorem 2 *The SAM prior converges to $\pi_1(\theta)$ if D_h and D_c are congruent (i.e., $\theta_h = \theta$), and converges to $\pi_0(\theta)$ if D_h and D_c are incongruent (i.e., $|\theta - \theta_h| = \delta$)*

Methodology: Example with binary endpoints

- Consider a binary endpoint $y_1, y_2, \ldots, y_n \sim Bernoulli(\theta)$.
  - Let $x = \sum_{i=1}^{n} y_i$ denote the number of responses among n subjects treated in the control arm
  - let x_h and n_h denote the corresponding number of responses and subject in the historical data

$$\pi_1(\theta) = Beta(a + x_h, b + n_h - x_h),$$

where the informative prior $\pi_1(\theta)$ is constructed based on a vague prior $\pi_0(\theta) = Beta(a, b)$.

- Let $\hat{\theta}_h = (a + x_h)/(a + b + n_h)$, the SAM prior is given by

$$\pi_{sam}(\theta) = wBeta(a + x_h, b + n_h - x_h) + (1 - w)Beta(a, b),$$

where $w = R/(1 + R)$ with

$$R = \frac{\hat{\theta}_h^x (1-\hat{\theta}_h)^{n-x}}{\max\left\{(\hat{\theta}_h+\delta)^x (1-\hat{\theta}_h-\delta)^{n-x}, (\hat{\theta}_h-\delta)^x (1-\hat{\theta}_h+\delta)^{n-x}\right\}}$$

Methodology: Example with binary endpoints

- Owing to its conjugacy, given $\pi_{sam}(\theta)$ and trial data D, the posterior of θ is given by

$$p(\theta|D, D_h) = w^* Beta(a + x_h + x, b + n_h + n - x_h - x) + (1 - w^*)Beta(a + x, b + n - x),$$

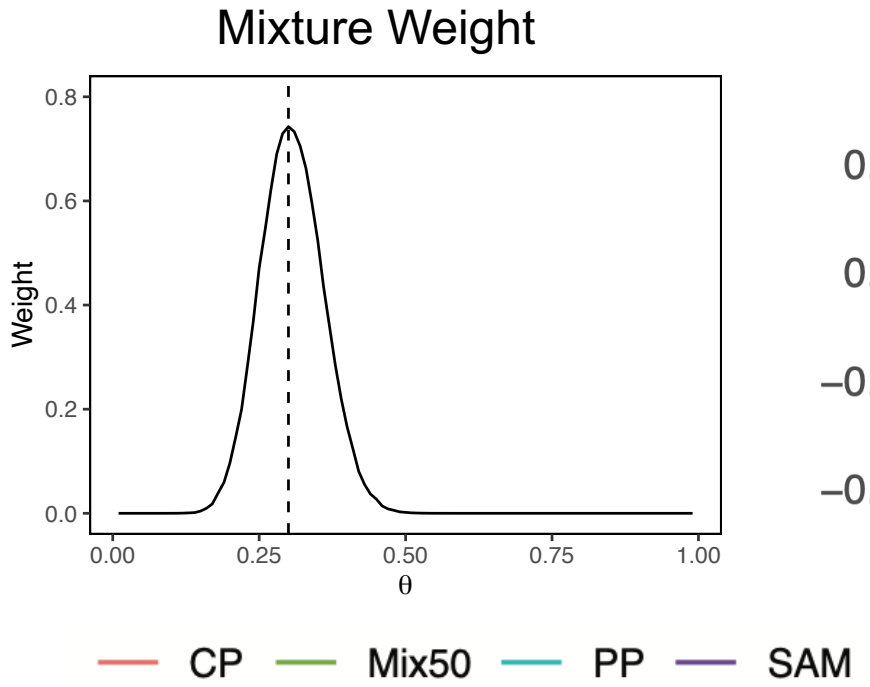
where w^* is the re-weighted w by the posterior normalizing constant associated with each mixture component. Specifically,

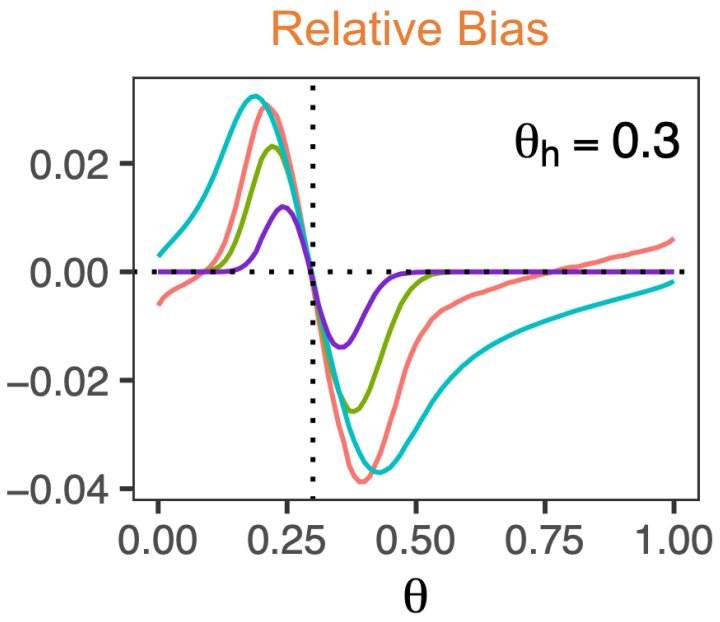
$$w^* = \frac{wz_1}{wz_1 + (1-w)z_0},$$
$$z_0 = \frac{B(a + x, n - x + b)}{B(a, b)},$$
$$z_1 = \frac{B(a + x_h + x, b + n_h + n - x_h - x)}{B(a + x_h, b + n_h - x_h)},$$

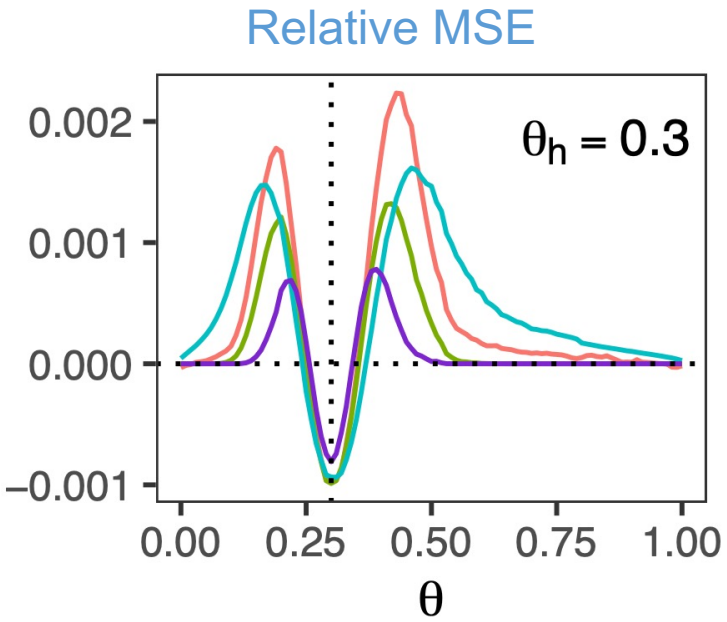
where $B(\cdot,\cdot)$ stands for beta function.

Methodology Application: Example with binary endpoints

- For binary case, we considered $\theta_h = 0.3, \delta = 0.1, n_h = n_t = 300; n = 150.$







Relative Bias: Bias of a method – Bias of NP

Relative MSE: MSE of a method – MSE of NP

Part 2: Self-adapting Mixture Prior to Dynamically Borrow Information from Historical Data

Methodology Application: Example with binary endpoints

- For binary case, we considered $\theta_h = 0.3, \delta = 0.1, n_h = n_t = 300; n = 150.$

Scenario	θ	θ_t	NP	SAM	Mix50	PP	CP	
Congruent								
2.1[a]	0.3	0.3	0.050	0.051	0.050	0.051	0.050	Type I error
2.2	0.3	0.4	0.657	0.888	0.894	0.890	0.902	Power
2.3	0.31	0.41	0.649	0.882	0.908	0.912	0.912	
2.4	0.28	0.38	0.667	0.852	0.854	0.839	0.840	
Incongruent								
2.5[a]	0.4	0.4	0.048	0.140	0.208	0.260	0.310	Type I error
2.6[a]	0.45	0.45	0.049	0.079	0.122	0.253	0.186	
2.7	0.2	0.3	0.720	0.711	0.544	0.554	0.453	Power
2.8	0.17	0.27	0.773	0.804	0.646	0.544	0.518	

Decision Making:

$$\Pr(\theta_t > \theta \mid D, D_t, D_h) > C$$

Message:
SAM prior preserves good power while maintaining better type I error control

Part 2: Self-adapting Mixture Prior to Dynamically Borrow Information from Historical Data

Software: CRAN

SAMprior: Self-Adapting Mixture (SAM) Priors

Implementation of the SAM prior and generation of its operating characteristics for dynamically borrowing information from historical data. For details, please refer to Yang et al. (2023) <doi:10.1111/biom.13927>.

Version:	1.1.1
Depends:	RBesT, assertthat, checkmate, Metrics, ggplot2
Suggests:	rmarkdown, knitr, testthat (≥ 2.0.0), foreach, purrr, rstanarm (≥ 2.17.2), scales, tools, broom, tidyr, parallel
Published:	2023-09-27
DOI:	10.32614/CRAN.package.SAMprior
Author:	Peng Yang [aut, cre], Ying Yuan [aut]
Maintainer:	Peng Yang <py11 at rice.edu>
License:	GPL (≥ 3)
NeedsCompilation:	no
Materials:	NEWS
CRAN checks:	SAMprior results

Documentation:

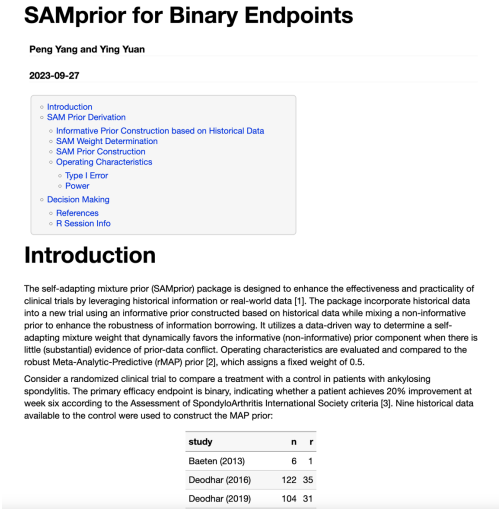
Reference manual:	SAMprior.pdf
Vignettes:	Getting started with SAMprior (binary) Getting started with SAMprior (continuous)

Downloads:

Package source:	SAMprior_1.1.1.tar.gz
Windows binaries:	r-devel: SAMprior_1.1.1.zip, r-release: SAMprior_1.1.1.zip, r-oldrel: SAMprior_1.1.1.zip
macOS binaries:	r-release (arm64): SAMprior_1.1.1.tgz, r-oldrel (arm64): SAMprior_1.1.1.tgz, r-release (x86_64): SAMprior_1.1.1.tgz, r-oldrel (x86_64): SAMprior_1.1.1.tgz
Old sources:	SAMprior archive

Linking:

Please use the canonical form https://CRAN.R-project.org/package=SAMprior to link to this page.


SAMprior for Binary Endpoints

Peng Yang and Ying Yuan

2023-09-27

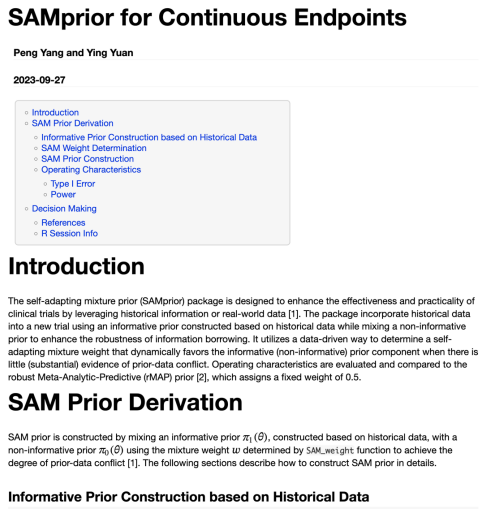
- Introduction
- SAM Prior Derivation
  - Informative Prior Construction based on Historical Data
  - SAM Weight Determination
  - SAM Prior Construction
  - Operating Characteristics
    - Type I Error
    - Power
- Decision Making
  - References
  - R Session Info

Introduction

The self-adapting mixture prior (SAMprior) package is designed to enhance the effectiveness and practicality of clinical trials by leveraging historical information or real-world data [1]. The package incorporate historical data into a new trial using an informative prior constructed based on historical data while mixing a non-informative prior to enhance the robustness of information borrowing. It utilizes a data-driven way to determine a self-adapting mixture weight that dynamically favors the informative (non-informative) prior component when there is little (substantial) evidence of prior-data conflict. Operating characteristics are evaluated and compared to the robust Meta-Analytic-Predictive (rMAP) prior [2], which assigns a fixed weight of 0.5.

Consider a randomized clinical trial to compare a treatment with a control in patients with ankylosing spondylitis. The primary efficacy endpoint is binary, indicating whether a patient achieves 20% improvement at week six according to the Assessment of SpondyloArthritis International Society criteria [3]. Nine historical data available to the control were used to construct the MAP prior:

study	n	r
Baeten (2013)	6	1
Deodhar (2016)	122	35
Deodhar (2019)	104	31


SAMprior for Continuous Endpoints

Peng Yang and Ying Yuan

2023-09-27

- Introduction
- SAM Prior Derivation
  - Informative Prior Construction based on Historical Data
  - SAM Weight Determination
  - SAM Prior Construction
  - Operating Characteristics
    - Type I Error
    - Power
- Decision Making
  - References
  - R Session Info

Introduction

The self-adapting mixture prior (SAMprior) package is designed to enhance the effectiveness and practicality of clinical trials by leveraging historical information or real-world data [1]. The package incorporate historical data into a new trial using an informative prior constructed based on historical data while mixing a non-informative prior to enhance the robustness of information borrowing. It utilizes a data-driven way to determine a self-adapting mixture weight that dynamically favors the informative (non-informative) prior component when there is little (substantial) evidence of prior-data conflict. Operating characteristics are evaluated and compared to the robust Meta-Analytic-Predictive (rMAP) prior [2], which assigns a fixed weight of 0.5.

SAM Prior Derivation

SAM prior is constructed by mixing an informative prior $\pi_1(\theta)$, constructed based on historical data, with a non-informative prior $\pi_0(\theta)$ using the mixture weight w determined by `SAM_weight` function to achieve the degree of prior-data conflict [1]. The following sections describe how to construct SAM prior in details.

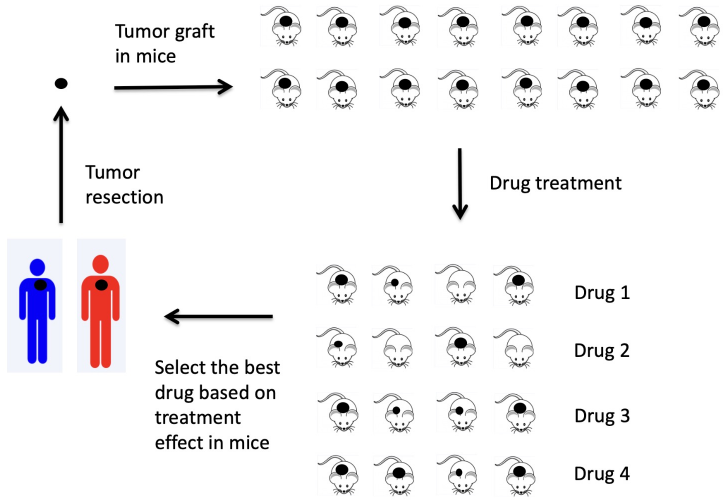
Informative Prior Construction based on Historical Data

Conclusion

- In this project, we proposed SAM prior to dynamically borrow information from historical data to current randomized clinical trials.
- SAM prior is an empirical Bayesian approach that determines the mixing weight using likelihood ratio test statistics or Bayes Factor based on outcome data.
  - SAM priors are data-driven and self-adapting, favoring the informative (noninformative) prior component when there is little (substantial) evidence of prior-data conflicts.
- The paper is published on *Biometrics.*
- SAM package is available on CRAN https://cran.r-project.org/web/packages/SAMprior/index.html;
- Propensity score-integrated (PS-SAM) will be released on CRAN soon.

Additional Thesis Research

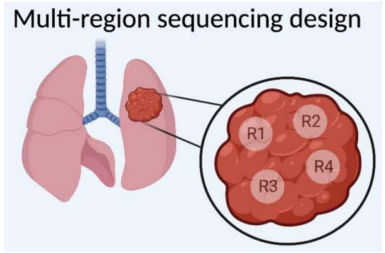
Precision Medicine and predictive biomarker identification

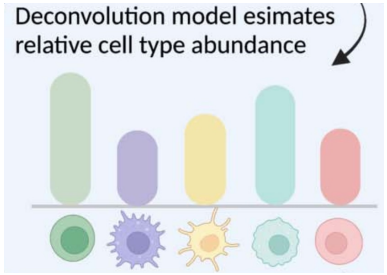


Avatar-driven clinical trials for precision medicine

- ❑ We proposed a Bayesian adaptive design for avatar-driven cancer clinical trials
  - ❑ Stage I: equal randomization (run-in phase)
  - ❑ Stage II: adaptive randomization (adaptive select optimal treatment for each patient)
- ❑ Predict the population that is sensitive to avatar responses

Quantifying intratumor heterogeneity from multi-region transcriptomic data





- ❑ We proposed a hierarchical Bayesian model for multi-region RNA data
  - ❑ Estimate the immune cell proportion while accounting for within subject correlation
  - ❑ Quantify the intratumor heterogeneity by the variability of immune cell proportions
- ❑ Utilize variance inference for optimization to enhance the scalability

Collaborative Research

Clinical Research and Genomic Study

- I served as Graduate Research Assistant at The Coordination and Data Management Center (CDMC) in MD Anderson Cancer Center
  - Maintenance and development of existing R programs to monitor enrollment, check data quality
  - Engaged in weekly meetings with investigators and physicians to address and resolve statistical inquiries
  - Conducting statistical analysis for collaborative publication in medical journals
- I collaborated with several projects on cancer genomic study
  - Performed deconvolution on matched whole genome sequencing and RNA sequencing data from TCGA and TRACERx study
  - Conducted sample integration, cell type annotation, differential gene expression analysis for single cell RNAseq data

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

Publication and Awards

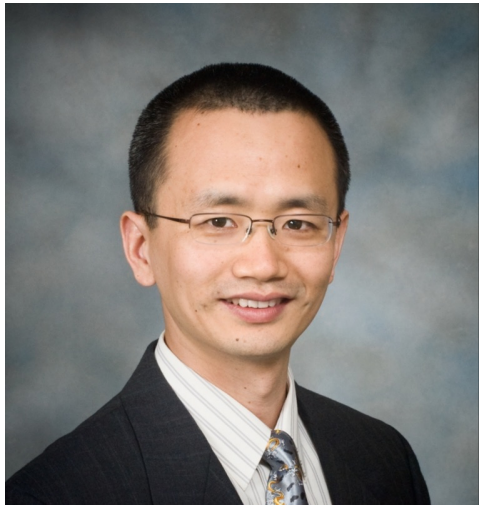
Thesis Research

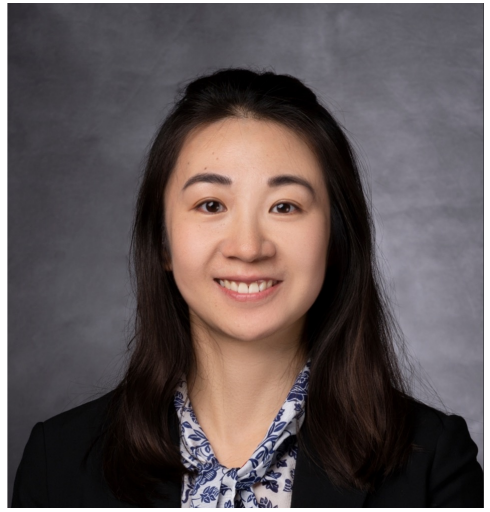
- Design and sample size determination for multiple-dose randomized phase II trials for dose optimization. *Statistics in Medicine*. 43, 2972-2986.
  - Selected for the 2023 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop Student Poster Award!
- SAM: Self-adapting mixture prior to dynamically borrow information from historical data in clinical trials. *Biometrics*. 79, 2857-2868.
- A novel Bayesian model for assessing intratumor heterogeneity of tumor infiltrating leukocytes with multi-region gene expression sequencing. *The Annals of Applied Statistics*. 18, 1879-1898.
  - Selected for the 2023 ASA Section on Statistics in Genomics and Genetics (SGG) Student Paper Award!
- A Bayesian Adaptive Design for Avatar-Driven Cancer Clinical Trials. Submitted to *The Journal of the American Statistical Association*.

Collaborative Research

- Estimation of tumor cell total mRNA expression in 15 cancer types predicts disease progression. *Nature Biotechnology*.
- Transcriptomic Profiling of Plasma Extracellular Vesicles Enables Reliable Annotation of the Cancer- specific Transcriptome and Molecular Subtype. *Cancer Research*.
- Transcriptome data analysis: Methods in Molecule Biology. *Humana Press.*
- Exocrine Pancreatic Dysfunction in Chronic Pancreatitis: Analysis of the PROCEED Study. *In Submission.*
- Single-Cell RNA Sequencing Identifies Molecular Biomarkers Predicting Late Progression to CDK4/6 Inhibition in Metastatic HR+/HER2- Breast Cancer. *In Submission.*

Thank you





THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

Reference

- Hobbs, B. P., Carlin, B. P., Mandrekar, S. J., and Sargent, D. J. (2011). Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. *Biometrics*, 67(3), 1047-1056.
- Ibrahim, J. G., and Chen, M. H. (2000). Power prior distributions for regression models. *Statistical Science*, 15(1), 46-60.
- Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D., and Neuenschwander, B. (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics*, 70(4), 1023-1032.
- Thall, P., Simon, R. & Ellenberg, S. (1988) Two-stage selection and testing designs for comparative clinical trials. *Biometrika*. 75, 303-310.
- Thall, P., Simon, R. & Ellenberg, S. (1989) A two-stage design for choosing among several experimental treatments and a control in clinical trials. *Biometrics*. pp. 537-547.
- U.S. Food And Drug Administration. (2022) Project Optimus: Reforming the dose optimization and dose selection paradigm in oncology. Silver Spring, MD: Food and Drug Administration.
- Whitehead, J. (1986) Sample sizes for phase II and phase III clinical trials: an integrated approach. *Statistics In Medicine*. 5, 459-464.