DOME: A Bayesian Optimal Design for Randomized Dose Expansion Cohorts in Oncology Trials

Jinjie Chen, PhD[1], Rong Liu, PhD[1]

[1]Bristol Myers Squibb

29TH ANNUAL BASS

Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Introduction

- Relapsed and refractory multiple myeloma (RRMM) caused considerable morbidity and mortality, a clear unmet medical need remains for novel therapies.
- In addition to mono-therapy, many RRMM patients receive drug combination of different classes.
- Identify promising combo-therapy is challenging, particularly for a novel-novel combination, where both agents haven't been approved by regulatory authority.
- When multiple combinations available, multi-arm group sequential designs are efficient to compare multiple treatments to a control, considerably lower average sample size and shorten trial period.

Bristol Myers Squibb™

Introduction cont.

- In early development (phase 1b/2a) of the combination therapy for RRMM, sample sizes are usually small (40-60 per arm), the target is screening out obviously futile treatment, rather than declaring efficacy.
- The small sample size means the heterogeneity between historical data and current data of the control arm is non-trivial.
- Recently FDA initiates Project Optimus to reform the dose optimization and dose selection paradigm in oncology drug development. Multiple doses optimization is required as early as possible.
- It is challenging for statistician to design a "good" trial under the constraints above.

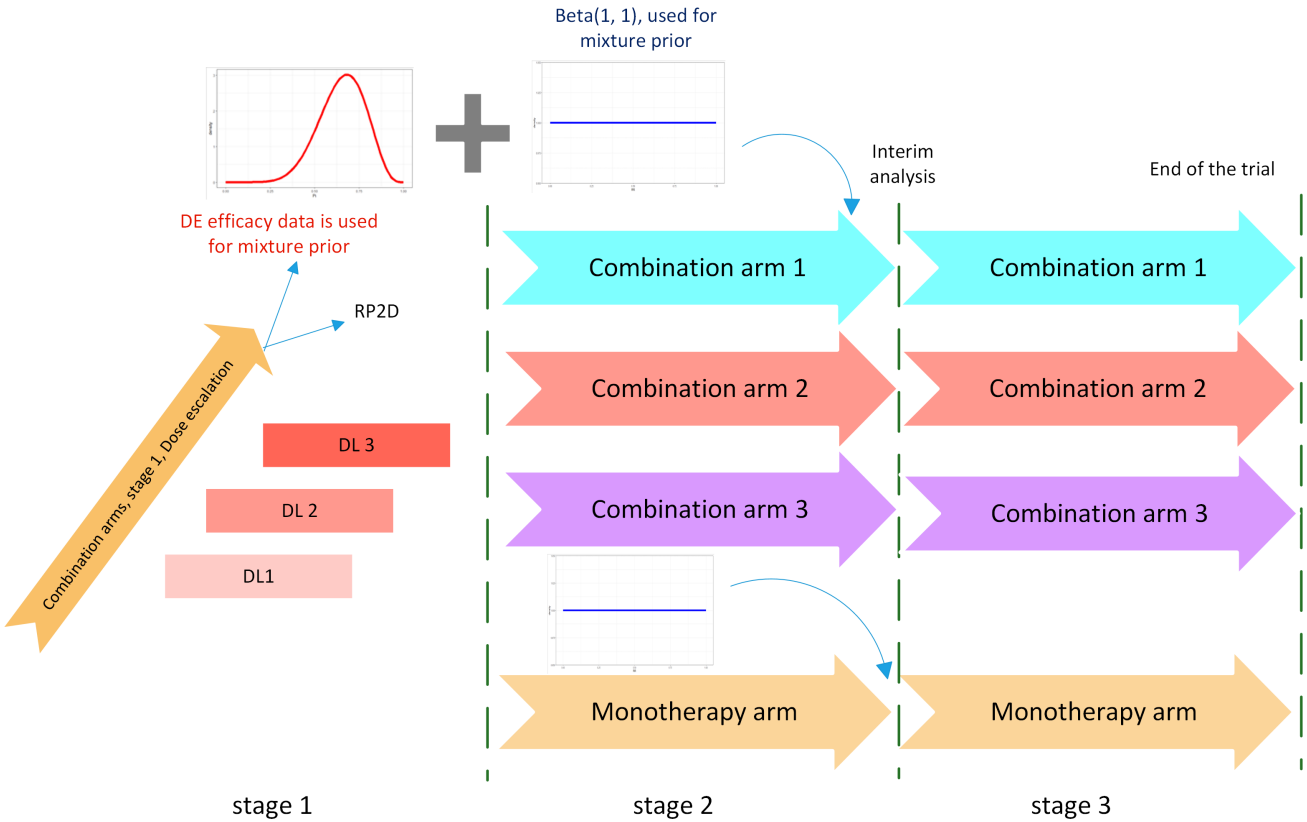
Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Schema of DOME design

We proposed a Bayesian Dual-criterion Optimal Design for Multi-arm randomized expansion cohorts (DOME design)



Bristol Myers Squibb™

Probability model for phase 1b dose escalation (DE)

Assume the primary efficacy endpoint is binary (e.g, overall response rate ORR). Let $Y_k = \sum X_k$ denotes the number of responses in treatment arm k with sample size N_k. We model Y_k with a binomial distribution

$$Y_k \sim \text{Binom}(\theta_k, N_k)$$

where θ_k represents the response rate. Assign a conjugate beta prior to θ_k, i.e., $\theta_k \sim \text{Beta}(a_k, b_k)$, the resulting posterior distribution of θ_k is another Beta distribution, of form

$$\theta_k | Y = y_k \sim \text{Beta}(a_k + y_k, b_k + N_k - y_k)$$

At phase 1b, we assume vague prior $a_k = b_k = 1$ or 0.5

Bristol Myers Squibb™

Probability model for phase 2a dose expansion

In phase 2a, patients will be randomly assigned to arm k, where $k = 0$ represents the control arm, and $k > 1$ correspond to treatment arms. The accrual sample size at each interim looks are denoted as n_{kl}, assuming

$$Y_{kl} \sim \text{Binom}(\theta_k, n_{kl})$$

The response rate θ_0 for the control arm is assigned a vague conjugate prior e.g, Beta(a_k, b_k), while for each treatment arm k in phase 2a, the corresponding response rate θ_k is assigned a mixture prior $\pi_{mix}(\theta_k)$ which combines the preliminary efficacy information for treatment arm k in phase 1b and a vague beta prior Beta(a_k, b_k).

Bristol Myers Squibb™

Mixture priors

Consider j individual priors for θ_k: $\pi = (\pi_1(\theta_k), \pi_2(\theta_k), \cdots, \pi_J(\theta_k))$, and a vector of weights $\omega = (\omega_1, \omega_2, \cdots, \omega_J)$ satisfying $\sum_j \omega_j = 1$, then a mixture prior distribution for θ_k is:

$$\pi_{mix}(\theta_k|\pi, \omega) = \sum_{j=1}^{J} \omega_j \pi_j(\theta_k), \tag{1}$$

The posterior density of $(\theta_k|y, \pi, \omega)$ is another mixture, of form

$$p(\theta_k|y, \pi, \omega) = \sum_{j=1}^{J} \omega_j^* \pi_j(\theta_k|y), \tag{2}$$

where $\pi_j(\theta_k|y)$ is the posterior of θ_k under each individual component $\pi_j(\theta_K)$. For Beta-binomial model,$\pi_j(\theta_k|y)$ is also a beta distribution

Bristol Myers Squibb™

Mixture priors with two beta distribution

Specifically, for a mixture prior with two beta components:

$$\pi_{mix}(\theta_k) = \omega_1 \beta(\theta_k | a_1, b_1) + \omega_2 \beta(\theta_k | a_2, b_2),$$

The posterior of θ_k is also a mixture of two beta distributions:

$$p(\theta_k | y, \pi, \omega) = \omega_1^* \beta(\theta_k | a_1, b_1, y) + \omega_2^* \beta(\theta_k | a_2, b_2, y),$$

with the updated weights of form:

$$\omega_1^* = \left[1 + \frac{\omega_2}{\omega_1} \frac{B(a_2 + x, b_2 + n - x)}{B(a_1 + x, b_1 + n - x)} \frac{B(a_1, b_1)}{B(a_2, b_2)} \right]^{-1} \tag{3}$$

$$\omega_2^* = 1 - \omega_1^*, \tag{4}$$

where B represents Beta function.

Bristol Myers Squibb™

Mixture priors with various weights

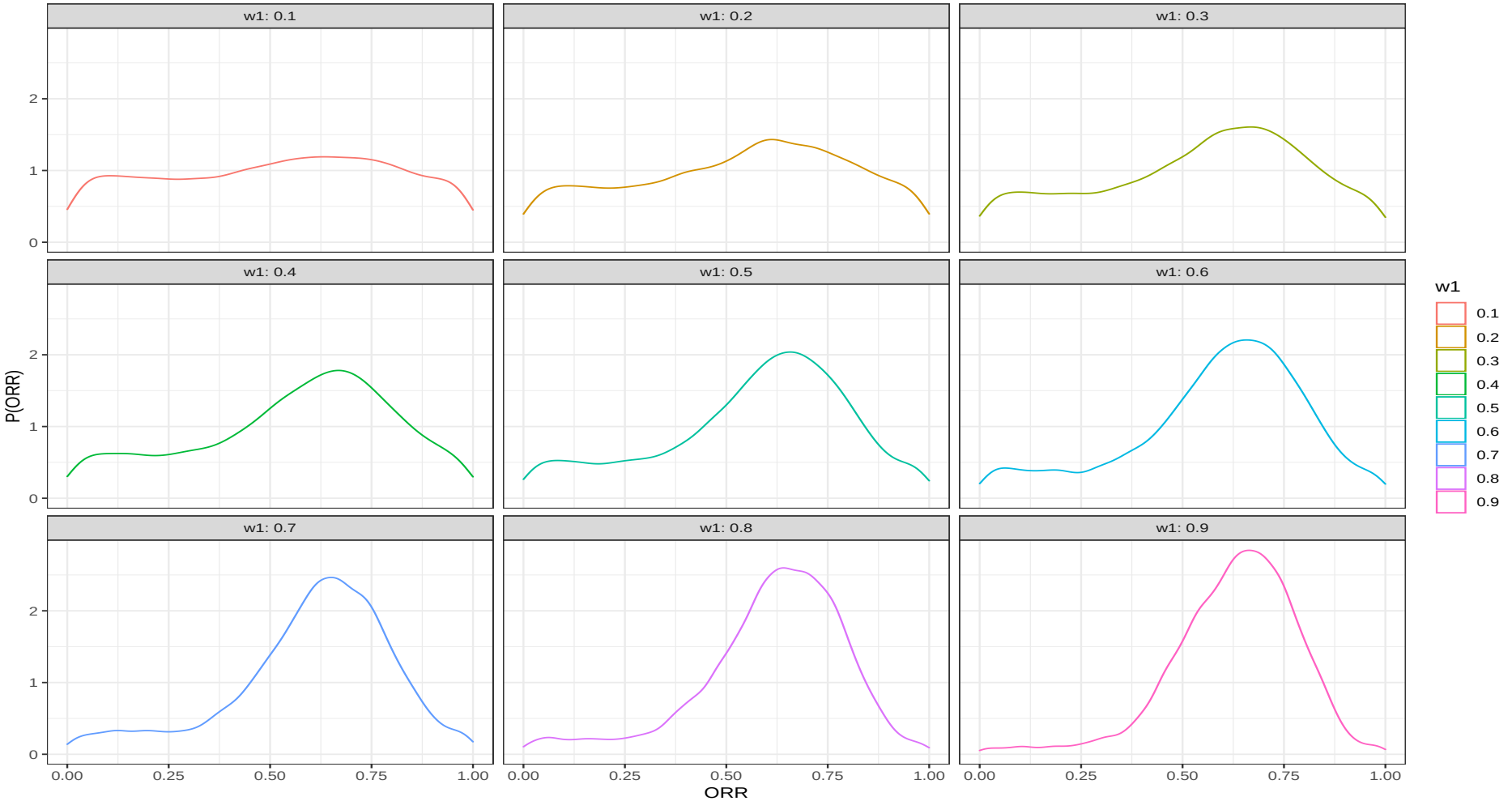


Figure 1: Density plot for the mixture of $\pi_1(\theta) = Beta(9, 5)$ and $\pi_2(\theta) = Beta(1, 1)$ with various weight ω_1, and $\omega_2 = 1 - \omega_1$

Bristol Myers Squibb™

Design based on posterior probability

- Zhou et al (2017) proposed a design that stops for futility when

$$Pr(\theta \leq p_0 | D_n) > C(n)$$

where θ is the response rate of a new treatment, p_0 is the uninteresting response rate, D_n is the data observed at an interim stage, and $C(n)$ is the cutoff value that changes with accumulated sample size n.

- The function of $C(n)$ is

$$C(n) = 1 - \lambda \left(\frac{n}{N}\right)^{\gamma}$$

where λ and γ are two tuning parameters, optimized to maximize the power.

Bristol Myers Squibb™

Dual-criterion of DOME

We propose dual-criterion for DOME design.
Criterion I corresponds to a single-arm comparison, comparing θ_k to a fixed benchmark θ_{Fixed}:

$$Pr[\theta_k \geq \theta_{Fixed} | y_{kl}, n_{kl}] > C_l, 1 < l \leq L, \tag{5}$$

where y_{kl} and n_{kl} are accumulated number of response and sample size by look l, C_l is a set of cutoff probabilities.
Criterion II concurrently compares treatment arms against the control arm, and the go-nogo decision is based upon:

$$Pr[\theta_k - \theta_0 \geq \delta | y_{kl}, n_{kl}, y_{0l}, n_{0l}, \pi_{mix}(\theta_k)] > P_l \ , 1 < l \leq L, k > 0, \tag{6}$$

where δ is a pre-specified margin. For a superiority trial, $\delta \geq 0$, and for a non-inferiority trial $\delta < 0$. Any treatment arm will be stopped at look l unless it meets the dual-criterion (5) and (6)

Bristol Myers Squibb™

Design parameters optimizing

A DOME design is optimized under null and alternative hypothesis of form:

$$H_0 : \theta_k - \theta_0 = 0$$
$$H_a : \theta_k - \theta_0 = \gamma,\ \gamma > 0,\ k \neq 0$$

The OC of DOME is determined by the cutoff probabilities C_l and P_l. We use two bi-parameter tuning functions of n_{kl} similar to Zhou, Lee, and Yuan (2017),

$$C_l(n_{kl}) = a_c(n_{kl}/N)^{b_c}, \tag{7}$$
$$P_l(n_{kl}) = a_p(n_{kl}/N)^{b_p}, \tag{8}$$

Tuning parameters $a_c, a_p \in (0, 1)$ and $b_c, b_p > 0$ are selected to guarantee that C_l and P_l are monotonically increasing with n_{kl}.

Bristol Myers Squibb™

The stopping cutoff function for C_i

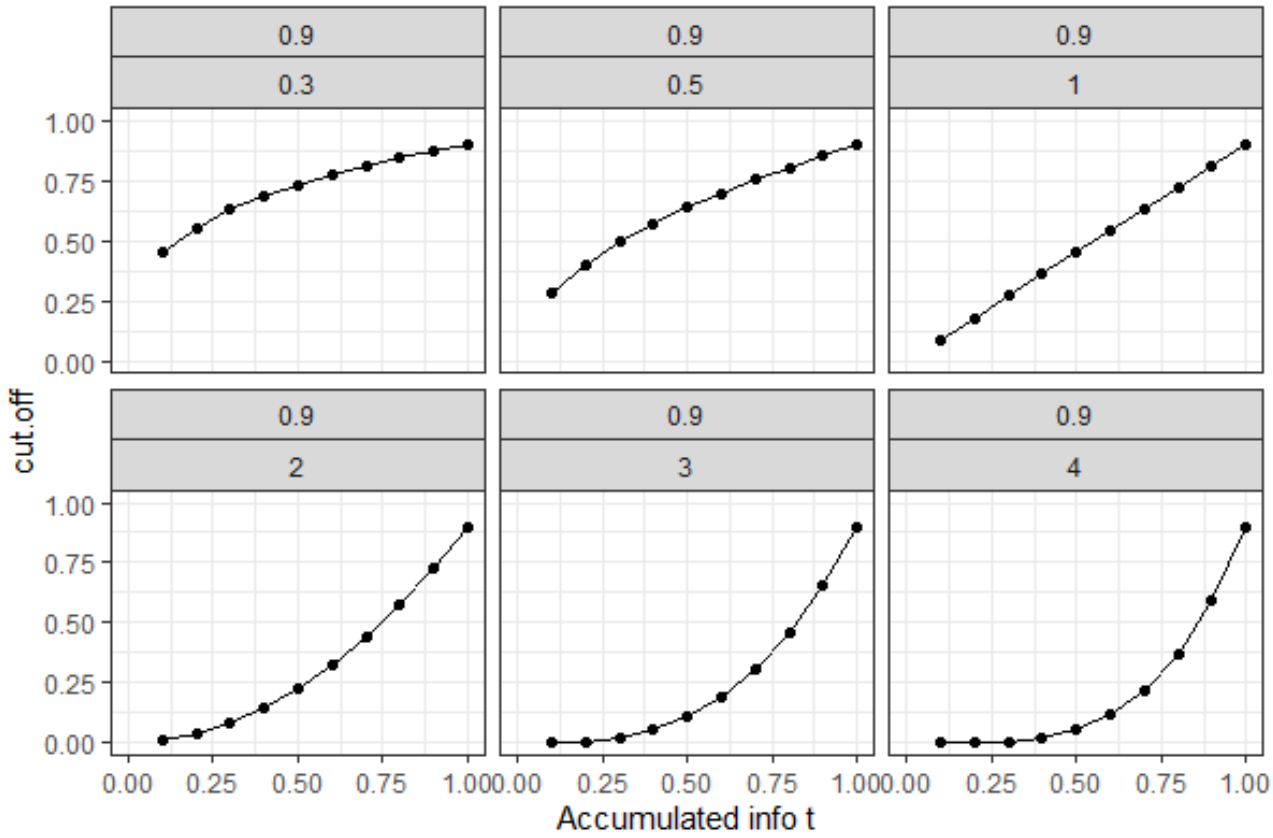


Figure 2: Different shapes of the stopping cutoff function for rule 1 when $a_c = 0.9,\ b_c = 0.3, 0.5, 1, 2, 3, 4$

Bristol Myers Squibb™

The steps and algorithm for an optimal DOME design

(1) Work with clinical team to specify k, I, n_{kI}, θ_0, γ, δ, θ_{Fixed}, and type I error α; specify the range and resolution for (a_c, b_c, a_p, b_p);

(2) Under $H_0 : \theta_k = \theta_0$, simulate a trial with parameters in step (1), compute the posterior probability on the left of inequality (5) and (6).

(3) For each grid of (a_c, b_c, a_p, b_p), derive C_I and P_I, then make the go-nogo decision of the dual-criterion

(4) Flag a trial as "final go" if it satisfies the dual-criterion at all looks, otherwise flag the trial as "final nogo".

(5) Repeat step (2-4) 10,000 times, compute the proportion of "final go" for each grid, obtain the Type I Errors.

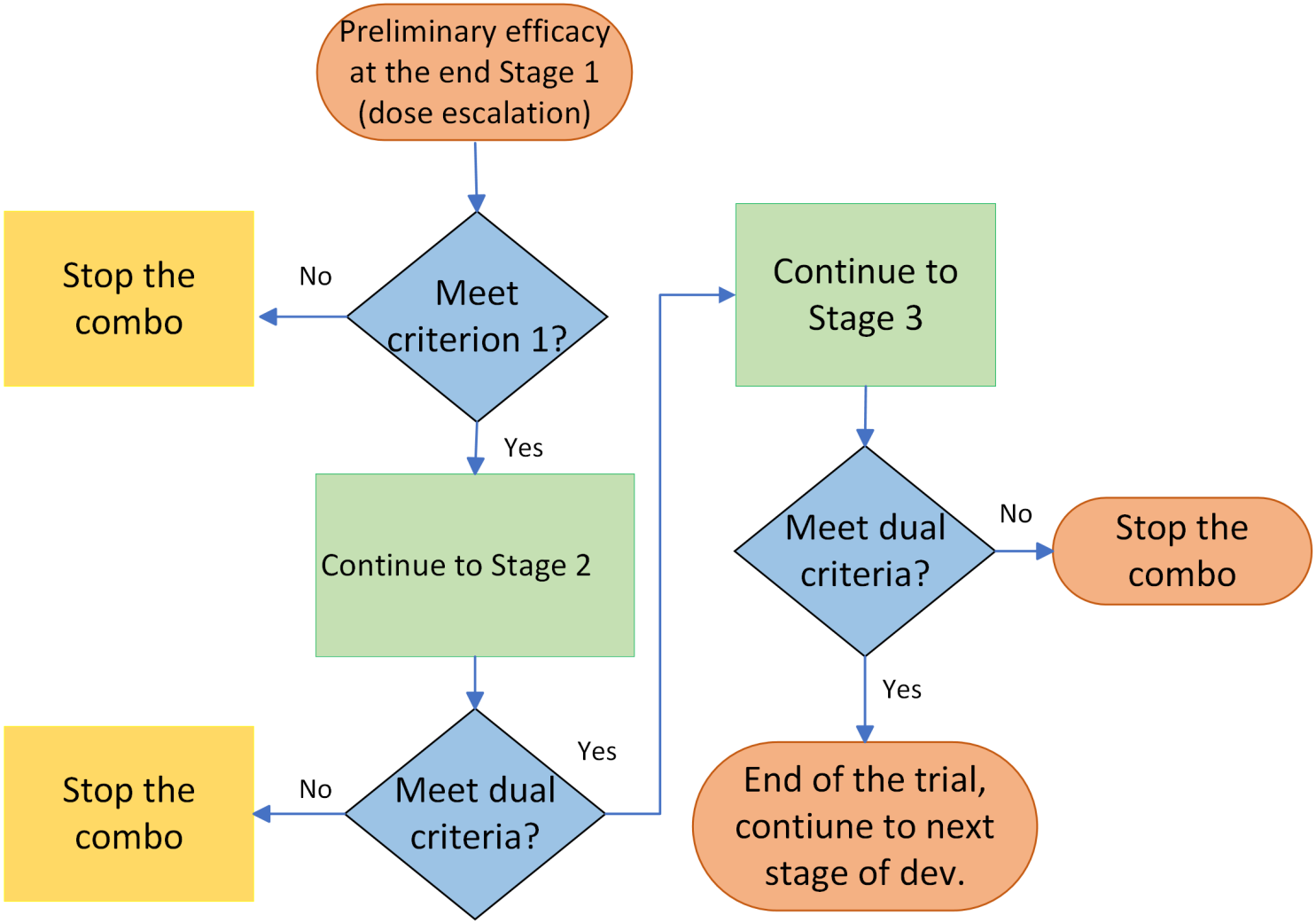
(6) Under $H_a : \theta_k = \theta_0 + \gamma$, repeat step (2-5), obtain the powers for each grid.

(7) Screen out those grids leading the type I error falling in a small range $(\alpha - \Delta, \alpha + \Delta)$, $\Delta > 0$ is a tiny margin.

(8) Among the grids in step (7), select those with maximum power and/or minimum expected sample size.

Bristol Myers Squibb™

Flowchart of decision making by DOME



Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Simulation setting

Table 1: Simulation setting of the proposed DOME design with 3 treatment arms and 3 stages: two sets of hypothesis 0.2 vs 0.4, and 0.45 vs 0.65 are considered.

Single-agent type I error	$\approx 5\%$			
Family-level type I error	$\leq 20\%$			
Treatment arms	3			
Early stop for futilty	Yes			
Early declaration for efficacy	No			
Sample size of stage 1 n_1	09	12	09	12
Sample size of stage 2 n_2	15	20	15	20
Sample size of stage 3 n_3	15	20	15	20
θ_{clt}	0.20		0.45	
θ_{trt}	0.40		0.65	

Bristol Myers Squibb™

Example of "optimal" selection: 0.20 vs 0.40, $n_1, n_2, n_3 = c(12, 20, 20)$

	C_1	C_2	C_3	P_2	P_3	α	EN	True Go
1	0.58	0.78	0.90	0.50	0.60	0.05	23.66	0.83
2	0.58	0.78	0.90	0.45	0.65	0.05	23.83	0.83
3	0.43	0.71	0.90	0.65	0.75	0.05	29.90	0.82
4	0.43	0.71	0.90	0.45	0.80	0.05	31.02	0.85
5	0.43	0.71	0.90	0.50	0.80	0.05	30.73	0.84
6	0.52	0.69	0.80	0.45	0.85	0.05	24.25	0.78
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	0.01	0.21	0.90	0.65	0.85	0.05	38.55	0.81
47	0.00	0.13	0.90	0.65	0.85	0.05	38.70	0.81
48	0.38	0.63	0.80	0.65	0.90	0.05	30.90	0.75

Table 2: Parameters selection: Row 4 has the maximum power 85%, Row 1's power is slightly smaller but Row 1's average sample size (EN) is much less than Row 4's, thus Row 1 is optimal.

Bristol Myers Squibb™

Simulation result part I

Case 1	θ_0	θ_1	θ_2	θ_3	go_1	go_2	go_3	family-go
$(n_1, n_2, n_3) = (12, 20, 20)$	0.2	0.4	0.3	0.2	85%	46%	6%	90%
	0.2	0.2	0.2	0.2	5%	5%	5%	16%
	0.2	0.4	0.4	0.4	85%	85%	85%	98%
Case 2								
$(n_1, n_2, n_3) = (9, 15, 15)$	0.2	0.4	0.3	0.2	78%	40%	6%	85%
	0.2	0.2	0.2	0.2	6%	6%	6%	16%
	0.2	0.4	0.4	0.4	78%	78%	78%	97%
Case 3								
$(n_1, n_2, n_3) = (12, 20, 20)$	0.45	0.65	0.55	0.45	81%	37%	5%	88%
	0.45	0.45	0.45	0.45	5%	5%	5%	13%
	0.45	0.65	0.65	0.65	81%	81%	81%	98%
Case 4								
$(n_1, n_2, n_3) = (9, 15, 15)$	0.45	0.65	0.55	0.45	74%	33%	5%	81%
	0.45	0.45	0.45	0.45	5%	5%	5%	15%
	0.45	0.65	0.65	0.65	74%	74%	74%	95%

Table 3: OC of DOME design under different sample sizes and underlying θ_0 and θ_k

Bristol Myers Squibb™

Simulation result part II

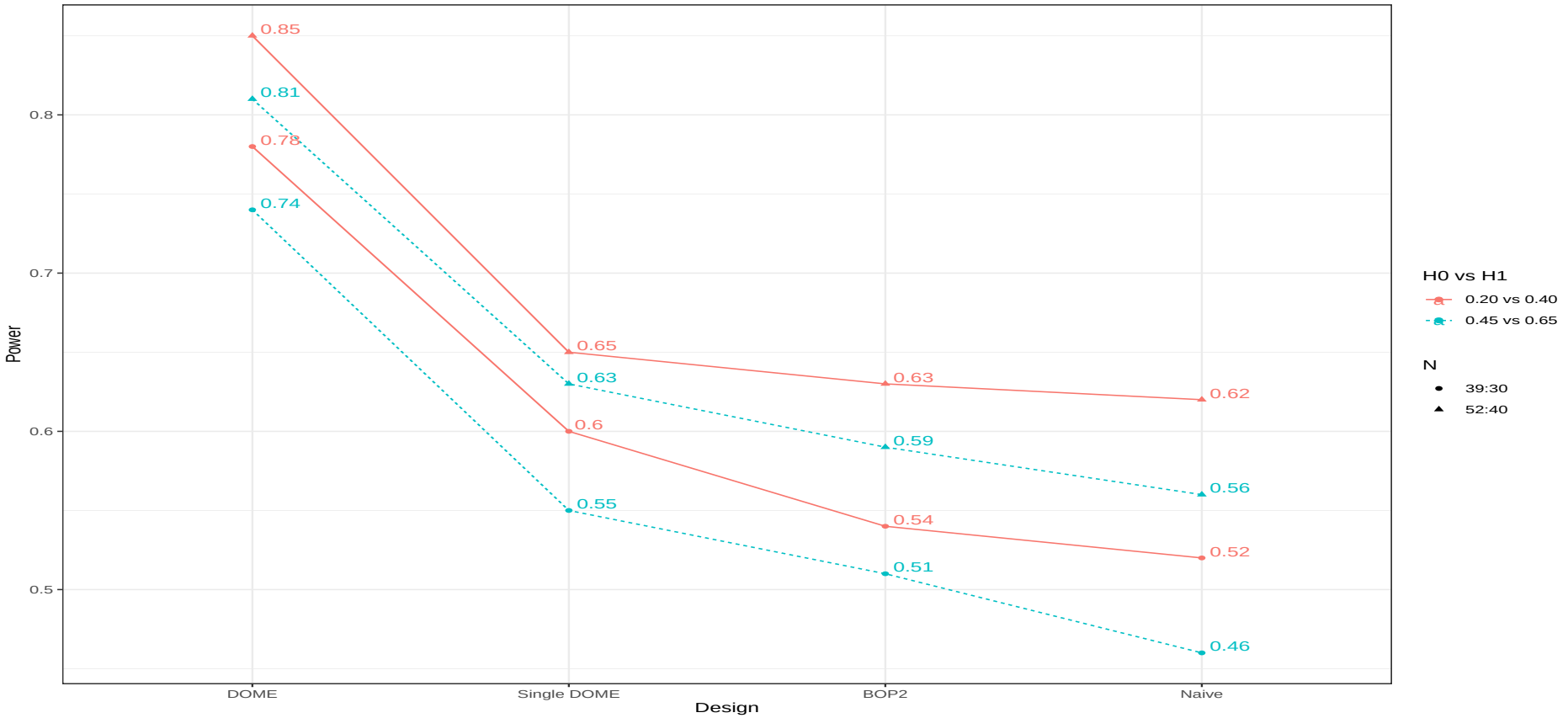


Figure 3: Model comparison grouped by $\theta_0 : \theta_k$ and sample size N. DOME is much more powerful compared with two-arm BOP2 and two-arm group-sequential design.

Bristol Myers Squibb™

Sample size saving, sensitivity analysis for mixture weight.

Table 4: Sample size savings compared with 2-look conventional two-arm design

	H0	*Ha*	Type I	Power	DOME	2-look Conventional[1]	Sample size Saving
1	0.20	0.40	0.05	0.85	92	150	58
2	0.45	0.65	0.05	0.81	92	160	68

Table 5: Sensitivity analysis for 0.45 vs 0.65 case, with $N = 12 + 20 + 20$

	Weight	Type I	Power
1	0.000	0.049	0.783
2	0.250	0.050	0.786
3	0.500	0.051	0.795
4	0.750	0.049	0.814
5	0.850	0.050	0.813
6	0.950	0.050	0.816

[1] 2-look conventional sample size is calculate by EAST software

Bristol Myers Squibb™

Table 6: Operating Characteristics of DOME two-arm vs BOP2 and group sequential designed based on gamma family.

ORR		Two-arm DOME		Two-arm BOP2[2]		Gamma family	
Con.	Trt.	Go rate	Average sample size	Go rate	Average sample size	Go rate	Average sample size
0.2	0.1	0	33.9	0.5	33.5	0.7	45.5
	0.2	9.9	48.8	90.9	49.8	9.8	62.6
	0.3	50.5	60.5	41.7	64.8	39.5	73.8
	0.4	84.5	66.3	75.2	74.2	74.4	78.4
	0.5	96.8	68.3	93.7	78.4	94.2	79.7
0.25	0.2	1.9	46.1	3.4	45.1	3.7	55.6
	0.25	10.1	53.1	9.9	52.3	10.0	62.9
	0.3	26.5	59.1	21.8	59.3	20.7	59.1
	0.4	68.9	66.1	54.6	70.6	54.7	76.3
	0.5	92.3	68.3	83.8	76.7	86.1	79.1

[2]Yujie Zhao et.al, BOP2: Bayesian Optimal Phase II Design for Randomized Clinical Trials

A case study

Example

Background

- In a FIH trial, the mono-therapy of agent A reaches the target ORR $\geq 60\%$
- Agent A will be combined with another agent B in a new phase 1b/2a study, as treatment arm, and agent A mono-therapy be the control
- Team would like to see a minimal 0.05 effect and over 70% power given the true ORR diff. is 0.2
- A dual-criterion with a lower reference value and a decision value following Roychoudhury et.al [a] would be of form,

[a] Beyond p-values: A phase II dual-criterion design with statistical significance and clinical relevance

Bristol Myers Squibb™

A case study

- $Pr(RR_1 - RR_2 \geq 0.05) \geq 80\%$
  $Pr(RR_1 - RR_2 \geq 0.12) \geq 50\%$
- With DOME design it would be of form:
- $Pr(RR_1 \geq 0.6) \geq C_l$,
  $Pr(RR_1 - RR_2 \geq 0) \geq P_l$;

ORR		Dual-criterion following Roychoudhury et.al			DOME design[3]		
Cont.	Trt.	Go rate	planned sample size	Average sample size	Go rate	planned sample size	Average sample size
	0.6	0.057	80	49.8	0.038	69	32.5
	0.65	0.126	80	54.52	0.125	69	39.6
0.6	0.7	0.242	80	59.68	0.2886	69	47.8
	0.75	0.408	80	65.32	0.5152	69	55.9
	0.8	0.589	80	70.6	0.7426	69	62.2
	0.9	0.911	80	78.12	0.9796	69	68.5

Table 7: Operating Characteristics of DOME vs Roychoudhury et.al

[3]With optimization $C_l = (0.501, 0.741, 0.9)$, $P_l = (0.30, 0.75)$

Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Summary

- DOME design incorporates partial information from DE phase by using mixture prior, saving some sample size.
- The dual-criterion in DOME represents a combination of single-arm phase II and two-arm randomized phase II. The latter remedies the heterogeneity between treatment arm and control arm; the former deals with the heterogeneity between historical and the current data, specifically useful when the sample size is small.
- Dual-criterion DOME design can be extended to triple-criterion by setting two δ_i ($\delta_1 = 0$, and $\delta_2 > \delta_1$), one for significance, one for clinical meaningful margin.
- The optimal result of DOME may not be straight-forward to understand by non-statistician, thus clearly interpreting the schema is important. One option is placing more restriction on the grids to search, though it may result in a sub-optimal solution.

Bristol Myers Squibb™

Outline

Bristol Myers Squibb™

Reference

- Liyun Jiang et al. "Comparing Bayesian early stopping boundaries for phase II clinical trials". In: *Pharmaceutical statistics* 19.6 (2020), pp. 928–939.
- Elias Laurin Meyer et al. "Decision rules for identifying combination therapies in open-entry, randomized controlled platform trials". In: *arXiv preprint arXiv:2103.09547* (2021).
- Steffen Ventz et al. "Designing clinical trials that accept new arms: an example in metastatic breast cancer". In: *Journal of Clinical Oncology* 35.27 (2017), pp. 3160–3168.

Bristol Myers Squibb™