LAIYA
Consulting Inc.

Bayesian Seamless Strategy for Early-Phase Drug Development

Yuan Ji, PhD

Outlines

- Introduction to the Bayesian Seamless Strategy
- The **B**ayesian **E**arly-Phase **S**eamless **T**ransformation (**BEST**) Platform
  - Platform overview
  - Bayesian Hierarchical Model (BHM) --- Key feature in BEST
  - Pilot simulation results for BHM
  - Bayesian multiplicity control

Bayesian adaptive design boost drug development

2006: “Improved utilization of adaptive and Bayesian methods” in clinical trials

2013: FDA will need to "turn the clinical trial paradigm on its head” to allow personalized drug therapies to get on the market faster

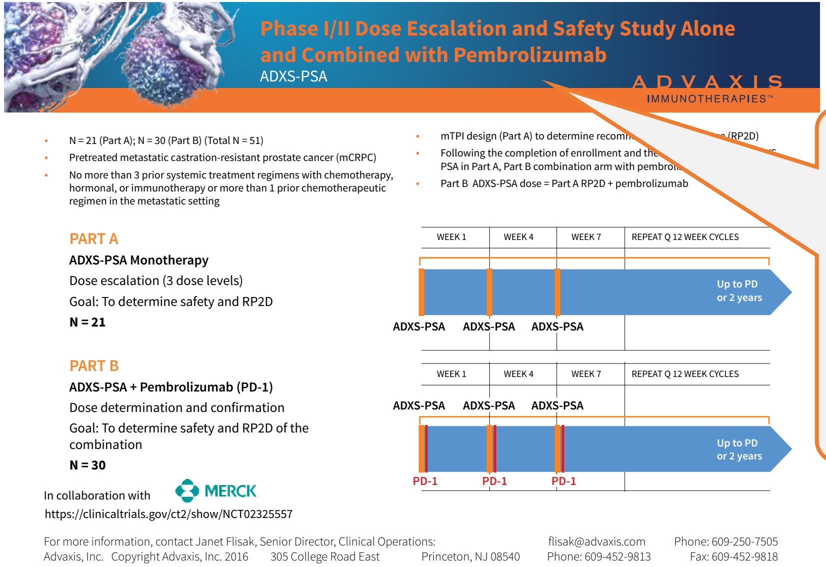
2015, Press launch of GBM AGILE: “This is the future.”

Janet Woodcock, Director CDER FDA

- The use of adaptive trial designs are well accepted in late-phase oncology drug development in US
- Early-phase drug development is even **MORE FLEXIBLE** and important: incorrect dose selection leads to disastrous late phase trials.
- Now is the time to deploy Bayesian adaptive designs in early-phase trials globally

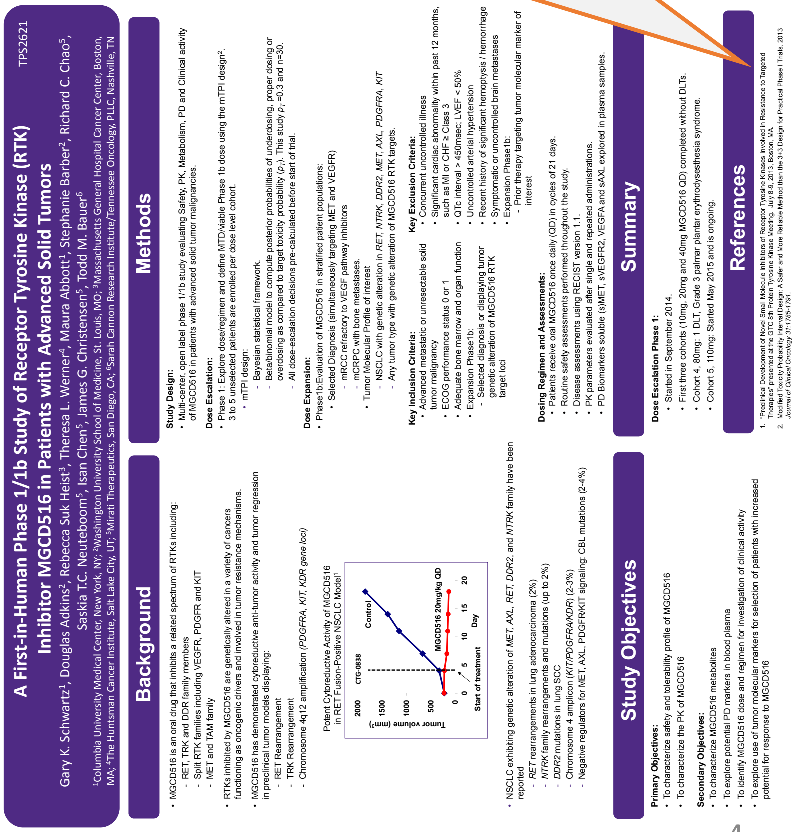
The mTPI design (Ji et al., 2010, 2013) is a success story in dose finding

EAST includes the mTPI design



The mTPI design for ADXS-PSA + PD-1

ASCO 2016 Protocol for FIH phase 1/1b study using the mTPI design



FDA's Woodcock: The Clinical Trials System is 'Broken’

Posted 20 September 2017

The clinical trials system is "broken" and there needs to be new ways to collect and utilize patient data, Janet Woodcock, director of FDA's Center for Drug Evaluation and Research, told a workshop on real world evidence (RWE) at the National Academies of Sciences, Engineering, and Medicine on Wednesday.

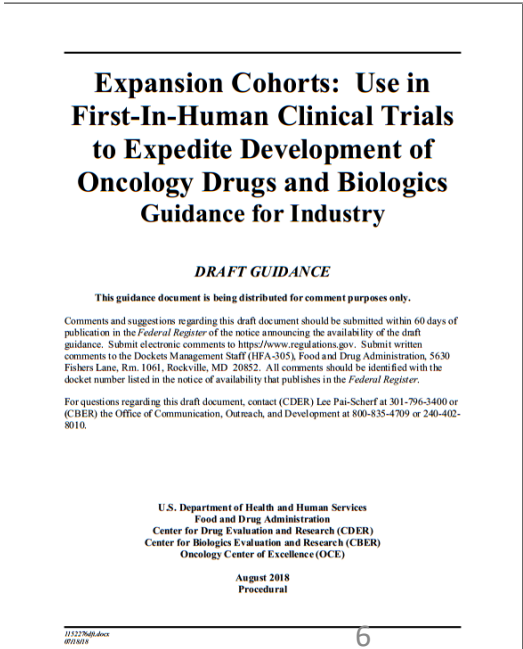
The comment came at the end of Woodcock's talk in which she also noted that **use of master protocols (ie. protocols for trials that look at multiple therapies in a single disease or a single treatment in multiple diseases) and the development of new clinical trial networks** "need to be the future."

FDA proposed “Seamless design and cohort expansion efficiently expedite the clinical development”

- “Approaches like basket trials, master protocols, and seamless trial designs allow us to learn more about new drugs, and even evaluate different drugs in the conduct of the same clinical trial. … Multiple, concurrently accruing and individual cohorts can allow us to better assess a lot of information in one large trial, … With more efficient development using a seamless design, the whole trial can be completed with a few hundred patients. ”

  ----- by Dr. Scott Gottlieb, M.D. Commissioner, Food and Drug Administration, Department of Health and Human Services on July 25, 2018

- FDA released a drafted guidance on multiple expansion cohorts in FIH trials on August 2018.



Expansion Cohorts: Use in First-In-Human Clinical Trials to Expedite Development of Oncology Drugs and Biologics Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Lee Pai-Scherf at 301-796-3400 or (CBER) the Office of Communication, Outreach, and Development at 800-835-4709 or 240-402-8010.

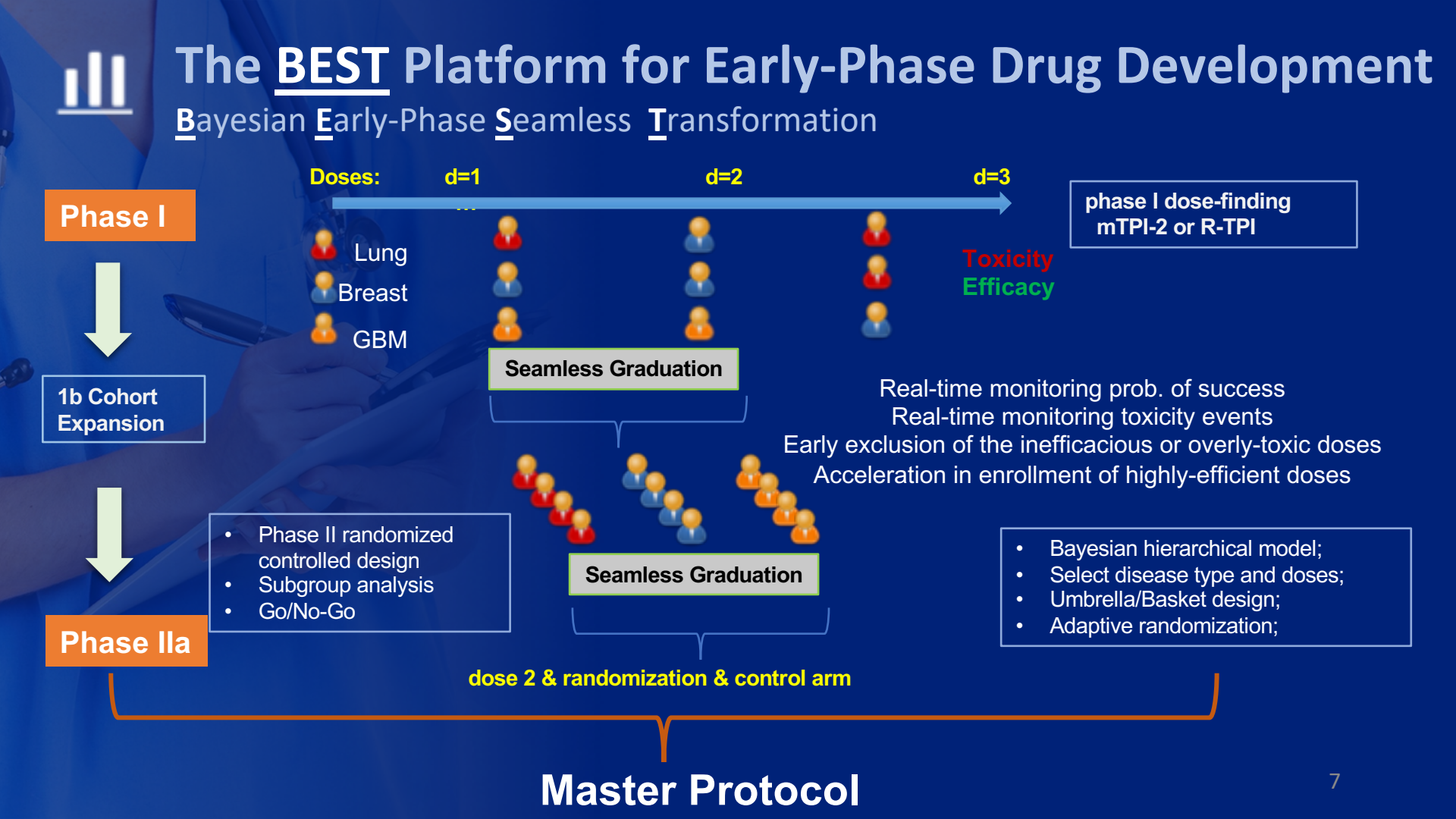
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

August 2018
Procedural

113227bdft.docx
08/16/18

The BEST Platform for Early-Phase Drug Development

Bayesian **E**arly-Phase **S**eamless **T**ransformation



Implementation Detail

Data set

Patient	Tox	Eff	Disease Type	Dose	…
1	0/1	0/1	Lung	d=1	…
2	0/1	0/1	Breast	d=2	…
3	0/1	0/1	GBM	d=2	…
…	…	…	…	…	…

$d = MTD$
$\hat{p} = 0.28$

	Lung	Breast	GBM
Resp.	10	3	2
No Resp.	3	5	7

$d = MTD\text{-}1$
$\hat{p} = 0.2$

	Lung	Breast	GBM
Resp.	…	…	…
No Resp.	…	…	…

$d = MTD+1$
$\hat{p} = 0.35$

	Lung	Breast	GBM
Resp.	…	…	…
No Resp.	…	…	…

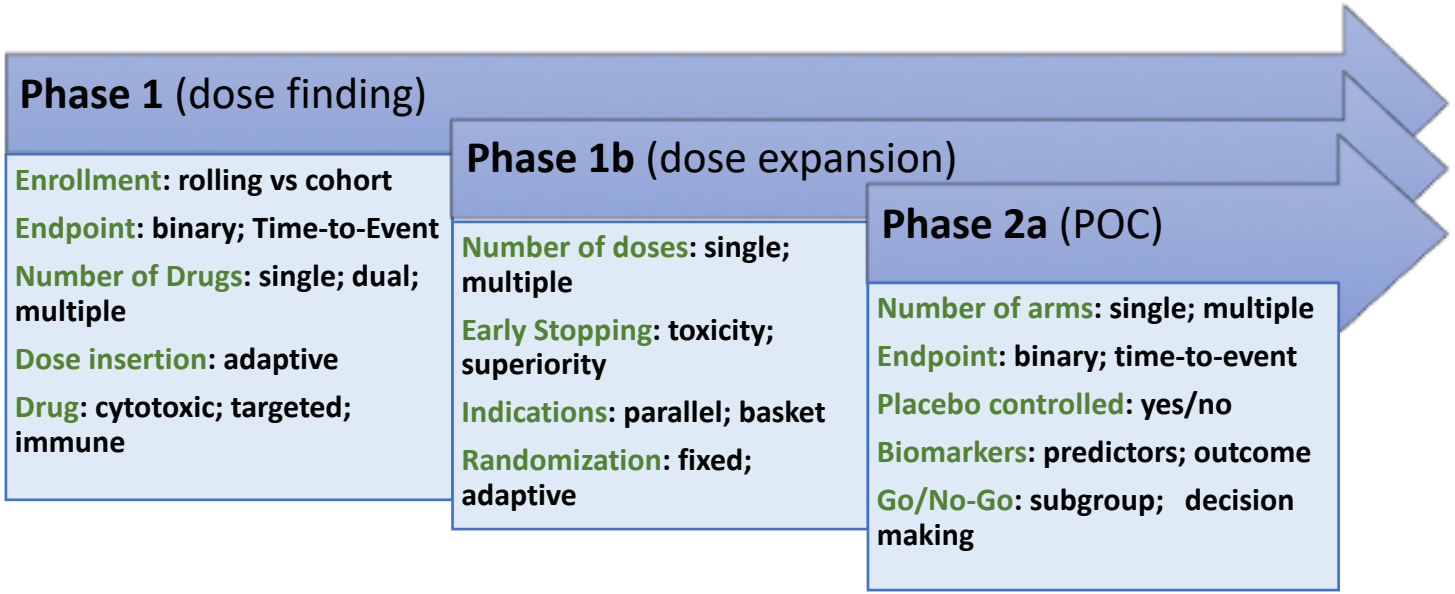
How to decide indication $T^* \in \{L, B, G\}$,
DOSE $D^* \in \{1, \cdots, MTD + 1\}$?

Basket Design
Bayesian Hierarchical Model

dose expansion

- Sample size of each D^*/T^* ?
- Randomization? Equal OR Adaptive?
- Control arm?
- Platform trial design, selection/exclusion of doses

The BEST Platform is a Flexible Module Design



Solution Suites

mTPI-2	single agent	Binomial/stopping boundary	single dose
AAA	dual agents	Parallel group/stopping boundary	multi doses
Laiya001	multi agents	Basket/Umbrella	biomarkers
R-TPI	fast enrollment	SEARS	seamless
TEAMS	dose insertion	Bayesian hierarchical model	AR, Go/No-Go
TEPI	Immune oncology	SCUBA	Subgroup analysis

BEST Scheme Advantages

1. **Multi doses versus single control**
2. **Borrow information across doses and indications**
3. Save sample size
4. Early stop inefficacious arms
5. Reduce time: seamlessly graduate efficacious arms
6. Higher probability of success

The advantages come from three key features

- **Bayesian hierarchical model**
- Seamless strategy
- mTPI-2 in dose finding

- A pilot simulation study: to demonstrate the superiority of the proposed Bayesian hierarchical model
- The advantages of the latter two have been shown in literatures (mTPI-2, Guo et al 2017; SEARS, Pan et al 2014)

Bayesian Hierarchical Model

- The objective of the trial is to test whether the any dose *i* is efficacious in any indication *j*

$$\boldsymbol{H_0: q_{ij} \le q_0 \;\; v.s. \;\; H_1: q_{ij} > q_0,}$$

where q_0 is the response rate that is considered as futile.

- Suppose n_{ij} patients from indication *j* have been enrolled and treated at dose *i*, *among which* y_{ij} patients responded to the treatment. We assume y_{ij} follows a hierarchical model

$$y_{ij} \mid n_{ij}, q_{ij} \sim Bin(n_{ij}, q_{ij})$$
$$q_{ij} = \text{logit}(\theta_{ij})$$
$$\theta_{ij} \mid \lambda_{ij} \sim f_1(\theta_{ij})^{I(\lambda_{ij}=1)} f_2(\theta_{ij})^{I(\lambda_{ij}=2)};$$
$$\lambda_{ij} = \begin{cases} 1; & if\ Z_{ij} \le 0 \\ 2; & if\ Z_{ij} > 0 \end{cases}$$
$$Z_{ij} \sim N(\xi_i + \eta_j, 1)$$

Inference based on λ_{ij} instead of θ_{ij} → automatically handle **multiplicity control**.

Introduce ξ_i and η_j → **borrow information** across doses and indications; the degree of information borrowed from doses and indications could be different.

Simulation Study Overview

- Cohort Expansion of four arms: Two doses and two indications
- Each cohort arm n=20
- Aim to identify efficacious arm for phase II and III trials

- **Power:** select true efficacious arms
- **Type I error:** select true inefficacious arms

- Three methods are compared regarding the power and sample size:
  - Bayesian hierarchical model in BEST (BEST BHM)
  - Bayesian inference based on independent beta binomial distribution for each dose-indication arm (BayInd)
  - Frequentist method for each dose-indication arm (Freq)
- BEST BHM uses a smaller sample size to achieve equal or higher power
- BEST BHM shows higher power than BayInd and Freq at the same Type I error rate

BEST reduces sample size without losing power

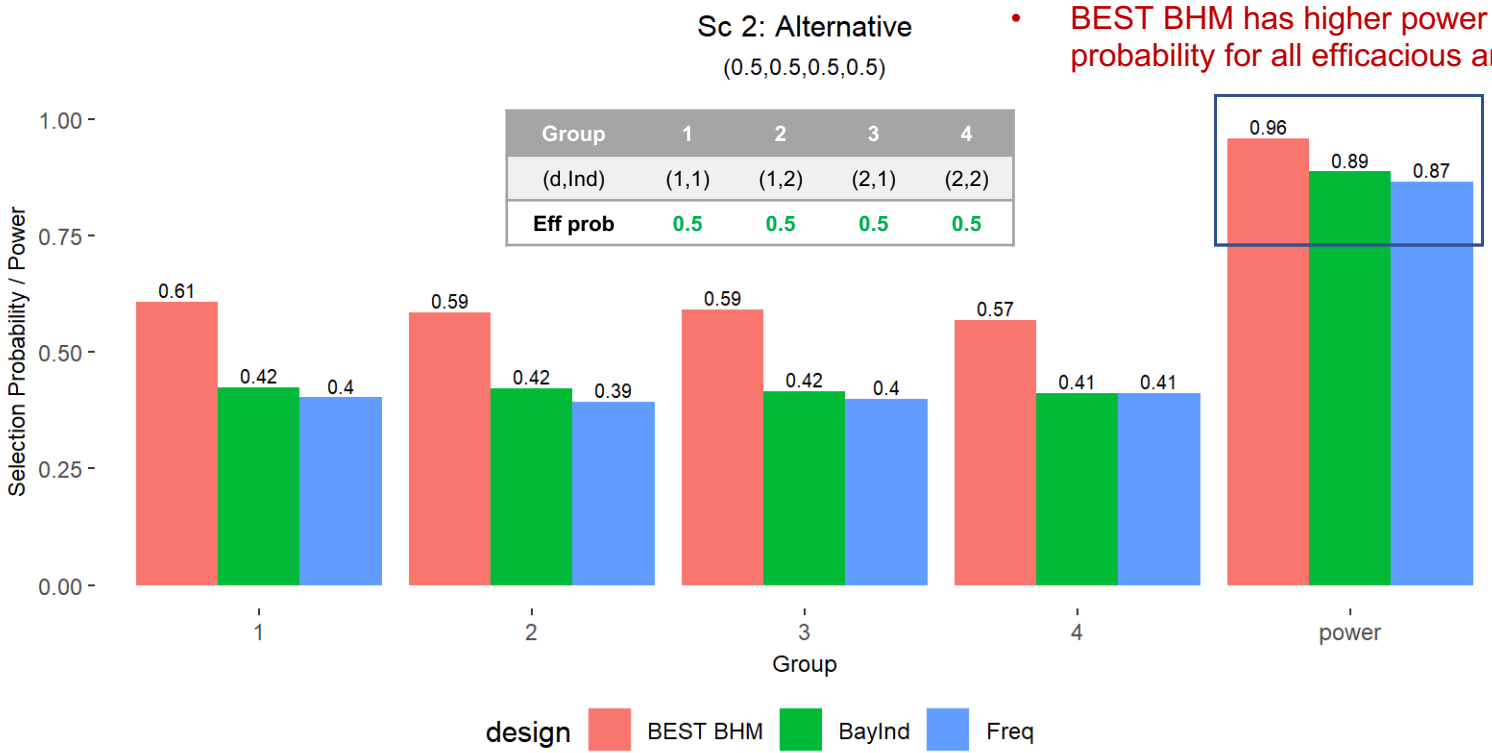
- Sample size reduction of BEST BHM – save 20 subjects; match the same power!

	Method	BEST BHM	BayInd	Freq
All four arms are better than control	Power	0.703	0.709	0.716
	Sample size	80	100	100
	Power	0.771	0.820	0.783
	Sample size	100	120	120

While maintaining similar power, BEST BHM saves about 20% sample size.

Compare power with 20 subjects per arm

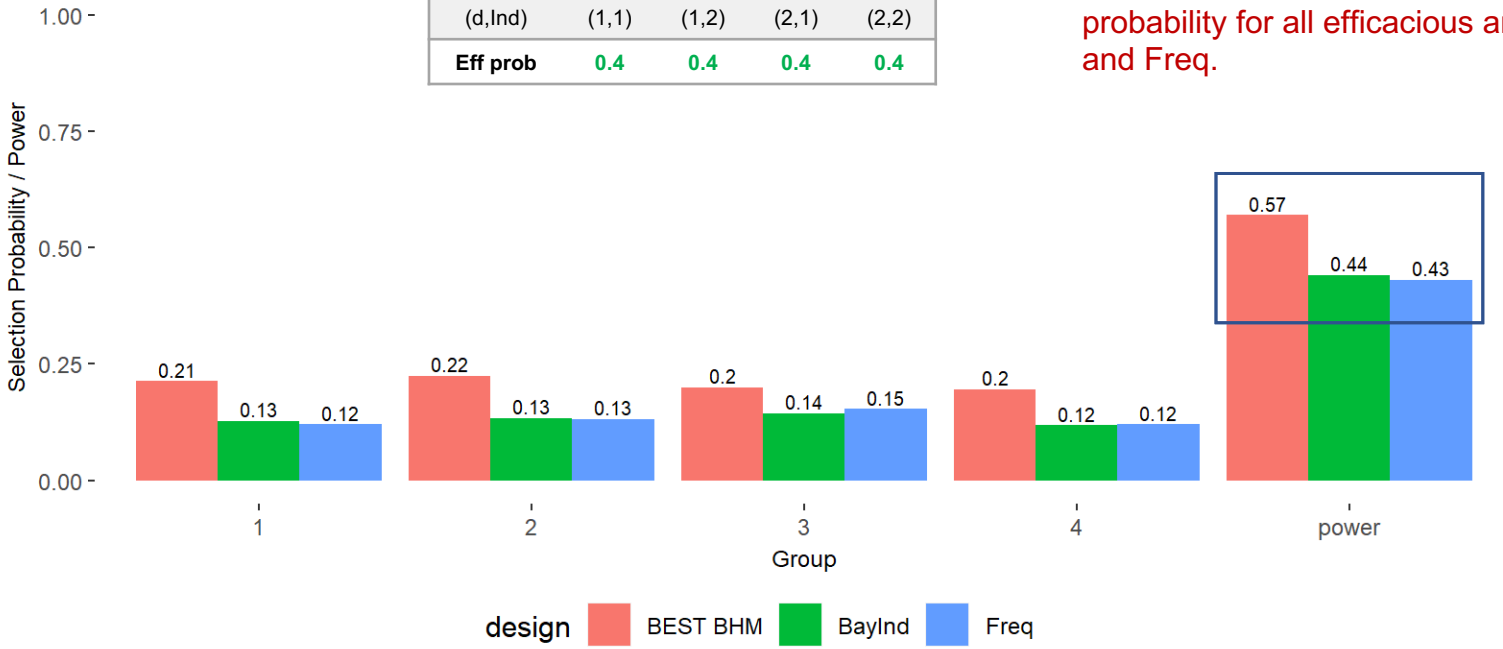
- All four arms are efficacious.
- BEST BHM has higher power and selection probability for all efficacious arms.



Compare power with 20 subjects per arm

Sc 3: All in middle

Group	1	2	3	4
(d,Ind)	(1,1)	(1,2)	(2,1)	(2,2)
Eff prob	**0.4**	**0.4**	**0.4**	**0.4**



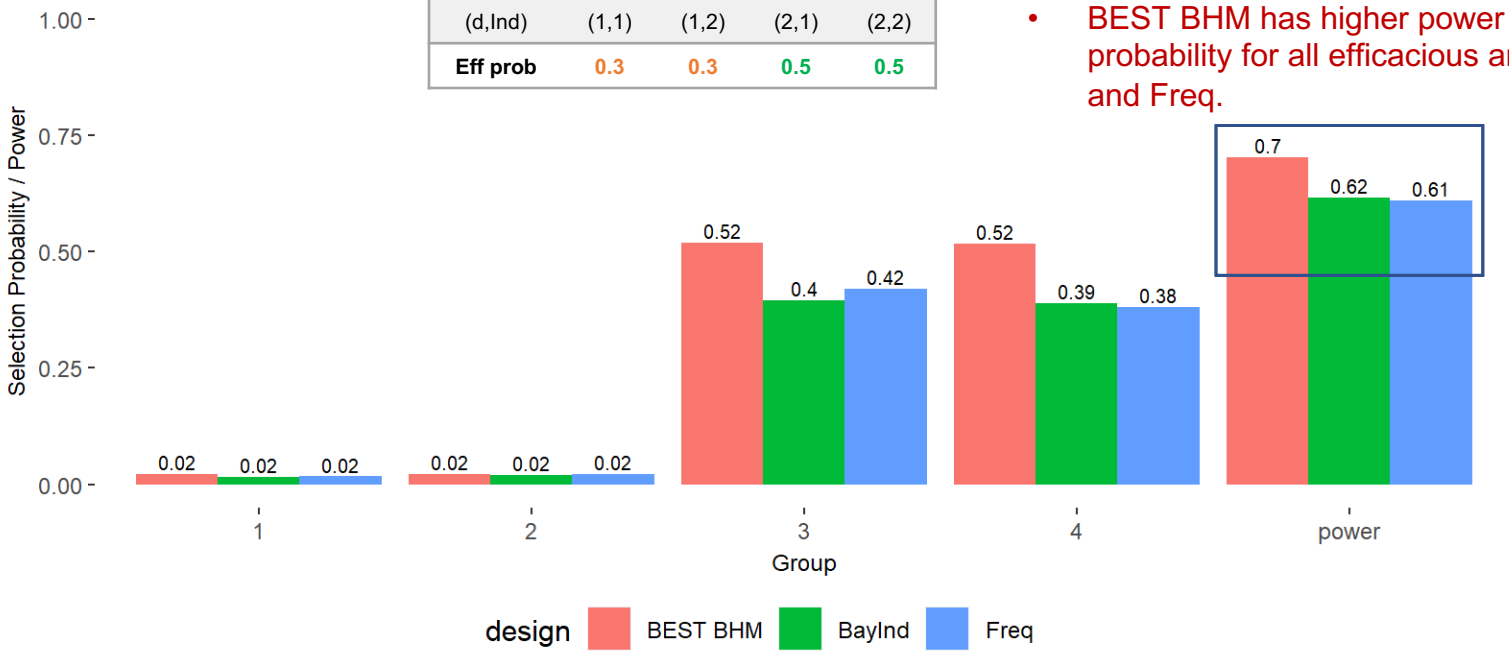
- All arms are efficacious. The effect size is smaller than Sc 2.
- BEST BHM has higher power and selection probability for all efficacious arms than BayInd and Freq.

Compare power with 20 subjects per arm

- One dose is efficacious for both indications, and the other is inefficacious for either indication.
- BEST BHM has higher power and selection probability for all efficacious arms than BayInd and Freq.

Sc 4: 2 null, 2 alternative

Group	1	2	3	4
(d,Ind)	(1,1)	(1,2)	(2,1)	(2,2)
Eff prob	**0.3**	**0.3**	**0.5**	**0.5**



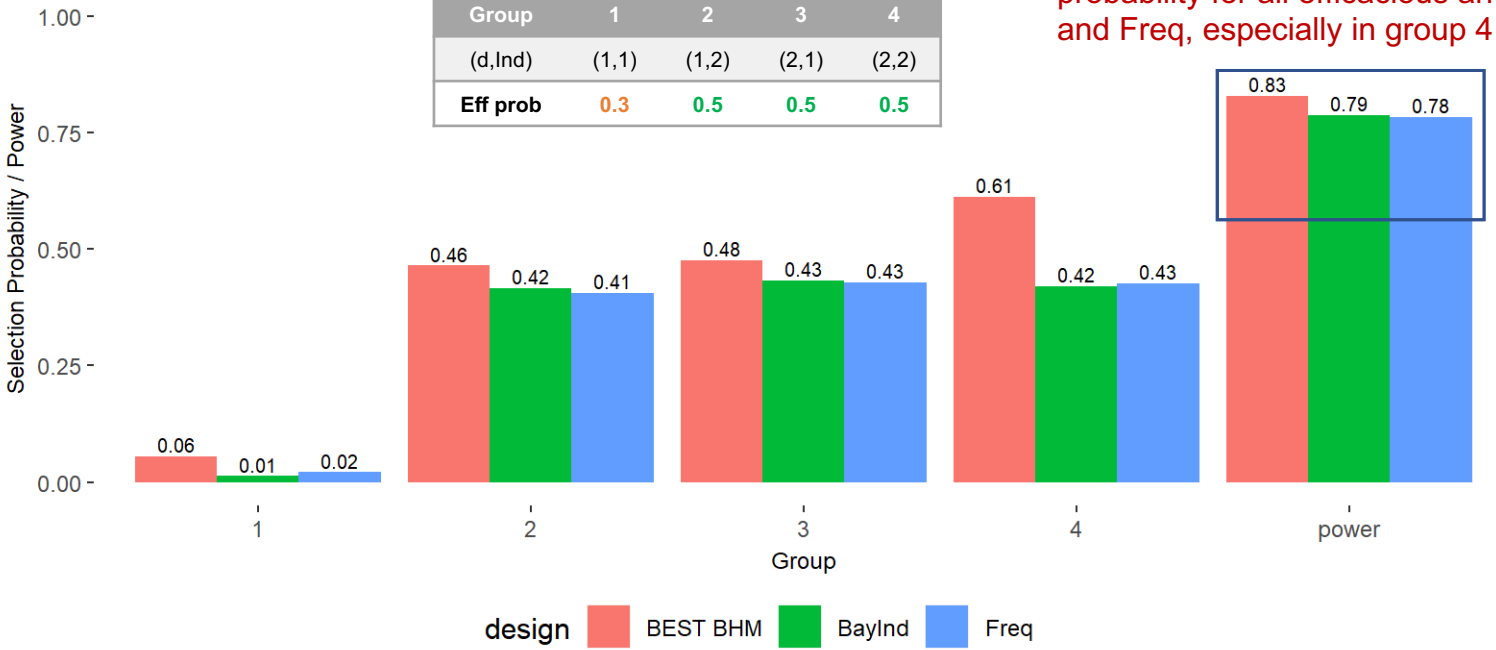
Compare power with 20 subjects per arm

- Other than one dose in one indication, all efficacious.
- BEST BHM has higher power and selection probability for all efficacious arms than BayInd and Freq, especially in group 4 (d2, Ind2)

Sc 5: 1 null, 3 alternative

(0.3,0.5,0.5,0.5)

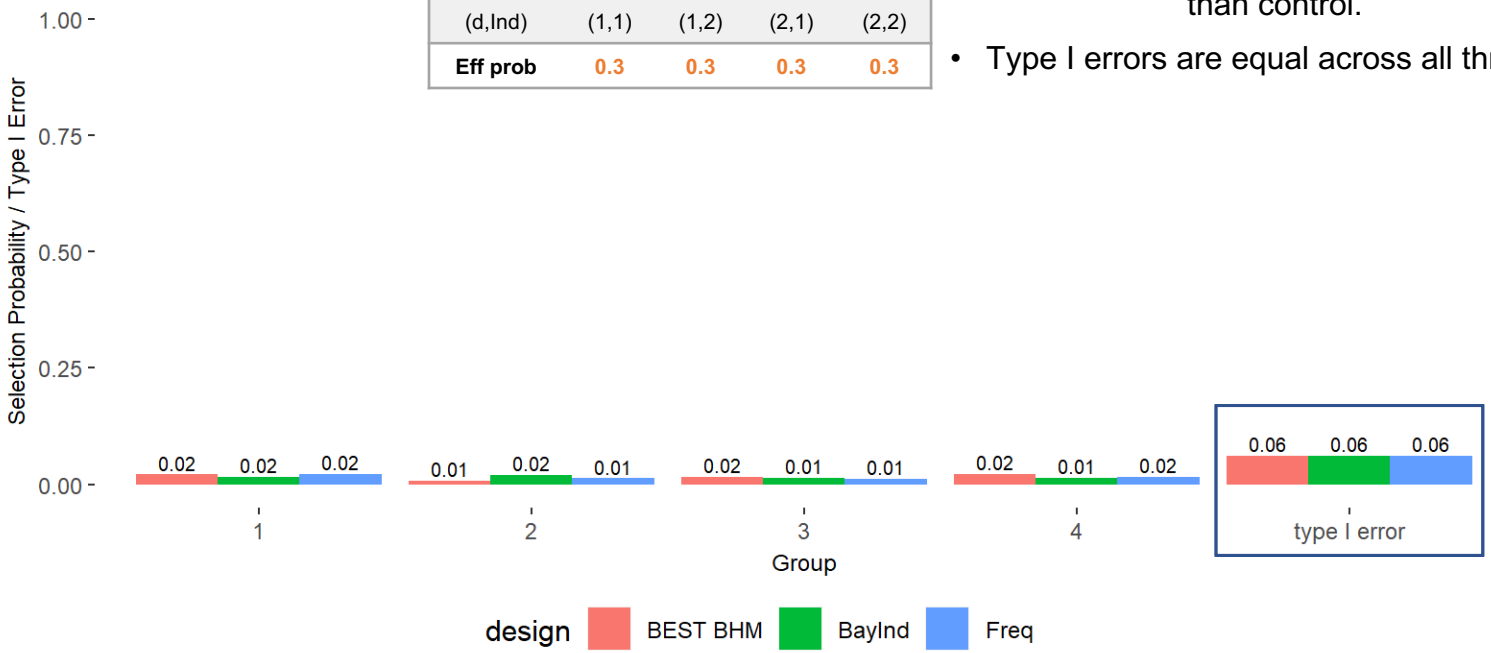
Group	1	2	3	4
(d,Ind)	(1,1)	(1,2)	(2,1)	(2,2)
Eff prob	0.3	0.5	0.5	0.5



Null Results
(n=20; control arm response rate = 0.3)

Sc 1: Null

Group	1	2	3	4
(d,Ind)	(1,1)	(1,2)	(2,1)	(2,2)
Eff prob	0.3	0.3	0.3	0.3



- Null scenario: No arms are more efficacious than control.
- Type I errors are equal across all three designs.

Pilot Simulation Study Summary

- BEST BHM shows higher power than BayInd and Freq when type 1 error is fairly controlled across all three methods
- BEST BHM shows higher probability of selecting efficacious arms than BayInd and Freq
- To achieve similar power, BEST BHM needs smaller sample size than BayInd and Freq.

Bayesian approach automatically handles multiplicity control

- James G. Scott and James O. Berger (2006) proposed that Bayesian approach could automatically handle multiple testing **by introducing a prior distribution for hypotheses** without introducing any penalty term. Because Bayesian testing has a built-in penalty or "Ockham's razor effect" (cf. Jefferys and Berger, 1992).
- They also developed a decision-theoretical approach to control false positive and false negative (i.e. missing the true promising arm) by using a loss function.
- Eg. Gene expression.
  - The hypotheses are $H_1: \mu_i \neq 0$. Define model index parameter $\lambda_i = 1$, if $\mu_i \neq 0$; otherwise, $\lambda_i = 0$.
  - Give prior for p that each $\mu_i = 0$, then make inference based on the posterior probability $1 - p_i = 1 - P(\lambda_i = 0|\boldsymbol{x})$.

The posterior probabilities of the 10 signal means being non-zero as n, the number of noise observations, increases, for two different priors on p

10 Signal observations										
n	−5.65	−5.56	−2.62	−1.20	−1.01	−0.90	−0.15	1.65	1.94	3.57
Uniform prior on p										
25	.97	.97	.71	.31	.28	.26	.20	.43	.51	.88
100	.99	.99	.47	.21	.20	.19	.16	.26	.31	.75
500	1	1	.34	.07	.06	.06	.04	.11	.15	.79
5000	1	1	.11	.02	.02	.02	.01	.03	.04	.42
Prior $\pi(p) = 11p^{10}$										
25	.91	.90	.37	.10	.08	.08	.06	.15	.20	.65
100	.98	.98	.23	.06	.05	.05	.04	.08	.11	.58
500	1	1	.26	.04	.04	.03	.02	.07	.10	.74
5000	1	1	.08	.01	.01	.01	.01	.02	.03	.36

"Objective" prior

"Subjective" prior

- the posterior inclusion probabilities, 1 - *pi*, decrease as the number of "noise" observations, n, grows, so that the same observation implies less evidence of a non-zero mean when more tests are simultaneously considered.
- the prior on *p* does have a strong effect on the *pi*'s. While this effect tends to fade as *n* increases, it becomes negligible only for extremely large *n*. So they recommend to use the subjective prior information, if available.

Thank you!!