BAYES
BAYESIAN BIOSTATISTICS
LYON 2019 21-24 MAY

TUESDAY 21ST
FRIDAY 24TH MAY 2019

MARRIOTT HOTEL CITÉ INTERNATIONALE,
LYON, FRANCE

Bayesian Model-Assisted Designs and Their Applications for Early Phase Clinical Trials – When Simplicity Meets Superiority

J. Jack Lee, Ph.D., Ying Yuan, Ph.D.
Department of Biostatistics
University of Texas MD Anderson Cancer Center

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

Primary Goals for Clinical Trials

- Find safe and efficacious agents
- Provide better treatments to patients enrolled in the trials
- Identify prognostic and predictive markers
- Make accurate and efficient inference

How to do it better?

Adaptive Designs

Adaptive Clinical Trials

Trials that use interim data to guide the study conduct

- Test the safety and efficacy of agents
  - *Adaptive dose finding*
  - *Adaptive estimation of treatment effect*
- Provide better trtms to patients enrolled in the trials
  - *Adaptive randomization*
  - *Adaptive drop/graduate treatments due to toxicity, futility, and/or efficacy*
- Identify prognostic and predictive markers
  - *Adaptive marker identification and validation*
- Make accurate and efficient inference
  - *Adaptive add/drop treatments*
  - *Adaptive decision making*
    - Utility-based

Class of Adaptive Designs

- Rule-based
  - Ad-doc, straightforward, simple to implement
  - No special software needed
- Model-based
  - Good statistical properties
  - More complex and difficult to implement
  - Require special software for the design and analysis
- Model-assisted
  - Based on the underlying statistical model
  - Straightforward, simple to implement
  - Get the best of the two worlds: When simplicity meets superiority
  - The new KISS principle: Keep it Simple and Smart!

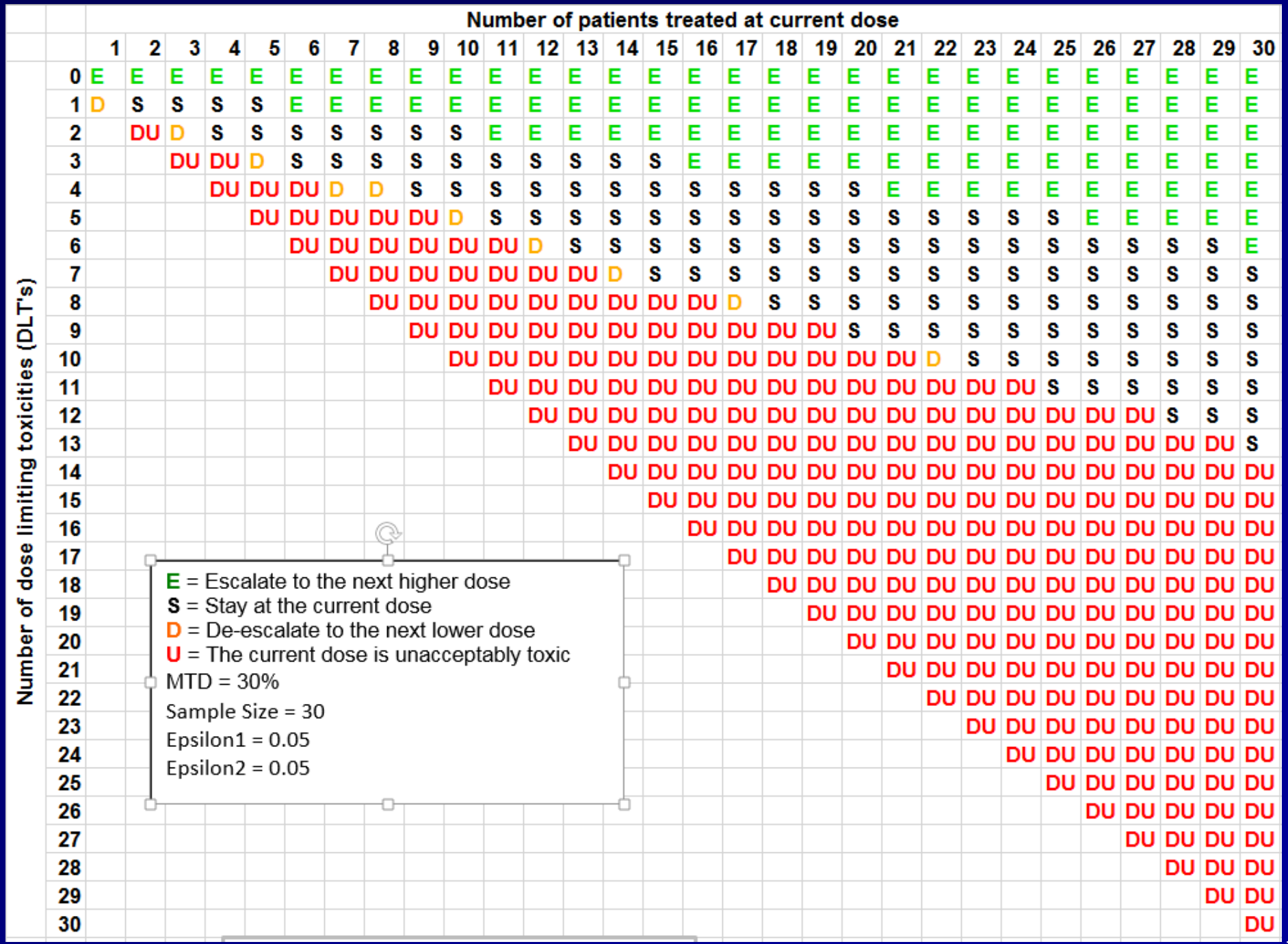
Phase I and Phase II Trial Designs

1. Phase I Designs
   - 3+3 Design
   - CRM, BMA-CRM
   - mTPI, mTPI-2, Keyboard, BOIN, BOIN-COMB, TITE-BOIN
2. Phase II Designs
   - Simon's 2-stage
   - Predictive probability Phase II design
   - BOP2

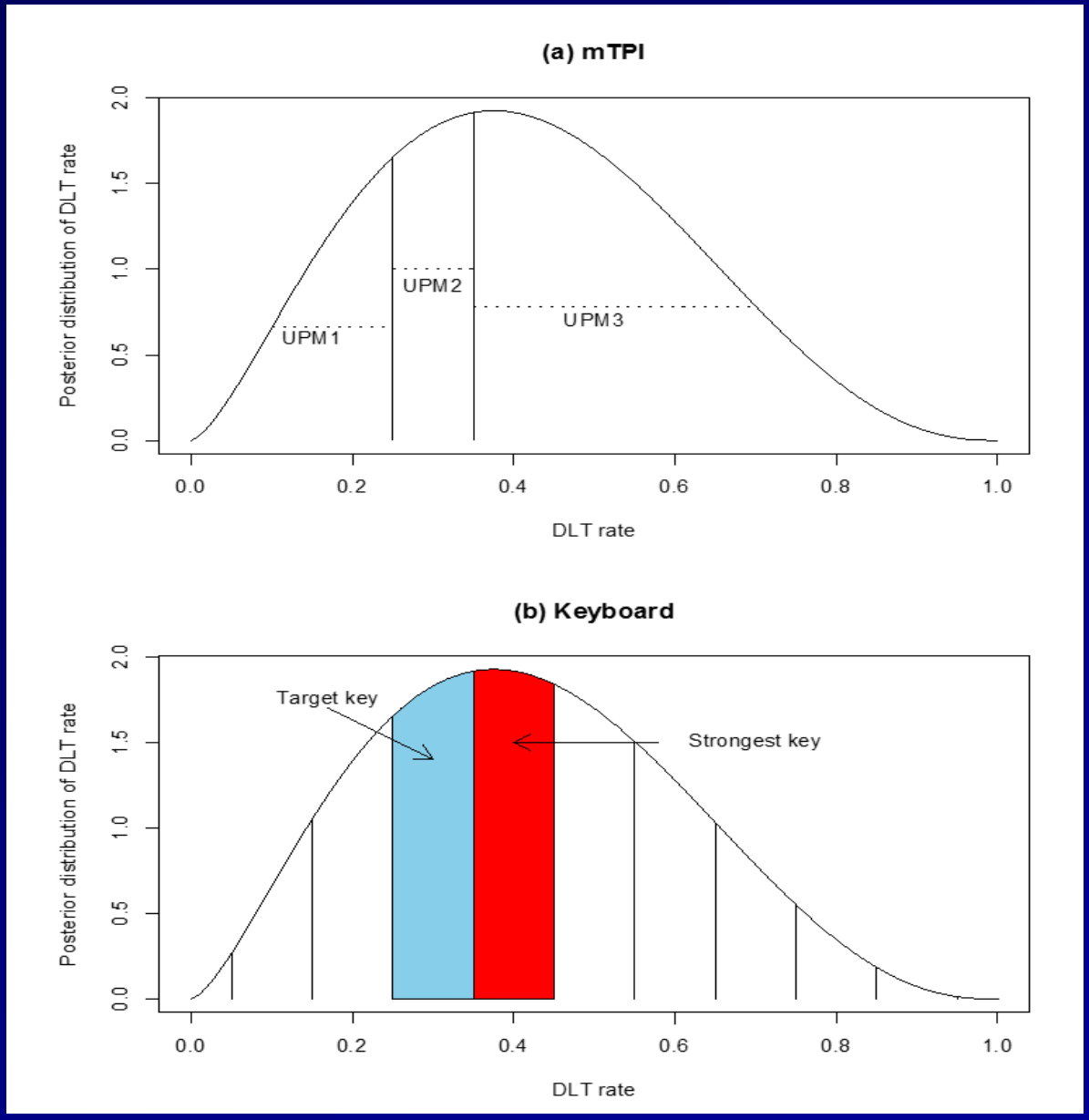
Model-Assisted Design: Modified Toxicity Probability Interval (mTPI), mTPI-2 Design

- Underlying Bayesian models
  - Middle ground between 3+3 design and model based designs, e.g. CRM design
- Posterior toxicity probability space is partitioned into 3 intervals: *high*, *acceptable*, *low* toxicity probability.
- Dose assignment rules determined for all possible outcomes before the trial begins
- Good operating characteristics (mTPI-2)
- Software (R and Excel) are available and easy to implement

Yang S, Wang S, Ji Y. "An integrated dose-finding tool for phase I trials in oncology." Contemporary Clinical Trials, 45, 426–434, 2015.
Guo, Wang, Yang, Lynn, Ji. “A Bayesian interval dose-finding design addressing Ockham's razor: mTPI-2, Contemporary Clinical Trials, 58, 23-33, 2017.


Number of patients treated at current dose
Number of dose limiting toxicities (DLT's)
E = Escalate to the next higher dose
S = Stay at the current dose
D = De-escalate to the next lower dose
U = The current dose is unacceptably toxic
MTD = 30%
Sample Size = 30
Epsilon1 = 0.05
Epsilon2 = 0.05

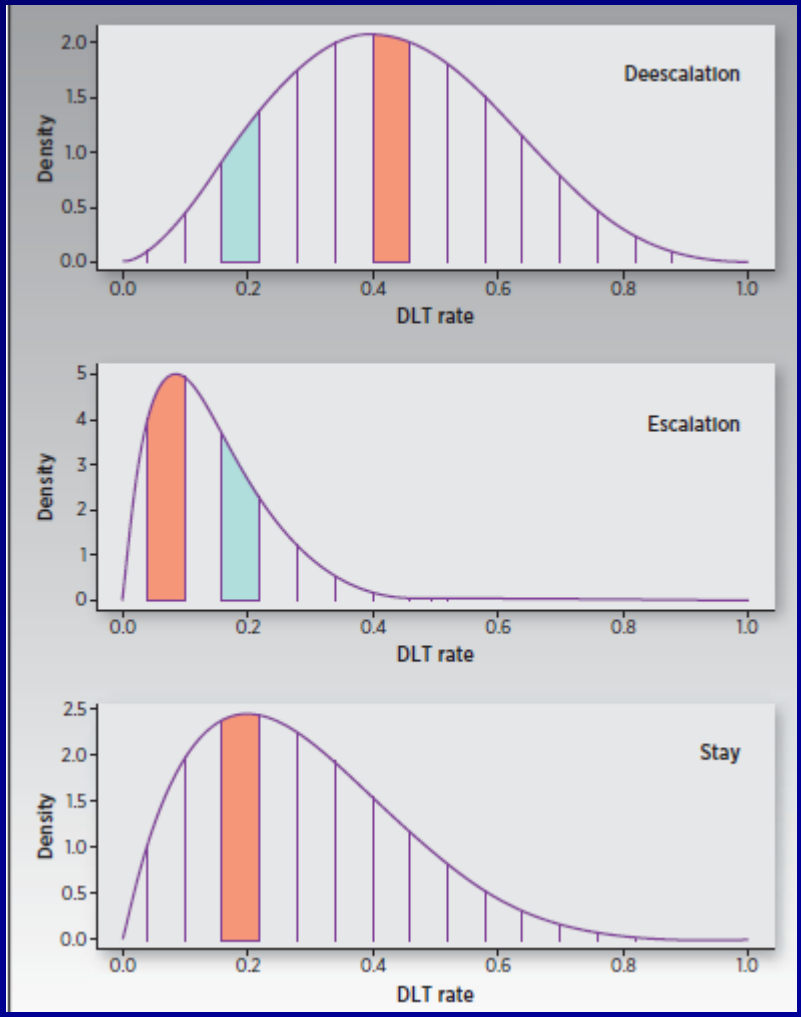
Compare mTPI and Keyboard (mTPI-2) Designs



Keyboard Design

- Define a series of equal-width dosing intervals (or keys) to guide the dose escalation and de-escalation
- E.g,: For targeting 20% DLT
- Target key of (0.15, 0.25) (cyan)
- Current dose (orange)
  - Too high → de-escalate
  - Too low → escalate
  - Same → stay
- Overdose control

  If the observed data shows > 95% prob that current dose is above the MTD, eliminate for further considerations



Keyboard Design
Dose Escalation/De-escalation Rule

Action	The number of patients treated at the current dose																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Target DLT rate =20%[8]																		
Escalate if no. of DLTs ≤	0	0	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	2
De-escalate if no. of DLTs ≥	1	1	1	1	2	2	2	2	3	3	3	3	3	4	4	4	4	5
*Eliminate if no. of DLTs ≥	NA	NA	2	3	3	3	4	4	4	5	5	5	5	6	6	6	7	7
Target DLT rate =30%[8]																		
Escalate if no. of DLTs ≤	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3	4	4
De-escalate if no. of DLTs ≥	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6	7
*Eliminate if no. of DLTs ≥	NA	NA	3	3	4	4	5	5	5	6	6	7	7	8	8	8	9	9

http://trialdesign.org

Clinical Trial Design Software

Filter by: ALL | PHASE I | PHASE II | PHASE I-II | BASKET & PLATFORM | SAMPLE SIZE CALCULATION | EDUCATION

Instructions: To access the software online click the red circle or the title. To download a desktop version, click the download arrow. To expand software description, mouse over the description.

BO

BOIN Suite

Bayesian optimal interval (BOIN) designs provide a novel platform to design phase ***more...***

CRM

CRM & BMA-CRM

The continual reassessment method (CRM) is a model-based dose-finding approach ***more...***

KB

Keyboard (mTPI-2) Design

The keyboard design provides an upgrade to the modified toxicity probability ***more...***

TTK

Time-to-Event Keyboard

The time-to-event keyboard design can handle toxicity data that are pending due ***more...***

S2S

Simon's Two Stage Design

The Simon's two stage design is a commonly used phase II design. It controlls type 1 ***more...***

BOP

Bayesian Optimal Phase 2 (BOP2) Design

The Bayesian optimal phase II (BOP2) design is a flexible Bayesian design that allows ***more...***

TOP

Time-to-Event Bayesian Optimal Phase II Trial Design

The time-to-event Bayesian Optimal Phase II (TOP) design is a flexible and efficient design for phase II clinical trials. ***more...***

PP

Bayesian Efficacy Monitoring with Predictive Probability

Bayesian efficacy monitoring with options of early futility ***more...***

DL

Bayesian Phase 2 Design with Delayed Outcomes

One practical impediment in adaptive phase II trials is that outcomes must be observed soon enough ***more...***

TM

Bayesian Toxicity Monitoring

Bayesian toxicity monitoring for evaluating drug safety.

PO

Bayesian Efficacy Monitoring with Posterior Probability

Bayesian efficacy monitoring with options of early futility and/or efficacy stopping using posterior probability.

OBD

Find Optimal Biological Dose for Immunotherapy

This design is used to find the optimal biological dose (OBD) for molecularly targeted agents and ***more...***

http://trialdesign.org

Keyboard: a novel Bayesian toxicity probability interval design for phase I clinical trial

V1.0.0 ; Last Updated: 6/17/2017

Yanhong Zhou and Ying Yuan

Department of Biostatistics, MD Anderson Cancer Center

Trial Setting | Simulation | Trial Protocol | Select MTD | Reference

How to Use the Keyboard App?

Doses

Number of doses: 5

Starting dose level: 1

Target Probability

Target Toxicity Probability ϕ : 0.3

Acceptable toxicity probability interval: 0.25 – 0.35

Sample Size

Cohort size: 3

Number of cohort: 10

Decision Table

Table 1: Dose escalation/de-escalation rule.

Copy | CSV | Excel | Print

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of patients treated	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Escalate if # of DLT <=	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4
Deescalate if # of DLT >=	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6	7	7	7
Eliminate if # of DLT >=	NA	NA	3	3	4	4	5	5	5	6	6	7	7	8	8	8	9	9	9	10

Dose Escalation/De-escalation Rule

Table 1. Dose escalation/deescalation rule for the keyboard design

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Number of patients treated at the current dose	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Escalate if # of DLT <=	0	0	0	0	1	1	1	1	2	2	2	2	3	3
Deescalate if # of DLT >=	1	1	2	2	2	3	3	3	4	4	4	5	5	5
Eliminate if # of DLT >=	NA	NA	3	3	4	4	5	5	5	6	6	7	7	8

	15	16	17	18	19	20	21	22	23	24	25	26
Number of patients treated at the current dose	15	16	17	18	19	20	21	22	23	24	25	26
Escalate if # of DLT <=	3	3	4	4	4	4	5	5	5	5	6	6
Deescalate if # of DLT >=	6	6	6	7	7	7	8	8	9	9	9	10
Eliminate if # of DLT >=	8	8	9	9	9	10	10	11	11	11	12	12

	27	28	29	30
Number of patients treated at the current dose	27	28	29	30
Escalate if # of DLT <=	6	6	7	7
Deescalate if # of DLT >=	10	10	11	11
Eliminate if # of DLT >=	12	13	13	14

Trial Setting | Simulation | Trial Protocol | Select MTD | Reference

Simulation

Method to enter simulation scenarios:

- (•) Type in
- ( ) Upload scenario file

Enter Simulation Scenarios

Add a Scenario | Remove a Scenario | Save Scenarios

Number of Simulations: 1000

Set Seed: 6

For each scenario, enter true toxicity rate of each dose level:

	D1	D2	D3	D4	D5
Scenario 1	0.30	0.47	0.53	0.58	0.64
Scenario 2	0.01	0.11	0.30	0.45	0.67
Scenario 3	0.02	0.07	0.13	0.30	0.47
Scenario 4	0.05	0.08	0.12	0.15	0.30

Run Simulation

Operating Characteristics

Copy | CSV | Excel | Print

Search:

	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Number of Patients	% Early Stopping
Scenario1							
True DLT rate	0.3	0.47	0.53	0.58	0.64		
Selection %	67.3	12.1	2.6	0.2	0		17.8
# Pts treated	18.86	6.52	1.12	0.14	0.02	26.6	
Scenario2							
True DLT rate	0.01	0.11	0.3	0.45	0.67		
Selection %	0.2	18.6	59.5	21.1	0.6		0
# Pts treated	3.32	8.37	12.16	5.46	0.69	30	
Scenario3							
True DLT rate	0.02	0.07	0.13	0.3	0.47		
Selection %	0.1	0.9	21	59.2	18.8		0
# Pts treated	3.28	4.26	7.75	10.13	4.58	30	
Scenario4							

Protocol Template

Template for Protocol Preparation

We will employ the keyboard design (Yan, Mandrekar and Yuan, 2017) to find the MTD. The keyboard design is a novel Bayesian interval design that can be implemented in a simple way similar to the traditional 3+3 design, but is more flexible and possesses superior operating characteristics that are comparable to those of the more complex model-based designs, such as the continual reassessment method (CRM). The keyboard design provides an upgrade to the modified toxicity probability interval (mTPI) design, with a substantially lower risk of overdosing and a better precision to identify the MTD.

The target toxicity rate for the MTD is $\phi = 0.3$, with the acceptable toxicity probability interval of (0.25,0.35). The maximum sample size is 30. We will enroll and treat patients in cohorts of size 3. The keyboard design is described as follows:

1. Patients in the first cohort are treated at dose level 1.
2. To assign a dose to the next cohort of patients, conduct dose escalation/de-escalation according to the rule displayed in Table 1. When using table 1, please note the following:
    a. "Eliminate" means that we eliminate the current and higher doses from the trial to prevent treating any future patients at these doses because they are overly toxic
    b. When we eliminate a dose, automatically de-escalate the dose to the next lower level. When the lowest dose is eliminated, stop the trial for safety. In this case, no dose should be selected as the MTD.
    c. If none of the actions (i.e., escalation, de-escalation or elimination) is triggered, we treat the new patients at the current dose.
    d. If the current dose is the lowest dose and the rule indicates dose de-escalation, treat the new patients at the lowest dose unless the number of DLTs reaches the elimination boundary, at which point terminate the trial for safety.
    e. If the current dose is the highest dose and the rule indicates dose escalation, treat the new patients at the highest dose.
3. Repeat step 2 until the maximum sample size of 30 is reached.

Operating Characteristics

Operation Characteristics

Table 2 shows the operating characteristics of the trial design based on 1000 simulations of the trial using shiny app "Keyboard" available at http://www.trialdesign.org. The operating characteristics show that the design selects the true MTD, if any, with high probability and allocates more patients to the dose levels with the DLT rate closest to the target of 0.3.

Table 2. Operating characteristics of the keyboard design

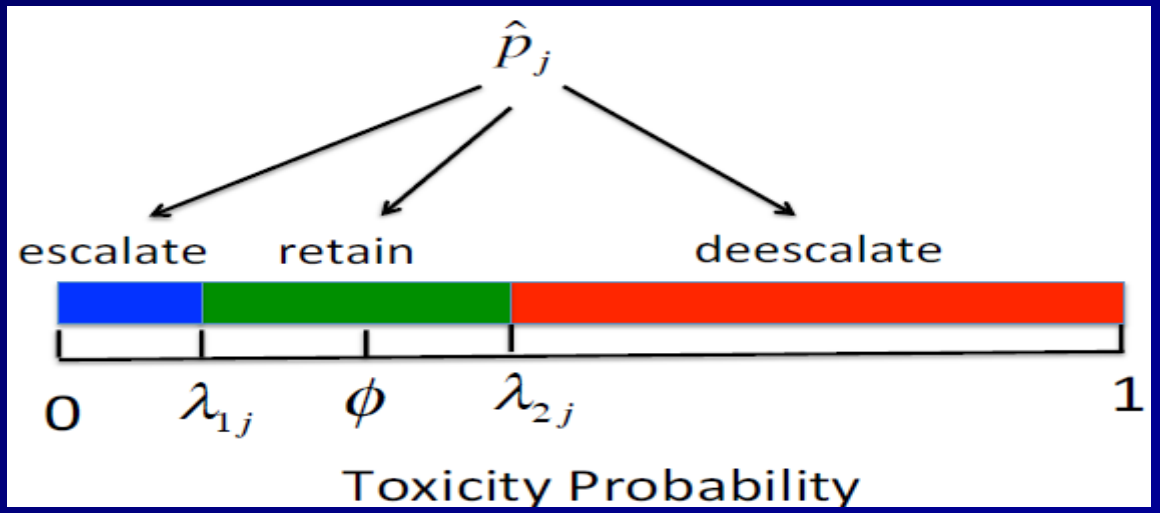
	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Number of Patients	% Early Stopping
Scenario 1							
True DLT rate	0.3	0.47	0.53	0.58	0.64		
Selection %	67.3	12.1	2.6	0.2	0		17.8
# Pts treated	18.86	6.52	1.12	0.14	0.02	26.6	
Scenario 2							
True DLT rate	0.01	0.11	0.3	0.45	0.67		
Selection %	0.2	18.6	59.5	21.1	0.6		0
# Pts treated	3.32	8.37	12.16	5.46	0.69	30	
Scenario 3							
True DLT rate	0.02	0.07	0.13	0.3	0.47		
Selection %	0.1	0.9	21	59.2	18.8		0
# Pts treated	3.28	4.26	7.75	10.13	4.58	30	

Reference

Yan, F., Mandrekar, S. J. & Yuan, Y. (2017). Keyboard: A Novel Bayesian Toxicity Probability Interval Design for Phase I Clinical Trials. Clinical Cancer Research, doi: 10.1158/1078-0432.CCR-17-0220.

Bayesian Optimal Interval (BOIN) Design

- With the target probability of toxicity ϕ, an interval design makes decision of dose escalation, stay, or de-escalation by comparing the estimated probability of toxicity $\hat{p}_j$ at dose j with a pre-specified toxicity interval.



- The interval boundaries λ_{1j} and λ_{2j} are selected to minimize the decision error of dosing.

Liu S, Yuan Y. Bayesian optimal interval designs for phase I clinical trials. *Appl. Statist.* (2015) 64, 507–523

Optimal Interval Boundaries

- Assume the prior probability of the 3 decisions are equal – A simple, yet, powerful result:

Table: The values of λ_{1j} and λ_{2j} under the BOIN design for different target toxicity rates.

Interval boundaries	Target toxicity rate ϕ					
	0.15	0.2	0.25	0.3	0.35	0.4
λ_{1j}	0.118	0.157	0.197	0.236	0.276	0.316
λ_{2j}	0.179	0.238	0.298	0.358	0.419	0.479

- The optimal dose escalation/deescalation boundaries are independent of n_j and j !!

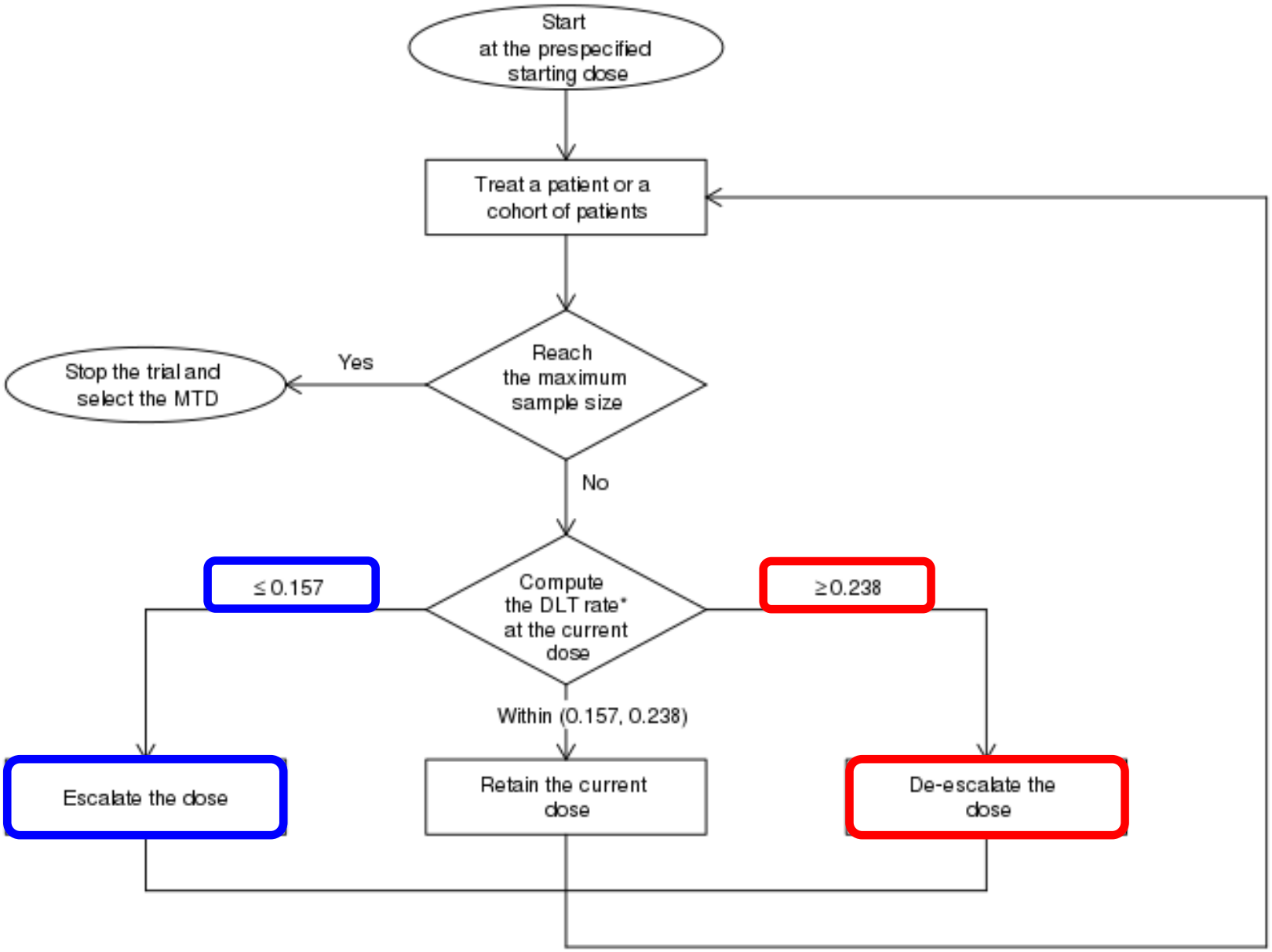
BOIN Design with ϕ= 0.2

- The first cohort are treated at the lowest dose
- At the current dose level *j*

Action	Number of patients treated at the current dose											
	1	2	3	4	5	6	7	8	9	10	11	12
Escalate if # of DLT $\leq$	0	0	0	0	0	0	1	1	1	1	1	1
Deescalate if # of DLT $\geq$	1	1	1	1	2	2	2	2	3	3	3	3

- Repeat step 2 until reaching the maximum sample size

The operating characteristics is much better than the 3+3 design and comparable to the continual reassessment method (CRM) design



$$* \text{ DLT rate} = \frac{\text{Total number of patients who experienced DLT at the current dose}}{\text{Total number of patients treated at the current dose}}$$

Bayesian Optimal Interval (BOIN) Design for Phase I Clinical Trials

PID: 979 ; V1.0.3.8 ; Last Updated: 02/20/2019

Yanhong Zhou, Suyu Liu, Ying Yuan and Heng Zhou

Department of Biostatistics, MD Anderson Cancer Center

Trial Setting | Simulation | Trial Protocol | Select MTD | Reference

How to Use the BOIN App

Doses

Number of doses: 5

Starting dose level: 1

Target Probability

Target Toxicity Probability ϕ : 0.3

☑ Use the default alternatives to minimize decision error (recommended).

Sample Size

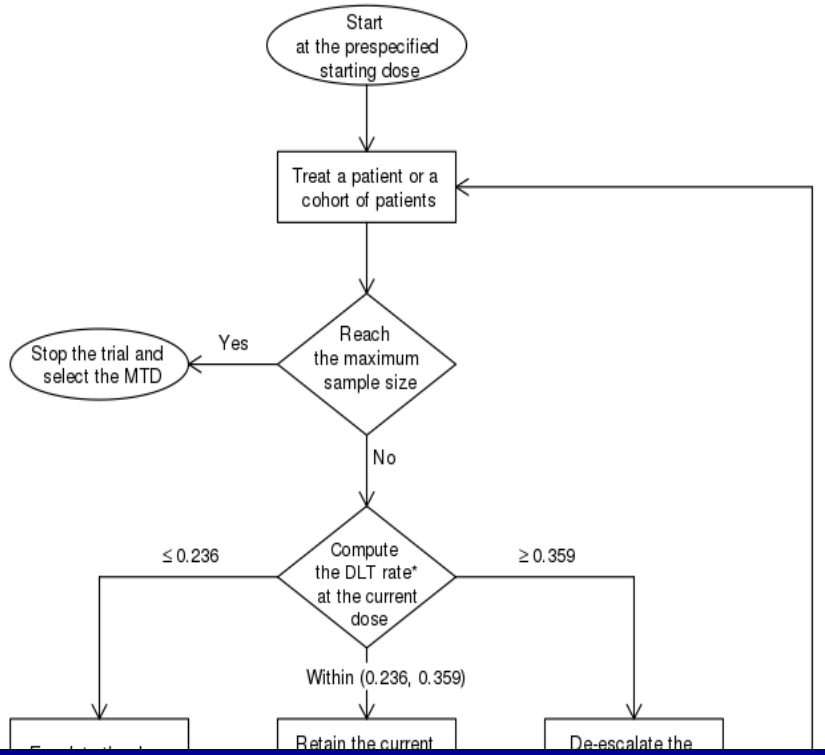
Cohort size: 3

Number of cohort: 10

Stop trial if number of patients assigned to single dose reaches: 30

Perform accelerated titration:

Design Flow Chart | Decision Table



Template for Protocol Preparation

We will employ the Bayesian optimal interval (BOIN) design (Liu and Yuan, 2015; Yuan et al., 2016) to find the MTD. The BOIN design is implemented in a simple way similar to the traditional 3+3 design, but is more flexible and possesses superior operating characteristics that are comparable to those of the more complex model-based designs, such as the continual reassessment method (CRM) (Zhou, Yuan and Nie, 2018).

The target toxicity rate for the MTD is $\phi = 0.2$ and the maximum sample size is 30. We will enroll and treat patients in cohorts of size 3. To guide dose-escalation decisions, if the observed DLT rate at the current dose is ≤ 0.157, the next cohort of patients will be treated at the next higher dose level; if it is ≥ 0.238, the next cohort of patients will be treated at the next lower dose level. For the purpose of overdose control, doses j and higher levels will be eliminated from further examination if $\Pr(p_j > 0.2 \mid \text{data}) > 0.95$, where p_j is the true DLT rate of dose level j, $j = 1, \cdots, 5$. When the lowest dose is eliminated, stop the trial for safety. The trial design is illustrated in Figure 1 and described through the following three steps:

1. Patients in the first cohort are treated at dose level 1.
2. To assign a dose to the next cohort of patients, conduct dose escalation/de-escalation according to the rule displayed in Table 1. When using Table 1, please note the following:
    a. "Eliminate" means that we eliminate the current and higher doses from the trial to prevent treating any future patients at these doses because they are overly toxic.
    b. When we eliminate a dose, automatically de-escalate the dose to the next lower level. When the lowest dose is eliminated, stop the trial for safety. In this case, no dose should be selected as the MTD.
    c. If none of the actions (i.e., escalation, de-escalation or elimination) is triggered, we treat the new patients at the current dose.
    d. If the current dose is the lowest dose and the rule indicates dose de-escalation, treat the new patients at the lowest dose unless the number of DLTs reaches the elimination boundary, at which point terminate the trial for safety.
    e. If the current dose is the highest dose and the rule indicates dose escalation, treat the new patients at the highest dose.
3. Repeat step 2 until the maximum sample size of 30 is reached.

Table 1. Dose escalation/deescalation rule for the BOIN Design

Number of patients treated at the current dose	1	2	3	4	5	6	7
Escalate if # of DLT <=	0	0	0	0	0	0	1
Deescalate if # of DLT >=	1	1	1	1	2	2	2
Eliminate if # of DLT >=	NA	NA	2	3	3	3	4

Number of patients treated at the current dose	8	9	10	11	12	13	14
Escalate if # of DLT <=	1	1	1	1	1	2	2
Deescalate if # of DLT >=	2	3	3	3	3	4	4
Eliminate if # of DLT >=	4	4	5	5	5	5	6

Number of patients treated at the current dose	15	16	17	18	19	20	21
Escalate if # of DLT <=	2	2	2	2	2	3	3
Deescalate if # of DLT >=	4	4	5	5	5	5	6
Eliminate if # of DLT >=	6	6	7	7	7	7	8

Number of patients treated at the current dose	22	23	24	25	26	27	28
Escalate if # of DLT <=	3	3	3	3	4	4	4
Deescalate if # of DLT >=	6	6	6	6	7	7	7
Eliminate if # of DLT >=	8	8	8	9	9	9	9

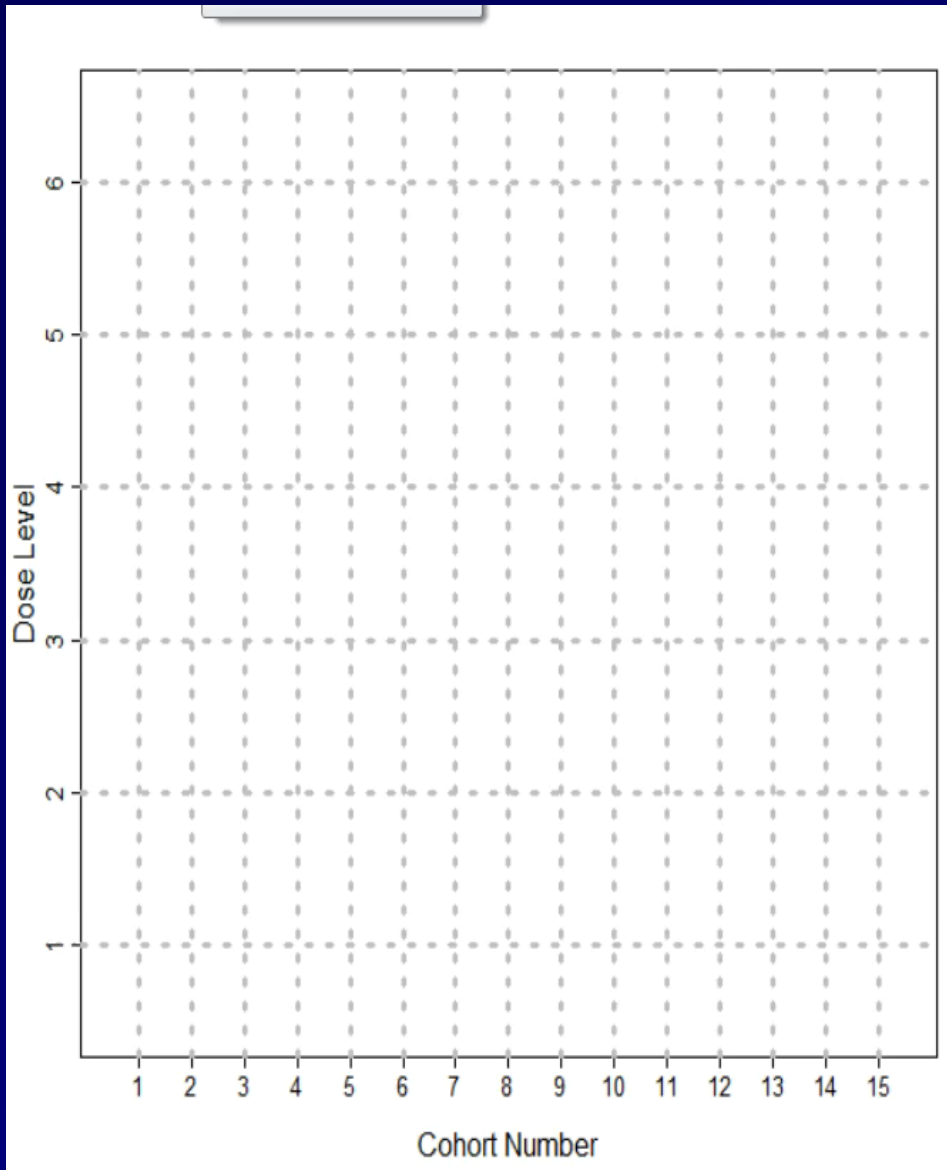
Number of patients treated at the current dose	29	30
Escalate if # of DLT <=	4	4
Deescalate if # of DLT >=	7	8
Eliminate if # of DLT >=	10	10

Note: # of DLT is the number of patients with at least 1 DLT.

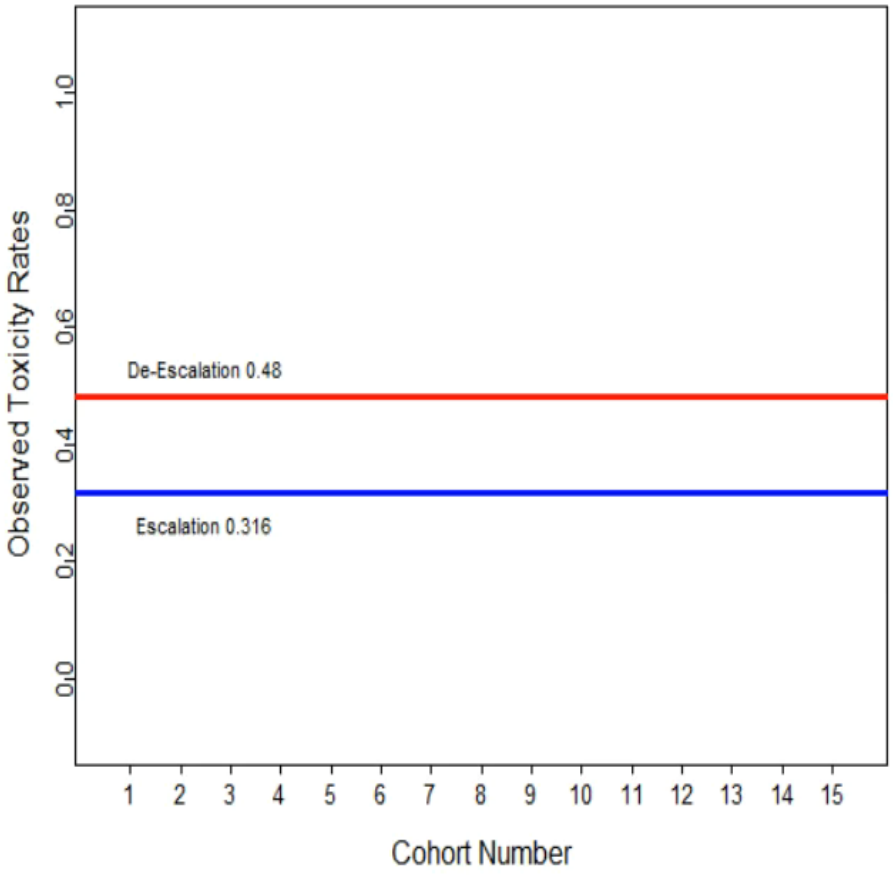
Table 2. Operating characteristics of the BOIN design

	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Number of Patients	% Early Stopping
Scenario 1							
True DLT rate	0.2	0.37	0.43	0.48	0.54		
Selection %	65.6	12.1	1.1	0	0		21.2
# Pts treated	19.17	5.21	0.97	0.13	0	25.5	
Scenario 2							
True DLT rate	0.01	0.07	0.2	0.35	0.57		
Selection %	2.1	26.2	55.3	15.8	0.6		0
# Pts treated	4.45	10.07	10.85	4.12	0.52	30	
Scenario 3							
True DLT rate	0.01	0.04	0.08	0.2	0.37		
Selection %	0.4	4.3	27.7	54.2	13.4		0
# Pts treated	3.64	5.34	8.7	9	3.33	30	
Scenario 4							
True DLT rate	0.02	0.04	0.07	0.09	0.2		
Selection %	0.5	3.8	7.8	32.5	55.2		0.2
# Pts treated	4.01	5.11	5.64	7	8.19	29.9	

Video 1: Illustrating BOIN Design



Number of Dose Level: 6
Target Toxicity: 0.4
Cohort Size: 3
Number of Cohorts: 15
Toxicity Level:
0.05, 0.1, 0.2, 0.3, 0.4, 0.6



Bayesian Optimal Interval (BOIN) Design for Drug Combination Trials: BOIN-COMB

- Similar to the BOIN design but allow two-dimensional dose escalation/stay/de-escalation.
- Treat the 1st cohort at the lowest dose (1, 1).
- To determine the next dose combination: Maximizing the posterior probability that the toxicity rate of the next dose falls inside a pre-specified probability interval based on the cumulative data.
- After the trial is completed, perform an isotonic regression to estimate toxicity rates satisfy the monotonicity assumption when fixing one drug at a certain dose level.

Lin R. and Yin, G. (2015). Bayesian Optimal Interval Design for Dose Finding in Drug-combination Trials, Statistical Methods in Medical Research

BOIN for Drug Combination Trials

Trial Setting | Simulation | Trial Protocol | Next Dose/Subtrial | Select MTD | Reference

How to Use the BOIN App for Combination Trial?

Doses

	Drug A	Drug B
Number of Doses:	3	5
Starting Dose Level:	1	1

Target Probability

Target toxicity probability ϕ :

0.3

☑ Use the default alternatives to minimize decision error (recommended).

Find:

◉ Single MTD

○ MTD Contour

Sample Size

Number of cohorts:

10

Cohort size:

3

Stop trial if the # of patients assigned to the current dose reaches:

15

Design Flow Chart | Decision Table

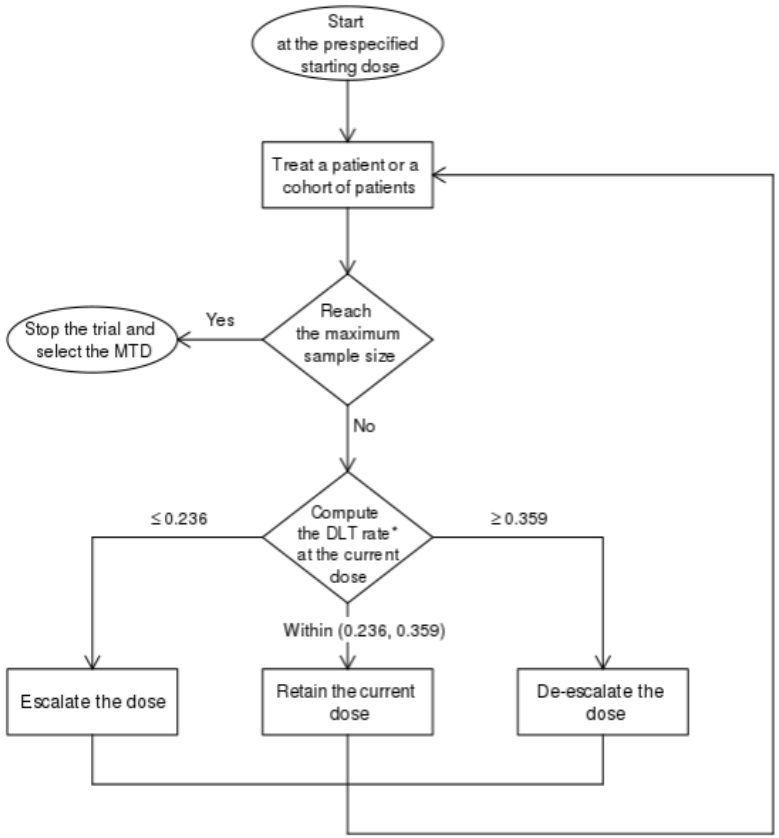
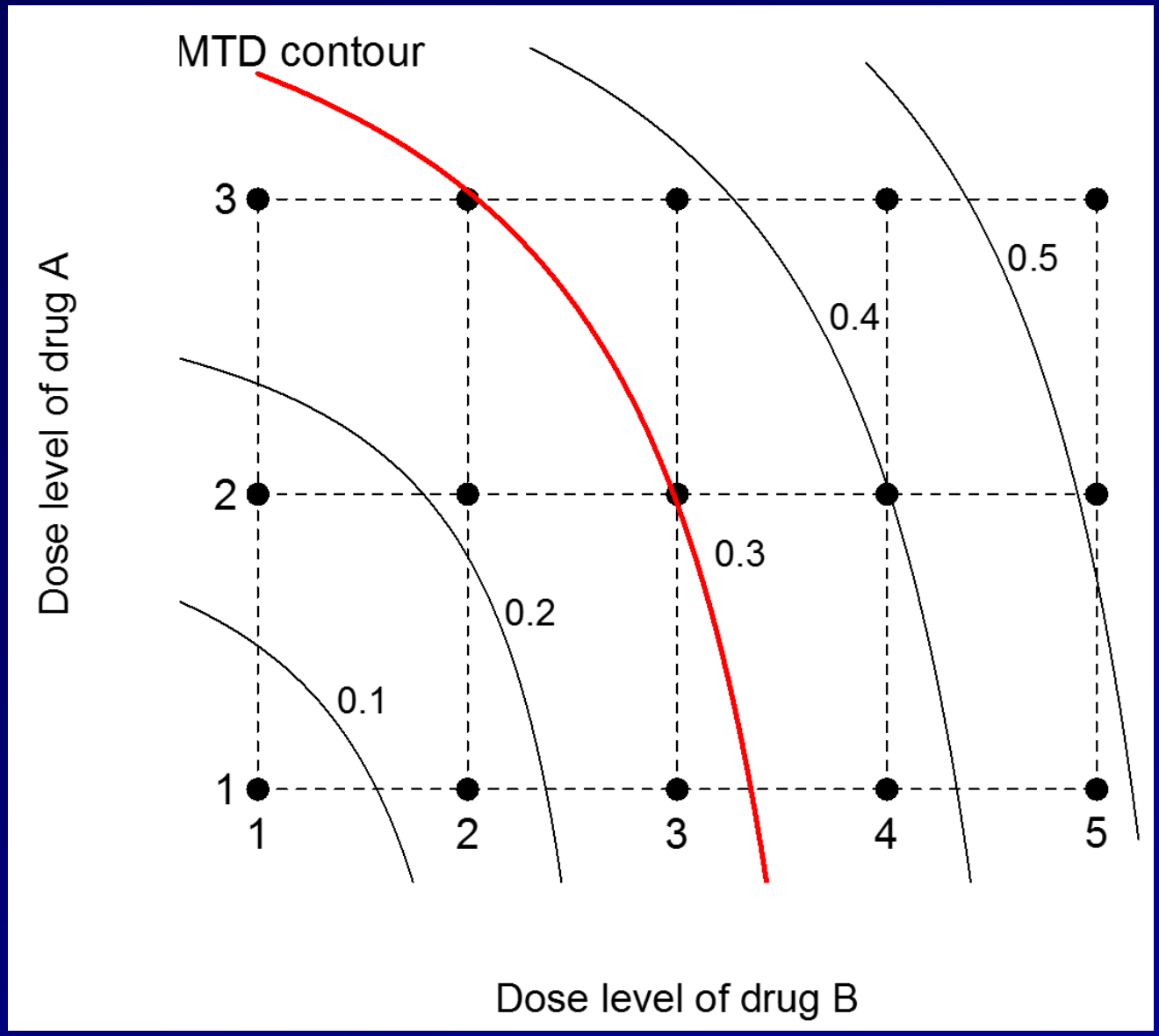


Figure 1. Flowchart for a trial conduct using the BOIN design



MTD contour in drug combination trials. Curved lines indicated the toxicity contours with true toxicity rates of 0.1, 0.2, 0.3, 0.4 and 0.5, respectively. Combinations located along the rows and columns are ordered in toxicity, but in other directions of the dose matrix (e.g., along the diagonals from the upper left corner to the lower right corner), the toxicity order is unknown due to unknown drug-drug interactions.

Bayesian Optimal Interval (BOIN) Design for Time-To-Event Endpoints: TITE-BOIN

- Enroll the first patient cohort at the lowest or prespecified starting dose.
- Based on the data observed at the current dose, make the dose-escalation/deescalation decision according to the pregenerated decision table for treating the next patient cohort.
- Repeat step 2 until the prespecified maximum sample size is reached and select the MTD using the statistical method isotonic regression.

Yuan, Y., Lin, R., Li, D., Nie, L. and Warren, K.E. (2018). Time-to-event Bayesian Optimal Interval Design to Accelerate Phase I Trials. Clinical Cancer Research, 24(20): 4921-4930.

BOIN for Time-To-Event Endpoint

Trial Setting | Simulation | Trial Protocol | STFT Calculator | Reference

How to Use the TITE-BOIN App

Doses

Number of doses: 5

Starting dose level: 1

Target Probability

Target Toxicity Probability ϕ : 0.3

☑ Use the default alternatives to minimize decision error (recommended).

Sample Size and Accrual Rate

Cohort size: 3

Number of cohort: 10

Stop trial if number of patients assigned to single dose reaches: 15

DLT assessment window: 3 month

Accrual rate / month: 1

☑ Use the default uniform prior for the time to toxicity.

Overdose Control

Eliminate dose j if $Pr(p_j > \phi \mid data) > p_E$

Use the default cutoff (recommended) p_E = 0.95

Decision Table | Design Flow Chart

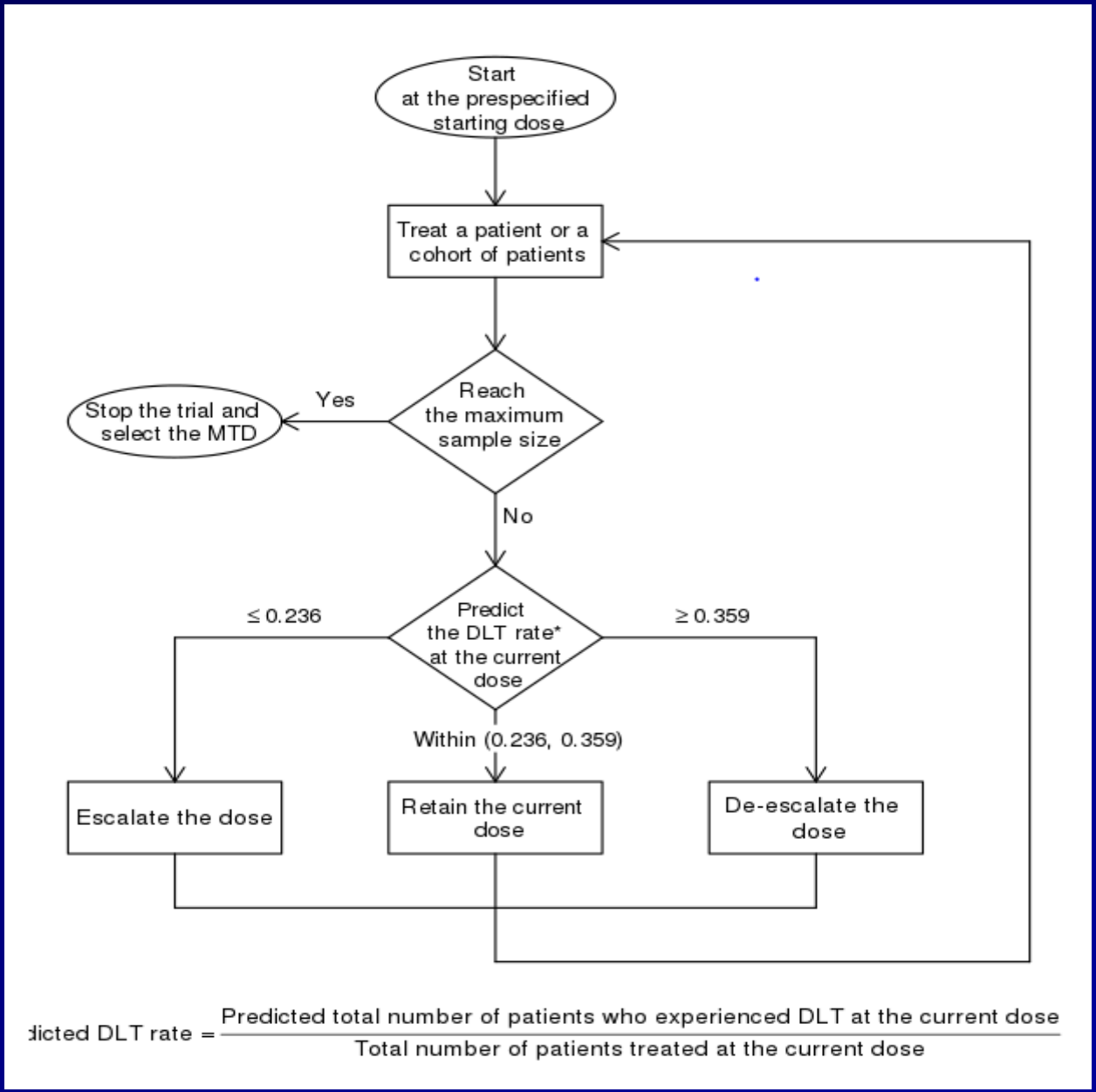
Table 1: Dose escalation/de-escalation rule.

Copy | CSV | Excel | Print

	# Patients	# DLTs observed	# Pending patients	Ecalation	Stay	De-escalation
1	3	0	<=1	Yes	No	No
2	3	0	>=2	Suspend	Suspend	Suspend
3	3	1	0	No	Yes	No
4	3	1	1	No	STFT>0.88	STFT<=0.88
5	3	1	>=2	Suspend	Suspend	Suspend
6	3	2	<=1	No	No	Yes
7	3	3	0	No	No	Yes & Eliminate
8	6	0	<=3	Yes	No	No
9	6	0	>=4	Suspend	Suspend	Suspend
10	6	1	<=1	Yes	No	No
11	6	1	2	STFT>=0.6	STFT<0.6	No
12	6	1	3	STFT>=1.96	STFT<1.96	No
13	6	1	>=4	Suspend	Suspend	Suspend
14	6	2	0	No	Yes	No
15	6	2	1	No	STFT>0.73	STFT<=0.73
16	6	2	2	No	STFT>1.8	STFT<=1.8
17	6	2	3	No	STFT>2.87	STFT<=2.87
18	6	2	>=4	Suspend	Suspend	Suspend
19	6	3	<=3	No	No	Yes

"STFT" is the standardized total follow-up time for the patients with data pending, defined as the total follow-up time for the patients with data pending divided by the length of the DLT assessment window.

BOIN for Time-To-Event Endpoint



Comparison of design characteristics among R6, TITE-CRM, and TITE-BOIN

Design characteristics	R6	TITE-CRM	TITE-BOIN
Can it target any prespecified DLT rate?	No	Yes	Yes
Allows to use a cohort size other than 3?	No	Yes	Yes
Uses follow-up time data from pending patients to make efficient decision of dose escalation and deescalation?	No	Yes	Yes
Can sample size be calibrated to ensure good operating characteristics?	No	Yes	Yes
Can the number of patients treated at the MTD be more than 6?	No	Yes	Yes
Can dose-escalation/deescalation rule be pretabulated for simple implementation?	Yes	No	Yes
Requires complicated, repeated estimation of the dose–toxicity curve model?	No	Yes	No

Software for the BOIN Designs

- The R package "BOIN" is available at CRAN.
- Standalone GUI based software is also available from MD Anderson Biostatistics software download website.

https://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware/Index/99 (BOIN Suite)

- Statistical tutorial and protocol templates are provided at

http://ibl.mdanderson.org/BOIN/ : Single agent

http://ibl.mdanderson.org/BOINComb/ : Combinations

https://ibl.mdanderson.org/TITE-BOIN/: Time to event

Comparisons of 3 Types of Designs

Design characteristics	Algorithm-based	Model-assisted	Model-based
Transparency & Simplicity			
Pre-determined dose escalation/de-escalation rule	Yes	Yes	No
Avoids computation-intensive, repeated estimation of the dose-toxicity curves for interim decisions	Yes	Yes	No
Flexibility			
Targets any prespecified DLT rate	No	Yes	Yes
Allows decision making when the cohort size deviates from the planned size	No	Yes	Yes
Number of patients treated at the MTD can be more than 6	No	Yes	Yes
Sample size can be calibrated to ensure good operating characteristics	No	Yes	Yes
Performance			
Identifies the MTD accurately	No	Yes	Yes
Allocates a high percentage of patients to the MTD	No	Yes	Yes
Provides good overdose control	Yes	Yes	Yes

Yuan Y, Lee JJ and Hilsenbeck SG. Model-Assisted Designs for Early Phase Clinical Trials: Simplicity Meets Superiority. JCO PO (In Press)

3. Phase II Designs

- Simon's 2-stage
- Predictive probability Phase II design
- BOP2

Phase IIA Design for A Single Treatment

- An efficacy screening trial
- Binary response endpoint with a response rate p.
- For testing H_0: $p \leq p_0$ vs. H_1: $p \geq p_1$
- Find the sample size to control
  - Type I (α) error
  - Type II (β) error
- Frequentist Designs
  - One-stage
  - Two-stage
    - Gehan's design
    - Simon's optimal and minimax designs
- Bayesian Design
  - Predictive probability design for continuous monitoring

BOP2: A Bayesian Optimal Design for Phase 2 Clinical Trials with simple & complex endpoints

- provides a unified framework for phase II trials with simple and complex efficacy and toxicity endpoints.
- explicitly controls the type I (and II) error rates.
- Is optimal by (i) maximizing power, given a fixed N and type I error; or (ii) minimizing the E(N|H0), given fixed type I and II error rates.
- Easy to use software is freely available to generate stopping boundaries, operating characteristics and protocol for the BOP2 design.

Zhou H, Lee JJ, Yuan Y. BOP2: Bayesian optimal design for phase II clinical trials with simple and complex endpoints. Stat Med. e-Pub 6/2017. PMID: 28589563.

BOP2: Statistical Model

- Multiple endpoints: $Y \sim Multinomial\ (\theta_1, \theta_2, \ldots, \theta_K)$
- $(\theta_1, \theta_2, \ldots, \theta_K) \sim\ Dir(a_1, a_2, \ldots., a_K)$
- Given data: $\theta|D_n \sim Dir(a_1{+}x_1, a_2{+}x_2, \ldots., a_K{+}x_K)$
- Decision rule: Stop the trial if
  - $Prob(b\theta \leq \phi|D_n) > C(n)$
  - $C(n) = 1\ - \lambda(n/N)^{\gamma}$
- Steps
  1. Elicit parameters under H_0, H_1 and desirable type I error
  2. Find the set of (λ, γ) yields type I error by grid search
  3. Among the set above, select the one optimize power
- An alternative is to find (λ, γ, N) to minimize $E(N|H_0)$

BOP2 Design, Examples

- Example 1: A treatment is
  - futile if ORR ≤ 0.2; promising if ORR ≥ 0.4.
- Example 2: A treatment is efficacious if
  - CR ≥ 0.15 or CR+PR ≥ 0.30.
- Example 3: A treatment
  - Fails if ORR ≤ 10% and PFS6 ≤ 20%.
  - Succeeds if ORR ≥ 30% or PFS6 ≥ 35%.
- Example 4: A treatment is safe and efficacious if
  - ORR ≥ 45% and toxicity rate ≤ 30%.

Stopping Boundaries for BOP2 Design

Trial	Stop the trial if		Number of patients treated						
			10	15	20	25	30	35	40
Example 1		# of OR $\leq$	1	2	4	5	7	9	10
Example 2	and	# of CR $\leq$	0	1	3	4	5	7	9
		# of CR/PR $\leq$	2	3	5	8	10	13	16
Example 3	and	# of OR $\leq$	0	1	2	3	4	5	7
		# of PFS6 $\leq$	1	2	4	5	7	9	12
Example 4	or	# of OR $\leq$	2	5	7	10	13	16	19
		# of Toxicities $\geq$	5	6	8	9	10	11	12

OR: objective response

1. $H_0 : Pr(OR) = 0.20; H_1 : Pr(OR) = 0.4$
2. $H_0 : Pr(CR) = 0.15, Pr(CR/PR) = 0.3; H_1 : Pr(CR) = 0.25, Pr(CR/PR) = 0.50.$
3. $H_0 : Pr(OR) = 0.1, Pr(PFS6m) = 0.2; H_1 : Pr(OR) = 0.3, Pr(PFS6m) = 0.35.$
4. $H_0 : Pr(OR) = 0.45, Pr(Toxicity) = 0.30; H_1 : Pr(OR) = 0.60, Pr(Toxicity) = 0.20.$

BOP2: Bayesian Optimal Phase II Design with Simple and Complex Endpoints

PID: 960; Version: V1.1.0.0 ; Last Updated: 03/22/2019

Heng Zhou, Ying-Wei Kuo, Ying Yuan and J. Jack Lee

Department of Biostatistics,MD Anderson Cancer Center

Trial Setting | Simulation | Protocol | Reference

Example 1: One efficacy endpoint

Endpoints:

- (•) Binary Efficacy
- ( ) Binary Toxicity
- ( ) Efficacy & Toxicity
- ( ) Co-Primary Efficacy
- ( ) Ordinal Efficacy

Interims: Help

Sample sizes when interim analyses to be performed, seperated by space. The last number must be the total sample size.

10 15 20 25 30 35 40

Null Hypothesis: Help

Response Rate

0.2

Alternative Hypothesis: Help

Response Rate

0.4

Type I error rate:

0.1

Simulation Parameters:

Number of Simulations

10000

Seed of the random number generator

1024

Stopping boundaries

Optimal stopping boundaries that maximize power

CSV | Excel | PDF | Print

Search:

# patients treated	Stop if # responses <=
10	1
15	2
20	4
25	5
30	7
35	9
40	10

Showing 1 to 7 of 7 entries Previous 1 Next

The power of this trial is: 0.8829

Futility stop if response is ≤ 1/10, 2/15, 4/20, 5/25, 7/30, 9/35 or 10/40 pts

Trial Setting | Simulation

Same as Simon's Optimal Two-Stage Design

Endpoints:

- (•) Binary Efficacy
- ( ) Binary Toxicity
- ( ) Efficacy & Toxicity
- ( ) Co-Primary Efficacy
- ( ) Ordinal Efficacy

Interims: Help

Sample sizes when interim analyses to be performed, seperated by space. The last number must be the total sample size.

12 25

Null Hypothesis: Help

Response Rate

0.2

Alternative Hypothesis: Help

Response Rate

0.4

Type I error rate:

0.1

Simulation Parameters:

Number of Simulations

10000

Seed of the random number generator

1024

Stopping boundaries

Optimal stopping boundaries that maximize power

CSV | Excel | PDF | Print

Search:

# patients treated	Stop if # responses <=
12	2
25	7

Showing 1 to 2 of 2 entries

Previous 1 Next

The power of this trial is: 0.8149

Futility stop if response is ≤ 2/12 or 7/25 pts

Variation of Simon's Optimal Two-Stage Design

Endpoints:

- (•) Binary Efficacy
- ( ) Binary Toxicity
- ( ) Efficacy & Toxicity
- ( ) Co-Primary Efficacy
- ( ) Ordinal Efficacy

Interims: Help

Sample sizes when interim analyses to be performed, seperated by space. The last number must be the total sample size.

12 18 25

Null Hypothesis: Help

Response Rate

0.2

Alternative Hypothesis: Help

Response Rate

0.4

Type I error rate:

0.1

Simulation Parameters:

Number of Simulations

10000

Seed of the random number generator

1024

Optimal stopping boundaries that maximize power

CSV Excel PDF Print Search:

# patients treated	Stop if # responses <=
12	2
18	4
25	7

Showing 1 to 3 of 3 entries Previous 1 Next

The power of this trial is: 0.8

Futility stop if response is ≤ 2/12, 4/18, or 7/25 pts

OC for Simon's 2-Stage and BOP2 Designs

Simulation scenarios:

Add a Scenario | Remove a Scenario

Scenario 1

Response Rate

0.2

Simon's 2-Stage

Scenario 2

Response Rate

0.4

Number of Simulated Trials

20000

BOP2 with 3-Stages

Simulate

Operating characteristics

Operating characteristics

CSV | Excel | PDF | Print

Search:

Response rate	Early stopping (%)	Claim promising (%)	Sample size
0.2	55.77	9.86	17.7
0.4	8.51	81.06	23.9

Showing 1 to 2 of 2 entries

Previous 1 Next

Response rate	Early stopping (%)	Claim promising (%)	Sample size
0.2	75.17	9.35	16.4
0.4	12.96	80.025	23.6

Example 4: One efficacy endpoint + one toxicity endpoint

Trial Setting | Simulation | Protocol | Reference

Endpoints:

- ○ Binary Efficacy
- ○ Binary Toxicity
- ◉ Efficacy & Toxicity
- ○ Co-Primary Efficacy
- ○ Ordinal Efficacy

Interims: Help

Sample sizes when interim analyses to be performed, seperated by space. The last number must be the total sample size.

10 15 20 25 30 35 40

Null Hypothesis: Help

Pr(Eff)	Pr(Tox)	Pr(Eff & Tox)
0.45	0.3	0.15

Alternative Hypothesis: Help

Pr(Eff)	Pr(Tox)	Pr(Eff & Tox)
0.6	0.2	0.18

Type I error rate:

0.1

Simulation Parameters:

Number of Simulations

10000

Seed of the random number generator

1024

Calculate stopping boundaries

Stopping boundaries

Optimal stopping boundaries that maximize power

CSV | Excel | PDF | Print

Search:

# patients treated	Stop if # response <=	OR # toxicity >=
10	2	5
15	5	6
20	7	8
25	10	9
30	13	10
35	16	11
40	19	12

Showing 1 to 7 of 7 entries

Previous 1 Next

The power of this trial is: 0.7464

Association Between Two Endpoints

- A treatment is safe and efficacious if ORR ≥ 45% and toxicity rate ≤ 30%
  - H_0: $p_{ORR} = 0.45, p_{TOX} = 0.30$

	No Tox	Tox	Total
No Resp			
Resp		?	0.45
Total		0.30	

	No Tox	Tox	Total
No Resp			
Resp		0.15	0.45
Total		0.30	

- Under the alternative hypothesis
  - H_1: $p_{ORR} = 0.60, p_{TOX} = 0.20$

	No Tox	Tox	Total
No Resp			
Resp		?	0.60
Total		0.20	

	No Tox	Tox	Total
No Resp			
Resp		0.18	0.60
Total		0.20	

Efficacy and Toxicity Monitoring

We simultaneously monitor efficacy and safety endpoints using the Bayesian optimal phase 2 (BOP2) design (Zhou, Lee and Yuan, 2017). Specifically, let n denote the interim sample size and N denote the maximum sample size. Let Y_{eff} and Y_{tox} respectively denote the efficacy and toxic endpoints, with $Y_{eff} = 1$ and $Y_{tox} = 1$ respectively indicating that patients experience efficacy and toxicity. We assume that the joint distribution of (Y_{eff}, Y_{tox}) follows a multinomial distribution with 4 elementary outcomes: $(Y_1, Y_2) = (1, 1)$, $(0, 1)$, $(1, 0)$ and $(0, 0)$. Let $p_{eff} = Pr(Y_1 = 1)$, $p_{tox} = Pr(Y_2 = 1)$ and define the null hypothesis H_0: $p_{eff} \leq 0.45$ and $p_{tox} > 0.3$, representing that the treatment is inefficacious or overly toxic. We will stop enrolling patients and claim that the treatment is not promising if

$$Pr(p_{eff} > 0.45|data) < \lambda(\frac{n}{N})^{\alpha},$$

or

$$Pr(p_{tox} \leq 0.3|data) < \lambda(\frac{n}{N})^{\alpha},$$

where λ=0.625 and α=1 are design parameters optimized to minimize the chance of incorrectly claiming that an efficacious and safe treatment is unacceptable (i.e., type II error) under the alternative hypothesis H_1: $p_{eff} = 0.6$ and $p_{tox} = 0.2$, while controlling the type I error rate at 0.1 (i.e., the chance of incorrectly claiming that an inefficacious or overly treatment is acceptable is no more than 10%). Assuming a Dirichlet prior distribution $Dir(0.15,0.3,0.15,0.4)$ for the treatment effect, the above decision rule corresponds to the following stopping boundaries and yields a statistical power of 0.7464 under H_1:

Table 1: Optimized stopping boundaries

# patients treated	Stop if # response <=	OR # toxicity >=
10	2	5
15	5	6
20	7	8
25	10	9
30	13	10
35	16	11
40	19	12

Based on Table 1, we perform the interim analysis when the number of enrolled patients reaches 10, 15, 20, 25, 30, 35. When the total number of patients reaches the maximum sample size of 40, we reject the null hypothesis and conclude that the treatment is acceptable if the number of responses in the efficacy endpoint are greater than 19, and the number of toxicities are less than 12; otherwise we conclude that the treatment is unacceptable.

Below are the operating characteristics of the design based on 10000 simulations using the BOP2 web application, which is available at http://www.trialdesign.org.

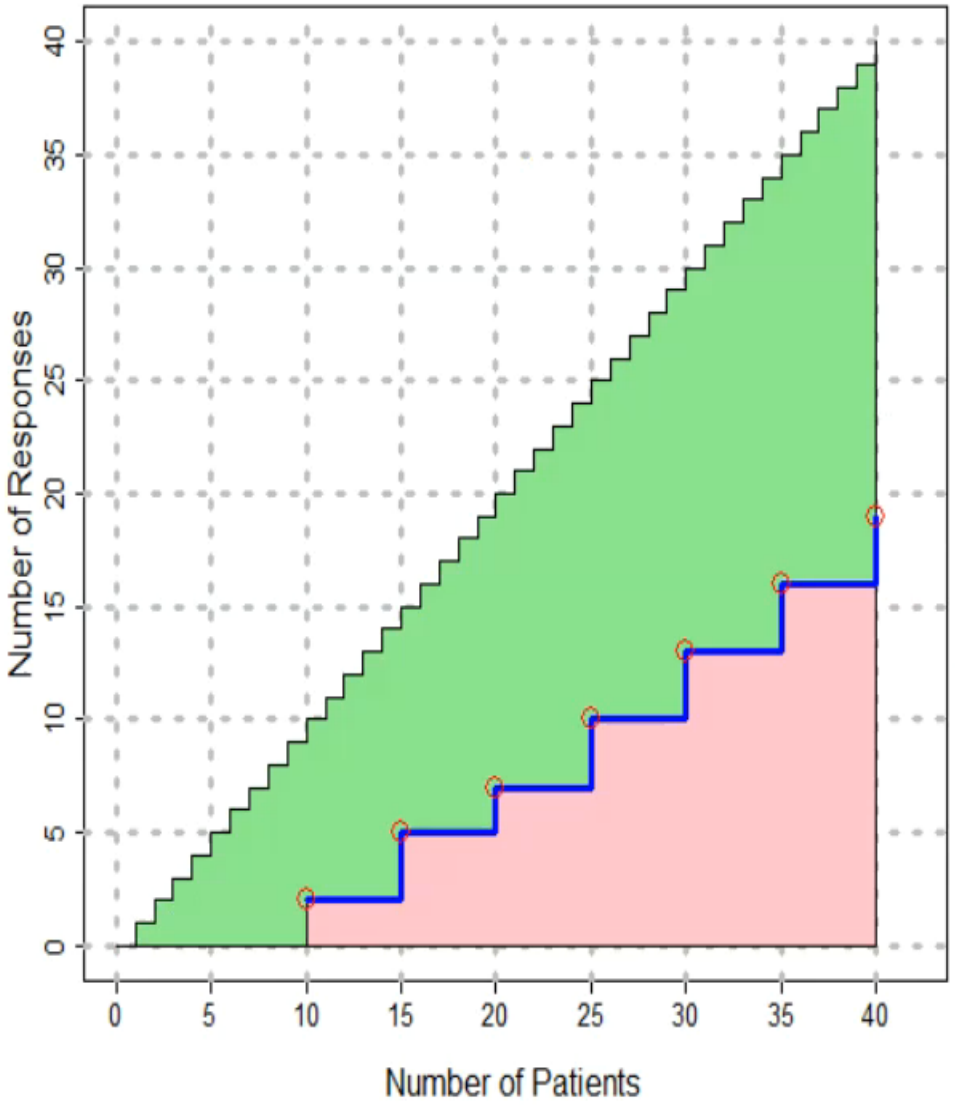
Table 2: Operating characteristics

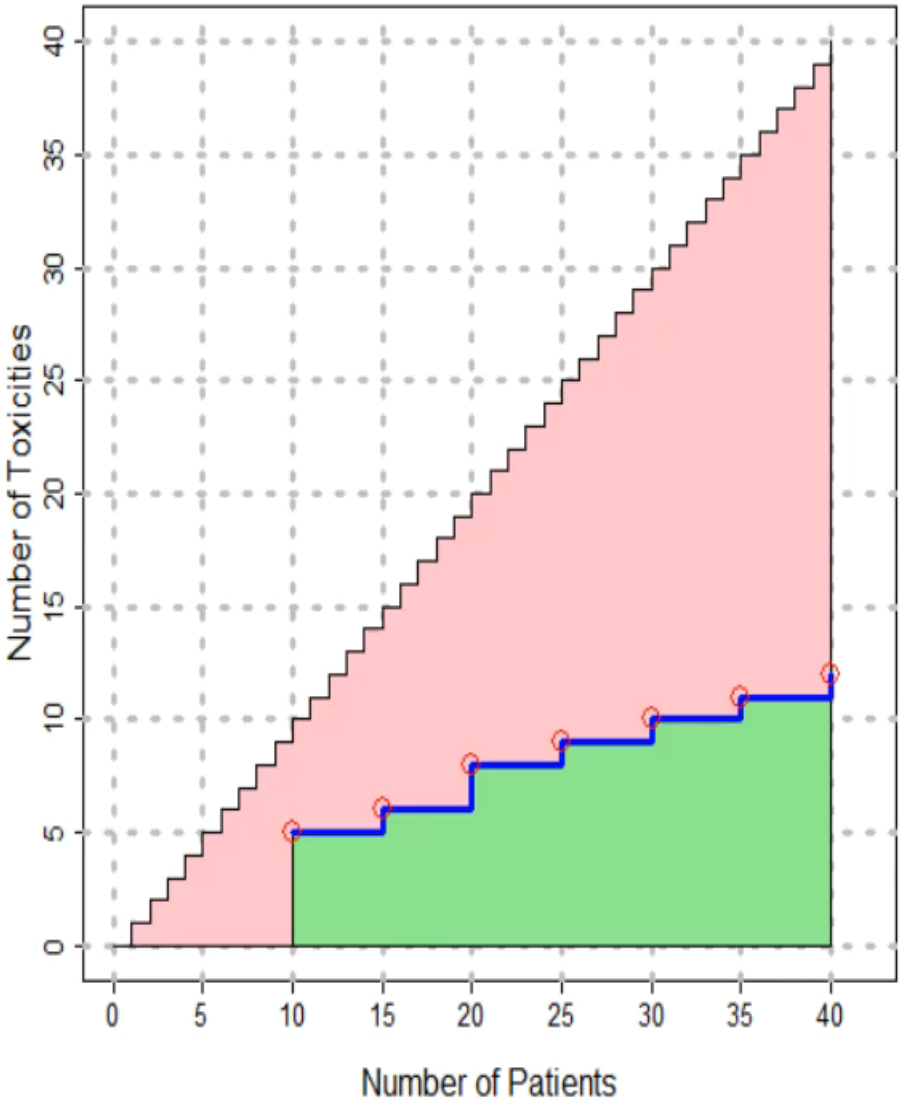
Pr(Eff)	Pr(Tox)	Pr(Eff & Tox)	Early stopping (%)	Claim acceptable (%)	Sample size
0.30	0.25	0.15	99.68	0.05	14.5
0.45	0.30	0.20	88.49	6.91	21.1
0.60	0.20	0.10	20.83	75.73	36.0
0.60	0.15	0.05	11.76	86.31	37.8

Reference

Zhou, H., Lee, J. J., & Yuan, Y. (2017). BOP2: Bayesian optimal design for phase II clinical trials with simple and complex endpoints. *Statistics in Medicine*, 36(21):3302-3314.

Video 2: Illustrating BOP2 Design





Summary

- Clinical trial is an adaptive learning process.
  - Bayesian framework provides an ideal platform for learning. Bayesian adaptive designs are flexible and efficient for adaptive learning.
  - "We learn as we go."
- BOIN, BOIN-COMB, TITE-BOIN combine the benefit of rule-based and model-based designs for Phase I studies.
- BOP2 design is useful in Phase II studies with complex endpoints.
- Model-assisted designs offer excellent statistical properties and are easy to conduct.
  - Simplicity meets superiority.
  - Keep It Simple and Smart!