



Should the choice of the BOIN design parameter $p.tox$ depend solely on the target DLT rate?

Rong Lu

ronglu@stanford.edu

<https://orcid.org/0000-0003-4321-9144>

05/20/2024

<https://doi.org/10.1101/2024.03.06.24303862>

BOIN Suite Designs



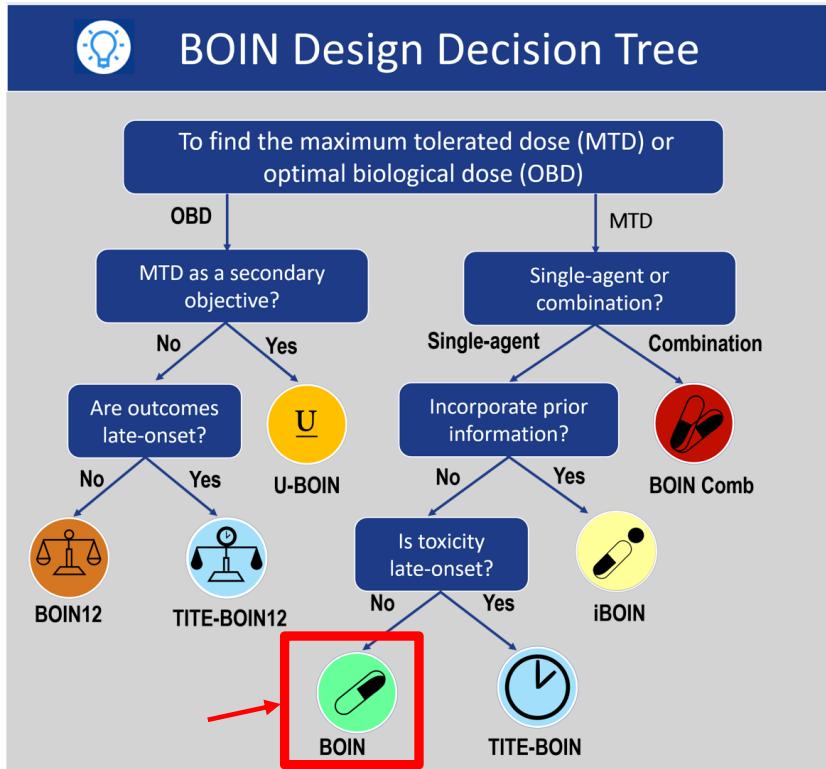
BOIN Suite [How to choose a design?](#)

Bayesian optimal interval (BOIN) designs provide a novel platform to design phase I trials with single agent, drug combination, platform [more...](#)

Bayesian optimal interval (BOIN) designs provide a novel platform to design phase I trials with single agent, drug combination, and late-onset toxicity under a unified framework. As a model-assisted design, the BOIN combines the simplicity of the algorithm-based design and the superiority of the model-based design. The BOIN can be implemented in a simple way similar to the 3+3 design, but yields superior performance comparable to more complex model-based designs, such as the continual reassessment method (CRM).

- <https://trialdesign.org>

How to choose a design?



BOIN Design

Appl. Statist. (2015)
64, Part 3, pp. 507–523

Bayesian optimal interval designs for phase I clinical trials

Suyu Liu and Ying Yuan

University of Texas MD Anderson Cancer Center, Houston, USA

[Received April 2014. Revised August 2014]

Summary. In phase I trials, effectively treating patients and minimizing the chance of exposing them to subtherapeutic and overly toxic doses are clinicians' top priority. Motived by this practical consideration, we propose Bayesian optimal interval (BOIN) designs to find the maximum tolerated dose and to minimize the probability of inappropriate dose assignments for patients. We show, both theoretically and numerically, that the BOIN design not only has superior finite and large sample properties but also can be easily implemented in a simple way similar to the traditional '3+3' design. Compared with the well-known continual reassessment method, the BOIN design yields comparable average performance to select the maximum tolerated dose but has a substantially lower risk of assigning patients to subtherapeutic and overly toxic doses. We apply the BOIN design to two cancer clinical trials.

Keywords: Bayesian adaptive design; Decision error; Dose finding; Maximum tolerated dose

Erratum published on
11Jan2022:

<https://rss.onlinelibrary.wiley.com/doi/full/10.1111/rssc.12541>

<https://trialdesign.org>

HOME SOFTWARE OUR TEAM PUBLICATIONS CONTACT

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

PID: 979 ; V3.0.7.0 ; Last Updated: 05/09/2024

Yanhong Zhou, Suyu Liu, Ying Yuan, J. Jack Lee, Heng Zhou, Nan Chen, and Ying-Wei Kuo

Department of Biostatistics, MD Anderson Cancer Center

Trial Setting Simulation Trial Protocol Animation Select MTD Reference

User guide for the BOIN App

Trial Name (Optional):

BOIN example

Design parameters:

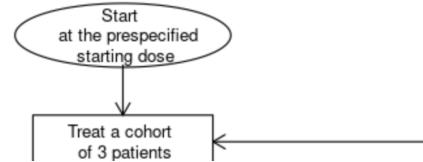
- Type in Upload using previously saved input

Enter design parameters in the following panels.

Design Flow Chart

Decision Table

 Download flowchart



R package

Package ‘BOIN’

October 12, 2022

Type Package

Title Bayesian Optimal INterval (BOIN) Design for Single-Agent and Drug- Combination Phase I Clinical Trials

Version 2.7.2

Date 2021-01-19

Author Ying Yuan and Suyu Liu

Maintainer Ying Yuan <yyuan@mdanderson.org>

Imports Iso

Description The Bayesian optimal interval (BOIN) design is a novel phase I clinical trial design for finding the maximum tolerated dose (MTD). It can be used to design both single-agent and drug-combination trials. The BOIN design is motivated by the top priority and concern of clinicians when testing a new drug, which is to effectively treat patients and minimize the chance of exposing them to subtherapeutic or overly toxic doses. The prominent advantage of the BOIN design is that it achieves simplicity and superior performance at the same time. The BOIN design is algorithm-based and can be implemented in a simple way similar to the traditional 3+3 design. The BOIN design yields an average performance that is comparable to that of the continual reassessment method (CRM, one of the best model-based designs) in terms of selecting the MTD, but has a substantially lower risk of assigning patients to subtherapeutic or overly toxic doses. For tutorial, please check Yan et al. (2020) <[doi:10.18637/jss.v094.i13](https://doi.org/10.18637/jss.v094.i13)>.

Fit-For-Purpose Tools and Supporting Information:

Disease Area	Submitter	Tool	Trial Component	Issuance Date and Supporting Information
Alzheimer's disease	The Coalition Against Major Diseases (CAMD)	Disease Model: Placebo/Disease Progression	Demographics, Drop-out	Issued June 12, 2013 <ul style="list-style-type: none"> Determination Letter
Multiple	Janssen Pharmaceuticals and Novartis Pharmaceuticals	Statistical Method: MCP-Mod	Dose-Finding	Issued May 26, 2016 <ul style="list-style-type: none"> Determination Letter Statistical Review Pharmacometric Review
Multiple	Ying Yuan, PhD The University of Texas MD Anderson Cancer Center Department of Biostatistics	Statistical Method: Bayesian Optimal Interval (BOIN) design	Dose-Finding	Issued: December 10, 2021 <ul style="list-style-type: none"> Determination Letter Statistical Review Publication Erratum ↗
Multiple	Pfizer	Statistical Method: Empirically Based Bayesian Emax Models	Dose-Finding	Issued: August 5, 2022 <ul style="list-style-type: none"> Determination Letter Multidisciplinary Review

<https://www.fda.gov/drugs/development-approval-process-drugs/drug-development-tools-fit-purpose-initiative>



IN THIS SECTION: Development & Approval Process | Drugs

[← Development & Approval Process | Drugs](#)

Drug Development Tools: Fit-for-Purpose Initiative



Background

The Fit-for-Purpose (FFP) Initiative provides a pathway for regulatory acceptance of dynamic tools for use in drug development programs. Due to the evolving nature of these types of drug development tools (DDTs) and the inability to provide formal qualification, a designation of ‘fit-for-purpose’ (FFP) has been established. A DDT is deemed FFP based on the acceptance of the proposed tool following a thorough evaluation of the information provided. The FFP determination is made publicly available in an effort to facilitate greater utilization of these tools in drug development programs.

Contact Us

For more information about the FFP Initiative, please contact DrugDevelopmentTools@fda.hhs.gov

R package BOIN Design Parameters I

This package allows the user to pre-specify **target dose limiting toxicity (DLT) rate (ϕ)** as well as the following 8 design parameters:

1. ***ncohort***: *The total number of cohorts.*
2. ***cohortsize***: *The cohort size.*
3. ***n.earlystop***: *The early stopping parameter. If the number of patients treated at the current dose reaches *n.earlystop*, stop the trial early and select the MTD based on the observed data. The default value of *n.earlystop* = 100 essentially turns off this type of early stopping.*

<https://trialdesign.org> BOIN Parameters

Doses & Sample Size ?

Number of doses:	Starting dose level:
5	1
Cohort size:	Number of cohort:
3	4

Stop trial if the number of patients assigned to single dose reaches m and the decision is to stay, where $m =$

The figure shows a horizontal scale with numerical markings from 6 to 12. A blue horizontal bar is positioned above the scale, starting at 6 and ending at 12. At the exact position of 12, there is a pink circular marker with the number '12' inside it.

- $ncohort = 4$
- $cohortsize = 3$
- $n.earlystop = 12$

R package BOIN Design Parameters II

4. **$p.saf(\phi_1)$:** The highest toxicity probability that is deemed subtherapeutic (i.e., below the MTD) such that dose escalation should be made. The default value of $p.saf = 0.6 * \text{target DLT rate}$.
5. **$p.tox(\phi_2)$:** The lowest toxicity probability that is deemed overly toxic such that dose de-escalation is required. The default value of $p.tox = 1.4 * \text{target DLT rate}$.

<https://tralldesign.org> BOIN Parameters

Target Probability ?

Target Toxicity Probability ϕ :

Use the default escalation and de-escalation boundaries (recommended).

The default boundaries are $\lambda_1 = 0.197$ and $\lambda_2 = 0.298$ for target toxicity probability 0.25

Escalation boundary λ_1 :

De-escalation boundary λ_2 :

The dose escalation/de-escalation boundaries correspond to the use of the following alternative hypotheses for minimizing the decision error of dose assignment: underdosing hypothesis $\phi_1 = 0.15$ and overdosing hypothesis $\phi_2 = 0.35$.

$$\lambda_d = \frac{\log\left(\frac{1-\phi}{1-\phi_2}\right)}{\log\left\{\frac{\phi_2(1-\phi)}{\phi(1-\phi_2)}\right\}}$$

$$\lambda_e = \frac{\log\left(\frac{1-\phi_1}{1-\phi}\right)}{\log\left\{\frac{\phi(1-\phi_1)}{\phi_1(1-\phi)}\right\}}$$

- **p.saf = $\phi_1 = 0.15$**
- **p.tox = $\phi_2 = 0.35$**

R package BOIN Design Parameters III

6. ***cutoff.elิ***: The cutoff to eliminate the overly toxic dose for safety. We recommend the default value $cutoff.elิ = 0.95$ for general use.
7. ***extrasafe***: Set $extrasafe = \text{TRUE}$ to impose a stricter stopping rule.
8. ***offset***: A small positive number (between 0 and 0.5) to control how strict the stopping rule is when $extrasafe = \text{TRUE}$. A larger value leads to a stricter stopping rule. The default value $offset = 0.05$ generally works well.

<https://trialdesign.org> BOIN Parameters

Overdose Control



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

Use the default cutoff (recommended) $p_E =$

0.95

- Check to impose a more stringent safety stopping rule on the lowest dose.

Stop the trial if $Pr(p_1 > \phi | data) > p_E - \delta$, where δ is

0.05

- Check to ensure $\hat{p}_{MTD} \leq$ de-escalation boundary, where \hat{p}_{MTD} is the isotonic estimate of the DLT probability for the dose selected as the MTD.

- cutoff.eli = 0.95***
- extrasafe = TRUE***
- offset = 0.05***

<https://trialdesign.org> BOIN Parameters

Modify the decision from de-escalation to stay when observing 1 DLT out of 3 patients:

- Yes No



Perform accelerated titration:

- Yes No



Cap the titration up to dose level:

5



2

3

4

5



Stanford
MEDICINE

3 + 3 Design

Design Flow Chart

Decision Table

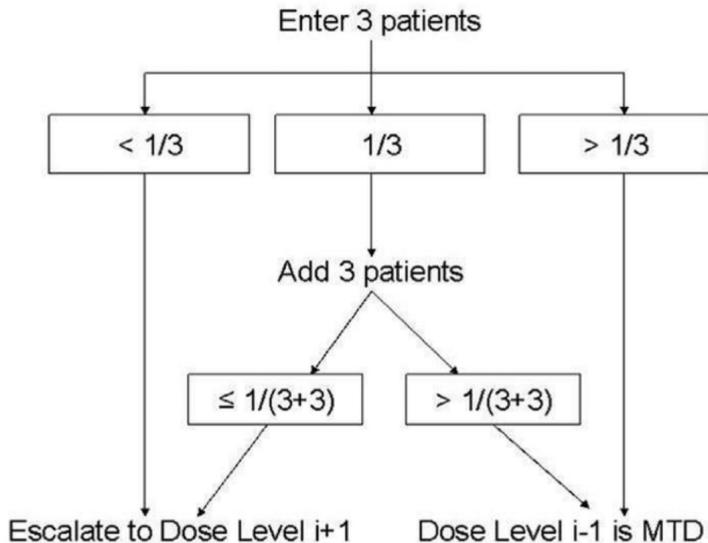


Table 1: Dose escalation/de-escalation rule.

Copy CSV Excel Print

Number of evaluable patients treated

Escalate if # of DLT \leq

Stay if # of DLT =

De-escalate if # of DLT \geq

Eliminate if # of DLT \geq

	3	4	5	6
0	0	0	0	1
1	1	1	1	NA
2	2	2	2	2
2	2	2	2	2

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

<https://doi.org/10.1016/j.conctc.2018.10.006>

3 + 3 Design

[Design Flow Chart](#)

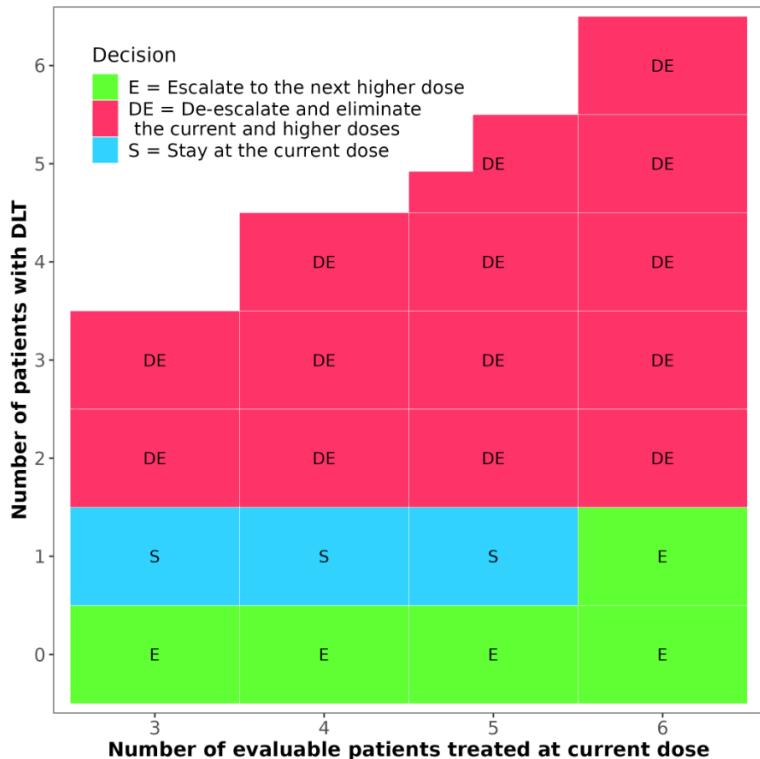
[Decision Table](#)

Table 1: Dose escalation/de-escalation rule.

[Copy](#) [CSV](#) [Excel](#) [Print](#)

Number of evaluable patients treated	3	4	5	6
Escalate if # of DLT <=	0	0	0	1
Stay if # of DLT =	1	1	1	NA
De-escalate if # of DLT >=	2	2	2	2
Eliminate if # of DLT >=	2	2	2	2

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



3 + 3 Design

Appendix 1: Trial and Design Specifications

Parameter	Value
Number of doses	5
Starting dose	1
Max sample size	6
Cohort size	3
Stop trial if # patients assigned to single dose reaches	6
Use accelerated titration	FALSE
Target toxicity probability	0.176
Use the default alternatives to minimize decision errors	FALSE
Alternative (unacceptable high toxicity) for optimization	0.9
Alternative (unacceptable low toxicity) for optimization	0.16
Escalation boundary	0.167
De-escalation boundary	0.559
Eliminate dose threshold	0.855
Impose a more stringent safety stopping rule	FALSE
Require the isotonic estimate of the DLT probability for the dose selected as the MTD less than the de-escalation boundary	FALSE
Number of repetitions per scenario	1000
Random number generator seed	6

```

> get.boundary(target=0.1761482, ncohort=10, cohortsize=3, n.earlystop = 6,
+               p.saf = 0.1582749, p.tox = 0.892814, cutoff.eliminate = 0.8548338,
+               extrasafe = F)
$lambda_e
[1] 0.1670842

$lambda_d
[1] 0.556843

$boundary_tab
Number of patients treated 3 6
Escalate if # of DLT <= 0 1
Deescalate if # of DLT >= 2 2
Eliminate if # of DLT >= 2 2

$full_boundary_tab
Number of patients treated 1 2 3 4 5 6
Escalate if # of DLT <= 0 0 0 0 0 1
Deescalate if # of DLT >= 1 2 2 2 2 2
Eliminate if # of DLT >= NA NA 2 2 2 2

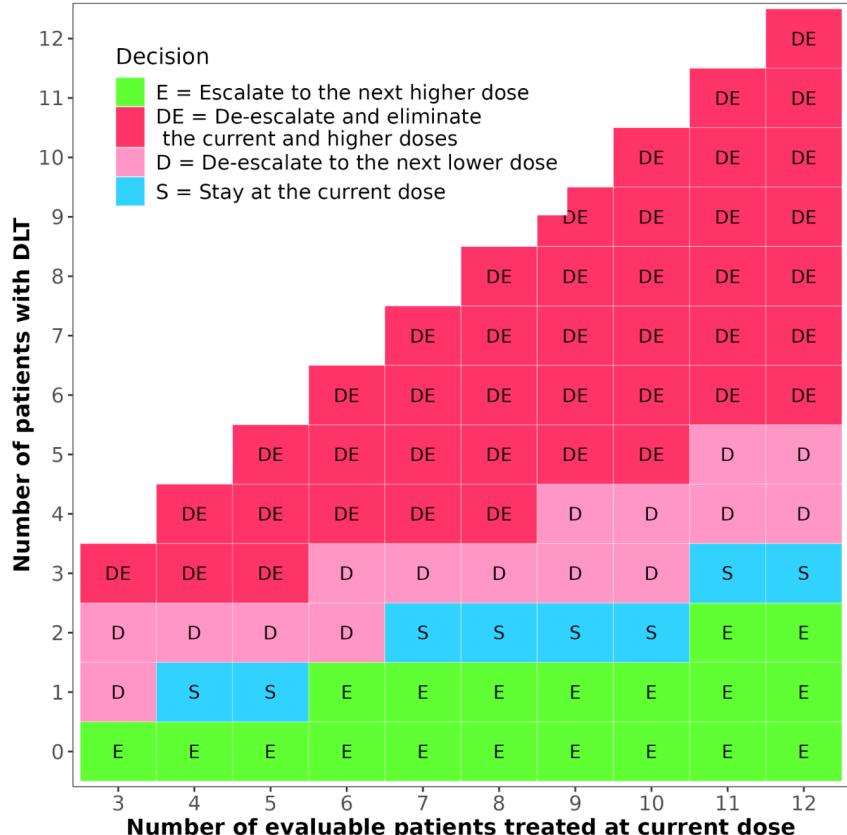
attr(),"class")
[1] "boin"
Warning message:
In get.boundary(target = 0.1761482, ncohort = 10, cohortsize = 3, :
  the value of n.earlystop is too low to ensure good operating characteristic
s. Recommend n.earlystop = 9 to 18.

```

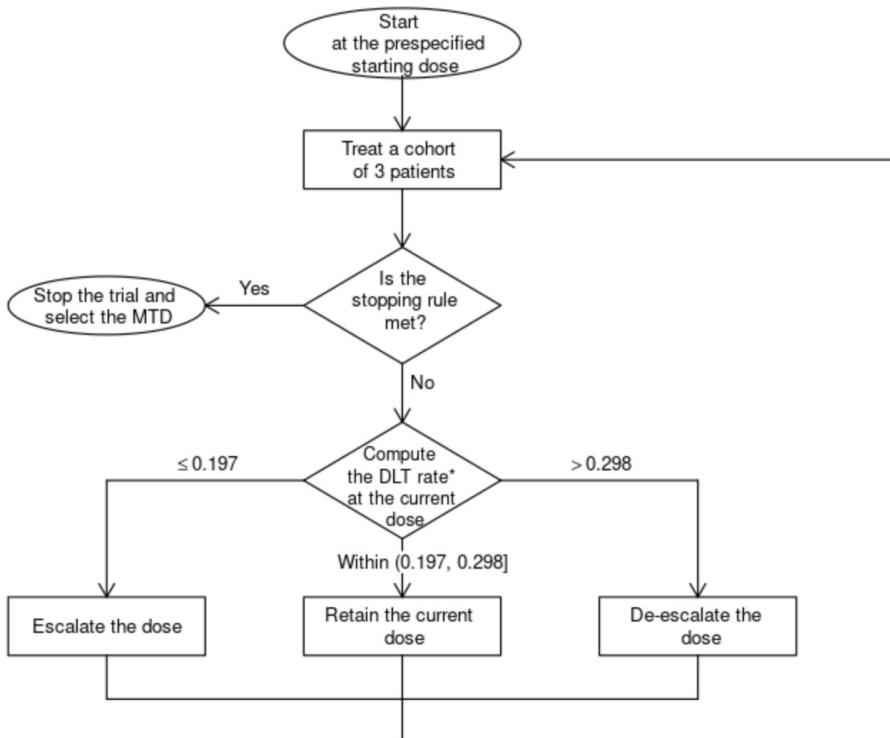
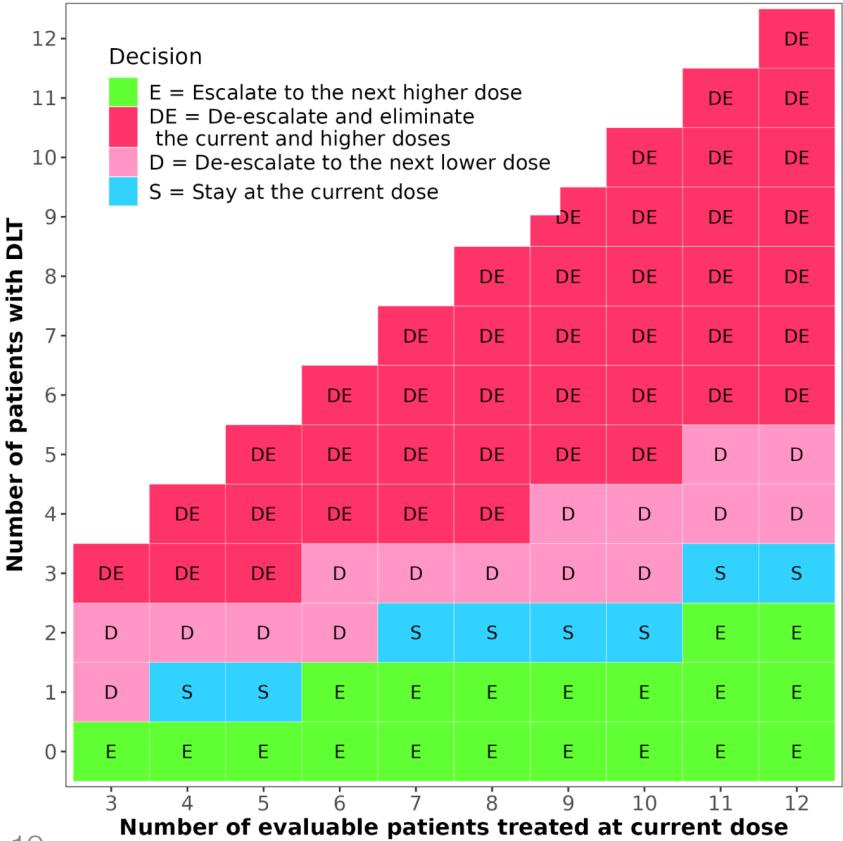
BOIN Example

Appendix 1: Trial and Design Specifications

Parameter	Value
Number of doses	5
Starting dose	1
Max sample size	12
Cohort size	3
Stop trial if # patients assigned to single dose reaches	12
Use accelerated titration	FALSE
Target toxicity probability	0.25
Use the default alternatives to minimize decision errors	FALSE
Alternative (unacceptable high toxicity) for optimization	0.35
Alternative (unacceptable low toxicity) for optimization	0.15
Escalation boundary	0.197
De-escalation boundary	0.298
Eliminate dose threshold	0.95
Impose a more stringent safety stopping rule	FALSE
Require the isotonic estimate of the DLT probability for the dose selected as the MTD less than the de-escalation boundary	FALSE
Number of repetitions per scenario	1000
Random number generator seed	6



BOIN Example



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

BOIN Example

Table 1. Dose escalation/de-escalation rule for the BOIN design

Decision	The number of evaluable patients treated at current dose									
	3	4	5	6	7	8	9	10	11	12
Escalate if # of DLT <=	0	0	0	1	1	1	1	1	2	2
Stay if # of DLT =	NA	1	1	NA	2	2	2	2	3	3
De-escalate if # of DLT >=	1	2	2	2	3	3	3	3	4	4
Eliminate if # of DLT >=	3	3	3	4	4	4	5	5	6	6

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

Question

- As with other model-based dose-finding algorithms, the operating characteristics of a BOIN design are greatly affected by the choice of its design parameters.
- Although setting the BOIN design parameter $p.tox = 1.4 * target.DLT.rate$ is recommended in almost all BOIN methodology articles and is the default value in the R package *BOIN*, it's unclear why the choice of $p.tox$ should only depend on the target DLT rate and whether certain range of $p.tox$ could produce the same BOIN boundary table.

Simulation Setup

- In this simulation study, following parameters will be varied one at a time, using R package *BOIN*, to explore each parameter's effect on the equivalence intervals of *p.saf* and *p.tox*:
 - 1) target DLT rate,
 - 2) *n.earlystop*,
 - 3) *cutoff.eli*,
 - 4) *cohortszie*
 - 5) *ncohort*.
- The 3+3 design boundary table will be used as a simple example to explore all equivalent sets of BOIN design parameters that can generate the same boundary table.

Simulation Setup

- All dose escalation/de-escalation boundary tables were calculated using the *get.boundddary()* function from R package *BOIN*
- *tryCatch()* function was used to handle *get.boundddary()* errors such as “the probability deemed safe cannot be higher than or too close to the target!”
- All *p.saf* and *p.tox* values that produce the same BOIN boundary table (*i.e.* the same *\$boundary_tab* output) are considered equivalent.
- While the *\$boundary_tab* output remain the same, the *\$full_boundary_tab* outputs may be different.

Equivalent p.saf and p.tox under varying target DLT rates

Equivalent values of $p.\text{saf}$ and $p.\text{tox}$ were explored via uniform search under varying target DLT rates: $\text{target} = 10\%, 15\%, 20\%, 25\%, 30\%, 35\%$, or 40% and fixed values of following parameters :

- ncohort = 10
- cohortsize = 3
- n.earlystop = 12
- cutoff.eli = 95%
- extrasafe = FALSE

For each target DLT rate under evaluation, 100,000 pairs of $p.\text{saf}$ and $p.\text{tox}$ values were randomly drawn from the following uniform distributions :

- $\text{p.saf} <- \text{runif}(1, \text{min}=0, \text{max}=\text{target}-0.0000001)$
- $\text{p.tox} <- \text{runif}(1, \text{min}=\text{target}+0.0000001, \text{max}=1)$

10 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, n.earlystop = 12)

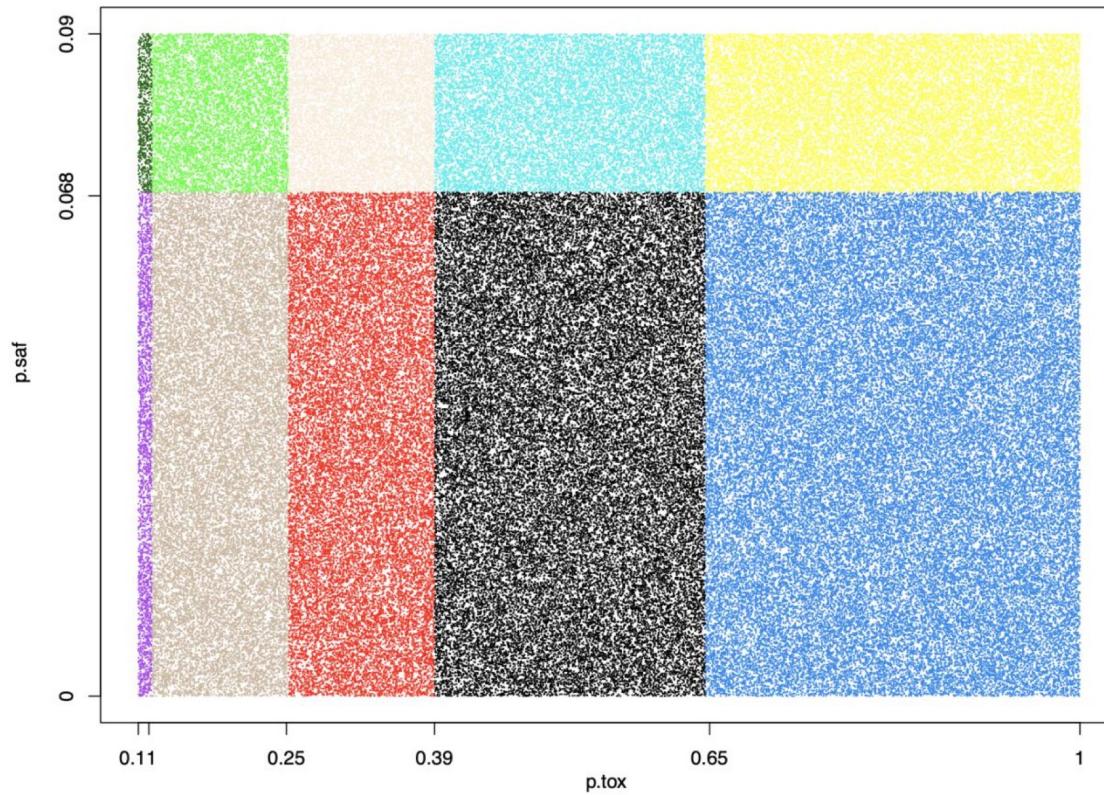


Figure 2a: equivalent intervals of $p.\text{saf}$ and $p.\text{tox}$ under target DLT rate = 10%,
cohortszie = 3, ncohorts = 10, n.earlystop = 12, cutoff.eli = 95%, and extrasafe = FALSE.

```

> get.boundary(target=0.1, ncohort=10, cohortsize=3, n.earlystop = 12,
+               p.saf = 0.00001, p.tox = 0.123, cutoff.elim = 0.95,
+               extrasafe = F)
$lambda_e
[1] 0.01130893

$lambda_d
[1] 0.1111531

$boundary_tab

Number of patients treated 3 6 9 12
Escalate if # of DLT <=    0 0 0 0
Deescalate if # of DLT >=  1 1 2 2
Eliminate if # of DLT >=  2 2 3 3

$full_boundary_tab

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

attr(,"class")
[1] "boin"

```

```

> get.boundary(target=0.1, ncohort=10, cohortsize=3, n.earlystop = 12,
+               p.saf = 0.067, p.tox = 0.14, cutoff.eli = 0.95,
+               extrasafe = F)
$lambda_e
[1] 0.08250041

$lambda_d
[1] 0.1190318

$boundary_tab
Number of patients treated 3 6 9 12
Escalate if # of DLT <=    0 0 0 0
Deescalate if # of DLT >=  1 1 2 2
Eliminate if # of DLT >=  2 2 3 3

$full_boundary_tab
Number of patients treated 1 2 3 4 5 6 7 8 9 10 11 12
Escalate if # of DLT <=   0 0 0 0 0 0 0 0 0 0 0 0
Deescalate if # of DLT >= 1 1 1 1 1 1 1 2 2 2 2 2
Eliminate if # of DLT >= NA NA 2 2 2 2 2 3 3 3 3 3

attr("class")
[1] "boin"

```

```

> get.boundary(target=0.1, ncohort=10, cohortsize=3, n.earlystop = 12,
+               p.saf = 0.067, p.tox = 0.25, cutoff.eli = 0.95,
+               extrasafe = F)
$lambda_e
[1] 0.08250041

$lambda_d
[1] 0.1659562

$boundary_tab

Number of patients treated 3 6 9 12
Escalate if # of DLT <= 0 0 0 0
Deescalate if # of DLT >= 1 1 2 2
Eliminate if # of DLT >= 2 2 3 3

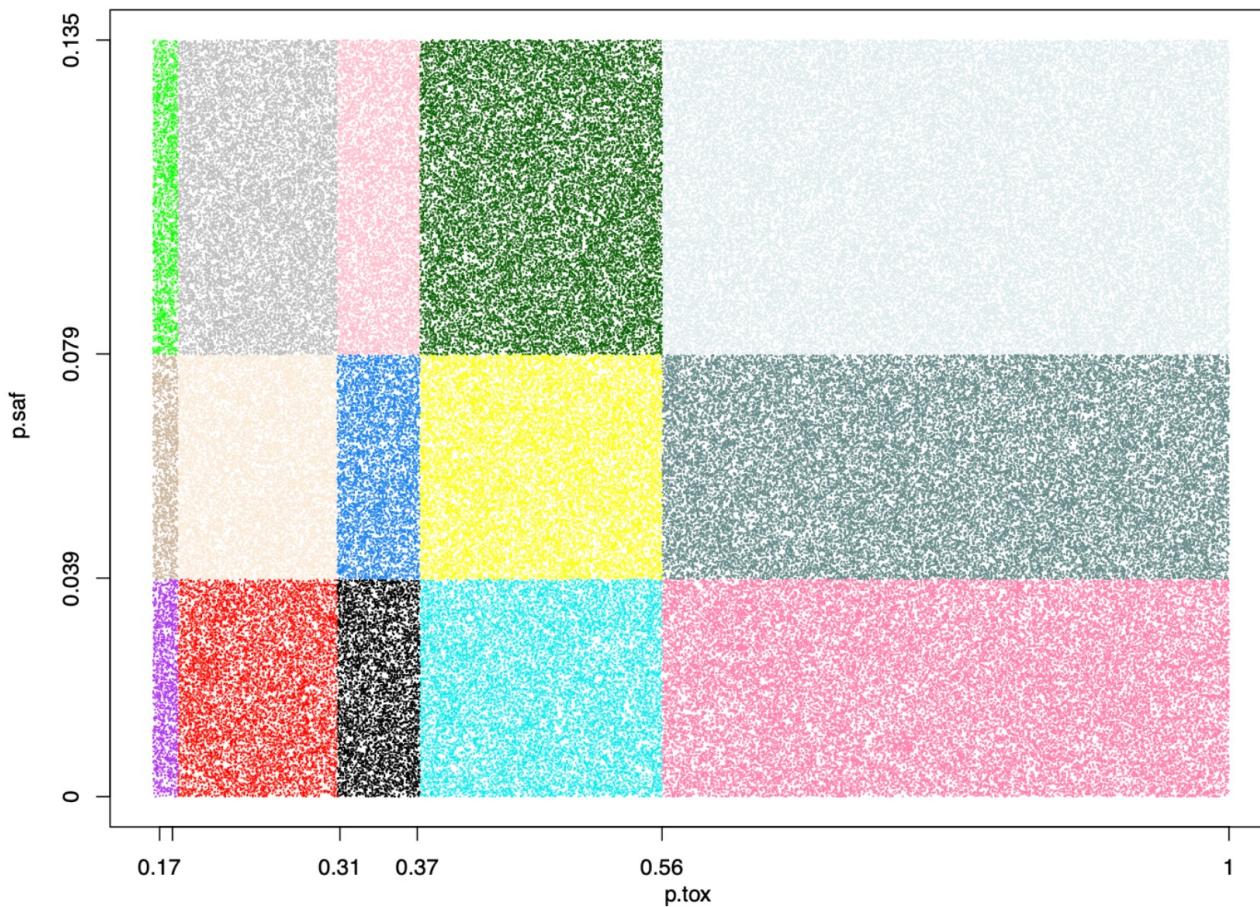
$full_boundary_tab

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

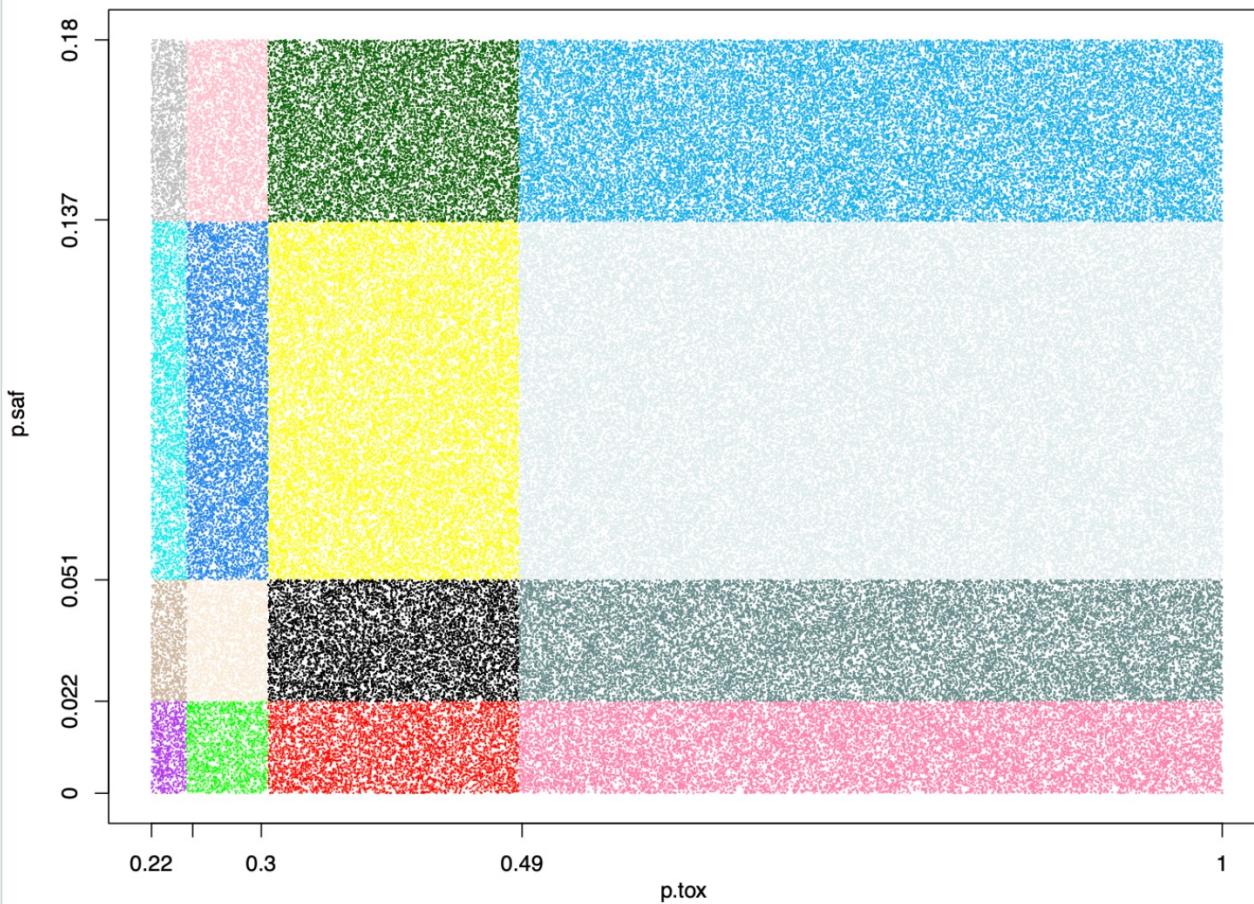
attr(),"class")
[1] "boin"

```

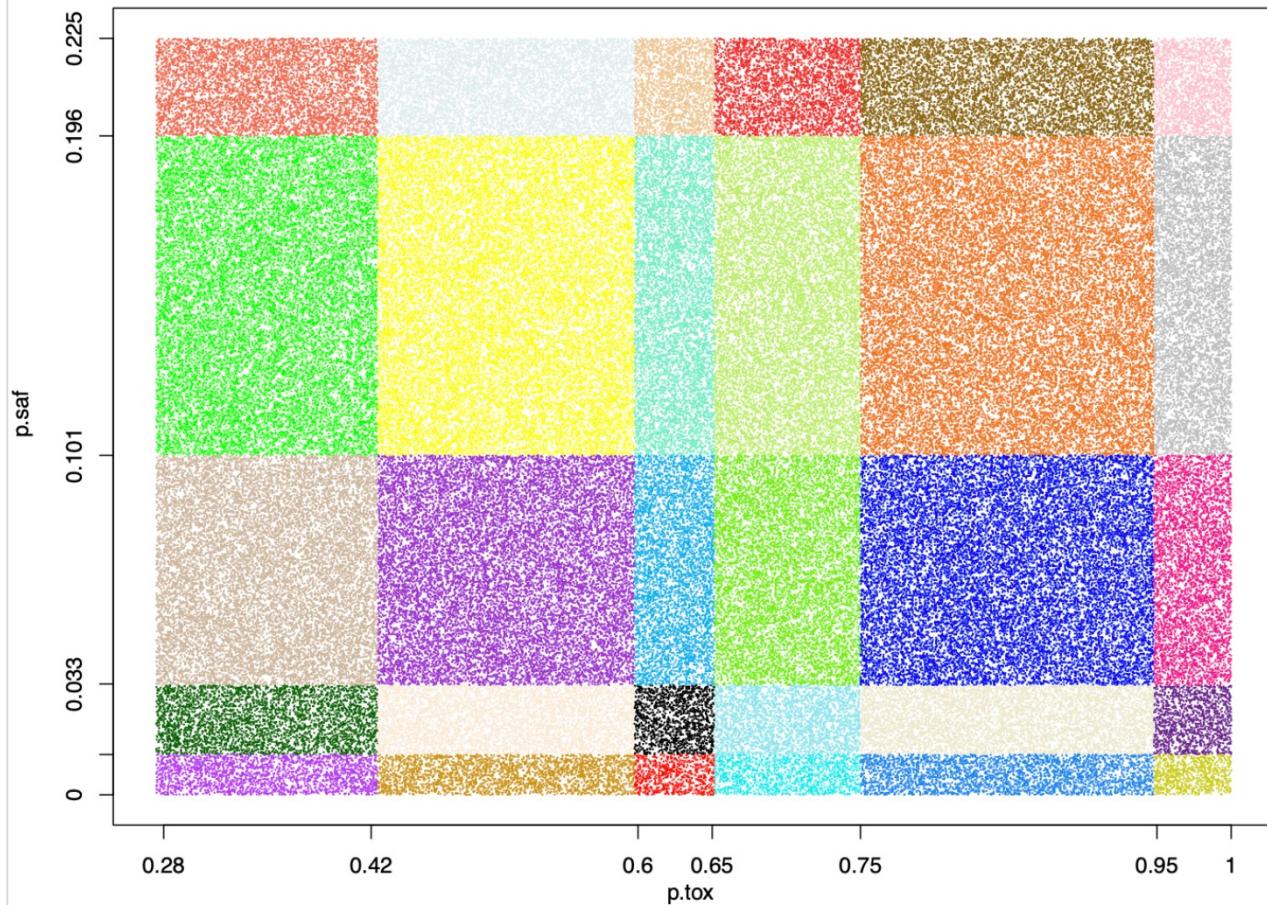
15 BOIN boundary tables indicated by different colors
(target DLT rate = 15%, n.earlystop = 12)



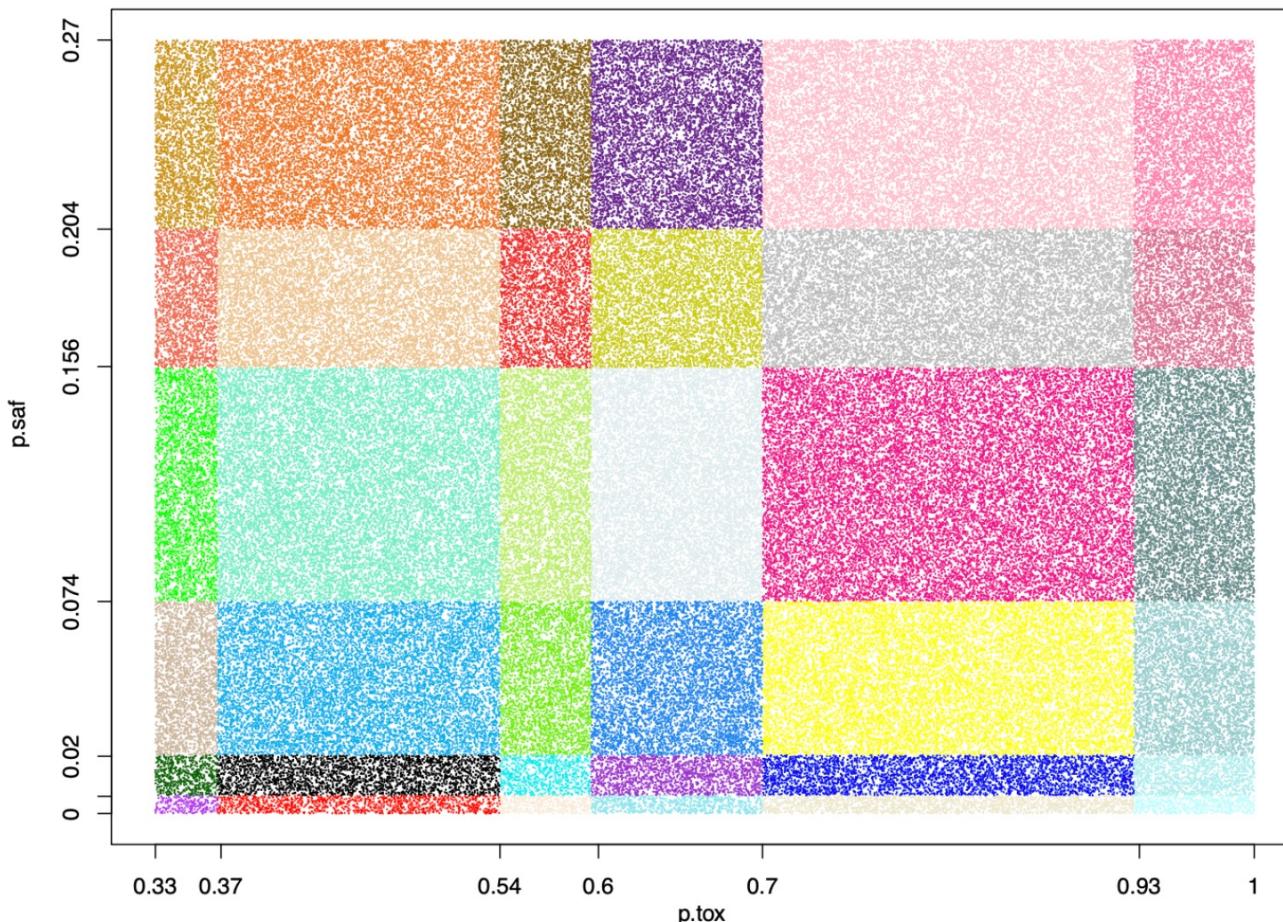
16 BOIN boundary tables indicated by different colors
(target DLT rate = 20%, n.earlystop = 12)



30 BOIN boundary tables indicated by different colors
(target DLT rate = 25%, n.earlystop = 12)



36 BOIN boundary tables indicated by different colors
(target DLT rate = 30%, n.earlystop = 12)



Equivalent p.saf and p.tox under varying n.earlystop

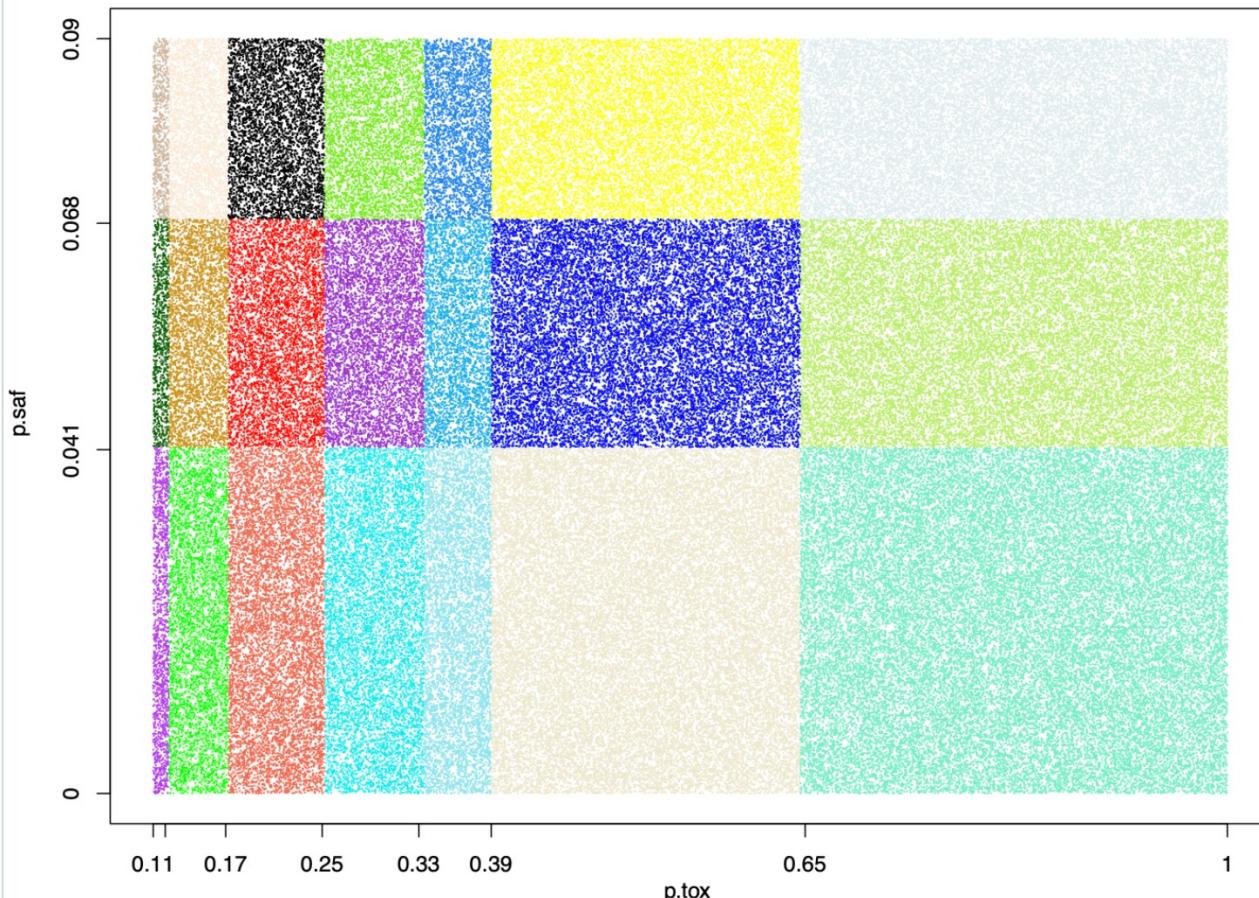
Equivalent values of p.saf and p.tox were explored via uniform search under varying n.earlystop = 15, 18, 21, 24, 27, or 30 and fixed values of following parameters:

- ncohort = 10
- cohortsize = 3
- target = 10% (target DLT rate)
- cutoff.elim = 95%
- extrasafe = FALSE

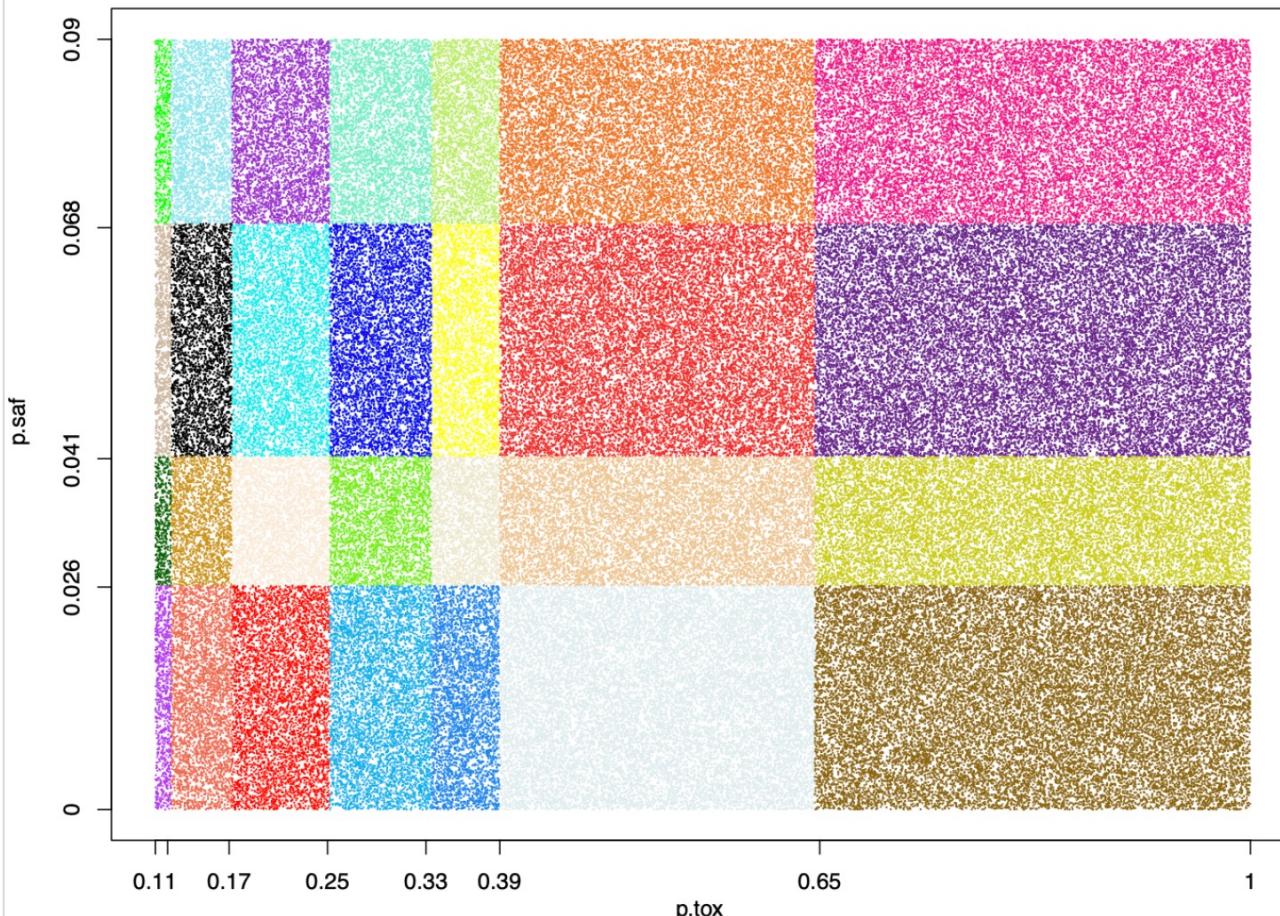
For each n.earlystop value, 100,000 pairs of p.saf and p.tox values were randomly drawn from the following uniform distributions:

- p.saf <- runif(1, min=0, max=target-0.0000001)
- p.tox <- runif(1, min=target+0.0000001, max=1)

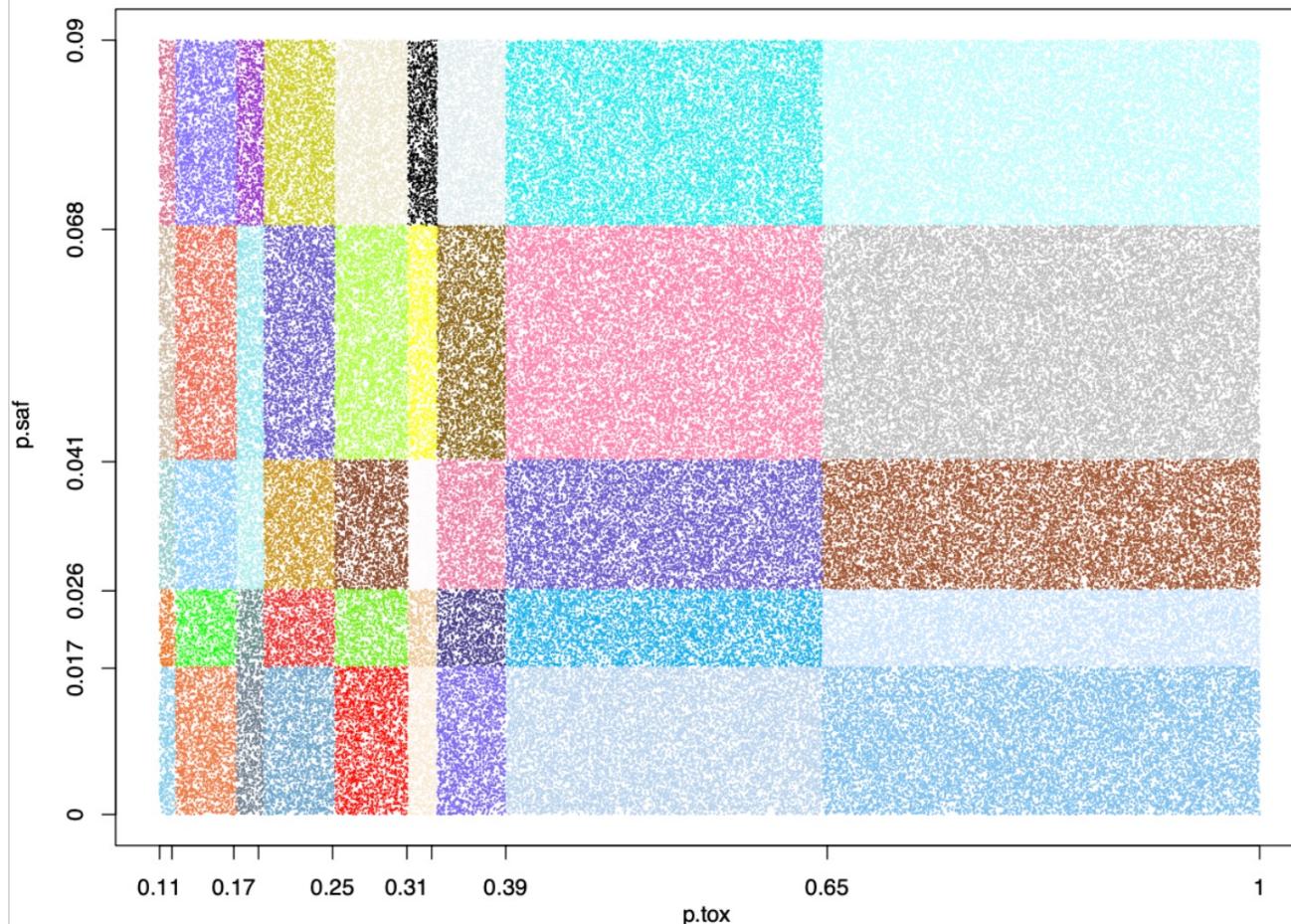
21 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, n.earlystop = 15)



28 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, n.earlystop = 18)



45 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, n.earlystop = 21)



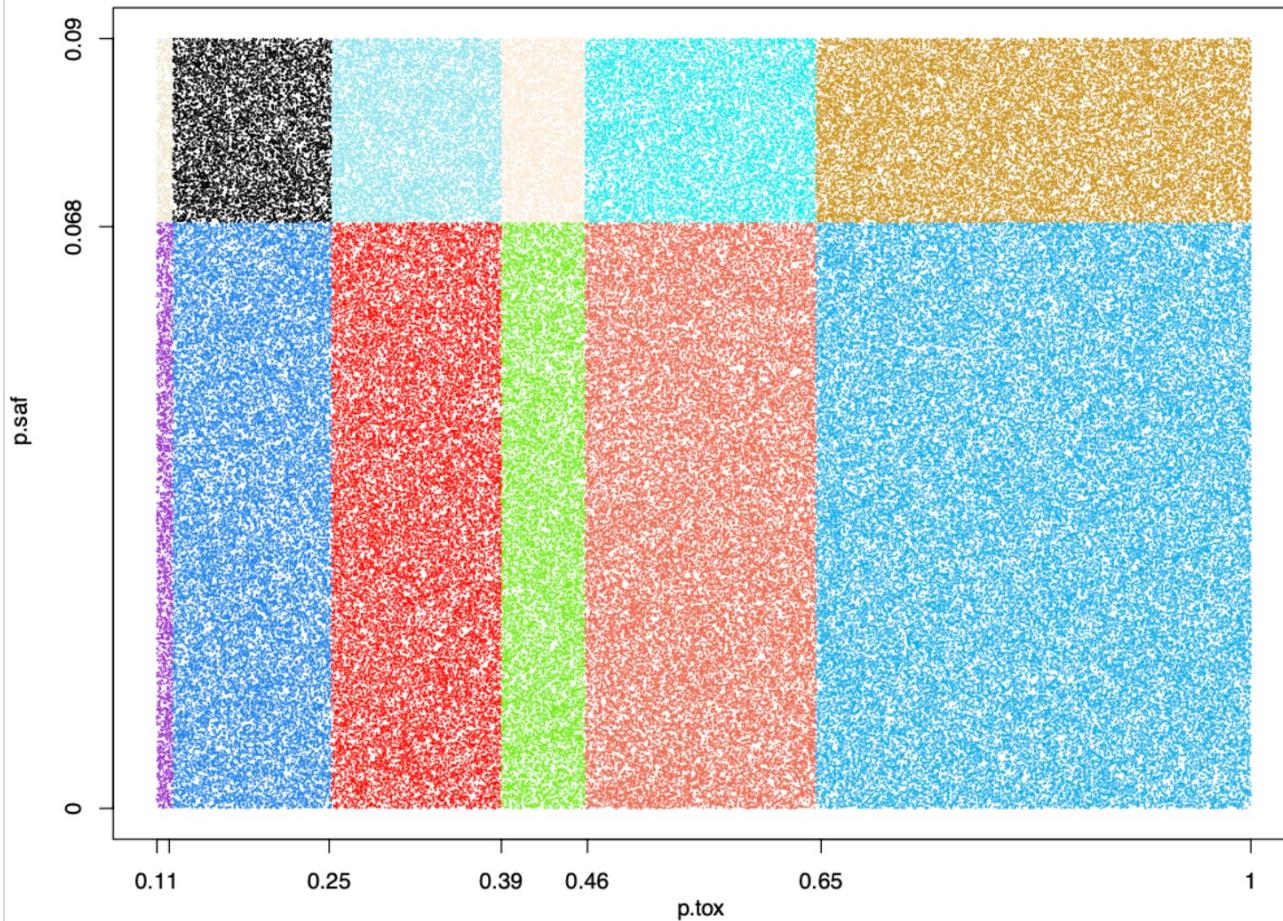
Equivalent p.saf and p.tox under varying cutoff.eli

Equivalent values of p.saf and p.tox were explored via uniform search under varying cutoff.eli = 70%, 80%, 90%, 97%, or 99% and fixed values of following design parameters:

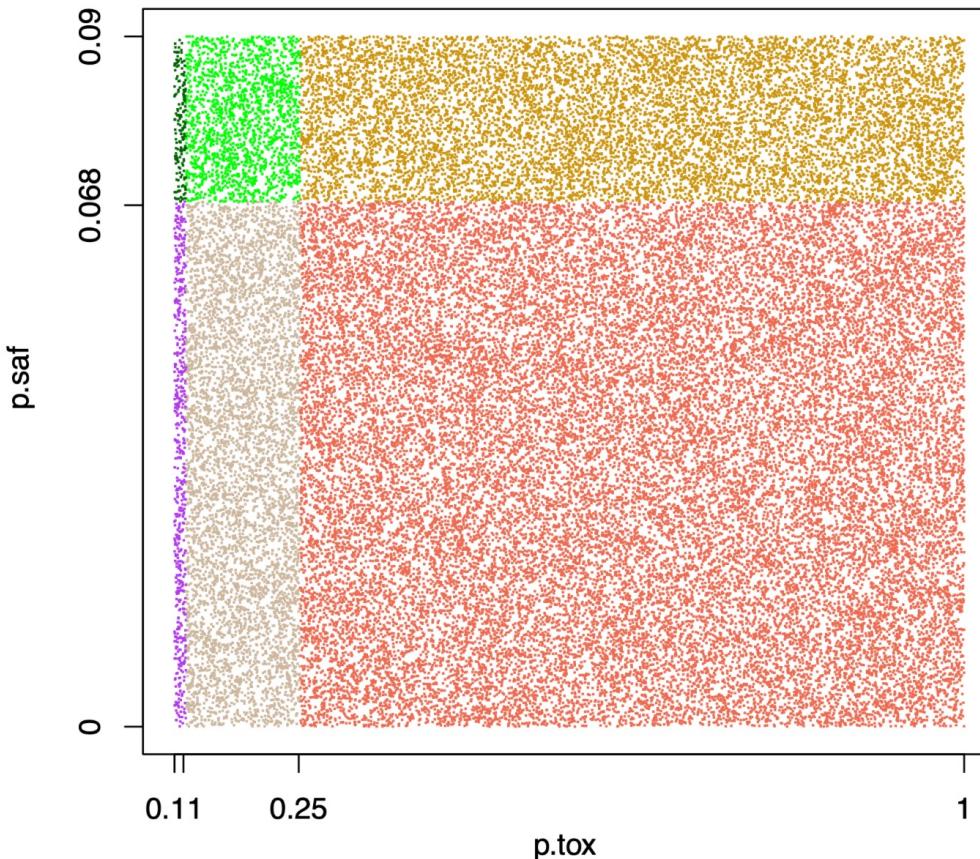
- ncohort = 10
- cohortsize = 3
- target = 10% (target DLT rate)
- n.earlystop = 12
- extrasafe = FALSE

For each cutoff.eli value, 100,000 pairs of p.saf and p.tox values were randomly drawn from the uniform distributions described in the previous sections.

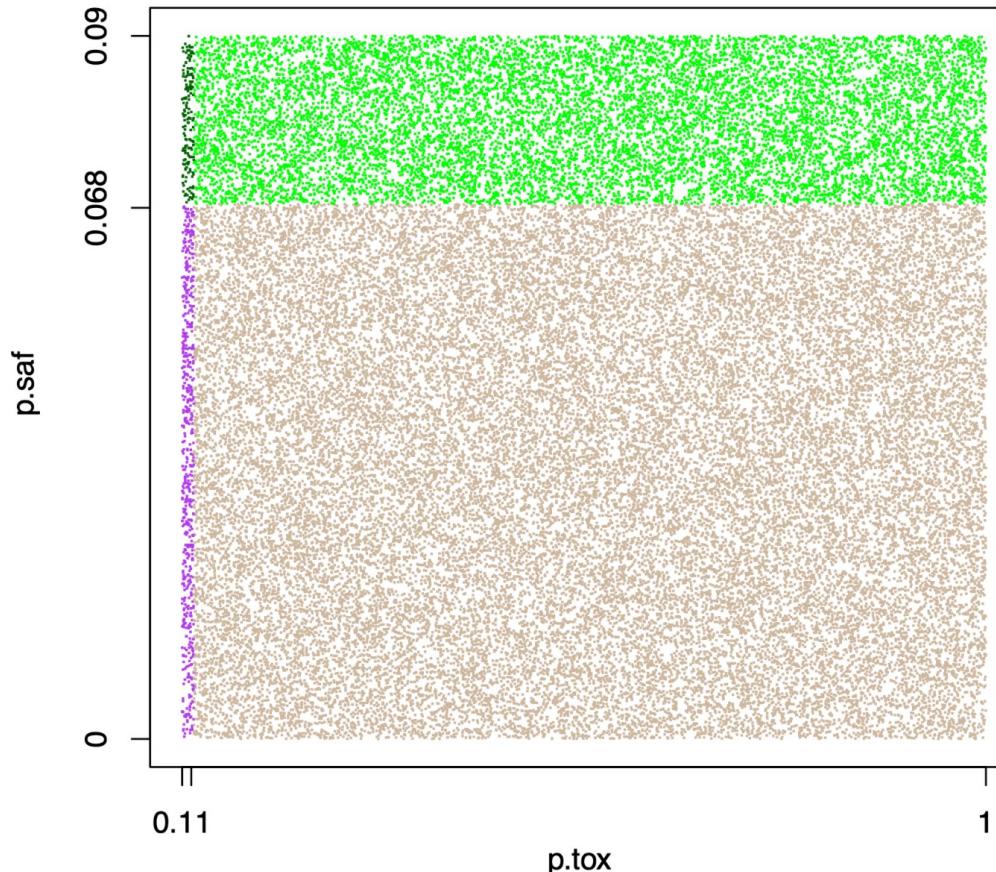
12 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cutoff.eli = 97%)



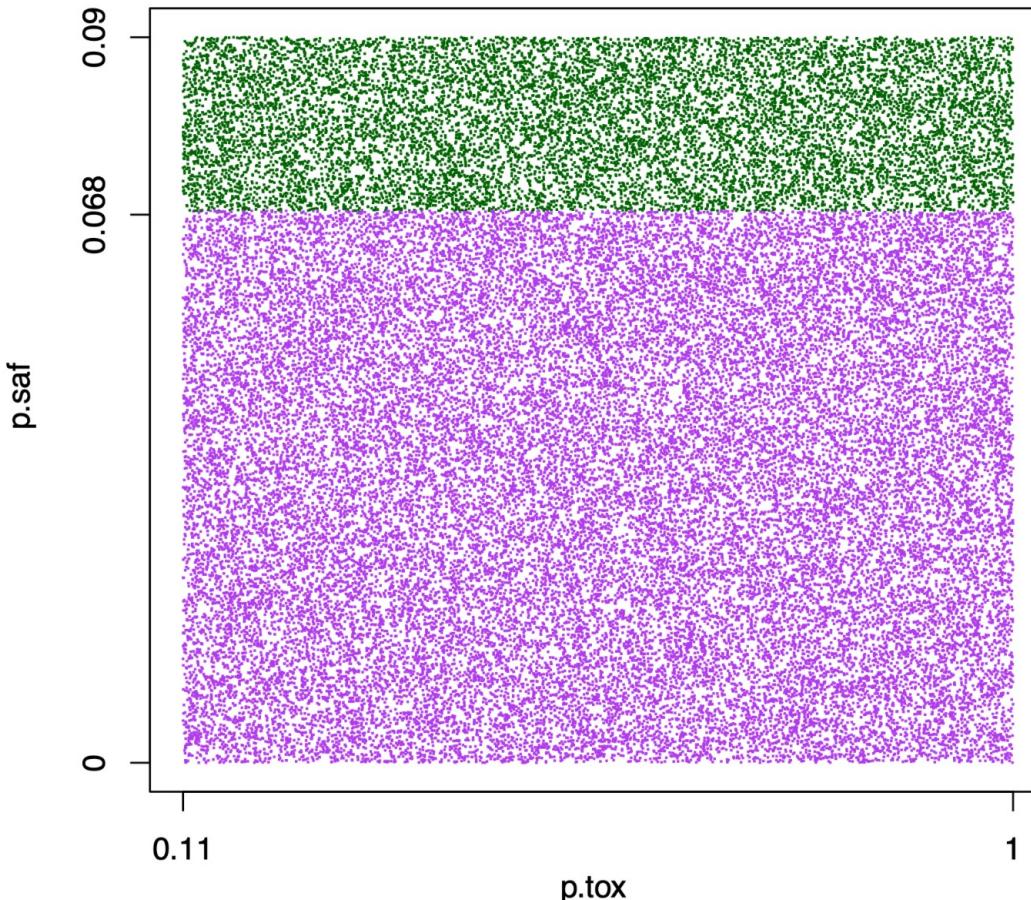
6 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cutoff.eli = 90%)



**4 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cutoff.eli = 80%)**



**2 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cutoff.eli = 70%)**



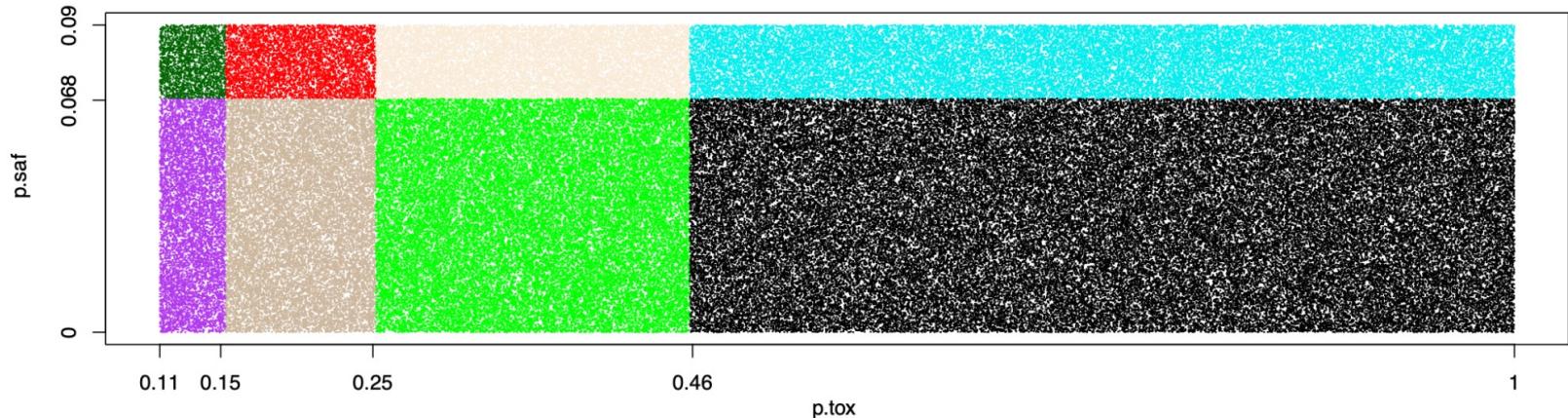
Equivalent p.saf and p.tox under varying cohortsize

Equivalent values of p.saf and p.tox were explored via uniform search under varying cohortsize = 4, 5, 6, 7, or 8 and fixed values of following design parameters:

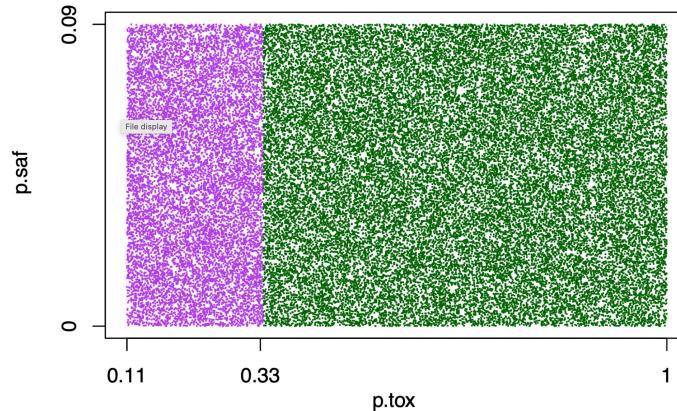
- ncohort = 10
- cutoff.elim = 95%
- target = 10% (target DLT rate)
- n.earlystop = 12
- extrasafe = FALSE

For each cohortsize, 100,000 pairs of p.saf and p.tox values were randomly drawn from the uniform distributions described in the previous sections.

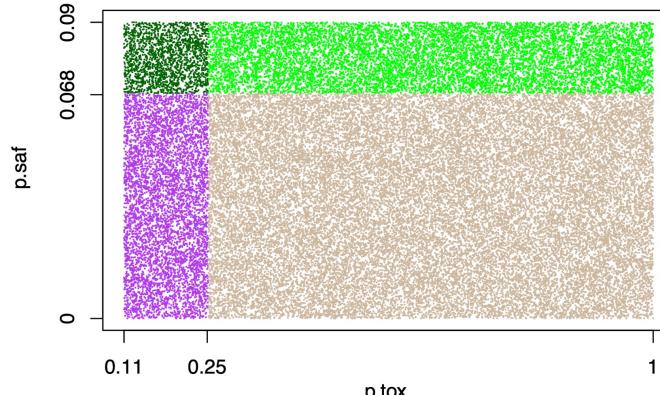
8 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cohortsize = 4)



2 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cohortsize = 5)



4 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, cohortsize = 6)



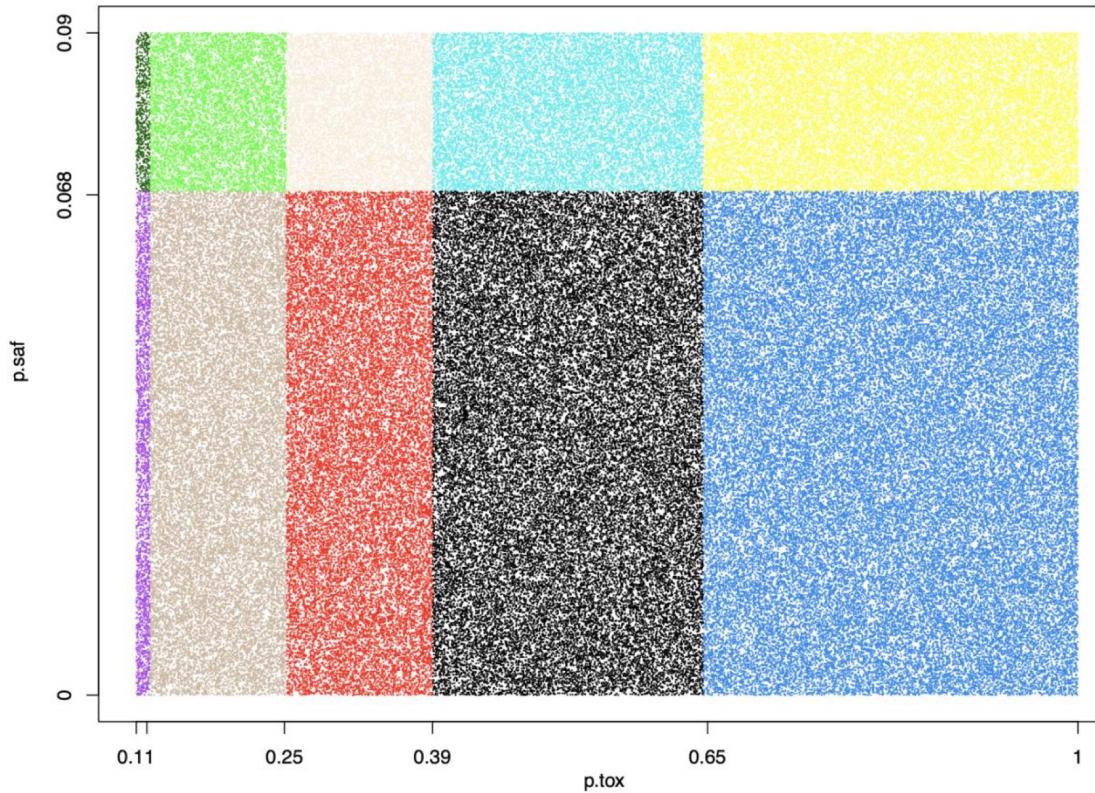
Equivalent p.saf and p.tox under varying ncohort

Equivalent values of p.saf and p.tox were explored via uniform search under varying ncohort = 5, 6, 7, 8, 9, 12, 20, 40, or and fixed values of following design parameters:

- target = 10% (target DLT rate)
- cohortsize = 3
- n.earlystop = 12
- cutoff.elim = 95%
- extrasafe = FALSE

For each ncohort value, 100,000 pairs of p.saf and p.tox values were randomly drawn from the uniform distributions described in the previous sections

10 BOIN boundary tables indicated by different colors
(target DLT rate = 10%, n.earlystop = 12)



When target DLT rate = 10%, cohortsize = 3, n.earlystop = 12, cutoff.eli = 95%, and extrasafe = FALSE, there are the same 10 possible BOIN boundary tables, regardless of the choices of ncohort $\in \{5, 6, 7, 8, 9, 10, 12, 20, 40, 100\}$.

```

> get.boundary(target=0.1, ncohort=5, cohortsize=3, n.earlystop = 12,
+               p.saf = 0.068, p.tox = 0.14, cutoff.eli = 0.95,
+               extrasafe = F)
$lambda_e
[1] 0.08306706

$lambda_d
[1] 0.1190318

$boundary_tab
Number of patients treated 3 6 9 12
Escalate if # of DLT <=    0 0 0 0
Deescalate if # of DLT >=  1 1 2 2
Eliminate if # of DLT >=  2 2 3 3

$full_boundary_tab
Number of patients treated 1 2 3 4 5 6 7 8 9 10 11 12
Escalate if # of DLT <=    0 0 0 0 0 0 0 0 0 0 0 0
Deescalate if # of DLT >=  1 1 1 1 1 1 1 2 2 2 2
Eliminate if # of DLT >=  NA NA 2 2 2 2 2 3 3 3 3 3

attr("class")
[1] "boin"

```

```

> get.boundary(target=0.1, ncohort=10000, cohortsize=3, n.earlystop = 12,
+               p.saf = 0.00001, p.tox = 0.14, cutoff.elim = 0.95,
+               extrasafe = F)
$lambda_e
[1] 0.01130893

$lambda_d
[1] 0.1190318

$boundary_tab
Number of patients treated 3 6 9 12
Escalate if # of DLT <=    0 0 0 0
Deescalate if # of DLT >=  1 1 2 2
Eliminate if # of DLT >=   2 2 3 3

$full_boundary_tab
Number of patients treated 1 2 3 4 5 6 7 8 9 10 11 12
Escalate if # of DLT <=    0 0 0 0 0 0 0 0 0 0 0 0
Deescalate if # of DLT >=  1 1 1 1 1 1 1 2 2 2 2
Eliminate if # of DLT >=   NA NA 2 2 2 2 2 3 3 3 3 3

attr(,"class")
[1] "boin"

```

BOIN parameters for generating the 3+3 boundary table

Equivalent values of BOIN design parameters were explored via uniform search under varying offset = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, or 0.49 and fixed values of following design parameters:

- cohortsize = 3
- ncohort = 10
- n.earlystop = 6
- extrasafe = TRUE
- target <- runif(1,min=0,max=0.5)
- p.saf <- runif(1,min=0,max=target-0.0000001)
- p.tox <- runif(1,min=target+0.0000001,max=1)
- cutoff.eli <- runif(1,min=0,max=1)

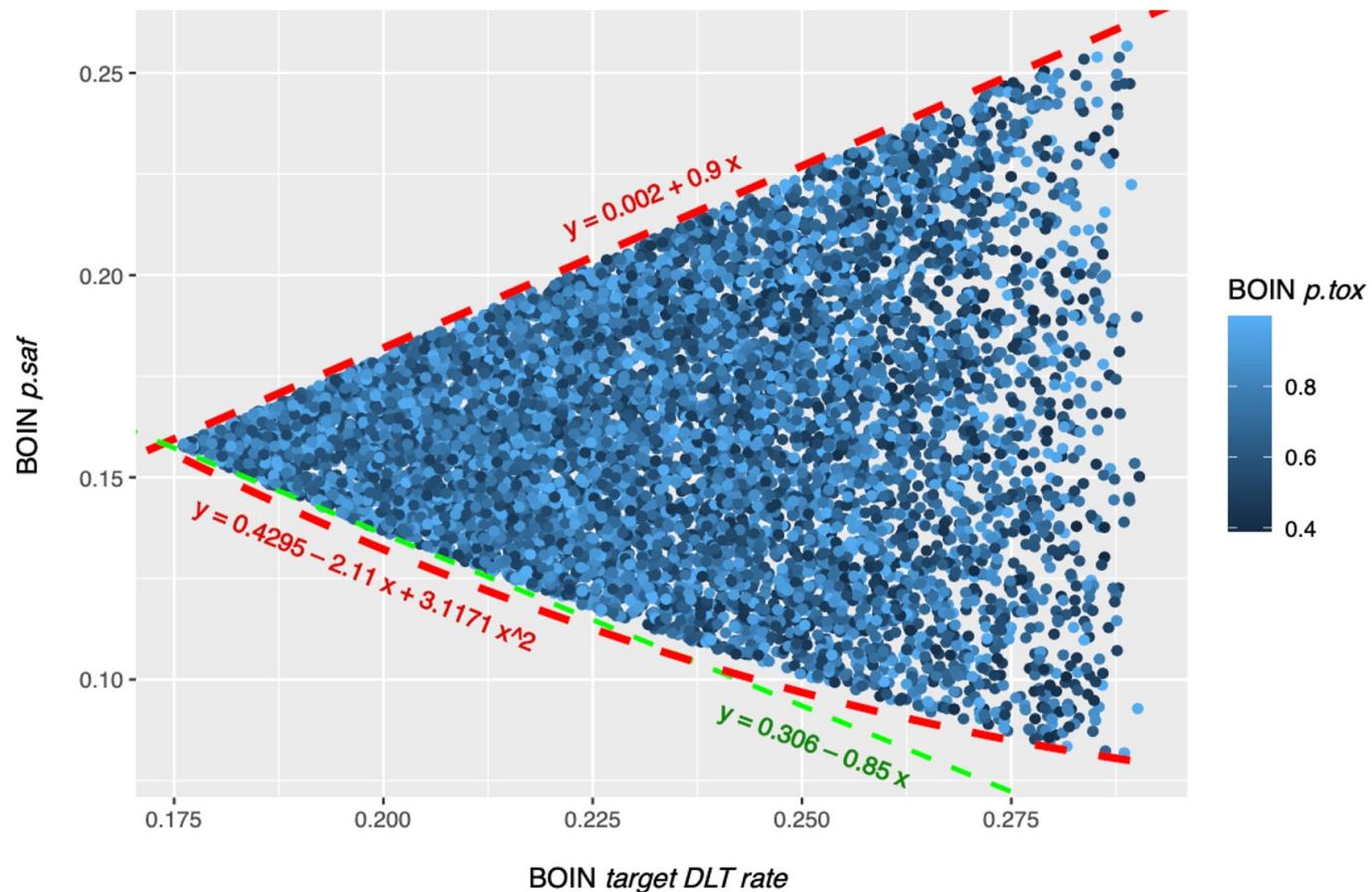
All parameter sets of BOIN design capable of yielding the 3+3 boundary table are viewed as being equivalent.

3 + 3 BOIN Design

A random search script was used to sample 8,127 sets of BOIN parameters that can generate the 3+3 design boundary table. These 8,127 sets of BOIN parameter values satisfy following conditions:

- $0.17 < \text{target.DLT.rate} < 0.30$
- $0.08 < p.saf < 0.26$
- $0.306 - 0.85 * \text{target.DLT.rate} < p.saf < 0.0002 + 0.9 * \text{target.DLT.rate}$
- $p.saf > 0.4295 - 2.11 * \text{target.DLT.rate} + 3.1171 * \text{target.DLT.rate}^2$

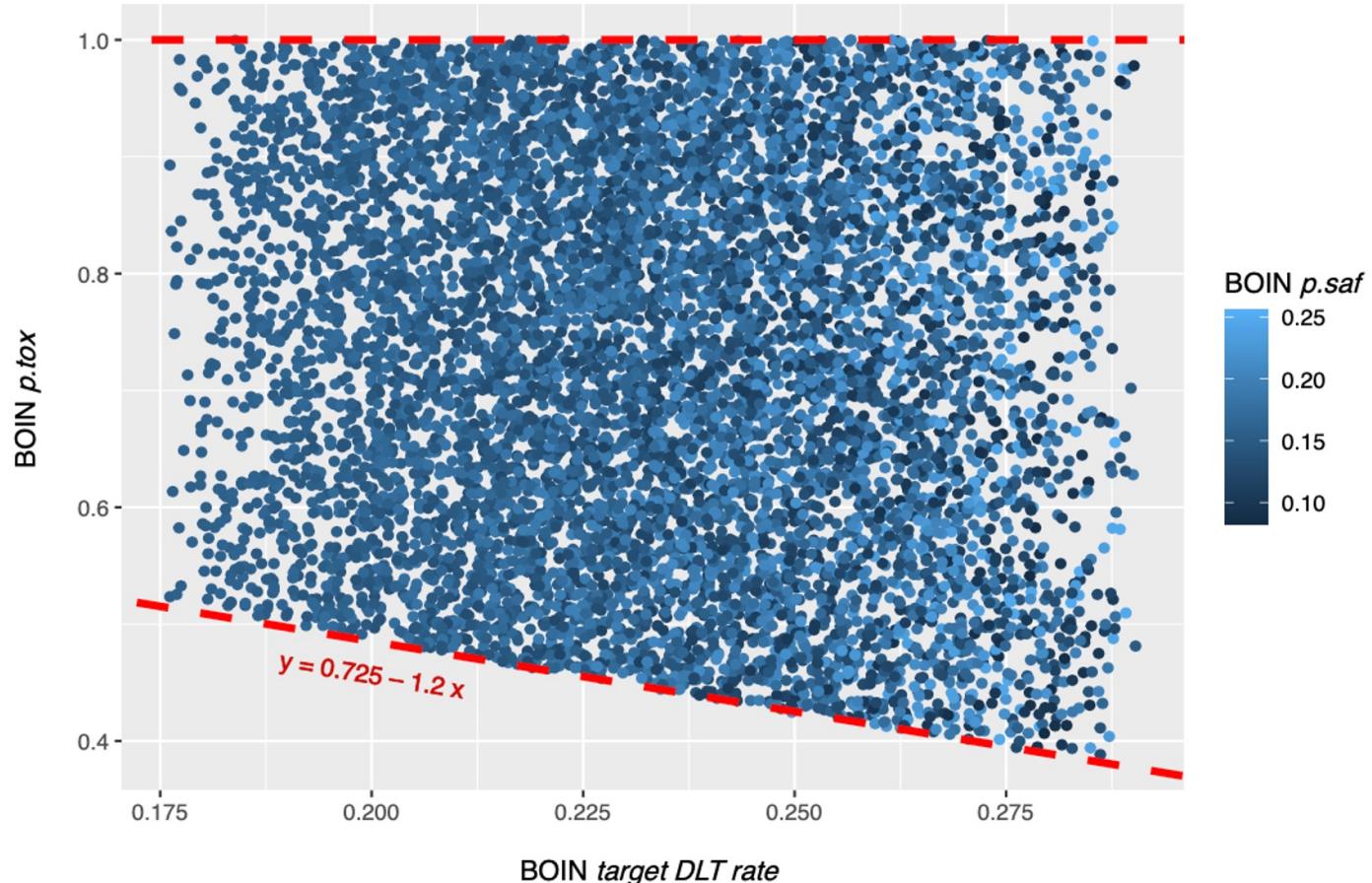
To generate the 3+3 design boundary table using BOIN designs:



3 + 3 BOIN Design

- $0.17 < \text{target.DLT.rate} < 0.30$
- $0.08 < p.saf < 0.26$
- $0.38 < p.tox < 1$
- $0.725 - 1.2 * \text{target.DLT.rate} < p.tox < 1$

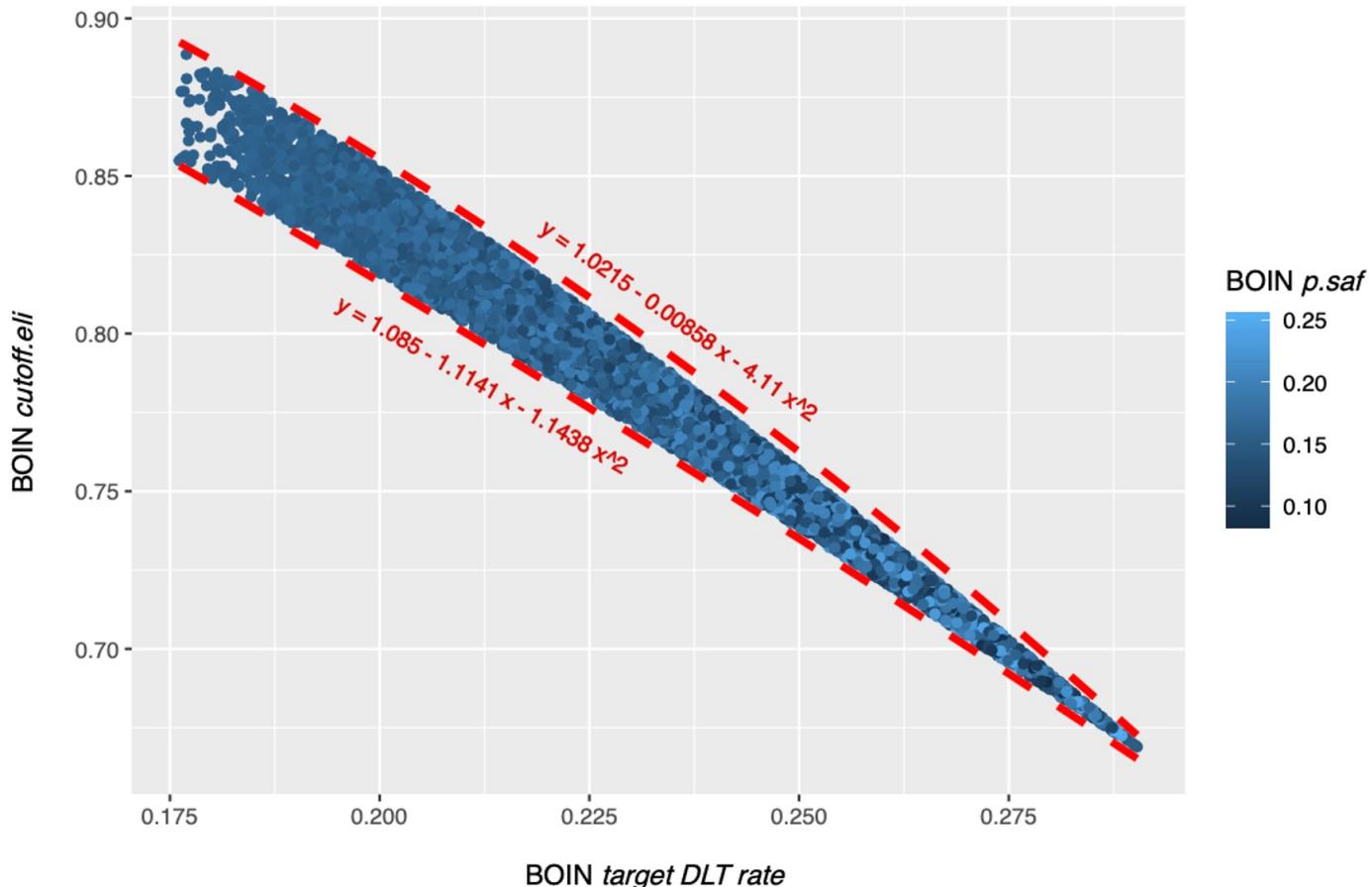
To generate the 3+3 design boundary table using BOIN designs:



3 + 3 BOIN Design

- $0.17 < \text{target.DLT.rate} < 0.30$
- $0.08 < p.saf < 0.26$
- $0.38 < p.tox < 1$
- $0.66 < \text{cutoff.eli} < 0.89$
- $\text{cutoff.eli} > 1.085 - 1.1141 * \text{target.DLT.rate} - 1.1438 * \text{target.DLT.rate}^2$
- $\text{cutoff.eli} < 1.0215 - 0.00858 * \text{target.DLT.rate} - 4.11 * \text{target.DLT.rate}^2$

To generate the 3+3 design boundary table using BOIN designs:



Summary/Discussion:

- In addition to target DLT rate, equivalent intervals of $p.tox$ also depend heavily on *cohortsize*, *cutoff.eli* and *n.earlystop*.
- Although *ncohort* is one of the required input parameter of *get.boundary()*, the choice of *ncohort* value has no effect on the calculation of BOIN boundary table. The main purpose of specifying *ncohort* is to terminate dose finding process when the sample size budget is reached.
- When the early stopping parameter *n.earlystop* is relatively small or the *cohortsize* value is not optimized via simulation, it might be better to use $p.tox < 1.4 * \text{target.DLT.rate}$, or try out different cohort sizes, or increase *n.earlystop*, whichever is both feasible and provides better operating characteristics.
 - When target DLT rate = 10%, *cohortsize* = 5, *ncohort* = 10, *n.earlystop* = 12, *cutoff.eli* = 95%, and *extrasafe* = FALSE, using $p.tox = 1.4 * \text{target.DLT.rate}$ to calculate BOIN boundary table is equivalent to using $p.tox > 3 * \text{target.DLT.rate}$.

Summary/Discussion:

- While changing target DLT rate, *cohortsize*, and *n.earlystop* will affect the equivalent intervals for both *p.saf* and *p.tox*, increasing or decreasing *cutoff.elิ* will only affect *p.tox* equivalent intervals.
- It appears that increasing *cutoff.elิ* will add more equivalent interval boundary points from both side of $p.tox = 0.5$ but won't be able to narrow *p.tox* intervals that are either close to target DLT rate (plus a small margin) or close to 1.
- And *cutoff.elิ* < 90% may need to be used with caution because the resulted equivalent interval of *p.saf* could be too wide for some pediatric trials.
 - when target DLT rate = 10%, *cohortsize* = 3, *ncohort* = 10, *n.earlystop* = 12, *cutoff.elи* = 80%, and *extrasafe* = FALSE, using *p.tox* = $1.4 * \text{target.DLT.rate}$ to calculate BOIN boundary table is equivalent to using any $p.tox \in (12.3\%, 99.9\%)$, as long as *p.saf* values fall into one equivalent interval of *p.saf*.

Summary/Discussion:

- This research highlights the importance of interpreting the BOIN design parameter p.tox as a range of toxicity rates regarded as excessively toxic, as opposed to one pre-determined value reflecting the lowest toxicity probability deemed overly toxic.
- It's also essential to perform simulation studies to recognize comparable sets of BOIN design parameters capable of producing an identical boundary table.

REFERENCES I

- Liu, Suyu, and Ying Yuan. "Bayesian optimal interval designs for phase I clinical trials." *Journal of the Royal Statistical Society: Series C: Applied Statistics* (2015): 507-523. (Erratum published online on 11Jan2022: <https://rss.onlinelibrary.wiley.com/doi/full/10.1111/rssc.12541>)
- Yuan, Ying, Kenneth R. Hess, Susan G. Hilsenbeck, and Mark R. Gilbert. "Bayesian optimal interval design: a simple and well-performing design for phase I oncology trials." *Clinical Cancer Research* 22, no. 17 (2016): 4291-4301.
- Lin, Ruitao, and Guosheng Yin. "Bayesian optimal interval design for dose finding in drug-combination trials." *Statistical methods in medical research* 26, no. 5 (2017): 2155-2167
- Yan, Fangrong, Liangcai Zhang, Yanhong Zhou, Haitao Pan, Suyu Liu, and Ying Yuan. "BOIN: an R package for designing single-agent and drug-combination dose-finding trials using Bayesian optimal interval designs." *Journal of Statistical Software* 94 (2020): 1-32
- Ananthakrishnan, Revathi, Ruitao Lin, Chunsheng He, Yanping Chen, Daniel Li, and Michael LaValley. "An overview of the BOIN design and its current extensions for novel early-phase oncology trials." *Contemporary Clinical Trials Communications* 28 (2022): 100943 .

REFERENCES II

- FDA Drug Development Tools: Fit-for-Purpose Initiative Determination Letter Issued December 10, 2021: <https://www.fda.gov/media/155363/download?attachment>
- FDA Drug Development Tools: Fit-for-Purpose Initiative Statistical Review Issued December 10, 2021: <https://www.fda.gov/media/155364/download?attachment>
- Storer, Barry E. "Design and analysis of phase I clinical trials." *Biometrics* (1989): 925-937
- Storer, Barry E. "An evaluation of phase I clinical trial designs in the continuous dose-response setting." *Statistics in medicine* 20, no. 16 (2001): 2399-2408
- Chen, Zhengjia, Youyun Zheng, Zhibo Wang, Michael Kutner, Walter J. Curran, and Jeanne Kowalski. "Interactive calculator for operating characteristics of phase I cancer clinical trials using standard 3+ 3 designs." *Contemporary Clinical Trials Communications* 12 (2018): 145-153.
- <https://www.medrxiv.org/content/10.1101/2024.03.06.24303862v2.full.pdf>
- All scripts and data produced are available online at https://github.com/ronglu-stanford/BOIN_p.tox_17Jan2024

ACKNOWLEDGEMENT

- **Funding Sources:** This work is partially supported by the Biostatistics Shared Resource (BSR) of the NIH-funded Stanford Cancer Institute (P30CA124435) and the Stanford Center for Clinical and Translational Research and Education (UL1- TR003142).



Thank you for your attention!