BOIN: An **R** Package for Designing Single-Agent and Drug-Combination Dose-Finding Trials Using Bayesian Optimal Interval Designs

Fangrong Yan
China Pharmaceutical
University

Liangcai Zhang
Janssen Pharmaceutica,
Inc.

Yanhong Zhou
The University of Texas
MD Anderson Cancer Center

Haitao Pan
St. Jude Children's
Research Hospital

Suyu Liu
The University of Texas
MD Anderson Cancer Center

Ying Yuan
The University of Texas
MD Anderson Cancer Center

Abstract

This article describes the R package **BOIN**, which implements a recently developed methodology for designing single-agent and drug-combination dose-finding clinical trials using Bayesian optimal interval designs (Liu and Yuan 2015; Yuan, Hess, Hilsenbeck, and Gilbert 2016). The BOIN designs are novel "model-assisted" phase I trial designs that can be implemented simply and transparently, similar to the $3 + 3$ design, but yield excellent performance comparable to those of more complicated, model-based designs. The **BOIN** package provides tools for designing, conducting, and analyzing single-agent and drug-combination dose-finding trials.

Keywords: maximum tolerated dose, dose finding, phase I trials, model-assisted design, Bayesian adaptive design.

1. Introduction

The objective of a phase I trial is to find the maximum tolerated dose (MTD), which is defined as the dose or dose combination with the dose limiting toxicity (DLT) probability closest to the target DLT rate. Numerous phase I trial designs have been proposed to find the MTD. The $3 + 3$ design (Storer 1989) has been dominant in phase I clinical trials for decades due to its simplicity, despite its poor ability to identify the MTD and tendency to

treat patients at low doses that are potentially subtherapeutic (Le Tourneau, Lee, and Siu 2009). The $3+3$ design and its variations are called algorithm-based designs because they use simple, prespecified rules to guide dose escalation. Model-based designs have been proposed to improve the efficiency of identifying the MTD. A typical example of model-based designs is the continual reassessment method (CRM; O'Quigley, Pepe, and Fisher 1990). The CRM yields better performance than the $3 + 3$ design, but is statistically and computationally complex, leading practitioners to perceive dose allocations as coming from a "black box", which limits its application in practice.

The Bayesian optimal interval (BOIN) design (Liu and Yuan 2015; Yuan *et al.* 2016) is a novel phase I trial methodology, belonging to a new class of designs known as model-assisted designs (Yan, Mandrekar, and Yuan 2017; Zhou, Murray, Pan, and Yuan 2018a), that combines the simplicity of algorithm-based designs with the superior performance of model-based designs. The BOIN design can be implemented in a simple way as the $3+3$ design, but yields excellent performance comparable to that of the more complicated, model-based CRM (Zhou, Yuan, and Nie 2018b; Zhou *et al.* 2018a). Another example of model-assisted designs is the keyboard design (Yan *et al.* 2017). Because of their good performance and simplicity, model-assisted designs have been increasingly used in practice. Clertant and O'Quigley (2017) proposed a semiparametric approach that provides the link between the BOIN and other designs, such as the cumulative cohort design (Ivanova, Flournoy, and Chung 2007).

The objective of this article is to provide practitioners a comprehensive, well-documented, easy-to-use R (R Core Team 2020) package to design phase I clinical trials using the BOIN design. The **BOIN** package (Yuan and Liu 2020) incorporates a number of practical considerations and functionalities that are not considered in the original methodological paper by Liu and Yuan (2015). For example, the BOIN R package allows users to early stop the trial when the number of patients accumulated in a dose reaches a certain cutoff. To the best of our knowledge, this is the first R package that provides a comprehensive implementation of the BOIN design, including simulation functions for generating operating characteristics of the design and implementation functions for conducting a real trial in practice. The existing dose-finding R packages, such as **TEQR** (Blanchard 2016) and **DoseFinding** (Bornkamp, Pinheiro, and Bretz 2009; Bornkamp 2019), cover algorithm-based designs (i.e., the cumulative cohort design) and model-based designs (i.e., MCP-Mod design), respectively, and focus on single-agent dose finding. In contrast, the **BOIN** R package focuses on the model-assisted design, and covers both single-agent and drug-combination dose findings. Drug combination trials are increasingly important in drug development.

2. Bayesian optimal interval designs

2.1. BOIN design for single-agent trials

Let $d_1 < \ldots < d_J$ denote the J prespecified doses of the new drug that is under investigation in the trial, p_j denote the dose-limiting toxicity (DLT) probability that corresponds to d_j, and ϕ denote the target DLT rate for the MTD. We use n_j to denote the number of patients who have been assigned to d_j, and y_j of them experienced DLTs, $j = 1, \ldots, J$. The BOIN design is illustrated in Figure 1 and summarized as follows:

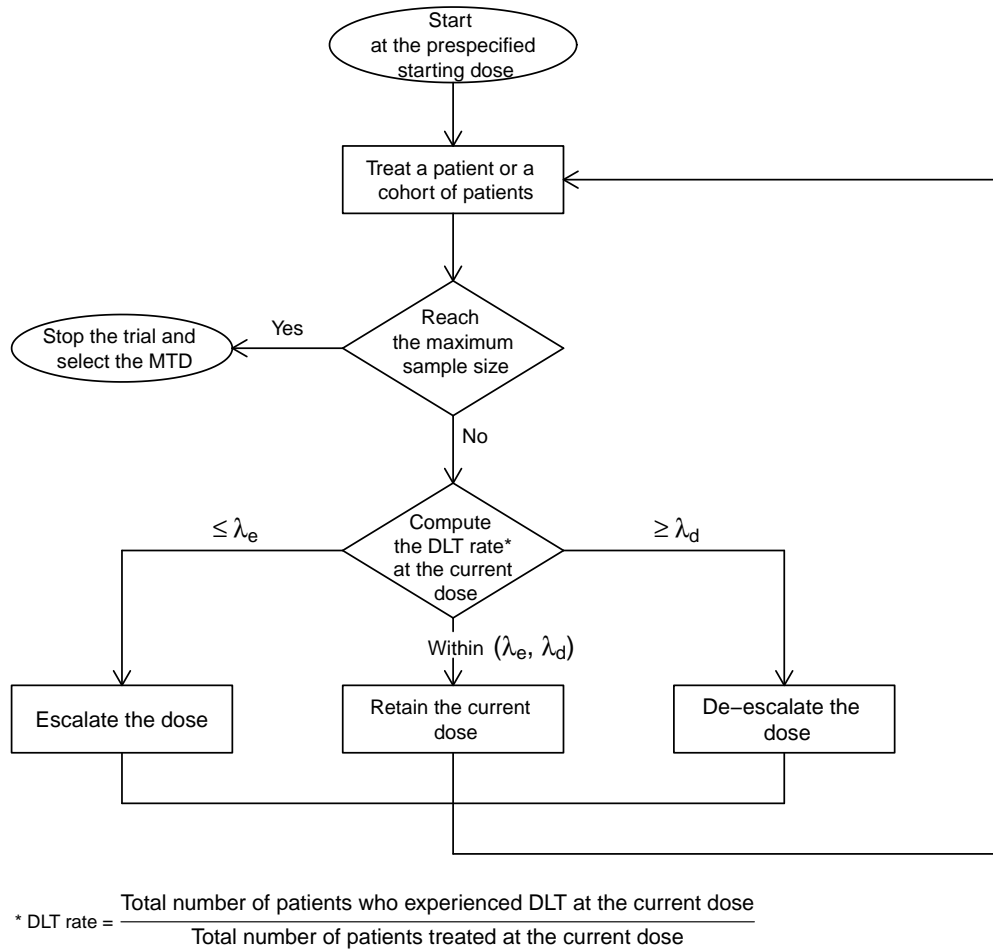 (a) Patients in the first cohort are treated at the lowest or a prespecified dose level.

Figure 1: The flowchart of the BOIN design.

(b) Let $\hat{p}_j = y_j/n_j$ be the observed DLT rate at the current dose level j, to assign a dose to the next cohort of patients,

- if $\hat{p}_j \leq \lambda_e$, escalate the dose level to $j + 1$,
- if $\hat{p}_j \geq \lambda_d$, de-escalate the dose level to $j - 1$,
- otherwise, i.e., $\lambda_e < \hat{p}_j < \lambda_d$, stay at the same dose level,

where λ_e and λ_d are prespecified dose escalation and de-escalation boundaries (see Table 1). The method of specifying λ_e and λ_d is described below.

(c) Repeat step (b) until the prespecified maximum sample size $N_{\max}$ is exhausted or the trial is terminated because of excessive toxicity, as described next. At that point, the MTD is selected as the dose for which the isotonic estimate of the DLT probability is closest to the target DLT rate ϕ. The isotonic estimates of the DLT probabilities can be obtained using the pooled adjacent violators algorithm (Barlow, Bartholomew, Bremner, and Brunk 1972). In the case that the trial is terminated because of excessive toxicity, no dose should be selected as the MTD.

Boundaries	Target DLT rate ϕ					
	0.15	0.20	0.25	0.30	0.35	0.40
λ_e	0.118	0.157	0.197	0.236	0.276	0.316
λ_d	0.179	0.238	0.298	0.359	0.419	0.480

Table 1: Dose escalation and de-escalation boundaries of the BOIN design for different target DLT rates.

For patient safety, during the trial conduct, the BOIN design imposes a dose elimination/safety stopping rule as follows:

> Suppose j is the current dose level. If $\mathsf{P}(p_j > \phi|y_j, n_j) > 0.95$ and $n_j \geq 3$, dose level j and higher doses are eliminated from the trial. The trial is terminated if the lowest dose is eliminated.

The posterior probability $\mathsf{P}(p_j > \phi|y_j, n_j)$ can be evaluated on the basis of a beta-binomial model, assuming $y_j \,|\, n_j, p_j \sim \mathrm{Binom}(n_j, p_j)$ and $p_j \sim \mathrm{Beta}(1,1) \equiv \mathrm{Unif}(0,1)$. Under the beta-binomial model, the posterior distribution of p_j arises as

$$p_j \,|\, y_j, n_j \sim \mathrm{Beta}(y_j + 1, n_j - y_j + 1), \text{ for } j = 1, \ldots, J. \tag{1}$$

The dose escalation and de-escalation boundaries (λ_e, λ_d) are obtained by minimizing the chance of making incorrect dose escalation and de-escalation (Liu and Yuan 2015). To do that, the BOIN design requires the investigator(s) to specify ϕ_1, which is the highest DLT rate that is deemed to be underdosing such that dose escalation is required, and ϕ_2, which is the lowest DLT rate that is deemed to be overdosing such that dose de-escalation is required. Liu and Yuan (2015) provided general guidance to specify ϕ_1 and ϕ_2, and recommended default values of $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$ for general use. When needed, the values of ϕ_1 and ϕ_2 can be calibrated to achieve a particular requirement of the trial at hand. For example, if more conservative dose escalation is required, setting $\phi_2 = 1.2\phi$ may be appropriate. We should avoid setting the values of ϕ_1 and ϕ_2 very close to ϕ. This is because the small sample sizes of typical phase I trials prevent us from differentiating the target DLT rate from the probabilities that are close to it. For example, at the significance level of 0.1, there is only 7% power to distinguish 0.25 from 0.35 with 15 patients at each dose, based on Fisher's exact test. Given ϕ_1 and ϕ_2, the optimal escalation and de-escalation boundaries (λ_e, λ_d) that minimize the incorrect decision of dose escalation and de-escalation arise as

$$\lambda_e = \frac{\log\left(\frac{1 - \phi_1}{1 - \phi}\right)}{\log\left\{\frac{\phi\left(1 - \phi_1\right)}{\phi_1\left(1 - \phi\right)}\right\}}, \qquad \lambda_d = \frac{\log\left(\frac{1 - \phi}{1 - \phi_2}\right)}{\log\left\{\frac{\phi_2\left(1 - \phi\right)}{\phi\left(1 - \phi_2\right)}\right\}}.$$

Table 1 provides the dose escalation and de-escalation boundaries (λ_e, λ_d) for commonly used target DLT rates ϕ using the recommended default values $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$.

For example, given the target DLT rate $\phi = 0.25$, the corresponding escalation boundary $\lambda_e = 0.197$ and the de-escalation boundary $\lambda_d = 0.298$, that is, escalate the dose if the observed DLT rate at the current dose $\hat{p}_j \leq 0.197$ and de-escalate the dose if $\hat{p}_j \geq 0.298$. Liu and Yuan (2015) showed that (λ_e, λ_d) are the boundaries corresponding to the likelihood

ratio test and Bayes factor, and thus the resulting BOIN design is optimal with desirable finite-sample and large-sample properties, i.e., long-memory coherence and consistency.

As the observed DLT rate $\hat{p}_j$ is the most natural and intuitive estimate of p_j that is accessible by non-statisticians, it is easy to explain the BOIN design to clinicians. In addition, thanks to the feature that the BOIN design guarantees de-escalating the dose when $\hat{p}_j$ is higher than the de-escalation boundary λ_d, it is particularly easy for clinicians and regulatory agents to assess the safety of a trial using the BOIN design. For example, given a target DLT rate $\phi = 0.25$, we know *a priori* that a phase I trial using the BOIN design guarantees de-escalating the dose if the observed DLT rate is higher than 0.298 (with the default values of ϕ_1 and ϕ_2). Accordingly, the BOIN design also allows users to easily calibrate the design to satisfy a specific safety requirement mandated by regulatory agents through choosing an appropriate target DLT rate ϕ or ϕ_2. For example, suppose for a phase I trial with a new compound, the regulatory agent mandates that if the observed DLT rate is higher than 0.25, the dose must be de-escalated. We can easily fulfill that requirement by setting the target DLT rate $\phi = 0.21$, under which the BOIN automatically guarantees de-escalating the dose if the observed DLT rate $\hat{p}_j > \lambda_d = 0.250$. If needed, the de-escalation boundary λ_d can be further fine tuned by calibrating the value of ϕ_2. Such flexibility and transparency gives the BOIN design an important advantage over many existing designs in practice.

2.2. BOIN designs for drug-combination trials

Drug combination therapy provides a critical approach to improve treatment efficacy and overcome resistance to monotherapies. One challenge in designing drug-combination trials is that, due to complicated drug-drug interactions, combinations are only partially ordered in toxicity probabilities. Consider a trial combining J doses of drug A, denoted as $A_1 < A_2 < \ldots < A_J$, and K doses of drug B, denoted as $B_1 < B_2 < \ldots < B_K$. Let A_jB_k denote the combination of A_j and B_k, and p_{jk} denote the true DLT rate of A_jB_k. We know a priori that $A_{j'}B_{k'}$ is more toxic than A_jB_k for $j' > j$ and $k' > k$, but typically do not know their toxicity order if $j' > j$ but $k' < k$. Another challenge in designing combination trials is the existence of the MTD contour in the two-dimensional dose space, as shown in Figure 2. In other words, multiple MTDs may exist in the $J \times K$ dose matrix. Depending on the application, combination trials may be interested in finding a single MTD or the MTD contour (i.e., multiple MTDs; Yuan and Zhang 2017). Because of these challenges, the single-agent phase I trial designs cannot be directly used for combination trials. In what follows, we first describe the BOIN drug-combination design (Lin and Yin 2017a) that aims to find a single MTD in the two-dimensional dose combination matrix, followed by a BOIN-waterfall design (Zhang and Yuan 2016) that aims to find the MTD contour (i.e., multiple MTDs)

BOIN drug-combination design for finding a single MTD

Let $\hat{p}_{jk} = y_{jk}/n_{jk}$ denote the observed DLT rate at dose combination A_jB_k, where n_{jk} denotes the number of patients treated at A_jB_k and y_{jk} denotes the number of patients who experienced DLT at A_jB_k. Given the current dose combination A_jB_k, define an admissible dose escalation set as $\mathcal{A}_E = \{A_{j+1}B_k, A_jB_{k+1}\}$ and an admissible dose de-escalation set as $\mathcal{A}_D = \{A_{j-1}B_k, A_jB_{k-1}\}$. The BOIN drug-combination design makes the decision of dose escalation/de-escalation based on the same rule as the single-agent BOIN design described previously. That is, escalate the dose if $\hat{p}_{jk} \leq \lambda_e$, and de-escalate the dose if $\hat{p}_{jk} \geq \lambda_d$. The
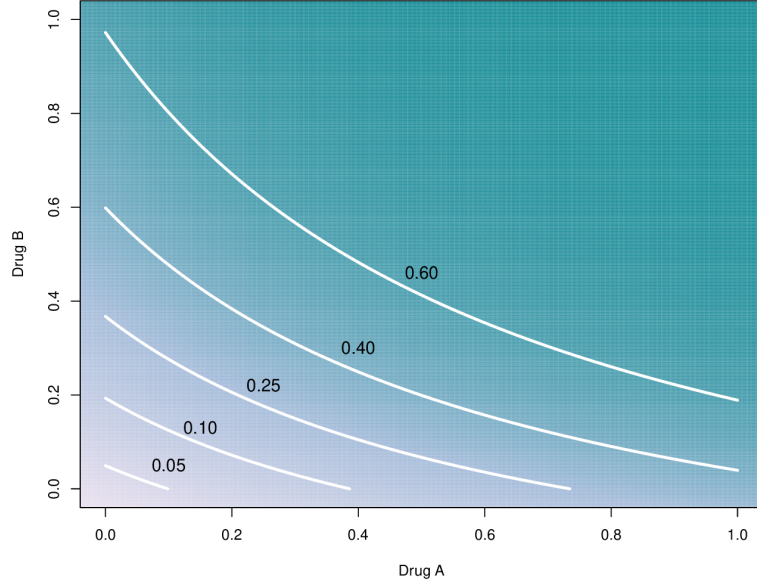
Figure 2: The MTD contour in the drug-combination space.

only difference is that, in combination trials, when we decide to escalate or de-escalate the dose, there is more than one neighbor dose to which we can move. For example, when we escalate/de-escalate the dose, we can escalate/de-escalate either the dose of drug A or the dose of drug B. The BOIN drug-combination design makes this choice based on $\mathsf{P}(p_{jk} \in (\lambda_e, \lambda_d)|D_{jk})$, which measures how likely a dose combination is located within the acceptable toxicity interval (λ_e, λ_d), where $D_{jk} = (y_{jk}, n_{jk})$. Specifically, the BOIN drug-combination design can be described as follows:

(a) Patients in the first cohort are treated at the lowest dose combination A_1B_1 or a pre-specified dose combination.

(b) Suppose the current cohort is treated at dose combination A_jB_k. Proceed as follows to assign a dose to the next cohort of patients.

 - If $\hat{p}_{jk} \leq \lambda_e$, we escalate the dose to the combination that belongs to $\mathcal{A}_E$ and has the largest value of $\mathsf{P}\{p_{j'k'} \in (\lambda_e, \lambda_d)|D_{jk}\}$.

 - If $\hat{p}_{jk} \geq \lambda_d$, we de-escalate the dose to the combination that belongs to $\mathcal{A}_D$ and has the largest value of $\mathsf{P}\{p_{j'k'} \in (\lambda_e, \lambda_d)|D_{jk}\}$.

 - Otherwise, i.e., $\lambda_e < \hat{p}_{jk} < \lambda_d$, the dose stays at the same combination A_jB_k.

(c) Step (b) is continued until the maximum sample size is reached or the trial is terminated because of excessive toxicity.

During dose escalation and de-escalation, if the two combinations in $\mathcal{A}_E$ or $\mathcal{A}_D$ have the same value of $\mathsf{P}\{p_{j'k'} \in (\lambda_e, \lambda_d)|\text{data}\}$, we randomly choose one with equal probability. If no dose combination exists in the sets of $\mathcal{A}_E$ and $\mathcal{A}_D$ (i.e., we are at the boundaries of the dose matrix), we retain the current dose combination. After the trial is completed, the MTD is

selected as the dose combination with the estimated DLT rate closest to ϕ. The estimates of DLT rates are obtained using isotonic regression as described previously, but in a matrix form. More details on the BOIN drug-combination design can be found in Lin and Yin (2017a).

BOIN-waterfall design for finding the MTD contour

Because of the existence of the MTD contour and the fact that doses on the MTD contour may have different efficacy due to drug-drug interactions, for many drug combination trials, it is of intrinsic interest to find multiple MTDs (or the MTD contour). The efficacy of the MTDs can be evaluated in subsequent phase II trials or simultaneously in phase I-II trials. Zhang and Yuan (2016) extended the BOIN design to achieve this design goal. For convenience, we refer to that design as the BOIN-waterfall design. As illustrated in Figure 3, the BOIN-waterfall design takes the divide-and-conquer strategy to divide the task of finding the MTD contour into a sequence of one-dimensional dose-finding tasks, known as "subtrials" (Yuan and Yin 2008). These subtrials are conducted sequentially from the top of the matrix to the bottom.

Specifically, the BOIN-waterfall design can be described as follows:

1. Divide the $J \times K$ dose matrix into J subtrials $S_J, \ldots, S_1$, according to the dose level of drug A:

$$
\begin{aligned}
S_J &= \{A_1B_1, \ldots, A_JB_1, A_JB_2, \ldots, A_JB_K\}, \\
S_{J-1} &= \{A_{J-1}B_2, \ldots, A_{J-1}B_K\}, \\
S_{J-2} &= \{A_{J-2}B_2, \ldots, A_{J-2}B_K\}, \\
&\ldots \\
S_1 &= \{A_1B_2, \ldots, A_1B_K\}.
\end{aligned}
$$

Note that subtrial S_J also includes lead-in doses $A_1B_1, A_2B_1, \ldots, A_JB_1$ (the first column of the dose matrix) to impose the practical consideration that the trial starts at the lowest dose combination. Within each subtrial, the doses are fully ordered with monotonically increasing toxicity.

2. Conduct the subtrials sequentially using the single-agent BOIN design as follows:

   (i) Conduct subtrial S_J, starting from the lowest dose combination A_1B_1, to find the MTD. We call the dose selected by the subtrial the "candidate MTD" to highlight that the dose selected by the individual subtrial may not be the "final" MTD that we will select at the end of the trial. The final MTD selection will be based on the data collected from all the subtrials. The objective of finding the candidate MTD is to determine which subtrial will be conducted next and the corresponding starting dose combination.

   (ii) Assuming that the current subtrial S_j, $j = J, \ldots, 2$, selects dose $A_{j^*}B_{k^*}$ as the candidate MTD, next, conduct subtrial S_{j^*-1} with the starting dose $A_{j^*-1}B_{k^*+1}$. That is, the next subtrial to be conducted is the one with the dose of drug A that is one level lower than the candidate MTD found in the previous subtrial. After identifying the candidate MTD of subtrial S_{j^*-1}, the same rule is used to determine the next subtrial and its starting dose combination. See Figure 3 for an example.
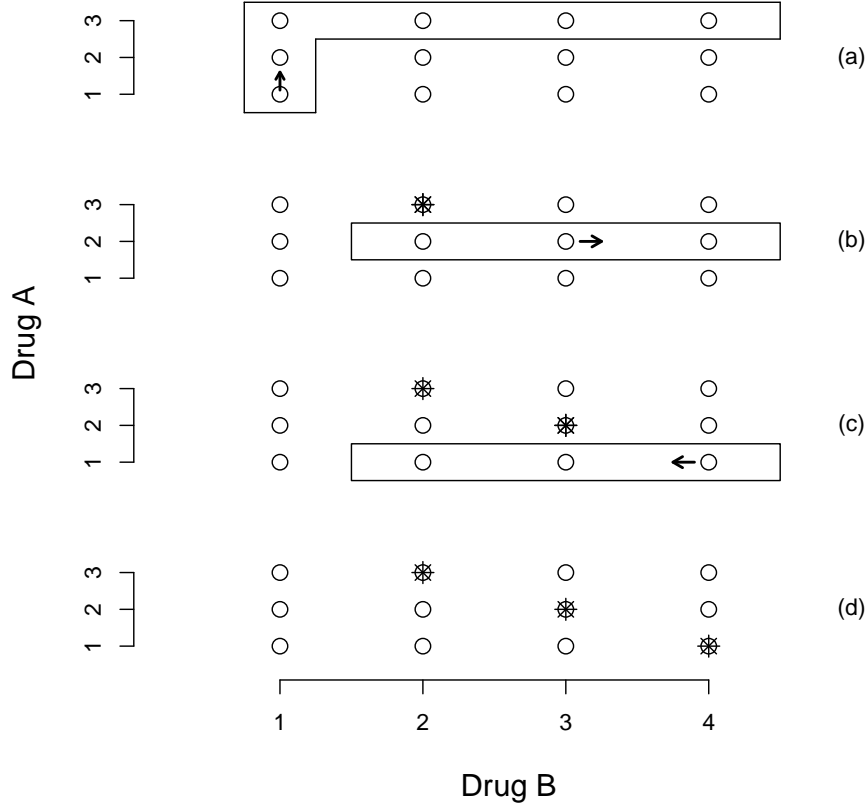
Figure 3: Illustration of the waterfall design for a combination trial of drug A with 3 levels and drug B with 4 levels. The doses in the rectangle form a subtrial, and the asterisk denotes the candidate MTD. As shown in panel (a), the trial started by conducting the first subtrial with the starting dose A_1B_1. After the first subtrial identified A_3B_2 as the candidate MTD, we conducted the second subtrial with the starting dose A_2B_3 (see panel (b)). After the second subtrial identified A_2B_3 as the candidate MTD, we conducted the third subtrial with the starting dose A_1B_4 (see panel (c)). All subtrials were conducted using the single-agent BOIN design. After all the subtrials were completed, we selected the MTD contour based on the data from all the subtrials, as shown in panel (d).

    (iii) Repeat step (ii) until subtrial S_1 is completed.

3. Estimate the DLT rate p_{jk} based on the toxicity data collected from all the subtrials using matrix isotonic regression. For each row of the dose matrix, select the MTD as the dose combination that has the estimate of DLT rate closest to the target DLT rate ϕ, unless all combinations in that row are overly toxic.

For the BOIN-waterfall design, one key issue is to determine when to end the current subtrial and initiate the next one. Zhang and Yuan (2016) recommended the following stopping rule to switch from one subtrial to the next one:

    At any time during the subtrial, if the total number of patients treated at the current dose reaches a certain prespecified number of patients, say n_{stop}, we stop the subtrial, select the candidate MTD and initiate the next subtrial.

The rationale for the stopping rule is that when patient allocation concentrates at a given dose, it indicates that the dose finding might have converged to the MTD, and thus we can stop the trial and claim the MTD. This stopping rule allows the sample size of the subtrials to be automatically adjusted according to the difficulty of the dose finding. Another attractive feature of the above approach is that it automatically ensures that a certain number of patients are treated at the MTD. Conventionally, we achieve this by adding cohort expansion after identifying the MTD. In practice, we recommend $n_{\mathrm{stop}} > 9$ to ensure reasonable operating characteristics. Although the above stopping rule provides an automatic, reasonable way to determine the sample size for a subtrial, in some cases, it is desirable to put a cap on the maximum sample size of the subtrials. This can be done by adding an extra stopping rule as follows: Stop the subtrial if its sample size reaches N_{max}, where N_{max} is the prespecified maximum sample size.

3. The R package BOIN

The R package **BOIN** is freely available from the Comprehensive R Archive Network (CRAN) at https://CRAN.R-project.org/package=BOIN. It contains functions for implementing the BOIN design for both single-agent and drug-combination dose-finding trials (see Table 2). We proceed with a small tutorial to present the use of the package. As an illustration, we load the package:

```
R> install.packages("BOIN")
R> library("BOIN")
```

3.1. Single-agent trials

To design and conduct a single-agent dose-finding trial, we follow 3 steps as described below, which involve functions `get.boundary()`, `get.oc()` and `selet.mtd()`, respectively. The objects (or results) returned by these functions can be displayed in a user friendly form by using the functions `summary()`, `print()`, and `plot()`.

Step 1. Obtain dose escalation and de-escalation boundaries

Like the algorithm-based design, the dose escalation and de-escalation rule of the BOIN design can be determined before the onset of the trial, making its implementation transparent and simple. To obtain the dose escalation and de-escalation boundaries (λ_e, λ_d), we run the function:

```
get.boundary(target, ncohort, cohortsize, n.earlystop = 100,
  p.saf = 0.6 * target, p.tox = 1.4 * target, cutoff.eli = 0.95,
  extrasafe = FALSE, offset = 0.05)
```

This function has the following arguments:

- `target`: The target DLT rate.

- `ncohort`: The total number of cohorts.

Function	Description	Object returned
`get.boundary`	Generate dose escalation and de-escalation boundaries.	A list including dose escalation and de-escalation boundaries.
`select.mtd`	Select the maximum tolerated dose (MTD) for single-agent trials.	A list including the selected MTD and estimated toxicity probability for each dose.
`get.oc`	Generate operating characteristics for single-agent trials.	A list including the selection percentage, the number of patients treated at each dose, and average total number of patients.
`next.comb`	Determine the dose combination for treating the next cohort of new patients for drug-combination trials that aim to find an MTD.	A list including the recommended dose combination for the next cohort of patients.
`next.subtrial`	Determine the starting dose combination and the dose-searching space for the next subtrial in BOIN-waterfall design.	A list including dose combinations included in the subtrial and the starting dose combination.
`select.mtd.comb`	Select the MTD or MTD contour for drug combination trials.	A list containing the selected MTD or MTD contour, and estimated toxicity probability for each dose combination.
`get.oc.comb`	Generate operating characteristics for drug-combination trials.	A list containing the selection percentage, the number of patients treated at each dose, and average total number of patients.
`summary`	Generate summary for the objects returned by other functions.	Descriptive results printed to the console.
`plot`	Generate flowchart of the design and barplot for simulation results.	Figures.
`print`	Print objects returned by other functions.	Objects returned from other functions.

Table 2: Overview of the user visible functions in **BOIN**. Please consult the documentation (e.g., `help("get.boundary")`) for function arguments and detailed return types.

- `cohortsize`: The cohort size.

- `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.

- `p.saf`: 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 * target`.

- `p.tox`: The lowest toxicity probability that is deemed overly toxic such that dose de-escalation is required. The default value of `p.tox = 1.4 * target`.

- `cutoff.eli`: The cutoff to eliminate the overly toxic dose for safety. We recommend the default value `cutoff.eli = 0.95` for general use.

- `extrasafe`: Set `extrasafe = TRUE` to impose a stricter stopping rule.

- `offset`: A small positive number (between 0 and 0.5) to control how strict the stopping rule is when `extrasafe = TRUE`. A larger value leads to a stricter stopping rule. The default value `offset = 0.05` generally works well.

Suppose a phase I trial with $J = 5$ dose levels and a target DLT rate of $\phi = 0.3$, the maximum sample size is 30 patients, and patients are treated in cohorts of size 3. Using the default values of `p.saf`, `p.tox`, and `cutoff.eli` automatically provided by the function, we can design the trial by running `get.boundary()`, as follows:

```
R> bound <- get.boundary(target = 0.3, ncohort = 10, cohortsize = 3)
R> summary(bound)


Escalate dose if the observed DLT rate at the current dose <=  0.2364907
Deescalate dose if the observed DLT rate at the current dose >=  0.3585195


This is equivalent to the following decision boundaries

Number of patients treated 3 6 9 12 15 18 21 24 27 30
Escalate if # of DLT <=    0 1 2  2  3  4  4  5  6  7
Deescalate if # of DLT >=  2 3 4  5  6  7  8  9 10 11
Eliminate if # of DLT >=   3 4 5  7  8  9 10 11 12 14


A more completed version of the decision boundaries is given by

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


Number of patients treated 18 19 20 21 22 23 24 25 26 27 28 29 30
```

```
Escalate if # of DLT <=     4  4  4  4  5  5  5  5  6  6  6  6  7
Deescalate if # of DLT >=   7  7  8  8  8  9  9  9 10 10 11 11 11
Eliminate if # of DLT >=    9  9 10 10 11 11 11 12 12 12 13 13 14

Default stopping rule: stop the trial if the lowest dose is eliminated.
```

R> plot(bound)

A few remarks are in order. First, the flowchart generated by `plot()` provides a transparent description of the trial design that can be easily understood by clinicians. We strongly recommend including it in the trial protocol. Second, as described previously, in practice, we should avoid setting the values of `p.saf` and `p.tox` very close to the `target` because the small sample sizes of typical phase I trials prevent us from being able to discriminate the target DLT rate from the rates close to it. The default values $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$ are strongly recommended and generally yield superior operating characteristics.

Lastly, the BOIN design has two built-in stopping rules: (i) stop the trial if the lowest dose is eliminated due to toxicity. In this case, no dose should be selected as the MTD; (ii) stop the trial and select the MTD if the number of patients treated at the current dose reaches `n.earlystop`. The first stopping rule is a safety rule to protect patients from the case in which all doses are overly toxic. The rationale for the second stopping rule is that when the number of patients assigned to a dose is large (i.e., reaches `n.earlystop`), this means that the dose-finding algorithm has approximately converged. Thus, we can stop the trial early and select the MTD to save sample size and reduce the trial duration. The default value `n.earlystop = 100` essentially turns off this type of early stopping rule. Setting `n.earlystop` at a value like 12 can potentially save sample size and finish the trial early. The trade-off is that it may affect the MTD selection percentage and decrease the rate of stopping for safety if the first dose is overly toxic. The value of `n.earlystop` should be calibrated by simulation to obtain desirable operating characteristics. In general, we recommend `n.earlystop = 9` to 18. Our experience is that this stopping rule is particularly useful when there is strong prior knowledge that the first dose is safe, since a major side effect of using the stopping rule is that it decreases the rate of stopping for safety when the first dose is actually overly toxic.

Although the BOIN design has a built-in safety stopping rule (i.e., stopping rule (i) described above), for some applications, investigators may prefer a stricter stopping rule for extra safety when the lowest dose is possibly overly toxic. Setting `extrasafe = TRUE` imposes the following stronger stopping rule:

> Stop the trial if (1) the number of patients treated at the lowest dose ≥ 3, and (2) P(DLT rate of the lowest dose $>$ `target` | data) $>$ `cutoff.eli - offset`.

As a trade-off, the stricter stopping rule will decrease the MTD selection percentage when the lowest dose actually is the true MTD. When using the option `extrasafe = TRUE`, we recommend the default value `offset = 0.05`, but users can calibrate the value of `offset` to obtain desired operating characteristics. In practice, `offset` is rarely greater than 0.2. If we set `extrasafe = TRUE` to turn on the `extrasafe` feature, the output will include the extra stopping boundaries, as follows,

```
R> bound2 <- get.boundary(target = 0.3, ncohort = 10, cohortsize = 3,
+    extrasafe = TRUE)
R> summary(bound2)

......

In addition to the default stopping rule (i.e., stop the trial if the
lowest dose is eliminated), the following more strict stopping safety
rule will be used for extra safety:
stop the trial if
(1) the number of patients treated at the lowest dose >= 3, AND
(2) Pr(the DLT rate of the lowest dose > 0.3 | data) > 0.9,
which corresponds to the following stopping boundaries:

The number of patients treated at the lowest dose    1  2 3 4 5 6 7 8 9
Stop the trial if # of DLT >=                        NA NA 2 3 3 4 4 4 5

The number of patients treated at the lowest dose   10 11 12 13 14 15 16
Stop the trial if # of DLT >=                        5  6  6  6  7  7  8

The number of patients treated at the lowest dose   17 18 19 20 21 22 23
Stop the trial if # of DLT >=                        8  8  9  9  9 10 10

The number of patients treated at the lowest dose   24 25 26 27 28 29 30
Stop the trial if # of DLT >=                       10 11 11 12 12 12 13
```

Step 2. Simulate operating characteristics

For protocol preparation, it is often useful to obtain the operating characteristics of the design. The function `get.oc()` can be used for this purpose.

```
get.oc(target, p.true, ncohort, cohortsize, n.earlystop = 100,
  startdose = 1, titration = FALSE, p.saf = 0.6 * target,
  p.tox = 1.4 * target, cutoff.eli = 0.95, extrasafe = FALSE, offset = 0.05,
  ntrial = 1000, seed = 6)
```

This function shares the same set of arguments as the function `get.boundary()` described previously, with five additional simulation related arguments:

- `p.true`: The true toxicity probabilities of the investigational dose levels.

- `startdose`: The starting dose level for treating the first cohort of patients. The default value is `startdose = 1`, i.e., starting from the lowest dose.

- `titration`: If `titration = TRUE`, dose titration is performed to accelerate dose escalation at the beginning of the trial, where patients are treated one by one (i.e., cohort size = 1), starting from `startdose`. If no DLT is observed, escalate the dose; otherwise switch to the specified cohort size = `cohortsize`.

- `ntrial`: The number of trials to be simulated.

- `seed`: Set a seed for the random number generator.

Dose titration, by setting `titration = TRUE`, accelerates dose escalation and improves the selection percentage of the MTD when the number of patients is small (e.g., `ncohort`/(number of doses) < 1.5) and low doses are expected to be safe. The tradeoff is that if the low doses are toxic, the dose titration increases the risk of overdosing patients.

Using the same setting as above and assuming that the true toxicity scenario is `p.true` = $(0.05, 0.15, 0.30, 0.45, 0.6)$, here we show how to obtain the operating characteristics based on 1000 simulated trials.

```
R> oc.single <- get.oc(target = 0.3, p.true = c(0.05, 0.15, 0.3, 0.45, 0.6),
+    ncohort = 10, cohortsize = 3, ntrial = 1000)
R> summary(oc.single)

selection percentage at each dose level (%):
1.1  23.4  54.2  20.2  1.1
average number of patients treated at each dose level:
4.2  9.3  11.0  4.9  0.7
average number of toxicity observed at each dose level:
0.2  1.4  3.3  2.2  0.4
average number of toxicities: 7.4
average number of patients: 30.0
percentage of early stopping due to toxicity: 0.0 %
risk of overdosing (>60% of patients treated above the MTD): 2.9 %
risk of overdosing (>80% of patients treated above the MTD): 0.0 %
```

We can use the `plot()` function to plot the operating characteristics, including the selection percentage (Figure 4), average number of patients treated, and average number of toxicities observed.

```
R> plot(oc.single)
```

Step 3. Select the MTD

When the trial is completed, based on the observed data, we can select the MTD using the function `select.mtd()`.

```
select.mtd(target, npts, ntox, cutoff.eli = 0.95, extrasafe = FALSE,
  offset = 0.05, verbose = TRUE)
```

This function has the following arguments:

- `npts`: The vector recording the total number of patients treated at each dose level.

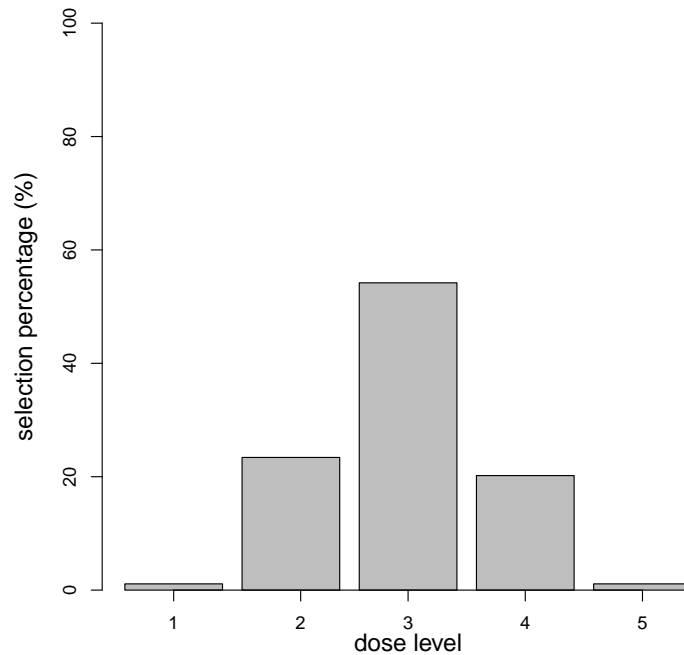- `ntox`: The vector recording the number of patients who experienced toxicity at each dose level.

Figure 4: Selection percentage of the doses in the simulation example using `get.oc()`.

Arguments `cutoff.eli`, `extrasafe` and `offset` are the same as (and should be consistent with) those in functions `get.boundary()` and `get.oc()`. Assume that when the trial is completed, the number of patients treated at the five doses is $n = (3, 3, 15, 9, 0)$ and the corresponding number of patients who experienced toxicity is $y = (0, 0, 4, 4, 0)$.

```
R> n <- c(3, 3, 15, 9, 0)
R> y <- c(0, 0, 4, 4, 0)
R> sel.single <- select.mtd(target = 0.3, ntox = y, npts = n)
R> summary(sel.single)


The MTD is dose level  3

Dose      Posterior DLT            95%
Level       Estimate        Credible Interval    Pr(toxicity>0.3|data)
  1           0.02           (0.00,0.20)                0.01
  2           0.02           (0.00,0.20)                0.01
  3           0.27           (0.09,0.51)                0.36
  4           0.45           (0.16,0.75)                0.81
  5           ----           (----,----)                ----
NOTE: no estimate is provided for the doses at which no patient was treated.
```

We can plot the estimates of DLT rates (Figure 5) by calling the `plot()` function.
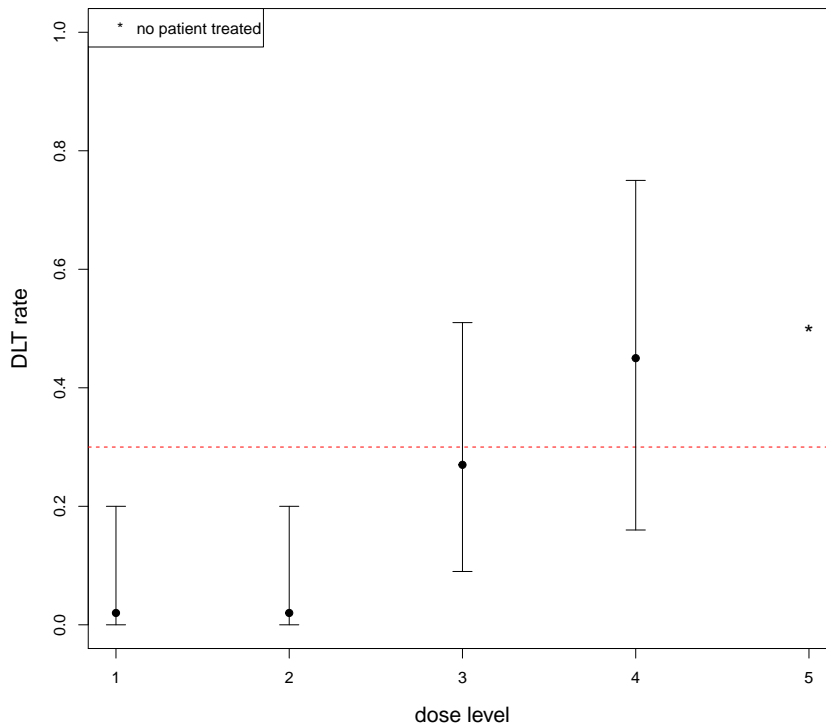
```
R> plot(sel.single)
```

Figure 5: Estimated DLT rate with a 95% confidence interval at each dose for the example using `select.mtd()`. Dotted line indicates the target DLT rate.

3.2. Drug-combination trials aiming to find a single MTD

Designing a drug-combination trial follows the similar three steps as those for single-agent trials. This involves four functions `get.boundary()`, `next.comb()`, `get.oc.comb()` and `selet.mtd.comb()`.

Step 1. Determine dose escalation and de-escalation

The BOIN drug-combination design uses the same dose escalation and de-escalation boundaries as the single-agent designs, which can be obtained using `get.boundary()` as described previously. For drug combination trials, the challenge is that when the number of observed DLTs crosses the dose escalation and de-escalation boundaries, we need to determine which drug (i.e., drug A or drug B) to be escalated or de-escalated. The function `next.comb()` is used to make such a decision.

```
next.comb(target, npts, ntox, dose.curr, n.earlystop = 100,
  p.saf = 0.6 * target, p.tox = 1.4 * target, cutoff.eli = 0.95,
  extrasafe = FALSE, offset = 0.05)
```

This function takes the data from patients who have been enrolled into the trial as the input, and outputs the dose combination for treating the next cohort of new patients. It shares a similar set of arguments with the function `get.boundary()` described previously, with three additional arguments:

- `npts`: The matrix recording the number of patients treated at each dose combination.

- `ntox`: The matrix recording the number of patients who experienced toxicity at each dose combination.

- `dose.curr`: The current dose combination, i.e., the dose combination that was used to treat the most recently enrolled cohort of patients.

Suppose that we conduct a 3×4 drug-combination trial with 3 dose levels of drug A and 4 dose levels of drug B, aiming to find a MTD that has the target DLT rate of 0.25. The maximum sample size is 48 patients, and patients are treated in cohort sizes of 3. To be consistent with the matrix notation of R, we now use (j, k), rather than $A_j B_k$, to denote the combination of the jth dose level of drug A and the kth dose level of drug B. The trial can be conducted as follows. We start the trial by treating the first cohort of 3 patients at the lowest dose $(1, 1)$. Assuming that none of the patients experienced DLT, the data from the first cohort of patients are given by

$$
n = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad\qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

where n records the number of patients treated at each dose combination, and y records the number of patients who experienced toxicity at each dose combination. In matrices y and n, entry (j, k) records the data associated with combination (j, k). To determine the dose for the second cohort of patients, we call the function `next.comb()`:

```
R> set.seed(4)
R> n <- matrix(c(3, 0, 0, 0,  0, 0, 0, 0,  0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0,  0, 0, 0, 0,  0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
R> nxt.comb <- next.comb(target = 0.25, npts = n, ntox = y,
+    dose.curr = c(1, 1))
R> summary(nxt.comb)
```

```
The recommended dose combination for the next cohort of patients is (1, 2).
```

Therefore, we escalate the dose and treat the second cohort of patients at dose combination $(1, 2)$. Suppose that one patient in the second cohort experiences DLT, the data matrices then become

$$
n = \begin{pmatrix} 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad\qquad y = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
$$

To determine the dose for the third cohort of patients, we again call `next.comb()` with updated `y`, `n` and `dose.curr`, as follows:

```
R> n <- matrix(c(3, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
R> y <- matrix(c(0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
```

```
R> nxt.comb2 <- next.comb(target = 0.25, npts = n, ntox = y,
+     dose.curr = c(1, 2))
R> summary(nxt.comb2)
```

```
The recommended dose combination for the next cohort of patients is (1, 1).
```

Therefore, we should de-escalate the dose and treat the third cohort of patients at dose (1, 1). We repeat this procedure until the maximum sample size is reached.

Step 2. Simulate operating characteristics

The function `get.oc.comb()` can be used to obtain the operating characteristics of the BOIN drug-combination design.

```
get.oc.comb(target, p.true, ncohort, cohortsize, n.earlystop = NULL,
  startdose = c(1, 1), titration = FALSE, p.saf = 0.6 * target,
  p.tox = 1.4 * target, cutoff.eli = 0.95, extrasafe = FALSE, offset = 0.05,
  ntrial = 1000, mtd.contour = FALSE, seed = 6)
```

This function shares the same set of arguments as the function `get.oc()` described previously (for single-agent trials), except that `p.true` is now a matrix (rather than a vector), `startdose` is a vector of length 2 (rather than a scalar), and an additional arguments `mtd.contour` is used to indicate whether we are interested in finding the MTD or the MTD contour. To find a single MTD, we should set `mtd.contour = FALSE`. Again, performing dose titration, by setting `titration = TRUE`, is useful to accelerate dose escalation and improve the selection percentage of the MTD when the number of patients is small (e.g., `ncohort`/(number of doses) < 1.5) and low doses are expected to be safe. However, when the low doses are toxic, dose titration may increase the risk of overdosing patients.

Consider a 3×4 combination trial with the true toxicity probabilities

$$\texttt{p.true} = \left(\begin{array}{cccc} 0.02 & 0.04 & 0.08 & 0.14 \\ 0.08 & 0.25 & 0.42 & 0.48 \\ 0.25 & 0.45 & 0.50 & 0.60 \end{array} \right),$$

and the target DLT rate of 0.25. Here, we show how to obtain the operating characteristics based on 1000 simulated trials.

```
R> p.true <- matrix(c(0.02, 0.04, 0.08, 0.14,  0.08, 0.25, 0.42, 0.48,  0.25,
+     0.45, 0.50, 0.60), ncol = 4, byrow = TRUE)
R> oc.comb <- get.oc.comb(target = 0.25, p.true = p.true, ncohort = 16,
+     cohortsize = 3, ntrial = 1000)
R> summary(oc.comb)
```

```
true DLT rate of dose combinations:
 0.02   0.04   0.08   0.14
 0.08   0.25   0.42   0.48
 0.25   0.45   0.50   0.60
```
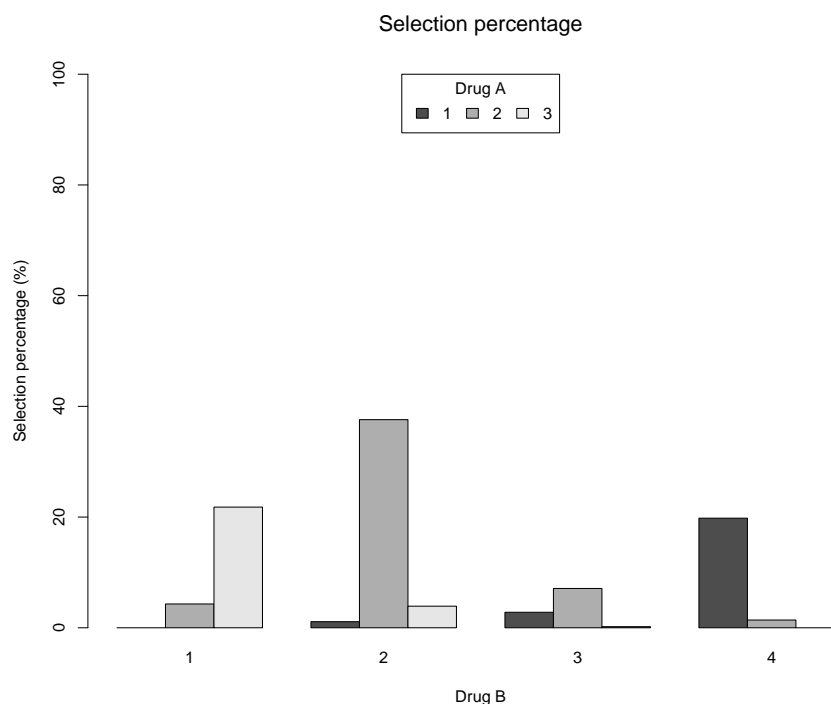
Figure 6: Selection percentage of the dose combinations in the example using `get.oc.comb()`.

```
selection percentage at each dose combination (%):
 0.00   1.10   2.80   19.80
 4.30   37.60  7.10   1.40
21.80   3.90   0.20   0.00


average number of patients treated at each dose combination:
 4.06   3.73   3.22   4.18
 6.08   10.04  4.19   1.90
 5.85   3.65   0.80   0.30


average number of toxicity observed at each dose combination:
 0.10   0.15   0.25   0.58
 0.47   2.57   1.80   0.89
 1.48   1.65   0.39   0.18


average number of toxicities: 10.5
average number of patients: 48.0
selection percentage of MTD: 59.4%
percentage of patients treated at MTD: 33.1%
```

We can use the `plot()` function to plot the figure illustrating dose escalation/de-escalation and operating characteristics such as selection percentage (Figure 6), average number of patient treated, and average number of toxicity observed at each dose combination.

```
R> plot(oc.comb)
```

Step 3. Select an MTD

When the trial is completed, based on the observed data, we can select the MTD using the function `select.mtd.comb()`.

```
select.mtd.comb(target, npts, ntox, cutoff.eli = 0.95,
  extrasafe = FALSE, offset = 0.05, mtd.contour = FALSE)
```

Since we are interested in finding a single MTD, we set `mtd.contour = FALSE`. The other arguments are the same as those in `get.boundary()` described previously for single-agent trials. Assume that the number of patients treated at each dose combination and the corresponding number of patients who experience toxicity at each dose combination are

$$
n = \begin{pmatrix} 6 & 3 & 0 & 0 \\ 6 & 24 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad\qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 5 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
$$

```
R> n <- matrix(c(6, 3, 0, 0, 6, 24, 9, 0, 0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 1, 5, 4, 0, 0, 0, 0, 0), ncol = 4,
+    byrow = TRUE)
R> sel.mtd.comb <- select.mtd.comb(target = 0.25, npts = n, ntox = y)
R> summary(sel.mtd.comb)

The MTD is dose combination (2, 2)

Isotonic estimates of the DLT rates are:
 0.01   0.02     NA     NA
 0.17   0.21   0.45     NA
   NA     NA     NA     NA

NOTE: no estimate is provided for the doses at which no patient was treated.
```

The result is that dose combination (2, 2) is selected as the MTD. Note that no estimate is provided for dose combinations that have never been used to treat patients, e.g., (1, 3) or (3, 3). The `plot()` function can be used to show the DLT estimates.

```
R> plot(sel.mtd.comb)
```

3.3. Drug-combination trials aiming to find the MTD contour

Designing a drug-combination trial for finding the MTD contour also follows the similar three steps as those for single-agent trials. This involves four functions `get.boundary()`, `next.subtrial()`, `get.oc.comb()` and `selet.mtd.comb()`.

Step 1. Obtain dose escalation and de-escalation boundaries and determine subtrials

The BOIN-waterfall design is used to find the MTD contour, which divides the task of finding the MTD contour into a sequence of subtrials. Each subtrial is conducted using the BOIN

single-agent design, whose dose escalation and de-escalation boundaries can be obtained using `get.boundary()` as described previously. The function `next.subtrial()` is used to determine the next subtrial based on the observed data.

```
next.subtrial(target, npts, ntox, p.saf = 0.6 * target, p.tox = 1.4 * target,
   cutoff.eli = 0.95, extrasafe = FALSE, offset = 0.05)
```

As described in Figure 3 and Zhang and Yuan (2016), the first subtrial is prespecified and starts from the lowest dose (1, 1). When the first subtrial is completed, we call the function `next.subtrial()` to determine the starting dose combination and the dose-searching space for next subtrial. The function `next.subtrial()` shares a similar set of arguments with the function `get.boundary()` described previously, with two additional arguments:

- `npts`: The matrix recording the number of patients treated at each dose combination.

- `ntox`: The matrix recording the number of patients who experienced toxicity at each dose combination.

Assume that when the current subtrial is completed, the number of patients treated at each dose combination (i.e., matrix n) and the corresponding number of patients who experienced toxicity at each dose combination (i.e., matrix y) are

$$
n = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 9 & 12 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \end{pmatrix}.
$$

We call `next.subtrial()` to determine the dose space for the next subtrial and its starting dose, as follows,

```
R> n <- matrix(c(6, 0, 0, 0, 6, 0, 0, 0, 9, 12, 0, 0), ncol = 4,
+    byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 1, 0, 0, 0, 2, 3, 0, 0), ncol = 4,
+    byrow = TRUE)
R> nxt.trial <- next.subtrial(target = 0.3, npts = n, ntox = y)
R> summary(nxt.trial)

Next subtrial includes doses:
(2, 2), (2, 3), (2, 4)


The starting dose for this subtrial is:
(2, 3)
```

That is, the next subtrial consists of three doses $\{(2, 2), (2, 3), (2, 4)\}$. As these doses are monotonically ordered, the BOIN design described in Section 3.1 for single-agent trial can be directly used here to run this subtrial with the starting dose (2, 3), i.e., the second dose level.

Step 2. Simulate operating characteristics

The function `get.oc.comb()` can be used to obtain the operating characteristics of the BOIN-waterfall design. This function shares the same set of arguments as the function `get.oc()`

described in Section 3.1 (for single-agent trials), except that `p.true` is now a matrix (rather than a vector), `startdose` is a vector of length 2 (rather than a scalar), and an additional argument `mtd.contour` is used to indicate whether we are interested in finding an MTD or the MTD contour.

Consider a 3×5 drug combination trial with true toxicity probabilities as follows

$$\texttt{p.true} = \left(\begin{array}{ccccc} 0.01 & 0.03 & 0.10 & 0.20 & 0.30 \\ 0.03 & 0.05 & 0.15 & 0.30 & 0.60 \\ 0.08 & 0.10 & 0.30 & 0.60 & 0.75 \end{array} \right),$$

and the target DLT rate is 0.30. We show below how to obtain the operating characteristics based on 1000 simulated trials.

```
R> p.true <- matrix(c(0.01, 0.03, 0.10, 0.20, 0.30,  0.03, 0.05, 0.15, 0.30,
+    0.60, 0.08, 0.10, 0.30, 0.60, 0.75), nrow = 3, ncol = 5, byrow = TRUE)
R> oc.comb2 <- get.oc.comb(target = 0.3, p.true, ncohort = c(10, 5, 5),
+    cohortsize = 3, n.earlystop = 12, startdose = c(1, 1), ntrial = 1000,
+    mtd.contour = TRUE)
R> summary(oc.comb2)


true DLT rate of dose combinations:
 0.01   0.03   0.10   0.20   0.30
 0.03   0.05   0.15   0.30   0.60
 0.08   0.10   0.30   0.60   0.75


selection percentage at each dose combination (%):
 0.00    0.00    1.80   26.40   71.80
 0.00    0.60   22.30   69.60    7.50
 3.00   21.30   68.60    6.90    0.00


average number of patients treated at each dose combination:
 3.10    0.00    0.45    3.46    9.67
 3.45    0.28    3.09    8.20    3.23
 4.11    6.05    8.82    3.12    0.11


average number of toxicity observed at each dose combination:
 0.03    0.00    0.05    0.66    2.89
 0.11    0.01    0.46    2.42    1.92
 0.35    0.56    2.67    1.85    0.09


average number of toxicities: 14.1
average number of patients: 57.1
percentage of patients treated at MTD contour: 46.7%
percentage of patients treated above MTD contour: 11.3%
percentage of patients treated below MTD contour: 42%
percentage of correct selection of the MTD contour: 36.3%
```
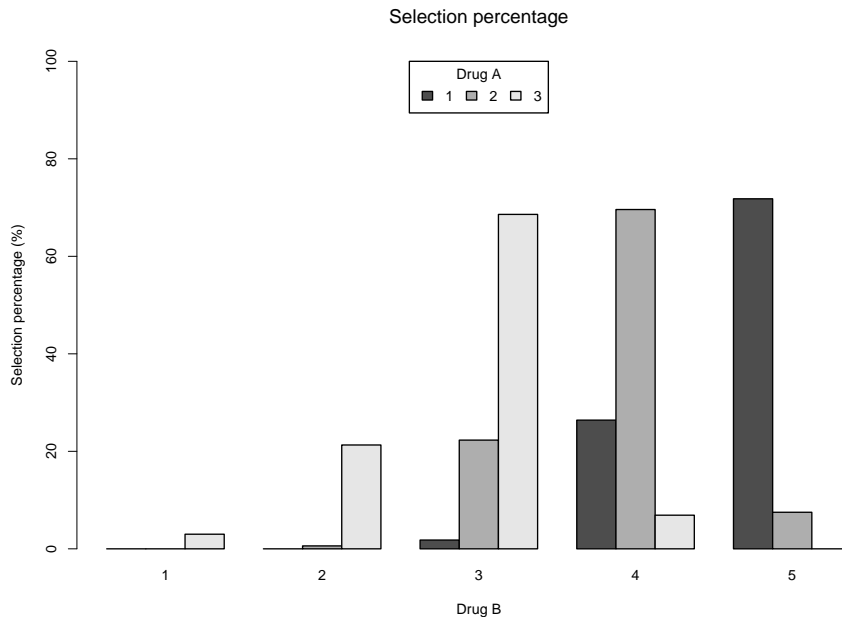
Figure 7: Selection percentage of the dose combinations in the example using `get.oc.comb()`.

We can use the `plot()` function to plot the figure illustrating the waterfall design (similar to Figure 3) and operating characteristics such as selection percentage (Figure 7), average number of patient treated, and average number of toxicity observed at each dose combination.

```
R> plot(oc.comb2)
```

Step 3. Select the MTD contour

When the trial is completed, based on the observed data, we can select the MTD contour using the function `select.mtd.comb()`, described previously. Since we are interested in finding the MTD contour, we should set argument `mtd.contour = TRUE`.

Assume that the number of patients treated at each dose combination (i.e., matrix n) and the corresponding number of patients who experienced toxicity at each dose combination (i.e., matrix y) are

$$
n = \begin{pmatrix} 6 & 9 & 24 & 0 \\ 6 & 24 & 9 & 0 \\ 12 & 18 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 1 & 5 & 0 \\ 1 & 5 & 4 & 0 \\ 1 & 5 & 0 & 0 \end{pmatrix}.
$$

```
R> n <- matrix(c(6, 9, 24, 0, 6, 24, 9, 0, 12, 18, 0, 0), ncol = 4,
+    byrow = TRUE)
R> y <- matrix(c(0, 1, 5, 0, 1, 5, 4, 0, 1, 5, 0, 0), ncol = 4,
+    byrow = TRUE)
R> sel.comb2 <- select.mtd.comb(target = 0.3, npts = n, ntox = y,
+    mtd.contour = TRUE)
R> summary(sel.comb2)
```

	Number of patients treated									
	3	6	9	12	15	18	21	24	27	30
Escalate if # of DLT $\leq$	0	1	2	2	3	4	4	5	6	7
De-escalate if # of DLT $\geq$	2	3	4	5	6	7	8	9	10	11
Eliminate if # of DLT $\geq$	3	4	5	7	8	9	10	11	12	14

Table 3: Dose escalation and de-escalation rule for the BOIN design.

```
The MTD contour includes dose combinations (1, 3) (2, 2) (3, 2)

Isotonic estimates of the DLT rates are:
0.01   0.12   0.21      NA
0.12   0.21   0.45      NA
0.12   0.28      NA     NA

NOTE: no estimate is provided for the doses at which no patient was treated.
```

The result is that dose combinations $(1, 3)$, $(2, 2)$ and $(3, 2)$ are selected as the MTD contour. Note that no estimate is provided for dose combinations that have never been used to treat patients, e.g., $(1, 4)$ or $(3, 3)$. We can use `plot()` function to show the estimates of the DLT rates.

```
R> plot(sel.comb2)
```

4. Trial examples

4.1. Single-agent phase I trial

Consider a single-agent phase I trial with 5 dose levels, in which the objective is to find the MTD with a target DLT rate of 0.3. The maximum sample size is 30 patients, treated in cohort sizes of 3. To design and conduct this trial, we first ran the following commands:

```
R> bound <- get.boundary(target = 0.3, ncohort = 10, cohortsize = 3)
R> summary(bound)
R> plot(bound)
```

This yields the dose escalation and de-escalation boundaries as shown in Table 3, and a flowchart of the trial design similar to Figure 1. The trial started by treating the first cohort of 3 patients at dose level 1 and none of the patients had dose limiting toxicity (DLT). According to the dose escalation and de-escalation rule provided in Table 3, we escalated the dose to level 2 to treat the second cohort of 3 patients, none of whom experienced DLT. Thus, we escalated the dose to level 3 and treated the third cohort of patients, 2 of whom experienced DLT. Based on Table 3, we de-escalated the dose back to level 2 and treated the fourth cohort of patients, one of whom experienced DLT. We then escalated the dose to level 3 and treated the fifth cohort of patients, none of whom experienced DLT. Therefore,
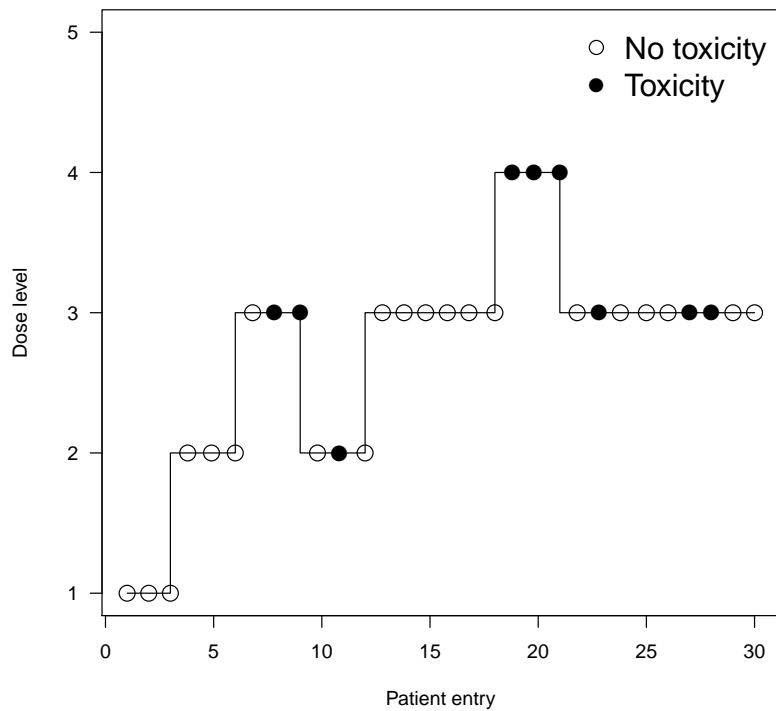
Figure 8: Illustration of the single-agent phase I trial using the BOIN design.

the sixth cohort was also treated at dose level 3. Figure 8 shows the dose assignment for all 30 patients. At the end of the trial, the number of patients and the number of DLTs at the 5 doses were `n = c(3, 6, 18, 3, 0)` and `y = c(0, 1, 5, 3, 0)`. To select the MTD, we ran the following commands:

```
R> n <- c(3, 6, 18, 3, 0)
R> y <- c(0, 1, 5, 3, 0)
R> sel.single <- select.mtd(target = 0.3, ntox = y, npts = n)
R> summary(sel.single)

The MTD is dose level  3

Dose      Posterior DLT              95%
Level       Estimate        Credible Interval    Pr(toxicity>0.3|data)
  1           0.02           (0.00,0.20)                0.01
  2           0.17           (0.01,0.53)                0.18
  3           0.28           (0.10,0.50)                0.39
  4           0.98           (0.80,1.00)                1.00
  5           ----           (----,----)                ----

NOTE: no estimate is provided for the doses at which no patient was treated.
```

which recommended the dose level 3 as the MTD, with the estimated DLT rate = 0.28 and the 95% CI = (0.10, 0.50). And the illustration of the single-agent phase I trial is shown in Figure 8.

4.2. Drug-combination trial to find a single MTD

Consider a drug-combination trial that combines 3 doses of drug A and 5 doses of drug B. The objective is to find an MTD with a target DLT rate of 0.3. The maximum sample size is 30 patients, treated in cohort size of 3. The trial started by treating the first cohort of 3 patients at the lowest dose combination $(1, 1)$, at which no DLT was observed. The observed data were

$$n = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where n records the number of patients treated at each dose combination, and y records the number of patients who experienced DLT at each dose combination. In matrices y and n, entry (j, k) records the data associated with dose combination (j, k). To determine the dose for the second cohort of patients, we used the following code:

```
R> set.seed(4)
R> n <- matrix(c(3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> nxt.comb <- next.comb(target = 0.3, npts = n, ntox = y,
+    dose.curr = c(1, 1))
R> summary(nxt.comb)
```

```
The recommended dose combination for the next cohort of patients is (1, 2).
```

which recommended to escalate the dose to combination $(1, 2)$. Therefore, we treated the second cohort of patients at dose combination $(1, 2)$. In the second cohort, no patient experienced DLT, so the updated data matrices became

$$n = \begin{pmatrix} 3 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To determine the dose for the third cohort of patients, we again called `next.comb()` with updated y, n and `dose.curr`.

```
R> n <- matrix(c(3, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> nxt.comb2 <- next.comb(target = 0.3, npts = n, ntox = y,
+    dose.curr = c(1, 2))
R> summary(nxt.comb2)
```

```
The recommended dose combination for the next cohort of patients is
(2, 2).
```
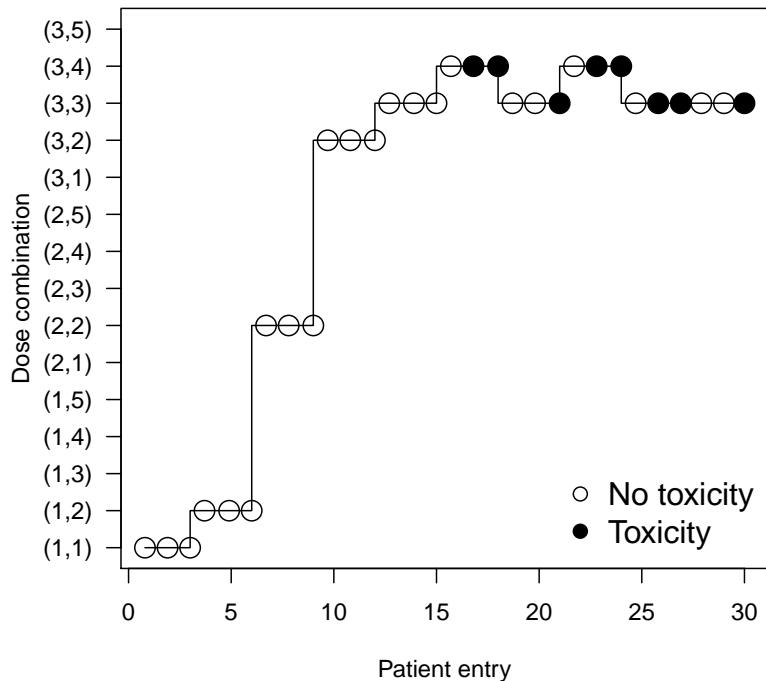
Figure 9: Illustration of the dual-agent combination design of finding the MTD for a 3×5 combination trial with a cohort of 3. Open circles indicate patients without toxicity, and solid circles denote patients with toxicity.

The function recommended escalating the dose to $(2, 2)$ for treating the third cohort of patients. We repeated this procedure until the maximum sample size was reached. Figure 9 shows the dose assignments for all 30 patients. For example, at dose combination $(3, 4)$ when completing the eighth cohort, there were two DLTs, based on the accumulating toxic information on this dose combination level, the function recommended de-escalating the dose to combination $(3, 3)$. When the trial was completed, the number of patients treated at each dose combination and the corresponding number of patients who experienced toxicity at each dose combination were

$$n = \begin{pmatrix} 3 & 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 3 & 12 & 6 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 4 & 0 \end{pmatrix}.$$

We called the function `select.mtd.comb()`, which recommended dose combination $(3, 3)$ as the MTD.

```
R> n <- matrix(c(3, 3, 0, 0, 0, 0, 3, 0, 0, 0, 0, 3, 12, 6, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 4, 0), nrow = 3,
+    ncol = 5, byrow = TRUE)
R> sel.comb <- select.mtd.comb(target = 0.3, npts = n, ntox = y)
R> summary(sel.comb)

The MTD is dose combination (3, 3)
```

```
Isotonic estimates of toxicity probabilities for combinations are
0.02    0.02    NA      NA      NA
NA      0.03    NA      NA      NA
NA      0.03    0.34    0.66    NA

NOTE: no estimate is provided for the doses at which no patient was treated.
```

4.3. Drug-combination trial to find the MTD contour

Consider a drug-combination trial that combines 3 doses of drug A and 5 doses of drug B. The objective is to find the MTD contour (multiple MTDs) with a target DLT rate of 0.25. The trial started with the first subtrial, which consisted of 7 ordered combinations $\{A_1B_1 \to A_2B_1 \to A_3B_1 \to A_3B_2 \to A_3B_3 \to A_3B_4 \to A_3B_5\}$. This subtrial was conducted using the BOIN design in a way similar to those as described in Example 4.1. The starting dose for this subtrial was A_1B_1 and `n.earlystop` was set as 12 such that the subtrial stopped when the number of patients treated at that dose reached 12. The first subtrial resulted in the following data:

$$
n = \begin{pmatrix} 6 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 \\ 9 & 12 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 0 \end{pmatrix}.
$$

Based on the data, we called the function `next.subtrial()` to obtain the doses for the next subtrial.

```
R> n <- matrix(c(6, 0, 0, 0, 0, 6, 0, 0, 0, 0, 9, 12, 0, 0, 0), ncol = 5,
+    byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 3, 0, 0, 0), ncol = 5,
+    byrow = TRUE)
R> nxt.trial <- next.subtrial(target = 0.25, npts = n, ntox = y)
R> summary(nxt.trial)

Next subtrial includes doses:
(2, 2), (2, 3), (2, 4), (2, 5)

The starting dose for this subtrial is:
(2, 3)
```

Therefore, we conducted the second subtrial with doses $\{A_2B_2 \to A_2B_3 \to A_2B_4 \to A_2B_5\}$ using the BOIN design with the starting dose A_2B_3. After the second subtrial was completed, the observed data were

$$
n = \begin{pmatrix} 6 & 0 & 0 & 0 & 0 \\ 6 & 0 & 3 & 12 & 0 \\ 9 & 12 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 \\ 1 & 3 & 0 & 0 & 0 \end{pmatrix}.
$$

We called `next.subtrial()` again using the above updated data to obtain the doses for the third subtrial.

```
R> n <- matrix(c(6, 0, 0, 0, 0, 6, 0, 3, 12, 0, 9, 12, 0, 0, 0), ncol = 5,
+     byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 0, 0, 1, 0, 0, 3, 0, 1, 3, 0, 0, 0), ncol = 5,
+     byrow = TRUE)
R> nxt.trial2 <- next.subtrial(target = 0.25, npts = n, ntox = y)
R> summary(nxt.trial2)

Next subtrial includes doses:
(1, 2), (1, 3), (1, 4), (1, 5)

The starting dose for this subtrial is:
(1, 5)
```

The third subtrial included doses $\{A_1B_2 \to A_1B_3 \to A_1B_4 \to A_1B_5\}$, with starting dose A_1B_5. After completing the third subtrial using the BOIN, the trial was completed and resulted in the following final data:

$$n = \begin{pmatrix} 6 & 0 & 0 & 6 & 12 \\ 6 & 0 & 3 & 12 & 0 \\ 9 & 12 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & 3 & 0 \\ 1 & 3 & 0 & 0 & 0 \end{pmatrix}.$$

Based on the final data, we ran `select.mtd.comb()` to select the MTD contour:

```
R> n <- matrix(c(6, 0, 0, 6, 12, 6, 0, 3, 12, 0, 9, 12, 0, 0, 0), ncol = 5,
+     byrow = TRUE)
R> y <- matrix(c(0, 0, 0, 1, 3, 1, 0, 0, 3, 0, 1, 3, 0, 0, 0), ncol = 5,
+     byrow = TRUE)
R> sel.comb2 <- select.mtd.comb(target = 0.25, npts = n, ntox = y,
+     mtd.contour = TRUE)
R> summary(sel.comb2)

The MTD contour includes dose combinations  (1, 5) (2, 4) (3, 2)

Isotonic estimates of toxicity probabilities for combinations are
 0.01     NA     NA   0.17   0.25
 0.12     NA   0.12   0.25     NA
 0.12   0.25     NA     NA     NA

NOTE: no estimate is provided for the doses at which no patient was treated.
```

Thus, we selected dose combinations A_1B_5, A_2B_4 and A_3B_2 as the MTD contour, which had the estimated DLT rate of 0.25.

5. Discussion

As a model-assisted design, the BOIN design combines the simplicity of the algorithm-based design and the superior performance of the model-based design, providing an attractive approach to designing phase I clinical trials. The R package **BOIN** provides an easy-to-use and

well-documented tool to implement the BOIN designs for single-agent and drug-combination phase I trials. The package allows users to generate the dose escalation and de-escalation boundaries, plot the flowchart, and simulate the operating characteristics of the design for trial protocol preparation. It also provides functions to allow users to make the real-time decision of dose assignment and MTD selection for conducting a real trial. By facilitating the use of novel adaptive designs, the **BOIN** package has potential to improve the efficiency of phase I trials and accelerate drug development. Recently, the BOIN design has been extended to account for toxicity grade, continuous, and semi-continuous toxicity endpoints (Mo, Yuan, Xu, Mandrekar, and Yin 2018), late-onset toxicity (Yuan, Lin, Li, Nie, and Warren 2018), and phase I-II trials (Lin and Yin 2017b; Takeda, Taguri, and Morita 2018). We plan to incorporate these new developments into the future version of the **BOIN** package.

Acknowledgments

References

Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD (1972). *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression.* John Wiley & Sons, New York.

Blanchard MS (2016). **TEQR**: *Target Equivalence Range Design.* R package version 6.0-0, URL `https://CRAN.R-project.org/package=TEQR`.

Bornkamp B (2019). **DoseFinding**: *Planning and Analyzing Dose Finding Experiments.* R package version 0.9-17, URL `https://CRAN.R-project.org/package=DoseFinding`.

Bornkamp B, Pinheiro J, Bretz F (2009). "**MCPMod**: An R Package for the Design and Analysis of Dose-Finding Studies." *Journal of Statistical Software*, **29**(7), 1–23. `doi:10.18637/jss.v029.i07`.

Clertant M, O'Quigley J (2017). "Semiparametric Dose Finding Methods." *Journal of the Royal Statistical Society B*, **79**(5), 1487–1508. `doi:10.1111/rssb.12229`.

Ivanova A, Flournoy N, Chung Y (2007). "Cumulative Cohort Design for Dose-Finding." *Journal of Statistical Planning and Inference*, **137**(7), 2316–2327. `doi:10.1016/j.jspi.2006.07.009`.

Le Tourneau C, Lee JJ, Siu LL (2009). "Dose Escalation Methods in Phase I Cancer Clinical Trials." *Journal of the National Cancer Institute*, **101**(10), 708–720. `doi:10.1093/jnci/djp079`.

Lin R, Yin G (2017a). "Bayesian Optimal Interval Design for Dose Finding in Drug-Combination Trials." *Statistical Methods in Medical Research*, **26**(5), 2155–2167. `doi:10.1177/0962280215594494`.

Lin R, Yin G (2017b). "STEIN: A Simple Toxicity and Efficacy Interval Design for Seamless Phase I/II Clinical Trials." *Statistics in Medicine*, **36**(26), 4106–4120. doi:10.1002/sim.7428.

Liu S, Yuan Y (2015). "Bayesian Optimal Interval Designs for Phase I Clinical Trials." *Journal of the Royal Statistical Society C*, **64**(3), 507–523. doi:10.1111/rssc.12089.

Mo R, Yuan Y, Xu J, Mandrekar SJ, Yin JY (2018). "gBOIN: A Unified Model-Assisted Phase I Trial Design Accounting for Toxicity Grades, Binary or Continuous Endpoint." *Journal of the Royal Statistical Society C*, **68**(2), 289–308. doi:10.1111/rssc.12263.

O'Quigley J, Pepe M, Fisher L (1990). "Continual Reassessment Method: A Practical Design for Phase 1 Clinical Trials in Cancer." *Biometrics*, **46**(1), 33–48. doi:10.2307/2531628.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Storer BE (1989). "Design and Analysis of Phase I Clinical Trials." *Biometrics*, **45**(3), 925–937. doi:10.2307/2531693.

Takeda K, Taguri M, Morita S (2018). "BOIN-ET: Bayesian Optimal Interval Design for Dose Finding Based on Both Efficacy and Toxicity Outcomes." *Pharmaceutical Statistics*, **17**(4), 383–395. doi:10.1002/pst.1864.

Yan F, Mandrekar JS, Yuan Y (2017). "Keyboard: A Novel Bayesian Toxicity Probability Interval Design for Phase I Clinical Trials." *Clinical Cancer Research*, **23**(15), 3994–4003. doi:10.1158/1078-0432.ccr-17-0220.

Yuan Y, Hess KR, Hilsenbeck SG, Gilbert MR (2016). "Bayesian Optimal Interval Design: A Simple and Well-Performing Design for Phase I Oncology Trials." *Clinical Cancer Research*, **22**(17), 4291–4301. doi:10.1158/1078-0432.ccr-16-0592.

Yuan Y, Lin R, Li D, Nie L, Warren KE (2018). "Time-to-Event Bayesian Optimal Interval Design to Accelerate Phase I Trials." *Clinical Cancer Research*, **24**(20), 4921–4930. doi:10.1158/1078-0432.ccr-18-0246.

Yuan Y, Liu S (2020). ***BOIN**: Bayesian Optimal INterval (BOIN) Design for Single-Agent and Drug- Combination Phase I Clinical Trials*. R package version 2.6.9, URL https://CRAN.R-project.org/package=BOIN.

Yuan Y, Yin G (2008). "Sequential Continual Reassessment Method for Two-Dimensional Dose Finding." *Statistics in Medicine*, **27**(27), 5664–5678. doi:10.1002/sim.3372.

Yuan Y, Zhang L (2017). "Designing Early-Phase Drug Combination Trials." In J O'Quigley, A Iasonos, B Bornkamp (eds.), *Handbook of Methods for Designing, Monitoring, and Analyzing Dose Finding Trials*, pp. 109–126. Chapman & Hall/CRC, New York.

Zhang L, Yuan Y (2016). "A Simple Bayesian Design to Identify the Maximum Tolerated Dose Contour for Drug Combination Trials." In *JSM Proceedings, Biopharmaceutical Section*, pp. 4924–4936. American Statistical Association, Alexandria, VA.

Zhou H, Murray T, Pan H, Yuan Y (2018a). "Comparative Review of Model-Assisted Designs for Phase I Clinical Trials." *Statistics in Medicine*, **37**(14), 2208–2222. doi:10.1002/sim.7674.

Zhou H, Yuan Y, Nie L (2018b). "Accuracy, Safety and Reliability of Novel Phase I Trial Designs." *Clinical Cancer Research*, **24**(18), 4357–4364. doi:10.1158/1078-0432.ccr-18-0168.

Affiliation:

Ying Yuan
Department of Biostatistics
Univeristy of Texas M.D.Anderson Cancer Center
1400 Pressler Street, Houston, TX 77030, United States of America
E-mail: yyuan@mdanderson.org
URL: http://odin.mdacc.tmc.edu/~yyuan/