U-Design

Version 1.4

Singe-Agent Dose-Finding Designs with Efficacy-Toxicity Endpoints and Cohort Enrollment

Your trial designs anywhere, anytime

November 2, 2020

U-Design® 1.4
Laiya Consulting Inc. Copyright 2020
# Contents

## Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

1. **Introduction** .................................................. 1
2. **User Interface and Tutorial** ............................... 2
   1.2.1 Overview .................................................. 2
   1.2.2 Simulation Setup ........................................... 4
   1.2.3 Simulation Results ....................................... 16
   1.2.4 Decision Tables ........................................... 22
   1.2.5 OBD Estimation ............................................ 25
3. **Statistical Methods Review** ............................... 28
   1.3.1 The Joint i3+3 (Ji3+3) Design ........................ 28
   1.3.2 The Probability Intervals of Toxicity and Efficacy (PRINTE) Design ............................ 32
   1.3.3 The Toxicity and Efficacy Probability Interval (TEPI) Design ................................. 38
   1.3.4 The Efficacy-Toxicity (EffTox) Trade-Offs-Based Design ...................................... 42
   1.3.5 The Utility-Based Bayesian Optimal Interval (U-BOIN) Design ............................... 46

## Reference

51
1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

1.1 Introduction

Gene therapies and adoptive cell therapies (ACTs), such as the chimeric antigen receptor (CAR) T-cell therapy, have demonstrated promising therapeutic effects in oncology patients. An important and distinct feature of some ACTs is that the probability of response may not increase with dose, which is normally seen for cytotoxic cancer therapeutics. For example, Porter et al. (2011) has shown that increased dose of CAR T-cells does not necessarily lead to higher efficacy. Because of the potential non-monotone relationship between response and dose, traditional phase 1 dose-finding designs searching for the maximum tolerated dose (MTD), like i3+3 (Liu et al., 2020) and mTPI-2 (Guo et al., 2017) designs, are not suitable to ACTs. For example, the best efficacious dose may be lower than the MTD as higher doses may not lead to higher efficacy.

To this end, the U-Design introduces the **Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment** module which consists of five novel statistical designs for gene and cell therapeutics dose-finding trials. The module performs trial simulations allowing head-to-head comparison of multiple designs, so that users may select the best design for their own clinical trials. The included novel designs are Ji3+3 (Lin and Ji, 2020b), PRINTE (Lin and Ji, 2020a), TEPI (Li et al., 2017), EffTox (Thall and Cook, 2004) and UBOIN (Zhou et al., 2019), all of which use joint toxicity and efficacy outcomes as endpoints for dose finding. The goal is to identify the optimal biological dose (OBD) that possesses high efficacy and safety simultaneously. As with all other U-Design modules involving trial simulation, below we provide detailed guidance
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

on setting up simulation for design comparison, and visualising simulation results (operating characteristics). In addition, the decision tables generation and the OBD selection are incorporated in this module so that users may generate the decision tables to guide trial conduct and estimate the OBD after trial completion. All the details are provided next.

1.2 User Interface and Tutorial

1.2.1 Overview

Entering the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment page, users will see four main tabs: Simulation Setup, Simulation Results, Decision table and OBD selection. These four tabs allow user to conduct simulations and visualize/download simulation results, generate decision tables for trial conduct, and select OBD after trial is completed. In the Simulation Setup tab, there are three steps (Figure 1.1): 1) Set trial parameters, 2) Select designs, and 3) Generate scenarios. Users need to complete the current step to get access to the next one. Upon completing steps 1-3, users click the “Launch Simulation” button at the bottom of the page to submit the simulations to the cloud for computation. Users may also click the “Reset” button next to Launch Simulation to clear all settings. After the simulation is launched, the results of simulations will be displayed in the Simulation Results tab. The simulation process can be monitored in real time at the top of the Simulation Results tab. Detailed steps of using this module are elaborated next in §1.2.2-§1.2.5.
1.2. User Interface and Tutorial

1.2.1. Overview

Figure 1.1: Simulation Setup in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

1.2.2 Simulation Setup

In the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module, U-Design provides five designs, Ji3+3, PRINTE, TEPI, Efftox, UBOIN, for simulation. Users can choose up to four designs for head-to-head comparison in the Simulation Setup tab each time. Three steps of simulation set up are needed.

1.2.2.1 Step 1: Set trial parameters

Specify the number of simulations \(n_{\text{sim}}\) and the random seed of simulation \(R_{\text{seed}}\). Specify the target toxicity probability \(p_T\) and minimum acceptable efficacy \(q_E\) for the simulated trials and select a number of doses \(n_{\text{dose}}\) from the dropdown box. Click the “Apply” button to apply the settings. See Figure 1.2. Hover mouse over each trial parameter, and a description will be displayed explaining the meaning of the parameter. The detailed explanation on U-Design interface of the above four input arguments is provided in Table 1.1.

![Figure 1.2: Set trial parameters in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.](image)

Upon clicking the “Apply” button, a table of actual dosage will be displayed. Specify the dosage of each dose level in the table. (Figure 1.3) This is only needed if the EffTox design is selected in Step 2 next. If EffTox is not going to be selected, leave the table unchanged and move to Step 2.

1.2.2.2 Step 2: Select designs

To select a design, click the button with the design’s name on it. Up to four designs may be selected for head-to-head comparison.

Check the “Apply Stopping Rule” box to apply an ad-hoc stopping rule that stops the trial if a maximum number of patients has been enrolled at a single dose. See the detailed rules in §1.3.
1.2. User Interface and Tutorial

1.2.2. Simulation Setup

Figure 1.3: Selecting actual dosage in the Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment module.

Click the “Draw” button to plot a contour map of the utility function. The horizontal axis represents efficacy and the vertical axis represents toxicity. See Figure 1.4.

Click the “Apply” button of all the designs before launching simulations to apply all settings.

Click the “Delete” button to remove the selected designs.

Design parameters can be modified in the input box. Hover mouse over each design parameter, and a description will be displayed explaining the meaning of the parameter. See detailed parameter descriptions in Table 1.2.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & ToxicityEndpoints and Cohort Enrollment

Table 1.1: Input parameters for trials parameters in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_{sim} )</td>
<td>The number of simulated trials</td>
<td>The maximum number of simulated trials allowed is 10,000. Default value is 1,000.</td>
</tr>
<tr>
<td>( R_{seed} )</td>
<td>The random seed of simulation</td>
<td>A random seed is a number used to initialize a pseudorandom number generator in the simulation. Default value is 32432.</td>
</tr>
<tr>
<td>( p_T )</td>
<td>Target toxicity probability</td>
<td>The target toxicity probability of the maximum tolerated dose (MTD). Default value is 0.3.</td>
</tr>
<tr>
<td>( q_E )</td>
<td>Minimum acceptable efficacy</td>
<td>The minimum acceptable efficacy used in the futility rule. A dose is considered not promising if the efficacy rate is unlikely to be larger than ( q_E ). Default value is 0.2.</td>
</tr>
<tr>
<td>( n_{dose} )</td>
<td>Number of doses</td>
<td>The number of doses in the trial.</td>
</tr>
</tbody>
</table>
1.2. User Interface and Tutorial
1.2.2. Simulation Setup

Figure 1.4: Select designs in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
### Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

**Table 1.2:** Input parameters for designs in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{\text{start}}$ (all designs)</td>
<td>Starting dose level</td>
<td>The starting dose level in the simulated trials. Default value is 1.</td>
</tr>
<tr>
<td>$n$ (all designs)</td>
<td>Sample size</td>
<td>The maximum number of patients to be treated in the trial. The upper limit is set at 100 since the number of patients that are enrolled in phase I clinical trial is typically small. Default value is 30.</td>
</tr>
<tr>
<td>$n_{\text{cohort}}$ (all designs)</td>
<td>Cohort size</td>
<td>The number of patients in each cohort. Default value is 3.</td>
</tr>
<tr>
<td>$K$ (all designs)</td>
<td>Maximum number of patients at a dose level</td>
<td>A number used in the “Stopping Rule” that stops a trial if 1) the dose-assignment decision is to escalate to the next higher dose and there has been $K$ patients enrolled at that dose; or 2) the dose-assignment decision is to stay at the current dose and there has been $K$ patients enrolled at that dose; or 3) if the dose-assignment decision is to de-escalate to the previous lower dose and there has been $K$ patients enrolled at that dose; Default value is 12.</td>
</tr>
<tr>
<td>$p_{\text{cut}}$ (all designs)</td>
<td>Cutoff probability for futility rule</td>
<td>A cutoff probability used in the safety rule. Exclude dose $d$ if $Pr(p_d &lt; p_T</td>
</tr>
<tr>
<td>$q_{\text{cut}}$ (all designs)</td>
<td>Cutoff probability for efficacy rule</td>
<td>A cutoff probability used in the futility rule. Exclude dose $d$ if $Pr(q_d &lt; q_E</td>
</tr>
<tr>
<td>$p_E$ (Ji3+3, PRINTE)</td>
<td>Target efficacy probability</td>
<td>The lower bound of the response probability for the treatment to be considered promising and warrant further clinical development. Default value is 0.4.</td>
</tr>
<tr>
<td>$\epsilon_1, \epsilon_2$ (Ji3+3, PRINTE)</td>
<td>$\epsilon_1$: lower margin, $\epsilon_2$: higher margin</td>
<td>Two small fractions used to define the equivalence interval of the MTD. Any dose with a toxicity probability falling into the interval $[p_T - \epsilon_1, p_T + \epsilon_2]$ is considered an acceptable dose MTD. Default values for both are 0.05.</td>
</tr>
</tbody>
</table>
1.2. User Interface and Tutorial
1.2.2. Simulation Setup

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1^<em>$, $p_2^</em>$ (Ji3+3, PRINTE, TEPI)</td>
<td>Prespecified cutoff values in utility function on toxicity</td>
<td>Cutoff values in utility function for toxicity. The toxicity utility score is 1 when $p &lt; p_1^<em>$, is 0 when $p &gt; p_2^</em>$ and linearly decreases when $p$ is between $(p_1^<em>, p_2^</em>)$. Default values are 0.2 and 0.4.</td>
</tr>
<tr>
<td>$q_1^<em>$, $q_2^</em>$ (Ji3+3, PRINTE, TEPI)</td>
<td>Prespecified cutoff values in utility function on efficacy</td>
<td>Cutoff values in utility function for efficacy. The efficacy utility score is 0 when $p &lt; p_1^<em>$, is 1 when $p &gt; p_2^</em>$ and linearly increases when $p$ is between $(p_1^<em>, p_2^</em>)$. Default values are 0.2 and 0.6.</td>
</tr>
<tr>
<td>$p_{grad}$ (Ji3+3, PRINTE)</td>
<td>Cutoff probability for a dose to be considered as OBD</td>
<td>A cutoff value used when choosing OBD. If the posterior probability of utility function lying in the admissible utility region is below $p_{grad}$, no OBD will be selected and the trial ends without selecting an optimal dose. Default value is 0.2.</td>
</tr>
<tr>
<td>$a_1$, $b_1$ (Ji3+3, PRINTE, TEPI)</td>
<td>Prior beta distribution parameters of toxicity rate</td>
<td>The parameters in the prior beta distribution of toxicity rate, $Beta(a_1, b_1)$. Default values for both are 1 to be conservative, since a $Beta(1,1)$ prior implies a prior a dose has a toxicity rate of 0.5 with effective sample size of 0.5.</td>
</tr>
<tr>
<td>$a_2$, $b_2$ (Ji3+3, PRINTE, TEPI)</td>
<td>Prior beta distribution parameters of efficacy rate</td>
<td>The parameters in the prior beta distribution of efficacy rate, $Beta(a_2, b_2)$. Default values for both are 0.5, which is Jefferey’s prior (Jeffreys, 1946).</td>
</tr>
<tr>
<td>$s_1$ (UBOIN)</td>
<td>Maximum sample size in one dose at stage 1</td>
<td>The maximum number of patients to be treated in one dose at stage 1. Move to stage 2 when the number of patients treated on one of the doses reaches $s_1$. A value between 9 and 15 generally yields good operating characteristics. Default value is 12.</td>
</tr>
<tr>
<td>$s_2$ (UBOIN)</td>
<td>Maximum sample size at one dose at stage 2</td>
<td>The maximum number of patients to be treated in one dose at stage 2. Stop the trial and choose OBD when the number of patients treated at one of the doses reaches $s_2$. For most trials, a value between 18 and 24 is a reasonable choice for $s_2$. Default value is 18.</td>
</tr>
</tbody>
</table>
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

<table>
<thead>
<tr>
<th>Pick The Winner, Adaptive Randomization (UBOIN)</th>
<th>Methods to select next dose</th>
<th>Pick The Winner: The pick-the-winner (PW) approach deterministically assigning the next cohort of patients to dose that has the largest posterior mean utility. Adaptive Randomization: The adaptive randomization (AR) approach adaptively randomizes the next cohort of patients to a dose with probability proportional to its posterior mean utility.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{1,E}^<em>, \pi_{2,T}^</em>, \pi_{3,E}^<em>, \pi_{3,T}^</em>$ (EffTox)</td>
<td>Parameters in the desirable trade-off target values</td>
<td>$\pi_{1,E}^<em>$ is the smallest efficacy probability that the physician would consider desirable if toxicity were impossible. $\pi_{2,T}^</em>$ is the maximum desirable value of toxicity if the efficacy were 1. Set $\pi_{1,E}^<em>, \pi_{2,T}^</em>, \pi_{3,E}^<em>, \pi_{3,T}^</em>$ so that $\pi_1^* = (\pi_{1,E}^<em>, 0), \pi_2^</em> = (1, \pi_{2,T}^<em>), \pi_3^</em> = (\pi_{3,E}^<em>, \pi_{3,T}^</em>)$ Default values are 0.15, 0.6, 0.25, 0.3.</td>
</tr>
</tbody>
</table>
1.2. User Interface and Tutorial

1.2.2. Simulation Setup

1.2.2.3 Step 3: Generate scenarios

There are two ways to generate scenarios, automatically (in the “Auto Generation” tab) or through manual construction (in the “Manual Construction” tab). Users could also manually add or delete scenarios. Once scenarios are generated, click the button “Submit” to notify the software that the scenarios are final, then click the “Launch Simulation” button at the bottom of the page to run $n_{sim}$ (set in step 1) simulations, for each scenario and selected design (set in step 2), using the $p_T$ and $q_E$ values. (set in step 1).

**Auto Generation**  Click the “Auto Generation” button and six diverse scenarios will be created automatically, each of which contains the true toxicity probabilities for $n_{dose}$ dose levels. These generated scenarios are displayed (Figure 1.5). One can click the button to delete any scenario.

**Manual Construction**  A list of toxicity/efficacy probabilities are displayed. Click “Add” to add an empty, editable row of toxicity or efficacy probabilities. Click the button to delete the row. Click “Delete All” to delete all the rows.

Check the “Select” box in the front to select the row of toxicity or efficacy probabilities. Click “Select All” to select all the toxicity or efficacy rows.

Upon selection, click “Generate” to generate scenarios which will combine existing rows of toxicity and efficacy probabilities. The scenarios will be displayed in. (Figure 1.7)

Once the scenarios are generated, clicking the button will delete a scenario. Clicking “Delete All” will delete all the scenarios. Click the “Submit” button to notify the software that all the scenarios are final (Figure 1.7). If there are duplicated scenarios in the list, a message will be displayed on the website to indicate that the duplicated scenarios have been removed. Click the “OK” button to proceed to launch simulation. (Figure 1.8)

1.2.2.4 Launch Simulation

Once the above Steps 1-3 are completed, users can conduct simulated clinical trials to examine the operating characteristics of the selected designs using the selected scenarios, by clicking the “Launch Simulation” button at the bottom of Simulation Setup tab (Figures 1.7). A green “Launch Successful” message will be displayed on the website (Figure 1.9) to indicate that the simulation has been successfully launched. Users may click the “Proceed to Simulation Results” button in the pop-up box to track the simulation processing status and simulation results.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

**Figure 1.5:** Automatically generated scenarios in the **Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment** module.
Figure 1.6: Selecting toxicity and efficacy in the **Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment** module.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

Figure 1.7: Selecting scenarios in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.

Figure 1.8: Removing the duplicated scenarios in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
1.2. User Interface and Tutorial

1.2.2. Simulation Setup

Figure 1.9: “Launch Successful” message after launching simulation in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
Module 1. Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment

1.2.3 Simulation Results

In the Simulation Results tab, users can view and delete the simulation progress and simulation results (§1.2.3.1), restore the simulation settings if needed (§1.2.3.2), and download intelligent simulation reports (§1.2.3.3). Specifically, all the simulation results (figures and tables) can be downloaded in Word format, accompanying the statistical sections in a trial protocol. Hereinafter, we use simulation results and operating characteristics interchangeably.

1.2.3.1 View simulation results

In the Simulation Results tab, the Running Simulations panel exhibits the progress of ongoing simulation (Figure 1.10). The ongoing simulations are displayed in ascending order by the launch time. Click the icon “×” to delete the corresponding simulation.

![Running Simulations](image)

**Figure 1.10:** Simulation progress in the Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment module.

Once the simulations are completed, the Running Simulations panel in Figure 1.10 will disappear, green “simulation result created” massages will appear instead and stay at the same place of the Running Simulations panel unless explicitly dismissed by clicking the icon “×” at the end of the corresponding row, and the simulation results will be automatically loaded into the Simulation History panel (Figure 1.11), with the blue bold Unread term shown to indicate new results. All the previously completed simulations are also listed in the Simulation History panel. Simulation results for other modules can also be viewed under the Simulation History by dropping down the “Select a module” button (Figure 1.11).
1.2. User Interface and Tutorial

1.2.3. Simulation Results

Figure 1.11: Simulation Results in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.

Click the button to unfold the simulation results (Figure 1.12). The design settings are firstly displayed at the top of each simulation study (Figure 1.12). Then the results of simulation are shown as plots and tables below. And one can also click the button to delete the selected simulation results.

Figure 1.12: View the simulation results in the Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.

Simulation Result Tables

Full simulation results are presented in tabular format arranged by scenarios (Figure 1.13).

In the upper part of Figure 1.13, the first three columns summarize dose levels, their true toxicity and true efficacy probabilities; the remaining columns report four dose-specific summary statistics from the simulations: selection probability, average number of patients treated, average number of toxicities (i.e. DLTs), along with their standard deviations, and average number of responses,
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

along with their standard deviations, at each dose level. Specifically, they are

1) **Selection Prob.**: The proportion of simulated trials that select each dose level as the MTD.
2) **Average # of Patients Treated (s.d.)**: The average number of patients treated at each dose level and its standard deviation.
3) **Average # of Toxicities (s.d.)**: The average number of patients experienced DLT at each dose level and its standard deviation.
4) **Average # of Responses (s.d.)**: The average number of patients observed efficacy response at each dose level and its standard deviation.

The true OBD(s) of the scenario is (are) highlighted by the orange bar. The true OBD is defined as the dose that achieves the highest utility, which could be calculated using true toxicity, efficacy probabilities and the utility function.

In the lower part of Figure 1.13, more trial-specific summary statistics are reported, mainly from five aspects: **OBD Selection, Subjects Assignment, Trial Toxicity, Trial Stopping and Trial Sample Size**. Specifically, they are

- **OBD Selection**
  - **Prob. of Selecting OBD**: The proportion of simulated trials that select the true OBD at the end of the trial. The higher the value, the better the design.
  - **Prob. of Selecting Does-over-MTD**: The proportion of simulated trials that select the doses higher than the true MTD at the end of the trial. The lower the value, the better the safety of the design.
  - **Prob. of No Selection**: The proportion of simulated trials in which none of the dose levels are selected as the OBD. If a scenario does not have any OBD, this values is treated as the probability of selecting the true OBD.

- **Subjects Allocation**
  - **Prob. of Correct Allocation (s.d.)**: The average proportion of patients who are correctly assigned to the true OBD by the design across all the simulated trials and its standard deviation. The higher the value, the better the design.
  - **Prob. of Overdosing Allocation (s.d.)**: The average proportion of patients who are assigned to doses higher than the MTD by the design across all the simulated trials and its standard deviation. The lower the number, the better the safety of the design.

- **Trial Toxicity**
  - **Prob. of Toxicity**: The proportion of patients experiencing DLT across all the simulated trials. The lower the number, the fewer patients having DLTs under the design.
1.2. User Interface and Tutorial
1.2.3. Simulation Results

- **Trial Stopping**
  - **Prob. of Early Stopping Trial due to No admissible dose**: The proportion of simulated trials in which the trial is stopped because there is no admissible dose left. This means that all the doses have unacceptable toxicity or efficacy and are excluded by safety rule or futility rule.
  - **Prob. of Early Stopping Trial due to Reaching $K$**: The proportion of simulated trials in which the trial is stopped because the dose-assignment decision is to escalate/stay/de-escalate to a dose level but that dose has enrolled at least $K$ patients ($K < n$, e.g., $K = 12$).
  - **Prob. of Stopping Trial due to Reaching $n$**: The proportion of simulated trials in which the trial is stopped because the total number of patients enrolled and treated in a trial has reached or exceeded the pre-specified maximum sample size $n$.

- **Trial Sample Size**
  - **Average # of Patients Treated (s.d.)**: The average number of patients treated in the simulated trials and its standard deviation. Due to early stopping, this number is lower than or equal to $n$.

- **Statistics of UBOIN**
  - **Prob. of Entering Stage II**: The proportion of simulated trials in which the trial enters Stage II because the number of patients at one dose has reached or exceeded the pre-specified maximum sample size $s_1$ in Stage I.
  - **Average # of Patients Treated in Stage I**: The average number of patients treated in Stage I in the simulated trials.
  - **Average # of Patients Treated in Stage II**: The average number of patients treated in Stage II in the simulated trials.

  When calculating the standard deviation, we use $n_{sim}$ as the denominator instead of $(n_{sim} - 1)$ in U-Design.
Module 1. Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment

Figure 1.13: Simulation result tables in the Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment module.
1.2.3.2 Restore simulation setup

Users can restore the simulation settings from the simulation results by clicking the button at the upper right corner of each simulation results panel (yellow arrow in Figure 1.14). Upon clicking, the display will switch to the Simulation Setup page with the same simulation settings restored. This is useful to restore the old simulation settings for Reproducible results.

![Figure 1.14: Restore simulation setup and download simulation results in the Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment module.](image)

1.2.3.3 Download simulation results

There is a button at the upper right corner of each simulation results panel (green arrow in Figure 1.14). Click it to download a word file, which includes four parts:

- Part A: Complete simulation results under the designs and scenarios users added in the Simulation Setup tab;
- Part B:
- Part C: Detailed technical descriptions of the designs users added in the Simulation Setup tab;
- Part D: Reference

Users may select the required parts and modify them tailored for their trials or contact us via email (admin@laiyaconsulting.com) for consulting services.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

1.2.4 Decision Tables

This function generates decision tables based on the Ji3+3, PRINTE, and TEPI designs, which can be used to conduct a dose-finding trial. Users can click the tabs to switch between the tables for the Ji3+3, PRINTE, and TEPI designs.

Manually type in the design settings for decision table generation (Figure 1.15). The parameters are the same as the ones in Step 2 (1.2.2.2) in the Simulation Setup tab. See detailed parameter descriptions in Table 1.2.

![Decision Table](image)

**Figure 1.15:** Input parameters in the Generate Decision Table tab of Single-Agent Dose-Finding Designs with Efficacy and Toxicity Endpoints and Cohort Enrollment module.

Click the “Generate” button to generate decision table (Figure 1.16). Decision tables are automatically generated for 3, 6, 9 and 12 patients at a dose in the panel below.

To generate a single decision table by specifying the number of patients treated at a dose \( d \), set \( n_d \) in the box and click the button “Add”. (Figure 1.16)
1.2. User Interface and Tutorial

1.2.4. Decision Tables

For each decision table, the column represents the number of patients responses among those treated at the dose, and the row represents the number of patients who have experienced dose-limiting toxicity (DLT) events. Note that these are the counts of patients, not DLT events or responses. For example, column 3 and row 1 means that among the patients that have been treated at the current dose 3 of them experiences DLT, and 1 of them responses.

Each cell in the decision table provides the dose-assignment decision based on the readouts from the corresponding row and column. For example, for column 3 and row 1, i.e., 3 patients experience DLTs, and 1 patient has efficacy response, the decision is “EU”. The letters in the decision table represent different dose-assignment decisions as shown below:

- “E” stands for escalating to the next higher dose,
- “S” stands for staying at the current dose, or escalate to dose $d + 1$ if $d$ is not the highest dose and $d + 1$ is untried
- “S*” stands for staying at the current dose
- “R” stands for escalate to the previous lower dose
- “EU” stands for escalate to the next higher dose and exclude the current dose from future use in the trial due to unacceptable low efficacy
- “DU” stands for escalate to the previous lower dose and exclude the current dose from future use in the trial due to unacceptable low efficacy
- “DUT” stands for escalate to the previous lower dose and exclude the current dose and higher doses from future use in the trial due to unacceptable toxicity

**Figure 1.16:** Decision tables in the Generate Decision Table tab of Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

- “D” stands for de-escalating to the previous lower dose,
- “DUT” stands for de-escalating to the previous lower dose, and the current dose and its higher doses is deemed unacceptable due to severe toxicity and will not be used again in the study. If at the first dose level, users can choose to early-terminate the trial or not based on their own discretion.
- “EUE” stands for escalating to the higher dose and marking the current dose as unacceptable (due to futility) so that it will never be used again in the remainder of the trial.
- “DUE” stands for de-escalating to the previous lower dose, and the current dose is deemed unacceptable due to futility and will not be used again in the study.

Some additional detailed explanation of the decisions are provided in the decision table report.

The meaning of the notations are shown below:

- The superscript * on DUE indicates that according to the Ji3+3 design, the decision is S and the current dose is deemed unacceptable due to futility. In this case, a decision S indicates a moderate or high toxicity probability, so the only sensible action is to de-escalate to the previous lower dose, and remove the current dose (due to futility) from the study.
- The superscript ** on DUT indicates that if the current dose is the first dose level, users can choose to early-terminate the trial or not based on their own discretion.

Click “DOWNLOAD ONE” to download a word file, which includes the design settings and the single decision table in the tab selected. Click “DOWNLOAD ALL” to download a word file, which includes the design settings and all the decision tables generated.
1.2.5 OBD Estimation

In this module, all designs aim to estimate the OBD when the trial is completed and the data is collected. The detailed statistical models for the included designs are described in §1.3.

First, select a design and provide corresponding model parameters. Second, select the number of doses \( n_{dose} \) from the dropdown box, and an editable table will be shown on the website (Figure 1.17). For the Ji3+3, PRINTE and TEPI design, provide the number of patients treated, the observed number of DLT events, and provide the observed number of efficacy events at each dose into the table; for the UBOIN and Efftox design, provide the observed number of patients who has no efficacy but DLT \( \left( Y_E, Y_T \right) = (0, 1) \), no efficacy and no DLT \( \left( Y_E, Y_T \right) = (0, 0) \), efficacy and DLT \( \left( Y_E, Y_T \right) = (1, 1) \) and the number of patients who has no efficacy and no DLT \( \left( Y_E, Y_T \right) = (1, 0) \). Click the “Generate” button to estimate the utilities of each dose and estimate the OBD for the trial. The estimated utility will be displayed in a table and the estimated OBD will be highlighted in green color as shown in Figure 1.18.

Hover mouse over each parameter, and a description will be displayed explaining the meaning of the parameter. See detailed parameter descriptions in Table 1.2 in §1.2.2.2.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

**Figure 1.17:** Input parameters in the OBD Estimation tab of Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment module.
1.2. User Interface and Tutorial

1.2.5. OBD Estimation

Figure 1.18: Determine the estimated OBD in the **Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment** module.
1.3 Statistical Methods Review

1.3.1 The Joint i3+3 (Ji3+3) Design

Ji3+3 is a rule-based phase I/II ACT dose-finding design proposed by Lin and Ji (2020b). Building upon i3+3 (Liu et al., 2020), Ji3+3 takes into account of both toxicity and efficacy outcomes in making dosing recommendations. Basically, the decision rules of the Ji3+3 design incorporate and extend the toxicity rules in i3+3 with a set of efficacy rules. Simulation results show that Ji3+3 outperforms existing designs when monotonic dose response assumption is violated, and achieves comparable performance when the assumption holds. Since Ji3+3 is a model-free design, it is transparent to physicians and simple to implement.

1.3.1.1 Dose-Finding Algorithm

Consider $D$ ascending doses in a single-agent ACT phase I trial. Due to ethical considerations, it is always assumed that the toxicity probability $p_d$ increases with dose level $d$, that is, $p_1 \leq \cdots \leq p_D$. However, the efficacy probability $q_d$ may increase initially and then reach a plateau from which minimal improvement or even decreasing efficacy may be seen with increasing dose. For this reason, we assume that $q_d$ is not monotone with $d$, and that $p_d$ and $q_d$ are independent. Suppose that dose $d$ is currently used in the trial and $n_d$ patients have already been allocated to dose $d$, with $x_d$ and $y_d$ patients experiencing toxicity and efficacy outcomes. Aggregating across all the doses, the trial data are denoted as $Data = \{(n_d, x_d, y_d), d = 1, \cdots, D\}$.

Denote $p_T$ as the target toxicity rate, which is the probability of toxicity at the MTD; denote $p_E$ as the target efficacy rate. In Ji3+3, $[p_T - \epsilon_1, p_T + \epsilon_2]$ is defined as the Equivalence Interval (EI), where $(\epsilon_1, \epsilon_2)$ are two small fractions that account for the uncertainty around $p_T$. This allows doses whose toxicity probabilities differ from $p_T$ to be considered as the MTD. Given the observed data $Data$, the dose-finding algorithm of the Ji3+3 design is shown in Table 1.3. The algorithm follows these principles:

1. If there is lack of evidence for efficacy, escalate to achieve higher efficacy; else, stay at the current dose because it is considered to have sufficient efficacy.

2. For toxicity, the idea is to compare the observed toxicity rate $\frac{x_d}{n_d}$ with the EI.

   - If $\frac{x_d}{n_d}$ is below the EI, the dose is considered safe; if $\frac{x_d}{n_d}$ is inside the EI, the dose is considered to be close to the MTD; if $\frac{x_d}{n_d}$ is above the EI, the dose is considered not safe except when $\frac{x_d - 1}{n_d}$ is below the EI.
1.3. Statistical Methods Review
1.3.1. The Joint i3+3 (Ji3+3) Design

Table 1.3: Schema of the Ji3+3 design.

<table>
<thead>
<tr>
<th>Eff cond.</th>
<th>Tox cond.</th>
<th>Next dose (Decision)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{x_d}{n_d} &lt; EI$</td>
<td>$\frac{x_d}{n_d} &lt; EI$</td>
<td>$d + 1 (E)$</td>
</tr>
<tr>
<td>$\frac{x_d}{n_d} \in EI$</td>
<td>$\frac{x_d}{n_d} \in EI$</td>
<td>$d (S)$</td>
</tr>
<tr>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} &lt; EI$</td>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} &lt; EI$</td>
<td>$d (S)$</td>
</tr>
<tr>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} \in EI$</td>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} \in EI$</td>
<td>$d - 1 (D)$</td>
</tr>
<tr>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} &gt; EI$</td>
<td>$\frac{x_d}{n_d} &gt; EI &amp; \frac{x_{d-1}}{n_d} &gt; EI$</td>
<td>$d - 1 (D)$</td>
</tr>
</tbody>
</table>

*: Escalate to dose $d + 1$ if $n_{d+1} = 0$.

- When $\frac{x_d}{n_d}$ is below the EI and $\frac{x_d}{n_d}$ is above the EI, the data is noisy since increment of one toxicity event renders the observed toxicity rate to jump from below the EI to above the EI. In other words, the observed data is not very informative because change of one toxicity event can greatly influence the toxicity estimate.

Consider an example. Suppose EI = [0.2, 0.3] with $x_d = 1$ and $n_d = 3$. Even though $\frac{x_d}{n_d} = \frac{1}{3}$ is above the EI, $\frac{x_{d-1}}{n_d} = 0$ is below the EI. And therefore, dose $d$ should not be considered as above the MTD.

3. Intersecting the two dosing principles for toxicity and efficacy, and taking the more conservative decision between the two, we arrive at the decisions in Table 1.3.

4. When $d$ is the highest dose or lowest dose, the above rules are modified as special cases,

- If the current dose is the highest dose, decision E (escalate and treat the next cohort of patients at the next higher dose) should be replaced with decisionS (stay and continue to enroll patients at the current dose), since there is no dose to escalate to.

- Similarly, if the current dose is the lowest dose, decision D (de-escalate to the next lower dose) should be replaced with S since there is no dose to de-escalate to.
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

Safety and futility rules

- Safety rule: if \( \Pr(p_d > p_T | x_d, n_d) > p_{\text{cut}} \) for a \( p_{\text{cut}} \) close to 1 (say, 0.95), exclude doses \( d, d + 1, \cdots, D \), from future use in the trial; treat the next cohort of patients at dose \( (d - 1) \).

- Futility rule: if \( \Pr(q_d < q_E | y_d, n_d) > q_{\text{cut}} \) for a \( q_{\text{cut}} \) close to 1 (say, 0.7), where \( q_E \) is the minimum acceptable probability of efficacy, then exclude dose \( d \) from future use in the trial.

Here, \( q_E \) is the reference efficacy rate, e.g., the efficacy rate of standard care.

Note that, here we assume the prior for each \( p_d \) follows an independent \( \text{beta}(a_1, b_1) \), and the prior for each \( q_d \) follows an independent \( \text{beta}(a_2, b_2) \), where \( \text{beta}(\alpha, \beta) \) denotes a beta distribution with mean \( \alpha / (\alpha + \beta) \). The posterior distributions for \( p_d \) and \( q_d \) in the above rules are \( \text{beta}(a_1 + x_d, b_1 + n_d - x_d) \) and \( \text{beta}(a_2 + y_d, b_2 + n_d - y_d) \), respectively.

Stopping rules

The trial is stopped if

1. the prespecified maximum total sample size \( n \) is reached; or
2. the lowest dose shows excessive toxicity according to the safety rule; In this case, the trial is early stopped and the MTD cannot be determined; or
3. optional:
   - the Ji3+3 decision is “S”, to stay at the current dose, and the current dose has enrolled \( K \) patients;
   - the Ji3+3 decision is “E”, to escalate to the next higher dose, and that next higher dose has enrolled \( K \) patients;
   - the Ji3+3 decision is “D”, to de-escalate to the previous lower dose, and that previous lower dose has enrolled \( K \) patients.

1.3.1.2 Dose Selection

At the end of the trial, Ji3+3 chooses the OBD using a joint utility score \( U(p, q) = f_1(p)f_2(q) \) (suppressing dose \( d \) in the notation), which takes the product of toxicity utility \( f_1(p) \) in (1.1) and efficacy utility \( f_2(q) \) in (1.2).

\[
f_1(p) = \begin{cases} 
1, & p \in (0, p_1^*), \\
1 - \frac{p - p_1^*}{p_2^* - p_1^*}, & p \in (p_1^*, p_2^*), \\
0, & p \in (p_2^*, 1).
\end{cases}
\] (1.1)
1.3. Statistical Methods Review

1.3.1. The Joint i3+3 (Ji3+3) Design

For toxicity, define two thresholds $p_1^*$ and $p_2^*$ such that the toxicity utility score is 1 when $p < p_1^*$, 0 when $p > p_2^*$, and linearly decreases when $p$ is between $(p_1^*, p_2^*)$. For efficacy, define two thresholds $q_1^*$ and $q_2^*$ such that the efficacy utility score is 0 when $q < q_1^*$, is 1 when $q > q_2^*$, and linearly increases when $q$ is between $(q_1^*, q_2^*)$. The OBD is selected according to the following process.

1. We generate a total of $T$ random samples, $\{p_{d}^{(t)}, t = 1, \cdots, T\}$ and $\{q_{d}^{(t)}, t = 1, \cdots, T\}$, from the posterior distributions $beta(a_0 + x_d, b_0 + n_d - x_d)$ and $beta(a_0 + y_d, b_0 + n_d - y_d)$ for each dose $d$, respectively. Here, U-Design sets $a_0 = b_0 = 0.005$ and $T = 1000$.
2. For toxicity probabilities of all doses in each sample $t$, $\mathbf{p}^{(t)} = (p_1^{(t)}, \cdots, p_D^{(t)})$, we perform isotonic transformation using the pool adjacent violators algorithm (PAVA; Mair et al. 2009) on $\mathbf{p}^{(t)}$ to obtain $\tilde{\mathbf{p}}^{(t)} = (\tilde{p}_1^{(t)}, \cdots, \tilde{p}_d^{(t)})$, where $\tilde{p}_i^{(t)} \leq \tilde{p}_j^{(t)}$ if $i < j$.
3. We propose a probabilistic inference for selecting the OBD and avoid selecting doses with low utility. Define an admissible probability region (APR) $A(p, q) = \{(p, q) \mid p \in (0, p_T], q \in [q_E, 1)\}$. Then the OBD is selected only from the candidate dose set $\mathcal{A}$,

$$\mathcal{A} = \{d \mid p_{d,n,d} \geq p_{\text{grad}}, n_d > 0, d = 1, \cdots, D\},$$

where $p_{d,n,d} = \Pr\{(p_d, q_d) \in \text{APR} \mid \text{Data}\}$ is the posterior probability that dose $d$ belongs to APR and $p_{\text{grad}}$ is a small value (say, 0.1). We use a simple a simple numerical approximation approach to compute $p_{d,n,d}$ given by

$$\hat{p}_{d,n,d} = \frac{1}{T} \sum_{t=1}^{T} 1\{ (\tilde{p}_d^{(t)}, q_d^{(t)}) \in \text{APR} \}.$$

4. The final selected dose $d^*$ is the one that maximizes the utility score $U(p_d, q_d)$. That is, $d^* = \arg\max_{d \in \mathcal{A}} \hat{E}[U(p_d, q_d) \mid \text{Data}]$, where

$$\hat{E}[U(p_d, q_d) \mid \text{Data}] = \frac{1}{T} \sum_{t=1}^{T} U(\tilde{p}_d^{(t)}, q_d^{(t)}).$$
1.3.2 The Probability Intervals of Toxicity and Efficacy (PRINTE) Design

PRINTE (Lin and Ji, 2020a) building upon previous work in TEPI (Li et al., 2017), is a dose-finding design which utilizes both toxicity and efficacy in making dosing decisions. Similar to TEPI, PRINTE partitions the unit intervals \((0, 1)\) for both the toxicity probability \(p_i\) and efficacy probability \(q_i\) into subintervals, and makes dosing-decisions based on the posterior probability of the interval combinations. Compared to TEPI, it does not require a physician-elicited decision table, the choice of which could be arbitrary and difficult, and might be subjective to Ockhams razor (Guo et al., 2017). Instead, PRINTE utilizes a decision principle that is simple and transparent, and is commonly applied in practice.

1.3.2.1 Probability Model

Consider \(D\) ascending doses in a single-agent ACT phase I trial. Due to ethical considerations, it is always assumed that the toxicity probability \(p_d\) increases with dose level \(d\), that is, \(p_1 \leq \cdots \leq p_D\). However, the efficacy probability \(q_d\) may increase initially and then reach a plateau from which minimal improvement or even decreasing efficacy may be seen with increasing dose. For this reason, we assume that \(q_d\) is not monotone with \(d\), and that \(p_d\) and \(q_d\) are independent. Suppose that dose \(d\) is currently used in the trial and \(n_d\) patients have already been allocated to dose \(d\), with \(x_d\) and \(y_d\) patients experiencing toxicity and efficacy outcomes. Aggregating across all the doses, the trial data are denoted as \(Data = \{(n_d, x_d, y_d), d = 1, \cdots , D\}\).

Let \(p_T\) be the target toxicity probability and \(p_E\) be the target efficacy rate. Define the equivalence interval (EI) as \([p_T - \epsilon_1, p_T + \epsilon_2]\) where \(\epsilon_1\) and \(\epsilon_2\) are two small fractions that allow toxicity probability of MTD to be in a range of values, rather than a single point \(p_T\).

Consider the unit square of \(Q = (0, 1) \times (0, 1)\) (here, operation \(\times\) represents the Cartesian product) representing the joint probability square of toxicity and efficacy probabilities. For toxicity, there are three probability intervals, \((0, p_T - \epsilon_1)\), \([p_T - \epsilon_1, p_T + \epsilon_2]\), and \([p_T + \epsilon_2, 1]\), which represent the under-dosing, equivalence, and over-dosing intervals. For efficacy, consider two probability intervals, \((0, p_E]\) and \((p_E, 1]\), which corresponds to low and high probability of efficacy. Denote \(S_{tox} = \{(0, p_T - \epsilon_1), [p_T - \epsilon_1, p_T + \epsilon_2], (p_T + \epsilon_2, 1]\}\) as the set of three toxicity probability intervals and \(S_{eff} = \{(0, p_E], (p_E, 1]\}\) as the set of two efficacy probability intervals. Taking a Cartesian product of the two sets, we obtain a set of six probability rectangles (PRs) in \(Q\), which is given by...
1.3. Statistical Methods Review

1.3.2. The Probability Intervals of Toxicity and Efficacy (PRINTE) Design

\[ S_{joint} = S_{tox} \times S_{eff} \]

\[ = \{ s_{ll} = (0, p_T - \epsilon_1) \times (0, p_E], \ s_{lh} = (0, p_T - \epsilon_1) \times (p_E, 1), \]

\[ s_{el} = [p_T - \epsilon_1, p_T + \epsilon_2] \times (0, p_E], \ s_{eh} = [p_T - \epsilon_1, p_T + \epsilon_2] \times (p_E, 1), \]

\[ s_{hl} = (p_T + \epsilon_2, 1) \times (0, p_E], \ s_{hh} = (p_T + \epsilon_2, 1) \times (p_E, 1) \} , \]

where the two letters \( l \) and \( h \) denotes low or high, respectively. See Figure 1.19a for a display of the probability rectangles in \( S_{joint} \).

\[ \text{Figure 1.19:} \text{ An example demonstrating the 2-dimensional probability rectangles and sub-rectangles of toxicity and efficacy. (a): The horizontal axis is the probability intervals of toxicity } (0, p_T - \epsilon_1), [p_T - \epsilon_1, p_T + \epsilon_2], \text{ and } (p_T + \epsilon_2, 1). \text{ The vertical axis is the probability intervals of efficacy } (0, p_E] \text{ and } (p_E, 1). \text{ The Cartesian product of both probability intervals is shown as the 6 probability rectangles (PRs) separated by dashed lines. (b): The horizontal axis is the probability sub-intervals of toxicity, where } (0, p_T - \epsilon_1), [p_T - \epsilon_1, p_T + \epsilon_2] \text{ and } (p_T + \epsilon_2, 1) \text{ are further divided into smaller intervals with the same length of } \epsilon_1 + \epsilon_2. \text{ The vertical axis is the probability sub-intervals of efficacy, where } (0, p_E] \text{ and } (p_E, 1) \text{ are further divided into multiple smaller intervals with the same length of their maximum common divisor. The Cartesian product of all probability sub-intervals is shown as the probability sub-rectangles (sub-PRs) separated by dashed lines.} \]

Divide the six PRs into sub-PRs with similar area, see Figure 1.19b for an illustration, which is realized by three steps.

1. For the toxicity interval set \( S_{tox} \), divide \( S_{tox} \) into sub-intervals given by the length of the
Module 1. Single-Agent Dose-Finding Designs with Efficacy&Toxicity Endpoints and Cohort Enrollment

equivalence interval \((\epsilon_1 + \epsilon_2)\). The division is done by keeping the equivalence interval \(m^t_e = [p_T - \epsilon_1, p_T + \epsilon_2]\) unchanged, and sub-divide the under-dosing interval \((0, p_T - \epsilon_1)\) and over-dosing interval \((p_T + \epsilon_2, 1)\) into sub-intervals with the length \(l_e = \epsilon_1 + \epsilon_2\), except for the sub-intervals on the boundary. Denote the set of all the resulting sub-intervals as \(M_{tox} = \{m^t_{ls}, m^e_{e}, m^t_{hs}\}\), in which \(m^t_{ls}\) and \(m^t_{hs}\) are the sub-intervals generated by dividing the under-dosing and over-dosing intervals, respectively.

2. For the efficacy interval set \(S_{eff}\), divide two intervals in \(S_{eff}\) into sub-intervals with the length \(l_e\),

\[
 l_e = \max \left\{ 0.10, \frac{\gcd(100 \cdot p_E, 100 \cdot (1 - p_E))}{100} \right\},
\]

where \(\gcd(a, b)\) is the greatest common divisor of \(a\) and \(b\). Denote the resulting set of sub-intervals by \(M_{eff} = \{m^e_{ls}, m^e_{hs}\}\), where

\[
 m^e_{ls} = \{(0, p_E - t_1 l_e), \ldots, (p_E - 2 l_e, p_E - l_e), (p_E - l_e, p_E)\},
\]

\[
 m^e_{hs} = \{(p_E, p_E + l_e), (p_E + l_e, p_E + 2 l_e), \ldots, (p_E + 2 l_e, 1)\}.
\]

Here, \(t_1\) and \(t_2\) are the maximum positive integers such that \(p_E - t_1 l_e > 0\) and \(p_E + t_2 l_e < 1\), respectively.

3. Take Cartesian product of the set of \(M_{tox}\) and \(M_{eff}\) to generate a set of two-dimensional sub-PRs of equal area, except for those on the boundary of the toxicity axis next to 0 or 1. These sets are denoted by \(M_{joint}\) as illustrated below, where \(k_{uv}, u \in \{l, e, h\}, v \in \{l, h\}\) denotes the number of sub-PRs in \(m_{uv}\).

\[
 M_{joint} = M_{tox} \times M_{eff}
\]

\[
 = \{m^t_{ls}, m^e_{e}, m^t_{hs}\} \times \{m^e_{ls}, m^e_{hs}\}
\]

\[
 = \{m_{ll} = \{m^1_{ll}, \ldots, m^k_{ll}\}, m_{lh} = \{m^1_{lh}, \ldots, m^k_{lh}\},
\]

\[
 m_{el} = \{m^1_{el}, \ldots, m^k_{el}\}, m_{eh} = \{m^1_{eh}, \ldots, m^k_{eh}\},
\]

\[
 m_{hl} = \{m^1_{hl}, \ldots, m^k_{hl}\}, m_{hh} = \{m^1_{hh}, \ldots, m^k_{hh}\}\}
\]

PRINTe treats each sub-PR as a model and considers a model indicator \(a\) that takes one of the sub-PRs. Denote \(m_{uv}\) as a sub-PR in the set \(M_{joint}\), and define \(\{a = m_{uv}\} = \{(p_d, q_d) \in m_{uv}\}\). Embedding the model indicator \(a\) into a Bayesian hierarchic model, we compute the posterior probability of each sub-PR given the observed toxicity and efficacy outcomes \(\{x_d, y_d\}\), given by

\[
 P(a = m_{uv} \mid x_d, y_d, n_d) = Pr((p_d, q_d) \in m_{uv} \mid x_d, y_d, n_d).
\]

From model selection perspective,
finding the optimal decision is equivalent to selecting the optimal model (sub-PR) that maximizes the marginal posterior model probability.

We further define dose-finding decisions as \( a^* \in \{ E, S, D \} \) and maps \( a \in \{ m_{ll}, m_{lh}, m_{el}, m_{eh}, m_{hl}, m_{hh} \} \) to \( a^* \in \{ E, S, D \} \) according to the following rule \( \mathcal{R} \).

\[
a^* = \mathcal{R}(a) = \begin{cases} 
E, & \text{if } a = m_{ll} \\
E, & \text{if } a = m_{el} \text{ and } n_{d+1} = 0 \\
S, & \text{if } a = m_{el} \text{ and } n_{d+1} > 0 \\
S, & \text{if } a \in \{ m_{lh}, m_{eh} \} \\
D, & \text{if } a \in \{ m_{hl}, m_{hh} \} 
\end{cases}
\]

The rule \( \mathcal{R} \) states that the dosing decisions \( \{ E, S, D \} \) correspond to the models that describe the toxicity and efficacy probabilities of the dose. According to \( \mathcal{R}(a) \), escalation (\( E \)) is recommended if toxicity and efficacy are both deemed low; Stay (\( S \)) is selected if \( n_{d+1} > 0 \), toxicity is near the MTD range and efficacy is low, while escalation (\( E \)) is recommended if \( n_{d+1} = 0 \), i.e., dose \((d+1)\) is untried; Stay (\( S \)) is selected if either 1) toxicity is low but efficacy is high \( m_{lh} \), or 2) toxicity is near the MTD range and efficacy is high; Lastly, de-escalation \( D \) is selected if toxicity is high regardless of efficacy. The goal is to seek an optimal \( a \) that leads to an optimal decision \( a^* \).

### 1.3.2.2 Dose-finding Algorithm

The implementation of PRINTE is simple and transparent. The only required input values are \( p_T \), \( p_E \), and the equivalence interval \([p_T - \epsilon_1, p_T + \epsilon_2]\). Once they are provided, optimal decisions \( a^{opt*} \) can be calculated for all possible toxicity and efficacy outcomes at a given dose. Suppose that the current dose is \( d \), \( d \in \{1, ..., D\} \). Record \( \{x_d, y_d, n_d\} \) and calculate the marginal model posterior probabilities \( \Pr(a \mid x_d, y_d, n_d) \), and then the optimal decision \( a^{opt*} \) can be determined. The next cohort of patients is allocated to \( \{\max(1, d - 1), d, \min(d + 1, D)\} \) according to \( a^{opt*} \).

### Safety and futility rules

- **Safety rule**: If \( \Pr(p_d > p_T \mid x_d, n_d) > p_{cut} \) for a \( p_{cut} \) close to 1 (say, 0.95), exclude doses \( d, d + 1, \ldots, D \), from future use in the trial; treat the next cohort of patients at dose \( (d - 1) \).

- **Futility rule**: If \( \Pr(q_d < q_E \mid y_d, n_d) > q_{cut} \) for a \( q_{cut} \) close to 1 (say, 0.7), where \( q_E \) is the minimum acceptable probability of efficacy, then exclude dose \( d \) from future use in the trial.

Here, \( q_E \) is the reference efficacy rate, e.g., the efficacy rate of standard care.
Note that, here we assume the prior for each $p_d$ follows an independent $\text{beta}(a_1, b_1)$, and the prior for each $q_d$ follows an independent $\text{beta}(a_2, b_2)$, where $\text{beta}(\alpha, \beta)$ denotes a beta distribution with mean $\alpha / (\alpha + \beta)$. The posterior distributions for $p_d$ and $q_d$ in the above rules are $\text{beta}(a_1 + x_d, b_1 + n_d - x_d)$ and $\text{beta}(a_2 + y_d, b_2 + n_d - y_d)$, respectively.

**Stopping rules**

The trial is stopped if

1. the prespecified maximum total sample size $n$ is reached; or
2. the lowest dose shows excessive toxicity according to the safety rule; In this case, the trial is early stopped and the MTD cannot be determined; or
3. optional:
   - the PRINTE decision is “S”, to stay at the current dose, and the current dose has enrolled $K$ patients;
   - the PRINTE decision is “E”, to escalate to the next higher dose, and that next higher dose has enrolled $K$ patients;
   - the PRINTE decision is “D”, to de-escalate to the previous lower dose, and that previous lower dose has enrolled $K$ patients.

### 1.3.2.3 Dose Selection

At the end of the trial, PRINTE chooses the OBD using a joint utility score $U(p, q) = f_1(p)f_2(q)$ (suppressing dose $d$ in the notation), which takes the product of toxicity utility $f_1(p)$ in (1.3) and efficacy utility $f_2(q)$ in (1.4).

$$f_1(p) = \begin{cases} 
1, & p \in (0, p_1^*), \\
1 - \frac{p - p_1^*}{p_2^* - p_1^*}, & p \in (p_1^*, p_2^*), \\
0, & p \in (p_2^*, 1). 
\end{cases} \quad (1.3)$$

$$f_2(q) = \begin{cases} 
0, & q \in (0, q_1^*), \\
\frac{q - q_1^*}{q_2^* - q_1^*}, & q \in (q_1^*, q_2^*), \\
1, & q \in (q_2^*, 1). 
\end{cases} \quad (1.4)$$

For toxicity, define two thresholds $p_1^*$ and $p_2^*$ such that the toxicity utility score is 1 when $p < p_1^*$, 0 when $p > p_2^*$, and linearly decreases when $p$ is between $(p_1^*, p_2^*)$. For efficacy, define
1.3. Statistical Methods Review

1.3.2. The Probability Intervals of Toxicity and Efficacy (PRINTE) Design

two thresholds $q_1^*$ and $q_2^*$ such that the efficacy utility score is 0 when $q < q_1^*$, is 1 when $q > q_2^*$, and linearly increases when $q$ is between $(q_1^*, q_2^*)$. The OBD is selected according to the following process.

1. We generate a total of $T$ random samples, $\{p_d^{(t)}, t = 1, \cdots, T\}$ and $\{q_d^{(t)}, t = 1, \cdots, T\}$, from the posterior distributions $\text{beta}(a_0 + x_d, b_0 + n_d - x_d)$ and $\text{beta}(a_0 + y_d, b_0 + n_d - y_d)$ for each dose $d$, respectively. Here, U-Design sets $a_0 = b_0 = 0.005$ and $T = 1000$.

2. For toxicity probabilities of all doses in each sample $t$, $p_d^{(t)} = (p_d^{(t)}_1, \cdots, p_d^{(t)}_D)$, we perform isotonic transformation using the pool adjacent violators algorithm (PAVA; Mair et al. 2009) on $p_d^{(t)}$ to obtain $\tilde{p}_d^{(t)} = (\tilde{p}_d^{(t)}_1, \cdots, \tilde{p}_d^{(t)}_D)$, where $\tilde{p}_i^{(t)} \leq \tilde{p}_j^{(t)}$ if $i < j$.

3. We propose a probabilistic inference for selecting the OBD and avoid selecting doses with low utility. Define an admissible probability region (APR) $A(p, q) = \{(p, q) \mid p \in (0, p_T], q \in [q_E, 1]\}$. Then the OBD is selected only from the candidate dose set $A$,

$$A = \{d \mid p_{\text{in}, d} \geq p_{\text{grad}}, n_d > 0, d = 1, \cdots, D\},$$

where $p_{\text{in}, d} = \Pr \{(p_d, q_d) \in \text{APR} \mid \text{Data}\}$ is the posterior probability that dose $d$ belongs to APR and $p_{\text{grad}}$ is a small value (say, 0.1). We use a simple a simple numerical approximation approach to compute $p_{\text{in}, d}$ given by

$$\hat{p}_{\text{in}, d} = \frac{1}{T} \sum_{t=1}^T 1 \{ (\tilde{p}_d^{(t)}, q_d^{(t)}) \in \text{APR} \}.$$

4. The final selected dose $d^*$ is the one that maximizes the utility score $U(p_d, q_d)$. That is, $d^* = \arg\max_{d \in A} \hat{E}[U(p_d, q_d) \mid \text{Data}]$, where

$$\hat{E}[U(p_d, q_d) \mid \text{Data}] = \frac{1}{T} \sum_{t=1}^T U(p_d^{(t)}, q_d^{(t)}).$$
1.3.3 The Toxicity and Efficacy Probability Interval (TEPI) Design

TEPI, proposed in Li et al. (2017), is a practical dose-finding design for ACT trials that incorporates both toxicity and efficacy data. It is a natural extension of mTPI by adding the efficacy interval into the dose-finding model. TEPI partitions the unit intervals \((0, 1)\) for both the toxicity probability \(p_i\) and efficacy probability \(q_i\) into subintervals, denoted as \((a, b)\) and \((c, d)\), respectively. Then it uses beta-binomial models to estimate the efficacy and toxicity probability and makes dosing-decisions based on the joint unit probability mass (JUPM) of the interval combinations \((a, b) \times (c, d)\). TEPI is transparent to clinicians and simple to implement in practice.

1.3.3.1 Elicited decision table

The dose-finding algorithm of TEPI is based on a clinician-elicited decision table in terms of efficacy and toxicity probability intervals. The procedures of eliciting the decision table are as follows.

Consider \(D\) ascending doses in a single-agent ACT phase I trial. Due to ethical considerations, it is always assumed that the toxicity probability \(p_d\) increases with dose level \(d\), that is, \(p_1 \leq \cdots \leq p_D\). However, the efficacy probability \(q_d\) may increase initially and then reach a plateau from which minimal improvement or even decreasing efficacy may be seen with increasing dose. For this reason, we assume that \(q_d\) is not monotone with \(d\), and that \(p_d\) and \(q_d\) are independent. Suppose that dose \(d\) is currently used in the trial and \(n_d\) patients have already been allocated to dose \(d\), with \(x_d\) and \(y_d\) patients experiencing toxicity and efficacy outcomes. Aggregating across all the doses, the trial data are denoted as \(Data = \{(n_d, x_d, y_d), d = 1, \cdots, D\}\).

Partition the unit intervals \((0, 1)\) for \(p_d\) and \(q_d\) into four subintervals. Denoting \((a, b)\) and \((c, d)\) a subinterval in the partition for \(p_d\) and \(q_d\) respectively, where

\[
(a, b) \in \{(0, t_1), (t_1, t_2), (t_2, t_3), (t_3, 1)\},
\]

\[
(c, d) \in \{(0, e_1), (e_1, e_2), (e_2, e_3), (e_3, 1)\}.
\]

The interval combinations \((a, b) \times (c, d)\) form the basis for dose-finding decisions, with each combination corresponding to a specific decision, such as dose escalation or de-escalation. U-Design uses a default fixed decision for each interval combination, see Table 1.4.

In order to formulate this table, it is required to determine: (i) bounds of efficacy rate interval, \(e_1, e_2, e_3\), and (ii) bounds of toxicity rate interval, \(t_1, t_2, t_3\).
1.3. Statistical Methods Review

1.3.3. The Toxicity and Efficacy Probability Interval (TEPI) Design

Table 1.4: An default decision table for each interval combination.

<table>
<thead>
<tr>
<th>Toxicity Rate</th>
<th>Efficacy Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low (0, e₁)</td>
</tr>
<tr>
<td></td>
<td>Moderate (e₁, e₂)</td>
</tr>
<tr>
<td></td>
<td>High (e₂, e₃)</td>
</tr>
<tr>
<td></td>
<td>Unacceptable (e₃, 1)</td>
</tr>
</tbody>
</table>

Note: "E", "S" and "D" denote escalation, stay and de-escalation, respectively.

1.3.3.2 Dose-finding Algorithm

Building upon the preset table, we set up a local decision-theoretic framework and derive a Bayes rule. Here, local means that the framework focuses on the optimal decision to be made for the current dose instead of the trial. We show that the Bayes rule is equivalent to computing the joint unit probability mass (JUPM) for the toxicity and efficacy probability intervals. For a given region A, the JUPM is defined as the ratio between the probability of the region and the size of the region. Considering the two-dimensional unit square (0, 1) × (0, 1) in the real space, the JUPM for each interval combination (a, b) × (c, d) is

\[ \text{JUPM}(c,d)_{(a,b)} = \frac{\Pr\{p_d \in (a,b), q_d \in (c,d) \mid D\}}{(b-a) \times (d-c)}; 0 < a < b < 1; 0 < c < d < 1. \] (1.5)

Here, the numerator, \( \Pr\{p_d \in (a,b), q_d \in (c,d) \mid D\} \), is the posterior probability of \( p_d \) and \( q_d \) falling in the interval \((a, b)\) and \((c, d)\), respectively.

Assume the prior for each \( p_d \) follows an independent \( \text{beta}(a_1, b_1) \), and the prior for each \( q_d \) follows an independent \( \text{beta}(a_2, b_2) \), where \( \text{beta}(\alpha, \beta) \) denotes a beta distribution with mean \( \frac{\alpha}{\alpha+\beta} \). The posterior distributions for \( p_d \) and \( q_d \) are \( \text{beta}(a_1 + x_d, b_1 + n_d - x_d) \) and \( \text{beta}(a_2 + y_d, b_2 + n_d - y_d) \), respectively.

Based on the posterior distributions, there exists a winning interval combination \((a^*, b^*) \times (c^*, d^*)\) that achieves the maximum JUPM among all the combinations in Table 1.4, and the corresponding decision for that combination is selected for treating the next cohort of patients.

The basic dose-finding concept of TEPI is as follows. Assume that the current patient cohort is treated at dose \( d \). After the current cohort completes DLT and response evaluation, compute the JUPMs for all the interval combinations in Table 1.4. The TEPI design recommends E,” S,” or D”, Your trial designs anywhere, anytime.
corresponding to the combination with the largest JUPM value according to Table 1.4.

In practice, the TEPI design needs to be calibrated according to physicians’ needs. This is transparent and requires some effort. The tuning is for the intervals in Table 1.4 so that the dosing decisions are satisfactory to the clinicians.

To enable ethical constraints, below are two additional rules as part of the dose-finding algorithm to exclude any dose with excessive toxicity and any dose with unacceptable efficacy.

Safety and futility rules

- Safety rule: if $\Pr(p_d > p_T | x_d, n_d) > p_{cut}$ for a $p_{cut}$ close to 1 (say, 0.95), exclude doses $d, d+1, \cdots, D$, from future use in the trial; treat the next cohort of patients at dose $(d-1)$.

- Futility rule: if $\Pr(q_d < q_{E} | y_d, n_d) > q_{cut}$ for a $q_{cut}$ close to 1 (say, 0.7), where $q_{E}$ is the minimum acceptable probability of efficacy, then exclude dose $d$ from future use in the trial.

Here, $q_{E}$ is the reference efficacy rate, e.g., the efficacy rate of standard care.

Note that, here we assume the prior for each $p_d$ follows an independent $\text{beta}(a_1, b_1)$, and the prior for each $q_d$ follows an independent $\text{beta}(a_2, b_2)$, where $\text{beta}(\alpha, \beta)$ denotes a beta distribution with mean $\alpha/(\alpha + \beta)$. The posterior distributions for $p_d$ and $q_d$ in the above rules are $\text{beta}(a_1 + x_d, b_1 + n_d - x_d)$ and $\text{beta}(a_2 + y_d, b_2 + n_d - y_d)$, respectively.

Stopping rules

The trial is stopped if

1. the prespecified maximum total sample size $n$ is reached; or
2. the lowest dose shows excessive toxicity according to the safety rule; In this case, the trial is early stopped and the MTD cannot be determined; or
3. optional:
   - the TEPI decision is “S”, to stay at the current dose, and the current dose has enrolled $K$ patients;
   - the TEPI decision is “E”, to escalate to the next higher dose, and that next higher dose has enrolled $K$ patients;
   - the TEPI decision is “D”, to de-escalate to the previous lower dose, and that previous lower dose has enrolled $K$ patients.
1.3.3. Dose Selection

At the end of the trial, TEPI selects the most desirable dose as the OBD based on a utility score that balances the toxicity and efficacy trade-off. The utility score function is defined as $U(p, q) = f_1(p)f_2(q)$ (suppressing dose $d$ in the notation), where $p$ denotes the toxicity rate, and $q$ denotes the efficacy rate.

Both $f_1(\cdot)$ and $f_2(\cdot)$ are truncated linear functions, given by

$$f_1(p) = \begin{cases} 
1, & p \in (0, p_1^*], \\
1 - \frac{p-p_1^*}{p_2^*-p_1^*}, & p \in (p_1^*, p_2^*), \\
0, & p \in [p_2^*, 1) 
\end{cases} \quad (1.6)$$

$$f_2(q) = \begin{cases} 
0, & q \in (0, q_1^*]. \\
\frac{q-q_1^*}{q_2^*-q_1^*}, & q \in (q_1^*, q_2^*), \\
1, & q \in [q_2^*, 1) 
\end{cases} \quad (1.7)$$

where $p^*$'s and $q^*$'s are prespecified cutoff values. The OBD is selected according to the following process.

1. We generate a total of $T$ random samples, $\{p_d^{(t)}, t = 1, \cdots, T\}$ and $\{q_d^{(t)}, t = 1, \cdots, T\}$, from the posterior distributions $\text{beta}(a_0 + x_d, b_0 + n_d - x_d)$ and $\text{beta}(a_0 + y_d, b_0 + n_d - y_d)$ for each dose $d$, respectively. Here, U-Design sets $a_0 = b_0 = 0.005$ and $T = 1000$.

2. For toxicity probabilities of all doses in each sample $t$, $p^{(t)} = (p_d^{(t)}, \cdots, p_D^{(t)})$, we perform isotonic transformation using the pool adjacent violators algorithm (PAVA; Mair et al. 2009) on $p^{(t)}$ to obtain $\tilde{p}^{(t)} = (\tilde{p}_1^{(t)}, \cdots, \tilde{p}_d^{(t)})$, where $\tilde{p}_i^{(t)} \leq \tilde{p}_j^{(t)}$ if $i < j$.

3. Let $\mathcal{A} = \{d \mid n_d > 0, d = 1, \cdots, D\}$ denote the candidate dose set from which doses have been excluded according to safety and futility rules, the final selected dose $d^*$ is the one that maximizes utility scores $U(p_d, q_d)$, that is, $d^* = \arg\max_{d \in \mathcal{A}} E[U(p_d, q_d) \mid \text{Data}]$, where

$$\hat{E}[U(p_d, q_d) \mid \text{Data}] = \frac{1}{T} \sum_{t=1}^{T} U(p_d^{(t)}, q_d^{(t)}).$$
Module 1. Single-Agent Dose-Finding Designs with Efficacy & Toxicity Endpoints and Cohort Enrollment

1.3.4 The Efficacy-Toxicity (EffTox) Trade-Offs-Based Design

EffTox, proposed in Thall and Cook (2004), is an outcome-adaptive, model-based Bayesian procedure that chooses doses of an experimental agent for successive patient cohorts in a clinical trial based on both efficacy (E) and toxicity (T) outcomes. EffTox models the dose-efficacy and dose-toxicity relationship respectively using two different dose-response curves. Based on accumulating efficacy and toxicity data over the trial, EffTox continuously updates the parameters of the dose-response models. The desirability of each dose $x$ is evaluated by using a family of contours characterizing the trade-off between E and T, and patients are assigned to the most desirable dose in cohorts.

1.3.4.1 Dose-Outcome Models

Assume $D$ dose $s_1, \cdots, s_D$ to be considered in the trial, and code dose as

$$x_d = \log(s_d) - D^{-1} \sum_{k=1}^{D} \log(s_k)$$

for use in the regression models. If $0 = s_1 < s_2$, first add $s_2$ to each $s_d$ before taking logs. Let $\pi(x, \theta) = \{\pi_E(x, \theta), \pi_T(x, \theta)\}$ be the probabilities of efficacy and toxicity, where $x$ denotes dose and $\theta$ is the model parameter vector.

Given the current interim trial data $D$, define $x$ to be an acceptable dose if

$$\Pr\{\pi_E(x, \theta) > q_E | D\} > 1 - q_{cut}$$

and

$$\Pr\{\pi_T(x, \theta) < p_T | D\} > 1 - p_{cut},$$

where $q_E$ and $p_T$ are fixed lower and upper limits specified by the physician, and $q_{cut}$ and $p_{cut}$ are fixed probability cutoffs.

For toxicity, assume $\text{logit}(\pi_T(x, \theta)) = \mu_T + x\beta_T$, in which we set $\beta_T > 0$ to meet the monotonic dose-toxicity assumption. For efficacy, to allow a wide variety of possible doseresponse relationships, assume $\text{logit}(\pi_E(x, \theta)) = \mu_E + x\beta_{E,1} + x^2\beta_{E,2}$. For simplicity, temporarily suppress $(x, \theta)$. The joint outcome model is given by

$$\pi_{a,b} = (\pi_E)^a(1-\pi_E)^{1-a}(\pi_T)^b(1-\pi_T)^{1-b} + (-1)^{a+b}\pi_E(1-\pi_E)\pi_T(1-\pi_T)(\frac{e^\psi - 1}{e^\psi + 1})$$

for $a, b \in \{0, 1\}$ and real-valued $\psi$. Thus, $\theta = (\mu_T, \beta_T, \mu_E, \beta_{E,1}, \beta_{E,2}, \psi)$. Since $\beta_T$ should be greater than 0, we assume that $\beta_T$ is lognormally distributed, with mean $\tilde{\mu}_{\beta_T}$ and standard deviation...
\[ \sigma_{\beta_T}. \] Except for \( \beta_T \), we assume that each component \( \theta_i \) of \( \theta \) is normally distributed with mean \( \mu_i \) and standard deviation \( \sigma_i \), denoted as \( \theta_i \sim N(\mu_i, \sigma_i) \).

The likelihood for a single patient treated at dose \( x \) is
\[ L(Y_i, x \mid \theta) = \prod_{a=0}^{1} \prod_{b=0}^{1} \left\{ \pi_{a,b}(x, \theta) \right\}^{I(Y=(a,b))}. \]

Denoting the data for the first \( n \) patients in the trial by \( D_n \), for \( 1 \leq n \leq N \), the likelihood is
\[ L_n(D_n \mid \theta) = \prod_{i=1}^{n} L(Y_i, x(i) \mid \theta), \]
where \( Y_i \) and \( x(i) \) denote the \( i \)th patient’s outcome and dose.

### 1.3.4.2 EfficacyToxicity Trade-Off Contours

To determine the desirability of each dose, the EffTox design constructs a efficacy-toxicity desirability contour, \( C \), in the two-dimensional domain \( \Pi = [0, 1]^2 \) by fitting a curve to target values of \( \pi \) elicited from the physician. The contour \( C \) is then used to construct a family of desirability contours such that all \( \pi \) on the same contour are equally desirable. Because the family of contours partitions \( \Pi \), this construction provides a basis for comparing doses in terms of their posterior means, \( E\{\pi(x, \theta) \mid D\} \).

To construct \( C \), we first elicit three target values, \( \{\pi_1, \pi_2, \pi_3\} \), which the physician considers equally desirable. First, elicit a desirable trade-off target, \( \pi_1^* = (\pi_{1,E}, \pi_{1,T}) = (\pi_{1,E}^*, 0) \), in the case where toxicity has probability 0. That is, elicit the smallest efficacy probability, \( \pi_{1,E}^* \), that the physician would consider desirable if toxicity were impossible. Next, elicit \( \pi_2^* \) having the same desirability as \( \pi_1^* \) by asking the physician what the maximum value of \( \pi_T \) may be if \( \pi_E = 1 \). Given these two equally desirable extremes, elicit a third pair, \( \pi_3^* \), that is equally desirable but is intermediate between \( \pi_1^* \) and \( \pi_2^* \).

The desirability function of \( (\pi_E, \pi_T) \in [0, 1]^2 \) is defined to be
\[
\delta(\pi_E, \pi_T) \equiv 1 - \|((\pi_E, \pi_T) - (1, 0))\|_p
\]
\[
= 1 - \left\{ \left( \frac{\pi_E - 1}{\pi_{1,E}^* - 1} \right)^p + \left( \frac{\pi_T - 0}{\pi_{2,T}^* - 0} \right)^p \right\}^{1/p} \tag{1.12}
\]
where \( p > 0 \). Solve \( \delta(\pi_{E,3}, \pi_{T,3}) = 0 \) for \( p \) using the bisection method, wherein intervals known to bracket the solution are successively refined (Peter et al., 2014). This gives \( \delta(\pi) = 0 \) on \( C \) with \( \delta(\pi) \) increasing as \( \pi \) moves along any straight line from a point in \( [0, 1]^2 \) to the ideal pair \( (\pi_E, \pi_T) = (1, 0) \). After solving for \( p \), the desirability measure can be computed for any point \( (\pi_E, \pi_T) \) using formula (1.12).

The following definition exploits this structure to induce an ordering on the set of doses.

**DEFINITION:** Given \( D \) and \( x \), the desirability, \( \delta(x, D) \), of \( x \) is the desirability of the posterior mean \( E\{\pi(x, \theta) \mid D\} \).
To apply this during the trial, after the most recent cohorts data have been incorporated into $\mathcal{D}$, for each $x, (\pi_E, \pi_T) = E\{\pi(x, \theta) \mid \mathcal{D}\}$ is first computed, and then the desirability of $x$ is computed by formula (1.12). Among the doses with acceptable efficacy and toxicity, the dose that maximizes $\delta(x, \mathcal{D})$ is selected.

1.3.4.3 The Trade-Off-Based Algorithm

Initially, the physician must provide a set of doses, a starting dose for the first cohort, $N$, $c$, and the limits $q_E$ and $p_T$ used in the acceptability criteria (1.9) and (1.10). The trade-off targets $\{\pi_1^*, \pi_2^*, \pi_3^*\}$ then must be elicited in order to construct $\mathcal{C}$ and the family of trade-off contours. The probability cut-offs $q_{cut}$ and $p_{cut}$ in (1.9) and (1.10) are determined, using preliminary computer simulation results, to obtain a design with desirable operating characteristics. Given this structure, the dose-finding algorithm proceeds as follows:

1. Treat the first cohort at the starting dose specified by the physician.
2. For each cohort after the first, $x \in A(\mathcal{D})$ if $x$ satisfies both (1.9) and (1.10), or if $x$ is the lowest untried dose above the starting dose and it satisfies (1.10).
3. If $A(\mathcal{D}) \neq \emptyset$, then the next cohort is treated at the most desirable $x \in A(\mathcal{D})$, subject to the constraint that no untried dose may be skipped when escalating.
4. If $A(\mathcal{D}) = \emptyset$, then the trial is terminated and no dose is selected.
5. If the trial is not stopped early and $A(\mathcal{D}_N) \neq \emptyset$ at the end of the trial, then the dose $x \in \mathcal{D}_N$ is selected.

Figure 1.20: Example of efficacy-toxicity desirability contours. The contour $\mathcal{C}$ is the line with desirability equals to 0 ($U = 0.0$).
A(\mathcal{D}_N) maximizing \delta(x, \mathcal{D}_N) is selected.
1.3.5 The Utility-Based Bayesian Optimal Interval (U-BOIN) Design

U-BOIN (Zhou et al., 2019) is a model-based design that jointly models toxicity and efficacy using a multinomial-Dirichlet model and employs a utility function to measure dose risk-benefit trade-off. The design consists of two seamless stages. In stage I, the Bayesian optimal interval (BOIN) design (Liu and Yuan, 2015) is used to quickly explore the dose space and collect preliminary toxicity and efficacy data. In stage II, the posterior estimate of the utility for each dose is continuously updated using accumulating efficacy and toxicity data, and the posterior estimate is used to direct patient allocation and OBD selection.

1.3.5.1 Efficacy-Toxicity Model

Consider a phase I/II trial with \( J \) doses under investigation. Let \( Y_E \) denote the binary efficacy endpoint, where \( Y_E = 1 \) denotes response, and 0 otherwise; let \( Y_T \) denote the binary toxicity endpoint, where \( Y_T = 1 \) denotes DLT, and 0 otherwise. The bivariate discrete outcome \((Y_E, Y_T)\) can be equivalently represented by a single variable \( Y \) with \( 2 \times 2 = 4 \) levels, with \( Y = 1 \) if \((Y_E, Y_T) = (0, 1)\); \( Y = 2 \) if \((Y_E, Y_T) = (0, 0)\); \( Y = 3 \) if \((Y_E, Y_T) = (1, 1)\); and \( Y = 4 \) if \((Y_E, Y_T) = (1, 0)\). Here \( Y = 1 \) is the least favorable clinical outcome (DLT, no efficacy), and \( Y = 4 \) denotes the most favorable clinical outcome (No DLT, efficacy).

Define \( \pi_{jk} = \Pr(Y = k | d = j) \), \( k = 1, \cdots, 4 \) and \( j = 1, \cdots, J \), with \( \sum_{k=1}^{4} \pi_{jk} = 1 \), where \( d \) denotes the dose level. Assume that \( Y \) follows a Dirichlet-multinomial model as follows:

\[
Y = k | d = j \sim \text{Multinomial}(\pi_{j1}, \cdots, \pi_{j4}) \tag{1.13}
\]

\[
(\pi_{j1}, \cdots, \pi_{j4}) \sim \text{Dirichlet}(a_1, \cdots, a_4) \tag{1.14}
\]

where \( a_1, \cdots, a_4 > 0 \) are hyperparameters. U-Design sets \( a_k = \frac{1}{4} \), \( k = 1, \cdots, 4 \), as the default values, such that the prior is vague and equivalent to an effective sample size of 1.

Assume that \( n_j \) patients have been treated at dose \( d = j \), among whom \( n_{jk} \) patients had outcome \( Y = k \), where \( n_j = \sum_{k=1}^{4} n_{jk} \). Denote \( D_j = (n_{j1}, \cdots, n_{j4}) \), and the posterior distribution of \( \pi_j = (\pi_{j1}, \cdots, \pi_{j4}) \) is

\[
\pi_j | D_j \sim \text{Dirichlet}(a_1 + n_{j1}, \cdots, a_4 + n_{j4}). \tag{1.15}
\]

1.3.5.2 Utility

Let \( \psi_k \) denote the utility value ascribed to outcome \( Y = k \), \( k = 1, \cdots, 4 \), which can be elicited from physicians to reflect the risk-benefit trade-off underlying their medical decisions using the following
1.3. Statistical Methods Review

1.3.5. The Utility-Based Bayesian Optimal Interval (U-BOIN) Design

procedures.

- Fix the value of the utility for the least desirable outcome $Y = 1$ as $\psi_1 = 0$, and for the most desirable outcome $Y = 4$ as $\psi_4 = 1$.

- Ask the clinician to use these two utilities as a reference to score the utility values $\psi_2, \psi_3$ for the other 2 possible outcomes $Y = 2, 3$ to quantify the risk-benefit trade-off under each outcome.

Table 1.5 shows two examples of the utility function.

**Table 1.5:** Examples of utility.

<table>
<thead>
<tr>
<th>(a) Example 1</th>
<th>(b) Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_T = 1$</td>
<td>$Y_T = 1$</td>
</tr>
<tr>
<td>$Y_E = 0$</td>
<td>$Y_E = 0$</td>
</tr>
<tr>
<td>$\psi_1 = 0$</td>
<td>$\psi_1 = 0$</td>
</tr>
<tr>
<td>$\psi_2 = 0.3$</td>
<td>$\psi_2 = 0.3$</td>
</tr>
<tr>
<td>$Y_E = 1$</td>
<td>$Y_E = 1$</td>
</tr>
<tr>
<td>$\psi_3 = 0.5$</td>
<td>$\psi_3 = 0.65$</td>
</tr>
<tr>
<td>$\psi_4 = 1$</td>
<td>$\psi_4 = 1$</td>
</tr>
</tbody>
</table>

Example 1 has utility values $\{\psi_1 = 0, \psi_2 = 0.3, \psi_3 = 0.5, \psi_4 = 1\}$ for the outcomes $\{(Y_E = 0, Y_T = 1), (Y_E = 0, Y_T = 0), (Y_E = 1, Y_T = 1), (Y_E = 1, Y_T = 0)\}$, respectively. Compared to example 1, example 2 rewards the response (i.e., $Y_E = 1$) more, in the presence of DLT (i.e., $Y_T = 1$), by assigning a larger value to $\psi_3$ (0.65 versus 0.50). This is appropriate for a trial where toxicity can be well managed and efficacy response is highly desirable (e.g., leading to long survival).

Given the values of $\psi_k$, the true mean utility for dose $j$ is given by

$$U_j = \sum_{k=1}^{4} \psi_k \pi_{jk}.$$  \hspace{1cm} (1.16)

Since the true mean utility $U_j$ depends on $\pi_{jk}$, which is unknown, it is estimated based on the observed data. Given the interim data $D = \{D_j\}$, the estimate of mean utility is given by

$$\hat{U}_j = \sum_{k=1}^{4} \psi_k E(\pi_{jk} | D).$$  \hspace{1cm} (1.17)

**1.3.5.3 Optimal Biological Dose**

Let $p_T$ denote the maximum tolerable DLT rate, and $q_E$ the lowest acceptable response rate. Let $\pi_{T,j} = \pi_{j1} + \pi_{j3} = Pr(Y_T = 1 | d = j)$ and $\pi_{E,j} = \pi_{j3} + \pi_{j4} = Pr(Y_E = 1 | d = j)$. Define
that dose $j$ is inadmissible, if it meets either one or both of the following two criteria:

$$\Pr(\pi_{T,j} > p_T | D) > p_{cut}$$  \hspace{1cm} (1.18)

$$\Pr(\pi_{E,j} < q_E | D) > q_{cut}$$  \hspace{1cm} (1.19)

where $p_{cut}$ and $q_{cut}$ are probability cutoffs. According to (1.13) and (1.14), $\pi_{T,j}$ and $\pi_{E,j}$ follow posterior beta distributions, given by

$$\pi_{T,j} | D \sim \text{Beta}(a_1 + a_3 + n_{j1} + n_{j3}, a_2 + a_4 + n_{j2} + n_{j4}),$$

$$\pi_{E,j} | D \sim \text{Beta}(a_3 + a_4 + n_{j3} + n_{j4}, a_1 + a_2 + n_{j1} + n_{j2}).$$

The admissible dose is then defined as the dose for which none of the criteria (1.18) and (1.19) is satisfied. Define the OBD as the dose that is admissible and has the highest utility value, i.e.,

$$\text{OBD} = \arg \max_{j \in A} (U_j)$$  \hspace{1cm} (1.20)

where $A$ denotes the set of admissible doses.

### 1.3.5.4 Dose-finding Algorithm

The U-BOIN design consists of two seamless stages (Figure 1.21). The objective of stage I is to quickly explore the dose space to identify a set of admissible doses that are reasonably efficacious and safe for stage II. In stage I, dose escalation is conducted based on only the toxicity outcome. However, efficacy data are also collected and will be used for decision making in stage II. Stage I dose escalation/de-escalation is guided by the BOIN design (Liu and Yuan, 2015). Due to very limited data and large uncertainty, for patient safety, set the target DLT rate $\phi_T = p_T - 0.05$, slightly lower than the maximum tolerable DLT rate $p_T$, to ensure that stage I dose exploration concentrates around up to, but not exceeding $p_T$. Let $\hat{\pi}_{T,j}$ denote the empirical (or maximum likelihood) estimate of $\pi_{T,j}$, given by $\hat{\pi}_{T,j} = \frac{m_j}{n_j}$ where $m_j$ is the number of patients who experienced DLT at the dose level $j$; and let $\lambda_e$ and $\lambda_d$ denote the predetermined optimal escalation boundary and de-escalation boundary. Table 1.6 provides the values of $\lambda_e$ and $\lambda_d$ for the commonly used target DLT rate $\phi_T$. See the work of Liu and Yuan (2015) for the derivation and formula to calculate $\lambda_e$ and $\lambda_d$. The dose-finding algorithm in stage I proceeds as follows.

Ia. Patients in the first cohort are treated at dose level 1 or a prespecified starting dose.

Ib. Suppose $j$ is the current dose; use the following rules to assign a dose to the next cohort of patients:
1.3.5. The Utility-Based Bayesian Optimal Interval (U-BOIN) Design

- Escalate the dose to \( j + 1 \) if \( \hat{\pi}_{T,j} \leq \lambda_e \).
- De-escalate the dose to \( j - 1 \) if \( \hat{\pi}_{T,j} \geq \lambda_d \).
- Otherwise, stay at the current dose \( j \).

Ic. Repeat step Ib until the number of patients treated on one of the doses reaches \( s_1 \), and then move to stage II.

In stage I, following the BOIN design, if \( \Pr(\pi_{T,j} > p_T \mid m_j, n_j) > 0.95 \) and \( n_j \geq 3 \), dose level \( j \) and higher are eliminated from the trial; the trial is terminated if the lowest dose level is eliminated, where \( \Pr(\pi_{T,j} \geq p_T \mid m_j, n_j) > 0.95 \) is evaluated based on a beta-binomial model with the uniform prior.

Stage II proceeds as follows.

IIa. Let \( j^* \) denote the highest dose level that has been tried. If \( \hat{\pi}_{T,j^*} \leq \lambda_e \) and \( j^* \) is not the highest dose in the trial, escalate the dose to \( (j^* + 1) \) for treating the next cohort of patients; otherwise, proceed to step IIb.

IIb. Given the observed interim data \( D \) collected in both stages I and II, determine the admissible dose set \( A \) from dose 1, \( \ldots \), \( j^* \), where none of the criteria (1.18) and (1.19) is satisfied for each dose in \( A \). If no dose is admissible, terminate the trial and no dose should be selected as the OBD. Otherwise, assign the next cohort of patients to a dose in \( A \). In U-Design, there are two methods to assign the next cohort,

- Pick The Winner, assigning to dose \( j \in A \) that has the largest posterior mean utility.
- Adaptive Randomization, adaptively randomizing the next cohort of patients to dose \( j \in A \), with probability \( \omega_j \) proportional to its posterior mean utility, i.e.,

\[
\omega_j = \frac{U_j}{\sum_{j \in A} U_j}.
\]

IIc. Repeat steps IIa and IIb until reaching the prespecified maximum sample size \( N \) or the number of patients treated at one of the doses in stage II reach \( s_2 \) (Zhou et al. (2019) recommends that \( s_2 > s_1 \)), and then select the OBD following the rules in §1.3.5.3.
Table 1.6: Dose escalation and de-escalation boundaries of the Bayesian optimal interval design

<table>
<thead>
<tr>
<th>Boundaries</th>
<th>Target DLT rate ((\phi_T))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>(\lambda_e) (escalation)</td>
<td>0.118</td>
</tr>
<tr>
<td>(\lambda_d) (de-escalation)</td>
<td>0.179</td>
</tr>
</tbody>
</table>

Figure 1.21: Diagram of the utility-based Bayesian optimal interval (U-BOIN) design.
Reference


Lin, X. and Ji, Y. (2020b). The joint i3+3 (Ji3+3) design for phase I/II adoptive cell therapy clinical trials. *Contemporary Clinical Trials*, In press.


Reference

