Supplementary Material: Learning for Dose Allocation in Adaptive Clinical Trials with Safety Constraints

Cong Shen, Zhiyang Wang, Sofía S. Villar, Mihaela van der Schaar

A. Preliminaries

Before presenting the technical proofs, we introduce some notations and regularity assumptions on the dose-toxicity model, which can be verified to hold for Eqn. (1). For a general toxicity function $p_k(a)$ of an unknown parameter $a \in \mathcal{A}$, the following regularities are imposed:

Assumption 2 *1) **Monotonicity:** For each $k \in \mathcal{K}$ and $a, a' \in \mathcal{A}$ there exists $C_{1,k} > 0$ and $1 < \gamma_{1,k}$, such that $|p_k(a) - p_k(a')| \geq C_{1,k}|a - a'|^{\gamma_{1,k}}$.*

*2) **Hölder continuity:** For each $k \in \mathcal{K}$ and $a, a' \in \mathcal{A}$ there exists $C_{2,k} > 0$ and $0 < \gamma_{2,k} \leq 1$, such that $|p_k(a) - p_k(a')| \leq C_{2,k}|a - a'|^{\gamma_{2,k}}$.*

We note that both monotonicity and continuity assumptions are mild and standard in the literature; see (Wang et al., 2018). Proposition 1 immediately follows with Assumption 2.

Proposition 1 *For functions $p_k(a), \forall k \in \mathcal{K}$ that satisfy Assumption 2, we have:*

1) $p_k(a)$ is invertible;

2) For each $k \in \mathcal{K}$ and $d, d' \in \mathcal{P}$, we have $|p_k^{-1}(d) - p_k^{-1}(d')| \leq \bar{C}_{1,k}|d - d'|^{\bar{\gamma}_{1,k}}$, where $\bar{\gamma}_{1,k} = \frac{1}{\gamma_{1,k}}$, $\bar{C}_{1,k} = (\frac{1}{C_{1,k}})^{\frac{1}{\gamma_{1,k}}}$.

For ease of exposition, we denote $C_1 = \min C_{1,k}$, $C_2 = \max C_{2,k}$, $\gamma_1 = \max \gamma_{1,k}$, $\gamma_2 = \min \gamma_{2,k}$, $\bar{\gamma}_1 = 1/\gamma_1$, and $\bar{C}_1 = C_1^{-\bar{\gamma}_1}$.

B. Select Design Parameters

The parameters appeared in Assumption 2 collectively determine the confidence interval in Eqn. (3). We take function (1) as an example to show how to select these parameters. We have

$$\begin{aligned} |p_k(a) - p_k(a')| &\geq C_{1,k}|a - a'|^{\gamma_{1,k}}, \\ \frac{|p_k(a) - p_k(a')|}{|a - a'|} &\geq C_{1,k}|a - a'|^{\gamma_{1,k}-1}, \\ \min_{a \in \mathcal{A}} p_k'(a) &\geq C_{1,k}|\mathcal{A}|^{\gamma_{1,k}-1}, \\ \log\left(\frac{\tanh(d_k) + 1}{2}\right) &\geq C_{1,k}|\mathcal{A}|^{\gamma_{1,k}-1}. \end{aligned}$$

Therefore, we can first set $\gamma_{1,k}$ as $\frac{3}{2}$ and find the corresponding $C_{1,k}$. Then, with the known function $p_k(a)$, parameters can be approximately calculated.

C. Proof of Lemma 1

$$\begin{aligned}
\mathbb{P}[\hat{a}(t)+\alpha(t)<p_i^{-1}(\theta)] &\leq \mathbb{P}[\hat{a}(t)+\alpha(t)<a^*] \\
&\leq \mathbb{P}[|a^*-\hat{a}(t)|>\alpha(t)] \\
&\leq \mathbb{P}\left[\sum_{k=1}^{K} w_k(t-1)\bar{C}_1|\hat{p}_k(t)-p_k(a^*)|^{\bar{\gamma}_1}>\alpha(t)\right] \\
&\leq \sum_{k=1}^{K}\mathbb{P}\left[|\hat{p}_k(t)-p_k(a^*)|>\left(\frac{\alpha(t)}{w_k(t-1)\bar{C}_1K}\right)^{\gamma_1}\right] \\
&\leq \sum_{k=1}^{K} 2\exp\left(-2N_k(t)\left(\frac{\alpha(t)}{w_k(t)\bar{C}_1K}\right)^{2\gamma_1}\right) \qquad (11) \\
&\leq 2K\exp\left(-2\left(\frac{\alpha(t)}{\bar{C}_1K}\right)^{2\gamma_1}t\right)=\delta. \qquad (12)
\end{aligned}$$

Inequality (11) is from the Hoeffding's inequality and (12) is derived from the definition of $N_k(t)=tw_k(t)$ and Assumption 2 with $\gamma_1>1$.

D. Proof of Lemma 2

From the Hoeffding's Inequality and Eqn. (6), we have:

$$\alpha(t)\leq p_k^{-1}(\theta)-a^*-\epsilon=\Delta_k-\epsilon,$$

where $\Delta_k=|a^*-p_k^{-1}(\theta)|$ denotes the gap between the true value of parameter a and the parameter corresponding to when the toxicity of dose level d_k is exactly at the MTD threshold θ. When $t>t_1$ and with the definition of $\alpha(t)$ in Eqn. (3), the lemma can be immediately derived.

E. Proof of Theorem 1

Depending on whether the optimal dose level is included in the admissible set or not, we can decompose the regret into two parts:

$$\begin{aligned}
R(n) &= \sum_{t=1}^{n}\mathbb{P}[k^*\notin\mathcal{D}_1(t)]Q+\mathbb{P}[k^*\in\mathcal{D}_1(t)]R_2(n) \\
&\leq n\delta Q+R_2(n).
\end{aligned}$$

The probability of the first error event $\{k^*\notin\mathcal{D}_1(t)\}$ can be bounded by Lemma 1, which indicates that at each step t the probability of a safe dose level being excluded from the admissible set is bounded by δ. For the second part, $R_2(n)$ represents the regret when the optimal dose is included in the admissible set. In this case, the error event is due to the inaccuracy of parameter estimation at the beginning as well as the limited efficacy information provided by each sample. Using Lemma 2, we have:

$$\begin{aligned}
R_2(n) &\leq t_1+(K-M)\sum_{t=1}^{n}\exp(-2t\epsilon^2)+\sum_{t=t_1+1}^{n}\sum_{d_k:p_k\leq\theta}\mathbb{1}\{I(t)=k\} \\
&\leq t_1+\frac{K-M}{2\epsilon^2}+\sum_{d_k:p_k\leq\theta}\frac{c\log(n)}{q^*-q_k}.
\end{aligned}$$

Putting the regret from both error events together leads to (7), which completes the proof.

F. Proof of Theorem 2

First we note:

$$\begin{aligned} p_{I(t)}(a^*) - \theta &\leq p_{I(t)}(a^*) - \theta + \theta - p_{I(t)}(a^* - \alpha(t)) \\ &\leq C_2 |a^* - \hat{a}(t) + \alpha(t)|^{\gamma_2}. \end{aligned}$$

Thus, the probability can be upper bounded as:

$$\mathbb{P}[\hat{a}(t) - a^* > \alpha(t) + \epsilon] \leq \exp(-2t(\alpha(t) + \epsilon)^2).$$

Reorganizing the terms, we finally have

$$\mathbb{P}\left[\frac{1}{n}\sum_{t=1}^{n} p_{I(t)}(a^*) - \theta < C_2 \epsilon^{\gamma_2}\right] \geq 1 - \exp(-2t(\alpha(t) + \epsilon)^2) \geq 1 - \delta.$$

G. Proof of Corollary 2

$$\begin{aligned} \mathbb{P}[|\hat{a}(n) - a^*| \geq \Delta_M] &\leq \sum_{k=1}^{K} \mathbb{P}\left[|\hat{p}_k(t) - p_k(a^*)| > \left(\frac{\Delta_M}{w_k(t)\bar{C}_1 K}\right)^{\gamma_1}\right] \\ &\leq \sum_{k=1}^{K} 2\exp\left(-2N_k(n)\left(\frac{\Delta_M}{w_k(t)\bar{C}_1 K}\right)^{2\gamma_1}\right) \\ &\leq 2K\exp\left(-2\left(\frac{\Delta_M}{\bar{C}_1 K}\right)^{2\gamma_1} n\right). \end{aligned}$$

H. Proof of Theorem 3

We first establish Lemma 3, whose proof directly follow Theorem C.1 in (Combes & Proutière, 2014).

Lemma 3 $\mathbb{E}[l_k(n)] = O(\log(\log(n)))$*, for each* $k \neq k^*$.

Then, following the similar proof steps in Theorem 1, we have the bound in (9).

I. Proof of Theorem 4

Since $k^* = \min\{M, N\}$ and $L_1(n)$ and $L_2(n)$ are the estimations for N and M respectively, $\{\hat{d}_r(n) \neq k^*\} \subseteq E_1 \bigcup E_2$, where $E_1 = \{L_1(n) \neq N\}$, $E_2 = \{L_2(n) \neq M\}$. The latter can be bounded by Corollary 2. With the notation $\beta_k(n) = \sqrt{\frac{c\log(n)}{N_k(n)}}$, the probability of E_1 can be bounded as follows:

$$\begin{aligned} \mathbb{P}[L_1(n) < M] &\leq \mathbb{P}[|\hat{q}_N(n) - \hat{q}_{N-1}(n)| \leq \beta_{N-1}(n) + \beta_N(n)] \\ &\leq \mathbb{P}[\hat{q}_{N-1}(n) - q_k + q_N - \hat{q}_N(n) \leq q_N - q_{N-1} - \beta_{N-1}(n) - \beta_N(n)] \\ &\leq 2\exp\left(-2N_{N-1}(n)\left(\frac{q_N - q_{N-1} - \beta_{N-1}(n) - \beta_N(n)}{2}\right)^2\right) \\ &\leq 2\exp\left(-2f(N-1)\log(n)\left(\frac{\Delta_{N-1,N} - \beta_{N-1}(n) - \beta_N(n)}{2}\right)^2\right) \\ &= o\left(n^{-\frac{5}{2}}\right). \end{aligned}$$

Furthermore,

$$\begin{aligned}\mathbb{P}[L_1(n) > M] &\leq \mathbb{P}[|\hat{q}_N(n) - \hat{q}_{N+1}(n)| > \beta_N(n) + \beta_{N+1}(n)] \\ &\leq \mathbb{P}[|\hat{q}_N(n) - q_N| + |q_{N+1} - \hat{q}_{N+1}(n)| > \beta_N(n) + \beta_{N+1}(n)] \\ &\leq \mathbb{P}[|\hat{q}_N(n) - q_N| > \beta_N(n)] + \mathbb{P}[|\hat{q}_{N+1}(n) - q_{N+1}| > \beta_{N+1}(n)] \\ &\leq \frac{2}{n^c}.\end{aligned}$$

Lastly, $f(N-1)$ is the coefficient of the lower bound of $N_{N-1}(n)$, and can be written as (see Theorem 4.1 in (Combes & Proutière, 2014))

$$f(N-1) = \frac{1}{I(q_{N-1}, q_N)}.$$

This completes the proof.

J. Baseline designs in the experiments

The following baseline designs are used for comparison to SEEDA and SEEDA-Plateau in the experiments.

- **KL-UCB** (Garivier & Cappè, 2011): This approach ignores the safety constraint and focuses entirely on efficacy during allocation, as for each patient it allocates the dose level with the highest efficacy index. The efficacy performance for each dose level is characterized by the KL-UCB index. However, at the end of the experiment, a dose level is recommended according to $\hat{d}(n) = \arg\max_{k:\hat{p}_k(n)\leq\theta} \hat{q}_k(n)$, where $\hat{q}_k(n)$ and $\hat{p}_k(n)$ are the last empirical estimations of toxicity and efficacy for dose level d_k. This suggests that safety constraint is considered in recommendation. Accordingly, type I and type II errors are defined as:

$$\begin{aligned}e_1 &= \sum_{k\in\mathcal{K}} \mathbb{1}\{p_k \leq \theta\}\mathbb{1}\{\hat{p}_k(n) > \theta\}, \\ e_2 &= \sum_{k\in\mathcal{K}} \mathbb{1}\{p_k > \theta\}\mathbb{1}\{\hat{p}_k(n) \leq \theta\}.\end{aligned}$$

- **UCB-1** (Auer et al., 2002): The allocation and recommendation rules are similar to KL-UCB above, with the only difference that the dose level with the highest UCB-1 index of efficacy is allocated to the patient.

- **Independent Thompson Sampling (TS)** (Thompson, 1933; Aziz et al., 2019): Toxicity and efficacy are estimated with Bayesian indices:

$$\tilde{p}_k(t) \sim Beta(S_k^p(t) + 1, N_k(t) - S_k^p(t) + 1),$$

and

$$\tilde{q}_k(t) \sim Beta(S_k^q(t) + 1, N_k(t) - S_k^q(t) + 1),$$

where $S_k^p(t)$ counts the number of toxic outcomes of dose level k among the first t patients and $S_k^q(t)$ counts the number of effective responses. The dose with maximum $\tilde{q}_k(t)$ is allocated to the t-th patient and $\hat{d}(n) = \arg\max_{k:\tilde{p}_k(n)\leq\theta} \tilde{q}_k(n)$ is recommended. Definitions of type I and type II errors are slightly modified to:

$$\begin{aligned}e_1 &= \sum_{k\in\mathcal{K}} \mathbb{1}\{p_k \leq \theta\}\mathbb{1}\{\tilde{p}_k(n) > \theta\}, \\ e_2 &= \sum_{k\in\mathcal{K}} \mathbb{1}\{p_k > \theta\}\mathbb{1}\{\tilde{p}_k(n) \leq \theta\}.\end{aligned}$$

- **CRM** (O'Quigley et al., 1990): We here employ the CRM algorithm with the same one-parameter toxicity model in our paper:

$$p_k(a) = \left(\frac{\tanh(d_k) + 1}{2}\right)^a.$$

We choose a typical prior distribution as $a \sim \exp(0.5)$. Therefore, d_k can be solved with $prior_{tox}$ and the prior mean of a. $\pi_t(a)$ denotes the posterior distribution of a after observing the outcomes of the first t patients. The allocation rule is a greedy one:

$$\begin{aligned} I_t^{CRM} &= \arg\min_{k\in\mathcal{K}} |\theta - p_k(\hat{a}(t))|, \\ \hat{a}(t) &= \int_0^\infty a \mathrm{d}\pi_t(a), \end{aligned}$$

where $\hat{a}(t)$ is the posterior mean value. With this estimation, the final recommendation rule can be written as:

$$\hat{d}(n) = \arg\min_{k\in\mathcal{K}} |\theta - p_k(\hat{a}(n))|.$$

- **3+3** (Storer, 1989): The lowest dose is first given to 3 patients. If none reports a toxic outcome, the next lowest dose level is given to the next 3 patients. If there are less than 2 among these 6 patients who report toxic outcome, the next lowest dose level is given to the next 3 patients; otherwise the experiment is stopped and the dose level used before stopping is recommended as MTD.

- **MCRM** (Neuenschwander et al., 2008): This algorithm classifies the probability of toxicity into four categories. For our simulated setting, the categories are set as:

Under-dosing:	$\pi_a(d) \in (0, 0.20]$
Targeted toxicity:	$\pi_a(d) \in (0.20, 0.35]$
Excessive toxicity:	$\pi_a(d) \in (0.35, 0.60]$
Unacceptable toxicity:	$\pi_a(d) \in (0.60, 1.00]$

  The recommendation and the allocation rules are to maximize the probability of targeted toxicity while controlling the probability of excessive or unacceptable toxicity at $P^{thre} = 25\%$. Based on the posterior distribution of the toxicity, the probability that the toxicity falls in the above four categories can be calculated. The probability that it falls in Targeted category is denoted as P_i^t while falls in Excessive and Unacceptable categories as P_i^e. The selection rule is therefore $I_t = \arg\max_{P_i^e \le P^{thre}} P_i^t$.

- **Multi-objective Bandits** (Yahyaa & Manderick, 2015): We implement the Pareto Thompson Sampling algorithm of (Yahyaa & Manderick, 2015) in our experiments. Specifically, after getting the estimations of toxicity and efficacy of each dose from running the Independent TS design, the algorithm computes the Pareto optimal dose level set $\mathcal{I}^*$, which means $\forall i \in \mathcal{I}^*, \forall j \notin \mathcal{I}^*, \tilde{p}_i(t) \le \tilde{p}_j(t)$ or $\tilde{q}_i(t) \ge \tilde{q}_j(t)$.

Other policies designed for MTA, such as MTA-RA, depend on a different truncated two-parameter logistic efficacy model (Riviere et al., 2018). In our setting, the exact efficacy model is assumed to be unknown – we only make the increase-then-plateau assumption.

K. Additional experiment results under the same setting as in Section 5

Due to space limitations, we were not able to include all the experiment results of the setting in Section 5. These additional results are provided here.

In particular, Table 2 only reports the recommendation and allocation percentages for a given $n = 100$. It is of interest to see how these metrics change with n. We plot the mean allocation and recommendation probabilities as a function of n in Fig. 4. It can be seen that SEEDA-Plateau outperforms all other methods across a large range of n.

L. Experiment of a new setting and its comprehensive results

In the main paper, a setting that has the efficacy reaching the maximal value (the optimal dose) before toxicity hits MTD threshold is used. A different setting can be considered when maximum efficacy dose exceeds the MTD threshold. The

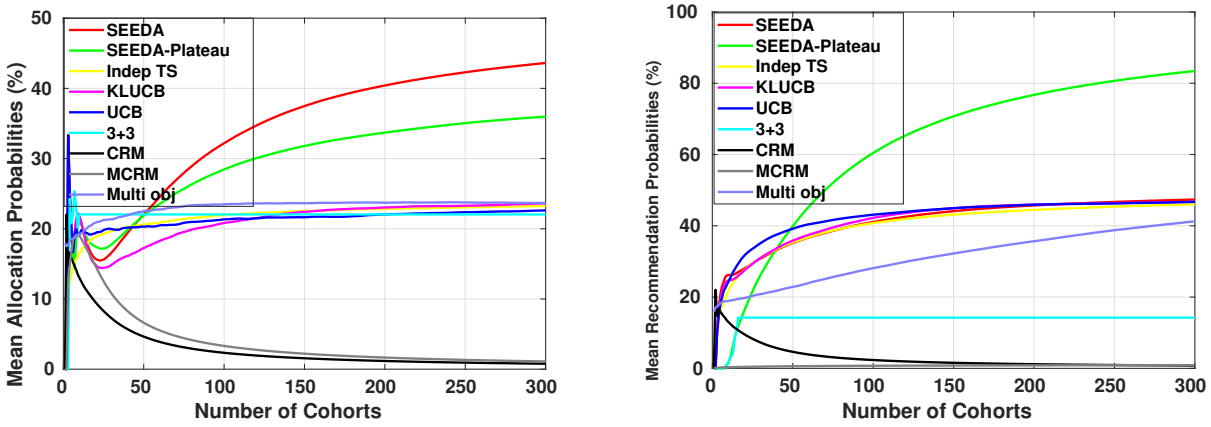

Figure 4: Mean allocation (left) and recommendation (right) probabilities versus number of patients n.

experiment results for this setting (called "setting 2") is reported in this section. Unless otherwise stated, the parameters are the same as in Section 5 of the main paper.

Table 5 presents the setting as well as the allocation and recommendation percentages of each dose for all considered algorithms. For this scenario, dose level 3 is the optimal one. We note that a large portion of the previous conclusions in the main paper still hold. However, the gain of SEEDA-Plateau is less significant over SEEDA, but still outperforms all the comparing designs. The corresponding Type I and Type II error rates are similarly plotted in Fig. 5.

Table 5: Recommendation & allocation percentages of different designs for setting 2.

	Recommended						Allocated					
Toxicity probabilities	0.1	0.2	**0.25**	0.4	0.5	0.6	0.1	0.2	**0.25**	0.4	0.5	0.6
Efficacy probabilities	0.3	0.4	**0.5**	0.7	0.7	0.7	0.3	0.4	**0.5**	0.7	0.7	0.7
SEEDA	9.54 (3.40)	19.34 (10.09)	52.66 (10.43)	16.00 (9.95)	2.12 (1.70)	0 (0)	6.82 (3.34)	17.61 (5.56)	48.99 (9.60)	21.77 (1.07)	3.47 (1.32)	1.33 (0.61)
SEEDA-Plateau	5.15 (3.72)	34.51 (5.96)	53.27 (6.80)	5.84 (2.64)	1.05 (0.50)	0.01 (0)	3.61 (2.28)	11.79 (1.79)	70.30 (7.51)	11.97 (5.12)	2.16 (0.42)	0.17 (0.12)
Independent TS	22.61 (5.61)	22.12 (7.43)	29.05 (8.24)	19.22 (5.96)	4.50 (2.41)	2.50 (2.01)	2.58 (1.90)	3.17 (2.23)	5.56 (3.72)	30.35 (4.73)	32.92 (4.62)	25.43 (3.82)
KL-UCB	19.72 (3.65)	21.03 (4.14)	29.19 (9.27)	24.02 (5.44)	5.46 (1.88)	0.59 (0.38)	2.13 (0.48)	2.50 (0.78)	3.37 (1.35)	32.80 (3.77)	30.63 (8.16)	28.58 (6.99)
UCB	8.95 (3.77)	22.45 (7.99)	41.04 (8.20)	21.61 (3.65)	4.83 (4.56)	1.11 (1.18)	8.12 (0.88)	10.31 (1.13)	13.20 (1.47)	22.90 (2.13)	22.58 (1.75)	22.89 (2.89)
3+3	6.80 (0.12)	20 (13.40)	23.80 (10.24)	29.80 (8.45)	16.40 (5.45)	3.20 (3.12)	26.99 (2.89)	27.50 (3.25)	19.59 (1.45)	13.14 (0.25)	5.01 (1.25)	0.76 (0.75)
CRM	0 (0)	0 (0)	0 (0)	0 (0)	99.10 (0.42)	0.90 (0.36)	0 (0)	0 (0)	0 (0)	0 (0)	99.11 (0.23)	0.89 (0.14)
MCRM	0 (0)	0.60 (0.93)	28.40 (13.29)	67.80 (13.95)	3.20 (3.06)	0 (0)	0.60 (0.09)	0.33 (0.12)	29.17 (9.47)	52.37 (13.95)	11.35 (4.34)	3.18 (1.92)
Multi-obj	6.57 (2.64)	13.38 (8.12)	50.95 (9.92)	22.71 (6.95)	4.44 (1.27)	1.95 (0.55)	20.17 (5.32)	14.78 (2.02)	19.05 (3.95)	20.29 (3.25)	14.57 (5.56)	11.17 (3.58)

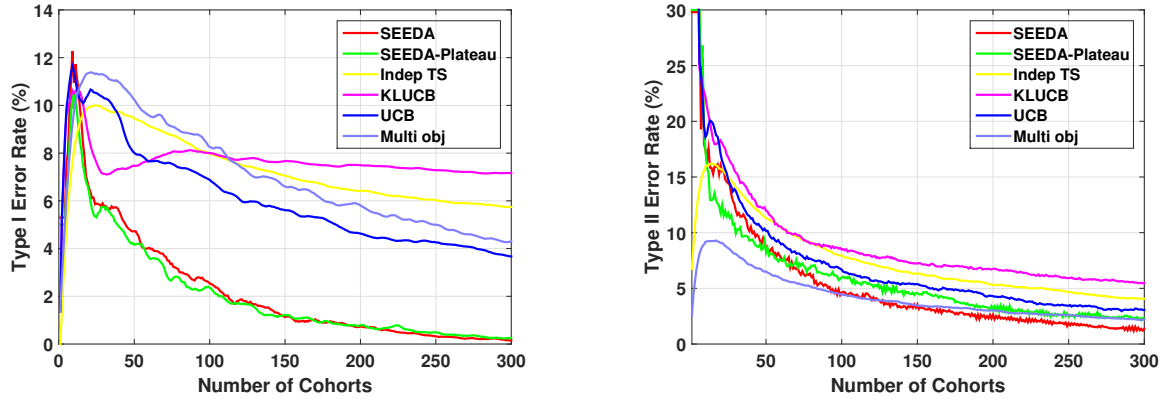

Figure 5: Type I and type II error rates in setting 2.

An in-depth look at the mean allocation and recommendation probabilities versus number of patients n for this new setting is given in Fig. 6. The same observation as in Section K holds.

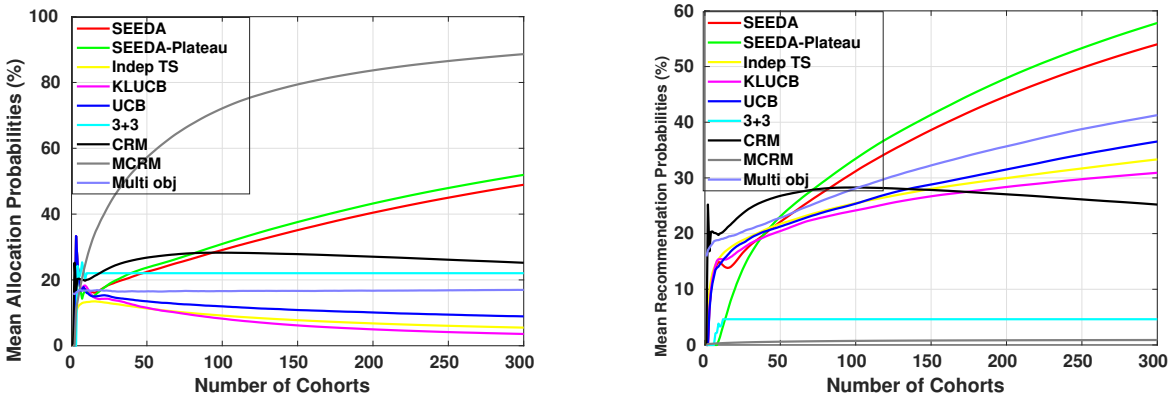

Figure 6: Mean allocation (left) and recommendation (right) probabilities versus number of patients n in setting 2.

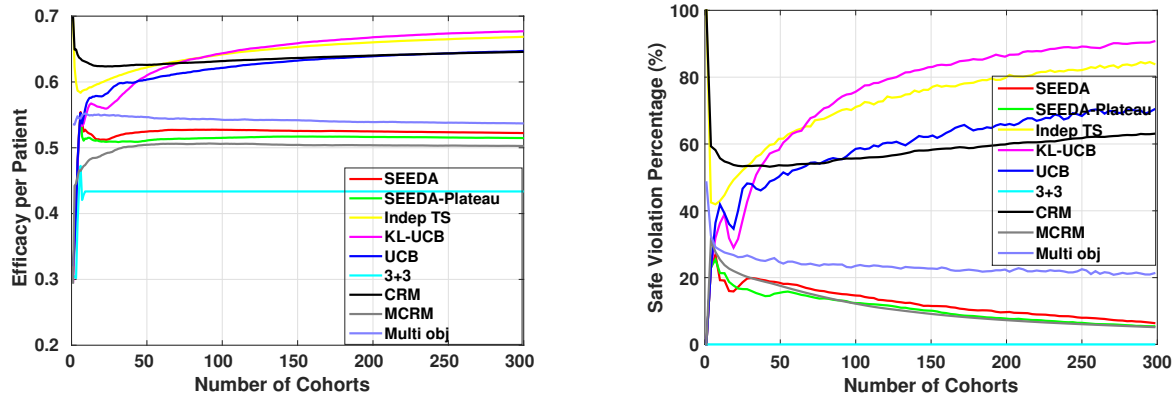

Figure 7: Comparison of efficacy per patient and the safety violation percentage in setting 2.

The convergence of efficacy and toxicity as t increases for setting 2 is plotted in Fig. 7. There is a notable difference to the previous result in Fig. 2, in that now SEEDA and SEEDA-Plateau converge to a different (but correct) dose than the other considered designs, which only emphasize maximum efficacy. It is clear that with such aggressive pursue of efficacy, they succeed in obtaining better treatment effect than SEEDA(-Plateau), but at the significant cost of frequent violation of the safety constraint: as opposed to safety violation percentage hovering between 40% and 50% in Fig. 2, now we face a violation in the range of 70% to 90% as shown in Fig. 7.

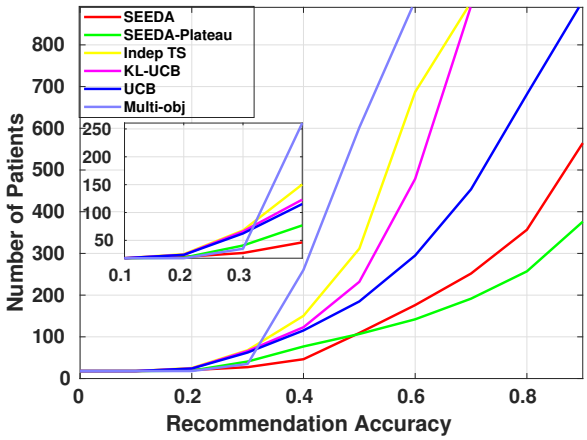

Figure 8: Sample size comparison in setting 2.

Lastly, the sample efficiency is evaluated. Fig. 8 plots the minimum number of patients to achieve a given a recommendation accuracy for different algorithms.

M. Experiment setting 3 to 8 with evaluation of allocation and recommendation percentages

This section reports the allocation and recommendation percentages of each dose for all considered algorithms under different toxicity/efficacy probabilities. We reuse the same 6 scenarios as those in the experiments of (Zang et al., 2014). See Table 6 to 11 for the detailed results. They are in line with the conclusions of the main paper.

Table 6: Recommended & allocated percentages for Scenario 1 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.08	0.12	0.2	**0.3**	0.4	0.08	0.12	0.2	**0.3**	0.4
Efficacy probability	0.2	0.4	0.6	**0.8**	0.55	0.2	0.4	0.6	**0.8**	0.55
SEEDA	2.72 (1.01)	4.88 (2.14)	21.72 (7.50)	69.52 (10.11)	1.16 (0.62)	2.84 (0.78)	4.67 (1.95)	18.55 (6.04)	71.20 (7.65)	2.74 (2.74)
Indep TS	2.34 (0.25)	4.38 (1.31)	12.91 (6.34)	76.83 (7.03)	3.54 (1.49)	1.67 (0.62)	2.99 (0.64)	7.93 (0.36)	81.18 (2.55)	6.23 (2.44)
KL-UCB	9.58 (1.57)	23.99 (3.53)	39.35 (8.10)	24.27 (9.13)	2.81 (2.28)	3.24 (0.34)	13.89 (0.51)	30.91 (1.64)	22.35 (2.14)	29.61 (1.12)
UCB	3.04 (0.91)	12.41 (3.11)	46.91 (8.68)	35.24 (7.68)	2.40 (1.99)	10.91 (0.72)	18.41 (1.31)	33.34 (2.10)	28.32 (2.67)	9.02 (1.85)
3+3	4 (2.65)	10.40 (4.73)	20 (5.94)	22.80 (2.73)	42.80 (6.95)	23.38 (5.79)	22.81 (1.22)	20.92 (4.63)	15.80 (2.14)	10.79 (1.26)
CRM	0.09 (0.02)	0.20 (0.02)	1.72 (0.02)	42.51 (2.38)	55.48 (2.38)	0.09 (0.02)	0.20 (0.02)	1.72 (0.02)	42.51 (2.38)	55.48 (2.38)
MCRM	1.09 (1.01)	2.26 (2.20)	26.69 (7.69)	65.68 (9.26)	4.28 (2.10)	2.09 (1.31)	2.26 (2.20)	26.50 (6.68)	64.88 (8.25)	4.28 (0.13)
Multi-obj	1.41 (1.13)	4.56 (3.97)	22.69 (8.44)	67.31 (9.93)	4.03 (3.29)	18.42 (1.31)	20.69 (2.40)	22.51 (6.67)	31.41 (8.25)	6.97 (1.23)

Table 7: Recommended & allocated percentages for Scenario 2 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.01	**0.05**	0.10	0.15	0.3	0.01	**0.05**	0.10	0.15	0.3
Efficacy probability	0.6	**0.8**	0.5	0.4	0.2	0.6	**0.8**	0.5	0.4	0.2
SEEDA	6.3 (0.90)	91.23 (3.18)	1.45 (1.02)	0.53 (0.34)	0.08 (0.08)	5.56 (3.11)	87.26 (3.94)	2.95 (2.09)	2.14 (1.43)	2.09 (0.63)
Indep TS	5.31 (4.95)	92.09 (1.32)	1.47 (1.08)	0.64 (0.56)	0.48 (0.16)	7.99 (2.55)	83.18 (5.34)	4.27 (4.34)	2.91 (2.30)	1.65 (1.05)
KL-UCB	9.68 (2.73)	87.66 (2.98)	1.91 (1.20)	0.66 (0.44)	0.09 (0.04)	7.01 (1.57)	81.93 (1.94)	3.03 (0.82)	2.31 (0.51)	5.72 (0.31)
UCB	8.58 (3.98)	89.80 (4.18)	1.26 (1.24)	0.34 (0.24)	0.03 (0.03)	21.06 (2.20)	46.31 (2.69)	15.07 (1.68)	11.16 (1.28)	6.40 (0.73)
3+3	0.20 (0)	1.80 (0.32)	5.40 (0.78)	13.80 (2.37)	78.80 (8.34)	16.71 (3.35)	18.81 (3.65)	19.40 (3.78)	19.88 (3.14)	19.75 (4.65)
CRM	0 (0)	0 (0)	0 (0)	9.98 (0.42)	90.02 (0.42)	0 (0)	0 (0)	0 (0)	9.97 (1.25)	90.03 (1.43)
MCRM	0.08 (0)	0.17 (0.02)	1.15 (1.00)	13.47 (0.44)	85.13 (0.04)	1.08 (0.27)	0.17 (0.09)	1.15 (0.61)	13.44 (5.76)	84.16 (6.73)
Multi-obj	6.07 (1.74)	90.85 (1.86)	1.93 (0.54)	0.92 (0.30)	0.22 (0.11)	34.88 (7.26)	51.74 (6.81)	7.11 (2.41)	4.34 (1.20)	1.93 (0.50)

Table 8: Recommended & allocated percentages for Scenario 3 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.06	0.08	0.14	**0.2**	0.3	0.06	0.08	0.14	**0.2**	0.3
Efficacy probability	0.2	0.4	0.6	**0.8**	0.55	0.2	0.4	0.6	**0.8**	0.55
SEEDA	1.84 (0.71)	1.97 (1.10)	6.15 (2.86)	88.12 (3.22)	1.58 (1.00)	2.27 (0.71)	2.54 (1.11)	6.27 (2.86)	85.46 (3.22)	3.46 (0.99)
Indep TS	0.76 (0.45)	1.55 (0.93)	5.49 (3.71)	89.85 (5.09)	2.35 (1.73)	1.67 (0.48)	2.98 (1.33)	8.17 (4.48)	81.28 (4.96)	5.89 (1.79)
KL-UCB	2.64 (0.54)	7.29 (1.15)	28.47 (3.22)	57.18 (3.58)	4.43 (1.41)	2.62 (0.54)	6.58 (1.15)	26.85 (3.22)	55.07 (3.58)	8.87 (1.41)
UCB	1.71 (0.48)	3.57 (1.33)	19.04 (4.48)	72.89 (4.96)	2.79 (1.79)	8.33 (0.48)	13.17 (1.33)	22.75 (4.48)	44.71 (4.96)	11.04 (1.79)
3+3	2.20 (1.93)	4.80 (2.10)	10.60 (3.22)	18.80 (3.92)	63.60 (9.33)	19.77 (3.54)	20.08 (5.93)	20.43 (5.12)	18.67 (3.95)	15.29 (3.45)
CRM	0 (0)	0 (0)	0 (0)	4.37 (0.69)	95.63 (0.69)	0 (0)	0 (0)	0 (0)	4.37 (0.66)	95.63 (0.66)
MCRM	0.60 (0.54)	0.87 (0.15)	3.57 (3.22)	31.89 (3.58)	63.07 (1.41)	1.60 (0.98)	0.87 (1.26)	3.57 (2.97)	31.68 (8.86)	62.28 (10.58)
Multi-obj	0.78 (0.20)	2.07 (0.45)	8.67 (1.97)	84.99 (2.40)	3.49 (1.02)	16.43 (0.20)	20.56 (0.45)	21.56 (1.97)	34.45 (2.41)	7.00 (1.02)

Table 9: Recommended & allocated percentages for Scenario 4 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.05	0.1	**0.25**	0.5	0.6	0.05	0.1	**0.25**	0.5	0.6
Efficacy probability	0.2	0.4	**0.6**	0.8	0.55	0.2	0.4	**0.6**	0.8	0.55
SEEDA	3.43 (1.26)	12.15 (3.69)	79.72 (4.25)	4.37 (1.90)	0 (0)	3.40 (1.24)	11.05 (3.48)	79.44 (4.28)	5.00 (1.75)	1.12 (0.45)
Indep TS	11.53 (9.17)	24.58 (10.80)	58.58 (12.42)	2.66 (1.53)	2.65 (3.42)	1.68 (0.99)	3.02 (2.39)	8.50 (5.40)	81.01 (16.00)	5.79 (6.50)
KL-UCB	24.60 (6.65)	37.78 (14.78)	28.34 (14.62)	6.91 (2.00)	2.37 (2.78)	1.91 (0.32)	2.43 (0.52)	3.41 (1.41)	51.61 (1.89)	40.64 (1.06)
UCB	4.87 (5.17)	32.53 (10.80)	60.34 (14.42)	1.84 (1.52)	0.42 (0.42)	14.29 (0.72)	26.93 (1.31)	40.69 (2.11)	9.15 (2.63)	8.94 (1.85)
3+3	3 (1.46)	6.20 (4.64)	34.20 (6.85)	40.40 (7.10)	16.20 (4.16)	22.57 (7.69)	22.82 (6.98)	26.70 (7.89)	17.10 (6.79)	4.29 (0.68)
CRM	0 (0)	0 (0)	0 (0)	95.56 (0.14)	4.44 (0.14)	0 (0)	0 (0)	0 (0)	95.23 (2.12)	4.77 (2.12)
MCRM	0.84 (0.83)	3.77 (1.73)	88.03 (3.92)	7.17 (3.58)	0.19 (0.01)	1.84 (0.83)	3.77 (1.73)	87.04 (3.91)	7.16 (3.57)	0.19 (0.01)
Multi-obj	3.64 (0.66)	19.66 (4.87)	70.79 (5.13)	4.80 (1.18)	1.11 (0.41)	19.93 (6.11)	23.96 (4.93)	23.97 (4.69)	26.16 (4.17)	5.98 (2.25)

Table 10: Recommended & allocated percentages for Scenario 5 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.1	**0.2**	0.4	0.5	0.6	0.1	**0.2**	0.4	0.5	0.6
Efficacy probability	0.1	**0.3**	0.5	0.5	0.5	0.1	**0.3**	0.5	0.5	0.5
SEEDA	7.20 (1.10)	74.95 (4.42)	15.01 (4.84)	2.50 (1.46)	0 (0)	6.86 (0.96)	67.46 (3.49)	21.21 (4.11)	3.22 (1.58)	1.26 (0.61)
SEEDA-Plateau	12.60 (2.12)	82.20 (5.45)	4.60 (2.12)	0.60 (0.40)	0 (0)	19.50 (5.12)	56.46 (9.23)	15.49 (4.56)	7.56 (1.23)	1.00 (0.54)
Indep TS	21.59 (7.05)	50.75 (10.00)	21.15 (11.41)	4.73 (1.91)	1.78 (1.42)	2.67 (1.52)	6.34 (6.28)	29.19 (6.32)	30.59 (6.41)	31.22 (6.14)
KL-UCB	23.64 (4.52)	40.01 (10.18)	21.58 (10.82)	10.92 (2.16)	3.85 (0.81)	3.80 (0.75)	2.24 (1.38)	23.69 (10.60)	40.19 (9.94)	30.10 (10.93)
UCB	13.71 (1.63)	75.24 (9.14)	8.66 (5.79)	1.85 (0.79)	0.54 (0.96)	18.75 (0.64)	36.38 (4.19)	16.49 (2.57)	14.23 (2.63)	14.14 (2.55)
3+3	7.40 (1.42)	21.20 (12.30)	42.60 (6.42)	21.80 (3.06)	7.00 (4.12)	29.03 (0.79)	29.38 (3.32)	23.97 (2.14)	8.35 (1.15)	1.76 (0.42)
CRM	0 (0)	0 (0)	0 (0)	94.72 (0.04)	5.28 (0.05)	0 (0)	0 (0)	0 (0)	94.39 (0.02)	5.61 (0.02)
MCRM	2.86 (0.80)	62.72 (1.66)	33.03 (4.13)	1.25 (4.05)	0.14 (0)	3.86 (0.80)	62.02 (1.66)	32.73 (4.11)	1.25 (0.42)	0.14 (0.02)
Multi-obj	9.56 (0.58)	60.18 (3.92)	23.51 (4.17)	5.38 (1.00)	1.38 (0.39)	23.42 (6.89)	25.22 (5.30)	22.55 (5.28)	16.27 (6.61)	12.54 (5.79)

Table 11: Recommended & allocated percentages for Scenario 6 of (Zang et al., 2014).

	Recommended					Allocated				
Toxicity probability	0.01	0.03	0.05	**0.1**	0.2	0.01	0.03	0.05	**0.1**	0.2
Efficacy probability	0.1	0.3	0.45	**0.6**	0.6	0.1	0.3	0.45	**0.6**	0.6
SEEDA	1.47 (0.45)	1.79 (1.16)	5.12 (3.94)	48.97 (10.31)	42.32 (12.35)	3.59 (0.56)	2.93 (1.55)	5.89 (3.19)	45.65 (6.51)	41.94 (6.62)
SEEDA-Plateau	0 (0)	0.20 (0.05)	3 (1.38)	96 (5.72)	0.80 (0.56)	4.20 (3.75)	5.64 (2.45)	13.73 (5.42)	40.22 (9.85)	36.18 (4.75)
Indep TS	0.42 (0.31)	1.24 (0.86)	5.20 (3.13)	47.46 (12.35)	45.67 (12.22)	13.71 (1.06)	18.37 (3.55)	22.33 (5.87)	28.10 (8.80)	17.48 (8.57)
KL-UCB	1.96 (0.50)	2.55 (1.46)	9.57 (3.46)	54.30 (10.30)	31.62 (10.06)	3.78 (0.77)	3.32 (0.76)	9.42 (2.14)	52.03 (10.56)	31.45 (10.53)
UCB	1.31 (0.37)	2.06 (1.22)	9.47 (4.06)	56.47 (10.82)	30.69 (10.74)	8.18 (0.58)	12.58 (1.30)	19.54 (2.02)	32.85 (2.83)	26.84 (2.93)
3+3	0 (0)	1.40 (0.23)	2.20 (1.23)	8.20 (1.27)	88.20 (7.21)	17.14 (6.79)	18.15 (7.90)	18.32 (7.45)	20.07 (6.52)	20.74 (6.48)
CRM	0 (0)	0 (0)	0 (0)	65.39 (2.29)	34.61 (2.29)	0 (0)	0 (0)	0 (0)	65.15 (6.79)	34.85 (6.41)
MCRM	0.06 (0.02)	0.08 (0.04)	0.49 (0.50)	2.92 (1.21)	96.45 (3.00)	1.06 (0.25)	0.08 (0.04)	0.48 (0.29)	2.92 (2.01)	95.45 (3.00)
Multi-obj	0.63 (0.17)	1.60 (0.36)	6.78 (1.52)	49.01 (10.01)	41.98 (10.03)	13.71 (7.55)	18.37 (8.18)	22.33 (8.00)	28.10 (7.23)	17.48 (7.41)