

# A Curve Free Method based on Bayesian Optimization for Oncology Phase I Clinical Trials

Ami Takahashi<sup>1,2</sup>, Taiji Suzuki<sup>3,4</sup>

<sup>1</sup>Tokyo Institute of Technology / <sup>2</sup>Pfizer R&D Japan  
2-12-1 Ookayama, Meguro City, Tokyo, Japan / 3-22-7 Yoyogi, Shibuya City, Tokyo, Japan  
ami.takata@pfizer.com; taiji@mist.i.u-tokyo.ac.jp

<sup>3</sup>The University of Tokyo / <sup>4</sup>RIKEN  
7-3-1 Hongo, Bunkyo City, Tokyo, Japan / 1-4-1 Nihonbashi, Chuo City, Tokyo, Japan

**Abstract** - The main goal of oncology phase I clinical trials is to identify a maximum tolerated dose. Although many authors have offered a variety of methods for the maximum tolerated dose, conventional methods have some potential shortcomings. Parametric model-based methods may provide incorrect results due to model misspecification because not only little is known about dose--toxicity relationship but also the sample size is small. In addition, most methods ignore uncertainty of estimated dose--toxicity relationship in their dose selection process. To address these concerns, we propose a Bayesian optimization design for estimating the maximum tolerated dose. Our proposed method utilizes a Bayesian optimization framework, so that a distribution of the dose--toxicity relationship is nonparametrically estimated and the dose is selected by balanced information between exploration and exploitation to avoid concentrating on a local optimum during a trial. Therefore, our proposed method corresponds to the model specification issue through utilizing a nonparametric model and enables us to account for uncertainty during dose selections through treating a distribution of an estimated curve. In this paper, we introduce a Bayesian optimization design under single-agent trials. We compare it with continual reassessment method, modified toxicity probability interval, and a curve-free method based on a product-of-beta prior. Simulation results suggest that our proposed method works successfully in the maximum tolerated dose estimation problem in terms of correct recommendation of a true maximum tolerated dose and it has a potential as being an alternative to existing methods.

**Keywords:** Bayesian optimization; Curve-free method; Dose-finding; Maximum tolerated dose; Nonparametric Model.

## 1. Introduction

In oncology phase I clinical trials, the primary objective is to identify a maximum tolerated dose (MTD). MTD is defined as the highest dose that does not cause an unacceptable level of dose limiting toxicity (DLT). The continual reassessment method (CRM) [1] is one of the most popular methods for estimating an MTD as a parametric model-based method. On the other hand, the CRM has a potential risk of model misspecification. The CRM identifies an MTD correctly regardless of model misspecification, if sample size is not small. However, dose-finding studies are conducted with a limited sample size in practice. Furthermore, little is known about the dose--toxicity relationship at early phase trials. Thus, an appropriate model should be carefully selected since the model generally affects its operating characteristics in dose-finding studies [2]. In addition, CRM utilizes only average information but ignores uncertainty of the estimated dose--toxicity curve in its dose selection. Thus, non-negligible variability is not reflected in the dose selections especially at the start of the trial [3].

Corresponding to the model specification issue, several nonparametric methods have been introduced to relax such strong parametric assumptions. A nonparametric Bayesian method based on a product-of-beta-prior (PBP) [4] can be often seen in many articles as a curve-free method. In the PBP, a toxicity probability at each dose level is reparametrized by another parameter with a beta prior distribution, so that a prior distribution of a toxicity probability becomes a product-of-beta-prior. The modified toxicity probability interval (mTPI) [5] is one of the most popular toxicity interval methods based on nonparametric approach. The mTPI can be implemented as simple as the 3+3 [6] without any logistic burden, while it is assisted by a beta-binomial model. Because of its simplicity and much better performance than the 3+3, the popularity of mTPI has been growing in both research and industry entities. The PBP and the mTPI estimate an MTD without strong assumptions, however, there are still some concerns. Like the CRM, the PBP does not reflect non-negligible

variability during its dose selection process. The mTPI utilizes a toxicity distribution during its dose selection but focuses on only the current test dose, therefore, dose--toxicity relationship connecting all doses is not estimated until the end of the trial.

In order to address these concerns, we propose a Bayesian optimization design for MTD estimation, which applies a Bayesian optimization framework [7]. Hereafter, the proposed method is expressed as ‘‘BO’’. BO is based on a different concept from existing nonparametric methods and realizes sophisticated dose selection procedures. BO can address the model specification issue by utilizing a nonparametric model for dose--toxicity relationship. It allows a much more flexible estimation than existing parametric model-based methods. In addition, BO updates dose--toxicity distribution for all doses once patient outcomes are obtained. Based on the updated distribution, the next dose is determined by accounting uncertainty of estimated dose-toxicity relationship.

We organize this paper as follows: We describe our proposed method in Section 2. We explain our simulation studies and show results in Section 3. Discussion and Conclusion are shown in Sections 4 and 5, respectively.

## 2. Bayesian Optimization for MTD Estimation

We establish an unknown objective function that derives from dose--toxicity relationship. An MTD candidate is estimated as the minimizer of the objective function in the explored dose range. As we cannot describe the objective function owing to lack of information about the dose--toxicity relationship, we estimate it through observation data.

### 2.1. Statistical Model for Dose--Toxicity Relationship

The dose--toxicity relationship is defined by a nonparametric model as follows:

$$f(x) = \text{logit}\{\pi(x)\} = \log[\pi(x)/\{1 - \pi(x)\}], \quad (1)$$

where  $\pi(x) = [1 + \exp\{-f(x)\}]^{-1}$  is a probability of DLT at a conceptual dose  $x$  ( $x \in \{x_1, \dots, x_J\}$ ). The logit transformation of  $\pi(x)$  guarantees that  $\pi(x)$  bounds within the range from 0 to 1 on a finite dose range.

The number of patients experienced DLT ( $Y$ ) in  $n$  patients treated at a dose  $x$  follows a binomial distribution  $Y \sim \text{Bin}(n, \pi(x))$ . Thus, the likelihood function at the  $t$ -th cohort is given by

$$L(D_{1:t}|f) = \prod_{j=1}^J \pi(x_j)^{y_j} \{1 - \pi(x_j)\}^{n_j - y_j}, \quad (2)$$

where  $D_{1:t} = \{(x_{(1)}, y_{(1)}), \dots, (x_{(t)}, y_{(t)})\}$ ;  $x_{(t)}$  and  $y_{(t)}$  denote the test dose and number of patients with DLT at the  $t$ -th test, respectively;  $n_j$  is the total number of patients treated with  $x_j$  ( $j \in \{1, \dots, J\}$ ); and  $y_j = \sum_{k=1}^t y_{(k)} 1[x_{(k)} = x_j]$ . The indicator function  $1[x_{(k)} = x_j]$  returns 1 when  $x_{(k)} = x_j$ . Otherwise, it returns 0.

### 2.2. Estimation of Dose--Toxicity Distribution and Next Dose Selection

The Bayesian optimization typically puts a Gaussian process for a prior belief of an unknown function. Therefore, BO also places a Gaussian process prior over the unknown function  $f$ , that is,  $f \sim \text{GP}(m, k)$ . The Gaussian process prior is specified by a mean function  $m(x)$  and a covariance function  $k(x, x')$ , where  $x \in \{x_1, \dots, x_J\}$ . The prior function for  $m(x)$  derives from pre-specified initial guess toxicities. In practice, little is known about the dose--toxicity relationship, therefore,  $m(x)$  can be generated by using the `getprior` function, which is in `dfcrm` package in R, as CRM does. For  $k(x, x')$ , we apply a squared exponential kernel described as follows:

$$k(x, x') = \sigma_f^2 \exp\{-|x - x'|^2 / (2\rho^2)\}, \quad (3)$$

where  $\sigma_f$  and  $\rho$  are a signal variance and a scale parameter, respectively. For calibration of the kernel parameters,  $\sigma_f$  is usually fixed as 1 and  $\rho$  is determined to fit the model. The value of  $\rho$  indicates the typical distance between turning points. Because it is unlikely that there are many turning points in dose--toxicity function, an appropriate search range would be from a half to the full length of conceptual doses. Conceptual doses can have any arbitrary values, however, the distance between doses should be an equal interval to avoid biased dose selections. In addition, a small value of  $\tau$  is added

on the diagonal elements of the covariance matrix for computational stability [8]. With this modification, each element is expressed as  $k(x, x') + \tau 1[x = x']$ . Once patient outcomes are obtained, a posterior distribution for  $f$  is estimated in Bayesian manner.

Our goal is to identify which dose level provides the closest toxicity to the target. For reaching the goal, the objective function  $g(x)$  is given by

$$g(x) = |\pi(x) - \theta|, \quad (4)$$

where  $\theta$  is a target toxicity rate. Because the exact form of  $g(x)$  is unavailable, we design an acquisition function  $\hat{g}(x)$ , which is an alternative of the true objective function of  $g(x)$ . We utilize one of the most popular acquisition functions called lower confidence bound [9]. It provides balanced information between exploitation and exploration on the posterior distribution of  $g(x)$ . In this paper, the lower confidence bound is defined as the 10-th percentile of the posterior distribution of  $g(x)$  as  $\text{LCB}(x)$ . The next dose is selected as follows:

$$x_{(t+1)} = \operatorname{argmin}_{x \in A_t} \{\text{LCB}(x)\}. \quad (5)$$

The admissible dose set  $A_t$  is determined by pre-specified rules and refreshed at each test. In this paper,  $A_t$  includes one dose level higher than the current tested dose as a maximum dose for the next candidate, because we assume that no dose skipping is allowed in the dose selection to ensure patient's safety.

### 2.3. Stopping rules of a Trial and the Final MTD Estimation

A trial is terminated when it meets either (1) - (3) conditions: (1) all planned patients are treated; (2)  $\text{P}\{\pi(x_1) > \theta | D_{1:t}\} > p_1$  when more than a total of  $n_1$  patients are treated; (3)  $\text{P}\{\pi(x_j) < \theta | D_{1:t}\} > p_2$  when more than  $n_2$  patients are treated at the highest dose. The values for  $n_1$ ,  $n_2$ ,  $p_1$  and  $p_2$  are specified before the trial.

Once the trial is completed or terminated, the MTD is determined by the final posterior distribution of  $g$  as follows:

$$\text{MTD} = \min_x [E_g\{g(x) | D_{1:t}\}]. \quad (6)$$

## 3. Evaluation of Operating Characteristics

We present the simulation frameworks and results for two different trials assuming single-agent trials in this section.

### 3.1. Simulation settings for each method

Each simulation was computed 1,000 trials. Simulation (a) focused on operating characteristics under lower target toxicity rate according to rigidity issues under a specific condition that has been pointed on PBP [10]. In simulation (b), MTD with toxicity of 0.3 was explored within 10 dose levels. We compared operating characteristics of BO with those of CRM, mTPI and PBP. Table 1 shows toxicity scenarios for simulation (a) and (b). The true MTD under each scenario is displayed as bold style. The trial started from the lowest dose. The trial was completed when the total number of patients reached the maximum sample size.

For BO, the initial guesses of the dose—toxicity relationship were generated by the `getprior` function with 4 parameters of  $(\delta = 0.03, \theta, v = J/2, J)$  by assuming a power model, where  $\delta$  is an indifference half width and  $v$  is the prior guess of MTD. The prior mean function in a Gaussian process was the logit transformed values of the generated initial guesses. The kernel parameters for a covariance function were  $\sigma_f = 1$  and  $\rho = 0.7$ . The value of  $\tau$  was 0.05 for simulation (a) and 0.07 for simulation (b) based on the computational speed. The conceptual doses were the values from 0 to 1 divided into  $J$  evenly. In BO, the trial was terminated early when either stopping rule of (2) or (3) was shown. We used  $p_1 = p_2 = 0.3$ ,  $n_1 = 9$  and  $n_2 = 3$ . We implemented BO by using R and the `rstan` package was used for MCMC.

A power model (CRM-p) and a logistic model with a fixed intercept of 3 (CRM-l) were evaluated as CRM. We set initial guesses for each model based on the `getprior` function. CRM-p had  $\delta = 0.05$  in both simulations (a) and (b). CRM-l had  $\delta = 0.06$  and 0.05 in simulation (a) and (b), respectively. Parameters in the `getprior` except for  $\delta$  were

the same as those in BO. The values for  $\delta$  were determined by [11] through 2,000 simulations. Also, CRM was implemented by `dfcrm` package in R. The MTD was determined based on the final posterior mean that had the closest toxicity to  $\theta$ .

For mTPI, a prior toxicity distribution at each dose followed a beta distribution  $\text{Beta}(1,1)$ . In the dose selection, unit probability mass was calculated for under, proper, and overdosing intervals by assigning 0.05 as a pre-specified value that determined the width of toxicity intervals. If  $P\{\pi_j > \theta | D_{1:t}\} > 0.95$  on a dose level  $j$  at which at least 6 patients treated, the dose level  $j$  and all higher doses were excluded in subsequent cohorts. If the excluded dose was the lowest dose, the MTD was not determined for the trial. At the end of a trial, the MTD was determined based on the isotonicity transformed posterior mean. The mTPI was conducted via `East`<sup>®</sup> 6.5.

For PBP, two prior inputs were utilized based on `getprior` function with the same parameters as those in BO except for  $\delta$ , while there is no direction if prior inputs from clinicians are unavailable in the original article. Hereafter, they were displayed as  $\text{PBP}(\delta)$ . The posterior expectations for each dose level were estimated by MCMC through the `rstan` package in R. The MTD was determined as a dose level with the closest toxicity level to the target toxicity regarding the posterior expectations at the end of a trial.

Table 1: Toxicity at Each Scenario.

Simulation	Scenario	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
(a) $\theta = 0.1$ , Maximum number of patients=30.	A1	0.05	<b>0.10</b>	0.30	0.70	0.80	0.85				
	A2	<b>0.10</b>	0.15	0.23	0.33	0.45	0.50				
	A3	0.02	<b>0.10</b>	0.15	0.20	0.24	0.30				
	A4	0.01	0.05	<b>0.10</b>	0.20	0.33	0.70				
	A5	0.01	0.02	0.05	<b>0.08</b>	0.13	0.23				
(b) $\theta = 0.3$ , Maximum number of patients=45.	B1	0.03	0.05	0.08	0.12	0.16	0.22	<b>0.33</b>	0.47	0.60	0.75
	B2	0.10	<b>0.33</b>	0.45	0.55	0.60	0.65	0.68	0.72	0.75	0.80
	B3	0.02	0.03	0.04	0.05	0.07	0.10	0.12	0.15	0.18	<b>0.26</b>
	B4	0.02	0.10	<b>0.30</b>	0.60	0.80	0.85	0.90	0.95	1.00	1.00
	B5	0.01	0.01	0.01	0.02	0.03	0.04	0.05	0.10	<b>0.30</b>	0.80
	B6	0.05	0.10	0.15	<b>0.30</b>	0.42	0.55	0.65	0.75	0.85	1.00

### 3.2. Simulation results

Table 2 shows the results in Simulation (a). As shown in Correct selection (%) in Table 2, BO selected the true MTD with higher percentages than other methods in most scenarios. Although PBP (0.08) at scenario A5 selected only subtherapeutic doses, BO did not show such rigidity. In terms of patient allocation to MTD (%), BO allocated more patients to the true MTD than other methods in all scenarios except for scenario A2. Scenario A2 places the MTD on the lowest dose. In this scenario, BO allocated patients to higher doses than  $x_1$ , while the average observed toxicity percentages (Observed toxicity (%)) was similar to that of mTPI.

Table 2: Simulation (a): Correctly selected percentages as an MTD by Scenario.

Scenario		BO	CRM-p	CRM-l	mTPI	PBP(0.05)	PBP(0.08)
A1	Total number of patients	28.8	30.0	30.0	29.2	30.0	30.0
	Observed toxicity (%)	15.7	12.4	11.8	14.5	10.4	10.0
	Patient allocation to MTD (%)	60.8	45.5	46.1	45.8	38.7	41.6
	Correct selection (%)	75.7	64.6	60.7	62.5	43.4	50.0
A2	Total number of patients	26.9	30.0	30.0	27.1	30.0	30.0
	Observed toxicity (%)	17.8	14.1	13.9	17.0	13.1	12.8
	Patient allocation to MTD (%)	22.4	54.5	55.8	51.1	67.2	61.6
	Correct selection (%)	40.8	57.1	57.8	39.6	75.9	69.6
A3	Total number of patients	29.7	30.0	30.0	29.9	30.0	30.0
	Observed toxicity (%)	12.4	10.9	10.8	11.7	8.4	8.5
	Patient allocation to MTD (%)	38.8	34.8	36.2	33.4	30.9	34.6
	Correct selection (%)	49.7	44.3	42.2	36.0	34.8	40.8
A4	Total number of patients	30.0	30.0	30.0	30.0	30.0	30.0
	Observed toxicity (%)	10.0	10.3	10.2	11.5	7.6	6.0
	Patient allocation to MTD (%)	45.6	34.4	30.3	31.1	32.8	44.4
	Correct selection (%)	57.1	47.6	49.3	49.8	40.0	48.5
A5	Total number of patients	30.0	30.0	30.0	30.0	30.0	30.0
	Observed toxicity (%)	6.0	7.6	7.7	8.3	4.7	3.6
	Patient allocation to MTD (%)	31.3	21.8	20.3	20.4	31.0	0.0
	Correct selection (%)	44.1	36.4	34.6	29.3	35.0	0.0

Table 3 shows the results in Simulation (b). BO provided good performance compared with other methods comprehensively in all scenarios in terms of correct selection of the MTD. The results of patient allocation to the MTD supported these results. On the other hand, the observed toxicity (%) in BO tended to be slightly higher than other methods depending on scenarios. Although the differences were minimal and most observed values were close to the target toxicity rate, it is expected to address overdose allocations in future work.

Table 3: Simulation (b): Correctly selected percentages as an MTD by Scenario.

Scenario		BO	CRM-p	CRM-l	mTPI	PBP(0.05)	PBP(0.08)
B1	Total number of patients	45.0	45.0	45.0	45.0	45.0	45.0
	Observed toxicity (%)	25.2	18.7	19.5	26.8	24.8	22.0
	Patient allocation to MTD (%)	30.9	17.5	18.8	28.1	33.3	21.5
	Correct selection (%)	48.1	38.1	40.2	41.5	51.1	34.5
B2	Total number of patients	45.0	45.0	45.0	45.0	45.0	45.0
	Observed toxicity (%)	36.2	31.6	31.0	31.1	33.2	33.6
	Patient allocation to MTD (%)	39.2	53.1	54.1	43.2	40.3	43.0
	Correct selection (%)	66.7	69.5	74.8	51.4	61.8	61.7
B3	Total number of patients	45.0	45.0	45.0	38.5	45.0	45.0
	Observed toxicity (%)	16.6	11.2	11.6	15.6	16.1	14.3
	Patient allocation to MTD (%)	34.7	9.2	10.2	41.9	34.6	23.4
	Correct selection (%)	67.6	33.3	37.3	36.0	74.7	64.7
B4	Total number of patients	45.0	45.0	45.0	45.0	45.0	45.0
	Observed toxicity (%)	32.6	30.4	29.5	28.6	30.7	31.0
	Patient allocation to MTD (%)	64.6	57.7	59.0	61.3	61.5	64.3
	Correct selection (%)	87.4	84.5	84.2	70.5	84.6	85.2
B5	Total number of patients	45.0	45.0	45.0	45.0	45.0	45.0
	Observed toxicity (%)	21.7	15.4	16.7	23.1	20.6	18.8
	Patient allocation to MTD (%)	42.3	26.9	29.2	55.6	39.0	36.1
	Correct selection (%)	77.0	70.5	81.4	71.2	70.0	75.8
B6	Total number of patients	45.0	45.0	45.0	45.0	45.0	45.0
	Observed toxicity (%)	31.1	26.6	26.1	29.9	29.8	29.1
	Patient allocation to MTD (%)	37.0	38.2	38.3	37.1	36.8	39.6
	Correct selection (%)	61.6	62.7	61.7	50.9	54.5	57.2

#### 4. Discussion

BO provided consistently higher correct selection percentages than other methods, however, additional overdose control might be needed depending on situations. For example, an admissible dose set based on a toxicity probability distribution can be determined by more conservative settings. We can also rearrange the final candidates for the MTD selection by using the updated toxicity probability distribution, while the current proposed method selects the MTD from all doses. In this paper, the simulation scenarios were relatively simple settings, therefore, it might be difficult to show dramatically differences among methods. If the dose--toxicity relationship is more complex, it becomes more difficult to apply strong assumptions before the trial. Therefore, as our next evaluation, we are addressing applications of a Bayesian optimization framework to identify a single MTD or MTD contour assuming trials for combination therapies.

#### 5. Conclusion

We introduced basic concept of applying a Bayesian optimization framework to MTD estimation under single-agent trials. BO is categorized to curve-free methods but a novel approach for MTD estimation. In this paper, we presented the results based on two different simulation studies to evaluate BO by comparing it with CRM, mTPI and PBP. In the simulation results, BO provided consistently better or comparable performance among almost all scenarios, while other methods yielded good results in one scenario but very poor results in another. Further numerical investigations would reveal more detailed properties of BO and it remains a challenge for future work to address overdose control and establish a calibration approach of design parameters, however, the simulation results lead to the conclusion that BO has a potential to be an alternative to the existing methods.

## Acknowledgements

TS was partially supported by MEXT Kakenhi (18K19793, 18H03201 and 20H00576), Japan Digital Design, and JST-CREST.

## References

- [1] J. O'Quigley, M. Pepe and L. Fisher, "Continual reassessment method: a practical design for phase I clinical trials in cancer," *Biometrics.*, vol. 46, no. 1, pp. 33-48, 1990.
- [2] X. Paoletti and A. Kramar, "A comparison of model choices for the Continual Reassessment Method in phase I cancer trials," *Stat Med.*, vol. 28, no. 24, pp. 3012-28, 2009.
- [3] B. Neuenschwander, M. Branson and T. Gsponer, "Critical aspects of the Bayesian approach to phase I cancer trials," *Stat Med.*, vol. 27, no. 13, pp. 2420-39, 2008.
- [4] M. Gasparini and J. Eisele, "A curve-free method for phase I clinical trials," *Biometrics.*, vol. 56, no. 2, pp. 609-15, 2000.
- [5] Y. Ji, P. Liu, Y. Li and B. N. Bekele, "A modified toxicity probability interval method for dose-finding trials," *Clin trials.*, vol. 7, no. 6, pp. 653-63, 2010.
- [6] B. E. Storer, "Design and analysis of phase I clinical trials," *Biometrics.*, vol. 45, no. 3, pp. 925-37, 1989.
- [7] J. Mockus, "On the Bayes Methods for Seeking the Extremal Point," in *Proceedings of the 6th IFAC World Congress*, Boston/Cambridge, MA, 1975, vol. 8, pp. 428-31.
- [8] R. M. Neal, "Monte Carlo implementation of Gaussian process models for Bayesian regression and classification," Dept. Statistics, Toronto Univ., Technical Report No. 9702, 1997.
- [9] J. Mockus, "Gaussian process optimization in the bandit setting: no regret and experimental design," in *Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel, 2010, pp. 1015–22.
- [10] Y. K. Cheung, "On the use of nonparametric curves in phase I trials with low toxicity tolerance," *Biometrics.*, vol. 58, no. 1, pp. 237-40, 2002.
- [11] S. M. Lee and Y. K. Cheung, "Model calibration in the continual reassessment method," *Clin Trials.*, vol. 6, no. 3, pp. 227-38, 2009.