Received: 18 October 2021 | Revised: 28 May 2022 | Accepted: 31 May 2022

DOI: 10.1002/pst.2247

MAIN PAPER

WILEY

CUSUMIN: A cumulative sum interval design for cancer phase I dose finding studies

Tomoyoshi Hatayama[1] | **Seiichi Yasui**[2]

[1]Statistics and Decision Sciences Japan, Janssen Pharmaceutical K.K., Tokyo, Japan

[2]Department of Industrial Administration, Tokyo University of Science, Tokyo, Japan

Correspondence
Tomoyoshi Hatayama, Statistics and Decision Sciences Japan, Janssen Pharmaceutical K.K., Tokyo, Japan.
Email: thatayam@its.jnj.com

Abstract

Recently, model-assisted designs, including the Bayesian optimal interval (BOIN) design with optimal thresholds to determine the dose for the next cohort, have been proposed for cancer phase I studies. Model-assisted designs are useful because of their good performance as model-based designs in addition to their algorithm-based simplicity. In BOIN, escalation and de-escalation based on boundaries can be understood as a type of change point detection based on a sequential test procedure. Notably, the sequential test procedure is used in a wide range of fields and is known for its application to control charts, statistical monitoring methods used for detecting abnormalities in manufacturing processes. In control charts, abnormalities are detected if the control chart statistics are observed to be outside of the optimal boundaries. The cumulative sum (CUSUM) statistic, which is developed for control chart applications, derives higher power under the same erroneous judgment rate. Hence, it is expected that a more efficient model-assisted design can be achieved by the application of CUSUM statistics. In this study, a model-assisted design based on the CUSUM statistic is proposed. In the proposed design, the dose for the next cohort is decided by CUSUM statistics calculated from the counts of the dose-limiting toxicity and pre-defined boundaries, based on the CUSUM control chart scheme. Intensive simulation shows that our proposed method performs better than BOIN, and other representative model-assisted designs, including modified toxicity probability interval (mTPI) and Keyboard, in terms of controlling over-dosing rates while maintaining similar performance in the determination of maximum tolerated dose.

KEYWORDS

BOIN, CUSUM, dose finding study, model-assisted design, phase I study

1 | INTRODUCTION

A phase I clinical trial design aims to identify the maximum tolerated dose (MTD) of a new drug, which is defined as the dose with a dose-limiting toxicity (DLT) probability closest to the target probability. Traditionally,

phase I dose-finding designs can be classified as algorithm-based and model-based methods.[1,2] Algorithm-based designs, including 3 + 3 and biased-coin designs[3] and their extensions,[4,5] are based on simple, easy to understand, and pre-specified rules to govern dose escalation and de-escalation. However, its poor operational characteristics, including a lower probability of MTD selection, are well known.[4–6] To overcome the poor operational characteristics of most algorithm-based designs, model-based designs have been proposed.[4–7] These methods are based on statistical models, such as logistic regression models, in which the dose is escalated and de-escalated based on DLT probabilities estimated using a statistical model. The most famous model-based method is the continual reassessment method (CRM).[7] The escalation with overdose control design[8] has also been implemented in practice. These model-based methods have been actively studied and have many extensions.[9–17] However, they are conceptually and computationally complicated. Model-based designs require computation for model fitting and MTD estimation as soon as new observations become available. This characteristic of model-based methods renders them a hurdle for implementation by investigators.[18–22]

Recently, as a new category of dose-finding design, model-assisted designs have been developed to combine the simplicity of algorithm-based designs with the superior performance of model-based designs. As reported in an earlier seminal study, Ji et al.[18] proposed modified toxicity probability interval (mTPI) method. However, mTPI has drawbacks that limit distinct interpretation and lead to a high risk of overdosing patients.[19] Keyboard design was also proposed, to overcome mTPI design drawbacks.[19] As another variation of model-assisted design, the Bayesian optimal interval (BOIN) design was proposed.[20] Compared with the mTPI and Keyboard designs, the BOIN design is more straightforward and transparent. The dose escalation and de-escalation in the BOIN design are determined by comparing the observed DLT rate at the current dose with a pair of fixed dose escalation and de-escalation boundaries. The BOIN design has been expanded to the design for dose finding in combination drug settings,[22] considers both efficacy and toxicity,[23,24] and expands to consider delayed toxicity outcomes.[25] Recently, Mu et al.[26] proposed a generalized BOIN design that accommodates various existing toxicity grade scoring systems under a unified framework. As mentioned above, several model-assisted designs have been studied. Notably, model-assisted designs have been implemented in actual trials, and their implementations are increasing because of their simplicity and transparency.[27] For model-assisted designs, Zhou et al.[1] summarized the features of model-assisted methods, including mTPI, Keyboard design, and BOIN, and compared the operation characteristics with CRM via extensive simulations.

In the BOIN design, dose escalation and de-escalation are achieved solely based on the pre-determined boundary and observed DLT rate at the current dose. Dose escalation or de-escalation occurs when the point estimate of the DLT is observed outside the boundary. The boundary for dose escalation and de-escalation in BOIN is determined by minimizing incorrect decisions regarding dose assignment. The incorrect decisions in the method are dose escalation or de-escalation, when the current dose is the true MTD, or staying at the current dose, when the current dose is not the true MTD. Because the dose assignment decision is made based on the point estimate of the DLT rate and the pre-specified boundary, the probability of making an incorrect decision can be defined as a function of the threshold values of the boundary. The boundary thresholds in the other BOIN extensions are determined using the same type of derivation.[20,22–26] These boundary-based escalation and de-escalation rules used in BOIN and its variants can be considered as a type of change point detection based on a sequential test procedure, in which the likelihood ratio is a test statistic.

The most famous sequential hypothesis test is the sequential probability ratio test (SPRT) that was proposed by Wald[28] for statistical quality control problems; that is, sampling inspection. The SPRT has a larger power than ordinal statistical testing with a fixed type I error.[28–30] This class of tests has been developed and used in a wide range of research areas. In quality control research, the cumulative sum (CUSUM) statistic was developed for statistical quality control charts to detect process abnormalities in the manufacturing process.[30–33] The CUSUM statistic has several desirable characteristics compared to the SPRT that lead to higher power, with the same probability of erroneous detection.

In this study, we propose a CUSUM interval (CUSUMIN) design for cancer phase I dose-finding studies as a new model-assisted design. In this study, we propose a method for cancer phase I dose finding studies of a single agent and discuss our proposed method. The advantages of the CUSUM statistic are expected to improve the operational characteristics of model-assisted designs for cancer phase I dose-finding studies. Intensive simulation shows that our proposed method is superior to BOIN, mTPI, and Keyboard in the sense that the proposed method provides the lowest probability of overdosing, while maintaining similar performance in MTD selection.

In the next section, we introduce CUSUM statistics and propose a CUSUMIN design. We present simulation studies to examine the operational characteristics of the new method in Section 3, followed by a discussion.

2 | CUSUMIN: CUSUM INTERVAL DESIGN

We propose CUSUMIN design for phase I cancer studies as a simple and transparent model-assisted design. CUSUMIN design is based on the theory of CUSUM control charts. Control charts, which are used in manufacturing processes for stabilizing the quality of products, are well known in the statistical quality control field, but not in clinical research. In Section 2.1, we briefly introduce CUSUM statistics, which are used in our proposed method, and we propose CUSUMIN design in Section 2.2.

2.1 | CUSUM statistics and control charts

CUSUM control charts were first proposed by Page[31] for quality control in the manufacturing process. In general, control charts are used to monitor processes and detect abnormal trends (shifts) in their quality characteristics. Control charts are generally run charts with decision lines that reflect whether or not a process is normal. The statistics calculated by process outcomes, (usually quality characteristics in the quality control field), are plotted on charts along the time course (or sequence of the products), and abnormal trends are detected if the statistic exceeds the thresholds, which are called control limits, to identify potential process abnormalities. In quality control, a process is called in-control if it is normal, and a process is called out-of-control if it is abnormal. There are many types of control charts, and CUSUM control charts are recommended by Hawkins and Olwell,[32] as a frequently used control chart due to its superior performance.

Although various control charts have been developed for various distributions, we assume a case in which the process outcomes have a binomial distribution. When we monitor the number of defectives as a quality characteristic, process outcomes are the number of defectives, and those are usually assumed to be independently distributed in the binomial distribution. In the context of clinical trials, it is usually assumed that the number of patients per cohort who experience DLT is independently distributed in the binomial distribution.

We assume an allowable defective rate of π_0 in a manufacturing process, which is considered in-control. The role of control charts is to detect when the defective rate of the process departs from π_0. When the defective rate departs from π_0, the process is considered out-of-control. The principle of this detection is basically statistical hypothesis testing. The in-control state is the null hypothesis $H_0 : \pi = \pi_0$. The out-of-control state is the alternative hypothesis $H_1 : \pi \neq \pi_0$. This two-sided hypothesis is used in order to detect insufficient defective rates as well as excessive defective rates, considering both fraud and low sensitivity of inspection. If process outcomes (binomial variates) are used directly, well-known hypothesis testing based on binominal distribution including normal approximations is used. Hence, in the context of process monitoring, if the i-th outcome (the number of defectives) is within the critical values, which are the upper and lower control limits, the process is in-control. Otherwise, the process is out-of-control. This control chart, in which the i-th outcome is directly plotted, is called the Shewhart type control chart, which is different from the CUSUM control chart. In CUSUM statistics, information from past outcomes are accumulated and plotted in the CUSUM control chart, which is more powerful than the Shewhart type control chart.[32]

In CUSUM control charts, CUSUM statistics are plotted along the time course. Here, let C_i^- and C_i^+ be the CUSUM statistics for the i-th outcome. C_i^- is the statistic for detecting a downward shift: $\pi < \pi_0$. C_i^+ is the statistic for detecting an upward shift: $\pi > \pi_0$. If outcomes from the process have binomial distributions, the CUSUM statistics C_i^- and C_i^+ are defined as:

$$C_i^- = \min\left(0, C_{i-1}^- + m_i - n_i k_u^-\right), \tag{1}$$

$$C_i^+ = \max\left(0, C_{i-1}^+ + m_i - n_i k_u^+\right), \tag{2}$$

where, m_i is the outcome (number of defectives) in the i-th sample, and n_i is the sample size. Here, $C_0^- = C_0^+ = 0$, and C_{i-1}^- and C_{i-1}^+ are the CUSUM statistics for the $(i-1)$-th binomial outcomes plotted in the CUSUM control chart. The hypothesis that a process is in-control, H_0, is rejected if $C_i^- < h^-$ or $C_i^+ > h^+$ is true, and the process is considered to be out-of-control.

Here, h^- and h^+ are prespecified thresholds (control limits) used to determine whether a process is in-control. The k_u^- and the k_u^+ are the reference values for effectively detecting the downward and upward step shifts, respectively.

These reference values are determined by the $\pi_0, \pi_1 (< \pi_0)$, and $\pi_2 (> \pi_0)$ parameters. The parameters π_1, and π_2 are the defective rates in which we are interested in quickly detecting the downward and upward step shifts, which are in point alternative hypotheses. The reference values are formulated based on the theory of SPRTs for binomial outcomes, as follows:

$$k_u^- = -\frac{\log\left(\frac{1-\pi_1}{1-\pi_0}\right)}{\log\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right)}, \tag{3}$$

$$k_u^+ = -\frac{\log\left(\frac{1-\pi_2}{1-\pi_0}\right)}{\log\left(\frac{\pi_2(1-\pi_0)}{\pi_0(1-\pi_2)}\right)}. \tag{4}$$

The detailed derivation of k_u^- and k_u^+ is provided in Appendix A.

Page[31] explicitly recognized that the detection rule in CUSUM control charts resulted in a sequence of Wald sequential tests. The primary difference between the SPRT and the detection rule in CUSUM charts is that the hypothesis that the process is in-control is never accepted in CUSUM charts. As with other control charts, the CUSUM control charts do not accept the in-control and stop-sampling hypotheses. Instead, the in-control decision can be considered evidence that the null hypothesis is favored, and the SPRT is restarted at that time.

The control limits h^- and h^+ play the same role as critical values in hypothesis testing. In control charts, the control limits h^- and h^+ are set to have the desired power within a specified erroneous detection rate. In a control chart, detection performance and erroneous detection rate are assessed by average number of plots by exceeding the control limits for the first time from the start of control chart, called average run length (ARL).[32,33] A larger value of ARL in the in-control state is desirable since it means that the control chart has a low probability of falsely detecting an out-of-control state in a truly in-control process. On the other hand, a smaller expected ARL in the out-of-control state is desirable since it means that the control chart can detect an out-of-control state quickly. Thus, ARL in the in-control state and ARL in the out-of-control state correspond to type I error and power in hypothesis testing. Notably, CUSUM has the smallest expected run length for out-of-control among all tests, with the same in-control ARL.[34] That is, CUSUM can detect an out-of-control state more quickly than other control charts while maintaining the same ARL in the in-control state.

Finally, we present an example of a CUSUM control chart for binomial outcomes. We assume that the in-control defective rate is $\pi_0 = 0.05$, and the out-of-control rates which we are interested in detecting quickly are $\pi_1 = 0.01$ and $\pi_2 = 0.10$. One hundred products are randomly drawn from the manufacturing process and inspected daily, so the sample size is $n_i = 100$, and then the reference parameters are $k_u^- = 0.02499$ and $k_u^+ = 0.07236$. The control limits are set as $h^- = -4$ and $h^+ = 4$, which holds that the ARL for in-control is 113.6. The CUSUM control chart is shown in Figure 1, and the individual values of the CUSUM statistics, along with the time course, are shown in Table 1. C_3^- departs from the zero line. However, because the plot is above h^-, the process is still considered in-control. C_i^+ rises at the sixth plot (C_6^+), and C_7^+ exceeds the upper control limit $h^+ = 4$. Then, the process is now considered out-of-control, and the defective rate increases again at the seventh plot.

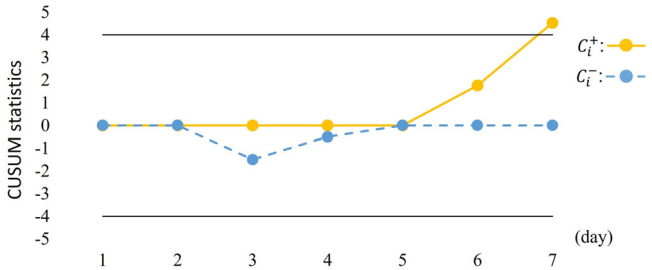

FIGURE 1 An example of a CUSUM control chart

TABLE 1 Individual values of CUSUM control chart example

Day	Sample size	*N* of defect	C_i^-	C_i^+
1	100	4	0	0
2	100	6	0	0
3	100	1	−1.49854222	0
4	100	2	−0.49854222	0
5	100	3	0	0
6	100	9	0	1.76416232
7	100	10	0	4.52832463

2.2 | CUSUMIN design framework

The dose finding study setting is similar to the control chart setting. In the dose-finding study, the assignment to the dose should be continued if the current dose is the MTD. However, if the current dose is not the MTD, dose escalation or de-escalation should be performed as soon as possible. In manufacturing process quality control, if processes depart from the in-control state, the out-of-control state should be detected as quickly as possible, and the corrective action should be conducted to restore the process to the in-control state. In addition, the in-control processes should be maintained. Thus, the treatments that achieve the MTD correspond to in-control, and the other doses correspond to out-of-control. As described in Section 2.1, CUSUM can detect an out-of-control state more quickly than other control charts while maintaining the same ARL in the in-control state. Therefore, the proposed method can be expected to provide a higher probability of detecting incorrect assignments early while maintaining a higher probability of correct assignments than the BOIN design during consecutive patient assignments, that is, the method provides a better balance of maintaining and changing doses.

Assuming that the single agent dose finding study includes J pre-specified doses, and has a target toxicity rate of ϕ, and a true toxicity rate that increases monotonically with dose levels, where p_j is the true toxicity rate at the respective dose level $j\,(j=1,...,J)$, then the true toxicity rate satisfies the relationship $p_1 \le p_2 \le ... \le p_J$, with respect to the dose level. If n_j is the total number of patients treated at dose level j and m_j patients are those who have experienced toxicity, then $\widehat{p}_j = m_j/n_j$ is the observed toxicity rate at dose level j.

The CUSUM statistic is defined at each dose level. Hence, we generalize Equations (1) and (2) to Equations (5) and (6) for the i-th cohort $\left(i=1,...,I_j\right)$ in the dose level j,:

$$C_{i,j}^- = \min\left(0, C_{i-1,j}^- + m_{i,j} - n_{i,j}k_u^-\right), \tag{5}$$

$$C_{i,j}^+ = \max\left(0, C_{i-1,j}^+ + m_{i,j} - n_{i,j}k_u^+\right). \tag{6}$$

Here, $C_{0,j}^- = C_{0,j}^+ = 0$, $C_{i-1,j}^-$ and $C_{i-1,j}^+$ are the CUSUM statistics for the $i-1$-th cohort in dose level j, $n_{i,j}$ is the number of patients of the i-th cohort for dose level j, and $m_{i,j}$ is the number of patients who experienced DLT in the i-th cohort for dose level j. The CUSUM statistics of the previous cohort, $C_{i-1,j}^-$ and $C_{i-1,j}^+$, are updated based on the newly observed outcomes of $n_{i,j}$ and $m_{i,j}$, and then CUSUM statistics for the i-th cohort in dose level j are obtained. This formulation is a simple expression fixed to dose level j for ease of understanding; a more strictly mathematical expression is shown in Appendix B.

In the dose-finding study setting, the hypothesis that the process is in-control, H_{0j}, is considered as $p_j = \phi$. The alternative two point-hypotheses are $H_{1j}: p_j = \phi_1$ and $H_{2j}: p_j = \phi_2$, where ϕ_1 and ϕ_2 are the values which we need to detect: subtherapeutic dose and too toxic dose, respectively. Though we have J CUSUM statistics, it is possible to detect whether the toxicity probability of each dose level departs from the target toxicity ϕ since the reference values k_u^-, k_u^+ and control limits h^-, h^+ are the same for all the dose levels. Thus, if the current dose is the MTD, the CUSUM statistics (5) and (6) are expected to be within the control limits. If the current dose level is lower than the MTD, the CUSUM statistic $C_{i,j}^-$ goes down and the current dose level is found to be lower than the MTD if the $C_{i,j}^-$ exceeds the lower control

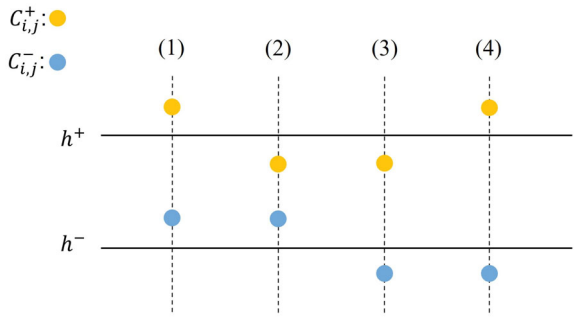

FIGURE 2 Relationship of CUSUM statistics and the boundaries

limit h^{-}. If the current dose level is higher than the MTD, $C_{i,j}^{+}$ goes up and the current dose level is found to be higher than the MTD if the $C_{i,j}^{+}$ exceeds the upper control limit h^{+}. Consequently, by replacing π_0, π_1, and π_2 of the reference values in (3) and (4) with ϕ, ϕ_1, and ϕ_2, respectively, the reference value equations become:

$$k_u^{-} = -\frac{\log\left(\frac{1-\phi_1}{1-\phi}\right)}{\log\left(\frac{\phi_1(1-\phi)}{\phi(1-\phi_1)}\right)}, \tag{7}$$

$$k_u^{+} = -\frac{\log\left(\frac{1-\phi_2}{1-\phi}\right)}{\log\left(\frac{\phi_2(1-\phi)}{\phi(1-\phi_2)}\right)}. \tag{8}$$

Note, throughout the study, we specify that $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$ which are the recommended thresholds in the original BOIN paper.[20]

We propose a model-assisted design based on CUSUM control charts theory, herein called CUSUMIN (CUSUM INterval) design. The algorithm is as follows:

i. Patients in the first cohort are treated at the lowest dose level 1.
ii. At the current dose level j, $n_{i,j}$ is the total number of patients in i-th cohort of dose level j; $m_{i,j}$ patients have experienced DLT in i-th cohort of dose level j. The CUSUM statistics, $C_{i-1,j}^{-}$ and $C_{i-1,j}^{+}$, are updated to $C_{i,j}^{-}$ and $C_{i,j}^{+}$ based on $n_{i,j}$ and $m_{i,j}$.
iii. The next dose level, j', is determined using the following rules:

(1) if ${C_{i,j}}^{+} \geq h^{+}$ and ${C_{i,j}}^{-} > h^{-}$, $\quad j' = j-1$

(2) if ${C_{i,j}}^{+} \leq h^{+}$ and ${C_{i,j}}^{-} \geq h^{-}$, $\quad j' = j$

(3) if ${C_{i,j}}^{+} \leq h^{+}$ and ${C_{i,j}}^{-} < h^{-}$, $\quad j' = j+1$

(4) if ${C_{i,j}}^{+} > h^{+}$ and ${C_{i,j}}^{-} < h^{-}$, $\quad j'$ is determined as :

if $\left|{C_{i,j}}^{+} - h^{+}\right| < \left|{C_{i,j}}^{-} - h^{-}\right|$ $\quad j' = j+1$

Otherwise $\quad j' = j-1$

Here, h^{-} and h^{+} are the optimal boundaries in CUSUMIN design, and are determined so that the design has good operational characteristics according to the study objectives. The relationships between ${C_{i,j}}^{-}$, ${C_{i,j}}^{+}$ and h^{-}, h^{+} for each above condition (1)–(4), are depicted in Figure 2.

(ii) and (iii) are repeated until the total number of treated patients reaches the maximum sample size or the trial is terminated because of excessive toxicity. Liu and Yuan[20] proposed implementing the following safety rule, to terminate the study early if the study agent is excessively toxic in BOIN.

If $pr\left(p_j > \phi | n_j, m_j\right) > 0.95$ and $n_j \geq 3$, dose levels j and higher are eliminated from the trial, and the trial is terminated if the first dose level is eliminated. We assume that m_j follows a binomial distribution and that p_j follows a vague beta prior $p_j \sim Beta(1,1)$.

We implement the same safety rule for CUSUMIN.

As in BOIN,[20] after the trial is completed, an isotonic regression is performed to ensure that the estimated toxicity rates satisfy the monotonicity assumption and that the MTD is selected based on the regression estimator. $\widetilde{p}_j$ refers to the isotonic regression estimate of the observed toxicity rate $\widehat{p}_j$ for dose level j. The MTD, j^*, is selected as the dose whose isotonic regression estimator of toxicity rate is closest to the target toxicity rate of ϕ, as described below:

$$j^* = arg \min_{j=(1,\ldots,J)} \left|\widetilde{p}_j - \phi\right|.$$

If multiple doses are tied for $\widetilde{p}_{j^*}$, we select the highest dose level that satisfies $\widetilde{p}_{j^*} < \phi$ or the lowest dose level that satisfies $\widetilde{p}_{j^*} > \phi$. The isotonic estimator can be obtained by applying the pooled adjunct violator algorithm[35,36] to $\widehat{p}_j$.

As shown by the algorithm above, the dose for the next cohort is decided only by CUSUM statistics and the predefined interval of the CUSUM control chart in CUSUMIN. The computation of CUSUM statistics and dose escalation/de-escalation are as simple as in BOIN.

Though h^- and h^+ are determined by the ARLs for in-control and out-of-control in the field of control charts, this is not appropriate in dose finding because there the primary indicator is not the ARL. In this study, the optimal thresholds h^- and h^+ are determined to ensure that the CUSUMIN design performs well in terms of the MTD selection probability, the probability of assigning patients to the MTD, and controlling the probability of overdosing. Thus, we propose to define a utility function $u\left(h^-, h^+\right)$ and search for the optimal interval via simulation studies. Generally, the utility function can be expressed as follows: The optimal thresholds are determined as the thresholds that maximize the utility function by grid search in terms of $\left(h^-, h^+\right)$.

$$u\left(h^-, h^+\right) = f_1\left(r_1 | h^-, h^+\right) + f_2\left(r_2 | h^-, h^+\right) + f_3\left(r_3 | h^-, h^+\right). \tag{9}$$

Here, r_1, r_2, and r_3 are the factors that comprise the utility, that is, the MTD selection probability, the probability of assigning patients to the MTD, and the overdosing probability. In addition, f_1, f_2, and f_3 are the functions of these

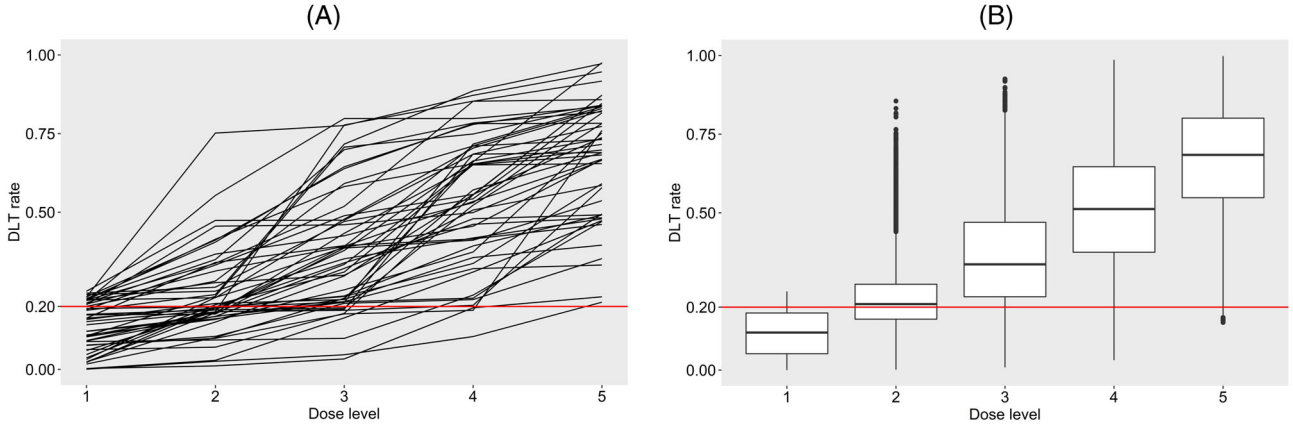

FIGURE 3 (A) Fifty randomly selected scenarios. (B) The distribution of DLT rates according to dose level from the 10,000 scenarios with five dose levels

factors, respectively. The functions must be determined through discussions with investigators, to reflect the importance of the factors. If we assume that the importance of the factors in the study are in order of the probability of assigning a patient the MTD, the probability of MTD selection, and the overdosing control, then the following functions can specify f_1,f_2, and f_3:

$$f_1 = -\frac{5\times\left(r'_1 - r_1\right)}{\sqrt{r'_1(1-r'_1)}},\ f_2 = -\frac{3\times\left(r'_2 - r_2\right)}{\sqrt{r'_2(1-r'_2)}},\ f_3 = \frac{\left(r'_3 - r_3\right)}{\sqrt{r'_3(1-r'_3)}}. \tag{10}$$

Here, r'_1,r'_2, and r'_3 are MTD selection, assigning patients to MTD, and the overdose probability in BOIN, respectively. Our proposed method is developed in the same framework as BOIN; the study phase, study aim, and importance of simplicity and transparency for investigators are included; and the utility functions are set as a form of comparison with the BOIN method. Practically, as described above, $u\left(h^-, h^+\right)$ and r'_1,r'_2, and r'_3 are evaluated by simulation in the same setting.

3 | SIMULATION STUDY

To examine the performance of our proposed method, we make comparisons of its operational characteristics with BOIN design via simulation studies. For fair comparison, we explore the optimal thresholds for the BOIN design in the simulation studies in the same manner as for the CUSUMIN design. As comparative model-assisted design methodologies, mTPI design and Keyboard design are also compared in this simulation.

3.1 | Simulation settings

In the simulation studies, we generate true dose-toxicity scenarios based on the pseudo-uniform algorithm proposed by Clertant and O′Quigley,[37] to avoid cherry-picking scenarios biased toward specific methods. Given a target toxicity rate ϕ and total dose level J, true dose-toxicity scenarios are generated as follows:

TABLE 2 Optimal boundaries and the value of utility function for CUSUMIN designs and BOIN designs

			Target ϕ		
Number of doses	**Method**		**0.2**	**0.25**	**0.3**
5 doses	BOIN	Optimal boundaries for Original BOIN (ϕ_1, ϕ_2)	(0.12, 0.28)	(0.15, 0.35)	(0.18, 0.42)
		Optimal boundaries for Original BOIN (ϕ_1, ϕ_2)	(0.065, 0.25)	(0.08125, 0.35)	(0.1425, 0.525)
		Value of Utility function $u(h^-, h^+)$	2.26	12.74	6.13
	CUSUMIN	Optimal boundaries (h^-, h^+)	(−1.3, 0.9)	(−1.2, 2.7)	(−0.8, 0.8)
		Value of Utility function $u(h^-, h^+)$	11.92	11.21	2.02
		r'_1, r'_2, r'_3 (%)	53.37, 50.59, 17.66	56.30, 45.48, 18.76	54.05, 44.09, 18.94
8 doses	BOIN	Optimal boundaries for Optimal BOIN (ϕ_1, ϕ_2)	(0.12, 0.28)	(0.15, 0.35)	(0.18, 0.42)
		Optimal boundaries for Optimal BOIN (ϕ_1, ϕ_2)	(0.04, 0.25)	(0.08125, 0.425)	(0.1425, 0.4875)
		Value of Utility function $u(h^-, h^+)$	10.66	16.42	4.80
	CUSUMIN	Optimal boundaries (h^-, h^+)	(−0.9, 0.6)	(−0.6, 0.8)	(−0.8, 1.9)
		Value of Utility function $u(h^-, h^+)$	17.17	4.12	5.81
		r'_1, r'_2, r'_3 (%)	44.87, 39.09, 18.39	46.65, 34.45, 17.26	42.99, 32.19, 18.31

TABLE 3 Average performance of each design across 10,000 scenarios with five dose levels

5 doses					
Target $\phi = 0.20$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	53.37	50.59	17.66	8.52	33.17
BOIN (Optimal)	53.12	50.79	16.29	8.35	33.17
mTPI	51.48	50.03	23.31	19.54	33.05
Keyboard	53.32	50.72	17.26	8.52	33.17
CUSUMIN	53.04	48.68	7.48	1.18	33.08
Target $\phi = 0.25$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	56.30	45.48	18.76	8.53	34.96
BOIN (Optimal)	55.69	45.97	12.53	5.09	34.94
mTPI	54.34	47.91	19.32	13.94	34.96
Keyboard	56.30	45.48	18.76	8.53	34.96
CUSUMIN	55.50	44.15	8.10	2.30	34.91
Target $\phi = 0.30$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	54.05	44.09	18.94	9.93	34.76
BOIN (Optimal)	53.55	45.11	16.98	10.00	34.76
mTPI	53.06	45.91	18.36	13.44	34.78
Keyboard	53.87	44.01	19.24	9.93	34.76
CUSUMIN	54.23	40.13	9.49	1.16	34.72

TABLE 4 Average performance of each design across 10,000 scenarios with eight dose levels

8 doses					
Target $\phi = 0.20$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	44.87	39.09	18.39	9.25	34.06
BOIN (Optimal)	44.83	39.17	14.31	7.92	34.06
mTPI	43.40	39.43	24.36	20.78	34.08
Keyboard	44.53	39.10	18.13	9.27	34.06
CUSUMIN	45.94	38.01	13.34	3.58	34.09
Target $\phi = 0.25$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	46.65	34.45	17.26	7.43	35.47
BOIN (Optimal)	46.33	37.40	16.89	10.97	35.41
mTPI	45.99	37.56	17.51	12.42	35.41
Keyboard	46.65	34.45	17.26	7.43	35.47
CUSUMIN	46.69	32.65	11.56	2.06	35.36
Target $\phi = 0.30$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	42.99	32.19	18.31	9.24	35.41
BOIN (Optimal)	42.72	32.35	15.79	8.76	35.41
mTPI	42.03	33.18	16.80	11.69	35.35
Keyboard	43.01	32.22	18.53	9.24	35.41
CUSUMIN	43.31	28.63	8.48	1.97	35.39

TABLE 5 Optimal boundaries and the value of utility function of CUSUMIN designs and BOIN designs for sensitivity analysis

			Target ϕ		
Number of doses	**Method**		**0.2**	**0.25**	**0.3**
5 doses	BOIN	Optimal boundaries for Optimal BOIN (ϕ_1, ϕ_2)	(0.04, 0.22)	**(0.08125, 0.35)**	(0.1125, 0.45)
		Value of Utility function $u(h^-, h^+)$	3.60	14.22	6.58
	CUSUMIN	Optimal boundaries (h^-, h^+)	**(−1.3, 0.9)**	**(−1.2, 2.7)**	**(−0.8, 0.8)**
		Value of Utility function $u(h^-, h^+)$	17.06	17.11	9.26
		r'_1, r'_2, r'_3 (%)	53.37, 50.59, 17.66	56.30, 45.48, 18.76	54.05, 44.09, 18.94
8 doses	BOIN	Optimal boundaries for Optimal BOIN (ϕ_1, ϕ_2)	**(0.04, 0.25)**	**(0.08125, 0.425)**	**(0.1425, 0.4875)**
		Value of Utility function $u(h^-, h^+)$	10.64	11.49	5.55
	CUSUMIN	Optimal boundaries (h^-, h^+)	(−1.0, 2.8)	**(−0.6, 0.8)**	**(−0.8, 1.9)**
		Value of Utility function $u(h^-, h^+)$	18.56	7.75	12.13
		r'_1, r'_2, r'_3 (%)	44.87, 39.09, 18.39	46.65, 34.45, 17.26	42.99, 32.19, 18.31

i. Select dose level k from J dose levels as the MTD with equal probabilities.
ii. Sample $M \sim Beta(\max\{J-k;0.5\},1)$, where, k is the selected dose level in the previous step (i) and set an upper bound $B = \phi + (1-\phi) \times M$ for the toxicity probabilities.
iii. Repeat sampling J toxicity probabilities uniformly on $[0,B]$ until these correspond to a scenario in which the dose level k is the MTD.

Repeat (i)–(iii) until a sufficient number of scenarios is generated. In these scenarios, the MTD can be defined as the dose whose DLT rate is closest to the target toxicity rate ϕ. In our simulation, MTD is defined as the dose whose DLT rate is closest to the target toxicity ϕ with $|p_k - \phi| < 0.05$.

In the simulation studies, we consider the target toxicity rates of $\phi = 0.20, 0.25,\ 0.30$ and dose levels of $J = 5$ and 8, with a maximum sample size of 36 patients, in 12 cohorts of three patients.

Under each setting, 10,000 scenarios are randomly generated, and trials are conducted once for each scenario. In Figure 3, 50 randomly selected scenarios and the distribution of the toxicity probabilities by dose level from the 10,000 scenarios with $\phi = 0.20$ and $J = 5$ are displayed. The results exhibit a variety of dose-toxicity curve shapes.

For CUSUMIN design, the optimal thresholds (h^-, h^+) are determined in the simulation. In this simulation study, we use the utility function Equation (9) with Equation (10) described in Section 2.2. To evaluate Equation (9), we first evaluate r'_1, r'_2, and r'_3 via simulation, and then evaluate Equation (9) in the range of $(h^-, h^+) = [-4,0] \times [0,4]$, with a width of 0.1 for each threshold. In the BOIN design, $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$ are used for the boundaries as recommended in the original BOIN paper.[20] We call this design the Original BOIN hereafter. In the simulation studies, for fair comparison, we evaluate the BOIN design with boundaries (ϕ_1, ϕ_2) optimized based on the same utility function used for the CUSUMIN design. For optimizing boundaries of the BOIN design, the utility function (9) can be described as the function of ϕ_1 and ϕ_2, $u(\phi_1, \phi_2) = f_1(r_1|\phi_1, \phi_2) + f_2(r_2|\phi_1, \phi_2) + f_3(r_3|\phi_1, \phi_2)$. We evaluate $u(\phi_1, \phi_2)$ in the range of $(\phi_1, \phi_2) = [0.025k\phi, (1 + 0.025k)\phi]^2, (k = 1, ..., 39)$. The optimal thresholds for the BOIN design are determined as the thresholds that maximize the utility function by grid search in terms of (ϕ_1, ϕ_2), the same way as for the CUSUMIN design. We call the BOIN design with optimized boundaries the Optimal BOIN hereafter to distinguish it from the Original BOIN. The optimal thresholds (h^-, h^+) and (ϕ_1, ϕ_2), the values of the utility function, and r'_1, r'_2, and r'_3, which are used for evaluation of Equation (9), are shown in Table 2. For the comparison, the optimal threshold for Original BOIN are also shown in Table 2.

We also evaluate mTPI design and Keyboard design to provide more information about the position of our proposed method. We apply the same safety rule and MTD selection rule as in CUSUMIN design also to Original BOIN, Optimal BOIN, mTPI, and Keyboard for the comparison of operational characteristics.

R for Windows, release 4.1.0 is used for the simulation studies.

TABLE 6 Sensitivity analysis of CUSUMIN designs across 10,000 scenarios with five dose levels

5 doses					
Target $\phi = 0.20$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	53.37	50.59	17.66	8.52	33.17
BOIN (Optimal)	52.42	50.10	13.36	5.57	33.08
mTPI	51.48	50.03	23.31	19.54	33.05
Keyboard	53.32	50.72	17.26	8.52	33.17
CUSUMIN	**53.04**	**48.68**	**7.48**	**1.18**	**33.08**
Target $\phi = 0.25$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	56.30	45.48	18.76	8.53	34.96
BOIN (Optimal)	**55.69**	**45.97**	**12.53**	**5.09**	**34.94**
mTPI	54.34	47.91	19.32	13.94	34.96
Keyboard	56.30	45.48	18.76	8.53	34.96
CUSUMIN	**55.50**	**44.15**	**8.10**	**2.30**	**34.91**
Target $\phi = 0.30$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	54.05	44.09	18.94	9.93	34.76
BOIN (Optimal)	53.35	44.68	15.64	8.75	34.76
mTPI	53.06	45.91	18.36	13.44	34.78
Keyboard	53.87	44.01	19.24	9.93	34.76
CUSUMIN	**54.23**	**40.13**	**9.49**	**1.16**	**34.72**

TABLE 7 Sensitivity analysis of CUSUMIN designs across 10,000 scenarios with eight dose levels

8 doses					
Target $\phi = 0.20$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	44.87	39.09	18.39	9.25	34.06
BOIN (Optimal)	**44.83**	**39.17**	**14.31**	**7.92**	**34.06**
mTPI	43.40	39.43	24.36	20.78	34.08
Keyboard	44.53	39.10	18.13	9.27	34.06
CUSUMIN	44.38	36.64	6.17	1.70	34.02
Target $\phi = 0.25$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	46.65	34.45	17.26	7.43	35.47
BOIN (Optimal)	**46.33**	**37.40**	**16.89**	**10.97**	**35.41**
mTPI	45.99	37.56	17.51	12.42	35.41
Keyboard	46.65	34.45	17.26	7.43	35.47
CUSUMIN	**46.69**	**32.65**	**11.56**	**2.06**	**35.36**
Target $\phi = 0.30$					
Method	**MTD Selection rate (%)**	**Patients treated at MTD (%)**	**Patients treated above MTD (%)**	**Risk of overdosing 60% (%)**	**Average sample size**
BOIN (Original)	42.99	32.19	18.31	9.24	35.41
BOIN (Optimal)	**42.72**	**32.35**	**15.79**	**8.76**	**35.41**
mTPI	42.03	33.18	16.80	11.69	35.35
Keyboard	43.01	32.22	18.53	9.24	35.41
CUSUMIN	**43.31**	**28.63**	**8.48**	**1.97**	**35.39**

3.2 | Results

The CUSUMIN design is evaluated by comparing it with the Original BOIN design and the Optimal BOIN design. In addition, for reference, mTPI design and Keyboard design are also compared with the proposed design. The measures we evaluate include the MTD selection rate, percentage of patients treated at MTD, percentage of patients treated above MTD, and percentage of the study in which over 60% of patients are treated above MTD (hereinafter, risk of overdosing 60%) for 10,000 randomly generated scenarios with target toxicity rates $\phi = 0.20, 0.25,$ and 0.30, and dose levels $J = 5$ and 8.

The simulation results of each method across 10,000 scenarios with five dose levels are shown in Table 3. Among all target toxicity rates, the percentages of patients treated above MTD and the values of risk of overdosing 60% for the CUSUMIN design are the lowest among the designs. The mTPI design is the worst in terms of overdosing: its values of risk of overdosing 60% are the largest among the designs, and the percentage of patients treated above MTD is the largest in almost cases. The percentages of patients treated above MTD of Optimal BOIN are the second lowest, and risk of overdosing 60% shows almost the same result (slightly higher when target toxicity is 0.25). The assignment of CUSUMIN designs is noticeably safer. For example, in the case of $\phi = 0.2$, the percentage of patients treated above MTD in the CUSUMIN design is 7.48, while those of Original BOIN, Optimal BOIN, mTPI and Keyboard are 17.66, 16.29, 23.31, and 17.26, respectively. This tendency is shown throughout the target toxicity rates. The MTD selection rates of the CUSUMIN design are similar to those of the other designs throughout the target toxicity rates. This shows that CUSUMIN design provides safer dose escalation and de-escalation while achieving similar performance on selection of MTD. It is also observed that its percentage of patients treated at MTD is slightly lower than those of the other designs, due to the safer dose assignment, which CUSUMIN provided.

We evaluate the case of eight dose levels in the same way. The results are shown in Table 4. The same trends which are shown with five dose levels is observed. The percentage of patients treated above MTD and risk of overdosing 60% of CUSUMIN are the lowest among the designs, and those of Optimal BOIN is the second lowest among the designs. MTD selection rates of the CUSUMIN design are similar to the other designs. The percentage of patients treated at MTD is slightly lower than for the other designs, due to the safer dose assignment which CUSUMIN provided. In the eight doses case, CUSUMIN design still provides safer dose escalation and de-escalation among the designs while providing similar MTD selection rates throughout the target toxicity rates.

Although we use the utility function (9) with (10), the optimal threshold depends on the utility function. Thus, the performance of the CUSUMIN design and of Optimal BOIN vary according to the specification of the utility function. To evaluate the sensitivity of the utility function specification, we conduct a sensitivity analysis using the utility function (9) with the different $f_1, f_2,$ and f_3 variables shown below.

$$f_1 = -\frac{3 \times (r'_1 - r_1)}{\sqrt{r'_1(1 - r'_1)}}, \; f_2 = -\frac{2 \times (r'_2 - r_2)}{\sqrt{r'_2(1 - r'_2)}}, \; f_3 = \frac{(r'_3 - r_3)}{\sqrt{r'_3(1 - r'_3)}}. \tag{11}$$

In this sensitivity analysis, we specify the utility function with importance in the order of the MTD selection rate, assigning patients to MTD, and the overdosing control. The optimal thresholds (h^-, h^+) and (ϕ_1, ϕ_2), the values of utility function, and $r'_1, r'_2,$ and r'_3 which are used for evaluation of Equation (11), are shown in Table 5. Almost all thresholds produced the same results as in Table 2 and are displayed in bold. In addition, actual performance results of Optimal BOIN and CUSUMIN are almost always the same as in Tables 3 and 4 and are displayed in Tables 6 and 7 in bold. These results further demonstrated that CUSUMIN design provides a safer and more accurate dose escalation/de-escalation rule, while maintaining a similar MTD selection rate. Our findings confirm that the CUSUMIN design is robust against the specification of the utility function.

4 | DISCUSSION

We developed CUSUMIN as a new method of model-assisted design based on a framework similar to that of BOIN. The proposed method is based on the CUSUM statistics, which was originally created for control charts for manufacturing process monitoring, as an extension of SPRT. Intensive simulation shows that our proposed method is superior to BOIN, mTPI, and Keyboard in the sense that the proposed method provides the lowest probability of overdosing among

the designs evaluated in this study, while maintaining similar performance in MTD selection. The characteristics of our proposed method, providing safer assignment while maintaining MTD selection performance, are attractive from the point of view of protecting the patient from the risk of overdosing. The proposed method can provide more accurate patient assignments than BOIN using CUSUM statistics. Theoretically, a more accurate patient assignment to the dose can be understood as an advantage of the characteristics of CUSUM statistics. This statistic, which is created for control charts, has proven to be optimal in the sense that CUSUM has the smallest expected run length of out-of-control among all tests with the same in-control ARL.[34] This is why the proposed method can provide a higher probability of detecting incorrect assignments early while maintaining a higher probability of correct assignments than the BOIN design during consecutive patient assignments, that is, the method provides a better balance of maintaining and changing the doses.

The dose determination in our proposed method is achieved by comparing CUSUM statistics and boundaries, and the calculation of CUSUM statistics is simple. The rule of dose determination is completely pre-defined and transparent, similar to those of other model-assisted designs.

To implement our proposed method, the CUSUM statistics boundaries must be determined in advance. In this study, we used the utility function to determine the optimal boundaries of the CUSUMIN design. The utility function in this study consists of the following factors: MTD selection probability, probability of assigning patients to MTD, and overdosing probability. The utility function varies among studies, indications, drug properties, and so forth, and it is necessary to determine the appropriate utility function by discussion with investigators. In this study, we evaluated the sensitivity of the utility function via a sensitivity analysis of the simulation studies. The sensitivity analysis confirms that the CUSUMIN design is robust against the specification of the utility function.

We evaluated the performance of our proposed method via intensive simulation studies, in which 10,000 random scenarios are used to avoid cherry-picking scenarios toward specific methods. The optimal boundaries for CUSUM designs are determined according to the simulation results of the 10,000 generated random scenarios to ensure that the design has good performance in terms of safety, accuracy of MTD selection, and assignment of a large portion of patients to MTD, on average, to the various 10,000 scenarios. We determined the optimal boundaries for BOIN design based on the same utility function used for CUSUMIN design, to perform a fair comparison. Though we examined the performance of our proposed method mainly based on random scenarios, if the feasible specific dose toxicity curve is available from an analogous prior study with similar drugs or pre-clinical information when the study is planned, the optimal boundaries can be determined based on several specific dose toxicity curve scenarios. We conducted an additional simulation to evaluate the performance of the CUSUMIN design with respect to specific dose toxicity curves. The results of the additional simulation are shown in Table S3 of the web-based appendix. As expected, the CUSUMIN design is shown to reduce the probability of overdosing compared with BOIN and the other model-assisted designs, while maintaining similar performance in MTD selection.

ACKNOWLEDGMENTS

We thank the referees for their thoughtful and constructive comments and suggestions.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Data available on request from the authors. The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Tomoyoshi Hatayama https://orcid.org/0000-0002-4384-5148

REFERENCES

1. Zhou H, Murray TA, Pan H, Yuan Y. Comparative review of novel model-assisted designs for phase I clinical trials. *Stat Med*. 2018; 37(14):2208-2222.
2. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. *J Natl Cancer Inst*. 2009;101(10):708-720.
3. Durham SD, Flournoy N, Rosenberger WF. A random walk rule for phase I clinical trials. *Biometrics*. 1997;53(2):745-760.
4. Yin G. *Clinical Trial Design: Bayesian and Frequentist Adaptive Methods*. John Wiley & Sons; 2012.
5. Storer BE. *Choosing a Phase I Design. Handbook of Statistics in Clinical Oncology*. Chapman & Hall/CRC; 2005:79-96.

6. Storer BE. Design and analysis of phase I clinical trials. *Biometrics*. 1989;45(3):925-937.
7. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics*. 1990; 46(1):33-48.
8. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med*. 1998;17(10):1103-1120.
9. Yin G, Zheng S, Xu J. Fractional dose-finding methods with late-onset toxicity in phase I clinical trials. *J Biopharm Stat*. 2013;23(4): 856-870.
10. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics*. 2000;56(4):1177-1182.
11. Yuan Z, Chappell R, Bailey H. The continual reassessment method for multiple toxicity grades: a Bayesian quasi-likelihood approach. *Biometrics*. 2007;63(1):173-179.
12. Yin G, Yuan Y. A latent contingency table approach to dose finding for combinations of two agents. *Biometrics*. 2009;65(3):866-875.
13. Yin G, Yuan Y. Bayesian dose finding in oncology for drug combinations by copula regression. *J R Stat Soc C-Appl*. 2009;58(2):211-224.
14. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*. 2004;60(3):684-693.
15. Yuan Y, Yin G. Bayesian dose finding by jointly modelling toxicity and efficacy as time-to-event outcomes. *J R Stat Soc C-Appl*. 2009; 58(5):719-736.
16. Yin G, Yuan Y. Bayesian model averaging continual reassessment method in phase I clinical trials. *J Am Stat Assoc*. 2009;104(487): 954-968.
17. Wages NA, Conaway MR, O'Quigley J. Continual reassessment method for partial ordering. *Biometrics*. 2011;67(4):1555-1563.
18. Ji Y, Liu P, Li Y, Bekele BN. A modified toxicity probability interval method for dose-finding trials. *Clin Trials*. 2010;7(6):653-663.
19. Yan F, Mandrekar SJ, Yuan Y. Keyboard: a novel Bayesian toxicity probability interval design for phase I clinical trials. *Clin Cancer Res*. 2017;23(15):3994-4003.
20. Liu S, Yuan Y. Bayesian optimal interval designs for phase I clinical trials. *J R Stat Soc C-Appl*. 2015;64(3):507-523.
21. Yuan Y, Hess KR, Hilsenbeck SG, Gilbert MR. Bayesian optimal interval design: a simple and well-performing design for phase I oncology trials. *Clin Cancer Res*. 2016;22(17):4291-4301.
22. Lin R, Yin G. Bayesian optimal interval design for dose finding in drug-combination trials. *Stat Methods Med Res*. 2017;26(5):2155-2167.
23. Lin R, Yin G. STEIN: a simple toxicity and efficacy interval design for seamless phase I/II clinical trials. *Stat Med*. 2017;36(26):4106-4120.
24. Takeda K, Taguri M, Morita S. BOIN-ET: Bayesian optimal interval design for dose finding based on both efficacy and toxicity outcomes. *Pharm Stat*. 2018;17(4):383-395.
25. Yuan Y, Lin R, Li D, Nie L, Warren KE. Time-to-event Bayesian optimal interval design to accelerate phase I trials. *Clin Cancer Res*. 2018;24(20):4921-4930.
26. Mu R, Yuan Y, Xu J, Mandrekar SJ, Yin J. gBOIN: a unified model-assisted phase I trial design accounting for toxicity grades, and binary or continuous end points. *J R Stat Soc C-Appl*. 2019;68(2):289-308.
27. Zhou H, Yuan Y, Nie L. Accuracy, safety, and reliability of novel phase I trial designs. *Clin Cancer Res*. 2018;24(18):4357-4364.
28. Wald A. Sequential tests of statistical hypotheses. *Ann Math Stat*. 1945;16(2):117-186.
29. Wald A. *Sequential Analysis*. Wiley; 1947.
30. Tartakovsky A, Nikiforov I, Basseville M. *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. CRC Press; 2014.
31. Page ES. Continuous inspection schemes. *Biometrika*. 1954;41(1/2):100-115.
32. Hawkins DM, Olwell DH. *Cumulative Sum Charts and Charting for Quality Improvement*. Springer Science and Business Media; 2012.
33. Montgomery DC. *Introduction to Statistical Quality Control*. 6th ed. Jhon Wiley & Sons, Inc; 2008.
34. Moustakides GV. Optimal stopping times for detecting changes in distributions. *Ann Stat*. 1986;14(4):1379-1387.
35. Barlow RE, Bartholomew DJ, Bremner J, Brunk HD. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. Wiley; 1972.
36. Bril G, Dykstra R, Pillers C, Robertson T. Algorithm AS 206: isotonic regression in two independent variables. *J R Stat Soc C-Appl*. 1984; 33(3):352-357.
37. Clertant M, O'Quigley J. Semiparametric dose finding methods. *J R Stat Soc B*. 2017;79(5):1487-1508.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hatayama T, Yasui S. CUSUMIN: A cumulative sum interval design for cancer phase I dose finding studies. *Pharmaceutical Statistics*. 2022;21(6):1324-1341. doi:10.1002/pst.2247

APPENDIX A

The CUSUM statistics and the derivation of k_u^- and k_u^+

In Appendix A, we explain that the CUSUM statistics can be understood as an expansion of the SPRT, and derive the parameters k_u^- and k_u^+.

The relationship of CUSUM and SPRT

Consider a simple null hypothesis, the state of in-control ($H_0 : \pi = \pi_0$) and a simple alternative hypothesis, the state of out-control ($H_1 : \pi = \pi_1 \neq \pi_0$). Associated with each hypothesis is a probability density function, $f_0(x)$ and $f_1(x)$, respectively. Let $\{X_i\}$ be a sequence of independent observations with length n. The log-likelihood ratio in terms of $f_0(x)$ and $f_1(x)$ based on $\{X_i\}$ can be expressed as follows:

$$\log \Lambda_n = \sum_{i=1}^{n} \log \frac{f_1(X_i)}{f_0(X_i)}.$$

Let A and B $(A < B)$ be the selected boundaries for the SPRT. The SPRT tests the hypothesis based on the following rule.[28,32]

(i) $\log \Lambda_n \leq A$	Accept H_0
(ii) $\log \Lambda_n \geq B$	Accept H_1
(iii) $A < \log \Lambda_n < B$	Samples new observation

The primary difference between CUSUM and SPRT is that the hypothesis of in-control, H_0, is never accepted in CUSUM. That is, it is never judged that the process is in-control and stops sampling. Instead, CUSUM restarts the test each time the evidence favors the null hypothesis H_0. This CUSUM rule, the restarting of the test, can be described algebraically, as below.[32]

$$C_i^- = \min\left(0, C_{i-1}^- + \log \frac{f_1(X_i)}{f_0(X_i)} - A\right), \quad (\text{if } \pi_0 < \pi_1)$$

$$C_i^+ = \max\left(0, C_{i-1}^+ + \log \frac{f_1(X_i)}{f_0(X_i)} - A\right). \quad (\text{if } \pi_0 > \pi_1)$$

$$\left(C_0^- = C_0^+ = 0\right)$$

In CUSUM, $A = 0$ is used as the boundary to favor H_0.[32] Therefore, C_i^+ and C_i^- are simplified as follows:

$$C_i^- = \min\left(0, C_{i-1}^- + \log \frac{f_1(X_i)}{f_0(X_i)}\right), \quad (\text{if } \pi_0 < \pi_1) \tag{A1}$$

$$C_i^+ = \max\left(0, C_{i-1}^+ + \log \frac{f_1(X_i)}{f_0(X_i)}\right). \quad (\text{if } \pi_0 > \pi_1) \tag{A2}$$

The derivation of k_u^- AND k_u^+

Here, we introduce the parameters π_1 and π_2 as the defective rates of interest for detecting downward and upward step shifts quickly, which are in point alternative hypotheses. Thus, the simple hypotheses can be expressed as follows:

$$H_0 : \pi = \pi_0$$

$$H_1 : \pi = \pi_1 \ (\pi_1 < \pi_0)$$

$$H_2 : \pi = \pi_2 \ (\pi_2 > \pi_0)$$

In this study, we analyze binary toxicity outcomes based on binomial distribution. The density functions $f_0(x)$, $f_1(x)$, and $f_2(x)$ associated with H_0, H_1, and H_2, can be written as $f_0(x) = binomial(n_i, \pi_0)$, $f_1(x) = binomial(n_i, \pi_1)$, and $f_2(x) = binomial(n_i, \pi_2)$, where, n_i is the sample size of i-th sequence (cohort size in this study). The likelihood ratio in terms of $f_0(x)$ and $f_1(x)$ based on the observation X_i is expressed as

$$\log\frac{f_1(X_i)}{f_0(X_i)} = n_i \log\frac{1-\pi_1}{1-\pi_0} + X_i \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)} = \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}\left(X_i + n_i \log\frac{1-\pi_1}{1-\pi_0} \Big/ \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}\right) \tag{A3}$$

Replace $\log\frac{1-\pi_1}{1-\pi_0} / \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}$ by k_u^- in (A3), the expression of (A4) is obtained.

$$\log\frac{f_1(X_i)}{f_0(X_i)} = \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}\left(X_i - n_i k_u^-\right) \tag{A4}$$

Similarly, the likelihood ratio in terms of $f_0(x)$ and $f_2(x)$ is given by (A5).

$$\log\frac{f_2(X_i)}{f_0(X_i)} = \log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}\left(X_i + n_i \log\frac{1-\pi_2}{1-\pi_0} \Big/ \log\frac{\pi_0(1-\pi_2)}{\pi_1(1-\pi_0)}\right) = \log\frac{\pi_0(1-\pi_2)}{\pi_2(1-\pi_0)}\left(X_i - n_i k_u^+\right) \tag{A5}$$

Since $\log\frac{\pi_0(1-\pi_1)}{\pi_1(1-\pi_0)}$ and $\log\frac{\pi_0(1-\pi_2)}{\pi_2(1-\pi_0)}$ are constants, it is equivalent to use $X_i - n_i k_u^-$ and $X_i - n_i k_u^+$ instead of $\log\frac{f_1(X_i)}{f_0(X_i)}$ and $\log\frac{f_2(X_i)}{f_0(X_i)}$ for the test, because the boundaries h^- and h^+ are optimized using $X_i - n_i k_u^-$ and $X_i - n_i k_u^+$. Then, (A1) and (A2) are simplified to (A6) and (A7), respectively.

$$C_i^- = \min\left(0, C_{i-1}^- + X_i - n_i k_u^-\right), \tag{A6}$$

$$C_i^+ = \max\left(0, C_{i-1}^+ + X_i - n_i k_u^+\right). \tag{A7}$$

The above confirms that k_u^- and k_u^+ are expressed by Equations (3) and (4), respectively. The constants h^- and h^+ are determined as the optimal boundaries for the normalized values of C_i^- and C_i^+.

APPENDIX B

To avoid complicated expressions in the CUSUMIN description, formulas (5) and (6) provide simple expressions. The strictly mathematical expression is as follows:

Let i_j be the index of the cohorts treated at dose level j. Thus, if k cohorts have been treated, then $k = \sum_{j=1}^{J} i_j$. Let $m_j(i_j)$ be the number of patients with DLT in the i_j-th cohort $(i_j = 0, ..., I_j)$ in dose level j, where $m_1(0) = \cdots = m_J(0) = 0$. We expand the CUSUM statistics C_i^- and C_i^+ to $C_j^-(i_j)$ and $C_j^+(i_j)$ for the i_j-th cohort at dose level j,

$$C_j^-(i_j) = \min\left(0, C_j^-(i_j - 1) + m_j(i_j) - n_j(i_j) k_u^-\right),$$

$$C_j^+(i_j) = \max\left(0, C_j^+(i_j - 1) + m_j(i_j) - n_j(i_j) k_u^+\right).$$

Here, $C_j^-(0) = C_j^+(0) = 0$, $C_j^-(i_j - 1)$ and $C_j^+(i_j - 1)$ are the CUSUM statistics for $(i_j - 1)$-th cohort at dose level j, $n_j(i_j)$ is the number of patients in the i-th cohort at dose level j, and $m_j(i_j)$ is the number of patients who experienced DLT in i-th cohort at dose level j. CUSUM statistics for i-th cohort at dose level j are updated based on the CUSUM statistics of the previous cohort $C_j^-(i_j - 1)$ and $C_j^+(i_j - 1)$, and the newly observed outcomes of $n_j(i_j)$ and $m_j(i_j)$.