Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

STOPPING RULES FOR SAMPLING IN PRECISION MEDICINE

Mingrui Ding[1,2], Siyang Gao[1], and Qiuhong Zhao[2]

[1]Dept. of Systems Eng., City University of Hong Kong, Hong Kong, CHINA
[2]School of Economics and Management, Beihang University, Beijing, CHINA

ABSTRACT

Precision medicine (PM) is an approach that aims to tailor treatments based on patient profiles (patients' biometric characteristics). In PM practice, treatment performance is typically evaluated through simulation models or clinical trials. Although these two methods have differences in sampling subjects and requirements, both are based on a sequential sampling process and require determining a stopping time for sampling to ensure that, with a prespecified confidence level, the best treatment is correctly identified for each patient profile. In this research, we propose unified stopping rules applicable to both simulation and clinical trial-based PM sampling processes. Specifically, we adapt the generalized likelihood ratio (GLR) test to determine when samples collected are sufficient and calibrate it using mixture martingales with a peeling method. Our stopping rules are theoretically grounded and can be integrated with different types of sampling strategies. Numerical experiments on synthetic problems and a case study demonstrate their effectiveness.

1 INTRODUCTION

Precision medicine is an emerging research area in healthcare. Conventionally, most medical treatments are designed for the "average patient" as a one-size-fits-all approach. However, since treatment outcomes usually depend heavily on individual patient profiles (patients' biometric characteristics such as age, gender, blood pressure, heart rate, BMI index, genetic information, etc.), this generalized approach may lead to significant variability in treatment effectiveness, yielding high efficacy in some patients but low or negligible effects in others. In contrast, PM seeks to tailor treatments to individual patient profiles. For each profile, PM aims to identify the most effective treatment from a predefined set of treatment alternatives. By doing so, PM can significantly increase the likelihood of achieving best treatment outcomes for all patients. Figure 1 provides a visual representation of this concept, where each color represents a group of patients with the same profiles.

To identify the best treatment for each patient profile, PM must first evaluate the performance of different treatment-profile combinations. This evaluation is typically conducted using one of two primary methods: simulation models or clinical trials. In simulation model-based methods, sampling subjects are simulation models of the diseases, and treatment performance is assessed through multiple simulation runs. Each time, the experimenter needs to specify both the patient profile and the treatment for testing in order to implement the simulation run. In clinical trials, the sampling subjects are patient volunteers, and treatment performance is evaluated using data collected from these trials. Since the profiles of the patients participating in the trials are beyond the experimenter's control, each time, the experimenter only needs to select a treatment for testing based on the given patient profile.

Despite their differences in sampling subjects and requirements, both the simulation model and clinical trial-based methods can be viewed as a sequential sampling process aimed at learning the best treatment for each patient profile from noisy samples. For this sampling process, on the one hand, we hope to use as few samples as possible to reduce resource consumption, as both simulation and clinical trial samples are costly. On the other hand, sufficient information must still be gathered to make high-quality treatment decisions. This trade-off raises a fundamental question for both methods: how early can we terminate the

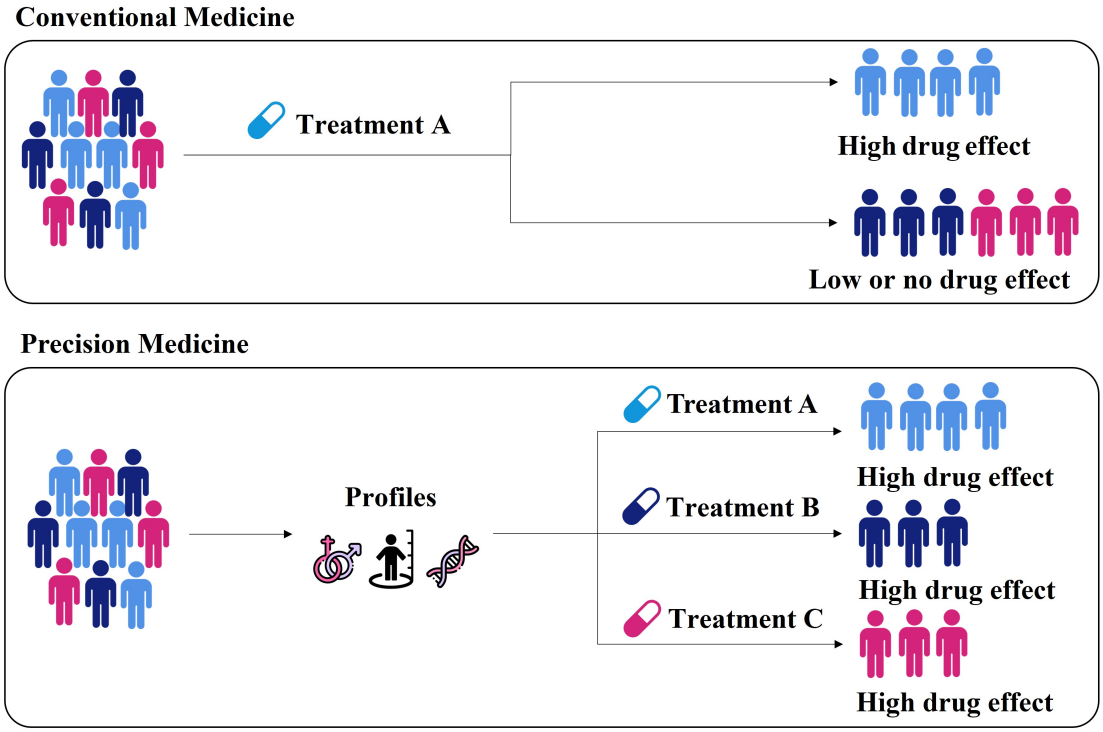

Figure 1: Illustration of precision medicine.

sampling process while still ensuring with a certain level of confidence the correct selection of the best treatment for every patient profile?

Similar problems have been considered in the literature. In the simulation environment, they are typically studied based on the contextual ranking and selection (CR&S) model and addressed using the indifference-zone (IZ) method. Shen et al. (2021) developed two-stage IZ procedures for treatment selection based on two probability of correct selection (PCS)-related measures: the expected PCS and the minimum PCS across all contexts (profiles). Keslin et al. (2022) decomposed the CR&S problem into several individual R&S problems, which could be solved using standard IZ methods such as the KN procedure (Kim and Nelson 2001). These IZ methods can provide quality guarantees for identifying the best treatment under each profile. However, their stopping rules must be integrated with certain types of two-stage or elimination-based sampling strategies. These sampling strategies have not been optimized to improve pairwise comparison efficiency, which often results in conservative stopping conditions that require an excessive number of samples before termination. This conservativeness of IZ-based methods has been noted in the literature, as observed in Branke et al. (2007).

In the clinical trial environment, related research is conducted within the contextual best-arm identification (BAI) model. Simchi-Levi et al. (2024) developed a stopping rule for the experimental design problem with randomly arriving contexts. Kato and Ariu (2024a) designed a stopping rule for identifying the treatment with the largest mean marginalized over the contextual distribution. Despite differences in problem settings, these stopping rules are designed under the assumption that the samples of the treatments follow a one-parameter exponential distribution family, typically normal distributions with known variances. However, the sampling variances of the treatments are typically unknown in practice. As a result, when such stopping rules are applied in real-world environments, they tend to underestimate the uncertainty of the samples and terminate the sampling process too early, leading to the selection of treatments that fail to meet the desired quality levels.

In this research, we develop stopping rules for the PM sampling process that are applicable in both the simulation and clinical trial environments. These stopping rules aim to identify the best treatment under each profile with a certain level of quality guarantee. To do it, we first employ two PCS-based quality measures that can reasonably summarize the PCS (i.e., our confidence) over the entire profile space, and seek to make sure that they reach the pre-specified level upon termination of the sampling process. We then utilize the generalized likelihood ratio (GLR) test to determine when the desired confidence level has been achieved. For each of the two quality measures, we calibrate the corresponding stopping thresholds. Once all the GLR statistics exceed the respective thresholds, the stopping rule is triggered and the sampling process terminates. To calibrate these thresholds, we derive new time-uniform concentration inequalities for both the sample mean and sample variance using mixture martingales and a peeling method. These concentrations allow us to quantify the deviation of estimators across all time stages, thus calibrating our stopping rules.

The development of the stopping rules in this research is nontrivial due to the need to incorporate sample variances into the GLR statistics. Unlike cases where variances are known, using sample variances makes the GLR statistics significantly less tractable. The sample variances generate additional uncertainty on the statistics and complicate their distributional properties over time. Calibrating the stopping rules requires confidence regions on both the sample mean and sample variance.

Another notable feature of our stopping rules is that they can be integrated to any static or adaptive PM sampling strategy, including those that have been extensively studied in the literature. In the simulation environment, Du et al. (2024) derive asymptotic optimality conditions for sampling strategies and propose a sampling algorithm within the Optimal Computing Budget Allocation (OCBA) framework. Li et al. (2020) develop a dynamic sampling policy under an approximate dynamic programming framework, while Shi et al. (2023) introduce a Top-Two Thompson Sampling policy. In the clinical trial environment, Alban et al. (2021) propose a sequential sampling algorithm, fEVI, based on the Expected Value of Information method and Kato et al. (2024b) propose an adaptive sampling strategy called PLAS. These sampling strategies are often supported by theoretical results about their superior properties, such as consistency and asymptotic optimality. When combined with these efficient sampling processes, our stopping rules have the potential to terminate earlier, saving samples while still maintaining the statistical guarantees. This is in contrast to IZ-type procedures developed in simulation environments (Shen et al. 2021; Keslin et al. 2022), which rely on less efficient elimination-based sampling. More importantly, our stopping rules remain flexible and adaptable to potential future developments of even more efficient sampling strategies (e.g., optimal strategies instead of asymptotic optimal ones), to allow for further reductions in sample usage.

The rest of the paper is organized as follows. In Section 2, we formulate the PM problem as a sampling process with stopping times. Section 3 defines the stopping rules for the two PCS-based measures using GLR statistics and calibrated thresholds. Section 4 presents numerical experiments, including synthetic problem instances and a case study. Section 5 concludes the paper.

2 PROBLEM FORMULATION

Suppose there are k treatments alternatives $\{1,2,...,k\}$ for the disease and m patient profiles $\{\boldsymbol{x}_1,\boldsymbol{x}_2,...,\boldsymbol{x}_m\}$ in total. A possible profile is denoted as the random vector $\boldsymbol{X} = (X_1,...,X_d)^{\mathrm{T}}$, where $\boldsymbol{X}$ has support $\mathscr{X} \in \mathbb{R}^d$ and each component X_i corresponds to a specific biometric characteristic under consideration. Let $y_i(\boldsymbol{x}_j)$ represent the mean performance of treatment i under profile $\boldsymbol{x}_j$, which is unknown and can be estimated by noisy samples from simulation or clinical trials, $i = 1,...,k,\ j = 1,...,m$. These samples are assumed to follow the normal distribution $N(y_i(\boldsymbol{x}_j),\sigma_i^2(\boldsymbol{x}_j))$, where the variances $\sigma_i^2(\boldsymbol{x}_j)$ are also unknown. The assumption that treatment performance follows a normal distribution is common in the clinical trials literature (Jiménez-Luna et al. 2020; Forster et al. 2021). The goal of PM is to identify the best treatment $i^*(\boldsymbol{x}_j) \in \arg\max_{i\in\{1,...,k\}} y_i(\boldsymbol{x}_j)$ for each patient profile $\boldsymbol{x}_j$. We further assume that $i^*(\boldsymbol{x}_j)$ is unique for each $\boldsymbol{x}_j$, which is also a common assumption in clinical trials (Robertson et al. 2023).

For a sampling process based on either simulation models or clinical trials, suppose t represents the sampling time and the sampling process ends at time $t = \tau$. We use the sample mean $\bar{y}_{t,i}(\boldsymbol{x}_j)$ to estimate the performance of treatment i for profile $\boldsymbol{x}_j$ at time t. When the sampling process ends, we select $\hat{i}_\tau(\boldsymbol{x}_j) = \arg\max_{i\in\{1,\ldots,k\}} \bar{y}_{\tau,i}(\boldsymbol{x}_j)$ as the estimated best treatment for profile $\boldsymbol{x}_j$. In practice, the total number of samples that can be used is finite, making it impossible to ensure that the estimated best treatment $\hat{i}_\tau(\boldsymbol{x}_j)$ always equals to the true best treatment $i^*(\boldsymbol{x}_j)$. Then, a common quality measure for the selected treatment is the probability of correct selection (PCS) for the best treatment. Given profile $\boldsymbol{x}_j$ and ending time τ, PCS is expressed as:

$$\mathrm{PCS}(\boldsymbol{x}_j) = \mathbb{P}\left(\hat{i}_\tau(\boldsymbol{x}_j) = i^*(\boldsymbol{x}_j)\right) = \mathbb{P}\left(\cap_{i=1,i\neq i^*(\boldsymbol{x}_j)}^{k}\left(\bar{y}_{\tau,i^*(\boldsymbol{x}_j)}(\boldsymbol{x}_j) > \bar{y}_{\tau,i}(\boldsymbol{x}_j)\right)\right),$$

where the probability is taken with respect to the randomness of the samples.

Ultimately, we aim for the PCS to not be limited to a specific profile $\boldsymbol{x}_j$ but instead to be summarized across all possible profiles. To achieve this, we consider two PCS-based measures, PCS_E and PCS_A, defined as follows:

$$\mathrm{PCS}_\mathrm{E} = \mathbb{E}(\mathrm{PCS}(\boldsymbol{X})) = \sum_{j=1}^{m} p_j \mathrm{PCS}(\boldsymbol{x}_j),$$
$$\mathrm{PCS}_\mathrm{A} = \mathbb{P}\left(\forall \boldsymbol{x} \in \mathscr{X}, \hat{i}_\tau(\boldsymbol{x}) = i^*(\boldsymbol{x})\right).$$

These two measures have been commonly used in the literature (Gao et al. 2019; Shi et al. 2023; Simchi-Levi et al. 2024). The measure PCS_E is the expectation of the PCS over the profile space, where p_j denotes the probability of profile $\boldsymbol{X} = \boldsymbol{x}_j$. It provides an average assessment of the PCS across all patient profiles. PCS_A corresponds to the probability that the best treatment is correctly selected simultaneously for all profiles. This is a natural measure to use because the best treatment being correctly selected for each profile directly describes the event we are interested in. It is obvious that $\mathrm{PCS}_\mathrm{E} \geq \mathrm{PCS}_\mathrm{A}$, with PCS_A setting a higher standard on the quality of the correct selection over the profile space.

Our goal is to design stopping rules for the sampling process that guarantee a pre-specified level of PCS_E or PCS_A. Suppose at time t, the decision maker chooses to sample treatment i_t under profile $\boldsymbol{x}_{j_t}$ according to some sampling strategy. Note that in the case of simulation models, $\boldsymbol{x}_{j_t}$ is part of the sampling decision made by the experimenter, while in clinical trials, $\boldsymbol{x}_{j_t}$ represents the profile of the volunteer who participates in the trial at that time. The observed outcome is denoted as $Y_{t,i_t}(\boldsymbol{x}_{j_t})$. We define the filtration $\mathscr{F}_t$ as the σ-algebra generated by the sequence $(Y_{1,i_1}(\boldsymbol{x}_{j_1}), Y_{2,i_2}(\boldsymbol{x}_{j_2}), \ldots, Y_{t,i_t}(\boldsymbol{x}_{j_t}))$, which contains all the information gathered up to time t. The stopping time τ will be a random variable adapted to the filtration $(\mathscr{F}_t)_{t\in\mathbb{N}}$, i.e., the decision to terminate sampling is based solely on the information accumulated up to that time point.

Let τ_α^E denote the stopping time that guarantees PCS_E to meet the confidence level $1-\alpha$, i.e., $\mathrm{PCS}_\mathrm{E} \geq 1-\alpha$. In other words, the selection error should be controlled such that

$$\sum_{j=1}^{m} p_j \mathbb{P}\left(\hat{i}_{\tau_\alpha^\mathrm{E}}(\boldsymbol{x}_j) \neq i^*(\boldsymbol{x}_j)\right) \leq \alpha. \tag{1}$$

Similarly, let τ_α^A denote the stopping time that guarantees PCS_A to meet the confidence level $1-\alpha$, i.e., $\mathrm{PCS}_\mathrm{A} \geq 1-\alpha$. It means that the selection error should be controlled such that

$$\mathbb{P}\left(\exists \boldsymbol{x} \in \mathscr{X}, \hat{i}_{\tau_\alpha^\mathrm{A}}(\boldsymbol{x}) \neq i^*(\boldsymbol{x})\right) \leq \alpha. \tag{2}$$

Note that the stopping times τ_α^E and τ_α^A depend on the sampling strategy, and in this research, we only consider sampling strategies for τ_α^E and τ_α^A to satisfy that $\mathbb{P}(\tau_\alpha^\mathrm{E} < +\infty) = 1$ and $\mathbb{P}(\tau_\alpha^\mathrm{A} < +\infty) = 1$ respectively.

3 STOPPING RULES

In this section, we develop stopping rules for the PM sampling process based on PCS_E and PCS_A. In Section 3.1, we construct the generalized likelihood ratio (GLR) statistic and propose a general stopping criterion. Section 3.2 further calibrates the stopping thresholds and proves the statistical guarantees.

3.1 The GLR Statistic

The generalized likelihood ratio (GLR) test method was originally introduced in Chernoff (1959). We extend it to the PM sampling process. Let $N_{t,i}(\boldsymbol{x}) = \sum_{s=1}^{t} \mathbf{1}\{i_s = i, \boldsymbol{x}_{j_s} = \boldsymbol{x}\}$ represent the total number of samples allocated to treatment i and profile $\boldsymbol{x}$ up to time t. The sample mean $\bar{y}_{t,i}(\boldsymbol{x}) = \frac{1}{N_{t,i}(\boldsymbol{x})} \sum_{s=1}^{t} \mathbf{1}\{i_s = i, \boldsymbol{x}_{j_s} = \boldsymbol{x}\} Y_{t,i}(\boldsymbol{x})$ and the sample variance $S^2_{t,i}(\boldsymbol{x}) = \frac{1}{N_{t,i}(\boldsymbol{x})-1} \sum_{s=1}^{t} \mathbf{1}\{i_s = i, \boldsymbol{x}_{j_s} = \boldsymbol{x}\} (Y_{t,i}(\boldsymbol{x}) - \bar{y}_{t,i}(\boldsymbol{x}))^2$ for treatment i under profile $\boldsymbol{x}$ at time t, and the estimated best treatment is given by $\hat{i}_t(\boldsymbol{x}) = \arg\max_{i \in \{1,\ldots,k\}} \bar{y}_{t,i}(\boldsymbol{x})$. For any treatment $i \neq \hat{i}_t(\boldsymbol{x})$, the GLR statistic $Z_i(\boldsymbol{x},t)$ quantifies the evidence against i being the best treatment under profile $\boldsymbol{x}$ at time t. It is defined as the log-likelihood ratio comparing $y_i(\boldsymbol{x})$ and $y_{\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})$:

$$Z_i(\boldsymbol{x},t) = \log \frac{\max_{\theta_{\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}) \geq \theta_i(\boldsymbol{x})} \mathbb{P}_{\theta_{\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}(\underline{Y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})) \mathbb{P}_{\theta_i(\boldsymbol{x})}(\underline{Y}_{t,i}(\boldsymbol{x}))}{\max_{\theta_{\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}) \leq \theta_i(\boldsymbol{x})} \mathbb{P}_{\theta_{\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}(\underline{Y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})) \mathbb{P}_{\theta_i(\boldsymbol{x})}(\underline{Y}_{t,i}(\boldsymbol{x}))},$$

where $\underline{Y}_{t,i}(\boldsymbol{x}) = (Y_{t,i}(\boldsymbol{x}) : i_l = i, l \leq t)$ denotes the vector of all observed samples for treatment i under profile $\boldsymbol{x}$ up to time t, and $\mathbb{P}_{\theta_i(\boldsymbol{x}) \sim \mathcal{N}(y_{t,i}(\boldsymbol{x}), \sigma_i^2(\boldsymbol{x}))}(\underline{Y}_{t,i}(\boldsymbol{x}))$ is the likelihood of $N_{t,i}(\boldsymbol{x})$ i.i.d. samples. For any $i \in \{1,\ldots,k\}$ and $\boldsymbol{x} \in \mathcal{X}$, the maximum likelihood estimate of its unknown mean is its sample mean $\bar{y}_{t,i}(\boldsymbol{x})$. It can be calculated, under normal sampling distributions,

$$\begin{aligned} Z_i(\boldsymbol{x},t) &= \inf_{u \in [\bar{y}_{t,i}(\boldsymbol{x}), \bar{y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})]} \left[\log \frac{\mathbb{P}_{\bar{y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}(\underline{Y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}))}{\mathbb{P}_u(\underline{Y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}))} + \log \frac{\mathbb{P}_{\bar{y}_{t,i}(\boldsymbol{x})}(\underline{Y}_{t,i}(\boldsymbol{x}))}{\mathbb{P}_u(\underline{Y}_{t,i}(\boldsymbol{x}))} \right] \\ &= \inf_{u \in [\bar{y}_{t,i}(\boldsymbol{x}), \bar{y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})]} \sum_{c \in \{i, \hat{i}_t(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) d(\bar{y}_{t,c}(\boldsymbol{x}), u), \end{aligned}$$

where $d(x,y) = \frac{(x-y)^2}{2\sigma^2}$ is defined as the KL-divergence between two normal distributions $\mathcal{N}(x,\sigma^2)$ and $\mathcal{N}(y,\sigma^2)$. We plug sample variances $S^2_{t,i}(\boldsymbol{x})$ to replace the unknown variances $\sigma_i^2(\boldsymbol{x})$ for each $i \neq \hat{i}_t(\boldsymbol{x})$, which yields

$$\begin{aligned} Z_i^s(\boldsymbol{x},t) &= \inf_{u \in [\bar{y}_{t,i}(\boldsymbol{x}), \bar{y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})]} \sum_{c \in \{i, \hat{i}_t(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) \frac{(\bar{y}_{t,c}(\boldsymbol{x}) - u)^2}{2S^2_{t,c}(\boldsymbol{x})} \\ &= \frac{1}{2} \frac{(\bar{y}_{t,i}(\boldsymbol{x}) - \bar{y}_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}))^2}{\frac{S^2_{t,i}(\boldsymbol{x})}{N_{t,i}(\boldsymbol{x})} + \frac{S^2_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}{N_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}}. \end{aligned}$$

Intuitively, for a given profile $\boldsymbol{x}$ and time t, a larger statistic $Z_i^s(\boldsymbol{x},t)$ indicates that treatment i is less likely to be the best under this profile at that time. Therefore, it is possible to design thresholds such that when $Z_i^s(\boldsymbol{x},t)$ exceeds its corresponding threshold for all $i \neq \hat{i}_t(\boldsymbol{x})$, PCS_E and PCS_A become large enough. Typically, these thresholds depend on t and α and are chosen to be on the order of α. In the PM setting with multiple profiles, we define the threshold for treatment i as $\phi_i(\boldsymbol{N}_t, \alpha, \boldsymbol{x})$, a function of confidence level α, profile $\boldsymbol{x}$ and $\boldsymbol{N}_t$ (the vector containing all the sample numbers $N_{t,i}(\boldsymbol{x}_j)$).

When all statistics $Z_i^s(\boldsymbol{x},t)$ exceed their corresponding thresholds, the stopping rule is triggered and the sampling process stops. Given these statistics and thresholds, the general form of the stopping rule is given by

$$\tau_\alpha = \inf\{t \in \mathbb{N} | \forall \boldsymbol{x} \in \mathcal{X}, \forall i \neq \hat{i}_t(\boldsymbol{x}), Z_i^s(\boldsymbol{x},t) > \phi_i(\boldsymbol{N}_t, \alpha, \boldsymbol{x})\}. \tag{3}$$

For different measures PCS_E and PCS_A, their stopping times τ_α^E and τ_α^A are different, and so are their thresholds ϕ_i in (3). We let ϕ_i^E be the thresholds for τ_α^E and ϕ_i^A be the thresholds for τ_α^A. These thresholds will be calibrated in the next subsection using newly developed concentration inequalities.

3.2 The Calibrated Thresholds

We first consider PCS_E. The error probability for each profile needs to be controlled. Therefore, we propose to bound the term $\mathbb{P}\left(\tau_\alpha^\text{E} < +\infty, \hat{i}_{\tau_\alpha^\text{E}}(\boldsymbol{x}) \neq i^*(\boldsymbol{x})\right)$ for each possible profile $\boldsymbol{x}$. We first have

$$\mathbb{P}\left(\tau_\alpha^\text{E} < +\infty, \hat{i}_{\tau_\alpha^\text{E}}(\boldsymbol{x}) \neq i^*(\boldsymbol{x})\right) \leq \mathbb{P}\left(\exists t \in \mathbb{N}, \exists i \neq i^*(\boldsymbol{x}), i = \hat{i}_t(\boldsymbol{x}), Z^s_{i^*(\boldsymbol{x})}(\boldsymbol{x},t) > \phi^\text{E}_{i^*(\boldsymbol{x})}(\boldsymbol{N}_t, \alpha, \boldsymbol{x})\right).$$

Since the GLR statistic satisfies, for all $i \neq i^*(\boldsymbol{x})$ and $i = \hat{i}_t(\boldsymbol{x})$,

$$Z^s_{i^*(\boldsymbol{x})}(\boldsymbol{x},t) \leq \sum_{c \in \{i, i^*(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) \frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{2S^2_{t,c}(\boldsymbol{x})},$$

we further obtain the upper bound for $\mathbb{P}\left(\tau_\alpha^\text{E} < +\infty, \hat{i}_{\tau_\alpha^\text{E}}(\boldsymbol{x}) \neq i^*(\boldsymbol{x})\right)$ as

$$\mathbb{P}\left(\exists t \in \mathbb{N}, \exists i \neq i^*(\boldsymbol{x}), \sum_{c \in \{i, i^*(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) \frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{2S^2_{t,c}(\boldsymbol{x})} > \phi^\text{E}_{i^*(\boldsymbol{x})}(\boldsymbol{N}_t, \alpha, \boldsymbol{x})\right). \tag{4}$$

Similarly, for PCS_A and τ_α^A, we propose to control the term $\mathbb{P}\left(\tau_\alpha^\text{A} < +\infty, \exists \boldsymbol{x} \in \mathscr{X}, \hat{i}_{\tau_\alpha^\text{A}}(\boldsymbol{x}) \neq i^*(\boldsymbol{x})\right)$ to ensure (2). It can be upper bounded by

$$\mathbb{P}\left(\exists t \in \mathbb{N}, \exists \boldsymbol{x} \in \mathscr{X}, \exists i \neq i^*(\boldsymbol{x}), \sum_{c \in \{i, i^*(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) \frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{2S^2_{t,c}(\boldsymbol{x})} > \phi^\text{A}_{i^*(\boldsymbol{x})}(\boldsymbol{N}_t, \alpha, \boldsymbol{x})\right). \tag{5}$$

The event described in the probability measure in either (4) or (5) represents a false selection of identifying treatment $i \neq i^*(\boldsymbol{x})$ as the best treatment. The threshold ϕ_i must be set sufficiently large to control the probability of such false selections uniformly throughout the time. We propose to determine these thresholds by tracking the time-uniform concentration behavior of the summation term

$$\sum_{c \in \{i, i^*(\boldsymbol{x})\}} N_{t,c}(\boldsymbol{x}) \frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{2S^2_{t,c}(\boldsymbol{x})}. \tag{6}$$

This summation term can be treated to have captured both the deviation of sample mean, represented by $\frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{\sigma^2_c(\boldsymbol{x})}$, and the deviation of the sample variance, represented by $\frac{S^2_{t,c}(\boldsymbol{x})}{\sigma^2_c(\boldsymbol{x})}$, for $c \in \{i, i^*(\boldsymbol{x})\}$. Given a significance level α, let $P_c(\boldsymbol{x})$ be the maximum value for $\frac{(\bar{y}_{t,c}(\boldsymbol{x}) - y_c(\boldsymbol{x}))^2}{\sigma^2_c(\boldsymbol{x})}$ and $Q_c(\boldsymbol{x})$ be the minimum value for $\frac{S^2_{t,c}(\boldsymbol{x})}{\sigma^2_c(\boldsymbol{x})}$, for $c \in \{i, i^*(\boldsymbol{x})\}$. Then, the thresholds ϕ_i for profile $\boldsymbol{x}$ can be determined as

$$\sum_{c \in \{i, i^*(\boldsymbol{x})\}} \frac{N_{t,c}(\boldsymbol{x})}{2} \frac{P_c(\boldsymbol{x})}{Q_c(\boldsymbol{x})}.$$

We derive a time-uniform tail concentration inequality on the sample mean to determine $P_c(\boldsymbol{x})$ and a time-uniform lower tail concentration inequality on the sample variance to determine $Q_c(\boldsymbol{x})$. The derivations utilize the mixture martingales constructed in Kaufmann and Koolen (2021) and Howard et al. (2020), along with a peeling method used by Jourdan et al. (2023).

For the concentration on the sample mean, let $s > 1$ and ζ be the Riemann ζ function. Let $\hat{\mu}_t$ be the sample mean of t i.i.d. normal samples parameters (μ, σ^2). Then, for all $V > 1$, we have

$$\mathbb{P}\left(\exists t \in \mathbb{N}, \frac{|\hat{\mu}_t - \mu|}{\sigma} > \frac{V}{\sqrt{t}}\right) \leq \exp\left(s\log(2s+\log t) - f(s) + \frac{1}{2}\log(V^2) - \frac{1}{2}V^2 + \frac{1}{2}\right),$$

where $f(s) = s\log(2s) - s - \log(\zeta(s))$.

For the concentration on the sample variance, let $\eta > 0$ and $\hat{\sigma}^2_{t+1}$ be the sample variance of $t+1$ i.i.d. normal samples with parameters (μ, σ^2). Then, for all $V \in (0,1)$, we have

$$\mathbb{P}\left(\exists t \in \mathbb{N}, \frac{\hat{\sigma}^2_{t+1}}{\sigma^2} \leq V\right) \leq \left(1 + \frac{\log(t)}{\log(1+\eta)}\right)^s \zeta(s) e^{-\frac{t}{2(1+\eta)}(V - \log(V) - 1)}.$$

Combining these results, we can establish stopping rules of the PM sampling process for measures $\mathrm{PCS_E}$ and $\mathrm{PCS_A}$.

Theorem 1 For all $t \in \mathbb{N}^+$, let $\boldsymbol{N}_{t,*} = (N_{t,\hat{i}_t(\boldsymbol{x}_1)}(\boldsymbol{x}_1), ..., N_{t,\hat{i}_t(\boldsymbol{x}_m)}(\boldsymbol{x}_m))^{\mathrm{T}}$ and $\boldsymbol{N}_{t,i} = (N_{t,i}(\boldsymbol{x}_1), ..., N_{t,i}(\boldsymbol{x}_m))^{\mathrm{T}}$ for all $i = 1, ..., k$. Define $g(x) = x - \log(x)$ for $x > 0$ and $h(s) = 1 + 2[\log(\zeta(s)) + s - s\log(2s)]$ for $s > 1$. Let $N(\boldsymbol{x})$ be a function that maps the profile space $\mathscr{X}$ to $\mathbb{N}^+$, with the vector $\boldsymbol{N} = (N(\boldsymbol{x}_1), ..., N(\boldsymbol{x}_m))^T$.

Let $\eta > 0, s > 1$ and ζ be the Riemann ζ function. For any $\boldsymbol{N} \in \mathbb{N}^{m\times 1}$ and $\alpha \in (0,1)$, define $\gamma^{\mathrm{E}}_{\mu}(\boldsymbol{N}, \alpha) > 0$ and $\gamma^{\mathrm{E}}_{\sigma}(\boldsymbol{N}, \alpha) > 0$ such that

$$g(\gamma^{\mathrm{E}}_{\mu}(\boldsymbol{N}, \alpha)) = h(s) + 2\log\left(\frac{4k}{\alpha}\right) + 2\log(\mathbb{E}[(2s + \log(N(\boldsymbol{X})))^s]),$$

$$\mathbb{E}\left[e^{-\frac{N(\boldsymbol{X})}{2(1+\eta)}(g(\gamma^{\mathrm{E}}_{\sigma}(\boldsymbol{N},\alpha)) - 1)}\left(1 + \frac{\log(N(\boldsymbol{X}))}{\log(1+\eta)}\right)^s\right] = \frac{\alpha}{4k\zeta(s)},$$

where the expectation is taken with respect to the distribution of $\boldsymbol{X}$. The thresholds

$$\phi^{\mathrm{E}}_i(\boldsymbol{N}_t, \alpha, \boldsymbol{x}) = \frac{\gamma^{\mathrm{E}}_{\mu}(\boldsymbol{N}_{t,i}, \alpha) N_{t,i}(\boldsymbol{x})}{2\gamma^{\mathrm{E}}_{\sigma}(\boldsymbol{N}_{t,i}, \alpha)\left(N_{t,i}(\boldsymbol{x}) - 1\right)} + \frac{\gamma^{\mathrm{E}}_{\mu}(\boldsymbol{N}_{t,*}, \alpha) N_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x})}{2\gamma^{\mathrm{E}}_{\sigma}(\boldsymbol{N}_{t,*}, \alpha)\left(N_{t,\hat{i}_t(\boldsymbol{x})}(\boldsymbol{x}) - 1\right)},$$

combined the stopping rule (3), ensure that $\mathrm{PCS_E} \geq 1 - \alpha$.

Theorem 2 Let $\eta > 0, s > 1$ and ζ be the Riemann ζ function. For any $t \in \mathbb{N}^+$ and $\alpha \in (0,1)$, define $\gamma^{\mathrm{A}}_{\mu}(t, \alpha) > 0$ and $\gamma^{\mathrm{A}}_{\sigma}(t, \alpha) > 0$ such that

$$g(\gamma^{\mathrm{A}}_{\mu}(t, \alpha)) = h(s) + 2\log\left(\frac{4m(k-1)}{\alpha}\right) + 2s\log(2s + \log(t)),$$

$$g(\gamma^{\mathrm{A}}_{\sigma}(t, \alpha)) = 1 + \frac{2(1+\eta)}{t}\left[s\log(1 + \frac{\log(t)}{\log(1+\eta)}) + \log(\zeta(s)) + \log\left(\frac{4m(k-1)}{\alpha}\right)\right].$$

The thresholds

$$\phi^{\mathrm{A}}_i(\boldsymbol{N}_t, \alpha, \boldsymbol{x}) = \sum_{c \in \{i, \hat{i}_t(\boldsymbol{x})\}} \frac{\gamma^{\mathrm{A}}_{\mu}(N_{t,c}(\boldsymbol{x}), \alpha) N_{t,c}(\boldsymbol{x})}{2\gamma^{\mathrm{A}}_{\sigma}(N_{t,c}(\boldsymbol{x}), \alpha)\left(N_{t,c}(\boldsymbol{x}) - 1\right)},$$

combined the stopping rule (3), ensure that $\mathrm{PCS_A} \geq 1 - \alpha$.

4 NUMERICAL EXPERIMENTS

We conduct numerical experiments to evaluate the performance of the proposed stopping rules τ_α^{E} for $\mathrm{PCS}_{\mathrm{E}}$ and τ_α^{A} for $\mathrm{PCS}_{\mathrm{A}}$. We will test them on two synthetic problems in Section 4.1 and a case study in Section 4.2.

Since our stopping rules need to be combined with a sampling strategy, in the tests, we choose the Contextual Optimal Computing Budget Allocation (C-OCBA) Algorithm proposed in Gao et al. (2019). C-OCBA is an adaptive sampling strategy that aims to select the best treatment for each profile and can be shown to be asymptotically optimal. For the synthetic problems, we also combine our stopping rules with the equal allocation (EA) strategy, where the samples are simply equally allocated to each pair of treatment and profile.

For comparison, we also evaluate the performance of two IZ-type procedures, which can also offer statistical guarantees for identifying the best treatment under each profile. These procedures differ in their mechanisms and the types of guarantees they provide:

- *Two-Stage Indifference-Zone Procedure (TS-IZ).* Shen et al. (2021) introduced a two-stage procedure for selecting the best treatment for each profile. This procedure has two versions, TS and TS+, where TS assumes homogeneous variances across treatment-profile pairs, while TS+ is based on heterogeneous variances. For both TS and TS+, in the first stage, each treatment-profile pair is allocated with a small number of samples to estimate sample variances. The second stage then determines the required number of additional samples for each treatment-profile pair based on these variance estimates and selects the best treatment for each profile based on their sample means. This procedure provides guarantees for $\mathrm{PCS}_{\mathrm{E}}$ and a new measure $\min_{\boldsymbol{x}\in\mathscr{X}} \mathrm{PCS}(\boldsymbol{x})$. In the tests, the performance of TS-IZ is evaluated based on $\mathrm{PCS}_{\mathrm{E}}$.
- *KN Indifference-Zone Procedure (KN-IZ).* Keslin et al. (2022) proposed a method for the CR&S problem by decomposing it into a number of independent R&S subproblems, each corresponding to a specific context. The KN procedure (Kim and Nelson 2001) is then used to select the best design for each context. This method is applicable to the PM problem, where the KN procedure allocates samples to independently select the best treatment (design) while ensuring a PCS guarantee for the selected treatment within each profile (context). These individual PCS guarantees jointly contribute to an overall guarantee across the profile space. The performance of KN-IZ is evaluated under both $\mathrm{PCS}_{\mathrm{E}}$ and $\mathrm{PCS}_{\mathrm{A}}$.

4.1 Synthetic Problems

We compare TS-IZ, KN-IZ and our stopping rules on the following two synthetic problems:

- *Problem 1.* Consider a problem with $k=5$ treatments and $m=10$ profiles. For each profile $\boldsymbol{x}_j$ $(j=1,\ldots,10)$, the mean performance of treatment i $(i=1,\ldots,5)$ is given by $y_i(\boldsymbol{x}_j)=i(1+0.05(j-1))$. The variance of the samples for treatment-profile pair $(i,\boldsymbol{x}_j)$ is given by $\sigma_i^2(\boldsymbol{x}_j)=1+0.3i+0.3j$. The indifference-zone (IZ) parameter is set as $\delta=1$.
- *Problem 2.* Consider a problem with $k=10$ treatments and $m=10$ profiles, where the means of the treatments are derived from the Branin function, a widely used benchmark function. Let $z_i=1.5(i-1)-5$ for $i=1,...,10$ be 10 treatments and $x_j=j-1$ for $j=1,...,10$ be 10 profiles. For each profile x_j $(j=1,\ldots,10)$, the mean performance of treatment i $(i=1,\ldots,10)$ is given by $y_i(x_j)=-(x_j-bz_i^2+cz_i-r)^2-10(1-v)cos(z_i)-10$, where the coefficients are set as $b=\frac{5.1}{4\pi^2}$, $c=\frac{5}{\pi}$, $r=6$, and $v=\frac{1}{8\pi}$. The variance of the samples for treatment-profile pair (i,x_j) is given by $\sigma_i^2(x_j)=10-0.05ij$. The probability of each profile $p_j\in\{0.03,0.07,0.2,0.1,0.15,0.2,0.02,0.08,0.1,0.05\}$. The IZ parameter is set as $\delta=1$.

For the C-OCBA and EA sampling strategies, we initialize the number of sampling replications as $n_0 = 5$, while for the two IZ-based methods, we set $n_0 = 20$. The confidence level is set at $1-\alpha = 95\%$. The results are presented in Tables 1 and 2. From the two tables, we observe that our stopping rules consistently achieve the prespecified confidence level, regardless of the sampling strategy they are combined with. When combined with C-OCBA, our stopping rules exhibit less conservativeness compared with the two IZ-based procedures, in the sense that their PCS_E and PCS_A tend to be lower.

Table 1: Performance comparisons on problem 1.

τ_α^E			τ_α^A		
Method	**PCS_E**	**Avg. SSize ± 95% CI**	**Method**	**PCS_A**	**Avg. SSize ± 95% CI**
C-OCBA	0.99	3,903 ± 160	C-OCBA	0.99	12,794 ± 92
KN-IZ	0.99	10,830 ± 47	KN-IZ	1	20,040 ± 87
TS-IZ	0.99	12,837 ± 43	EA	1	51,269 ± 676
EA	1	44,610 ± 622			

Table 2: Performance comparisons on problem 2.

τ_α^E			τ_α^A		
Method	**PCS_E**	**Avg. SSize ± 95% CI**	**Method**	**PCS_A**	**Avg. SSize ± 95% CI**
C-OCBA	0.99	4,409 ± 928	C-OCBA	0.99	16,605 ± 2,645
KN-IZ	1	142,196 ± 407	KN-IZ	1	245,637 ± 704
TS-IZ	1	180,389 ± 382	EA	0.99	343,013 ± 6,713
EA	0.99	292,553 ± 6,254			

Regarding the samples used, when combined with C-OCBA, our stopping rules require fewer samples than the two IZ-based procedures. This efficiency gain arises because the computation of our stopping rules is independent of the sampling strategy, which allows us to fully utilize the efficient sample allocation (if any) of the sampling strategy to compare different treatments more effectively.

This intuition is illustrated in Figure 2, which shows the sample allocations on 5 treatments in profile 1 in Problem 1. In this problem, the best and second-best treatments under profile 1 are respective treatments $i = 1, 2$ respectively. As can be observed, when the sampling process ends, C-OCBA primarily allocates samples to these two treatments, which is an efficient sample allocation to distinguish the treatments. In contrast, both KN-IZ and TS-IZ allocate a large number of samples to the other three inferior treatments. This allocation is less efficient since these treatments can be confidently identified as non-best with few samples. Specifically, TS-IZ is significantly less efficient as it utilizes only sample variance information without leveraging sample mean. Due to the effective allocation strategy of C-OCBA, our stopping rules can terminate early and save samples. If combined with even more efficient sampling strategies, our method could further reduce sample usage.

Additionally, in Problem 2, C-OCBA requires significantly fewer samples than the two IZ-based procedures. This is because, in Problem 2, the minimum performance differences among treatments vary across profiles, and the IZ parameter is set based on the smallest differences. As a result, in profiles with larger performance gaps, the IZ-based procedures will consume a large number of samples.

4.2 PM for Esophageal Cancer Prevention

Esophageal cancer is a highly aggressive malignancy with significant morbidity and mortality around the world. Early intervention strategies, guided by precision medicine (PM), have the potential to improve patient outcomes by tailoring prevention and treatment plans based on individual risk profiles. In this case

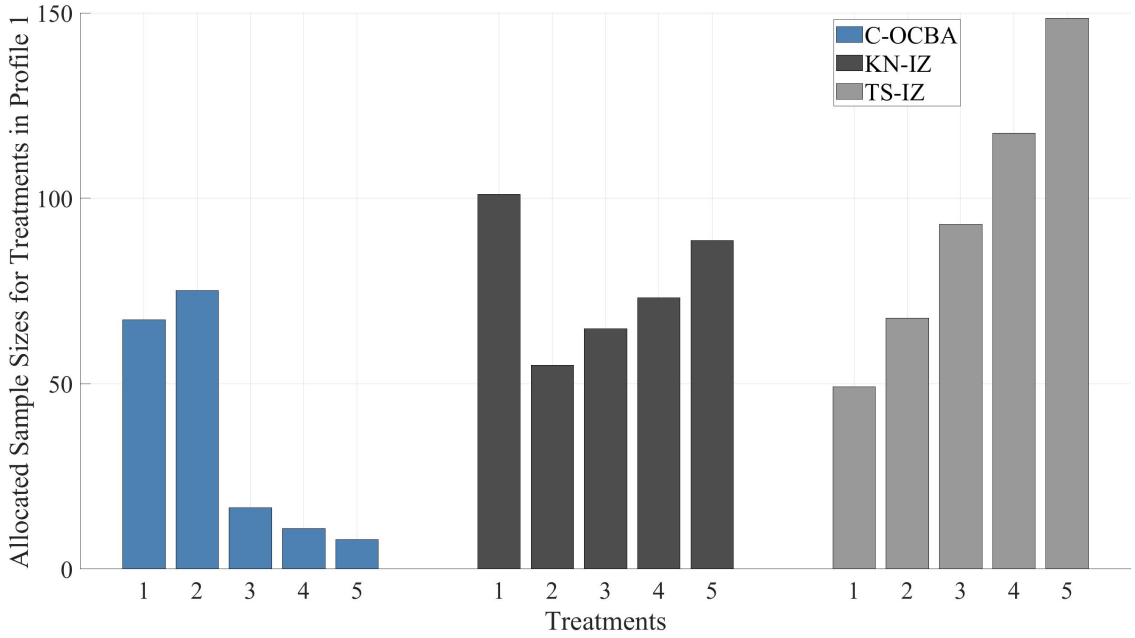

Figure 2: Sample allocations on 5 treatments in profile 1 in Problem 1.

study, we apply our proposed stopping rules to a simulation-assisted PM framework for esophageal cancer prevention. This case has also been studied in Shen et al. (2021) and Li et al. (2024). We test it in the simulation environment. The Markov Simulation Model for this disease is firstly proposed by (Hur et al. 2004) and can also be seen in Shen et al. (2021).

We consider four treatments and eight patient profiles. This setting is consistent with medical applications (Alban et al. 2021). Our method is evaluated against TS-IZ and KN-IZ under PCS_E, and against KN-IZ under PCS_A. The four treatment options include: (1) endoscopic surveillance only, (2) aspirin chemoprevention with endoscopic surveillance, (3) statin chemoprevention with endoscopic surveillance, and (4) combined aspirin and statin chemoprevention with endoscopic surveillance. Each treatment differs in both drug efficacy and associated complications, including their types and annual incidence rates. The effects of aspirin, statin and their combination are set at 0.54, 0.53, and 0.78, respectively, according to Kastelein et al. (2011). The complications' coefficients are set the same as Shen et al. (2021).

Patient profiles are defined by age groups ([55, 60, 65, 70]) and risk levels ([2.5%, 5%]), which represents the annual progression rate of Barrett's Esophagus to Esophageal Adenocarcinoma. The IZ parameter is set as $\delta = 0.1$. We conduct experiments with 250 replications to compute PCS_E and PCS_A. The other parameter settings for the procedures remain the same as those used in the synthetic problems. The results are summarized in Table 3.

Table 3: Performance comparison of stopping rules and IZ methods of case study.

τ_α^E			τ_α^A		
Method	**PCS_E**	**Avg. SSize ± 95% CI**	**Method**	**PCS_A**	**Avg. SSize ± 95% CI**
C-OCBA	0.99	271,163 ± 9,988	C-OCBA	0.99	340,644 ± 11,257
KN-IZ	0.98	581,597 ± 18,174	KN-IZ	0.96	1,054,615 ± 32,976
TS-IZ	0.98	632,948 ± 18,874			

From Table 3, we observe that our stopping rules achieve the predefined confidence level in this case study. When combined with the C-OCBA sampling strategy, their performances remain consistent with the

results from synthetic problems. In precision medicine applications, the simulation models are typically complicated, requiring a long time to run for one replication. Using our stopping rules combined with a high efficiency sampling strategy can significantly reduce the sample size required and save computational resources in practice. Further, in the clinical trial environment, we anticipate that our method can similarly achieve a lower sample size to meet the desired confidence level compared with IZ-based methods.

5 CONCLUSION

In this paper, we develop stopping rules for the PM sampling process. These rules effectively terminate the sampling process once the prespecified confidence level is met. Combined with efficient sampling strategies, our approach achieves high sampling efficiency while maintaining statistical guarantees for best treatment selection over the profile space. Through numerical experiments, we demonstrate the effectiveness of our stopping rules.

Although our stopping rules exhibit significant sample reductions for the PM sampling process compared to IZ-based procedures, the empirical error is still much lower than α. This conservation has been commonly seen in the literature. It is an important future research direction to optimize the calibration of stopping thresholds. Additionally, extending these stopping rules to PM applications with larger profile spaces is of interest.

REFERENCES

Alban, A., S. E. Chick, and S. I. Zoumpoulis. 2021. "Expected Value of Information Methods for Contextual Ranking and Selection: Clinical Trials and Simulation Optimization". In *2021 Winter Simulation Conference (WSC)*, 1–12 https://doi.org/10.1109/WSC52266.2021.9715303.

Branke, J., S. E. Chick, and C. Schmidt. 2007. "Selecting a Selection Procedure". *Management Science* 53(12):1916–1932 https://doi.org/10.1287/mnsc.1070.0721.

Chernoff, H. 1959. "Sequential Design of Experiments". *The Annals of Mathematical Statistics* 30(3):755–770 https://doi.org/10.1214/aoms/1177706205.

Du, J., S. Gao, and C.-H. Chen. 2024. "A Contextual Ranking and Selection Method for Personalized Medicine". *Manufacturing & Service Operations Management* 26(1):167–181 https://doi.org/10.1287/msom.2022.0232.

Forster, M., S. Brealey, S. Chick, A. Keding, B. Corbacho, A. Alban, *et al*. 2021. "Cost-effective Clinical Trial Design: Application of a Bayesian Sequential Model to the ProFHER Pragmatic Trial". *Clinical trials* 18(6):647–656 https://doi.org/10.1177/17407745211032909.

Gao, S., J. Du, and C.-H. Chen. 2019. "Selecting the Optimal System Design under Covariates". In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 547–552. Piscataway, USA: Institute of Electrical and Electronics Engineers https://doi.org/10.1109/COASE.2019.8842957.

Howard, S. R., A. Ramdas, J. McAuliffe, and J. Sekhon. 2020. "Time-Uniform Chernoff Bounds via Nonnegative Supermartingales". *Probability Surveys* 17:257–317 https://doi.org/10.1214/18-PS321.

Hur, C., N. S. Nishioka, and G. S. Gazelle. 2004. "Cost-Effectiveness of Aspirin Chemoprevention for Barrett's Esophagus". *JNCI Journal of the National Cancer Institute* 96(4):316–325 https://doi.org/10.1093/jnci/djh039.

Jiménez-Luna, J., F. Grisoni, and G. Schneider. 2020. "Drug Discovery with Explainable Artificial Intelligence". *Nature Machine Intelligence* 2(10):573–584 https://doi.org/10.1038/s42256-020-00236-4.

Jourdan, M., D. Rémy, and K. Emilie. 2023. "Dealing with Unknown Variances in Best-Arm Identification". In *Proceedings of The 34th International Conference on Algorithmic Learning Theory*, edited by S. Agrawal and F. Orabona, Proceedings of Machine Learning Research, 776–849. Singapore: Proceedings of Machine Learning Research.

Kastelein, F., M. C. Spaander, K. Biermann, E. W. Steyerberg, E. J. Kuipers, and M. J. Bruno. 2011. "Nonsteroidal Anti-Inflammatory Drugs and Statins have Chemopreventative Effects in Patients With Barrett's Esophagus". *Gastroenterology* 141(6):2000–2008 https://doi.org/10.1053/j.gastro.2011.08.036ddd.

Kato, M., and K. Ariu. 2024a. "The Role of Contextual Information in Best Arm Identification". *arXiv preprint arXiv:2106.14077*.

Kato, M., K. Okumura, T. Ishihara, and T. Kitagawa. 2024b. "Adaptive Experimental Design for Policy Learning". *arXiv preprint arXiv:2401.03756*.

Kaufmann, E., and W. M. Koolen. 2021. "Mixture Martingales Revisited with Applications to Sequential Tests and Confidence Intervals". *Journal of Machine Learning Research* 22(246):1–44.

Keslin, G., B. L. Nelson, M. Plumlee, B. K. Pagnoncelli, and H. Rahimian. 2022. "A Classification Method for Ranking and Selection with Covariates". In *2022 Winter Simulation Conference (WSC)*, 1–12 https://doi.org/10.1109/WSC57314.2022.10015235.

Kim, S.-H., and B. L. Nelson. 2001, July. "A Fully Sequential Procedure for Indifference-Zone Selection in Simulation". *ACM Transactions on Modeling and Computer Simulation* 11(3):251–273 https://doi.org/10.1145/502109.502111.

Li, H., H. Lam, Z. Liang, and Y. Peng. 2020. "Context-Dependent Ranking and Selection under a Bayesian Framework". In *2020 Winter Simulation Conference (WSC)*, 2060–2070 https://doi.org/10.1109/WSC48552.2020.9383965.

Li, H., H. Lam, and Y. Peng. 2024. "Efficient Learning for Clustering and Optimizing Context-Dependent Designs". *Operations Research* 72(2):617–638 https://doi.org/10.1287/opre.2022.2368.

Robertson, D. S., B. Choodari-Oskooei, M. Dimairo, L. Flight, P. Pallmann, and T. Jaki. 2023. "Point Estimation for Adaptive Trial Designs I: A Methodological Review". *Statistics in medicine* 42(2):122–145 https://doi.org/10.1002/sim.9605.

Shen, H., L. J. Hong, and X. Zhang. 2021. "Ranking and Selection with Covariates for Personalized Decision Making". *INFORMS Journal on Computing*:ijoc.2020.1009 https://doi.org/10.1287/ijoc.2020.1009.

Shi, X., Y. Peng, and G. Zhang. 2023. "Top-Two Thompson Sampling for Selecting Context-Dependent Best Designs". In *2023 Winter Simulation Conference (WSC)*, 3400–3411 https://doi.org/10.1109/WSC60868.2023.10407595.

Simchi-Levi, D., C. Wang, and J. Xu. 2024. "On Experimentation With Heterogeneous Subgroups: An Asymptotic Optimal δ-Weighted-PAC Design". *SSRN preprint SSRN:4721755*.

AUTHOR BIOGRAPHIES

MINGRUI DING is a Ph.D. candidate of the Department of Systems Engineering at City University of Hong Kong and the School of Economics and Management at Beihang University. His research interests include simulation optimization, combinatorial optimization and their applications to healthcare. His email address is mingrding2-c@my.cityu.edu.hk.

SIYANG GAO is an Associate Professor of the Department of Systems Engineering at City University of Hong Kong. He holds a Ph.D. in industrial engineering from University of Wisconsin-Madison, Madison. His research interests include simulation optimization, large-scale optimization and their applications to healthcare. He is a senior member of IEEE and a member of INFORMS. His email address is siyangao@cityu.edu.hk and his website is https://www.cityu.edu.hk/stfprofile/siyangao.htm.

QIUHONG ZHAO is a Professor of the School of Economics and Management at Beihang University. She holds a Ph.D. in management science from Beihang University, Beijing. Her research interests include emergency management, supply chain management, and metaheuristics algorithms. Her email address is qhzhao@buaa.edu.cn.