PH 240C/STATS 245:
Clinical Trial Design (1)

Jingshen Wang

October 20, 2021

1 Clinical Trial Design

As discussed in Pocock (1979), patient assignment in a clinical trial for the evaluation of one or more new treatments raises several fundamental statistical issues:

- *Size and representation of sample* A new treatment should be given to a sufficiently large and representative sample of patients in order to make inferences about its effectiveness for the population with a particular disease.
- *Comparative Experimentation* The results for new treatment(s) need to be compared with a control group of patients receiving a standard treatment (or placebo/no treatment if no effective standard exists).
- *Randomization* The avoidance of bias in the comparison of treatment groups may best be achieved by adopting some random mechanism for patient assignments.
- *Stratification* Randomization may need to be restricted in some way to ensure that treatment groups are comparable as regards their pre-treatment state, as summarized by one or more prognostic factors.

In practice, in most clinical trials, patients become available for entry one at a time. Therefore, patient assignment is a developing process requiring careful monitoring rather than an entirely preplanned exercise. There used to be many debates in the light of medical practice:

1. Small trials lead to unreliable conclusions.
2. Phase I trial is aimed at finding an acceptable does schedule with emphasis on toxic effects rather than disease response, thus comparison with other treatments is normally unnecessary. Phase II trial is a screening study to determine whether a treatment has activity worthy of further investigation. Phase III trial is the full scale evaluation of any new treatment, but there used to be many instances where uncontrolled studies took place. Because investigators find less ethical difficulty and quicker results and more frequent, optimistic publications by carrying out uncontrolled studies.
3. Many clinicians find randomization an unnatural interference with individual clinical judgement (Gehan and Freireich, 1974).

Randomization servers as the basis of valid statistical inference, and today it is commonly accepted randomization (with controlled and uncontrolled units) is needed to guarantee the trial result validity.

1.1 Randomization strategies

When two treatments are compared in a clinical trial, simple randomization, also called complete randomization, allocates patients randomly into two treatment groups. In medical studies, however, patients arrive subsequently and often need be treated immediately. Hence, the simple randomization (or pure randomization) procedure [1] may suffer from the at least two issues: (1) this simple randomization does not balance patients' prognostic factors such as age, gender and disease stage that may influence the outcome (although statistical inference remains valid), and (2) sampling variations may lead to unequal sized treatment groups.

Many different randomization methods have been proposed. We now discuss some of the famous ones in the long history.

1.1.1 Covariate-adaptive randomization

Covariate-adaptive randomization method, which refers to a randomized treatment allocation scheme that depends on covariates or prognostic factors but is conditionally independent of the outcomes, given the covariates used in randomization. We start with introducing some notations.

We consider two-treatment clinical trials. Suppose the patients come to the clinic sequentially and respond to treatments without delay. Let n be the total number of patients under treatments, let $D_i \in \{0, 1\}$ be the treatment assignment status, $Y_i(t)$ be the potential outcome, and

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0),$$

be the observed outcome. For patient i, let X_i be a vector of covariates used in the construction of the test procedure, and Z_i denotes the covariates or prognostic factors used in covariate-adaptive randomization. Suppose we are interested in a two-sided test

$$H_0 : \ \mathbb{E}\big[Y(1) - Y(0)\big] = 0.$$

A common characteristic of all covariate-adaptive randomization method is that, given Z_i's and X_i's, D_i is independent withs the potential outcomes.

1. **Biased coin method.** Efron (1971) proposed a "biased coin" method that assigns the ith patient to treatment with probability $p > 1/2$ according to:

$$\mathbb{P}(D_i = 1) = \begin{cases} p & G_{i-1} < 0 \\ 1/2 & G_{i-1} = 0 \\ 1 - p & G_{i-1} > 0 \end{cases}, \quad i = 1, \ldots, n,$$

   where $G_0 = 0$, and G_{i-1} is the difference between the number of patients in treatment 1 and the number of patient in the control after $i - 1$ assignments have been made. This assignment rule tends to achieve balance between the numbers of patients in two treatment groups, since $p > 1/2$ and G_{i-1}

[1] In Pocock (1979), the author mentioned a commonly used method to ensure clinicians do not know treatment assignments in advance is to: Transfer the list of randomly assigned treatment to an ordered set of sealed envelopes; after each patient is entered on trial the investigator opens the next envelope to discover which treatment is to be given to that patient. This system is not infallible and if practicable it may be better to have the randomization list kept in secret by a person not involved with the patients, who is consulted each time an assignment is needed. In multi-centre trials one should have a central assignment office which can be contacted by telephone.

is an imbalance metric. Efron (1971) shows that the difference between treatment and control armed patients after n assignments vanishes asymptotically. Note that the biased coin method does not make sure of any covariate, and thus is not a covariate-adaptive randomization method.

2. **Stratified block randomization** of Pocock and Simon (1975) applies block randomization to patients grouped by prognostic factors, are the most popular randomization methods in clinical trials. Advantages of these methods, such as minimizing imbalance between treatment groups, reducing selection bias, minimizing accidental bias and improving efficiency in inference.

   Concretely, Pocock and Simon (1975) apply biased coin method in a covariate-adaptive fashion. In a special case, suppose Z_i is discrete and takes value z_k $k = 1, \ldots, K$, a special case is to apply the biased coin method within the category of patients with $Z_i = z_k$. The motivation is to achieve balance between treatment groups for each prognostic factor. When Z_i is a continuous covariate, we can form a discrete function $D(Z_i)$ and then apply the biased coin method. This method is also called *"covariate-adaptive" biased coin method*. More generally, Pocock and Simon (1975) generalized the the minimization procedure (Taves, 1974) so as to replace G_{i-1} by

$$G_{i-1} = \sum_j w_j \Big[\sum_{i \in [j]} |D_i| - \sum_{i \in [j]} |1 - D_i| \Big],$$

   where j is the stratum formed by a prognostic factor.

3. Taves (1974)'s minimization procedure is its special case of Pocock and Simon (1975) with $p = 1$. Note that one of the oldest covariate-adaptive randomization methods is the *minimization* procedure proposed by Taves (1974).

Although covariate-adaptive randomization is commonly adopted in medical research, valid statistical testing procedures associated with it are not rigorously studied until the past decade. It is stated in for Proprietary Medicinal Products et al. (2004) that "it remains controversial whether the analysis adequately reflects the randomization scheme." Two question are of critical importance:

1. Can we develop a test procedure valid under covariate-adaptive randomization?

2. If we use covariate-adaptive randomization and a test procedure under simple randomization, will the Type-I error of the test be inflated?

3. Is a test under covariate-adaptive randomization more powerful than it is under simple randomization?

Rosenberger and Sverdlov (2008) writes "'Very little theoretical work has been done in this area, despite the proliferation of papers. The original source papers are fairly uninformative about theoretical properties of the procedures."

Shao et al. (2010) show that if the covariate Z used in covariate-adaptive randomization is a function of the covariates used to construct a test T valid under **any** fixed treatment allocation, then T is valid under covariate-adaptive randomization. But this also says T is a very conservative test given it remains valid under any fixed treatment allocation. Furthermore, Shao et al. (2010) show that two sample t-test is conservative under covariate-adaptive biased coin randomization, which is quite intuitive. Because the averaged potential outcomes are correlated between treatment groups, and the simple t-test ignores this correlation when constructing the test statistics. A less conservative test is carried out via bootstrap. Ma

et al. (2015) further provide theoretical foundation of hypothesis testing under covariate-adaptive designs based on linear models.

1.1.2 Response-adaptive randomization

Response-adaptive randomization procedures are desirable for ethical and efficiency reasons (Hu and Rosenberger, 2003). The key component of the response-adaptive randomization are the allocation proportion. Consider a binary response clinical trial where treatment A is assumed to have probability of success (failure) p_A ($q_A = 1 - p_A$), and treatment B is assumed to have probability of success (failure) p_B ($q_B = 1 - p_B$). It is well known that for fixed sample size n, we can maximize the power of the test

$$H_0 : \ p_A - p_B = 0,$$

of the simple difference:

$$R(p_A, p_B) = \frac{n_A}{n_B} = \sqrt{\frac{p_A(1-p_A)}{p_B(1-p_B)}}.$$

Such an allocation ratio $R(p_A, p_B)$, called Neyman allocation, cannot be implemented directly in a clinical trial, because we d not know the values of p_A and p_B. Even if we could implement Neyman allocation, when $p_A + p_B > 1$, we would be assigning more patients to the inferior treatment, which would compromise certain ethical objectives.

Rosenberger et al. (2001) have argued for the criteria that fix the variance of the estimator under an alternative hypothesis, to minimize the expected number of treatment failures. That is, for a binary response trial suing $p_A - p_B$ as the measure of the treatment effect, the variance of the estimator $\widehat{p}_A - \widehat{p}_B$ is fixed, say equals K:

$$\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B} = K,$$

and the expected number of treatment failtures, $n_A(1-p_A) + n_B(1-p_B)$ is minimized. This optimization problem (!again) leads to the optimal allocation ratio:

$$R(p_A, p_B) = \sqrt{\frac{p_A}{p_B}}.$$

Such an allocation deals with both the power objective and the objective favoring the individual patients' experience.

Although particular allocation ratio may be optimal in terms of power and other criteria, response-adaptive randomization induces correlation among treatment assignments that leads to extra binomial variability when performing inference. This additional variability can have severe adverse effects on power, as has been demonstrated by simulation studies in a number of papers.

1.1.3 Modern adaptive design?

While classical RCTs are often designed to maximize statistical power to detect clinically relevant differences in treatment outcomes, the objectives of RCTs are modernized in the hope of improving the overall welfare

of program participants (**?**Kasy and Sautmann, 2021) and better incorporating real world evidence (**??**). As documented in the literature (Thall and Wathen, 2007), in a clinical trial when a physician favours one treatment over another based on personal experience or published data, it may be more appropriate ethically for that physician to use the favoured treatment, rather than enrolling patients on a randomised trial. Thall and Wathen (2007) provided an example: "Mm. Fornier, I have two possible treatments for your cancer, A and B, but I do not know which is better. So I would like to enroll you in a clinical trial aimed at comparing these treatments to each other. If you agree to enter the trial, your treatment will be chosen by flipping a coin." This statement reflects the physician's equipoise, but it also reflects the fact that, outside the scientific community, randomisation is a rather strange idea. Patients entrust physicians with their well-being, and sometimes their lives, based on the assumption that physicians are highly knowledgeable and have their patients' best interests foremost in mind when choosing treatment regimens. Many physicians feel that admitting complete uncertainty, as illustrated above, may damage the bond of trust underlying the physician-patient relationship.

Now, suppose you are a trial designer and facing such dilemmas, what options do you have? To partially answer this question, we will introduce multi-armed bandit problems widely studied in computer science.

2 Multi-arm Bandit

2.1 Greedy search

The multi-armed bandit problem has been the subject of decades of intense study in statistics, operations research, electrical engineering, computer science, and economics. A "one-armed bandit" is a somewhat antiquated term for a slot machine, which tends to "rob" players of their money. The colourful name for our problem comes from a motivating story in which a gambler enters a casino and sits down at a slot machine with multiple levers, or arms, that can be pulled. When pulled, an arm produces a random payout drawn independently of the past. Because the distribution of payouts corresponding to each arm is not listed, the player can learn it only by experimenting. As the gambler learns about the arms' payouts, she or he faces a dilemma: in the immediate future one expects to earn more by **exploiting** arms that yielded high payouts in the past, but by continuing to **explore** alternative arms she may learn how to earn higher payouts in the future. Can she develop a sequential strategy for pulling arms that balances this trade-off and maximizes the cumulative payout earned? See the illustrative example in Figure 1.

Example 1. *Suppose there are K actions, and when played, any action yields either a success or a failure. Action $k \in \{1, \ldots, K\}$ produces a success with probability $\theta_k \in [0, 1]$. The success probabilities $\theta_1, \ldots, \theta_K$ are unknown to the agent, but are fixed over time. Therefore, these probabilities can be learned by experimentation. The objective, roughly speaking, is to maximize the cumulative number of successes over T periods, where T is relatively large compared to the number of arms K.*

The "arm" in different applications represents different objectives. For example, in clinical trials, an arm represents a treatment among several therapeutic interventions. Patients arrive at the clinic sequentially and receive different treatment. A success is associated with some beneficial health outcomes. In *online* advertisement, an arm represent different banner ads that can be displayed on a website. Users arriving at the site are shown versions of the website with different banner ads. A success is associated either with a click on the ad, or with a conversion (a sale of the item being advertised). The parameter θ_k represents

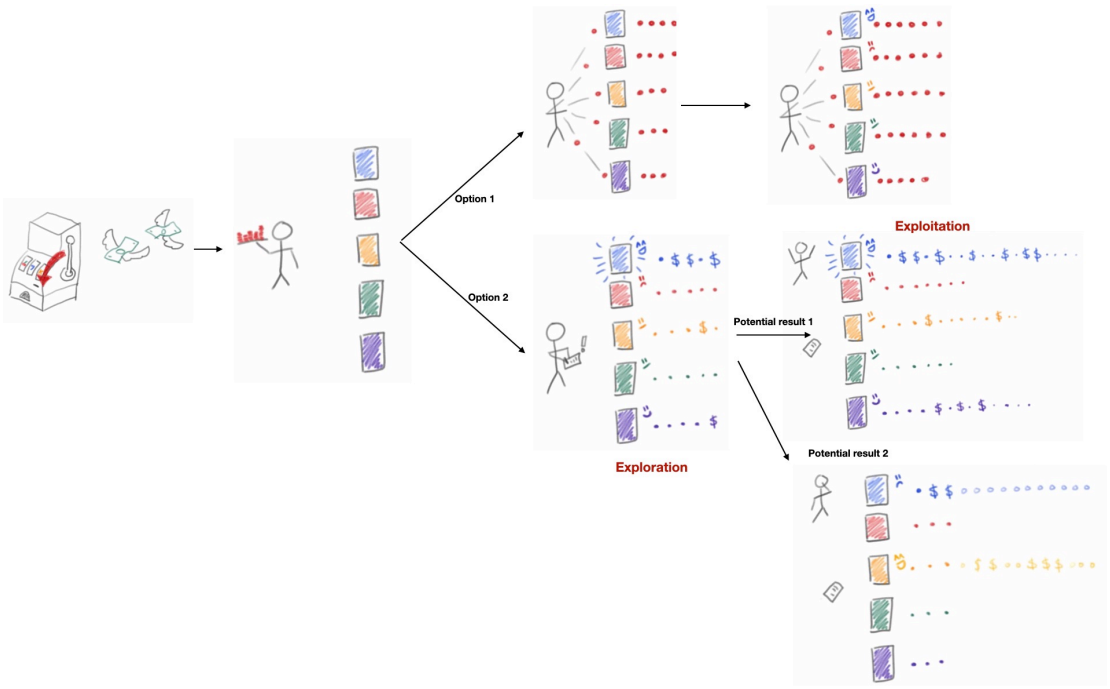

Figure 1: Illustration of multi-armed bandit problems with the slot machine example.

either the click-through- rate or conversion-rate among the population of users who frequent the site. The website hopes to balance exploration and exploitation in order to maximize the total number of successes.

A naive approach shown in Figure 1 to this problem involves allocating some fixed fraction of time periods to **exploration** (figuring out the success rate of each slot machine), and in each such period sampling an arm uniformly at random. Then select successful actions in the next time period. Then one can iterate this procedure multiple times in the hope of finding the most successful action after sufficient long time. This algorithm is also called greedy algorithms. Greedy algorithms serve as perhaps the simplest and most common approach to online decision problems. Suppose at each action step t, we are given K different choices of actions. The following two steps are taken to generate each action:

1. Estimate a model from historical data $H_{t-1} = \{(A_1, Y_1), \ldots, (A_{t-1}, Y_{t-1})\}$, where A_{t-1} is the action taken, and Y_{t-1} is the observed outcome at $t-1$ step

2. Select the action A_t that is optimal for the estimated model, breaking ties in an arbitrary manner. The action $A_t = \arg\max_k \{\text{Reward estimated for action k using } H_{t-1}\}$

Such an algorithm is greedy in the sense that an action is chosen solely to maximize immediate reward. A shortcoming of the greedy approach, which can severely curtail performance, is that it does not actively explore.

To see its drawback more concretely, suppose there are three actions with mean rewards $\theta \in \mathbb{R}^3$. In particular, each time an action k is selected, a reward of 1 is generated with probability θ_k. Otherwise, a reward of 0 is generated. The mean rewards are not known to the agent. Instead, the agent's knowledge in any given time period about these mean rewards can be expressed in terms of conditional distributions. Suppose, conditional on the observed history, the distribution of rewards is plotted in Figure 2.

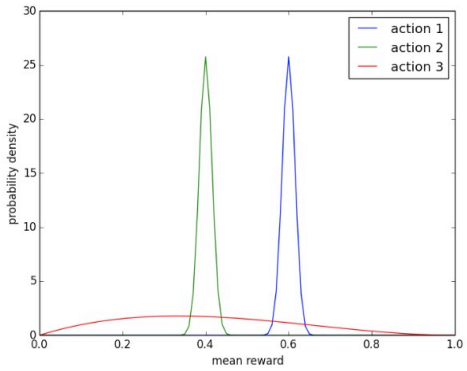

Figure 2.2: Probability density functions over mean rewards.

Figure 2: Illustration of multi-armed bandit problems with greedy algorithms.

The greedy algorithm would select action 1, since that offers the maximal expected mean reward. The algorithm is also likely to avoid action2, since it is extremely unlikely that $\theta_2 > \theta_1$. Although there is some chance that $\theta_3 > \theta_1$, the agent would need to try action 3, but the greedy algorithm will unlikely ever do that. The algorithm fails to account for uncertainty in the mean reward of action 3, which should entice the agent to explore and learn about that action.

$\varepsilon-$greedy exploration is one method to force the algorithm to explore. It applies the greedy action with probability $1-\varepsilon$ and otherwise selects an action uniformly at random. Though this form of exploration can improve behavior relative to a purely greedy approach, it wastes resources by failing to "write off" actions regardless of how unlikely they are to be optimal. In the example in Figure 2, we can see that action 2 has almost no chance of being optimal, and therefore, does not deserve experimental trials, while the uncertainty surrounding action 3 warrants exploration. However, $\varepsilon-$greedy exploration allocates an equal number of experimental trials to each action. Though only half of the exploratory actions are wasted in this example, the issue is exacerbated as the number of possible actions increases. Thompson sampling, introduced more than eight decades ago, provides an alternative to dithering that more intelligently allocates exploration effort.

Note that the above part of lecture note for greedy search is largely based on the comprehensive overview of Thompson sampling in Russo et al. (2017).

Let's now take a pause here. In healthcare, in what kind of applications, do you think multi-arm bandits will be helpful?

2.2 Thompson sampling

Thompson sampling is an algorithm for online decision problems where actions are taken sequentially in a manner that must balance between exploiting what is known to maximize immediate performance and investing to accumulate new information that may improve future performance. As it depends on bayesian statistics, we will first introduce Bayesian analysis in our next class.

References

Bradley Efron. Forcing a sequential experiment to be balanced. *Biometrika*, 58(3):403–417, 1971.

Committee for Proprietary Medicinal Products et al. Committee for proprietary medicinal products (cpmp) point to consider on adjustment for baseline covariates. *Statistics in medicine*, 23:701–709, 2004.

Edmund A Gehan and Emil J Freireich. Non-randomized controls in cancer clinical trials. *New England Journal of Medicine*, 290(4):198–203, 1974.

Feifang Hu and William F Rosenberger. Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association*, 98 (463):671–678, 2003.

Maximilian Kasy and Anja Sautmann. Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1):113–132, 2021.

Wei Ma, Feifang Hu, and Lixin Zhang. Testing hypotheses of covariate-adaptive randomized clinical trials. *Journal of the American Statistical Association*, 110(510):669–680, 2015.

Stuart J Pocock. Allocation of patients to treatment in clinical trials. *Biometrics*, pages 183–197, 1979.

Stuart J Pocock and Richard Simon. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*, pages 103–115, 1975.

William F Rosenberger and Oleksandr Sverdlov. Handling covariates in the design of clinical trials. *Statistical Science*, 23(3):404–419, 2008.

William F Rosenberger, Nigel Stallard, Anastasia Ivanova, Cherice N Harper, and Michelle L Ricks. Optimal adaptive designs for binary response trials. *Biometrics*, 57(3):909–913, 2001.

Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, and Zheng Wen. A tutorial on thompson sampling. *arXiv preprint arXiv:1707.02038*, 2017.

Jun Shao, Xinxin Yu, and Bob Zhong. A theory for testing hypotheses under covariate-adaptive randomization. *Biometrika*, 97(2):347–360, 2010.

Donald R Taves. Minimization: a new method of assigning patients to treatment and control groups. *Clinical Pharmacology & Therapeutics*, 15(5):443–453, 1974.

Peter F Thall and J Kyle Wathen. Practical bayesian adaptive randomisation in clinical trials. *European Journal of Cancer*, 43(5):859–866, 2007.