PLAY-THE-WINNER RULE AND ADAPTIVE DESIGNS OF CLINICAL TRIALS

XIKUI WANG and DARYL PULLMAN

ABSTRACT. In another paper, we have argued that the traditional randomized design of clinical trials is ethically infeasible in desperate medical situations and adaptive designs are morally required. We have also argued that in such situations, the appropriate designs must satisfy what we call the "Principle of interchangeability." In this statistics paper, we show that the statistical model of bandit processes satisfies this principle of interchangeability. Moreover, we demonstrate that when such a model is used as an adaptive design, the total regret of successes lost is smaller when compared with simple randomization. We illustrate the results by the simple deterministic play-the-winner design.

2000 Mathematics Subject Classification. 62K05, 62L05.

1. Introduction. In our ethics paper (Pullman and Wang [17]) on the ethical necessity to use adaptive designs in clinical research, we have argued that the key consideration in clinical research depends on the nature of the relationship between the clinician/researcher and the patient/subject, instead of on the tension between individual and collective ethics. Our study on the dynamics of the researcher's moral responsibility shows that the process of informed consent is central to this relationship. Randomization is justified as long as the patient/subject understands the nature of the proposed clinical research and the risks involved, and is capable of providing a fully informed consent. The ethical justification for randomization hence lies in the process of informed consent.

However, in desperate medical situations, the patient/subject is justifiably concerned with his/her own physical well-being, and from his/her point of view the physician's dual role as a researcher is irrelevant to the clinical decision that must be made. The patient is no longer capable of comprehending the nature of the proposed clinical research and the risks involved, and his/her capacity to provide a fully informed consent is almost entirely compromised.

As subjects/patients lose their capacity for autonomous choice, it is incumbent upon the researchers to assume greater responsibility for their care and well-being. The appropriate designs of the clinical research must then be such, so as to minimize the risk to individual patients. Adaptive clinical trials are designed for this purpose. Moreover, informed consent does not provide the ethical justification for adaptive designs as it does for randomization. Hence adaptive designs are ethically justified and may be morally required in desperate medical situations.

We have argued that the appropriate adaptive design must satisfy what we call the "Principle of interchangeability." Suppose that there are N patients to be treated both

in and after the trial, and each patient is treated by one and only one of two treatments. We say that a design satisfies the principle of interchangeability if any two of these N patients are ethically interchangeable. That is, at the point of enrollment in the clinical trial, the intent is to provide the best treatment available to each patient given current information. This principle meets the ethical imperative that clinicians must always endeavour to provide the best overall treatment for each individual patient and so each patient's fate is determined not by the particular design of the trial but instead by the chance and timing of getting the disease.

Randomized trials do not satisfy this principle of interchangeability because a patient in the trial and another patient after the trial are not ethically interchangeable. In Section 2, we demonstrate that the principle of interchangeability is satisfied if we utilize the statistical model of bandit processes as an adaptive design. We also compare this adaptive model with randomization and show that more lives can be saved under adaptive designs. The comparison is illustrated by the simple deterministic play-the-winner design. Finally, we conclude with a discussion on relevant studies and further research directions.

2. Bandit processes and the play-the-winner rule. For our discussion, assume immediate and dichotomous responses and two treatments A and B with probabilities of success P_A and P_B, respectively. We denote Z_i, $i = 1, 2, \ldots, N$, as the response from the ith patient, which is either 1 for a success or 0 for a failure. As noted by Berry and Eick [3], N depends on the prevalence of the disease and is normally unknown to us.

The principle of interchangeability is satisfied when we maximize the total expected responses from all N patients. Such an objective is particularly desirable from the subject/patients' perspective. Following Berry and Eick [3], we maximize $W_\pi(P_A, P_B) = E_\pi(Z_1 + Z_2 + \cdots + Z_N \mid P_A, P_B)$, conditionally on P_A, P_B, and N, or unconditionally, where π is a strategy for allocating treatments to these N patients. Any optimal strategy π^* which maximizes $W_\pi(P_A, P_B)$ is characterized by the dynamic programming equation which states that the current patient is offered the best treatment under the current information, given that all future patients are treated optimally. Such a recursive property of the optimality equation demonstrates the satisfaction of the principle of interchangeability.

This optimization problem is a typical bandit problem (Berry and Fristedt [4], Gittins [11], and Presman and Sonin [16]). The fundamental characteristic of an optimal strategy is that it enables us to compromise between the need to gather information about the unknown effectiveness of the treatments in order to provide better treatments in the future and the imperative ethics to maximize the immediate response for the current patient (Berry and Fristedt [4]). By nature such a strategy must be adaptive. Randomization aims only at gathering information and ignores immediate responses. On the other hand, the myopic strategy focuses on immediate responses and ignores information gathering. These are two extreme strategies and are not optimal in general.

Suppose that $n(\leq N)$ patients are included in the clinical trial, and the rest $N - n$ patients are treated with the superior treatment identified at the end of the trial. n may be either fixed or random. If P_A and P_B were known, the total expected number of successes is $N \max\{P_A, P_B\}$. The conditional regret of successes lost for strategy π

when not knowing the treatment effectiveness P_A, P_B is defined as

$$R_\pi(P_A, P_B) = N \max\{P_A, P_B\} - E_\pi(Z_1 + Z_2 + \cdots + Z_N \mid P_A, P_B)$$

$$= (N - n)[\max\{P_A, P_B\} - \max\{P_A^\pi, P_B^\pi\}]$$

$$+ E_\pi\left[\sum_{i=1}^n (\max\{P_A, P_B\} - Z_i) \mid P_A, P_B\right] \tag{2.1}$$

$$= R_{\pi,1}(P_A, P_B) + R_{\pi,2}(P_A, P_B),$$

where P_A^π and P_B^π are, respectively, the estimated probabilities of success on treatments A and B at the conclusion of the trial.

$R_{\pi,1}(P_A, P_B) = (N - n)[\max\{P_A, P_B\} - \max\{P_A^\pi, P_B^\pi\}]$ is the total regrets from all patients after the trial and is determined by the trial conclusion. On the other hand, $R_{\pi,2}(P_A, P_B) = E_\pi[\sum_{i=1}^n (\max\{P_A, P_B\} - Z_i) \mid P_A, P_B]$ is the total regrets from all patients in the trial and is determined by the trial design and treatment allocations. It is ideal to minimize both $R_{\pi,1}(P_A, P_B)$ and $R_{\pi,2}(P_A, P_B)$ simultaneously in order to minimize $R_\pi(P_A, P_B)$.

Compare the performance of randomization π_1 and an adaptive design π_2 with respect to $R_\pi(P_A, P_B)$. Conditioning on (P_A, P_B), $E_{\pi_1}(Z_i)(P_A, P_B) = (1/2)P_A + (1/2)P_B$ for all $i = 1, \ldots, n$, and

$$R_{\pi_1,2}(P_A, P_B) = \sum_{i=1}^n \left[\max\{P_A, P_B\} - \left(\frac{1}{2}P_A + \frac{1}{2}P_B\right) \Big| P_A, P_B \right]$$

$$= \frac{n}{2}(\max\{P_A, P_B\} - \min\{P_A, P_B\}). \tag{2.2}$$

Fixed-size randomized trials use fixed sample sizes for n and minimize $R_{\pi,1}(P_A, P_B)$ only. In a sequential trial, n is adaptive and minimized such that the trial is stopped as soon as there is strong evidence to indicate the superiority or inferiority of one treatment. Hence a sequential trial minimizes in general both $R_{\pi,1}(P_A, P_B)$ and $R_{\pi,2}(P_A, P_B)$. This clearly indicates that sequential trials are both ethically and statistically superior to fixed-size randomized trials, especially when there is a large difference between the two treatments.

Let $p_i = P_{\pi_2}$ (treatment A is used for patient i) for $i = 1, \ldots, n$ under the strategy π_2. Then $E_{\pi_2}(Z_i \mid P_A, P_B) = p_i P_A + (1 - p_i)P_B$ and

$$R_{\pi_2,2}(P_A, P_B) = n \max\{P_A, P_B\} - \left(\sum_{i=1}^n p_i\right)P_A - \sum_{i=1}^n (1 - p_i)P_B$$

$$= \begin{cases} \left(n - \displaystyle\sum_{i=1}^n p_i\right)(P_A - P_B) & \text{if } P_A \geq P_B, \\ \left(\displaystyle\sum_{i=1}^n p_i\right)(P_B - P_A) & \text{if } P_A < P_B. \end{cases} \tag{2.3}$$

To minimize $R_{\pi_2,2}(P_A, P_B)$, it is desired to maximize $\sum_{i=1}^n p_i$ when $P_A > P_B$ or to minimize $\sum_{i=1}^n p_i$ when $P_A < P_B$. The conditional advantage of adaptive design π_2 over

randomization π_1 for R_2 is

$$R_{\pi_2,2}(P_A,P_B) - R_{\pi_1,2}(P_A,P_B) = \left(\frac{n}{2} - \sum_{i=1}^{n} p_i\right)(P_A - P_B), \tag{2.4}$$

and $R_{\pi_2,2}(P_A,P_B) < R_{\pi_1,2}(P_A,P_B)$ if and only if

$$P_A > P_B, \qquad \sum_{i=1}^{n} p_i > \frac{n}{2}, \tag{2.5}$$

or

$$P_A < P_B, \qquad \sum_{i=1}^{n} p_i < \frac{n}{2}. \tag{2.6}$$

If P_A and P_B have a joint density function

$$f(P_A,P_B) = \begin{cases} q_1 & \text{on } 0 \leq P_B \leq P_A \leq 1, \\ q_2 & \text{on } 0 \leq P_A < P_B \leq 1, \end{cases} \tag{2.7}$$

then the unconditional advantage of adaptive designs over randomization for R_2 is

$$R_{\pi_2,2} - R_{\pi_1,2} = \frac{1}{6}\left(\frac{n}{2} - \sum_{i=1}^{n} p_i\right)(q_1 - q_2), \tag{2.8}$$

which is negative if and only if

$$q_1 > q_2, \qquad \sum_{i=1}^{n} p_i > \frac{n}{2}, \tag{2.9}$$

or

$$q_1 < q_2, \qquad \sum_{i=1}^{n} p_i < \frac{n}{2}. \tag{2.10}$$

In both conditional and unconditional cases, the total regret for all patients in the trial is less if the total probability of using the superior treatment for all patients in the trial is more than $n/2$. This is exactly what is expected through the use of adaptive designs.

It is hypothesized that in general, the probability that the superior treatment is allocated to each patient would be more than 0.5 under the use of an adaptive design. Although its demonstration may be difficult in general, we show it for the deterministic play-the-winner design (Zelen [28]). Under this design, the first patient is allocated to either treatment by a simple randomization. Then the same treatment is applied after a success and the other treatment is used after a failure. Assuming clinical equipose (Freedman [10]) (i.e., there is genuine uncertainty at the beginning of the trial as to which treatment is superior), the principle of interchangeability is satisfied when the initial patient is enrolled.

Suppose the deterministic play-the-winner rule is followed. Let $\Delta = P_A - P_B$ and $K = P_A + P_B$. To avoid triviality, assume that $K \neq 2$. Then $p_1 = 1/2$ and for any integer $n \geq 0$, by mathematical induction,

$$p_{n+1} = \frac{1}{2} + \frac{1}{2}\Delta + \frac{1}{2}\Delta(K-1) + \cdots + \frac{1}{2}\Delta(K-1)^{n-1} = \frac{1}{2} + \frac{1}{2}\Delta\frac{1 - (K-1)^n}{2 - K}. \tag{2.11}$$

This sequence of probabilities of allocations on treatment A has the following properties.

(a) $p_{n+1} > 1/2$ for any integer $n \geq 0$ if $P_A > P_B$, $p_{n+1} < 1/2$ if $P_A < P_B$, and $p_{n+1} = 0.5$ if $P_A = P_B$. Moreover, when $K > 1$, p_{n+1} is increasing in n when $P_A > P_B$, and is decreasing when $P_A < P_B$.

(b) $P(p_{n+1} > 1/2) > P(P_A > P_B)$.

(c) $p = \lim_{n \to \infty} p_{n+1} = (1 - P_B)/((1 - P_A) + (1 - P_B)) = q_B/(q_A + q_B)$ exists since p_n is a Cauchy sequence, where $q_A = 1 - P_A$, $q_B = 1 - P_B$.

(d) $p > 1/2$ if $P_A > P_B$.

(e) $P(p > 1/2) = P(P_A > P_B)$.

(f) $\lim_{P_A \to 1} p = 1$, $\lim_{\Delta \to 0} p = 1/2$, $\lim_{P_B \to 1} p = 0$.

(g) If $K = 1$, then $p_{n+1} = p = P_A$.

These results have intuitive interpretations. (a), (b), (d), and (e) say that the probability of using the superior treatment is more than 50%. (c) indicates that p is simply the asymptotic fraction of patients on treatment A. Zelen [28] achieved the same result based on a different approach. Wei and Durham [26] observed the same result for randomized play-the-winner rule. (f) shows that if P_A (or P_B) is sufficiently large, treatment A (or B) will eventually be identified as the superior treatment. On the other hand, if there is no difference between the two treatments, we eventually randomize between them. (g) assumes that $P_B = 1 - P_A$. Suppose that $P_A > 0.5$. If we call a "success" if treatment A is allocated and a "failure" if B is used, we essentially have a binomial experiment with P_A as the probability of success. The probability of using the superior treatment remains a constant larger than 0.5, and the expected number of successes is then $N \cdot P_A > N/2$.

For $n < N$, we have

$$\sum_{i=1}^{n} p_i = \frac{n}{2} + \frac{\Delta}{2(2 - K)}\left[n - \frac{1 - (K-1)^n}{2 - K}\right]. \tag{2.12}$$

Then conditioning on (P_A, P_B),

$$R_{\pi_2}(P_A, P_B) - R_{\pi_1}(P_A, P_B) = (R_{\pi_2,2} - R_{\pi_1,2}) + (R_{\pi_2,1} - R_{\pi_1,1})$$
$$\leq (N - n) + \frac{\Delta^2}{(2 - K)^2} - \frac{n\Delta^2}{2(2 - K)} = B. \tag{2.13}$$

For given values of P_A and P_B, this upper bound B is decreasing in $n < N$ and is negative for sufficiently large n. In fact, if $n/N > [1 + \Delta^2/(N(2-K)^2)][2(2-K)/(2(2-K)+\Delta^2)]$, then the upper bound B is negative. The values of $[1 + \Delta^2/(N(2-K)^2)][2(2-K)/(2(2-K)+\Delta^2)]$ when $N = 100$ is given in Table 2.1. The values become smaller when N is increased.

Therefore if the majority of all patients are recruited into the trial (i.e., n/N is large), deterministic play-the-winner design is statistically superior to randomization in the sense that the maximum number of patients in and after the trial are treated successfully. This proportion n/N becomes smaller for larger difference between the treatments.

TABLE 2.1

	P_A								
P_B	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	1.000	0.997	0.988	0.971	0.947	0.914	0.872	0.821	0.762
0.2		1.000	0.997	0.986	0.967	0.939	0.900	0.851	0.791
0.3			1.000	0.996	0.984	0.961	0.927	0.881	0.821
0.4				1.000	0.996	0.981	0.953	0.911	0.853
0.5					1.000	0.995	0.976	0.941	0.886
0.6						1.000	0.993	0.969	0.921
0.7							1.000	0.991	0.953
0.8								1.000	0.985
0.9									1.000

3. Conclusion and discussions. The deterministic play-the-winner rule has been used in some clinical trials such as Bjerkeset et al. [5], Larsen et al. [14], and Reiertsen et al. [18, 19, 20]. Since its use in practice may introduce selection bias, it is later generalized to the randomized play-the-winner design (Wei and Durham [26]) which has been used in Bartlett et al. [2] and Tamura et al. [25].

Many simulations have demonstrated the superiority of adaptive designs over randomization. For example, Berry and Eick [3] proposed and examined the objective of optimization which we have discussed here. Based on simulations, they suggest that adaptive designs may be appropriate when the majority of the patients with a particular disease are recruited into the trial. Our statistical arguments are consistent with their observations. Yao and Wei [27] have simulated the AZT trial for reducing the risk of maternal-to-infant HIV transmission (Connor et al. [7], Rosenberger [21]), and concluded that more newborns could have been saved with no jeopardy to the statistical power if randomized play-the-winner design were used instead of randomization. Day [8] shows that adaptive clinical trials are better than sequential clinical trials, which are in turn better than randomized clinical trials. Simon [24] reports that the gain of adaptive clinical trials over sequential clinical trials is relatively modest, and Hallstrom et al. [12] indicate that the power is largely unaffected by using the play-the-winner rule in a typical chronic disease mortality trial. More recently, Coad and Rosenberger [6] have reported reduction of failures when the randomized play-the-winner design is combined with a fully sequential triangular test. Also see Flehinger and Louis [9], Louis [15], and Rosenberger and Seshaiyer [23] about adaptive clinical trials for survival trials.

The use of adaptive designs have also been suggested by Hardwick [13], Rosenberger and Lachin [22], and others. Yet adaptive designs have never become part of the mainstream clinical research methodology even though they seem to be desirable in desperate medical situations from ethical point of view. Pullman and Wang [17] (and references quoted) have discussed the many reasons behind, including ideological resistance.

What kind of adaptive design is both statistically optimal and practically feasible remains an open problem. It is also a major difficulty to develop appropriate methods

of statistical inferences for adaptive designs. We use the deterministic play-the-winner design for illustration purpose only.

Despite these challenges in both the design and analysis of adaptive clinical trials, we are confident that adaptive designs will remain an important and active research area. After all, this seems to be the kind of contribution we statisticians should be making (Armitage [1]). This area has a bright future because when adaptive trials are properly designed and analyzed, we statisticians may save more lives than the medical doctors.

REFERENCES

[1] P. Armitage, *The search for optimality in clinical trials*, Internat. Statist. Rev. **53** (1985), no. 1, 15–24. MR 89i:62172. Zbl 0586.62129.

[2] R. H. Bartlett, D. W. Roloff, R. G. Cornell, A. F. Andrews, P. W. Dillon, and J. B. Zwischenberger, *Extracorporeal circulation in neonatal respiratory failure: a prospective randomized study*, Pediatrics **76** (1985), 479–487.

[3] D. A. Berry and S. G. Eick, *Adaptive assignment versus balanced randomization in clinical trials: a decision analysis*, Stat. Med. **14** (1995), 231–246.

[4] D. A. Berry and B. Fristedt, *Bandit Problems. Sequential Allocation of Experiments*, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1985. MR 87i:62137. Zbl 0659.62086.

[5] O. Bjerkeset, S. Larsen, and O. Reiertsen, *Evaluation of enoxaparin given before and after operation to prevent venous thromboembolism during digestive surgery: play-the-winner designed study*, World J. Surg. **21** (1997), 584–589.

[6] D. S. Coad and W. F. Rosenberger, *A comparison of the randomized play-the-winner rule and the triangular test for clinical trials with binary responses*, Stat. Med. **18** (1999), 761–769.

[7] E. M. Connor, R. S. Sperling, R. Gelber, P. Kiselev, G. Scott, M. J. O'Sullivan, R. Van Dyke, M. Bey, W. Shearer, R. J. Jacobson, E. Jimenez, E. O'Neill, B. Bazin, J.-F. Delfraissy, M. Culnane, R. Coombs, M. Elkins, J. Moye, P. Stratton, and J. Balsley for the Pediatric AIDS Clinical Trials Group Protocol 076 Study Group, *Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovidine treatment*, N. Engl. J. Med. **331** (1994), 1173–1180.

[8] N. E. Day, *A comparison of some sequential designs*, Biometrika **56** (1969), 301–311.

[9] B. J. Flehinger and T. A. Louis, *Sequential treatment allocation in clinical trials*, Biometrika **58** (1971), 419–426. MR 47#1217.

[10] B. Freedman, *Equipoise and the ethics of clinical research*, N. Engl. J. Med. **317** (1987), 141–145.

[11] J. C. Gittins, *Multi-Armed Bandit Allocation Indices*, John Wiley & Sons, Chichester, 1989. MR 90e:62113. Zbl 0699.90068.

[12] A. Hallstrom, M. M. Brooks, and M. Peckova, *Logrank play the winner, power and ethics*, Stat. Med. **15** (1996), 2135–2142.

[13] J. Hardwick, *Comment: recent progress in clinical trial designs that adapt for ethical purposes*, Statist. Sci. **4** (1989), 327–336.

[14] S. Larsen, O. Reiertsen, P. Mowinckel, H. Lund, and M. Osnes, *"Play-the-winner". A step-by-step process design to compare and control the efficacy and safety of treatment*, Pharm. Med. **8** (1994), 11–23.

[15] T. A. Louis, *Sequential allocation in clinical trials comparing two exponential survival curves*, Biometrics **33** (1977), no. 4, 627–634. MR 57#4443.

[16] È. L. Presman and I. N. Sonin, *Sequential Control with Incomplete Information*, Economic Theory, Econometrics, and Mathematical Economics, Academic Press, California, 1990. MR 94d:93004. Zbl 0721.93003.

[17] D. Pullman and X. Wang, *Adaptive designs, informed consent, and the ethics of research*, Control. Clin. Trials **22** (2001), 203–210.

[18] O. Reiertsen, S. Larsen, and J. H. Solhaug, *Development of play-the-winner design: weighting and accumulation of information*, Scand. J. Clin. Lab. Invest. **58** (1998), 241–250.

[19] O. Reiertsen, S. Larsen, R. Storkson, E. Trondsen, T. Lovig, O. K. Andersen, H. Lund, and P. Mowinckel, *Safety of enoxaparin and dextran-70 in the prevention of venous thromboembolism in digestive surgery. A play-the-winner-designed study*, Scand. J. Gastroenterol. **28** (1993), 1015–1020.

[20] O. Reiertsen, P. Mowinckel, O. Bjerkeseth, T. Lovig, G. Thorsen, T. Gerner, T. Lotveit, and S. Larsen, *Characterization of "winners" to enoxaparin in the prevention of postoperative venous thromboembolism in digestive surgery*, Scand. J. Gastroenterol. **31** (1996), 616–621.

[21] W. F. Rosenberger, *New directions in adaptive designs*, Statist. Sci. **11** (1996), 137–149.

[22] W. F. Rosenberger and J. M. Lachin, *The use of response-adaptive designs in clinical trials*, Control. Clin. Trials **14** (1993), 471–484.

[23] W. F. Rosenberger and P. Seshaiyer, *Adaptive survival trials*, J. Biopharm. Stat. **7** (1997), 617–624.

[24] R. Simon, *Adaptive treatment assignment methods and clinical trials*, Biometrics **33** (1977), 743–749.

[25] R. N. Tamura, E. F. Douglas, J. S. Andersen, and J. H. Heiligenstein, *A case study of an adaptive clinical trial in the treatment of out-patients with depressive disorder*, J. Amer. Statist. Assoc. **89** (1994), 768–776.

[26] L. J. Wei and S. Durham, *The randomized play-the-winner rule in medical trials*, J. Amer. Statist. Assoc. **73** (1978), 840–843. Zbl 0391.62076.

[27] Q. Yao and L. J. Wei, *Play the winner for phase II/III clinical trials*, Stat. Med. **15** (1996), 2413–2423.

[28] M. Zelen, *Play the winner rule and the controlled clinical trial*, J. Amer. Statist. Assoc. **64** (1969), 131–146. MR 39#2283.

XIKUI WANG: DEPARTMENT OF STATISTICS, UNIVERSITY OF MANITOBA, WINNIPEG, MANITOBA, CANADA R3T 2N2
E-mail address: xikui_wang@umanitoba.ca

DARYL PULLMAN: FACULTY OF MEDICINE, MEMORIAL UNIVERSITY OF NEWFOUNDLAND, ST. JOHN'S, NEWFOUNDLAND, CANADA A1B 3V6
E-mail address: dpullman@morgan.ucs.mun.ca