Statistica Neerlandica (2007) Vol. 61, nr. 3, pp. 329–344

Adaptive two-treatment two-period crossover design for binary treatment responses

Uttam Bandyopadhyay*

Department of Statistics, University of Calcutta, 35 Ballygunge Circular Road, Kolkata – 700 019, India

Atanu Biswas†

Applied Statistics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata – 700 108, India

Shirsendu Mukherjee‡

Department of Statistics, University of Calcutta, 35 Ballygunge Circular Road, Kolkata – 700 019, India

Adaptive designs are sometimes used in a phase III clinical trial with the aim, of allocating a larger number of patients to the better treatment. In the present paper, we use some adaptive designs in a two-treatment two-period crossover design where the treatment responses are binary. We use some simple designs to choose between the possible treatment combinations AA, AB, BA or BB. The goal is to use the better treatment a larger proportion of times. We calculate the allocation proportions to the possible treatment combinations and their standard deviations. We also study related inferential problems. Related asymptotics are derived. The proposed procedure is compared with some possible competitors. Finally, we use real data to illustrate the applicability of our proposed design.

Keywords and Phrases: asymptotic power, conditional test, ethics, limiting proportion of allocation, maximum likelihood estimates, play-the-winner rule, randomized clinical trials, randomized play-the-winner rule, repeated measurement designs.

1 Introduction

Repeated measurement designs (sometimes referred to as crossover designs or change-over designs) are frequently used in clinical trials to compare the efficacy of several different treatments. Brown (1980) reported that in a survey of studies investigating the effects of anti-anxiety drugs on humans, 68% of the studies used the crossover approach, which is evidence of the popularity of repeated measurement designs, even

quarter of a century back. Reasons for the widespread use of repeated measurement designs are discussed in Hedayat and Afsarinejad (1978) and Carriere and Reinsel (1992).

Many researchers (e.g. Hedayat and Afsarinejad, 1978; Cheng and Wu, 1980; Laska and Meisner, 1985; Matthews, 1987; Kunert, 1991; Kushner, 1997) have constructed optimal designs under various assumptions about the model. Once adopted, clinicians are to allocate subjects to one of several study sequences on the basis of a randomization scheme provided by the protocol. However, these randomized clinical trials ignore the long-standing tension between individual and collective ethics (Clayton, 1982; Palmer and Rosenberger, 1999); that is, doing what is best for individual patients in the current trial versus doing what is best for future patients who would benefit from the knowledge gained as a result of the current randomized trial. The goal of randomized trials is usually to acquire statistically valid information to benefit future patients.

Adaptive clinical trials, on the other hand, give priority to individual ethics by allocating a treatment to each patient on the basis of accumulated information. The objective is to treat as many patients as effectively as possible. The pros and cons of adaptive designs are discussed in Rosenberger (1996) and Biswas (2001). They identify situations when adaptive designs are clearly beneficial from an ethical point of view. However, although several investigators (e.g. Atkinson, 1982; Rosenberger *et al.*, 2001; Atkinson and Biswas, 2005) have worked on optimal adaptive designs, there are no guidelines on how to design for experiments with repeated measures data, especially when the experiments involve binary responses. This paper investigates response-adaptive crossover designs for correlated binary responses based on a play-the-winner principle.

The paper is organized as follows. Section 2 describes the design. The present paper considers the simplest possible set-up, having two treatments and two time points, and a very simple adaptive design. Section 3 provides the inferences under independence of parameters. Specifically, we obtain the asymptotic joint distribution of the estimators of the parameters, and the asymptotic distribution of the test statistics. We then discuss the performance characteristics of the design and the follow-up test in section 4. We examine the power of an exact conditional test, the power of a naive test and asymptotic power, as well as the exact proportion of allocation in different treatment combinations in section 4. The proposed design is compared with the corresponding equal allocation design and that is also discussed in section 5. Section 6 gives an example of the application of our approach. Section 7 provides the overall conclusions.

2 The proposed design

We consider a two-treatment, two-period crossover design where the treatment responses are binary, say of the success/failure type. Suppose the two treatments are

denoted by A and B. Let n patients be treated by this design where n is any prefixed positive integer. Each patient is treated in any one of the following sequences: AA, AB, BA, BB. We consider a simple hypothetical set-up where we have a sequential entry of patients and the ith patient enters the system at time i, and is treated at the time points i and $i+1$. Again we assume that the responses are instantaneous, i.e. the response of the patients at time point i is available before time point $(i+1)$.

For the ith patient, we define the indicator variables $\{\delta_i, Z_i, \tau_i, U_i\}$ such that $\delta_i = 1$ or 0 as the ith patient receives the first dose of treatment A or B; $Z_i = 1$ or 0 as success or failure occurs with the ith patient after receiving the first dose of treatment; $\tau_i = 1$ or 0 as the ith patient receives the second dose of treatment A or B; $U_i = 1$ or 0 as success or failure occurs with the ith patient after receiving the second dose of treatment.

We now define the following notations.

$$\delta_j^c = 1 - \delta_j, \quad \tau_j^c = 1 - \tau_j,$$
$$P(Z_i = 1 | \delta_i = 1) = p_{\mathrm{A}}, \quad P(Z_i = 1 | \delta_i = 0) = p_{\mathrm{B}},$$
$$P(U_i = 1 | \tau_i = 1) = \phi_{\mathrm{A}}, \quad P(U_i = 1 | \tau_i = 1) = \phi_{\mathrm{B}},$$

with $q_k = 1 - p_k$ and $\psi_k = 1 - \phi_k$, $k = \mathrm{A}, \mathrm{B}$.

For any patient entering the study, let the first treatment (either A or B) be decided by a design $\mathcal{D}_1$ and the second treatment by another design $\mathcal{D}_2$. Both $\mathcal{D}_1$ and $\mathcal{D}_2$ are adaptive in the sense that we use all or some of the earlier data to determine the treatment to be allocated. Thus the adaptive design used in this two-period crossover design is $\mathcal{D} \equiv \mathcal{D}_1 \cup \mathcal{D}_2$. Our method is flexible enough to accomodate any reasonable $\mathcal{D}_1$ and $\mathcal{D}_2$. The degree of mathematical difficulty will depend on this choice. For the sake of simplicity, and as an illustration, here we consider the simple randomized play-the-winner (RPW) rule as $\mathcal{D}_1$ and the simple play-the-winner (PW) rule as $\mathcal{D}_2$. The intuitive appeal and importance of these designs are apparent, in that most of the real-life adaptive clinical trials (not crossover trials) to date are either by PW (unpublished application of Zelen in a lung cancer trial, reported by Iglewicz, 1983; Rout *et al.*, 1993) or by RPW (Bartlett *et al.*, 1985; Tamura *et al.*, 1994; Biswas and Dewanji, 2004).

The PW rule was introduced by Zelen (1969). For a two-treatment set-up, if any patient had a successful response, this rule allocates the next patient to the same treatment. On the other hand, if the treatment failed, the next patient receives the other treatment.

Wei and Durham (1978) and Wei (1979) introduced the RPW rule, a modification of Zelen's PW rule. The RPW rule can be illustrated by an urn model as follows. We start with an urn having 2α balls, α balls of type A and α balls of type B. We treat any entering patient by drawing a ball from the urn and replacing the ball immediately to the urn. If the patient had a successful response, we add β balls of the same type to the urn. On the other hand, if the treatment fails, we add β balls of

the opposite type to the urn. The idea is to skew the allocation proportion in favor of the treatment that is doing better.

For our design, we use RPW for the first treatment and build up the urn using the allocation and response history of the first period of the earlier patients. Information about the second period is not incorporated at this stage, but will be incorporated in a forthcoming article. We use PW for the second treatment in the sense that, if the response to the first treatment is a success, we use the successful treatment as the second treatment in the second period; if the first treatment fails, we use the second treatment as the other treatment.

3 Inferences under independence of parameters

If p_A, p_B, ϕ_A and ϕ_B are mathematically independent, the likelihood function is proportional to

$$L = p_A^{S_A} q_A^{N_A - S_A} p_B^{S_B} q_B^{N_B - S_B} \phi_A^{T_A} (1-\phi_A)^{S_A + F_B - T_A} \phi_B^{T_B} (1-\phi_B)^{S_B + F_A - T_B},$$

where S_A, S_B, T_A and T_B denote the total number of successes in the first (S_A and S_B) and second (T_A and T_B) periods by the subscripted treatment. Similarly, N_A and N_B denote the total number of patients and F_A and F_B denote the total number of failures in the first period. Consequently, the maximum likelihood estimates of p_A, p_B, ϕ_A and ϕ_B are given by

$$\hat{p}_A = \frac{S_A}{N_A}, \quad \hat{p}_B = \frac{S_B}{N_B}, \quad \hat{\phi}_A = \frac{T_A}{S_A + F_B}, \quad \hat{\phi}_B = \frac{T_B}{S_B + F_A}.$$

These estimates would have the same expression if the sampling were non-adaptive, but the distribution of these statistics depends on the particular adaptive scheme. Here

$$\hat{p}_A - p_A = \frac{\sum_{i=1}^n \delta_i (Z_i - p_A)}{N_A}, \quad \hat{p}_B - p_B = \frac{\sum_{i=1}^n \delta_i^c (Z_i - p_B)}{N_B},$$
$$\hat{\phi}_A - \phi_A = \frac{\sum_{i=1}^n \tau_i (U_i - \phi_A)}{S_A + F_B}, \quad \hat{\phi}_B - \phi_B = \frac{\sum_{i=1}^n \tau_i^c (U_i - \phi_B)}{S_B + F_A}.$$

Theorem 1 gives the asymptotic joint distribution of $\hat{p}_A$, $\hat{p}_B$, $\hat{\phi}_A$ and $\hat{\phi}_B$.

THEOREM 1. *As* $n \to \infty$,

$$\frac{1}{\sqrt{n}} \begin{pmatrix} S_A - N_A p_A \\ S_B - N_B p_B \\ T_A - (S_A + F_B)\phi_A \\ T_B - (S_B + F_A)\phi_B \end{pmatrix} \xrightarrow{d} N_4(0, \Sigma),$$

where

$$\Sigma = (q_A + q_B)^{-1} \operatorname{diag}(p_A q_A q_B, p_B q_B q_A, \phi_A \psi_A q_B, \phi_B \psi_B q_A).$$

PROOF. The proof is given in the Appendix A. □

The elements of the vector $(p_A, p_B, \phi_A, \phi_B)$ are mathematically independent. But this is not true in general. These vectors can be, for example, $(p_A, p_B, \theta p_A, \theta p_B)$ or (p_A, p_B, p_A, p_B) or $(p_A, p_B, p_A + \theta, p_B + \theta)$. To get asymptotics under such situations, we let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ or $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ be a vector of parameters belonging to an admissible set Θ. Let $\hat{\boldsymbol{\theta}}$ be a consistent solution of the likelihood equation that can be used to test any hypothesis concerning $\boldsymbol{\theta}$. Then, setting $\theta_1 = p_A$, $\theta_2 = p_B$ and $\theta_3 = \theta$, we, in place of p_A, p_B, ϕ_A and ϕ_B, model the respective success probabilities by $\pi_l = g_l(p_A, p_B, \theta)$, $l = 1, 2, 3, 4$, where g_l values are some specified functions. Then, as in section 2, we write

$$L = \pi_1^{S_A}(1-\pi_1)^{N_A - S_A}\pi_2^{S_B}(1-\pi_2)^{N_B - S_B}\pi_3^{T_A}(1-\pi_3)^{S_A + F_B - T_A}\pi_4^{T_B} \times (1-\pi_4)^{S_B + F_A - T_B}.$$

We consider the following three cases.

CASE 1. $\pi_1 = p_A$, $\pi_2 = p_B$, $\pi_3 = \theta p_A$ and $\pi_4 = \theta p_B$.

CASE 2. $\pi_1 = p_A$, $\pi_2 = p_B$, $\pi_3 = p_A$ and $\pi_4 = p_B$ (i.e. $\theta = 1$ in the earlier case).

CASE 3. $\pi_1 = p_A$, $\pi_2 = p_B$, $\pi_3 = \theta + p_A$ and $\pi_4 = \theta + p_B$.

The asymptotic distributions of the parameter vectors under each of these three models are derived in Appendix B.

4 Performance characteristics

In this section, we provide numerical simulation results for the proportion of AA, AB, BA and BB (with the corresponding standard error). Moreover, we provide the power of different tests under the models described in section 3.

For the simulation, we consider the case $\theta = 1$. We carry out the test of $H_0 : p_A = p_B$ against the one-sided alternative $H_1 : p_A > p_B$ based on the statistic

$$\hat{p}_A - \hat{p}_B = \frac{S_A + T_A}{S_A + N_A + F_B} - \frac{S_B + T_B}{S_B + N_B + F_A},$$

where we reject H_0 at level of significance α if

$$\hat{p}_A - \hat{p}_B > k_\gamma,$$

where k_γ should be the upper $100\gamma\%$ point of the null distribution of $\hat{p}_A - \hat{p}_B$. Certainly, the test depends on the common value p of $p_A = p_B$ under H_0. The values of k_γ for $\gamma = 0.05$, $n = 16$, $\alpha = \beta = 1$ are given in Table 1. Clearly, these values depend somewhat on p, which is unknown. As an approximation, we can take some prior distribution for p and take the average value of k_γ with respect to the prior distribution of p. We consider a uniform prior of p over $(0, 1)$ and the average value of p is taken as 0.1008. Taking this average value of k_γ, one can carry out the naive

Table 1. Cut-off points for the naive test procedure for different values of p.

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
k_α	0.069	0.090	0.105	0.115	0.119	0.117	0.112	0.100	0.080

Table 2. Power of the naive test procedure and the conditional test procedure for different p_A and p_B for $n=16$, $\alpha=\beta=1$, $\gamma=0.05$.

		Power of the tests	
p_A	p_B	Naive test	Conditional test
0.3	0.2	0.4953	0.1947
0.3	0.15	0.8015	0.3183
0.3	0.1	0.9635	0.4719
0.4	0.37	0.1483	0.0896
0.4	0.3	0.4928	0.1625
0.4	0.2	0.9382	0.3485
0.5	0.47	0.1625	0.0686
0.5	0.4	0.4988	0.1349
0.5	0.3	0.9269	0.2844
0.6	0.5	0.4994	0.1044
0.6	0.4	0.9214	0.2160
0.7	0.6	0.4984	0.0865
0.7	0.5	0.9270	0.1767
0.8	0.7	0.5067	0.0750

test procedure which is not exactly size γ, but should work well in practice, at least in the absence of concrete information about p, the common value under H_0. Furthermore, unless p is too small or too large, the average value will work well. Table 2 provides the power of the naive test procedure for different (p_A, p_B). It is observed that the power of the test is moderately high when the treatment difference is large. Again, the power of the procedure is quite robust on the value of α/β.

The question naturally comes up as to whether one can find any exact test procedure that under H_0, does not depend on p. Note that

$$P(S_A+T_A=j|S_A+T_A+S_B+T_B=k)=\frac{P(S_A+T_A=j,S_B+T_B=k-j)}{P(S_A+T_A+S_B+T_B=k)}$$

$$=c\times\frac{p_A^j(1-p_A)^{N_A+M_A-j}p_B^{k-j}(1-p_B)^{N_B+M_B-(k-j)}}{\sum_l p_A^l(1-p_A)^{N_A+M_A-l}p_B^{k-l}(1-p_B)^{N_B+M_B-(k-l)}},\qquad(1)$$

where c is a constant, free of the parameters, and $M_A=S_A+F_B$ and $M_B=S_B+F_A$ are the second-stage number of allocations to A and B, respectively. Under H_0, the expression in (1) is free of all the parameters. Thus, one can use the conditional test statistic $U=S_A+T_A$ given $Z=S_A+T_A+S_B+T_B$, for a distribution-free test. We then carry out a detailed simulation study to conduct a distribution-free test. In each simulation we find the value of Z and use the cut-off point (possibly randomized) obtained from the null conditional distribution of U, given that value of Z. The power of the procedure is also reported in Table 2 for $n=16$. It is observed that the power of both the test procedures increases in p_A-p_B. But, the power of the naive test procedure is much larger than that of the conditional test.

Table 3. Proportion of allocation (with standard error in parantheses) for $n=16$, $\alpha=\beta=1$.

		Proportion (SE) of:			
p_A	p_B	AA	AB	BB	BA
0.3	0.3	0.1496	0.3497	0.1506	0.3501
		(0.0356)	(0.0472)	(0.0352)	(0.0476)
0.5	0.3	0.2898	0.2891	0.1271	0.2940
		(0.0453)	(0.0448)	(0.0336)	(0.0450)
0.8	0.3	0.5988	0.1501	0.0756	0.1755
		(0.0496)	(0.0359)	(0.0264)	(0.0385)
0.4	0.4	0.1995	0.3001	0.2004	0.3000
		(0.0400)	(0.0455)	(0.0398)	(0.0457)
0.6	0.4	0.3550	0.2357	0.1635	0.2458
		(0.0481)	(0.0421)	(0.0366)	(0.0434)
0.8	0.4	0.5744	0.1433	0.1131	0.1692
		(0.0488)	(0.0350)	(0.0316)	(0.0377)
0.5	0.5	0.2501	0.2500	0.2495	0.2504
		(0.0433)	(0.0432)	(0.0436)	(0.0438)
0.7	0.5	0.4268	0.1821	0.1950	0.1961
		(0.0497)	(0.0385)	(0.0400)	(0.0398)
0.9	0.5	0.6885	0.7579	0.1179	0.1178
		(0.0465)	(0.0263)	(0.0322)	(0.0322)
0.6	0.6	0.3010	0.1988	0.2988	0.2014
		(0.0462)	(0.0405)	(0.0457)	(0.0402)
0.7	0.6	0.3920	0.1675	0.2642	0.1762
		(0.0490)	(0.0377)	(0.0446)	(0.0382)
0.8	0.6	0.5066	0.1264	0.2201	0.1469
		(0.0500)	(0.0332)	(0.0411)	(0.0356)

Considering the small sample size, the power is satisfactory. In Table 3, we provide the simulated proportion of cases of AA, AB, BA and BB (with their standard errors in parenthesis) for different (p_A, p_B) with $\alpha=\beta=1$, $p_A=\phi_A$, $p_B=\phi_B$ and $n=100$. It is observed that a considerably larger proportion of A occurs when p_A is reasonably larger than p_B. Thus, the ethical point is justified in this procedure.

Under the sequence of local alternatives

$$p_A = p + \frac{\eta}{\sqrt{n}}, \quad p_B = p,$$

the asymptotic power (AP) of the test is given by

$$P(\eta) = 1 - \Phi\left(\tau_\gamma - \frac{\eta}{v}\right),$$

where τ_γ is the upper γ percentile point of the standard normal distribution, as under $H_0: p_A = p_B$, as $n \to \infty$,

$$\sqrt{n}\frac{(\hat{p}_A - \hat{p}_B)}{v} \xrightarrow{d} N(0,1),$$

where $v = v(p, \theta)$. Following is a discussion of the three cases studied in section 4.

Case 1. It is easy to check that

$$v^2 = \frac{2pq^2(1-\theta p)}{(1-\theta p+\theta q)^2}[4(1-\theta p)+q+2\theta q^2+\theta q^3].$$

Table 4. AP for model (1) with $\eta = 1, 1.5$ and 2, for different values of p and θ.

	p				
θ	0.1	0.2	0.3	0.4	0.5
0.8	0.5106	0.3838	0.3556	0.3648	0.4033
	0.8057	0.6477	0.6053	0.6195	0.6756
	0.9552	0.8540	0.8171	0.8299	0.8760
1.0	0.5621	0.4301	0.4052	0.4247	0.4821
	0.8547	0.7118	0.6782	0.7046	0.7748
	0.9748	0.9020	0.8781	0.8971	0.9402
1.2	0.6119	0.4793	0.4620	0.4992	0.5902
	0.8942	0.7718	0.7516	0.7937	0.8779
	0.9865	0.9383	0.9270	0.9496	0.9822

Table 5. AP for model (2) for different values of p and η.

	p				
η	0.1	0.2	0.3	0.4	0.5
1	0.7611	0.5478	0.4602	0.4207	0.4090
2	0.9988	0.9706	0.9251	0.8925	0.8809
3	0.9999	0.9998	0.9985	0.9963	0.9953

Table 6. AP for model (3) with $\eta = 0.4, 0.7$ and 0.9, for different values of p and θ.

	p				
θ	0.1	0.2	0.3	0.4	0.5
0.01	0.5653	0.3990	0.3366	0.3103	0.3058
	0.9359	0.7840	0.6899	0.6435	0.6313
	0.9923	0.9306	0.8658	0.8271	0.8162
0.03	0.5192	0.3824	0.3294	0.3077	0.3045
	0.9062	0.7612	0.6776	0.6385	0.6325
	0.9847	0.9167	0.8560	0.8228	0.8173
0.05	0.4856	0.3690	0.3234	0.3058	0.3057
	0.8789	0.7416	0.6671	0.6349	0.6349
	0.9758	0.9038	0.8475	0.8195	0.8195

Table 4 gives the AP of the test for $\gamma = 0.05$, and for different values of p and θ. The three entries corresponding to each combination of (θ, p) correspond to $\eta = 1, 1.5$ and 2, respectively.

Case 2. Here $v^2 = 2pq$. Table 5 gives the values of AP of the test for different values of p and η.

Case 3. Here we obtain

$$v^2 = \frac{2(\theta+p)^2(q-\theta)^2 pq}{[pq+(\theta+p)(q-\theta)]^2}.$$

Table 6 gives the AP of the test for $\gamma = 0.05$, and for different values of p and θ. The three entries for each (θ, p) correspond to $\eta = 0.4, 0.7, 0.9$, respectively.

We get large values of AP in most of the cases. Moreover, for model (1), the AP increases in η for fixed p and θ. On the other hand, if p and η are kept fixed, the

AP increases in θ. But, for fixed θ and η, the AP is not monotonic in p. For model (2), AP is increasing in η for fixed p, and decreasing in p for fixed η. For model (3), the AP increases in η for fixed p and θ, decreases in θ if p and η are kept fixed, and decreases in p for fixed θ and η.

5 A comparative study

We now compare with the two-treatment, two-period, randomized allocation design where each of the sequences AA, AB, BA and BB has probability 0.25 for any patient. Let us write $T_{k,k'}(j_1,j_2)$ as the number of responses of the type (j_1,j_2) corresponding to the treatment combination (k,k'), $j_1,j_2=S,F;k,k'=\mathrm{A},\mathrm{B}$, where S (F) denotes a success (failure). Suppose m is the number of patients corresponding to each treatment combination (k,k') with a total sample size $4m=n$. The maximum likelihood estimates of p_A and p_B under the assumption $p_\mathrm{A}=\phi_\mathrm{A}$ and $p_\mathrm{B}=\phi_\mathrm{B}$ are

$$\begin{aligned}\hat{p}_\mathrm{A}&=n^{-1}[2T_{\mathrm{AA}}(S,S)+T_{\mathrm{AA}}(S,F)+T_{\mathrm{AA}}(F,S)+T_{\mathrm{AB}}(S,S)\\&\quad+T_{\mathrm{BA}}(S,S)+T_{\mathrm{AB}}(S,F)+T_{\mathrm{BA}}(F,S)],\\\hat{p}_\mathrm{B}&=n^{-1}[2T_{\mathrm{BB}}(S,S)+T_{\mathrm{BB}}(S,F)+T_{\mathrm{BB}}(F,S)+T_{\mathrm{BA}}(S,S)\\&\quad+T_{\mathrm{AB}}(S,S)+T_{\mathrm{BA}}(S,F)+T_{\mathrm{AB}}(F,S)].\end{aligned}$$

The exact test for $H_0: p_\mathrm{A}=p_\mathrm{B}$ can be derived as follows. Let S_A and S_B be the total number of successes with treatments A and B, respectively. Then

$$P(S_\mathrm{A}=s_\mathrm{A}|S_\mathrm{A}+S_\mathrm{B}=s)=\frac{\binom{2m}{s_\mathrm{A}}\binom{2m}{s-s_\mathrm{A}}}{\binom{n}{s}},$$

free of $p_\mathrm{A}=p_\mathrm{B}$. An exact test for H_0 can be obtained by using this conditional distribution. Such a test is given by

$$S_\mathrm{A}\geq c,$$

where c is obtained from

$$P_{H_0}(S_\mathrm{A}\geq c|S_\mathrm{A}+S_\mathrm{B}=s)\leq\gamma.$$

The power of the test depends on p_A and p_B only through $\lambda=p_\mathrm{A}(1-p_\mathrm{B})/p_\mathrm{B}(1-p_\mathrm{A})$, the odds ratio. We carried out a detailed numerical study, but are not reporting the details for the sake of brevity. We observe that the power of the randomized (non-adaptive) design for $n=16$ is 0.1989 for $(p_\mathrm{A},p_\mathrm{B})=(0.7,0.5)$, 0.6713 for $(p_\mathrm{A},p_\mathrm{B})=(0.6,0.2)$, where these figures are 0.1767 and 0.6577, respectively, for our adaptive design. From the computation, it is clear that we have a little loss in power in the adaptive procedure, but we have an ethical gain in terms of giving a larger number of patients a better treatment.

We also compare the asymptotic power. It is easy to observe that under H_0: $p_A = p_B$,

$$V(\hat{p}_A) = V(\hat{p}_B) = \frac{pq}{n}$$

and

$$\mathrm{cov}(\hat{p}_A, \hat{p}_B) = 0,$$

so that

$$V(\hat{p}_A - \hat{p}_B) = \frac{2pq}{n}$$

$p = 1 - q$ is the common value of p_A and p_B under H_0. Thus, it immediately follows that, as $n \to \infty$,

$$\sqrt{n}(\hat{p}_A - \hat{p}_B) \xrightarrow{d} N(0, 2pq).$$

Hence, under the sequence of local alternatives

$$p_A = p + \frac{\eta}{\sqrt{n}}, \quad p_B = p,$$

the asymptotic local power is given by

$$\mathrm{AP} = 1 - \Phi\left(\tau_\gamma - \frac{\eta}{\sqrt{pq/2}}\right),$$

which is the same for the adaptive procedure.

6 An example

In this section, we describe the applicability of the proposed adaptive crossover design in practice. Matthews (1989) provided some data on a three-period crossover trial of two anti-hypertensive agents. This data set was also analyzed earlier by Ebbutt (1984). The design allocated $m = 17$ patients to each of the sequences ABB, BAA, ABA, BAB, where A and B denote treatment with metroprolol and metroprolol with chlorthalidone, respectively. Thus the total number of patients is $n = 68$. The response is the systolic blood pressure at the end of each of the treatment period. Our design is a two-period design with a possibility of each of AA, AB, BA and BB. Hence we consider the data for the last two time periods of the three-period data provided by Matthews (1989). As the original data are continuous, we dichotomize them by setting a threshold of 135, i.e. a response ≤ 135 is treated as a success, and otherwise it is a failure. From the data, we have $\hat{p}_A = 0.235$ and $\hat{p}_B = 0.294$. Treating these as the true values of p_A and p_B, we carry out a simulation of 10,000 repetitions. The simulated expectation (SD) of percentage of S_{AA} is 11.275 (3.161), that of S_{AB} is 36.746 (4.799), that of S_{BB} is 15.330 (3.593), and that of S_{BA} is 36.649 (4.864). If we set the threshold as 140, the estimates become

$\hat{p}_A = 0.353$ and $\hat{p}_B = 0.515$. Consequently, our simulation provides the expectations (SD) of percentages as S_{AA}: 15.226 (3.615), S_{AB}: 28.119 (4.503), S_{BA}: 29.107 (4.574), S_{BB}: 27.548 (4.535). Thus, we observe that our proposed design could allocate ethically for the present data set, as the proportion of patients receiving BB is much larger than the proportion of patients receiving AA (here B is the better treatment as $\hat{p}_B > \hat{p}_A$). Note that the proportion of patients receiving AB or BA has one allocation to A and one allocation to B. Thus, in total, a considerably larger allocation could be to treatment A. An allocation of 2 SD less than the expected allocation for S_{AA} is still much better than equal allocation, from an ethical point of view. Thus, such an ethically appropriate adaptive crossover design can easily be used in practice.

7 Discussions

In this adaptive allocation scheme, while updating the RPW urn for the first time point, we did not use information about the PW outcome for the second time point. Without doing this, we still get an adequate benefit in terms of skewed allocation. If we did this, we might face considerable additional difficulty in terms of mathematical development. This is under study and will be reported in a separate issue. Again incorporating carry-over effect in the present set up is to be pursued in a separate communication.

The two-period RPW + PW crossover design can easily be extended for more than two periods. For example, with three periods, one can implement an RPW + PW + PW sampling scheme that is the same as the present sampling scheme up to the first two periods. For the third period, we use the same or other treatment as the second period according to whether a success or a failure occurs at the second time point. Let ξ_i be the indicator of assignment for the third period for the ith patient. Then

$$\begin{aligned}\xi_i &= \tau_i U_i + (1-\tau_i)(1-U_i)\\ &= \delta_i Z_i U_i + (1-\delta_i)(1-Z_i)U_i + (1-\delta_i Z_i)(1-U_i) - (1-\delta_i)(1-Z_i)(1-U_i).\end{aligned}$$

Subsequent development may be along similar lines. The details are under study.

Acknowledgements

The authors wish to thank the referee for careful reading and valuable suggestions which led to some improvement over an earlier version of the manuscript. The authors acknowledge some fruitful discussions with Prof. K.C. Carriere and her interest during the preparation of the manuscript. The work of the third author was supported by a research grant from the Council for Scientific and Industrial Research, India (Sanction No. 9/28(552)/2001-EMR-I).

Appendix A

PROOF OF THEOREM 1. Our objective is to obtain the asymptotic joint distribution of $\hat{p}_A$, $\hat{p}_B$, $\hat{\phi}_A$ and $\hat{\phi}_B$. For this purpose, for any fixed constants a, b, c and d, let us write

$$W_{ni} = \frac{a}{\sqrt{n}}\delta_i(Z_i - p_A) + \frac{b}{\sqrt{n}}\delta_i^c(Z_i - p_B) + \frac{c}{\sqrt{n}}\tau_i(U_i - \phi_A) + \frac{d}{\sqrt{n}}\tau_i^c(U_i - \phi_B).$$

Then we have to find the limiting distribution of

$$W_n = \sum_{i=1}^{n} W_{ni}.$$

Let $\mathcal{F}_n$ be the σ-field generated by the random variables (δ_i, Z_i), $i = 1, \cdots, n$. Then, given $\mathcal{F}_n$, the asymptotic distribution of W_n is

$$N\left(\frac{a}{\sqrt{n}}\sum_{i=1}^{n}\delta_i(Z_i - p_A) + \frac{b}{\sqrt{n}}\sum_{i=1}^{n}\delta_i^c(Z_i - p_B), c^2\phi_A\psi_A \times \frac{1}{n}\sum_{i=1}^{n}\tau_i + d^2\phi_B\psi_B\frac{1}{n}\sum_{i=1}^{n}\tau_i^c\right).$$

But, as in WEI (1979) and WEI *et al.* (1990), we have, as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n}\delta_i, \frac{1}{n}\sum_{i=1}^{n}\tau_i \xrightarrow{P} \frac{q_B}{q_A + q_B} = \rho,$$

$$\frac{1}{n}\sum_{i=1}^{n}\delta_i^c, \frac{1}{n}\sum_{i=1}^{n}\tau_i^c \xrightarrow{P} \frac{q_A}{q_A + q_B} = 1 - \rho = \rho^c,$$

and hence

$$\frac{a}{\sqrt{n}}\sum_{i=1}^{n}\delta_i(Z_i - p_A) + \frac{b}{\sqrt{n}}\sum_{i=1}^{n}\delta_i^c(Z_i - p_B) \xrightarrow{d} N\left(0, a^2p_Aq_A\rho + b^2p_Bq_B\rho^c\right).$$

Hence, as in HAJEK and SIDAK (1967, Chap. V, pp. 194–195), we get, as $n \to \infty$,

$$W_n \xrightarrow{d} N(0, a^2p_Aq_A\rho + b^2p_Bq_B\rho^c + c^2\phi_A\psi_A\rho + d^2\phi_B\psi_B\rho^c).$$

The required result then follows by using the Cramer–Wold device.

Appendix B

The maximum likelihood (ML) equations are

$$\frac{dL(\boldsymbol{\pi})}{d\boldsymbol{\theta}} = 0,$$

that is

$$\frac{S_A - N_A\hat{\pi}_1}{\hat{\pi}_1(1-\hat{\pi}_1)}\frac{\partial \pi_1}{\partial \hat{\theta}_1} + \frac{T_A - (S_A + F_B)\hat{\pi}_3}{\hat{\pi}_3(1-\hat{\pi}_3)}\frac{\partial \pi_3}{\partial \hat{\theta}_1} = 0,$$
$$\frac{S_B - N_B\hat{\pi}_2}{\hat{\pi}_2(1-\hat{\pi}_2)}\frac{\partial \pi_2}{\partial \hat{\theta}_2} + \frac{T_B - (S_B + F_A)\hat{\pi}_4}{\hat{\pi}_4(1-\hat{\pi}_4)}\frac{\partial \pi_4}{\partial \hat{\theta}_2} = 0,$$
$$\frac{T_A - (S_A + F_B)\hat{\pi}_3}{\hat{\pi}_3(1-\hat{\pi}_3)}\frac{\partial \pi_3}{\partial \hat{\theta}_3} + \frac{T_B - (S_B + F_A)\hat{\pi}_4}{\hat{\pi}_4(1-\hat{\pi}_4)}\frac{\partial \pi_4}{\partial \hat{\theta}_3} = 0,$$

where $\hat{\pi}_l = g_l(\hat{\boldsymbol{\theta}})$. Expanding $(\hat{\pi}_l - \pi_l)$ for $l = 1, 2, 3, 4$ at θ, we get

$$\hat{\pi}_l - \pi_l = \sum_{j=1}^{3}(\hat{\theta}_j - \theta_j)\frac{\partial \pi_i}{\partial \theta_j'}, \quad \theta_j' \in (\hat{\theta}_j, \theta_j).$$

Also note that

$$\sum_j \frac{1}{\hat{\pi}_j(1-\hat{\pi}_j)}\frac{\partial \pi_j}{\partial \hat{\theta}_r}\frac{\partial \pi_j}{\partial \hat{\theta}_s} \longrightarrow \sum_j \frac{1}{\pi_j(1-\pi_j)}\frac{\partial \pi_j}{\partial \theta_r}\frac{\partial \pi_j}{\partial \theta_s}, \quad \text{as } \hat{\pi}_l \to \pi_l \text{ and } \frac{\partial \pi_l}{\partial \hat{\theta}_r} \to \frac{\partial \pi_l}{\partial \theta_r}.$$

Thus we get

$$\frac{1}{\sqrt{n}}\frac{\mathrm{d}L(\boldsymbol{\pi})}{\mathrm{d}\boldsymbol{\theta}} = A_n\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

no for some A_n. Now we discuss three particular model assumptions given earlier.

Case 1. If $\pi_1 = p_A$, $\pi_2 = p_B$, $\pi_3 = \theta p_A$ and $\pi_4 = \theta p_B$, then

$$A_n \xrightarrow{p} A,$$

where

$$A = (q_A + q_B)^{-1}\begin{pmatrix} \frac{q_B}{p_A q_A} + \frac{\theta q_B}{p_A(1-\theta p_A)} & 0 & \frac{q_B}{1-\theta p_A} \\ 0 & \frac{q_A}{p_B q_B} + \frac{\theta q_A}{p_B(1-\theta p_B)} & \frac{q_A}{1-\theta p_B} \\ \frac{q_B}{1-\theta p_A} & \frac{q_A}{1-\theta p_B} & \frac{p_A q_B}{\theta(1-\theta p_A)} + \frac{q_A p_B}{\theta(1-\theta p_B)} \end{pmatrix},$$

since

$$\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}\boldsymbol{\theta}} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & \theta_3 \\ 1 & 1 & \theta_1 & \theta_2 \end{pmatrix},$$

and it can be easily shown that

$$\frac{S_A}{n} \xrightarrow{P} \frac{p_A q_B}{q_A + q_B}, \quad \frac{S_B}{n} \xrightarrow{P} \frac{p_B q_A}{q_A + q_B}.$$

Let us denote

$$A^{-1} = (q_A + q_B)((a^{ij})).$$

Consequently, we get

$$\frac{1}{\sqrt{n}}\begin{pmatrix} \frac{1}{\pi_1(1-\pi_1)}\frac{\partial\pi_1}{\partial\theta_1} & 0 & \frac{1}{\pi_3(1-\pi_3)}\frac{\partial\pi_3}{\partial\theta_1} & 0 \\ 0 & \frac{1}{\pi_2(1-\pi_2)}\frac{\partial\pi_2}{\partial\theta_2} & 0 & \frac{1}{\pi_4(1-\pi_4)}\frac{\partial\pi_4}{\partial\theta_2} \\ 0 & 0 & \frac{1}{\pi_3(1-\pi_3)}\frac{\partial\pi_3}{\partial\theta_3} & \frac{1}{\pi_4(1-\pi_4)}\frac{\partial\pi_4}{\partial\theta_4} \end{pmatrix}\begin{pmatrix} S_A - N_A\pi_1 \\ S_B - N_B\pi_2 \\ T_A - (S_A+F_B)\pi_3 \\ T_B - (S_B+F_A)\pi_4 \end{pmatrix}$$
$$\stackrel{a}{=} A\sqrt{n}\begin{pmatrix} \hat{\theta}_1-\theta_1 \\ \hat{\theta}_2-\theta_2 \\ \hat{\theta}_3-\theta_3 \end{pmatrix}.$$

This implies that

$$\sqrt{n}\begin{pmatrix} \hat{\theta}_1-\theta_1 \\ \hat{\theta}_2-\theta_2 \\ \hat{\theta}_3-\theta_3 \end{pmatrix} \stackrel{a}{=} L\frac{1}{\sqrt{n}}\begin{pmatrix} \hat{\pi}_1-\pi_1 \\ \hat{\pi}_2-\pi_2 \\ \hat{\pi}_3-\pi_3 \\ \hat{\pi}_4-\pi_4 \end{pmatrix},$$

where

$$L=(q_A+q_B)\begin{pmatrix} \frac{a^{11}}{p_Aq_A} & \frac{a^{12}}{p_Bq_B} & \frac{a^{11}}{p_A(1-\theta p_A)}+\frac{a^{13}}{\theta(1-\theta p_A)} & \frac{a^{13}}{\theta(1-\theta p_B)}+\frac{a^{12}}{p_B(1-\theta p_B)} \\ \frac{a^{12}}{p_Aq_A} & \frac{a^{22}}{p_Bq_B} & \frac{a^{23}}{\theta(1-\theta p_A)}+\frac{a^{12}}{p_A(1-\theta p_A)} & \frac{a^{22}}{p_B(1-\theta p_B)}+\frac{a^{23}}{\theta(1-\theta p_B)} \\ \frac{a^{13}}{p_Aq_A} & \frac{a^{23}}{p_Bq_B} & \frac{a^{13}}{p_A(1-\theta p_A)}+\frac{a^{33}}{\theta(1-\theta p_A)} & \frac{a^{23}}{p_B(1-\theta p_B)}+\frac{a^{33}}{\theta(1-\theta p_B)} \end{pmatrix}.$$

Hence, as $n\to\infty$,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \xrightarrow{d} N_3(0, L\Sigma^{-1}L^T).$$

Case 2. If we take $\pi_1=p_A$, $\pi_2=p_B$, $\pi_3=p_A$ and $\pi_4=p_B$, and the vector of parameter $\boldsymbol{\theta}=(p_A,p_B)$, then

$$\frac{d\boldsymbol{\pi}}{d\boldsymbol{\theta}}=\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Consequently, we get

$$\frac{1}{\sqrt{n}}\begin{pmatrix} \frac{S_A-N_A\pi_1}{\pi_1(1-\pi_1)} & \frac{T_A-(S_A+F_B)\pi_3}{\pi_3(1-\pi_3)} \\ \frac{S_B-N_B\pi_2}{\pi_2(1-\pi_2)} & \frac{T_B-(S_B+F_A)\pi_4}{\pi_4(1-\pi_4)} \end{pmatrix}=\left(\frac{1}{n}A_n\right)\sqrt{n}\begin{pmatrix} \hat{\theta}_1-\theta_1 \\ \hat{\theta}_2-\theta_2 \end{pmatrix},$$

where

$$A_n=\begin{pmatrix} \frac{N_A}{\pi_1(1-\pi_1)}+\frac{S_A+F_B}{\pi_3(1-\pi_3)} & 0 \\ 0 & \frac{N_B}{\pi_2(1-\pi_2)}+\frac{S_B+F_A}{\pi_4(1-\pi_4)} \end{pmatrix}$$
$$\xrightarrow{P}(q_A+q_B)^{-1}\begin{pmatrix} \left(\frac{1}{q_A}+\frac{\theta}{1-\theta p_A}\right)\frac{q_B}{p_A} & 0 \\ 0 & \left(\frac{1}{q_B}+\frac{\theta}{1-\theta p_B}\right)\frac{q_A}{p_B} \end{pmatrix}=A.$$

Noting that

$$A^{-1}=(q_A+q_B)\begin{pmatrix} a^{11} & 0 \\ 0 & a^{22} \end{pmatrix},$$

with

$$a^{11}=\frac{q_A p_A}{2q_B},\quad a^{22}=\frac{q_B p_B}{2q_A},$$

we have

$$L=(q_A+q_B)\begin{pmatrix}\frac{a^{11}}{p_A q_A} & 0 & \frac{a^{11}}{p_A q_A} & 0\\ 0 & \frac{a^{22}}{p_B q_B} & 0 & \frac{a^{22}}{p_B q_B}\end{pmatrix}.$$

As

$$\Sigma=(q_A+q_B)^{-1}\mathrm{diag}(p_A q_A q_B, p_B q_B q_A, p_A q_A q_B, p_B q_B q_A),$$

we get

$$\sqrt{n}\begin{pmatrix}\hat{p}_A-p_A\\ \hat{p}_B-p_B\end{pmatrix}\xrightarrow{d}N_2(0, L\Sigma L^{\mathrm{T}}),$$

where

$$L\Sigma L^{\mathrm{T}}=\begin{pmatrix}\frac{p_A q_A(q_A+q_B)}{2q_B} & 0\\ 0 & \frac{p_B q_B(q_A+q_B)}{2q_A}\end{pmatrix}.$$

Case 3. If we take $\pi_1=p_A$, $\pi_2=p_B$, $\pi_3=\theta+p_A$ and $\pi_4=\theta+p_B$, we immediately get

$$\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}\boldsymbol{\theta}}=\begin{pmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\end{pmatrix}.$$

In that case, we have

$$A=(q_A+q_B)^{-1}\times\begin{pmatrix}\frac{q_B}{p_A q_A}+\frac{q_B}{(\theta+p_A)(q_A-\theta)} & 0 & \frac{q_B}{(\theta+p_A)(q_A-\theta)}\\ 0 & \frac{q_A}{p_B q_B}+\frac{q_A}{(\theta+p_B)(q_B-\theta)} & \frac{q_A}{(\theta+p_B)(q_B-\theta)}\\ \frac{q_B}{(\theta+p_A)(q_A-\theta)} & \frac{q_A}{(\theta+p_B)(q_B-\theta)} & \frac{q_B}{(\theta+p_A)(q_A-\theta)}+\frac{q_A}{(\theta+p_B)(q_B-\theta)}\end{pmatrix}.$$

Consequently, the limiting distribution routinely follows.

References

Atkinson, A. C. (1982), Optimal biased coin designs for sequential clinical trials with prognostic factors, *Biometrika* **69**, 61–67.

Atkinson, A. C. and A. Biswas (2005), Bayesian adaptive biased-coin designs for clinical trials with normal responses, *Biometrics* **61**, 118–125.

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

Biswas, A. (2001), Adaptive designs in phase III clinical trials: controversies and progress, *Statistical Methods in Medical Research* **10**, 353–364.

Biswas, A. and A. Dewanji (2004), A randomized longitudinal play-the-winner design for repeated binary data, *Australian and New Zealand Journal of Statistics*, **46**, 675–684.

Brown, B. W. (1980), The crossover experiment for clinical trial, *Biometrics* **36**, 69–79.

Carriere, K. C. and G. C. Reinsel (1992), Investigation of dual-balanced crossover designs for two treatments, *Biometrics* **48**, 1157–1164.

Cheng, C. S. and C. F. Wu (1980), Balanced repeated measurements designs, *The Annals of Statistics* **8**, 1272–1283.

Clayton, D. G. (1982), Ethically optimized designs, *British Journal of Clinical Pharmacology* **13**, 479–480.

Ebbutt, A. F. (1984), Three-period crossover designs for two treatments, *Biometrics* **40**, 219–224.

Hadayat, A. and K. Afsarinejad (1978), Repeated measurements design II, *The Annals of Statistics* **6**, 619–628.

Hajek, J. and Z. Sidak (1967), *Theory of rank tests*, Academic Press, New York.

Iglewicz, B. (1983), Alternative designs: sequential, multi-stage, decision theory and adaptive designs, in: M. E. Buyse, J. Staquet and R. J. Sylvester (eds.), *Cancer clinical trials: methods and practice*, Oxford University Press, Oxford, 312–334.

Kunert, J. (1991), Crossover designs for two treatments and correlated errors, *Biometrika* **78**, 315–324.

Kushner, H. B. (1997), Optimality and efficiency of two-treatment repeated measurements design, *Biometrika* **84**, 455–468.

Laska, E. M. and M. Meisner (1985), A variational approach to optimal two-treatment crossover designs: application to carrover effect models, *Journal of the American Statistical Association* **80**, 704–710.

Matthews, J. N. S. (1987), Optimal crossover designs for the comparison of two treatments in the presence of carryover effects and autocorrelated errors, *Biometrika* **74**, 311–320.

Matthews, J. N. S. (1989), Estimating dispersion parameters in the analysis of data from crossover trials, *Biometrika* **76**, 239–244.

Palmer, C. R. and W. F. Rosenberger (1999), Ethics and practice: alternative designs for phase III randomized clinical trials, *Controlled Clinical Trials* **20**, 172–186.

Rosenberger, W. F. (1996), New directions in adaptive designs, *Statistical Science* **11**, 137–149.

Rosenberger, W. F., N. Stallard, A. Ivanova, C. N. Harper and M. L. Ricks (2001), Optimal adaptive designs for binary response trials, *Biometrics* **57**, 909–913.

Rout, C. C., D. A. Rocke, J. Levin, E. Gouws and D. Reddy (1993), A reevaluation of the role of crystalloid preload in the prevention of hypotension associated with spinal anesthesia for elective cesarean section, *Anesthesiology* **79**, 262–269.

Tamura, R. N., D. E. Faries, J. S. Andersen and J. H. Heiligenstein (1994), A case study of an adaptive clinical trials in the treatment of out-patients with depressive disorder, *Journal of the American Statistical Association* **89**, 768–776.

Wei, L. J. (1979), The generalized Polya's urn for sequential medical trials, *The Annals of Statistics* **7**, 291–296.

Wei, L. J. and S. Durham (1978), The randomized play-the-winner rule in medical trials, *Journal of the American Statistical Association* **73**, 838–843.

Wei, L. J., R. T. Smythe, D. Y. Lin and T. S. Park (1990), Statistical inference with data-dependent treatment allocation rule, *Journal of the American Statistical Association* **85**, 156–162.

Zelen, M. (1969), Play-the-winner rule and the controlled clinical trial, *Journal of the American Statistical Association* **64**, 131–146.

Received: June 2005. Revised: September 2006.