Sankhyā : The Indian Journal of Statistics
2017, Volume 79-B, Part 1, pp. 156-169

Optimal Test Statistics for Minimising not Cured Proportion in Adaptive Clinical Trial

Optimal Test Statistics for Minimising not Cured Proportion

Anupam Kundu, Nabaneet Das, Sayantan Chakraborty and Subir Kumar Bhandari
Indian Statistical Institute, Kolkata, India

Abstract

In last several decades adaptive sequential binary design has been used with a goal to increase performance in estimation, testing of parameters and to reduce expected number of non-cured patients in the context of clinical trials.The procedures have been studied theoretically and also using simulation techniques in many papers. As for example play the winner rule, randomised play the winner rule, adaptive randomised play the winner rule have been studied extensively. Rosenberger et al. (*Biometrics* **57**, 3, 909–913, 2001) considered different types of difference function of $\hat{p_A}$ and $\hat{p_B}$ as the test statistic in adaptive sequential design for the purpose of better inference along with decreasing number of non-cured patients. In this paper we considered how to choose the optimal function to achieve the goals with better performance. Using extensive simulation studies we have supported our claim and have shown that our methods perform better than existing methods. Also in the methods given by us we have a choice on the proportion of non-cured patients which we can vary with the inferential goal in mind. This is a new approach in adaptive sequential design.

AMS (2000) *subject classification.* Primary 62L05; Secondary 62F03.
Keywords and phrases. Adaptive sequential design, Not-cured proportion, Optimal test, Clinical trial

1 Introduction

In case of a clinical trial problem with binary responses the trade-off between the error rate and proportion of treatment failures is always a big question. For optimal allocation one option is Neyman Allocation, which allocates in proportional to standard deviation of the parameter estimates, as it minimises the variance. Some papers claim that it maximises the power but ignores the information that becomes available about the effectiveness of the drugs as the trial progresses.

Response-Adaptive designs are proposed to overcome the ethical problem of assigning worse treatment to less number of subjects and hence it depends on the outcome of the previous subject. Optimal allocations are affected by the statistic used to perform the hypothesis testing. In many cases to study this design, urn models are widely used. For example: ARPW (Adaptive Randomised Play the Winner), RPW (Wei and Durham, 1978) (Randomised Play the Winner), which were modified from Zelen (1969)'s play the winner rule.Also we can mention ARPW with prognostic factors (Biswas and Bandyopadhyay, 1999). These rules are not generated based on any optimality criterion but they allocate according to the relative risk to the treatments.

Sometimes it is proposed that 50 percent allocation is useful as it maximises the power and increases the relative efficiency (Azriel, Mandel & Rinott, 2012) which from an inferential point of view makes the adaptive model questionable. Therefore in between the ethical ground and the necessity of the adaptive model, it is natural to suspect that use of some optimal statistic may improve the inferential goal which is indeed the case.

In most of the work they considered a fixed test statistic $\hat{p_A} - \hat{p_B}$ and searched for optimal allocation (Rosenberger et al., 2001). For the testing goal, here we redefined the goal and considered a different test statistic on the basis of that. In this process we have also tried to reduce the expected number of non-cured patients. Thus our paper considers to optimise two goals i.e. inferential goal and also the goal to reduce the expected number of non-cured patients. We have discussed a class of procedures by redefining the null hypothesis and using corresponding test statistic with the suitable allocation rule. We tried to choose the best procedure among a rich class of procedures with the above two goals in mind. In this context we also find that our best procedure always beats the existing procedures.

2 Prerequisites and Main Results

Let $X_i, i = 1, 2, \ldots$ are random sample from $Ber(p_A)$ and $Y_i, i = 1, 2, \ldots$ are from $Ber(p_B)$. We have two face values p_1 and p_2 where $(p_1 \geq p_2)$. We are to test the simple hypothesis $H_0 : (p_A, p_B) = (p_1, p_2)$ vs $H_1 : (p_A, p_B) = (p_2, p_1)$. For this purpose we proceed sequentially. Let at stage n, we have drawn a total of n number of samples from the two populations with a number n_A from the first population and a number n_B from the second population such that $n_A + n_B = n$. Let $R = \frac{n_A}{n_B}$, that is the ratio of the respective number of allocation. At stage n on the basis of the observations $X_1, X_2, \ldots X_{n_A}$ and $Y_1, Y_2, \ldots Y_{n_B}$ we will determine from which population we shall draw the next sample at the $(n + 1) - th$ stage. Then we shall

have $(n+1)$ samples and on the basis of that we shall decide from which population we shall draw the next sample. We repeat the process until a total of N samples are exhausted. This is called adaptive sequential sampling design.

At each stage we shall get the function, on the basis of which allocation rule will be determined. Here we consider

$$S_T = T(\hat{p_A}) - T(\hat{p_B})$$

as the statistics, with the help of which we will specify the allocation rule, where $\hat{p_A}$ and $\hat{p_B}$ are the success proportion in the two samples and T is some increasing function with continuous second derivative in the interval $(0,1)$. As we have seen in case of a normal distribution, inference does not improve for using any other function T instead of using the sufficient statistics. Here in this case of adaptive binomial design, there is a chance to improve the inference sufficiently by choosing suitable function T, so as to consider S_T for the purpose of better inference along with decreasing expected number of non-cured patients.

In the previous works some particular test statistic were used and no adequate justification were provided for using those statistics (Rosenberger et al., 2001) for example $\hat{p_A} - \hat{p_B}$, $\frac{\hat{p_A}}{\hat{p_B}}$, $log\frac{\hat{p_A}}{\hat{p_B}}$ etc. as the test statistic and considered optimal limiting allocation that will minimise expected proportion of non-cured patients keeping the asymptotic variance of the function fixed. Here in this paper we have tried to generalise that fact.

Now by delta method:

$$\sqrt{n_A}(T(\hat{p}_A) - T(p_A)) \to N(0, p_A q_A (T'(p_A)))$$

Therefore asymptotic variance of S_T (which is assumed to be a constant K) becomes

$$Avar(S_T) = \frac{1+R}{nR}(p_A q_A (T'(p_A))^2 + R p_B q_B (T'(p_B))^2) \tag{2.1}$$

According to our assumption of constant asymptotic variance we get n as a function of K, R, p_A and p_B. Now

$$R^* = \underset{R}{\operatorname{argmin}}(n_A q_A + n_B q_B)$$

Differentiating this we get the optimal allocation as

$$R^* = \sqrt{\frac{p_A (T'(p_A))^2}{p_B (T'(p_B))^2}} \tag{2.2}$$

This is our optimal allocation rule for any arbitrary T. Now using this optimal rule we calculated the proportion of not cured persons (denoted by p_{NC}) which turns out to be the weighted average of failure probabilities with weights as the proportion of the optimal allocation.

$$p_{NC} = \frac{T'(p_A)q_A\sqrt{p_A} + T'(p_B)q_B\sqrt{p_B}}{T'(p_A)\sqrt{p_A} + T'(p_B)\sqrt{p_B}} \tag{2.3}$$

Now let, at the optimal allocation,

$$Avar(S_T) = K(n)$$

and again we can see that $K(n) \to 0$ as $n \to \infty$ as $K(n) = O(\frac{1}{n})$
As it turns out that the P($type - I$ error) and P($type - II$ error) are same, taking the symmetric critical region, we only calculate the P($type - I$ error) given by

$$PICS = P([T(\hat{p_A}) - T(\hat{p_B})] < 0 | [T(p_A) - T(p_B)] = a) \qquad a > 0.$$

Now to get an idea of a measure of efficiency, we tried to calculate the error rate of this test. To calculate the efficiency we have taken the limit of the n-th root of P($type - I$ error).

$$\lim_{n\to\infty} P([T(\hat{p_A}) - T(\hat{p_B})] < 0 | [T(p_A) - T(p_B)] = a)^{\frac{1}{n}}$$

gives us the error rate of this test while using T.

We calculated this quantity which is

$$B(a,t) = \max_{u\in(-\infty,0]} \left[exp\left(-\frac{(u-a)^2}{2t} \right) \right] \tag{2.4}$$

where $t = \lim_{n\to\infty} t_n$ and t_n's are defined as $K(n) = \frac{t_n}{n}$ and hence $O(1)$. We have

$$log(PICS) = nlog(B(a,t)) + o(n)$$

$$log(PICS) = ASN \times log(B(a,t)) + o(ASN) \tag{2.5}$$

Now to test the efficiency of two competitive procedures, we are using the concept of Bahadur efficiency, which is the limit of the ratio of average sample numbers (ASN) for same PCS i.e. proportion of correct selection. Here we are using 0 as a cut off for calculating the error and $P[Error] = PICS$ i.e. proportion of incorrect selection. In this adaptive sequential Bernoulli

trials, symmetric tests (i.e. permutation invariant tests) are used, where corresponding estimates are $\hat{p_A}$ and $\hat{p_B}$. Using T (function) we consider the test statistic S_T and corresponding permutation invariant critical region $\{S_T < 0\}$. For clarity we consider the following:

$$H_0 : T(p_A) - T(p_B) = a$$

$$H_1 : T(p_A) - T(p_B) = -a$$

For any two functions T_1 and T_2 we have two competing procedures : procedure 1 and procedure 2 respectively. Here $P(type - I \text{ error})$=$P(type - II$ error)=PICS (from the permutation invariance). Using expansion given by Eq. 2.5 and keeping same PICS (with two values of ASN's : ASN_1 and ASN_2 respectively) we have

$$\begin{aligned} log(PICS) &= ASN_1 \times log(B_1(a,t)) + o(ASN_1) \\ &= ASN_2 \times log(B_2(a,t)) + o(ASN_2) \end{aligned}$$

Hence efficiency of procedure 1 with respect to procedure 2 becomes

$$\lim_{n\to\infty} \frac{ASN_2}{ASN_1} = \frac{-log(B_1(a,t))}{-log(B_2(a,t))}$$

The above equation follows from the concept of Bahadur efficiency. Therefore to verify the performance of some procedure with function T we have to consider the values of $\{-log(B(a,t))\}$.

Using the optimal allocation rule we have calculated the value of t which is $C^2 p_{NC}$ where $C = (T'(p_A)\sqrt{p_A} + T'(p_B)\sqrt{p_B})$. From these results it follows trivially that our error rate is $e^{\left(-\frac{a^2}{2C^2 p_{NC}}\right)}$. In order to minimise this error rate we maximise the fraction $\frac{a^2}{C^2}$, because of the reasons stated in the following paragraph.

Here $-log(B(a,t)) = \frac{a^2}{C^2 p_{NC}}$ and the bigger value of this quantity will indicate better performance of the procedure. Under condition $p_{NC} \leq c_0$ from a restricted class of functions T (to get a procedure depending on T) we find the best procedure i.e. best $\{\log(B(a,t))\}$. Hence with restriction on p_{NC} we maximise $\frac{a^2}{C^2}$ for better performance of the procedure as we have noted that with the condition $p_{NC} \leq c_0$, actually $p_{NC} = c_0$ gives us the minimum error rate. So even if p_{NC} involves T, we are maximising the fraction because here our goal is constrained optimisation with restriction on p_{NC}.

3 Two Possible Methods

Though the following two cases lead to the same conclusion, we mainly focus on case 2 because case 1 does not provide some good form of result.

3.1. Method 1. Now we consider the case where $p_B = p_A - \delta$ and in this case

$$\left(\frac{a^2}{C^2}\right) \approx \left(\frac{\delta}{2\sqrt{p_A}} + \frac{\delta^2}{2\sqrt{p_A}}\left(\frac{T''(p_A)}{2T'(p_A)} + \frac{1}{4p_A} - \delta\frac{T''(p_A)}{4p_AT'(p_A)}\right)\right) \tag{3.1}$$

where we assumed that

$$\left|\delta\left(\frac{T''(p_A)}{2T'(p_A)} + \frac{1}{4p_A} - \delta\frac{T''(p_A)}{4p_AT'(p_A)}\right)\right| < 1$$

Therefore if δ is so small that we can neglect the terms which are $O(\delta^2)$ then we can see that T has no effect in error rate at all. For a slight difference in two treatments any nice increasing differentiable function T will provide the same result. But T plays an important role when it is believed that one treatment is significantly better than the other and of course by significance we mean that δ is not too small.

If we consider the effect of T when δ is significant then we need to use a function T which has a large second derivative and small first derivative at p_A to minimise the error.Small probability of success in the first treatment actually minimise the error rate.

Here we also calculated the value of p_{NC} which is a very complicated expression involving δ, δ^2. We came to this expression using Taylor series expansion around p_A. Here we have observed the fact that if δ is small enough then proportion of not cured will be approximately equal to q_A. Even if δ is significant and we do not neglect the terms involving δ, δ^2, we fail to manipulate the algebraic expression for p_{NC} obtained by using general form of the function T and its derivatives. But in the second method using particular forms of T we have been able to settle the question of minimisation of p_{NC}.

3.2. Method 2. In this case we have considered the case when $p_A = kp_B$. In this case also we have found out the error rates, proportion of not cured and optimal allocation. The Expressions are the following:

$$\frac{a}{C} = \frac{T(kp_B) - T(p_B)}{\sqrt{p_B}(\sqrt{k}T'(kp_B) + T'(p_B))} \tag{3.2}$$

$$R^* = \sqrt{k}\frac{T'(kp_B)}{T'(p_B)} \tag{3.3}$$

$$p_{NC} = \frac{T'(kp_B)q_A\sqrt{k} + T'(p_B)q_B}{T'(kp_B)\sqrt{k} + T'(p_B)} \tag{3.4}$$

In this method we tried to do the same thing as method 1 because method 1 produces complicated expressions by the use of Taylor series theorem which can be made simple in this method.

4 Two Special Examples of Method 2

4.1. Example 1. We have taken $T(p) = log(p)$ and calculated corresponding $\frac{a}{C}$, R^* and p_{NC} as a function of k and p_B.

$$\begin{aligned} \frac{a}{C} &= \sqrt{p_B}\frac{\sqrt{k}log(k)}{\sqrt{k}+1} \\ R^* &= \frac{1}{\sqrt{k}} \\ p_{NC} &= (1-\sqrt{k}p_B) \end{aligned}$$

In this case p_{NC} generates a constraint i.e. p_B has to be less than or equal to $\frac{1}{\sqrt{k}}$ which is always satisfied because $\sqrt{p_A}, \sqrt{p_B} \in (0,1)$. So $\sqrt{kp_B}\sqrt{p_B} = \sqrt{k}p_B \in (0,1)$.

4.2. Example 2. Here we have taken the example $T(p) = p^\alpha$. Here we have also focused those above mentioned three quantities.

$$\begin{aligned} \frac{a}{C} &= \frac{\sqrt{p_B}(k^\alpha - 1)}{\alpha(k^{\alpha-0.5}+1)} \\ R^* &= k^{(\alpha-0.5)} \\ p_{NC} &= \left[1 - p_B\left(\frac{k^{\alpha+0.5}+1}{k^{\alpha-0.5}+1}\right)\right] \end{aligned}$$

As $\alpha \to 0$

- $\frac{a}{C} \to \frac{\sqrt{kp_B}log(k)}{\sqrt{k}+1}$
- $R^* \to \frac{1}{\sqrt{k}}$
- $p_{NC} \to (1-\sqrt{k}p_B)$

It turns out that in case of small α these quantities behave like in the case for $log(p)$. If we compare the two cases where $\alpha \to 0$ and $\alpha = 1$ the first one performs better in minimising the error but later performs better in minimising the proportion of not cured.

4.2.1. Remarks. **1) Sign of** α p^α is no longer increasing for $\alpha < 0$ however if we still use $T(p) = p^\alpha$ with $\alpha < 0$, p_{NC} is close to the highest value i.e. $(1-p_B)$ for all k at the negative region of the values of α. It is not preferable. Figure 1 shows the p_{NC} curves (as α varies) for various values of p_B and k.

If we look $\frac{a}{C}$ as a function of α it becomes evident that it decreases on both sides of 0. Here negative as well as positive α will maximise $\frac{a}{C}$ near 0. Positive α will give higher value than the corresponding reflected point with respect to the origin on the negative side which supports the use of positive α in this case. We focus our attention on positive α only. Figure 2 shows $\frac{a}{C}$ as a function of α.

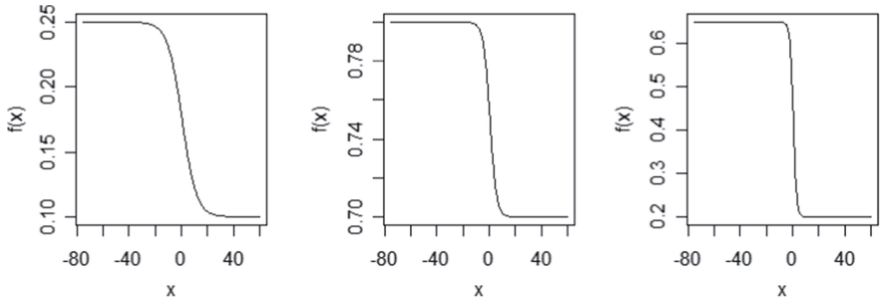
2) To find Optimal α

Here we consider the best α which gives maximum value of $\frac{a}{C}$ under the constraint $p_{NC} \leq c_0$, where c_0 is a positive real number in the interval $[(1-p_A), (1-p_B)]$. It reduces to a simple optimisation problem that will maximise the fraction $\frac{a}{C}$ and simultaneously keep p_{NC} within a predetermined tolerable value c_0(say). The fact $\frac{a}{C}$ and p_{NC} both decreases when α increases implies that $p_{NC} = c_0$ will give us the optimal $\alpha = \hat{\alpha}(k, c_0, p_B)$.

3) Expression of α

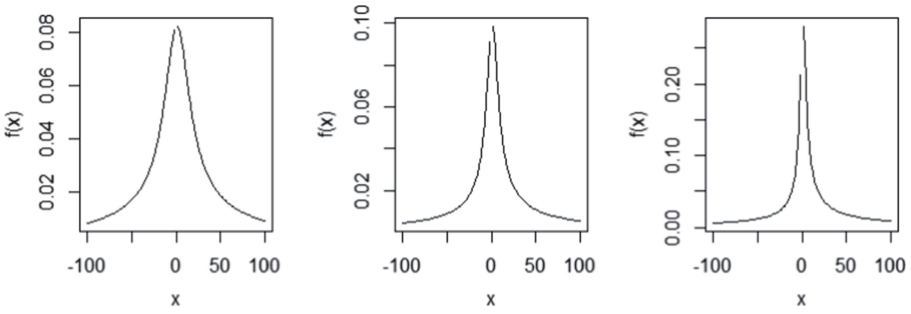
Considering Example 2 in Method 2 under the constraint $p_{NC} \leq c_0$ we get

$$\alpha = log_k \left(\frac{\sqrt{k}(1 - c_0 - p_B)}{(p_A + c_0 - 1)} \right) \tag{4.1}$$



(a) $p_B = 0.75$ and $k = 1.2$ (b) $p_B = 0.2$ and $k = 1.5$ (c) $p_B = 0.35$ and $k = 2.286$

Figure 1: Plots of curves of p_{NC} (as α varies) for different p_B and k



(a) $p_B = 0.75$ and $k = 1.2$ (b) $p_B = 0.2$ and $k = 1.5$ (c) $p_B = 0.35$ and $k = 2.286$

Figure 2: Plots of curves of $\frac{a}{C}$ for different p_B and k

For a given value of p_A and p_B, k is determined and $(1 - c_0)$ is a convex combination of p_A and p_B from which we determine $\alpha = \hat{\alpha}(k, c_0, p_B)$. At each step of the adaptive sequential method we use this alpha to determine T to get the allocation proportion of the next step as the limiting allocation proportion. The expected proportion of cured patients i.e. $(1 - c_0)$ is a convex combination of p_A and p_B and we can choose the fixed weights of the convex combination.

Note that in the application of the methodology, p_A and p_B will not be known. Hence we will not know the value of k and c_0 and also α. At each step of the adaptive sequential design, we consider $\hat{p_A}$ and $\hat{p_B}$ in place of p_A and p_B respectively to get the estimated values of k, c_0 and α. Then estimated values are used to determine the estimated limiting optimal allocation proportion. It is used as the allocation probability of the next step allocation.

5. Simulation

For simulation purpose we have only considered the second method. As we do not have real data, we have simulated examples and showed that our optimal statistics is actually better than the existing three statistics namely : $T(p) = p$, $T(p) = log(p)$ and Neymann allocation method, which allocates in the ratio of variances.

In the simulation procedure at first we needed an idea of $\hat{p_A}$ and $\hat{p_B}$. Given the number of total available subjects we have allocated first 5 percent of them randomly to treatment A or treatment B with the success probability as given in the procedure. As in some of the cases, the allocation rule may be undefined or alpha cannot be calculated, we used equal allocation in treatments for those initial steps. Then using that $\hat{p_A}$, $\hat{p_B}$ and α (calculated

from c_0, p_A, p_B) we determined the allocation rule for the rest 95 % subjects. In each step $\hat{p_A}$ and $\hat{p_B}$ gets updated, providing a new allocation rule based on $\hat{p_A}, \hat{p_B}$ and α (obtained from success probabilities) throughout the process until the total number of available subjects gets exhausted. This experiment is repeated 10000 times to get a good estimate of p_A and p_B (which is $\hat{p_A}$ and $\hat{p_B}$) and the number of times our hypothesis turns out to be correct i.e. $\hat{p_A}$ is greater than $\hat{p_B}$.

In the above process we have calculated alpha from the success probabilities of the procedure which is not a very good idea in practice because in the first place we have a very little knowledge of p_A and p_B. Instead if we use $\hat{p_A}$ and $\hat{p_B}$ to evaluate $\hat{\alpha}$ in each step of the experiment and use the combination $(\hat{p_A}, \hat{p_B}, \hat{\alpha})$ while determining the allocation rule adaptively, it will reflect the real life situation. In the tables below we have simulated in both ways described above and compared them with the existing methods (Tables 1 and 2).

In the existing procedures while choosing the statistic only error rates were taken into account and there was no way to control the proportion of not cured in the experiment which is a very important aspect to observe. In our case we have devised a procedure to control the proportion of not cured along with optimising the error rate. "70–30 Constant α" is the case where $(1 - c_0)$ is selected in the ratio $7 : 3$ between p_A and p_B and remains fixed throughout the experiment keeping α fixed. Similarly "60-40 Adaptive" denotes that we have chosen $(1 - \hat{c_0})$ between $\hat{p_A}$ and $\hat{p_B}$ in the ratio $3 : 2$. Here $(1 - \hat{c_0})$ is chosen closer to larger among $\hat{p_A}$ and $\hat{p_B}$ because it will make $\hat{\alpha}$ positive like the constant α set up. Similarly "85-15 Adaptive" and

Table 1: Simulated proportion of not cured for existing statistics

			$\alpha = 1$		$log(p)$		Neymann Allocation	
p_A	p_B	N	p_{NC}	PCS	p_{NC}	PCS	p_{NC}	PCS
0.1	0.05	350	0.917	0.963	0.924	0.963	0.917	0.960
0.3	0.2	150	0.739	0.920	0.747	0.920	0.741	0.919
0.4	0.25	160	0.662	0.979	0.677	0.981	0.667	0.978
0.4	0.35	1500	0.624	0.975	0.626	0.977	0.624	0.979
0.6	0.4	150	0.487	0.992	0.507	0.992	0.500	0.992
0.7	0.55	250	0.370	0.991	0.379	0.993	0.379	0.991
0.8	0.35	80	0.380	1	0.461	0.999	0.442	1
0.8	0.65	350	0.271	0.999	0.278	0.999	0.284	0.999
0.9	0.5	80	0.271	1	0.324	1	0.342	1
0.9	0.75	150	0.171	0.994	0.178	0.994	0.194	0.993

Table 2: Simulated proportion of not cured for optimal statistic

			60–40 adaptive		85–15 adaptive		70–30 constant α		85–15 constant α	
p_A	p_B	N	p_{NC}	PCS	p_{NC}	PCS	p_{NC}	PCS	p_{NC}	PCS
0.1	0.05	350	0.915	0.957	0.903	0.851	0.913	0.963	0.909	0.96
0.3	0.2	150	0.733	0.914	0.709	0.819	0.727	0.912	0.72	0.893
0.4	0.25	160	0.655	0.976	0.62	0.923	0.643	0.976	0.63	0.962
0.4	0.35	1500	0.619	0.976	0.607	0.919	0.615	0.931	0.614	0.809
0.6	0.4	150	0.478	0.993	0.43	0.959	0.460	0.992	0.438	0.98
0.7	0.55	250	0.359	0.991	0.323	0.958	0.343	0.992	0.33	0.95
0.8	0.35	80	0.379	0.999	0.276	0.998	0.340	1	0.286	0.999
0.8	0.65	350	0.260	0.999	0.223	0.985	0.242	0.997	0.227	0.965
0.9	0.5	80	0.261	0.999	0.170	0.998	0.222	1	0.173	0.999
0.9	0.75	150	0.161	0.993	0.125	0.960	0.144	0.987	0.129	0.968

"85-15 Constant α" are also defined. PCS is defined as the proportion of correct selection i.e. number of cases where our hypothesis turns out to be true (Table 2).

From the table we can easily see that in our method proportion of not cured is significantly less than the existing procedures stated above and it is optimal in the sense that if we fix the bound for proportion of not cured at a particular constant (c_0), then this allocation will give us the optimal rule for allocation while simultaneously minimising the error rate among a large class of functions $T(p) = p^\alpha$. Now if the success probabilities of two treatments are close then for two competing treatment procedures we need a large number of subjects to reach a decision as expected. Small and close success probabilities need higher number of subjects than high and close success probabilities which is evident from the table. So using our optimal statistic we can improve the inferential goal at the same time controlling the proportion of not cured subjects in our experiment.

In the simulation procedure we have used the quantity 'a' tacitly. This is because the way 'a' is defined, it has become a function of k (i.e. $\frac{p_A}{p_B}$) and α (i.e. T). In the total process of simulation studies all through k and α have been used. Hence the use of $'a'$ is implicit in the process.

6. Conclusion

In our case we have achieved a lower proportion of not cured persons by using a optimal statistic and designing the allocation rule accordingly. In most cases controlling the proportion of not cured is more important than maintaining the power. Though it is indeed the case that we have lost power a little, still this procedure maintains a good balance in between the two. On that ground we can say that this procedure performs better than the existing procedures. At each step we need little computation which takes negligible computer time. So in real life application our methods in the process of costly clinical trials will be helpful.

Acknowledgments. We are grateful to an anonymous referee for his helpful comments and suggestions that helped to improve the write-up substantially.

References

ROSENBERGER, W.F., STALLARD, N., IVANOVA, A., HARPER, C.N. and RICKS, M.L (2001). Optimal adaptive designs for binary response trials. *Biometrics* **57**, 3, 909–913.

WEI, L.J. and DURHAM, S (1978). The randomized play-the-winner rule in medical trials. *J. Amer. Statist. Assoc.* **73**, 364, 840–843.

ZELEN, M. (1969). Play-the-winner rule and the controlled clinical trial. *J. Amer. Statist. Assoc.* **64**, 131–146.

BISWAS, A. and BANDYOPADHYAY, U (1999). Allocation by randomised play-the-winner rule in the presence of prognostic factors. *Sankhya: The Indian Journal of Statistics* **61**, Series B, Pt. 3, 397–412.

AZRIEL, D., MANDEL, M. and RINOTT, Y (2012). Optimal allocation to maximise the power of two sample tests for binary response. *Biometrica* **99**, 101–113.

Appendix

In this section we just provide some calculations that may not be that obvious.

Equation 2.1 : Directly follows from the result of the delta method stated above.

Equation 2.2 : We have assumed that asymptotic variance is constant. Let

$$K = Avar(S_T) = \frac{1+R}{nR}(p_A q_A (T'(p_A))^2 + R p_B q_B (T'(p_B))^2)$$

From the definition of R it follows that

$$n_B = \frac{n}{1+R}$$

Substituting the value of n obtained from the previous equation we get

$$n_A q_A + n_B q_B = \frac{R q_A + q_B}{KR}[(p_A q_A (T'(p_A))^2 + R p_B q_B (T'(p_B))^2)]$$

taking logarithm and differentiating with respect to R we obtain the desired result.

Equation 2.3 : Expression for proportion of not cured persons is the following:

$$p_{NC} = \frac{n_A q_A + n_B q_B}{n}$$

Substituting the value of n_B, n and optimal value of R obtained from the previous equation we have this result.

Equation 2.4 : From the delta method we can say that $T(\hat{p_A}) - T(\hat{p_B})$ is asymptotically normal.

$$\lim_{n\to\infty} P([T(\hat{p_A}) - T(\hat{p_B})] < 0 | [T(p_A) - T(p_B)] = a)^{\frac{1}{n}}$$

$$= \lim_{n\to\infty} \left(\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi K}} e^{\left(-\frac{(u-a)^2}{2K}\right)} du \right)^{1/n}$$

$$
\begin{aligned}
&= \max_{u\in(-\infty,0)} \lim_{n\to\infty} \frac{n^{\frac{1}{2n}}}{(2\pi t_n)^{\frac{1}{2n}}} e^{\left(-\frac{(u-a)^2}{2t_n}\right)} \\
&= e^{\left(-\frac{a^2}{2C^2 p_{NC}}\right)}
\end{aligned}
$$

Equation 3.1 : Using Taylor series expansion (taking two terms) only we get the following:

$$
\frac{a}{C} \approx \frac{\delta}{2\sqrt{p_A} - \delta\left(\frac{1}{2\sqrt{p_A}} + \frac{T''(p_A)\sqrt{p_A}}{T'(p_A)} - \frac{\delta T''(p_A)}{2\sqrt{p_A}T'(p_A)}\right)}
$$

From this the result follows by infinite series expansion.

Equations 3.2, 3.3 and 3.4 : Follows from the definition of the corresponding quantities.

Equation 4.1: In our optimisation problem the optimum occurs when the following equation is satisfied. Hence the result follows.

$$
\begin{aligned}
p_{NC} &= c_0 \\
k^{\alpha} &= \frac{(1 - c_0 - p_B)\sqrt{k}}{p_A - 1 + c_0}
\end{aligned}
$$

Anupam Kundu
Nabaneet Das
Sayantan Chakraborty
Subir Kumar Bhandari
Indian Statistical Institute,
Kolkata, India
E-mail: subir@isical.ac.in
E-mail: anupam.kundu8@gmail.com

Paper received: 29 May 2015.