Information and Management Sciences
Volume 18, Number 1, pp. 1-16, 2007

An Adaptive Design for Categorical Responses in Phase III Clinical Trials

Atanu Biswas
Indian Statistical Institute
India

Wen-Tao Huang
Tamkang University
R.O.C.

Rahul Bhattacharya
Asutosh College
India

Abstract

An adaptive design is provided for phase III clinical trials where the treatment responses are categorical. The proposed design extends the drop-the-loser rule (Ivanova [10]) which is proposed for binary treatment responses only. It is illustrated that the proposed design is an improvement over the existing design of Bandyopadhyay and Biswas [3] for such categorical responses in terms of low variability. Some probability generating functions of the proposed design are obtained. The applicability of the proposed design is illustrated by using some real data from an trial of patients of rheumatoid arthritis.

Keywords: Immigration Ball, Limiting Proportion of Allocation, Ordinal Categorical Responses, Proportion of Allocation, Randomisation, Response-Driven Adaptive Design, Treatment Difference, Urn Model.

1. Introduction

Response-driven adaptive designs are used in phase III clinical trials with an objective to treat a larger number of patients by the eventual better treatment. The objective of a phase III trial is to compare the performances of two or more competing treatments where the patients often arrive sequentially into the study. Quite often the patients are treated one after another and thus it allows to use the past allocation-and-response history up to that patient to determine his/her treatment. Thus, the adaptive designs have their role to play in such a scenario to help us achieve some ethical gain by treating a larger number

Received December 2005; Revised June 2006; Accepted July 2006.
This paper has been partially supported by National Science Council of Taiwan with grant number NSC 93-2118-M-032-002. The work of the third author was supported by a Research Fellowship from the Council of Scientific and Industrial Research (Sanction No: 9/28(567)/2002-EMR-I).

of patients by the better treatment. At the same time, we also need some significant amount of allocation to the worse treatment as well to enable us to make meaningful inference about the treatment difference in an efficient manner. Adaptive design is all about the trade-off between ethical gain (which is achieved if a larger number of patients are treated by the better treatment) and efficiency of the follow-up inference (which is achieved if the allocation is balanced in a 50:50 way).

Quite a few real applications of adaptive designs are there with an increasing frequency in the recent days. Some real applications of adaptive clinical trials for dichotomous responses are due to Professor M. Zelen (in a breast cancer trial, reported by Iglewicz [9]), Bartlett et al. [4], Tamura et al.[16], Ware [17], Rout et al. [15], Müller and Schäfer [14] and Biswas and Dewanji [5]. Several adaptive designs are available in literature, although most of them are suitable for binary treatment responses. Some of the well-known designs are the play-the-winner rule (see Zelen [20]), the randomized play-the-winner rule (see Wei and Durham [18]), the success driven design (see Durham et al. [6]), the birth and death urn design (see Ivanova et al. [11]). For such designs, the expected proportion of allocation to the better treatment arm is more than 50%, and this proportion increases with the increase in treatment difference. However, most of these designs are birth processes and accordingly the variability is too high. In fact, the standard deviations of the proportion of allocation for these designs are so high that an allocation which is less than one or two standard deviation(s) from the expectation leads less than 50% patients treated by the better treatment, in case of a two treatment experiment. Recently Ivanova [10] introduced a new adaptive design for two-treatment allocation, called the drop-the-loser rule, which is a death process. Consequently, the variation is quite low as it is known from the results of stochastic processes that death processes have less variability than the birth processes. Hu and Rosenberger [8] observed that the drop-the-loser rule has the smallest variability among the available adaptive designs for binary responses.

It is not quite difficult to generalize the urn designs for a more complicated treatment responses, say, when the responses are ordinal categorical. In several biomedical studies the responses are pain, post-operative conditions, etc., which are often measured in an ordinal categorical scale like nil, mild, moderate, severe, etc. Recently an adaptive trial was conducted in the Indian Statistical Institute, Kolkata, which was a trial of the pulsed electro-magnetic field therapy (PEMF) versus placebo on patients of rheumatoid

arthritis. See Biswas and Dewanji [5] for details. It was a longitudinal trial as the number of responses from each patient was more than one. Again the responses were multivariate categorical in the sense that there were a few responses like pain, tenderness, swelling, joint stiffness, and each of them were measured in ordinal categorical scale. But, due to the unavailability of suitable designs, we transformed that multivariate categorical responses to some univariate binary responses by using some prefixed norm and carried out an urn model based adaptive design for that longitudinal binary responses. The design was a generalization of the popular randomized play-the-winner rule (see Wei and Durham [18]). Certainly, the design could be much better if the complete information on the categorical responses could be used.

Bandyopadhyay and Biswas [3] provided a generalization of the randomized play-the-winner rule to incorporate the categorical responses with possible values $0, 1, \ldots, k$, in the design where initially the urn contains α balls of both types A and B, and for a response j from treatment A (B), we add an additional $j\beta$ balls of kind A (B) along with $(k-j)\beta$ balls of kind B (A). But that is also a birth process and has the same problem of high variability. In the present paper, we provide a version of the drop-the-loser rule applicable for categorical responses. We present the design for a single response only (not longitudinal) and we assume that the response is univariate.

The rest of the paper is organized as follows. In Section 2, a drop-the-loser rule for categorical treatment responses, which we abbreviate as CatDL. Section 3 deals with some properties, exact and limiting, of the design. The exact properties include the proportion of allocation to the better treatment, its standard deviation, some probability generating functions and also some inferential issues which are evaluated by numerical simulations. Limiting properties include limiting proportion of allocation and limiting distribution. A comparison with some existing competitors are done in Section 5. Section 5 provides an illustration with the real data on the PEMF trial, discussed earlier. Section 6 concludes.

2. Drop-the-loser Rule for Categorical Responses

Suppose we have the two competing treatments, say A and B, in a phase III clinical trial. We have a set up where the patients enter into the set up sequentially and each entering patient is treated either by A or by B using some randomisation where the probability of allocating any treatment is adaptively determined according to the *state*

of art based on the data up to that stage. Here we have a set up where the responses are ordinal categorical.

Suppose we have n patients in the trial. Let T_i be an indicator which takes the value 1 or 0 according as the ith patient is treated by A or B. Consequently, let Y_i be the response. Here Y_i can take the values $0, 1, \ldots, k$, where we assume that a higher value of response indicate a better performance of the treatment. Assume that $P(Y_i = j|T_i) = p_{Aj}T_i + p_{Bj}(1 - T_i)$, which means that the conditional probability of Y_i taking value j is p_{Aj} or p_{Bj} depending on the ith patient is treated by A or B. Clearly, $\sum_{j=0}^{k} p_{Aj} = \sum_{j=0}^{k} p_{Bj} = 1$.

Note that, in our model above, the treatment difference (see Ware [17]; Wei et al.[19]) is $\mu_A - \mu_B$, where $\mu_A = \sum_j jp_{Aj}$ and $\mu_B = \sum_j jp_{Bj}$, the expected responses for the two treatments under consideration. Our allocation design should be such that it will allocate a larger number of patients to treatment A if $\mu_A - \mu_B > 0$, and the allocation proportion to treatment A should increase with the increase in the difference $\mu_A - \mu_B$. So the allocation design will depend on the definition of treatment difference. Our proposed allocation design is as follows.

We start with an urn having one ball each of type A, B and I, where I is the immigration ball. For the $(i + 1)$ st entering patient, $i \geq 0$, we draw a ball from the urn, and treat the patient by treatment A or B if the drawn ball is of type A or B. On the other hand, if the drawn ball is of type I, we add one ball each of the types A and B to the urn, replace the I ball, and draw one ball from the urn afresh. We continue this procedure until we get a ball of A or B to treat the patient accordingly. Let the response of the patient be Y_{i+1}, and the indicator of allocation is T_{i+1}. We then replace the drawn ball with a probability $\pi_u = \pi_u(Y_{i+1})$ if the patient is treated by treatment u, $u = A, B$. We then carry out the same procedure for the next entering patient. The all important problem lies in determining $\pi_u(j)$. Due to the ordinal nature of the responses, we need $0 = \pi_u(0) < \pi_u(1) < \cdots < \pi_u(k-1) < \pi_u(k) = 1$. In particular, looking at the definition of treatment difference $\mu_A - \mu_B$, we set $\pi_u(j) = j/k$. We denote this rule as CatDL rule. In case $k = 1$, the rule reduces to the standard drop-the-loser rule of Ivanova [10].

3. Properties and Comparison: Exact

In the urn design, let $\mathbf{Z}_w = (Z_{Iw}, Z_{Aw}, Z_{Bw})$ be the urn composition after w draws, where Z_{jw} be the number of type j balls, $j = I, A, B$. According to the urn design, we

start with the initial urn composition $\mathbf{Z}_0 = (1, 1, 1)$.

Since the proposed urn scheme is not analyzed in the usual way, the only specialized technique required is that of embedding the urn into a continuous time 3-type Markov Branching process (see Athreya and Ney [2], p.221). We can use some known results on the continuous time process to characterize our urn process.

Let $Z_i(t)$ be the number of type i balls in the urn at time t for $i = I, A, B$. Suppose p_i be the probability of replacing a ball of type i and q_i be the probability that the number of type i ball is reduced by unity. In addition, let r_i be the probability of increasing the number of type i balls by unity. Clearly, $p_i + q_i + r_i = 1$ for $i = A, B$. [For the proposed design, $r_i = 0$ for all i.] Note that, for the proposed allocation design, in any trial, for any response Y by treatment i, the conditional probaility of replacing the ball to the urn is $\pi_i(Y)$, $i = A, B$. Consequently, $p_i = E(\pi_i(Y)) = \sum_{j=0}^{k} \pi_i(j)p_{ij}$, $i = A, B$. We also note that our design is such that the immigration rate remains unchanged throughout the trial.

Defining τ_w as the wth draw time, $w = 0, 1, 2, \ldots$ with $\tau_0 = 0$, we assume that the time intervals between draws are exponentially distributed with rate parameters equal to the total number of balls in the urn at that time. Also assume that $Z_{i0} = Z_i(0)$, $i = A, B$. Then we have the following theorem.

Theorem 1. *The discrete time stochastic process* $\{\mathbf{Z}_w,\ w = 0, 1, 2, \ldots\}$ *and the continuous time stochastic process* $\{\mathbf{Z}(\tau_w),\ w = 0, 1, 2, \ldots\}$ *are equivalent.*

Proof. The proof follows exactly in the same way of Athreya and Ney [2], p.221.

Note that the time parameter is an artificial constant and has no particular relation to the real time. We call it *virtual time*. As a consequence, we can consider the following two sampling schemes. (a) Stop the sequence of trials at a certain *virtual time* t, and (b) Continue the sequence of trials until a certain number of subjects are treated.

Joint probability generating functions:

We start with the embedded rule when there is no immigration ball in the urn. We want to obtain the probability generating function in this set up. In this set up, we define the following. Suppose $Z_i(t)$ be the number of type i balls by time t, $N_i(t)$ be the number of trials on treatment i by time t, and $X_{ij}(t)$ be the number of trials resulted in j on treatment i by time t, $i = A, B$, $j = 0, \ldots, k$. As the process corresponding to

each ball type in continuous time are independent, we only consider treatment i. Note that $Z_i(t)$ is a linear death process with immigration. The following theorem gives the differential equation for the joint probability generating function. This equation is solved only when there is no immigration and effect of immigration is considered later on.

Theorem 2. *Given that $Z_i(0) = 1$, the joint generating function G_i of $X_{ij}(t)$, $j = 0, \ldots, k$, $Z_i(t)$, $N_i(t)$ satisfies*

$$\frac{\partial}{\partial t}G(\mathbf{s},t) = a(s_{k+1}-1)G_i(\mathbf{s},t) + \sum_{j=0}^{k} p_{ij}\pi_i(j)s_j s_{k+1}s_{k+2}\frac{\partial}{\partial s_{k+1}}G_i(\mathbf{s},t)$$
$$+ \sum_{j=0}^{k} p_{ij}(1-\pi_i(j))s_j s_{k+2}\frac{\partial}{\partial s_{k+1}}G_i(\mathbf{s},t) - s_{k+1}\frac{\partial}{\partial s_{k+1}}G_i(\mathbf{s},t), \quad (3.1)$$

where

$$\mathbf{s} = (s_0, s_1, \ldots, s_{k+2}),\ |s_j| \le 1, \quad \textit{for all } j,$$
$$G_i(\mathbf{s},t) = \sum_{s_0=0}^{\infty}\cdots\sum_{s_{k+2}=0}^{\infty}\left\{\prod_{j=0}^{k} s_j^{x_j}\right\} s_{k+1}^{z}s_{k+2}^{n}P(x_0,\ldots,x_k,z,n,t),$$

with $P(x_0, \ldots, x_k, z, n, t)$ be the joint probability function and a is the number of immigration balls in the urn [$a = 1$ in our case].

Proof. Here we consider the possible transitions which can occur in the time interval Δt and results in x_j responses of type j, $j = 0, \ldots, k$, z type i balls and n trials on treatment i. Define $P(x_0, \ldots, x_k, z, n, t)$ as the probability of getting x_j responses of type j, $j = 0, \ldots, k$, z type i balls, n trials on treatment i, starting with x_{j0} responses of type j, $j = 0, \ldots, k$, z_0 type i balls and n_0 trials on treatment i. We have the following possible cases.

(i) First, consider an immigration, i.e.

$$(x_0, x_1, \ldots, x_k, z-1, n) \to (x_0, x_1, \ldots, x_k, z, n)$$

with probability of occurrence $a\Delta t$.

(ii) Next, consider a response of type j and the ball is replaced, i.e.

$$(x_0, \ldots, x_{j-1}, x_j - 1, x_{j+1}, \ldots, x_k, z, n-1) \to (x_0, x_1, \ldots, x_k, z, n)$$

with probability $zp_{ij}\pi_i(j)\Delta t$, $j = 0, \ldots, k$.

(iii) Also consider a response of type j, but the ball is not replaced, i.e.

$$(x_0, \ldots, x_{j-1}, x_j - 1, x_{j+1}, \ldots, x_k, z+1, n-1) \to (x_0, x_1, \ldots, x_k, z, n)$$

with probability $(z+1)p_{ij}(1-\pi_i(j))\Delta t$, $j = 0, \ldots, k$.

(iv) In addition, a trial can result in occurrence of no composite events, i.e.

$$(x_0, x_1, \ldots, x_k, z, n) \to (x_0, x_1, \ldots, x_k, z, n)$$

with probability $1 - (z+a)\Delta t$.

Since $Z_i(t)$ is a linear death process with immigration, it can be shown that (see Karlin and Taylor [12], p. 189), starting from $(\mathbf{x}^*, z^*, n^*)$,

$$P(\mathbf{x}, z, n, \Delta t)/\Delta t = o(1)$$

for $z^* \notin \{z, z+1\}$. Then, for $z > 0$, we have

$$\begin{aligned}
&\frac{\partial}{\partial t} P(x_0, \ldots, x_k, z, n, t) \\
&\quad = a\ P(\mathbf{x}, z, n, t) + \sum_{j=0}^{k} z p_{ij} \pi_i j P(x_0, \ldots, x_{j-1}, x_j - 1, x_{j+1}, \ldots, x_k, z, n-1, t) \\
&\qquad + \sum_{j=0}^{k} (z+1) p_{ij} (1 - \pi_i(j)) P(x_0, \ldots, x_{j-1}, x_j - 1, x_{j+1}, \ldots, x_k, z+1, n-1, t) \\
&\qquad - (z+a) P(x_0, \ldots, x_k, z, n, t), \qquad (3.2)
\end{aligned}$$

whereas, for $z = 0$, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t} P(\mathbf{x}, 0, n, t) = &\sum_{j=0}^{k} p_{ij} (1 - \pi_i(j)) P(x_0, \ldots, x_{j-1}, x_j - 1, x_{j+1}, \ldots, x_k, 0, n, t) \\
&- a P(x_0, \ldots, x_k, 0, n, t). \qquad (3.3)
\end{aligned}$$

Multiplying both sides of (3.2) and (3.3) by $\left\{\prod_{j=0}^{k} s_j^{x_j}\right\} s_{k+1}^z s_{k+2}^n$ and summing over all possible values, we obtain the equation in (3.1).

Writing

$$\alpha = \sum_{j=0}^{k} p_{ij} \pi_i(j) s_j s_{k+2},$$

$$\beta = \sum_{j=0}^{k} p_{ij} (1 - \pi_i(j)) s_j s_{k+2},$$

and $s = s_{k+1}$ and taking $a = 0$, we have from (2.1),

$$\frac{\partial}{\partial t} G_i(\mathbf{s}, t) = (-\gamma s + \delta) \frac{\partial}{\partial s} G_i(\mathbf{s}, t)$$

with $G_i(\mathbf{s}, 0) = s$ and $-\gamma = \alpha - 1$, $\delta = \beta$. Thus, using the algorithm given in Anderson [1], pp. 104-105, we obtain

$$G_i(\mathbf{s}, t) = \left(\frac{\delta}{\gamma}\right)\left(1 - e^{-\gamma t}\right) + s\ e^{-\gamma t}.$$

Consequently, we find the probability generating function of $N_i(t)$ as

$$\begin{aligned} G^{N_i(t)}(w, t) &= G_i(\mathbf{s}, t | s_j = 1, j = 0, \ldots, k+1; s_{k+2} = w) \\ &= \frac{w \sum_{j=0}^{k} p_{ij}(1 - \pi_i(j))}{1 - w \sum_{j=0}^{k} p_{ij} \pi_i(j)} \left\{1 - e^{-(1 - w \sum_{j=0}^{k} p_{ij} \pi_i(j))t}\right\} + e^{-(1 - w \sum_{j=0}^{k} p_{ij} \pi_i(j))t}. \end{aligned}$$

Note that the above expression is the same as that of Durham and Ivanova [7] with p_i replaced by $\sum_{j=0}^{k} p_{ij} \pi_i(j)$ and q_i replaced by $\sum_{j=0}^{k} p_{ij}(1 - \pi_i(j))$. Using the technique given in Durham and Ivanova [7], we obtain the joint generating function for $a > 0$ in terms of the following theorem.

Theorem 3. *The joint generating function for $X_{ij}(t)$, $j = 0, \ldots, k$, $Z_i(t)$, $N_i(t)$ for the embedded CatDL rule with initial composition $\mathbf{Z}_0 = (a, 1, 1)$ is*

$$G_i(\mathbf{s}, t | a) = e^{-at} \exp\left\{a \int_0^t G_i(\mathbf{s}, u) du\right\} G_i(\mathbf{s}, t).$$

In this Section, we also provide a detailed simulation study to examine the nature of ethical gain obtained through this allocation design. In fact, we study the proportion of allocation to the two treatments for different distributions of the responses. Some of the simulation results are presented in Table 1. It is observed that we have considerable larger allocation to the better treatment. We compare our results with the standard allocation design for categorical responses provided by Bandyopadhyay and Biswas [3] (henceforth called B&B rule) as that seems to be the only comparable adaptive design with categorical responses in this case. This proportion is denoted by Prop(A). We also study the standard deviation (SD) of the proportion, as the initial goal of the DL rule is to reduce the variability. These are presented in Table 1 for different parametric values. We also present the same for the 50:50 randomized rule, where every patient is randomly allocated to either treatment irrespective of the accumulated data. The table

also provides a comparison of the limiting proportion of allocation of the CatDL rule with the same for the 50:50 rule (in which case it is always 0.5) and the B&B design. Note that the limiting proportion of the B&B rule is same as that of the CatDL rule.

The natural question following the allocation is to carry out the inference. Here we want to carry out a test for the null hypothesis $H_0 : \mu_A = \mu_B$ against the one-sided alternative $H_1 : \mu_A > \mu_B$. For simplicity, we carry out the test for a fixed-sample size n. Quite naturally, a right-tailed test based on the test statistic $\widehat{\mu}_{An} - \widehat{\mu}_{Bn}$, the estimate of the treatment difference based on the first n sample observations, is recommended. We present the power of the test in Table 2 for the design CatDL. The null distribution of the test statistic is symmetric about '0', but the distribution depends on the null value of $\mu_A = \mu_B$ through the vecors $\mathbf{p}_A = (p_{A0}, \ldots, p_{Ak})$ and $\mathbf{p}_B = (p_{B0}, \ldots, p_{Bk})$. But still we compare the powers of the tests from different competitive designs for the sake of comparison just to see the comparative power structure for different designs. In practice, one should arrive at a fixed cut-off point, irrespective of the null hypothetical parametric value. This can be done theoretically either by integrating the cut-off points over the empirical distribution or the prior distribution of the parameter. The details are under study.

As a natural comparison, we compare the performance of our designs and the follow-up tests with a test procedure which randomises the patients among the two treatments in a 50:50 way. Note that the expected allocation proportion and the limiting allocation proportion is 0.5 in such a randomized 50:50 procedure, whatever be the treatment difference. Thus, although we can have more power in the follow-up test, we have ethical loss in terms of treating a larger number of patients to the worse treatment than the corresponding CatDL rule.

There seems only one adaptive design available in the literature which considers continuous treatment responses and also covariates. This is the design introduced and studied by Bandyopadhyay and Biswas [3]. We compare the performances our proposed designs in Section 2 with the design of B&B. The numerical computations are provided in Tables 1-2. It is observed that our proposed design works well in terms of allocation and power and too with a much lower variability. This establishes the superiority of the proposed design.

4. Properties: Limiting Distribution and Proportion

We are interested in the proportion of subjects assigned to the treatment i, $i = A, B$, as $t \to \infty$. As in Ivanova [10], in the present categorical response set up, it can be shown that, as $t \to \infty$,

$$\frac{N_i(t)}{at} \xrightarrow{P} \frac{1}{\sum_{j=0}^{k}(1-\pi_i(j))p_{ij}}, \ i = A, B.$$

Hence, as $t \to \infty$,

$$\frac{N_i(t)}{N_A(t)+N_B(t)} \xrightarrow{P} \frac{\frac{1}{\sum_{j=0}^{k}(1-\pi_i(j))p_{ij}}}{\frac{1}{\sum_{j=0}^{k}(1-\pi_A(j))p_{Aj}}+\frac{1}{\sum_{j=0}^{k}(1-\pi_B(j))p_{Bj}}} = \bar{D}_i, \text{ say.}$$

Consequently, as in Ivanova [10],

$$\sqrt{t}\left(\frac{N_i(t)}{N_A(t)+N_B(t)} - \bar{D}_i\right) \xrightarrow{d} N\left(0, \bar{D}_A^2\bar{D}_B^2\frac{\bar{p}_A+\bar{p}_B}{a}\right)$$

as $t \to \infty$, where $\bar{p}_i = \sum_{j=0}^{k}\pi_i(j)p_{ij}$.

In clinical trials, sampling is often made until a prefixed sample size, say n, predetermined by using some power condition. In case of CatDL rule, the conditional distribution of $N_i(t)$ given $N_A(t) + N_B(t) = n$ is not straightforward, but the limiting proportion of allocation can be obtained by using the embedding theorem as

$$\lim_{n\to\infty}\frac{N_{in}}{n} = \lim_{t\to\infty}\frac{N_i(t)}{N_A(t)+N_B(t)}$$

where N_{in} is the number of allocations to the ith treament. Consequently,

$$\frac{N_{in}}{n} \xrightarrow{P} \bar{D}_i, \ i = A, B. \tag{3.1}$$

It is interesting to note that the above CatDL rule can be interpreted as the usual drop-the-loser rule (Durham and Ivanova [7]; Ivanova [10]) with the unconditional probability of replacing the ball $\sum_{j=0}^{k}\pi_A(j)p_{Aj} = \mu_A/k$ or $\sum_{j=0}^{k}\pi_B p_{Bj} = \mu_B/k$ depending on the patient is treated by the treatment A or B. Consequently, the limiting proportion of allocation to treatment A becomes

$$\pi^* = (k-\mu_B)/(2k-\mu_A-\mu_B).$$

We also consider the asymptotic distribution of $\widehat{\Delta}$. Note that a natural estimator of μ_s, $s = A, B$, is

$$\widehat{\mu}_s = \sum_{j=0}^{k} w_j \nu_{sjn}/N_{sn},$$

where $w_j = j$, ν_{sjn} is the number of patients responded j in the treatment group s, and N_{sn} is the total number of patients allocated to treatment s out of the first n patients. Clearly, we can write $\nu_{sjn} = \sum_{i=1}^{n} W_{ij}^s$, where W_{ij}^s takes the values 1 or 0 according as the response from the ith patient by the treatment s is j or not, $i = 1, \ldots, n$, $s = A, B$. Writing $\mathbf{W}_i^s = (W_{i0}^s, \ldots, W_{ik}^s)^T$, for the ith patient we write the vector of observation $\mathbf{W}_i = T_i \mathbf{W}_i^A + (1 - T_i)\mathbf{W}_i^B$, $i = 1, \ldots, n$, with $\sum_{j=1}^{k} W_{ij}^s = 1$ for all (i, s). Clearly, $Y_i = j$ if $W_{ij} = 1$ and all other components of $\mathbf{W}_i$ are zero. Note that $\mathbf{W}_i^s$ follows a multinomial distribution with parameters $(1; p_{s0}, p_{s1}, \ldots, p_{sk})$.

Now we can write

$$\begin{aligned}\widehat{\Delta} - \Delta &= \sum_{j=0}^{k} w_j \sum_{i=1}^{n} T_i (W_{ij} - p_{Aj})/N_{An} - \sum_{j=0}^{k} w_j \sum_{i=1}^{n} (1 - T_i)(W_{ij} - p_{Bj})/N_{Bn} \\ &= \sum_{j=0}^{k} w_j \widehat{\Delta}_{Aj} - \sum_{j=0}^{k} w_j \widehat{\Delta}_{Bj},\end{aligned}$$

where

$$\widehat{\Delta}_{Aj} = \sum_{i=1}^{n} T_i (W_{ij} - p_{Aj})/N_{An}, \quad \widehat{\Delta}_{Bj} = \sum_{j=0}^{k} w_j \sum_{i=1}^{n} (1 - T_i)(W_{ij} - p_{Bj})/N_{Bn}.$$

Note that, for the corresponding non-adaptive estimators $\tilde{\Delta}_{Aj}$ and $\tilde{\Delta}_{Bj}$, say, $j = 0, 1, \ldots, k$, writing $\tilde{\boldsymbol{\Delta}}_s = (\tilde{\boldsymbol{\Delta}}_{s0}, \ldots, \tilde{\boldsymbol{\Delta}}_{sk})^T$, we have

$$\sqrt{n}\tilde{\boldsymbol{\Delta}}_s \xrightarrow{d} N_{k+1}(\mathbf{0}, \Sigma_s),$$

where $\Sigma_s = (\sigma_{s,jj'})$ with $\sigma_{s,jj} = p_{sj}(1 - p_{sj})$ and $\sigma_{s,jj'} = -p_{sj} psj'$ when $j \neq j'$, $s = A, B$. Now, using (4.1), it follows from a simple extension of Theorem 3.2 of Melfi and Page [13] that

$$(\sqrt{n}\widehat{\boldsymbol{\Delta}}_A, \sqrt{n}\widehat{\boldsymbol{\Delta}}_B) \xrightarrow{d} (\mathbf{Z}_1, \mathbf{Z}_2),$$

where $\widehat{\boldsymbol{\Delta}}_s = (\widehat{\boldsymbol{\Delta}}_{s0}, \ldots, \widehat{\boldsymbol{\Delta}}_{sk})^T$, $s = A, B$, and $\mathbf{Z}_1 \sim N_{k+1}(\mathbf{0}, \Sigma_A)$ independently of $\mathbf{Z}_2 \sim N_{k+1}(\mathbf{0}, \Sigma_B)$. Using Cramer-Wold device, we immediately obtain that

$$\sqrt{n}(\widehat{\Delta} - \Delta) \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = \frac{\mathbf{w}^T \Sigma_A \mathbf{w}}{\bar{D}_A} + \frac{\mathbf{w}^T \Sigma_B \mathbf{w}}{\bar{D}_B},$$

with $\mathbf{w} = (w_0, \ldots, w_k)^T$, which is the vector $(0, 1, \ldots, k)^T$ in the present situation.

5. Illustration with Real Data

Here consider the PEMF trial described in the Introduction. We just want to illustrate the applicability of our present approach. We, for simplicity, consider only one response variable, pain, and moreover, we consider only one response by clubbing the longitudinal responses. Our response is the worst pain-level (classified into nil, mild, low moderate and high moderate) in the first two weeks. In the study, we have data from 22 patients of which 16 are treated by the PEMF and 6 by placebo. We find the empirical distributions of the treatment responses from the data and treat them as the true ones. Let these responses-distributions for PEMF be C_A and that for placebo be C_B. From the PEMF data, we obtain $C_A : \mathbf{p}_A = (0, 2/6, 2/6, 2/6)$, $C_B : \mathbf{p}_B = (2/16, 8/16, 6/16, 0)$. Using these we carry out a simulation study of 10000 simulations to find the expected number of allocations to the two treatments and the SD using our proposed CatDL rule. We then carry out the same exercise using the generalized randomized play-the-winner rule of B&B [3] and also for the 50:50 allocations. Also we find the expected allocations with SD for a randomized play-the-winner (RPW) rule and the drop-the-loser (DL) rule where the responses nil and mild are clubbed together as success and low moderate and high moderate are clubbed together as failure and these probabilities of successes and failures are obtained from the estimated distributions C_A and C_B. We also report the same for a 50:50 randomized rule. The results are reported in Table 3. From the Table 3, we observe that the proposed CatDL rule performs much better than the other two adaptive alternatives with categorical responses in the sense that it has much less SD. Expected proportion of allocation to the better treatment arm in case of CatDL is almost same as that of the B&B rule. Each of the adaptive rule is better than the 50:50 randomized rule in the ethical sense as more allocations are likely to the better treatment. But the proposed CatDL is the best in the sense that the SD is minimum. If the first and last two categories are clubbed together, the responses can be transformed to binary ones. The corresponding results for DL, RPW (with $\alpha = \beta = 1$) and 50:50 randomized rule are also given in Table 2. But, the adaptive designs using the categorical responses are

more sensible than the designs with transformed binary responses in any case as those use complete categorical responses.

Table 1. Comparison of Prop(A) and SD (in parantheses) of the CatDL, B&B and 50:50 designs. Here $k = 3$ and $\mathbf{p}_B = (0.2, 0.3, 0.3, 0.2)$. Limiting proportion of allocation (π^*) for CatDL and B&B (the same) is also given.

	$n = 40$			$n = 100$			π^*
p_A	CatDL	B&B	50:50	CatDL	B&B	50:50	
$(0.2, 0.3, 0.3, 0.2)$	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	(0.069)	(0.102)	(0.078)	(0.047)	(0.067)	(0.050)	
$(0.2, 0.2, 0.3, 0.3)$	0.526	0.531	0.500	0.531	0.534	0.500	0.536
	(0.072)	(0.110)	(0.079)	(0.050)	(0.072)	(0.050)	
$(0.2, 0.2, 0.2, 0.4)$	0.542	0.552	0.500	0.548	0.553	0.500	0.556
	(0.073)	(0.117)	(0.079)	(0.051)	(0.079)	(0.050)	
$(0.1, 0.2, 0.3, 0.4)$	0.569	0.587	0.500	0.586	0.593	0.500	0.600
	(0.075)	(0.122)	(0.080)	(0.053)	(0.080)	(0.050)	
$(0.1, 0.1, 0.2, 0.6)$	0.613	0.654	0.500	0.646	0.667	0.500	0.682
	(0.075)	(0.133)	(0.078)	(0.053)	(0.091)	(0.050)	

Table 2. Comparison of power for the CatDL, B&B and 50:50 designs. Here $k = 3$ and $\mathbf{p}_B = (0.2, 0.3, 0.3, 0.2)$.

	$n = 40$			$n = 100$		
p_A	CatDL	B&B	50:50	CatDL	B&B	50:50
$(0.2, 0.3, 0.3, 0.2)$	0.050	0.050	0.050	0.050	0.050	0.050
$(0.2, 0.2, 0.3, 0.3)$	0.153	0.161	0.159	0.261	0.260	0.254
$(0.2, 0.2, 0.2, 0.4)$	0.245	0.253	0.248	0.434	0.422	0.434
$(0.1, 0.2, 0.3, 0.4)$	0.448	0.446	0.451	0.799	0.773	0.780
$(0.1, 0.1, 0.2, 0.6)$	0.781	0.765	0.795	0.989	0.984	0.986

Table 3. Simulated number of allocations and SD's for CatDL, B&B, RPW, DL and 50:50 rules for 22 patients using the empirical response distributions from the PEMF data.

	Allocation to PEMF	
Design	Expectation	SD
CatDL	0.621	0.089
B&B	0.613	0.129
50:50 (Cat)	0.500	0.106
DL	0.590	0.024
RPW	0.607	0.145
50:50 (binary)	0.500	0.102

6. Conclusions

The design depends on the definition of the treatment difference. If the treatment difference is defined in some other way, we need to modify the design accordingly. For example, if the treatment difference is defined as

$$\mu_A(\alpha) - \mu_B(\alpha) = \sum_{j=0}^{k} \alpha_j p_{Aj} - \sum_{j=0}^{k} \alpha_j p_{Bj},$$

where $\alpha_0 < \alpha_1 < \cdots < \alpha_k$, then we set

$$\pi_A(j) = \pi_B(j) = \alpha_j^* = \frac{\alpha_j - \alpha_0}{\alpha_k - \alpha_0},$$

so that we have $0 = \alpha_0^* < \alpha_1^* < \cdots < \alpha_k^* = 1$.

In this paper we introduced drop-the-loser type designs for categorical responses. These designs yield adaptive allocation for categorical responses with smaller variability than the existing adaptive design of Bandyopadhyay and Biswas [3]. The present work assumes a very simple structure where there is no delayed responses, no staggered entry. With the presence of all these practical logistics the method will be much more complicated and we need to adjust the rules sensibly to carry out response-adaptive allocation. The details are under study.

Acknowledgement

The authors wish to thank the referee for careful reading.The work was carried out when the first author was visiting the Department of Management Sciences and Decision Making, Tamkang University, Tamsui, Taiwan. The first author thanks the department for excellent hospitality during the visit.

References

[1] W. J. Anderson, Continuous-Time Markov Chains, Springer Verlag, New York, 1991.

[2] K. B. Athreya and P. E. Ney, Branching Processes, Springer Verlag, Berlin, 1972.

[3] U. Bandyopadhyay and A. Biswas, *Selection procedures in multi-treatment clinical trials*, Metron, Vol. 60, pp.143-157, 2001.

[4] 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 trial*, Pediatrics, Vol. 76, pp.479-487, 1985.

[5] A. Biswas and A. Dewanji, *A randomized longitudinal play-the-winner design for repeated binary data.* Australian and New Zealand Journal of Statistics, to appear, 2004.

[6] S. Durham, N. Fluornoy and W. Li, A sequential design for maximizing the probability of a favourable response. *The Canadian Journal of Statistics*, Vol. 26, pp.479-495, 1998.

[7] S. D. Durham and A. Ivanova, Drop-the-loser. Technical Report. University of North Carolina, 2001.

[8] F. Hu and W. F. Rosenberger, optimality, variability, power: Evaluating response-adaptive randomization procedures for treatment comparisons, *Journal of the American Statistical Association*, Vol. 98, pp.671-678, 2003.

[9] B. Iglewicz, Alternative designs: sequential, multi-stage, decision theory and adaptive designs. In *Cancer Clinical Trials: Methods and Practice*, eds. M. E. Buyse, J. Staquet and R. J. Sylvester, pp. 312-334. Oxford University Press, 1983.

[10] A. Ivanova, *A play-the-winner-type urn design with reduced variability*, Metrika, Vol.58, pp.1-13, 2003.

[11] A. Ivanova, W. F. Rosenberger, S. D. Durham and N. Flournoy, *A birth and death urn for randomized clinical trials*, Sankhya, Series B, Vol. 62, pp.104-118, 2000.

[12] S. Karlin and H. M. Taylor, A First Course in Stochastic Processes, Academic Press, New York, 1975.

[13] V. F. Melfi and C. Page, *Estimation after adaptive allocation*, *Journal of Statistical Planning and Inference*, Vol.87, pp.353-363, 2000.

[14] H. H. Müller and H. Schäfer, *Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and of classical group sequential approaches*, Biometrics, Vol.57, pp.886-891.

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

[16] R. N. Tamura, D. E. Faries, J. S. Andersen and J. H. Heiligenstein, *A case study of an adaptive clinical trials in the treatment of out-patients with depressive disorder*, J. Amer. Statist. Assoc., Vol. 89, pp.768-776, 1994.

[17] J. H. Ware, Investigating therapies of potentially great benefit: ECMO. *Statistical Science*, Vol.4, pp.298-340, 1989.

[18] L. J. Wei and S. Durham, *The randomized play-the-winner rule in medical trials*, J. Am. Statist. Assoc., Vol.73, pp.838-843, 1978.

[19] L. J. Wei, R. T. Smythe, D. Y. Lin and T. S. Park, *Statistical inference with datadependent treatment allocation rules*, J. Amer. Statist. Assoc. Vol. 85, pp.156-162, 1990.

[20] M. Zelen, *Play-the-winner rule and the controlled clinical trial*, J. Am. Statist. Assoc. Vol.64, pp.131-146, 1969.

Authors' Information

Atanu Biswas is an Associate Professor in the Applied Statistics Unit, Indian Statistical Institute, Kolkata, India. He got his PhD from the Calcutta University. His research interest includes sequential adaptive designs in clinical trials, categorical data, survival analysis, nonparametrics and sequential nalysis.

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

E-mail: atanu@isical.ac.in TEL : +91-33-2575-2818.

Wen-Tao Huang is a full professor of Department of Management Sciences and Decision Making, Tamkang University. He got his PhD from Purdue University. He is interested in the Statistical Inferences, Empirical Bayes Decisions and Indutrial Statistics.

Department of Management Sciences and Decision Making, Tamkang University, Tamsui, Taipei County, Taiwan 251, Republic of China.

E-mail: 005697@mail.tku.edu.tw TEL : +886-2-26215656 ext 3395.

Rahul Bhattacharya is a Lecturer in the Department of Statistics, Asutosh College, Kolkata, India, and a graduate student in the Department of Statistics, Calcutta University, India. His research area is sequential adaptive designs and two stage designs.

Department of Statistics, Asutosh College, 92 S.P.Mukherjee Road, Kolkata – 700 026, India.

E-mail: rahul_bhattya@yahoo.com