

# Randomly reinforced urns for clinical trials with continuous responses

*Urne con rinforzo aleatorio per esperimenti clinici con risposte continue*

Pietro Muliere<sup>1</sup>, Anna Maria Paganoni<sup>2</sup>, Piercesare Secchi<sup>2</sup>

<sup>1</sup> Istituto di Metodi Quantitativi, Università “L. Bocconi”, Milano

<sup>2</sup> MOX- Dipartimento di Matematica, Politecnico di Milano

e-mail: piercesare.secchi@polimi.it

**Riassunto:** Un disegno sperimentale adattivo alle risposte modifica le probabilità di allocazione ai trattamenti, mano a mano che si rendono disponibili i dati dell’esperimento, con una duplice finalità: (a) generare informazione utile all’analisi inferenziale per il confronto dei trattamenti e (b) allocare un maggior numero di unità sperimentali al trattamento migliore. In questo lavoro passiamo in rassegna le proprietà di una classe di disegni adattivi utili anche nel caso di esperimenti con risposte continue e costruiti per mezzo di un’urna con rinforzi aleatori che estende quella classica di Pólya. Questi schemi d’urna hanno come obiettivo quello di allocare asintoticamente le unità sperimentali al trattamento migliore, diversamente da altre procedure che mirano invece ad allocare asintoticamente le unità ai trattamenti secondo una proporzione altrimenti ottimale.

**Keywords:** response-adaptive designs, randomly reinforced urns, clinical trials.

## 1. Introduction: response-adaptive designs

An experiment is being conducted to compare two treatments. Observations accrue sequentially; when a new statistical unit enters the experiment, the experimenter faces the control problem of allocating the unit to a treatment based on information generated by past observations and relative to the unknown response distributions. The experimenter simultaneous goals are (a) to collect evidence for determining the superior treatment and (b) to maximize the number of units in the experiment that receive the superior treatment; moreover, as part of the problem constraints, the experimenter’s allocation strategy must be random.

This control problem is encountered in industrial applications, but it is more easily illustrated within the framework of clinical trials where it is motivated by ethical reasons; units to be treated are patients and the experimenter’s desire to bias along the experiment the allocation probabilities toward the better treatment sounds immediately appealing. Moreover, the constraint that the allocation strategy be randomized is easily sustained in this setting.

Under the names of response-adaptive designs or response-driven designs (we will not make a distinction), a vast number of allocation strategies for this problem have been proposed in recent years: an informed review is found in Rosenberger (2002) and Rosenberger and Lachin (2002). Most response-adaptive designs deal with the case where responses are dichotomous. Only a few procedures have been studied for clinical trials with continuous outcomes: the literature includes Bandyopadhyay and Biswas (2001), Biswas and Basu (2001), Melfi *et al.* (2001), Hu and Zhang (2004), Atkinson and Biswas

(2005), Baldi Antognini and Giovagnoli (2005), Zhang and Rosenberger (2006). These response-adaptive designs have been proposed within the context of optimal allocation: i.e. their goal is to allocate patients to treatments targeting (asymptotically) a proportion  $\rho \in (0, 1)$  which is optimal in a sense to be specified. This is different from the goal of the designs object of this paper which target the treatment generating the superior response. To clarify this important issue at the very onset, after setting the statistical model, in section 3 we briefly describe a very general design for targeting an optimal allocation: the implementation studied in Hu and Zhang (2004) of the doubly-adaptive coin design of Eisele (1994). Section 4 introduces two designs targeting the optimal treatment: one has been studied in Muliere *et al.* (2006) and in Paganoni and Secchi (2005), the other one is new. For comparing the performance of different response-adaptive designs when response distributions are normal, Hu and Rosenberger (2003) and Zhang and Rosenberger (2006) introduce a method based on the Taylor expansion of the noncentrality parameter of a standard z-test; this method does not work for designs targeting the optimal treatment. A guideline for their evaluation and comparison has been set in Paganoni and Secchi (2005); we sketch it in section 5. Numerical illustrations along this guideline are presented in the last section of the paper.

## 2. The model

Consider a clinical trial with two treatments, say  $R$  and  $W$ . Patients enter the trial sequentially and are allocated to either treatment according to a *response-adaptive design*  $\delta = (\delta_0, \delta_1, \dots, \delta_n, \dots)$  where  $\delta_0 \in [0, 1]$  while, for  $n \geq 1$ ,  $\delta_n$  is a measurable function from the space  $(\{0, 1\} \times \mathbb{R}^k)^n$  to  $[0, 1]$ .

To understand  $\delta$ , for  $n = 1, 2, \dots$  indicate with  $X_n$  the random variable which takes value 1 if patient  $n$  is allocated to treatment  $R$  and value 0 if patient  $n$  is allocated to treatment  $W$ . The response of patient  $n$  is represented by a random vector  $Y_n \in \mathbb{R}^k$ . Let  $\mathcal{F}_n$  be the sigma-field generated by  $X_1, Y_1, \dots, X_n, Y_n$ , i.e. the sigma-field representing the information available to the experimenter before the allocation of the  $(n + 1)$ -th patient. Then we assume that  $X_1$  has Bernoulli( $\delta_0$ ) distribution, while, for  $n \geq 1$ , the conditional distribution of  $X_{n+1}$ , given  $\mathcal{F}_n$ , is Bernoulli( $\delta_n(X_1, Y_1, \dots, X_n, Y_n)$ ).

To complete the model we need to specify the conditional distribution of  $Y_{n+1}$  given  $\mathcal{F}_n$  and  $X_{n+1}$ . In accordance with the literature on response-adaptive designs, let  $(R_1, R_2, \dots)$  and  $(W_1, W_2, \dots)$  be two independent sequences of i.i.d random vectors describing the potential responses of patients if allocated to treatment  $R$  or  $W$  respectively: we indicate with  $\nu_R$  the distribution of  $R_1$  and with  $\nu_W$  the distribution of  $W_1$ . Then, for  $n = 0, 1, 2, \dots$ , we assume that

$$Y_{n+1} = X_{n+1}R_{n+1} + (1 - X_{n+1})W_{n+1}.$$

Some final notations: when patient  $n + 1$  enters the trial, a total of  $\sum_{i=1}^n X_i$  patients has been already allocated to treatment  $R$ . For shortness, we indicate with  $N_R(n)$  this random number;  $N_W(n) = n - N_R(n)$  indicates the number of patients already allocated to treatment  $W$ .

### 3. Targeting an optimal allocation: the doubly biased coin design

The response distributions  $\nu_R$  and  $\nu_W$  are usually unknown; almost all the literature on response-adaptive designs consider the parametric case where the specification of  $\nu_R$  and  $\nu_W$  depends on the value of two finite vectors of parameters, say  $\theta_R$  and  $\theta_W$  respectively. Within the optimal allocation approach, the design's goal is to randomly allocate to treatment  $R$  a certain proportion  $\rho \in (0, 1)$  of patients among the  $n$  patients involved in the trial; in point of fact, the design is said to be *optimal* if  $\lim_{n \rightarrow \infty} N_R(n)/n = \rho$  almost surely. The *target allocation*  $\rho = \rho(\theta_R, \theta_W)$  is usually a function of the unknown parameters  $\theta_R$  and  $\theta_W$  and is determined by means of some optimality criteria. A very general sequential design that can target any proportion  $\rho$  satisfying some regularity conditions, is the doubly-adaptive coin (DBCD) of Eisele (1994) implemented with the allocation function proposed by Hu and Zhang (2004).

Let  $g : [0, 1] \times [0, 1] \rightarrow [0, 1]$  be defined as

$$g^{(\alpha)}(\pi, \rho) = \begin{cases} \frac{\rho(\rho/\pi)^\alpha}{\rho(\rho/\pi)^\alpha + (1-\rho)((1-\rho)/(1-\pi))^\alpha} & \text{for } (\pi, \rho) \in (0, 1) \times [0, 1] \\ 1 - \pi & \text{for } (\pi, \rho) \in \{0, 1\} \times [0, 1], \end{cases}$$

where  $\alpha$  is a nonnegative real number that controls the degree of randomness for the procedure. The DBCD studied in Hu and Zhang (2004) is now defined as follows: let  $\delta_0 = 1/2$  and, for  $n = 1, 2, \dots$ , set

$$\delta_n(X_1, Y_1, \dots, X_n, Y_n) = g^{(\alpha)}\left(\frac{N_R(n)}{n}, \rho(\hat{\theta}_R, \hat{\theta}_W)\right)$$

where  $(\hat{\theta}_R, \hat{\theta}_W)$  is an estimator of  $(\theta_R, \theta_W)$  measurable with respect to  $\mathcal{F}_n$ . (For a discussion on the type of estimator see Hu and Zhang (2004): MLEs or moment estimators are usually fine). If  $\rho$  is a continuous function and it is twice continuously differentiable in a small neighborhood of  $(\theta_R, \theta_W)$ , Hu and Zhang (2004) proved that the DBCD can target  $\rho$  in the sense that, when the number  $n$  of patients involved in the trial grows to infinity, both  $N_R(n)/n$  and  $\rho(\hat{\theta}_R, \hat{\theta}_W)$  converge to  $\rho$  almost surely: moreover the rate of convergence is computed and the allocation proportion is shown to be asymptotically normal.

To illustrate the difference with the response-adaptive designs considered in the next sections, suppose that responses are normally distributed real numbers and let  $N(\mu_R, \sigma_R^2)$  be the outcome distribution after treatment  $R$  and  $N(\mu_W, \sigma_W^2)$  be the outcome distribution after treatment  $W$ ; the parameters,  $\mu_R, \mu_W, \sigma_R^2$  and  $\sigma_W^2$  are all unknown. Consider testing the hypotheses

$$H_0 : \Delta = \mu_R - \mu_W = 0 \text{ vs. } H_1 : \Delta \neq 0.$$

For a fixed number  $n$  of patients involved in the trial, the power of the test based on the usual test statistic

$$Z_0 = \frac{\hat{\mu}_R - \hat{\mu}_W}{\sqrt{\frac{\hat{\sigma}_R^2}{\rho n} + \frac{\hat{\sigma}_W^2}{(1-\rho)n}}}$$

is maximized when the proportion of patients allocated to treatment  $R$  is equal to

$$\rho = \frac{\sigma_R}{\sigma_R + \sigma_W} \quad (\text{Neyman allocation}).$$

When patients allocation follows the DBCD with this target allocation, Hu and Zhang (2004) prove that  $n^{1/2}(N_R(n)/n - \rho)$  is asymptotically distributed as a

$$N(0, \frac{2 + \alpha}{1 + 2\alpha} \frac{\sigma_R \sigma_W}{(\sigma_R + \sigma_W)^2}).$$

The same testing problem is studied also in Zhang and Rosenberger (2006), where it is assumed that a small response is desirable and the total expected response from all  $n$  patients involved in the trial should be minimized. The target allocation  $\rho$  is found in this case by solving the following optimization problem:

$$\begin{cases} \min_{\rho} [\rho n \mu_R + (1 - \rho) n \mu_W], \\ \frac{\sigma_R^2}{\rho n} + \frac{\sigma_W^2}{(1 - \rho) n} = K, \end{cases}$$

where  $K$  is some positive constant. With the additional constraint that it is not appropriate to allocate more patients to treatment  $W$  when  $\mu_R < \mu_W$ , Zhang and Rosenberger (2006) find that the optimal target allocation is in this case

$$\rho = \begin{cases} \frac{\sigma_R \sqrt{\mu_W}}{\sigma_R \sqrt{\mu_W} + \sigma_W \sqrt{\mu_R}} & \text{if } (\mu_R < \mu_W, \sigma_R \sqrt{\mu_W} > \sigma_W \sqrt{\mu_R}) \\ \text{or } (\mu_R > \mu_W, \sigma_R \sqrt{\mu_W} < \sigma_W \sqrt{\mu_R}) \\ 1/2 & \text{otherwise.} \end{cases}$$

Again, by using the DBCD with the previous target allocation and by referring to the results in Hu and Zhang (2004), Zhang and Rosenberger (2006) prove that  $n^{1/2}(N_R(n)/n - \rho)$  is asymptotically normal with mean zero and variance a known function of the parameters  $\mu_R, \mu_W, \sigma_R^2, \sigma_W^2$  and  $\alpha$ .

#### 4. Targeting the optimal treatment

Assume that the response distributions  $\nu_R$  and  $\nu_W$  belong to a class  $\mathcal{P}$  of probability distributions on  $\mathfrak{R}^k$ . The experimenter is able to express preferences among the elements of  $\mathcal{P}$ ; the preference pattern is represented by a bounded utility function  $u : \mathfrak{R}^k \rightarrow \mathfrak{R}$  such that if  $P_1$  and  $P_2$  are in  $\mathcal{P}$  then  $P_1$  is preferred to  $P_2$  if and only if

$$\int_{\mathfrak{R}^k} u(y) P_1(dy) > \int_{\mathfrak{R}^k} u(y) P_2(dy)$$

while  $P_1$  and  $P_2$  are equivalent if and only if  $\int u dP_1 = \int u dP_2$ . Conditions which guarantee the existence of a bounded utility function  $u$  such that expected utilities of the elements of  $\mathcal{P}$  are ordered in the same way as the true preferences among the  $P \in \mathcal{P}$  can be found, for instance, in DeGroot (1970) and Fishburn (1981). It is to be noted that if  $u$  is such a utility, then  $u_1 = bu + c$ , with  $b > 0$ , is also a utility function that represents

the same preference pattern among the elements of  $\mathcal{P}$  : hence, without loss of generality, we assume that  $0 < \lambda \leq u(y) \leq v < \infty$  for all  $y \in \mathfrak{R}^k$ .

The experimenter's ethical goal is to bias allocations along the trial toward the optimal treatment, i.e. the treatment between  $R$  and  $W$  generating the response with maximum expected utility. By setting

$$\theta_R = \int u d\nu_R \text{ and } \theta_W = \int u d\nu_W,$$

we see that, in the language of the previous section, this goal corresponds to a target allocation  $\rho$  such that

$$\rho(\theta_R, \theta_W) = \begin{cases} 1 & \text{if } \theta_R > \theta_W, \\ 0 & \text{if } \theta_R < \theta_W. \end{cases}$$

The discontinuity of  $\rho$  cannot be eliminated by setting an appropriate value when  $\theta_R = \theta_W$ ; hence this target allocation does not belong to those considered in Hu and Zhang (2004).

We describe in this section two sequential designs targeting the optimal treatment. The first one, called randomly reinforced urn design or RRUD, is based on an urn scheme introduced in Muliere *et al.* (2006). This design generalizes to continuous responses the urn design studied in Durham *et al.* (1998) and in Li *et al.* (1997): an initial analysis of its performance appears in Paganoni and Secchi (2005). The second design is new and still needs to be studied; it is proposed here only as a sparring partner of the RRUD.

#### 4.1. The randomly reinforced urn design

Let  $(r_0, w_0)$  be a couple of positive real numbers. The randomly reinforced urn design  $\delta$  is defined by setting  $\delta_0 = r_0/(r_0 + w_0)$  and, for  $n \geq 1$  and  $(x_1, y_1, \dots, x_n, y_n) \in (\{0, 1\} \times \mathfrak{R}^k)^n$ ,

$$\delta_n(x_1, y_1, \dots, x_n, y_n) = \frac{r_0 + \sum_{j=1}^n x_j u(y_j)}{r_0 + w_0 + \sum_{j=1}^n u(y_j)}.$$

That is, according to  $\delta$ , given past allocations  $X_1, \dots, X_n$  and observed responses  $Y_1, \dots, Y_n$ , the experimenter allocates the next patient  $n+1$  to treatment  $R$  with probability proportional to  $r_0 + \sum_{j=1}^n X_j u(Y_j)$ . The design  $\delta$  is randomized and response-adaptive; it is implemented by an ideal urn that generalizes the classical one by Pólya. The urn initially contains  $r_0$  balls of color  $R$  and  $w_0$  balls of color  $W$ . When patient  $n = 1, 2, \dots$  enters the trial, the experimenter allocates him or her to a treatment by sampling a ball from the urn. After the treatment effect  $Y_n$  is observed, and before allocating the  $(n+1)$ -th patient, the sampled ball is returned to the urn together with  $u(Y_n)$  balls of the same color (this is called *reinforcement*).

In Muliere *et al.* (2006) it is proved that, if

$$\int u d\nu_R \neq \int u d\nu_W, \tag{1}$$

then, whatever the initial composition  $(r_0, w_0)$ , with probability one

$$\lim_{n \rightarrow \infty} \delta_n(X_1, Y_1, \dots, X_n, Y_n) = \begin{cases} 1 & \text{if } \int u d\nu_R > \int u d\nu_W, \\ 0 & \text{if } \int u d\nu_R < \int u d\nu_W; \end{cases} \tag{2}$$

that is with probability one the RRUD asymptotically assigns patients to the treatment with maximum expected utility.

Simulations show that convergence in (2) is typically slow, although the actual rate of convergence is still an open and intriguing problem; in any case, for small  $n$  the effect of the initial composition  $(r_0, w_0)$  on the allocation probabilities will be highly relevant. For a Bayesian, different choices of  $r_0$  and  $w_0$  may incorporate different prior opinions about the preferred treatment. A frequentist believing in initial equipoise, may choose  $r_0$  and  $w_0$  both equal to the same nonnegative number; an alternative is to initially allocate at random an equal number  $k$  of patients to treatment  $R$  and to treatment  $W$ , observe the responses  $Y_1, \dots, Y_{2k}$ , and then allocate patients, from the  $(2k + 1)$ -th on, according to a RRUD with initial composition

$$r_0 = \sum_{i=1}^{2k} X_i u(Y_i), \quad w_0 = \sum_{i=1}^{2k} (1 - X_i) u(Y_i) \quad (3)$$

where, for  $i = 1, \dots, 2k$ ,  $X_i$  indicates treatment allocation for patient  $i$ . In Paganoni and Secchi (2005) the integer number  $k$  is called the *initialization parameter* and it is set equal to a small fraction of the actual trial sample size.

The case when  $\int u d\nu_R = \int u d\nu_W$  is still partially unknown. The situation at the opposite extreme of (1) when  $\int u^m d\nu_R = \int u^m d\nu_W$  for all  $m = 1, 2, \dots$ , i.e. when the reinforcement distributions  $\nu_R u^{-1}$  and  $\nu_W u^{-1}$  are the same, has been studied in May *et al.* (2005) where it is proved that

$$P[\lim_{n \rightarrow \infty} \delta_n(X_1, Y_1, \dots, X_n, Y_n) = Z_\infty] = 1 \quad (4)$$

with  $Z_\infty$  a random variable. The distribution  $F(r_0, w_0)$  of  $Z_\infty$  has support equal to  $[0, 1]$ , no point masses and expected value equal to  $r_0(r_0 + w_0)^{-1}$ ; albeit its analytical form is still unknown, its functional dependency on the initial parameters  $(r_0, w_0)$  is analyzed in Aletti *et al.* (2005) where it is also proved that  $F(r_0, w_0)$  is the unique solution satisfying some regularity conditions of a functional equation involving unknown probability distributions on  $[0, 1]$ .

An interesting Bayesian interpretation of the RRUD is given in Muliere *et al.* (2006). Suppose that the response probability distributions  $\nu_R$  and  $\nu_W$  are random and assume them to be independent Dirichlet processes with parameters  $\alpha_R$  and  $\alpha_W$  respectively ( $\alpha_R$  and  $\alpha_W$  are known finite measures on  $\mathfrak{R}^k$ .) If the urn initial composition is

$$r_0 = \int_{\mathfrak{R}^k} u(y) \alpha_R(dy), \quad w_0 = \int_{\mathfrak{R}^k} u(y) \alpha_W(dy),$$

then, for  $n \geq 1$ ,

$$\delta_n(X_1, Y_1, \dots, X_n, Y_n) = \frac{\alpha_R(\mathfrak{R}^k) + N_R(n)}{\int u d\alpha_R + \int u d\alpha_W + \sum_{i=1}^n u(Y_i)} E\left[\int u d\nu_R | \mathcal{F}_n\right].$$

Hence the experimenter may consider the RRUD as a randomizing device that allocates the next patient  $n + 1$  to treatment  $R$  according to a probability proportional to the predictive utility of the response generated by  $R$  weighted by the “total” number of patients who experienced treatment  $R$  in the past. (We are here interpreting the prior information contained in  $\alpha_R(\mathfrak{R}^k)$  and  $\alpha_W(\mathfrak{R}^k)$  as if these were sample sizes weighting the experimenter’s prior expected values of  $\nu_R$  and  $\nu_W$  respectively.)

## 4.2. The mean difference mapping design

The RRUD is not the unique sequential randomized design targeting the optimal treatment. A competitor, inspired by Bandyopadhyay and Biswas (2001) and Biswas and Basu (2001), is constructed by selecting a continuous distribution function  $G$  such that  $0 < G(x) < 1$  for all  $x \in \mathfrak{R}$  and  $G$  is symmetric in 0, i.e.  $G(x) + G(-x) = 1$  for all  $x \in \mathfrak{R}$ . The idea generating the mean difference mapping design (MDMD) is to obtain the allocation probability for the next patient by transforming with  $G$  the standardized difference between sample estimates of the expected utilities of the response probability distributions.

Here is an implementation of the MDMD: let  $k \geq 2$  be the initialization parameter and indicate with  $\delta$  the design that initially allocates at random an equal number  $k$  of patients to treatment  $R$  and to treatment  $W$ , as in (3). For  $n \geq 2k + 1$ , given  $\mathcal{F}_n$  the design  $\delta$  assigns patient  $n + 1$  to treatment  $R$  with conditional probability

$$\delta_n(X_1, Y_1, \dots, X_n, Y_n) = G\left(\frac{\hat{u}_R(n) - \hat{u}_W(n)}{\sqrt{\frac{\hat{\sigma}_R^2(n)}{N_R(n)} + \frac{\hat{\sigma}_W^2(n)}{N_W(n)}}}\right),$$

where

$$\begin{aligned}\hat{u}_R(n) &= \frac{1}{N_R(n)} \sum_{i=1}^n X_i u(Y_i), \\ \hat{u}_W(n) &= \frac{1}{N_W(n)} \sum_{i=1}^n (1 - X_i) u(Y_i), \\ \hat{\sigma}_R^2(n) &= \frac{1}{N_R(n) - 1} \sum_{i=1}^n X_i (u(Y_i) - \hat{u}_R(n))^2, \\ \hat{\sigma}_W^2(n) &= \frac{1}{N_W(n) - 1} \sum_{i=1}^n (1 - X_i) (u(Y_i) - \hat{u}_W(n))^2,\end{aligned}$$

are estimators of the means  $u_R = \int u d\nu_R$  and  $u_W = \int u d\nu_W$  and the variances  $\int (u - u_R)^2 d\nu_R$  and  $\int (u - u_W)^2 d\nu_W$  respectively. (Of course, if, for instance, the variances of  $\nu_R u^{-1}$  and  $\nu_W u^{-1}$  are known, their values are used instead of the estimators  $\hat{\sigma}_R^2$  and  $\hat{\sigma}_W^2$  and the initialization parameter  $k$  can be set equal to 1.)

In applications, obvious choices for  $G$  are the inverse of a probit or the inverse of a logit: that is,  $G$  may be set equal to the standard normal distribution or the logistic distribution

$$G(x) = \frac{\exp(bx)}{1 + \exp(bx)}, \quad (5)$$

for  $x \in \mathfrak{R}$ , with  $b > 0$ . In the latter case a unit increase in the standardized difference between the sample estimates  $\hat{u}_R(n)$  and  $\hat{u}_W(n)$  corresponds to an increase of  $b$  units in the log-odds of the allocation probabilities.

From general results in Melfi and Page (2000) on consistency of estimators based on samples selected by means of a response-adaptive design, it follows that  $\hat{u}_R(n)$ ,  $\hat{u}_W(n)$ ,  $\hat{\sigma}_R^2(n)$  and  $\hat{\sigma}_W^2(n)$  are consistent. Moreover, if (1) holds, then (2) is true also for the MDMD: this claim is supported by simulations and will be the object of future research.

## 5. Evaluating the performance of a response-adaptive design

At the end of a clinical trial the experimenter compares response distributions, in most cases by means of a test. The extra variability introduced in the trial by the response-adaptive design has a great impact on power; within the context of optimal allocation, for binary or normal responses Hu and Rosenberger (2003) and Hu and Zhang (2004) provide an explicit relationship between the average power of the test and the variance of the proportion  $N_R(n)/n$  targeting a given allocation proportion  $\rho \in (0, 1)$ . When the asymptotic distribution of  $N_R(n)/n$  is known, different procedures targeting the same  $\rho$  are ranked based on their smaller asymptotic variability. Inspired by this work, in Paganoni and Secchi (2005) a guideline is proposed for evaluating and comparing different response-adaptive designs targeting the optimal treatment.

When considering a response-adaptive design for a particular clinical trial, a standard, non-adaptive alternative is usually available; this is the *default design* that the experimenter would implement for carrying on the trial in the absence of a response-adaptive competitor, to be followed by a *default inferential analysis* applied to the data generated by the design. Default design and default inferential analysis represent the experimenter's *default plan*. The experimenter might be persuaded to use a response-adaptive design, instead of the default design, if:

- (i) the response-adaptive design makes it possible to perform an inferential analysis with the same optimality characteristics as those guaranteed by the default plan;

and

- (ii) the number of patients allocated in the trial to the worse treatment by the response-adaptive design is less than that provided by the default design.

Informed by (i) and (ii), a guideline for the evaluation and comparison of a response-adaptive design focuses on a *benchmark*, a basic inferential problem often encountered in practice, and bases the design's evaluation on the conditions for which requirements (i) and (ii) are simultaneously satisfied. For instance, consider a univariate normal and homoscedastic model with variance known for responses after treatment and assume that the trial's goal is to test

$$H_0 : \Delta = \mu_R - \mu_W = 0 \text{ vs. } H_1 : \Delta > 0 \quad (6)$$

where  $\mu_R$  and  $\mu_W$  are the unknown expected values of the response distributions after treatment  $R$  and  $W$  respectively. The default plan is to randomly allocate  $n_R = n/2 = n_W$  patients to the two treatments (Neyman allocation) and, at the end of the experiment, to perform a one-sided z-test. Assign the level  $\alpha$  of the test and choose  $n$  so that the balanced, one-sided z-test has a given power  $1 - \beta$  when  $\Delta$  is greater or equal to a specific, clinically relevant difference  $\Delta_0 > 0$ . To convince the experimenter to switch from the default design to a competitor response-adaptive design, one needs:

- (i') to elicit an  $\alpha$ -level test, function of the experimental data generated by the response-adaptive design, for proving  $H_0$  versus  $H_1$ . For a trial sample size  $n^*$ , the power of the test must be at least  $1 - \beta$  when  $\Delta \geq \Delta_0$ ;



(ii') to show that, when the trial sample size is  $n^*$  and  $\Delta \geq \Delta_0$ , the random number  $N_W(n^*)$  of patients allocated to treatment  $W$  is less than or equal to  $n/2$ , with high probability.

If (i') and (ii') hold simultaneously, the experimenter adopting the response-adaptive design, knows that the probability that a significant result will be obtained if a clinically relevant difference between the two treatments exists (i.e. the power of the test), is not less than the power of the test in the default plan at the smallest clinically relevant difference; this might happen at the cost of a number  $n^*$  of patients in the trial greater than or equal to  $n$ , but with the assurance that, if there is a clinically relevant difference between treatments, with high probability less patients than those allocated by the default design will in fact experience the worse treatment. In the next section we will expand on this example by means of simulations.

## 6. Numerical illustrations

To illustrate the guideline sketched in the previous section we conducted a simulation study for comparing the RRUD and the MDMD when the benchmark is the test of hypotheses described in (6). Hence responses follow two univariate normal distributions with unknown means  $\mu_R$  and  $\mu_W$  respectively and variances equal and known:  $\sigma_R^2 = \sigma_W^2 = (0.25)^2$ . For generating the simulated responses we set the mean  $\mu_W = 1$  and the mean  $\mu_R = 1 + \Delta$ , with  $\Delta \in [0, 0.8]$ . Higher responses are preferred and the utility function is assumed to be

$$u(x) = \begin{cases} 0.1, & \text{if } x \leq 0.1, \\ x, & \text{if } 0.1 < x < 10, \\ 10, & \text{if } x \geq 10; \end{cases}$$

i.e.  $u$  is virtually the identity.

We compared three different designs targeting the optimal treatment, the RRUD with initialization parameter  $k = 3$  and two logistic MDMDs: for both designs the initialization parameter  $k = 3$  and the function  $G$  is the logistic distribution represented in (5). For the first MDMD the parameter  $b = 0.915$  while the second has  $b = 1.83$ : these parameters have been selected after noticing that for  $b = 0.915$  the MDMD allocates a patient to treatment  $R$  with probability 0.975 when the standardized difference between sample utilities is close to 4, while for  $b = 1.83$  the MDMD allocates a patient to treatment  $R$  with probability equal to 0.975 when the same difference is close to 2.

For  $\Delta > 0$ , simulated responses after treatment  $R$  have greater expected utility: hence the allocation probabilities for both the RRUD and the MDMDs converge to 1 as  $n$  grows to infinity, as well as the proportion  $N_R(n)/n$  of patients allocated to the better treatment  $R$ . This is illustrated by the two pictures in the upper part of Figure 1 where, for  $\Delta = 0.2, 0.6$  and increasing values of  $n$ , the quartiles of the distribution of  $N_R(n)/n$  are plotted. In point of fact, for different  $n$ , a sample of 1000 values of  $N_R(n)/n$  has been obtained by simulations and the sample quartiles have been computed: the curves appearing in the pictures are *loess* regressions of the sample quartiles on the sample size  $n$ . (For details on *loess* regression see Cleveland *et al.* (1992).) Note how both MDMDs tend to allocate to the preferred treatment a larger proportion of patients than the RRUD, even for small sample sizes  $n$ ; moreover the variability of the distribution of  $N_R(n)/n$

is smaller when the MDMDs are in force. Not surprisingly the MDMD with higher parameter  $b$  is quicker to unbalance allocations in favor of  $R$ .

Next we compared the three designs according to the first analysis in Paganoni and Secchi (2005). Hence assume that the default plan is a balanced design with sample size  $n = 50$  followed by a one-sided z-test of level  $\alpha = 0.05$  for testing the hypotheses in (6). As a function of the difference  $\Delta = \mu_R - \mu_W$ , let  $1 - \beta(\Delta)$  be the power function of this z-test. Then, for each response-adaptive design and for different values of  $\Delta \in [0, 0.8]$ , we computed through simulation (1000 replications for each different value of  $\Delta$ ) the smallest trial sample size  $n^* = n^*(\Delta)$  such that a given 0.05-level test has power greater than or equal to the power  $1 - \beta(\Delta)$ ; in fact, for all three response-adaptive designs we used a one-sided z-test based on the test statistic:

$$Z_0(n^*) = \frac{\hat{u}_R(n^*) - \hat{u}_W(n^*)}{\sqrt{(0.25)^2 \left( \frac{1}{N_R(n^*)} + \frac{1}{N_W(n^*)} \right)}}.$$

When  $H_0$  is true,  $Z_0(n^*)$  is asymptotically standard normal: the claim has been proved for the RRUD in Paganoni and Secchi (2005) while it is supported by simulations for the two MDMDs considered in this section. Finally we obtained by simulation the distribution of the random number  $N_W$  of patients allocated to the worse treatment  $W$  by the three response-adaptive designs when the trial sample size is  $n^* = n^*(\Delta)$ .

The values of  $n^* = n^*(\Delta)$  for the three response-adaptive designs are shown in the bottom-left picture of Figure 1: *loess* regressions are used to smooth the data points. The quartiles of the distribution of  $N_W(n^*(\Delta))$  for the three designs appear in the bottom-right picture of Figure 1: again, for different values of  $\Delta$  a sample of size 1000 from the distribution of  $N_W(n^*(\Delta))$  has been generated by simulation and the sample quartiles curves have been smoothed by *loess* regressions.

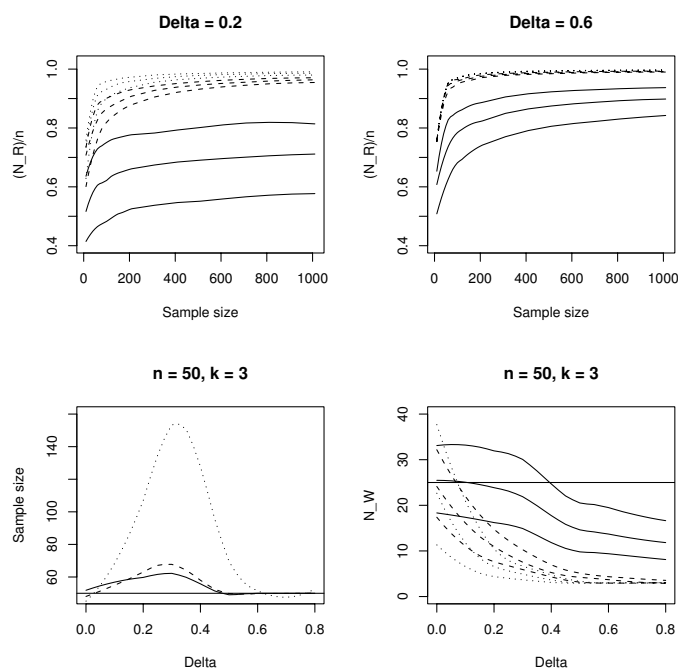
Inspection of the two pictures at the bottom of Figure 1 show the existence, for all three response-adaptive designs, of three different regions of  $\Delta$  values, as predicted in Paganoni and Secchi (2005). For small values of  $\Delta$ ,  $n^*$  is larger than  $n$  but  $N_W$  is not smaller than  $n/2$  with high probability: this is the “red zone” where condition (i') of the guideline sketched in section 5 is met, but not condition (ii'). That is: in order to get the same power as that of the default plan with sample size  $n$ , the response-adaptive design needs a sample size  $n^* > n$ , but the higher cost due to a larger sample size is not compensated by a gain in terms of less patients allocated to the worse treatment. Moderate values of  $\Delta$  fall in the “yellow zone”: (i') and (ii') are met at the cost of a larger sample size  $n^*$  for the response-adaptive design. Finally, large values of  $\Delta$  belong to the “green zone”: (i') and (ii') are satisfied and  $n^*$  is not significantly larger than  $n$ .

If the experimenter believes that values of  $\Delta$  in the “red zone” are clinically relevant, he shouldn't exchange the default design for the response-adaptive design. If the smallest clinically relevant value  $\Delta_0$  falls in the “yellow zone”, the experimenter might switch from the default design to the response-adaptive design, at a cost of a larger trial sample size. Finally, when  $\Delta_0$  falls in the “green zone” it seems unreasonable not to use the response-adaptive design. In fact, a qualitative estimate of the yellow zone based on inspection of the *loess* curves for the three response-adaptive designs is:

Design	Yellow zone
RRUD	(0.4,0.45)
MDMD, $b = 0.915$	(0.1,0.5)
MDMD, $b = 1.83$	(0.1,0.6)

For this example, the previous analysis seems to favor the logistic MDMD with  $b = 0.915$  : its zones are preferable to those of the RRUD and comparable to those of the other MDMD with  $b = 1.83$ . However the costs in terms of sample size  $n^*$  of the MDMD with  $b = 1.83$  are much higher than those of the MDMD with  $b = 0.915$  when  $\Delta$  belongs to the yellow zone. This conclusion raises the MDMD to a role more significant than simple sparring partner of the RRUD and forces us to study its properties in more detail in the future.

**Figure 1:** Graphic comparison of three response-adaptive design: the RRUD with  $k = 3$  (continuous lines), the logistic MDMD with  $k = 3$  and  $b = 0.915$  (dashed line) and the logistic MDMD with  $k = 3$  and  $b = 1.83$  (point line) .



## References

- Aletti G., May C., Secchi P. (2005). On the distribution of the limit proportion for a two-color, randomly reinforced urn with equal reinforcement distributions, *MOX report no. 74*, MOX-Dipartimento di Matematica, Politecnico di Milano.
- Atkinson A.C., Biswas A. (2005). Adaptive biased coin designs for skewing allocation proportions in clinical trials with normal proportions, *Statistics in Medicine*, 24, 2477-2492
- Baldi Antognini A., Giovagnoli A. (2005). On the large sample optimality of sequential designs for comparing two or more treatments, *Sequential Analysis*, 24, 205-217.

- Bandyopadhyay U., Biswas A. (2001). Adaptive designs for normal responses with prognostic factors. *Biometrika*, 88, 409-419.
- Biswas A., Basu A. (2001). Robust adaptive designs in clinical trials for continuous responses, *Sankhya*, Series B, 63, 27-42.
- Cleveland W. S., Grosse E., Shyu W.M. (1992). Local Regression Models. In *Statistical Models in S*, Chambers J., Hastie T. eds., Wadsworth & Brooks/Cole.
- DeGroot M. H. (1970). *Optimal Statistical Decisions*, McGraw-Hill, New York.
- Durham S.C., Flournoy N., Li W. (1998). A sequential design for maximizing the probability of a response, *The Canadian Journal of Statistics*, 26(3), 479-495.
- Eisele J. (1994). The doubly adaptive biased coin design for sequential clinical trials, *Journal of Statistical Planning and Inference*, 38, 249-261.
- Fishburn P.C. (1981). Subjective expected utility: a review of normative theories, *Theory and Decision*, 13, 139-199.
- Hu F., Rosenberger W.F. (2003). Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons, *Journal of the American Statistical Association*, 98, 671-678.
- Hu F., Zhang L. (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials, *The Annals of Statistics*, 32, 268-301.
- Li W., Durham S.D., Flournoy N. (1997). Randomized Pólya urn design, *Proceedings of the Biometric Section of the American Statistical Association*, 166-170.
- May C., Paganoni A.M., Secchi P. (2005). On a two-color generalized Polya urn, *Metron*, LXIII(1), 115-134.
- Melfi V.F., Page, C. (2000). Estimation after adaptive allocation, *Journal of Statistical Planning and Inference*, 87, 353-363.
- Melfi V.F., Page C., Geraldes M. (2001). An adaptive randomized design with application to estimation, *The Canadian Journal of Statistics*, 29(1), 107-116.
- Muliere P., Paganoni A.M., Secchi P. (2006). A randomly reinforced urn, *Journal of Statistical Planning and Inference*, 136, 1853-1874.
- Paganoni A.M., Secchi P. (2005). A numerical study for comparing two response-adaptive designs for continuous treatment effects, *MOX report no. 65*, MOX-Dipartimento di Matematica, Politecnico di Milano.
- Rosenberger, W.F. (2002). Randomized urn models and sequential design, *Sequential Analysis*, 21 (1&2), 1-28.
- Rosenberger W.F., Lachin J.M. (2002). *Randomization in clinical trials: theory and practice*, Wiley, New York.
- Zhang L., Rosenberger F. (2006). Response-adaptive randomization for clinical trials with continuous responses, To appear in *Biometrics*.