MOX–Report No. 38/2010

Response-adaptive designs in clinical trials for targeting the best treatment: an overview

Nancy Flournoy, Caterina May, Piercesare Secchi

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Response-adaptive designs in clinical trials for targeting the best treatment: an overview

Nancy Flournoy[⋆], Caterina May[♮] and Piercesare Secchi[♯]

October 18, 2010

[⋆]Department of Statistics
University of Missouri - Columbia
Columbia, MO 65211, USA
flournoyn@missouri.edu

[♮]Dipartimento SEMEQ
Università del Piemonte Orientale
via Perrone 18, 28100 Novara, Italy
caterina.may@eco.unipmn.it

[♯] MOX–Dipartimento di Matematica
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
piercesare.secchi@polimi.it

Keywords: Response-adaptive designs, Clinical trials, Optimal allocation, Randomly reinforced urns, Ethical allocation, Experimental Design, Ethical allocation, Sequential design, Migrating urn

Abstract

Response-adaptive designs are increasingly being implemented in clinical trials, particularly early phase trials, and they have increasingly stimulated the work of researchers. This paper reviews a particular class of response-adaptive designs, which have a different property from the most adaptive designs in literature. These are response-adaptive designs targeting asymptotically the superior response, that is, treating with the superior treatment with probability converging to one. The model underlying such designs is a randomly reinforced urn. In the context of clinical trials, this property is particularly attractive from an ethical point of view. This overview starts from the early paper of [8] until the recent work by [9].

1 Introduction

In recent years, there has been an increasing interest, in the context of clinical trials research, in response-adaptive designs. This is because response-adaptive designs are sequential procedures that can skew, along the experiment, the allocation probabilities of statistical units on the base of previous allocations and responses. In a clinical trial to compare two or more treatments, the experimenter faces two simultaneous goals: collecting evidence to determine the superior treatment, and skewing the allocations toward the superior treatment in order to reduce the proportion of patients that receive the worst treatment. The first is an inferential goal and concerns future patients' interest; the second is an ethical responsibility and concerns the current study patients' interest.

An informed review about the theory of response-adaptive designs can be found in [23] and in [12]. A particular role in response-adaptive designs has been played by urn models. As addressed in the review work of [22], urn models are particularly attractive for clinical experiments because they guarantee the randomization of allocations.

In this review paper, we focus on a particular urn model that we call the Randomly Reinforced Urn (RRU). The RRU for binary experiments (success/failure) was introduced by [8] and [14] as a modification of the randomized-play-the-winner scheme, and applied to select an optimal dosage in [7]; further, it has been extended to experiments with general responses by [18]. It is important to know that the RRU model differentiates from the generalized Polya urn (GPU) described in [22] and in the references therein, because its reinforcement matrix is not irreducible. This gives the RRU-designs different properties from the other urn designs. In a different context, the RRU has an interesting potential for applications since it can be used to describe a general model for reinforcement learning ([10] and [4]).

Most of response-adaptive designs presented in literature allocate patients targeting asymptotically a certain proportion $\rho \in (0, 1)$, which may be ad hoc or may be determined by some optimality criteria that are usually a function of the unknown parameters of the outcomes and has to be estimated. The RRU-designs are different because they have the ethically optimal property of assigning patients to the best treatment with a proportion that converges almost surely to 1, while the proportion of patients allocated to the inferior treatment converges to 0.

In this work, we provide an overview of response-adaptive designs generated by a RRU, with particular attention to recent developments and open problems. We focus on the case of $K = 2$ treatments, since the

results can be extended straightforwardly to more then two treatments; further comments about the $K > 2$ case are postponed in Section 7. Hopefully, this review will stimulate a further developments in this important area of research.

The paper is organized as follows: after the presentation of the model, we motivate its application in clinical trials. Then, we review the asymptotic results and we concentrate on problems and results concerning inference. Finally, we present simulation studies and some concluding comments.

2 Model and designs

An experiment is conducted to compare two treatments, say B and W. Patients enter the experiment sequentially and are allocated randomly to a treatment according to a RRU-design, whose model can be described in the following way.

Let $\{(Y_B(n), Y_W(n)) : n \geq 1\}$ be a sequence of independent and identically distributed random response vectors with marginal distributions $\mathcal{L}_B$ and $\mathcal{L}_W$, discrete or continuous on $\mathbb{R}$. Only one response, $Y_B(n)$ or $Y_W(n)$, will be observed for each subject n depending on their treatment assignment. Consider an urn containing initially b black balls and w white balls, where b and w are two strictly positive real numbers. With the arrival of the first patient ($n = 1$), a ball is drawn at random from the urn and its color is observed: we define a random variable δ_1 that we assume to be independent of the potential response vector $(Y_B(i), Y_W(i))$ for every $i \geq 1$ such that $\delta_1 = 1$ if the extracted ball is black; while $\delta_1 = 0$ if the extracted ball is white. So δ_1 is a Bernoulli random variable with parameter $Z_0 = b/(b + w)$. After the ball is extracted, if it is black, it is replaced in the urn together with $U(Y_B(1))$ black balls. Otherwise, if it is white, is replaced in the urn together with $U(Y_W(1))$ white balls, where U is an arbitrary measurable function such that $U(Y_B(1))$ and $U(Y_W(1))$ have distributions on a nonnegative and bounded real set. In typical applications, U will be a monotone function. (Note that U can be the identity function when the distributions $\mathcal{L}_B$ and $\mathcal{L}_W$ have nonnegative and bounded support).

This process is then iterated for each new patient $n + 1$, $n \geq 1$: a ball is extracted and δ_{n+1} indicates its color: $\delta_{n+1} = 1$ if the ball extracted is black, while $\delta_{n+1} = 0$ if the ball extracted is white. We always assume that δ_{n+1} is independent of the potential response vector $(Y_B(i), Y_W(i))$ for every $i \geq n+1$. After the ball is extracted, it is replaced in the urn together

with

$$\delta_{n+1}U(Y_B(n+1)) + (1-\delta_{n+1})U(Y_W(n+1))$$

balls of the same color. So, given the σ-algebra

$$\mathcal{F}_n = \sigma\left(\delta_1, \delta_1 Y_B(1) + (1-\delta_1)Y_W(1), \ldots, \delta_n, \delta_n Y_B(n) + (1-\delta_n)Y_W(n)\right),$$

δ_{n+1} is conditionally Bernoulli distributed with parameter

$$Z_n = \frac{B_n}{B_n + W_n},$$

where

$$\begin{cases} B_n = b + \sum_{i=1}^{n} \delta_i U(Y_B(i)) \\ W_n = w + \sum_{i=1}^{n} (1-\delta_i) U(Y_W(i)). \end{cases}$$

The *RRU* procedure drives the allocations $\{\delta_n\}$: when δ_n is 1, allocate the n-th patient to the first treatment, say treatment B, and let the random variable $Y_B(n)$ be the potential response of n-th patient to treatment B; when δ_n is 0, allocate the n-th patient to the second treatment, say treatment W, and let $Y_W(n)$ be the potential response of n-th patient to treatment W. The one response for the n-th patient that is actually observed can be written as $Y(n) = \delta_n Y_B(n) + (1-\delta_n)Y_W(n)$.

Thus the *RRU* generates the following processes: the sequence $\{\delta_n : n \geq 1\}$ of conditionally Bernoulli random variables, corresponding to the treatment allocations, and the sequence $\{Z_n : n \geq 0\}$ of random variables in $[0,1]$ representing the proportion of black balls present in the urn at every stage. Now the total number of patients that have been assigned to treatment B and to treatment W through the nth treatment allocation can be written as $N_B(n) = \sum_{i=1}^{n} \delta_i$ and $N_W(n) = \sum_{i=1}^{n}(1-\delta_i)$, respectively; clearly $N_B(n) + N_W(n) = n$. Also note that B_n and W_n are the cumulative (possibly transformed) observed responses to treatment B and W, respectively, augmented by the initial numbers of balls in the urn.

3 Motivation for applying a *RRU*-design to clinical trials

Let $m_B = \int U(y)\mathcal{L}_B(dy)$ and $m_W = \int U(y)\mathcal{L}_W(dy)$ be the means of the transformed responses, that is, the urn reinforcements; a fundamental result, proved by [14] for dichotomous responses and extended to general responses by [18] and [2], is that

$$\text{if } m_B > m_W, \text{ then } \lim_{n\to\infty} Z_n = 1, \text{almost surely.} \tag{1}$$

As a consequence, suppose that the sequences of responses $\{Y_B(n)\}$ and $\{Y_W(n)\}$ have finite means $\mu_B = \int y\mathcal{L}_B(dy)$ and $\mu_W = \int y\mathcal{L}_W(dy)$ and that, for instance, the treatment B is preferred to the treatment W if $\mu_B > \mu_W$. Then, choosing a function U such that $\mu_B > \mu_W$ if and only if $m_B > m_W$ and $\mu_B = \mu_W$ if and only if $m_B = m_W$, it follows that a RRU-design allocates patients to the superior treatment with probability converging to one as n goes to infinity. This is the reason why these designs are optimal from an ethical point of view.

As argued in [19], the existence of a suitable function U is guaranteed from the theory of utility. Different choices of the function U imply different properties for the RRU-design, in terms of rate of convergence and skewness of allocations. More needs to be done to identify good choices of U. This choice will be influenced by the trade-off between inferential and ethical goals that is considered preferable for a particular experiment.

Note that, notwithstanding the fact that in a RRU-design the conditional probability of allocating the nth patient to the inferior treatment W converges to zero, both the random numbers of patients

$$N_B(n) \text{ and } N_W(n) \text{ converge almost surely to infinity,} \tag{2}$$

as the total number n of patients goes to infinity. This is a basic result, proved in [15], allowing their development of inferential procedures for RRU-designs.

Also, when the two treatments are equivalent, that is, when $m_B = m_W$, the sequence of proportions

$$\{Z_n : n \geq 0\} \text{ converges almost surely to a random limit } Z_\infty \text{ in } [0,1],$$

since it is eventually a bounded super or sub-martingale, as proved by [18]. An important property, proved by [2], is that

$$P(Z_\infty = p) = 0 \quad \text{for all } p \in [0,1] \tag{3}$$

in this case; while on the contrary, if $m_B > m_W$, then $P(Z_\infty = 1) = 1$ from (1). The distribution of Z_∞ when $m_B = m_W$ is in general unknown and it is a non-trivial, open problem; further discussion of this is postponed to Section 4.1.

We remark also that, both when treatments are equivalent and when one treatment is superior, the proportion of patients allocated to B and W has the same limit as the urn composition:

$$\lim_{n\to\infty} \frac{N_B(n)}{n} = Z_\infty, \quad a.s. \quad \text{and} \quad \lim_{n\to\infty} \frac{N_W(n)}{n} = 1 - Z_\infty, \quad a.s., \tag{4}$$

as proved in [15]. As a consequence, also the proportion of patients allocated to the best treatment converges to one.

4 Asymptotic results

[15] generalize [7] in proving the following asymptotic results. The first one concerns the relative convergence rates of the cumulative responses:

$B_n/({W_n}^{m_B/m_W})$ *converges almost surely to a random variable* ψ *with support in* $(0, \infty)$.

The following two regard the case when treatment B is superior to treatment W: if $m_B > m_W$, then there exist a random variable η^2 with $P(0 < \eta^2 < \infty) = 1$ such that

$$\lim_{n \to +\infty} \frac{N_W(n)}{n^{m_W/m_B}} = \eta^2, \quad a.s., \tag{5}$$

$$\lim_{n \to +\infty} \frac{1 - Z_n}{n^{m_W/m_B - 1}} = \frac{m_W}{m_B}\, \eta^2, \quad a.s. \tag{6}$$

Since η^2 is an almost sure finite random variable with no mass at zero, from (5) we obtain the exact rate of convergence to infinity of the number of patients assigned to the worst treatment, while from (1) and (4) we know that this rate for the number of patients allocated to B is n.

Simulations on the distribution of η^2 are provided in [16] and in [13]; in particular, it is shown that η^2 is not a point mass.

4.1 The limiting urn composition when treatments are equivalent

The study of the distribution of the limit urn composition Z_∞ in a *RRU*-design when treatments are equivalent, motivated by applications, has resulted several theoretical works. In fact, for testing hypotheses about treatment effects, it is important to know the limiting urn composition under the null hypothesis that there is no difference between treatment effects.

The exact distribution of Z_∞ is unknown except in a few particular cases. Consider first the case in which $\mathcal{L}_B = \mathcal{L}_W$ so that the urn reinforcements $U(Y_B(n))$ and $U(Y_W(n))$ have the same distribution, say μ. When μ is a point mass at a non-negative real number m, the *RRU* degenerates to Polya's urn, in which case Z_∞ has a Beta$(b/m, w/m)$ distribution. This is also the case for binary responses (success/failure) when m balls are added to the urn after each success is obtained; in fact, [1] prove that only the non null part of the reinforcement distribution needs to be considered.

[17] have proved, in the case with $\mathcal{L}_B = \mathcal{L}_W = \mu$, that the sequence $\{\delta_n\}$ of treatment allocations is asymptotically exchangeable, having De

Finetti measure equal to the distribution of Z_∞; [18] have extended this result to the general case of $m_B = m_W$. Notice that, when $m_B = m_W$, it may happen that $\int U(y)^k \mathcal{L}_B(dy) \neq \int U(y)^k \mathcal{L}_W(dy)$ for some $k \geq 2$ and then $U(Y_B(n))$ and $U(Y_W(n))$ may have different distributions; this is of particular interest because it corresponds to a situation in which the two treatments are considered equivalent but reinforcement distributions are not the same.

For the general RRU, [1] and [3] have characterized the distribution of Z_∞ as the unique continuous solution, satisfying some boundary conditions, of a specific functional equation in which the unknowns are distribution functions on $[0, 1]$.

[2] have proved a central limit theorem for the sequence of random compositions when the means of the reinforcement distributions are the same. As a consequence, they have been able to show that the limiting urn composition has no point masses. This gives also a new drive to the open problem concerning the absolute continuity of the distribution of Z_∞.

5 Inference

Consider estimation of the means $\{\mu_B, \mu_W\}$ and the variances $\{\sigma_B^2, \sigma_W^2\}$ of the responses to treatments. In [15], the following estimators based on the observed responses through patient n are defined, with random sample sizes $N_B(n)$ and $N_W(n)$, respectively:

$$\hat{Y}_B(n) = \frac{\sum_{i=1}^n \delta_i Y_B(i)}{N_B(n)} \quad \text{and} \quad \hat{Y}_W(n) = \frac{\sum_{i=1}^n (1-\delta_i) Y_W(i)}{N_W(n)} \ ,$$

$$\hat{\sigma}_B^2(n) = \frac{\sum_{i=1}^n \delta_i (Y_B(i) - \hat{Y}_B(n))^2}{N_B(n)} \quad \text{and} \quad \hat{\sigma}_W^2(n) = \frac{\sum_{i=1}^n (1-\delta_i)(Y_W(i) - \hat{Y}_B(n))^2}{N_W(n)} \ .$$

[5] prove independence properties of the sequences of observed responses, so that the strong consistency of estimators based on those sequences can be deduced:

$\hat{Y}_B(n)$, $\hat{Y}_W(n)$ and $\hat{\sigma}_B^2(n)$, $\hat{\sigma}_W^2(n)$ are strong consistent estimators for μ_B, μ_W and σ_B^2, σ_W^2, respectively.

Moreover, these estimators, appropriately standardized, are jointly asymptotically normal, both when the two treatments are equivalent and when one treatment is superior; this holds despite the randomness of $N_B(n)$ and

$N_W(n)$, their dependence, and the fact that they don't satisfy the classical assumption that $N_B(n)/n$ and $N_W(n)/n$ converge in probability to a constant in $(0,1)$:

$$\left(\frac{\sqrt{N_B(n)}}{\sigma_B} \left(\hat{Y}_B(n) - \mu_B \right), \frac{\sqrt{N_W(n)}}{\sigma_W} \left(\hat{Y}_W(n) - \mu_W \right) \right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad \text{(mixing)}.$$

The *mixing* property of the convergence (for details see [15]), is essential to obtain the asymptotic distribution of the classical two-sample t-test statistic applied to RRU data; it can be derived from the following asymptotic result:

$$\zeta(n) := \frac{\hat{Y}_B(n) - \hat{Y}_W(n) - (\mu_B - \mu_W)}{\sqrt{{\sigma_B}^2/N_B(n) + {\sigma_W}^2/N_W(n)}} \xrightarrow{d} N(0,1) \quad \text{(mixing)}, \tag{7}$$

both when $\mu_B = \mu_W$ and $\mu_B \neq \mu_W$. Result (7) holds also with $\hat{\sigma}_B^2(n)$ and $\hat{\sigma}_W^2(n)$ instead of σ_B^2 and σ_W^2.

5.1 Testing hypothesis

Suppose interest is in the hypothesis test of the mean responses:

$$H_0 : \mu_B = \mu_B \quad \text{versus} \quad H_1 : \mu_B > \mu_W,$$

which is equivalent, for a suitable choice of the function U, to the test on the mean reinforcements: $H_0 : m_B = m_B$ versus $H_1 : m_B > m_W$. From (7), it follows that the following natural extension of the t-test statistic

$$\zeta_0(n) = \frac{\hat{Y}_B(n) - \hat{Y}_W(n)}{\sqrt{\dfrac{\hat{\sigma}_B^2(n)}{N_B(n)} + \dfrac{\hat{\sigma}_W^2(n)}{N_W(n)}}} \tag{8}$$

is asymptotically normal when H_0 is true. Hence, fixing an asymptotic significance level α and denoting by $z_{1-\alpha}$ the quantile of order $1-\alpha$ of a standard normal distribution, one can consider the following critical region:

$$C_\alpha = \{\zeta_0(n) > z_{1-\alpha}\}.$$

Note that arguments about asymptotic normality and power comparisons in [11] and in [24] apply when response-adaptive allocation proportions converge to a determined proportion $\rho \in (0,1)$, and thus don't apply to *RRU* data. [15] establish also that, under the alternative hypothesis H_1, the test statistic $\zeta_0(n)$ is a mixture of normal distributions, where the mixing variable is the positive square root η of the random variable η^2 defined in (5):

the conditional distribution of $\zeta_0(n)$ given η, is asymptotically normal with mean equal to $\sqrt{n^{m_W/m_B}}\,\eta\,\frac{\mu_B-\mu_W}{\sigma_W}$ and unit variance.

It follows that the power of the t-test can be approximated, for a large number of patients n, by

$$1-\beta_1 = P\left(\mathcal{N} + n^{m_W/(2m_B)}\eta\frac{\mu_B-\mu_W}{\sigma_W} > z_{1-\alpha}\right),$$

where $\mathcal{N}$ is a standard normal random variable independent of η.

[9] investigate also a different test statistic given by the proportion Z_n of black balls. In fact, for an asymptotic significance level α, they can consider the critical region

$$C^*_\alpha = \{Z_n > c_{1-\alpha}\},$$

where $c_{1-\alpha}$ is the quantile of order $1-\alpha$ for the distribution of the limit variable Z_∞. Moreover, they can approximate the power of this test, for large n, by

$$1-\beta^* = P\left(\eta^2 < (1-c_{1-\alpha})\frac{m_B}{m_W}n^{1-m_W/m_B}\right).$$

We believe that a promising future area of research, to improve the performance of tests of hypothesis for *RRU*-designs, could be to perform permutation tests in this case; for an overview on these tests see, for instance, [21].

6 Simulation studies

[20] propose as new guideline for the evaluation of response-adaptive designs considering their performance in competion with a non-adaptive benchmark. As a seminal example, in their paper they select their benchmark to be the problem of comparing the means of two normal responses (with same known variance), with reference to the t-test statistic defined by (8). The default design is a balanced, non-adaptive design of level α, and having n choosen such that the power is a given value $1-\beta$ when the difference δ between the two mean responses is greater or equal to a clinically relevant difference $\delta_0 > 0$. (Recall that the balanced allocation, assigning the same number of patients to treatments B and W, maximizes the power of the one-sided z-test.) Numerical simulations show that for clinical trials where δ_0 is sufficiently large, we expect that a *RRU*-design to be effective alternative for the experimenter, since

(a) to obtain an α-level test with power at least $1-\beta$ when $\delta \geq \delta_0$, it needs a trial of sample size n^* slightly larger then n;

(b) when the sample size is n^* and $\delta \geq \delta_0$, the random number $N_W(n^*)$ of patients allocated to the worst treatment is less then $n/2$ with high probability.

For instance, for $n = 100$ and known variance of the responses $v_0^2 = 0.25^2$, of one can conclude that, at the possible cost of a larger sample size for the trial, the RRU-design becomes a viable alternative to a balanced design for values of δ_0 greater then 0.25.

A different approach to the numerical evaluation of the performance of a RRU-design is proposed in [9]. In this work, they simulate the number of patients assigned to the superior treatment in a fixed sample size experiment modeled by a RRU-design, in comparison with other competitive response-adaptive designs proposed in literature. Then they evaluate the empirical power of the two-sample t-test in the RRU-design for different values of the mean responses. Their numerical simulations are performed in the basic case when treatments have dichotomous (success/failure) responses, substituting p_B and p_W for m_B and m_W, respectively. Their analysis could be extended to a more general situation. Simulations show that for sample size $n = 400$, the RRU-design provides the empirical distribution of N_B most skewed towards the best treatment and the empirical distribution of the total number of failures most skewed towards zero, which are the desirable ethical properties. Unfortunately, the variability is high, but it decreases if one increases the initial parameter $b = w$ (a good compromise seems to be reached for $b = w = 3$) and for smaller values of p_W/p_B. As expected, simulations show that power decreases when the success probabilities are close. Moreover, the t-test seems to be more powerful than the test based on the proportion of black balls. However, since neither statistic alone is sufficient, they suggest the development of a judicious combination of the two test statistics to further increase power.

In order to study the properties of the non-centrality parameter of the t-test statistic $\zeta_0(n)$, which determines the power of the test, [16] have summarized some numerical simulations regarding the distribution of η^2. In particular, they have considered some RRU-designs with different continuous treatment responses, iterating simulations to obtain an approximation of η^2 in terms of the empirical distribution of $N_W(n)/n^{m_W/m_B}$ for large n. Simulations suggest that the distribution is weakly dependent on the particular shape and variance of the responses, while it seems to be strongly dependent on the difference between m_B and m_W. Also the location of η^2,

which has an asymmetric distribution between zero and infinity, seems to vary linearly with such difference.

7 $K > 2$ treatments

After the work of [8], the first extension of the *RRU*-designs to $K \geq 2$ treatments was presented by [14] for the case of dichotomous responses. In particular, they find the limits of the urn proportions Z_{ni}, for any $i = 1, \ldots, K$, and they show that a proportion converges to one when treatment-i has a unique maximum probability of success. [7] also show when success probabilities are equal and the initial numbers of each ball type are equal, allocation proportions converge to a Dirichlet distribution. More generally, if success probabilities are equal but the initial numbers of each ball type are not the same, allocation proportions converge to the ratio of a gamma random variable divided by a linear combination of gammas, where the coefficients in the linear combination depend on the initial numbers of balls of each type. The results by [18], for the *RRU* with general responses, also are proved for $K \geq 2$ treatments in their work. The asymptotic results in [15] and results concerning the distribution of the test statistic, is extended to the K-case in the private communication by [25].

We believe that the characterization of the limiting urn composition for more general (non-binary) random variables when treatments are equivalent that is presented in [1] and [3] could be extended for $K > 2$. A central limit theorem for the sequence of random compositions is proved for a multi-color *RRU*, generalizing the result of [2], by [6].

We note that K group comparisons did not provide the original motivation for *RRU* development. [7] were interested in an randomized exploratory procedure in a dose response setting. In the 1980's, Flournoy proposed using the *RRU* with binary responses whenever a ball was drawn whose color indicated a dose on the interior of the current design space. But if a ball was drawn whose color indicated a dose on the boundary of the current design space, the procedure would differ: if a success was observed, the design space would be enlarged by adding a ball whose color indicated a neighboring dose; if a failure was observed, the ball at that boundary dose would be removed. Flournoy called this procedure the *migrating urn*, and conducted several simulation studies of its properties. While it remains intriguing as a randomized conservative dose finding procedure, it's feasibility for clinical practice is limited by slow convergence. So finding ways to speed convergence while maintaining convergence to the best treatment with probability one is important for future work.

8 Final comments

As already addressed by [22], sequential analysis has been tenuously developed for response-adaptive treatment allocation procedures. For what concerns the designs described in this overview, an initial contribution can be found in [14] and [7], where they propose a simple random stopping rule for the experiment. It is obtained by inserting in the urn an additional ball type, called *control balls*, and they are able to derive the expected number of successes and failures correspondent to each treatment after this stopping rule, and the expected sample size of the experiment. We believe a significant improvement to RRU-designs, and response-adaptive treatment allocation procedures in general, could be obtained by the development of optimal stopping rules and sequential tests. We hope to see work in this important area of research in the future.

In our knowledge, no other results on randomized response-adaptive designs targeting the best treatment have appeared in literature. In particular, we wonder if it is possible to modify the RRU-design or build other response-adaptive designs targeting the ethically optimal allocation with probability one, maintaining randomization, but having faster convergence and smaller variability.

References

[1] Aletti, G., May, C., Secchi, P., 2007. On the distribution of the limit proportion for a two-color, randomly reinforced urn with equal reinforcement distributions. Adv. in Appl. Probab. 39, 690–707.

[2] Aletti, G., May, C., Secchi, P., 2009a. A central limit theorem, and related results, for a two-color randomly reinforced urn. Adv. in Appl. Probab. 41, 829–844.

[3] Aletti, G., May, C., Secchi, P., 2009b. A functional equation whose unkonwn is P[0,1]-valued. Technical Report. Mox-Report 16/2009, http://arxiv.org/abs/0905.3310.

[4] Beggs, A.W., 2005. On the convergence of reinforcement learning. J. Econom. Theory 122, 1–36.

[5] Bélisle, C., Melfi, V., 2008. Independence after adaptive allocation. Statist. Probab. Lett. 78, 214–224.

[6] Berti, P., Crimaldi, I., Pratelli, L., Rigo, P., 2010. Central limit theorems for multicolor urns with dominated colors. Stochastic Processes and their Applications 120, 1473–1491.

[7] Durham, S.D., Flournoy, N., Li, W., 1998. A sequential design for maximizing the probability of a favourable response. Canad. J. Statist. 26, 479–495.

[8] Durham, S.D., Yu, K.F., 1990. Randomized play-the leader rules for sequential sampling from two populations. Probability in Engineering and Information Science 26, 355–367.

[9] Flournoy, N., May, C., Moler, J.A., Plo, F., 2010. On testing hypotheses in response-adaptive designs targeting the best treatment, in: mODa 9—Advances in model-oriented design and analysis. Physica-Verlag/Springer, Heidelberg. Contrib. Statist., pp. 81–88.

[10] Hopkins, E., Posch, M., 2005. Attainability of boundary points under reinforcement learning. Games Econom. Behav. 53, 110–125.

[11] Hu, F., Rosenberger, W., 2003. Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons. Journal of the American Statistical Association 98, 671–678.

[12] Hu, F., Rosenberger, W.F., 2006. The theory of response-adaptive randomization in clinical trials. Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.

[13] Huang, P., 2007. A numerical study of a randomized reinforced urn design for binary responses. Technical Report. University of Missouri-Columbia. A Masters' project available from N. Flournoy upon request.

[14] Li, W., Durham, S.D., Flournoy, N., 1996. Randomized polya urn designs. Proceedings of the Biometric Section of the American Statistical Association , 166–170.

[15] May, C., Flournoy, N., 2009. Asymptotics in response-adaptive designs generated by a two-color, randomly reinforced urn. Ann. Statist. 37, 1058–1078.

[16] May, C., Paganoni, A., Secchi, P., 2007. Response-adaptive designs targeting the best treatment for clinical trials with continuous responses, in: S.Co.2007 Fifth Conference - Complex Models and Com-

putational intensive methods for estimation and prediction - Book of short papers, Cluep. pp. 326–331.

[17] May, C., Paganoni, A.M., Secchi, P., 2005. On a two-color generalized Pólya urn. Metron 63, 115–134.

[18] Muliere, P., Paganoni, A., Secchi, P., 2006a. A randomly reinforced urn. J. Statist. Plann. Inference 136, 1853–1874.

[19] Muliere, P., Paganoni, A., Secchi, P., 2006b. Randomly reinforced urns for clinical trials with continuous responses., in: SIS - Proceedings of the XLIII Scientific Meeting, Cluep Padova. pp. 403–414.

[20] Paganoni, A.M., Secchi, P., 2007. A numerical study for comparing two response-adaptive designs for continuous treatment effects. Stat. Methods Appl. 16, 321–346.

[21] Pesarin, F., Salmaso, L., 2010. Permutation Tests for Complex Data. Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons].

[22] Rosenberger, W.F., 2002. Randomized urn models and sequential design. Sequential Analysis 21, 1–28.

[23] Rosenberger, W.F., Lachin, J.M., 2002. Randomization in clinical trials. Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], New York. Theory and practice.

[24] Zhang, L., Rosenberger, W.F., 2006. Response-adaptive randomization for clinical trials with continuous responses. Biometrics 62, 562–569.

[25] Zhang, L.X., Hu, F., Cheung, S.H., Chan, W.S., 2008. Multi-color randomly reinforced urn for adaptive designs. Technical Report. Private communication.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

38/2010 Nancy Flournoy, Caterina May, Piercesare Secchi:
Response-adaptive designs in clinical trials for targeting the best treatment: an overview

37/2010 Marco Discacciati, Alfio Quarteroni,
Samuel Quinodoz:
Numerical approximation of internal discontinuity interface problems

36/2010 Gianni Arioli, Hans Koch:
Non-Symmetric low-index solutions for a symmetric boundary value problem

35/2010 Gianni Arioli, Hans Koch:
Integration of dissipative PDEs: a case study

34/2010 Antonella Abba', Luca Bonaventura:
A mimetic finite difference method for Large Eddy Simulation of incompressible flow

33/2010 Giovanni Migliorati, Alfio Quarteroni:
Multilevel Schwarz Methods for Elliptic Partial Differential Equations

32/2010 A. Cristiano I. Malossi, Pablo J. Blanco,
Simone Deparis, Alfio Quarteroni:
Algorithms for the partitioned solution of weakly coupled fluid models for cardiovascular flows

31/2010 Andrea Manzoni, Alfio Quarteroni, Gianluigi Rozza:
Shape optimization for viscous flows by reduced basis methods and free-form deformation

30/2010 Pablo J. Blanco, Marco Discacciati,
Alfio Quarteroni:
Modeling dimensionally-heterogeneous problems: analysis, approximation and applications

29/2010 Matteo Lesinigo, Carlo D'Angelo,
Alfio Quarteroni:
A multiscale Darcy-Brinkman model for fluid flow in fractured porous media