A simulation study on the use of response-adaptive randomized designs

Uno studio di simulazione sull'uso di disegni casuali adattivi alla risposta

Anna Maria Paganoni, Andrea Ghiglietti, Maria Giovanna Scarale, Rosalba Miceli, Francesca Ieva, Luigi Mariani, Cecilia Gavazzi and Valeria Edefonti

Abstract Response-adaptive designs have been proposed in randomized clinical trials to achieve ethical advantages by using sequential accrual information collected during the trial to update probabilities of treatment assignments. We propose the use of a response adaptive design based on urn models in a simulation study on a randomized clinical trial on the efficacy of home enteral nutrition in cancer patients after major gastrointestinal surgery. We compare results with the adaptive design with those previously obtained with the non-adaptive approach.

Abstract *I disegni adattivi alla risposta sono stati proposti nell'ambito degli studi clinici per ottenere vantaggi etici utilizzando le informazioni sequenziali raccolte durante lo studio per aggiornare le probabilità di assegnazione ai trattamenti. In questo lavoro si propone l'uso di un disegno adattivo di risposta basato su modelli d'urna in uno studio di simulazione basato su uno studio clinico dell'efficacia della nutrizione enterale domiciliare in pazienti oncologici dopo un intervento chirurgico*

Anna Maria Paganoni, Francesca Ieva
Politecnico di Milano, Milano (Italy)
e-mail: anna.paganoni@polimi.it; e-mail: francesca.ieva@polimi.it

Andrea Ghiglietti
Università Cattolica del Sacro Cuore, Milano (Italy)
e-mail: andrea.ghiglietti@unicatt.it

Maria Giovanna Scarale
IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (Italy)
e-mail: mgiovanna.scarale@operapadrepio.it

Rosalba Miceli, Luigi Mariani, Cecilia Gavazzi
Fondazione IRCCS Istituto Nazionale Tumori, Milano (Italy)
e-mail: Rosalba.Miceli@istitutotumori.mi.it; e-mail: luigi.mariani@istitutotumori.mi.it;
e-mail: Cecilia.Gavazzi@istitutotumori.mi.it

Valeria Edefonti
Università degli Studi di Milano, Milano (Italy)
e-mail: valeria.edefonti@unimi.it

gastrointestinale maggiore. Si confrontano i risultati del disegno adattivo con quelli precedentemente ottenuti con l'approccio non adattivo.

Key words: Randomly Reinforced Urn model, Randomized trials, Response-adaptive randomization, Simulation study

1 Introduction

In statistical literature, urn models have been widely studied as mathematical tools to implement randomization in the context of clinical trials. These designs randomly assign those subjects that sequentially enter the trial to the treatment arms according to to the proportion of balls of different color sampled from a virtual urn. Recently, interest has been increased in the use of urn models, in which the probability to sample a ball of a certain type depends on the treatment performance observed on the subjects previously randomized [2, 1]. A popular class of such designs is the Randomly Reinforced Urn (RRU) model, which has been introduced in [2] for binary treatment responses and extended in [5] to handle continuous responses. In the RRU model, an urn containing balls of two colors is sequentially sampled and the subjects in the trial receive the treatments associated to the colors of the sampled balls. In addition, the urn composition is sequentially updated by adding new balls of the same color of the sampled ones, whose number depend each time on the response observed by the correspondent patient. For the purposes of this paper, we simply remind that a RRU design assigns patients to the superior treatment with a probability that converges to one as the sample size increases. Although the theoretical result of assigning most of the patients to the superior treatment is very attractive from the ethical point of view, the RRU design have rarely been implemented in clinical trials or in simulation studies based on a real set-up. In detail, we will simulate a large number of trials that follow the RRU model starting from the real-life data collected in a previously published Home Enteral Nutrition (HEN) randomized trial [3], where a non-adaptive design was originally adopted (see [4]). Comparing the performance of the RRU with that of the original non-adaptive design, we expect that the RRU design will: 1) assign fewer patients to the inferior treatment; 2) maintain similar inferential properties. This will turn out in an advantage in terms of both statistical performance and ethical responsibility.

2 Materials and Methods

The RRU model was here implemented in a simulation study based on results from a multicenter, controlled, open-label, two-parallel groups, randomized clinical trial conducted at the Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy, and at the European Institute of Oncology, Milan, Italy, between December

2008 and June 2011 [3]. The enrolled subjects were adult ($>$ 18 years) patients with documented upper gastrointestinal cancer who were candidates for major elective surgery and showed a preoperative nutritional risk score that indicated a potential benefit from any nutritional intervention. A random permuted block design (stratified for referring center) randomly assigned patients before discharge to receive either HEN to cover the basal energy requirement (experimental group), or nutritional counselling by an expert dietitian, including oral supplements only when needed (Control Group - CG), in a 1:1 ratio. The treatment effect was defined as the difference between the mean "weight change" (weight after two months - weight at baseline) in the HEN and nutritional counselling arms (primary end-point). In total, 79 patients were initially randomized; however, as 11 patients had a missing two-months weight, the final analysis was performed on 68 patients, of which 33 patients were allocated to the HEN group and 35 to the CG. The main result of the primary end-point analysis was that the mean weight loss in the patients undertaking the HEN treatment was significantly lower than that in the CG, with a treatment effect estimated by the corresponding ANOVA model coefficient (95% confidence interval) of 3.2 (1.1-5.3) and a p-value from the corresponding two-sided t-test equal to $0.31\% < 2.94\%$. For this reason, the trial was stopped at the interim analysis and results from this analysis were published in [3]. So, the HEN was found to be the superior treatment in this trial.

To simulate the RRU design starting from the HEN trial data to derive the results for comparing the RRU design with the non-adaptive one, we performed the following main steps:

(A) using the HEN trial dataset [3]:

(1) we estimated the parameters of the Gaussian distribution of the responses to the HEN group;
(2) we estimated the parameters of the Gaussian distribution of the responses in the CG;
(3) we computed the empirical distribution of the difference between arrival times of consecutive subjects;

(B) we simulated N independent trial samples based on the RRU model; for each sample, responses to both treatments and intervals between arrival times were randomly generated from distributions introduced in point (A);

(C) we computed from these N trials:

(1) the empirical distribution of the number of subjects assigned to the inferior treatment $\mathscr{W}$;
(2) the empirical power of the corresponding t-test.

The previous steps are detailed in the following.

To start, we considered three alternative set-ups of trial sample sizes equal to (a): $n = 58$;(b): $n = 68$; (c): $n = 78$, where the total sample size 68 of the HEN trial was used as the reference set-up and we moved $\pm 15\%$ from that to get other two reasonable sample sizes.

For each set-up, we performed $N = 10,000$ simulations of independent trials based

on the RRU design: in each run we have a virtual urn to be sampled and reinforced. Formally, we denote by (R_i^j, W_i^j) the urn composition and by $R_i^j/(R_i^j + W_i^j)$ the urn proportion in simulation $j = \{1,..,N\}$ at time $i \in \{1,..,n\}$.
All the urns start with the same (fixed) initial composition, i.e. $(R_0^j, W_0^j) = (R_0, W_0)$ for any $j = \{1,..,N\}$. Then, the urn composition (R_i^j, W_i^j) is updated in the following way

$$\begin{cases} R_i^j = R_{i-1}^j + \sum_{k \in (A_{i+1}^j \setminus A_i^j)} X_k^j \xi_{Rk}^j \\ \\ W_i^j = W_{i-1}^j + \sum_{k \in (A_{i+1}^j \setminus A_i^j)} (1 - X_k^j) \xi_{Wk}^j, \end{cases}$$

where X_k^j is a Bernoulli random variable with parameter $R_{k-1}^j/(R_{k-1}^j + W_{k-1}^j)$ indicating the treatment assignement, the set A_i^j here includes all the patients who arrived two months earlier than subject i and ξ_{Rk}^j, ξ_{Wk}^j are the subjects responses to treatment R and W respectively. Indeed, in the HEN trial, responses were available only two months after treatment administration.

In addition, as normality assumptions in the original data were not rejected, responses to both treatments were generated as independent Gaussian random variables with arm-specific means and variances computed using the HEN dataset and given by: $m_R = -0.315$ and $\sigma_R = 3.868$ for treatment $\mathscr{R}$ (HEN group), $m_W = -3.571$ and $\sigma_W = 4.789$ for treatment $\mathscr{W}$ (CG). Formally, we generated the following quantities:

(1)$\xi_{R1}^j, .., \xi_{Rn}^j \sim \mathscr{N}(m_R, \sigma_R^2)$ potential responses to treatment $\mathscr{R}$ (HEN group);
(2)$\xi_{W1}^j, .., \xi_{Wn}^j \sim \mathscr{N}(m_W, \sigma_W^2)$ potential responses to treatment $\mathscr{W}$ (CG),

where either ξ_{Ri}^j or ξ_{Wi}^j is observed, as each subject just receives one treatment.

3 Results

Table 1 shows some descriptive statistics of the empirical distribution of the number of subjects assigned to the inferior treatment, $N_W(n)$, and the empirical power of the t-test, $1 - \hat{\beta}$, for the different sample sizes n, in comparison with the corresponding results for the non-adaptive design, n_W and $1 - \beta$.

For all sample sizes under consideration [cases (a)-(b)-(c)], the mean and the median of $N_W(n)$ were smaller than n_W, the number of subjects assigned to the inferior treatment by the non-adaptive design. It follows that the RRU design provided 50% of probability (or more) to assign fewer subjects to the inferior treatment, as compared to the non-adaptive design. Although higher than n_W for all the sample sizes considered, the third quartile of $N_W(n)$ in the RRU design was very close to n_W for any n under consideration. In addition, the obtained values for the t-test empirical power under the RRU design were close, but slightly smaller than, the corresponding power values derived in the non-adaptive design.

Further information on the distribution of $N_W(n)$ is provided by the boxplots re-

n	$N_W(n)$				n_W	$1-\beta$	$1-\hat{\beta}$
	1^{st} quartile	Mean	Median	3^{rd} quartile			
(a) 58	19	25.6	25	31	29	0.88	0.83
(b) 68	**22**	**29.6**	**29**	**36**	**35**	**0.92**	**0.88**
(c) 78	25	33.6	33	41	38	0.94	0.92

Table 1 Summary statistics (1^{st} and 3^{rd} quartiles, mean, and median) of the empirical distribution of the number of subjects assigned to the inferior treatment, $N_W(n)$, and empirical power, $1-\hat{\beta}$, of the t-test for equal mean weight changes (corresponding to the treatment coefficient in the ANOVA model) for the different sample sizes n in the Randomly Reinforced Urn design, in comparison with the corresponding results for the non-adaptive design, n_W and $1-\beta$. We reported in bold typeface the results obtained with the same sample size of the original Home Enteral Nutrition trial.

ported in Figure 1. For any sample size, the median of $N_W(n)$ was below the dashed line indicating the number of subjects assigned to the inferior treatment by the non-adaptive design. Similarly, we confirmed that, although higher, the third quartile was closer than the median to the dashed line for the three cases under consideration. In addition, the probability that $N_W(n)$ was less than n_W was close to 75% for any sample size under consideration. Finally, although mostly symmetric, the empirical distributions of the number of subjects assigned to the inferior treatment showed a high level of variability. This variability increases, as the total sample size increases.

References

1. Atkinson, A.C., Biswas, A., 2014. Randomised response-adaptive designs in clinical trials. Chapman and Hall/CRC.
2. Durham, S.C., Flournoy, N. and Li, W., 1998. A sequential design for maximizing the probability of a response. The Canad. J. Stat., 26 (3), 479-495.
3. Gavazzi, C., Colatruglio, S., Valoriani, F., Mazzaferro, V., Sabbatini, A., Biffi, R., Mariani, L., Miceli, R. 2016. Impact of home enteral nutrition in malnourished patients with upper gastrointestinal cancer: a multicenter randomized clinical trial. Europ. J. Cancer, 64, 107-112.
4. Ghiglietti, A., Scarale, M.G., Miceli, R., Ieva, F., Mariani, L., Gavazzi, C., Paganoni, A.M., Edefonti, V. 2018. Urn models for response-adaptive randomized designs: a simulation study based on a non-adaptive randomized trial *J Biopharm Stat*, In press.
5. Muliere, P., Paganoni, A.M., Secchi, P., 2006. A randomly reinforced urn. J. Stat. Plann. Inference, 136, 1853-1874.
6. Wei, L.J. 1978. An application of an urn model to the design of sequential controlled clinical trials. JASA, 73, 559-563.

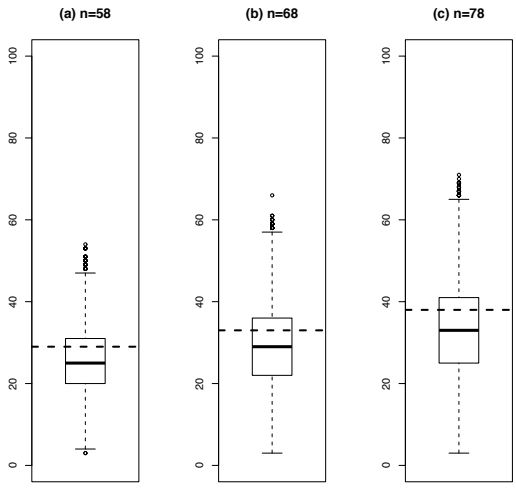

Fig. 1 Boxplots of the number of subjects assigned to the inferior treatment (Control Group) in the three cases reported above each picture: (a) $n = 58$, (b) $n = 68$, (c) $n = 78$. The dashed line indicated the number of subjects assigned to the control group in the non-adaptive trial in the three cases.