Comparison of allocation procedures in clinical trials in small population groups with respect to accidental and selection bias

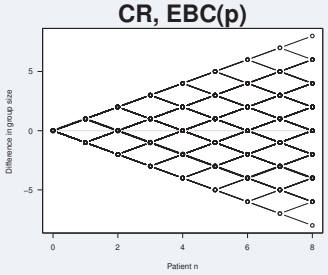
D. Schindler[1], D. Uschner[2], M. Tamm, N. Heussen, R.-D. Hilgers
Department of Medical Statistics, RWTH Aachen University, Pauwelsstraße 30, 52074 Aachen, Germany
[1,2]presenting authors: [1]dschindler@ukaachen.de, [2]duschner@ukaachen.de

Introduction

For statistically proving the effectiveness of a medical intervention, the randomised controlled clinical trial is considered the "gold standard". The aim of this investigation is to compare established randomisation procedures with respect to selection and accidental bias. Latter is investigated in form of a linear time trend. We consider the situation of open (i.e. unmasked) two-armed clinical trials with parallel group design. We present six randomisation procedures and give a statistical model for selection bias and time trend. The randomisation procedures are compared concerning their susceptibility to both biases. We conduct a simulation study with an analysis unadjusted for this bias. We focus on small clinical trials with total sample size $N \leq 40$.

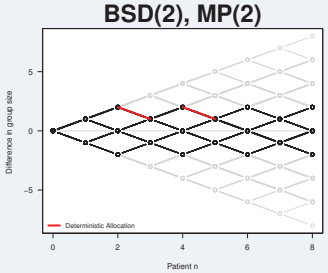
Randomisation Procedures

CR, EBC(p)



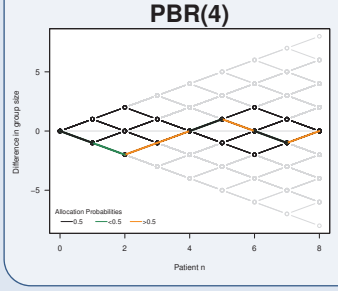
Method	Property
Complete Randomisation (CR)	Fair coin flip
Efron's Biased Coin (EBC(p))	Biased coin flip in order to produce balance with $p \in [0.5,1]$

BSD(2), MP(2)



Method	Property
Big Stick Design (BSD(MTI))	Fair coin flip with a deterministic allocation when imbalance boundary MTI[1] is hit
Maximal Procedure (MP(MTI))	Equally probable sequences with boundary MTI[1] for the imbalance and final balance

PBR(4)



Method	Property
Permuted Block Randomisation (PBR(b))	Equally probable sequences with forced balance after multiples of b allocations
Random Allocation Rule (RAR)	Equal probable sequences with final balance

[1]MTI= Maximum Tolerated Imbalance during the trial

Model

Let $T \in \Gamma = \{-1,1\}^N$ be the vector of treatment assignments with total sample size N. Let $T_n = 1$ (resp. -1) if patient n is assigned to treatment E (resp. C). For the number of patients assigned to E after n allocations, we write $N_E(n) := 0.5 \cdot \sum_{i=1}^{n}(T_i + 1)$ and, respectively, $N_C(n) := N - N_E(n)$. Suppose there is no difference in treatment effects of the two groups. Without loss of generality, we assume

$$\mu_E = \mu_C = 0.$$

Let Y_n denote the normally distributed *response* of the n-th patient. We assume that Y_m, Y_n are stochastically independent (for $m \neq n$) and have the same variance σ^2.

Under the assumptions that the investigator favours E, knows the past assignments and that the chosen randomisation procedure forces balance (i.e. $N_E(N) = N_C(N)$), it is opportune for the investigator to select the next patient according to his expected response:

$$Y_n \sim \begin{cases} \mathcal{N}(-\eta, \sigma^2) & N_E(n-1) > N_C(n-1) \\ \mathcal{N}(0, \sigma^2) & N_E(n-1) = N_C(n-1) \\ \mathcal{N}(\eta, \sigma^2) & N_E(n-1) < N_C(n-1) \end{cases}.$$

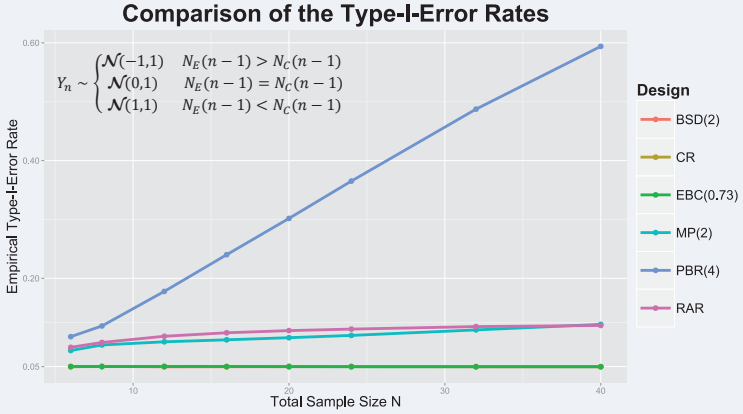
If the patient responses are influenced by a linear time trend we assume that they are shifted according to the index of inclusion (with $\vartheta > 0$):

$$Y_n \sim \mathcal{N}\left(\vartheta \cdot \frac{n-1}{N}, \sigma^2\right).$$

Simulation Study

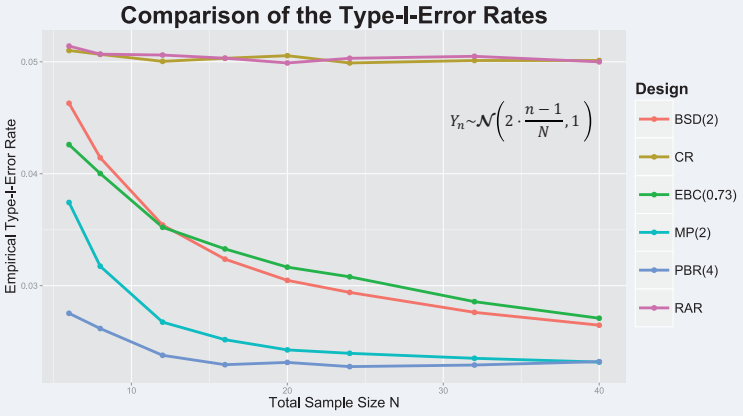
For the simulation study we generated 10^6 sequences for each randomisation procedure, each $N \in \{6,8,12,16,20,24,32,36,40\}$ and both settings selection bias and time trend. For each generated sequence we simulated a response vector according to the model. Then we conducted Student's t-test without adjusting for bias. Finally, we calculated the proportion of times that $H_0: \mu_E - \mu_C = 0$ was rejected to the total number of repetitions 10^6, setting $\sigma^2 = 1, \gamma = 1, \vartheta = 2$ and $\alpha = 5\,\%$. We call this proportion *(empirical) type-I-error rate*.

Selection Bias



Observe that Student's t-test rejects $H_0: \mu_E = \mu_C$ in 5% of cases if we use BSD, EBC or CR. For all other designs, the type-I-error rate grows with N. RAR and MP are less liberal to selection bias than PBR. Using PBR for N $\leq$ 40, Student's t-test rejects H_0 in $\approx$ 60% of cases.

Linear Time Trend



We observe that Student's t-test attains an empirical type-I-error rate of 5% if we use CR or RAR. All other designs react more conservatively the greater N is. BSD and EBC behave very similarly and generally less conservatively than MP and PBR .

Discussion

For $N \leq 40$, selection bias and linear time trend can strongly influence the test decision of an unadjusted t-test. The randomisation method that is used for a trial should thus be chosen carefully according to the type of bias that is expected. The evaluation of selection bias and time trend seems to reveal that CR doesn't react to the investigated types of bias. This is only true on average. In small clinical trials the asymptotically good properties of CR don't hold. We suggest the use of MP if final balance is required, and the use of BSD otherwise.

References

Blackwell, D., and Hodges J. L. (1957). Annals of Math. Statist., 28, 449-460.
Efron, B. (1971). Biometrika, 58, 403-417.
Kennes, L. N. (2013). *The effect of and adjustment for selection bias in randomized clinical trials.* Doctoral thesis, RWTH Aachen.
Kennes, L. N., Cramer E., Hilgers, R.-D. and Heussen N. (2011). Statist. Med., 30, 2573-2581.
Proschan, M. (1994). Statistica Sinica, 4, 219-231.
Rosenberger, W. F., and Lachin, J. M. (2002). *Randomization in Clinical Trials.* New York: Wiley.
Soares, J. F., and Wu, C. F. J. (1983). Commun. Statist.-Theor. Meth., 12, 2017-2034.
Tamm, M., Cramer, E., Kennes, L. N., and Heussen, N. (2012). Meth. Inf. Med., 51, 138-143.
Tamm, M., and Hilgers, R.-D. (2014). (accepted for publication in Meth. Inf. Med.)

This project has received funding from the European Union's 7th Framework Programme for research, technological development and demonstration under Grant Agreement no 602552.