A comparison of the performance of randomization methods in small-size clinical trials

Kanae Takahashi* and Kouji Yamamoto *†

***Abstract* — There are several methods of randomization those are widely used. However, the performance of randomization methods for small-size clinical trials has not been clearly. Thus, the objective in this study is to compare the performance of the existing randomization methods by simulation for small-size clinical trials. We conducted numerical simulations to investigate (a) the difference of the behavior of the empirical power and the empirical type I error rate among randomization methods and statistical analysis methods, (b) the behavior of the empirical power when we analyze using baseline characteristics that have no effect on the outcome, and (c) the behavior of the empirical power when a prognostic factor has an interaction. Assuming total sample size N was 20, 40 and 50, we compared four randomization methods (simple randomization, permuted block design, stratified blocked randomization, and minimization) and three statistical analysis methods (Student's t-test, permutation test, and covariate-adjusted analysis).**

Keyword: *Randomization; Simulation study; Small-size trial.*

1 Introduction

Clinical trials are the most definitive method of determining whether an intervention has the postulated effect [1]. Those should also be designed to reduce the bias and confounding to isolate the effect of an intervention, and establish cause and effect. The pivotal component of clinical trials is randomization that is a technique for assigning patients to the experimental treatment(s) or control. Randomization promotes comparability among the study groups with respect to not only known covariates but also unknown important covariates [2], and it provides an unbiased and precise estimate of the intervention's effect.

There are several methods of randomization that are commonly used. For example, simple randomization (SP), permuted block design (PB), stratified blocked randomization (ST) and minimization (MI). The selection of randomization method is important at the time of protocol planning. There are several articles that compare the performance of these methods [3-8]. The performance of randomization methods may be associated with trial sample size, but these articles did not consider the case of small-size clinical trial. Thus, the objective in this study is to compare the performance of the existing randomization methods by simulation for small-size clinical trials.

2 Methodology

We performed a simulation study to investigate (1) the difference of the behavior of the empirical power and the empirical type I error rate among four randomization methods and three statistical analysis methods, (2) the behavior of the empirical power when we analyze using baseline characteristics that have no effect on the outcome, and (3) the behavior of the empirical power when a prognostic factor has an interaction. We considered a small-size randomized controlled trial in which patients were allocated to the treatment group or the control group. The total sample size N was assumed 20, 40 and 50. We generated simulation data from the following model:

$$Y_i = \mu + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_{71} X_{71i} + \beta_{72} X_{72i} + \varepsilon_i \quad (1)$$

where Y_i is the continuous outcome from the ith patient, X_{1i} is a binomial variable that represents treatment groups (treatment or control), X_{2i}-X_{6i} are prognostic factors that have effects on the outcome.

* Department of Medical Innovation, Osaka University Hospital, 2-2 Yamadaoka, Suita, Osaka, 565-0871, Japan

E-mail: kanae.takahashi@dmi.med.osaka-u.ac.jp, Tel: +81-6-6210-8373

† Department of Clinical Epidemiology and Biostatistics, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka, 565-0871, Japan

X_{2i}, X_{3i}, X_{4i} and X_{5i} are binomial variables distributed as Bernoulli distribution. X_{6i} is a continuous variable distributed as normal distribution with mean=5, variance=5. Assuming X_{7i} is a categorical variable that has three categories, X_{71i} and X_{72i} are dummy variables of X_{7i}. ε_i is a random variable that has mean=0 and variance=1. In addition, we considered Z_{1i}, Z_{2i} and Z_{3i} as baseline characteristics that had no effect on the outcome. These are binomial variables distributed as Bernoulli distribution. When investigating purpose (3) the behavior of the empirical power when a prognostic factor has an interaction, we added the interaction term $\beta_8 X_{1i} X_{2i}$ to the model (1).

To compare the performance of the randomization methods, we generated X_{1i} by using one of four randomization methods: SP, PB, ST and MI. The number of stratification factors considered in ST was N/15 [9]. For MI, the number of stratification factors was N/15 or six that included X_{2i}-X_{7i} or nine that included X_{2i}-X_{7i} and Z_{1i}-Z_{3i}. The continuous variables were binarized by using mean when conducting MI. In PB and ST, the block size was set to four. In MI, we considered 0.80 for allocation probability because some articles recommended it as about 0.80 [6, 7]. The weight for all stratification factors was defined as one to simplify the results. We generated 100,000 data sets with 20, 40 and 60 patients for each randomization scenario.

To investigate the difference of the performance of randomization methods among statistical analysis techniques, we conducted three analyses: Student's t-test, permutation test and covariate-adjusted analysis that adjusted the effect of stratification factors in a regression model. We defined the number of stratification factors used in covariate-adjusted analysis as N/15 in the simulation [9]. We could not use the stratification factors when N=20 because we could only use N/15 = 1 factor in the covariate-adjusted analysis and the factor had got to be assigned as treatment group (X_{1i}).

We defined performance measures as follows: the empirical power, the empirical type I error rate, the difference in the number of patients between groups, point estimate and mean squared error (MSE) of treatment effect.

References

[1] Friedman LM, Furberg CD, DeMets DL, et al. Introduction to Clinical Trials. In: Fundamentals of Clinical Trials. Cham: Springer International Publishing, pp. 1–23.

[2] Rosenberger WF, Lachin JM. Randomization and the Clinical Trial. In: Randomization in Clinical Trials Theory and Practice. 2002, pp. 1–14.

[3] Therneau TM. How many stratification factors are 'too many' to use in a randomization plan? Control Clin Trials 1993; 14: 98–108.

[4] Weir CJ, Lees KR. Comparison of stratification and adaptive methods for treatment allocation in an acute stroke clinical trial. Stat Med 2003; 22: 705–726.

[5] Hagino A, Hamada C, Yoshimura I, et al. Statistical comparison of random allocation methods in cancer clinical trials. Control Clin Trials 2004; 25: 572–584.

[6] Brown S, Thorpe H, Hawkins K, et al. Minimization - Reducing predictability for multi-centre trials whilst retaining balance within centre. Stat Med 2005; 24: 3715–3727.

[7] Toorawa R, Adena M, Donovan M, et al. Use of simulation to compare the performance of minimization with stratified blocked randomization. Pharm Stat 2009; 8: 264–278.

[8] Kahan BC, Morris TP. Adjusting for multiple prognostic factors in the analysis of randomised trials. BMC Med Res Methodol 2013; 13: 99.

[9] Harrell FE. Regression Modeling Strategies. In: Regression Modeling Strategies, pp. 501–507.