Statistica Sinica Preprint No: SS-2016-0330.R2

Title	Sequential Monitoring of Covariate-Adaptive Randomized Clinical Trials
Manuscript ID	SS-2016-0330.R2
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0330
Complete List of Authors	Feifang Hu and Hongjian Hu
Corresponding Author	Feifang Hu
E-mail	feifang@gwu.edu

Notice: Accepted version subject to English editing.

Sequential Monitoring of Covariate-Adaptive Randomized Clinical Trials

Abstract: The sequential monitoring of covariate-adaptive randomized clinical trials is standard in modern clinical studies. However, the validity of this sequential procedure is not well studied in the literature. Clinical trialists therefore implement the procedure and perform data analysis based on the theory of the sequential monitoring of fixed designs, and many clinical trials are open to question. In this paper, we study the theoretical properties of the sequential procedure and propose some important adjustments to classical statistical inference. Under different scenarios, we derive the asymptotic joint distribution of the sequential test statistics. Further, we estimate the decreased variability of the estimated treatment effect due to covariate-adaptive randomization, so that the sequential test statistics can be adjusted to be an asymptotic Brownian motion and the type I error rate can be controlled in real trials. Numerical results from simulation and the redesign of a clinical trial support our theoretical findings, showing that our procedure can control the type I error rate well, and also demonstrating the advantages of our method in terms of power and early stopping. Both theoretical and numerical results provide important guidance for future practical clinical trials using covariate-adaptive randomization procedures.

Key words and phrases: Brownian motion, linear regression, personalized medicine, stratified permuted block randomization, Pocock–Simon's randomization, type I error rate.

1. Introduction

Clinical trials are usually complex, involving multiple covariates of interest in addition to the treatment effects. In particular, with the development of bioinformatics, the association between biomarkers and disease has become widely accepted. In the era of personalized medicine, it is desirable to incorporate covariates into clinical trial designs that investigate the heterogeneity of patients' responses to a treatment (Hu, 2012; Hu et al, 2015). The study results may be invalid if there is treatment imbalance over the covariates. Covariate-adaptive randomization (CAR) procedures, which sequentially assign the next patient based on previous assignments and covariates, and the current covariate profile, have been developed to mitigate such imbalances and are extensively used in clinical trials. Stratified permuted block (SPB) randomization and Pocock and Simon's design (1975) are the most popular CAR procedures. Other CAR designs have been developed by Taves (1974), Wei (1978), Nordle and Brantmark (1977), Signorini et al. (1993), Heritier et al. (2005), and Hu and Hu (2012). Clinical trials that use these designs include Iacono et al. (2006), Jakob et al. (2012), Anderson et al. (2000), Gridelli et al. (2003), Krueger et al. (2007), Molander et al. (2007), and Ohtori et al. (2012). A detailed discussion of CAR procedures can be found in Rosenberger and Sverdlov (2008). The theoretical properties of hypothesis testing based on CAR procedures have recently been developed by Shao et al. (2010) and Ma et al. (2015). However, both papers focused on the final test statistic instead of the sequential statistics (a stochastic process).

While CAR procedures are very popular in clinical trials, interim analysis is also common because of its ethical, administrative, and economic advantages (Jennison and Turnbull, 2000). Sequential monitoring arose from the sequential probability ratio test proposed by Wald (1947) for quality control, and its use in medical research was pioneered by Armitage (1975). Influential papers on sequential monitoring in clinical trial designs include Pocock (1977), O'Brien and Fleming (1979), and Lan and DeMets (1983). Further, Jennison and Turnbull (1997) discussed a series of group sequential analysis methods incorporating covariate information through linear models, general parametric regression models and survival models. However, they did not take into account the problems caused by covariate adaptive designs and the scenario where not all the design covariates were used in the analysis. Tsiatis et al. (1985) and Gu and Ying (1995) derived the joint distribution of sequential parameter estimators from proportional hazards models. More details of sequential monitoring can be found in Jennison and Turnbull (2000). Note that these studies considered the scenarios where non-adaptive designs are implemented in clinical trials.

Despite the widespread popularity of the combination of CAR procedures with sequential monitoring in real trials and the advantages mentioned above, there have been few theoretical investigations of the sequential procedure. The CAR procedure has two limitations: the complicated correlation structure of the within-stratum imbalances and the discreteness of the allocation function. Furthermore, a special situation often arises in real clinical trials: only some of the covariates used in the randomization procedures are included in the data analysis. For example, Lai et al. (2006) investigated the influences of music on maternal anxiety in kangaroos in a randomized controlled trial. Under similar conditions, female infants are believed to have a significantly greater chance of surviving than male infants, hence permuted block randomization stratified on gender was used to allocate the patients. In the data analysis, a t-test was used to analyze the maternal-anxiety outcomes. The reasons for not using all the covariates include, but are not limited to, (i) it is not easy to explain the practical significance of including certain covariates such as investigation sites in the model; (ii) using too many covariates will lead to theoretical difficulties; (iii) the correct model specification is usually unknown. Consequently, theoretical investigation into the sequential monitoring of CAR procedures has been hindered for decades. More importantly, the clinical trials that employ this procedure lack complete theoretical support, and many of these trials may be open to question.

In this paper, we study clinical trials with the CAR design for randomization and linear regression models for analysis. We obtain the joint distribution of the sequential statistics for the following three scenarios: (1) all the covariates used in the CAR are included in the data analysis; (2) some of the covariates are included; and (3) no covariates are included, which is Student's t-test. We find that for scenario (1) the joint distribution of the commonly used sequential statistics discussed in Section 2 is asymptotically Brownian motion, which is the asymptotic joint distribution for complete randomization and fixed designs. As mentioned before, clinical trial practitioners often perform data analysis following the sequential monitoring of CAR procedures, assuming that the data are from the sequential monitoring of complete randomization. This finding, for the first time to our knowledge, theoretically justifies and validates all such clinical trials for this scenario.

We also derive the joint distribution of the sequential statistics for scenarios (2) and (3), and we can see its difference from standard Brownian motion. As a result, trials that ignore the difference between CAR procedures and complete randomization could give misleading conclusions. The above theoretical results provide guidance for practical clinical trials, and they are one of the major contributions of this paper. In addition, the asymptotic variances of the sequential statistics for scenarios (2) and (3) indicate that the CAR design shrinks the variability of the estimated treatment effect. We propose an approach to estimate the decreased variance and adjust the sequential statistics, so that the critical values for Brownian motion can still be used, which offers clinical trialists practical steps to deal with these complex situations.

Finally, we perform extensive numerical studies for the above three scenarios in terms of the type I error, power, and early stopping. We also redesign a double-blind randomized two-arm clinical trial conducted by Tilley et al. (1995) to study the properties of the proposed methods. The numerical results support our theoretical findings and demonstrate the advantages of our methods.

In Section 2, we introduce the notation, describe the framework, and formulate the main theorems. In Section 3, we use generated data to numerically study the sequential monitoring of CAR procedures. Numerical results from the redesign of a clinical trial are discussed in Section 4. Conclusion remarks are in Section 5, and the proofs are given in the online

domized clinical trials is not well studied in the literature. In this paper, we have derived the joint distribution of the sequential statistics for three common scenarios in clinical trials that use CAR procedures for randomization and linear regression models for analysis. Based on these theoretical properties, we have proposed practical approaches to make use of existing critical values to control the type I error rate. We have also numerically studied different procedures and redesigned a clinical trial. The results demonstrated that the type I error rate can be protected as indicated by our theoretical conclusions, and they also showed the advantages of the combination of sequential monitoring and covariate adaptive designs.

There are several important directions for future research. First, we have studied data analysis for continuous responses with linear regression. Binary responses with logistic regression are a natural generalization. Many other types of responses and models deserve study; difficulties could be introduced by the nonexistence of a closed form of the parameter estimators. Second, we have made use of the α-spending function to control the type I error rate. Other methods may provide diverse advantages; these include optimal spending functions (Anderson, 2007) and beta spending functions. Third, a generalized structure of covariates could be investigated for other scenarios in real clinical trials. Fourth, other approaches to adjust the sequential statistics could be developed. Finally, Hu and Rosenberger (2006) classified adaptive randomization procedures into four categories, i.e., restricted randomization, response-adaptive randomization (RAR), CAR and covariate-adjusted response-adaptive (CARA) randomization. Zhu and Hu (2010, 2012) studied sequential monitoring of RAR in clinical trials. The sequential monitoring of CARA is worth an investigation.

Supplementary Materials

The proofs are in the online supplementary materials.

Acknowledgements

We highly appreciate the constructive suggestions from the referees and the editors. Research is supported by grant DMS-1612970 from the National Science Foundation (USA) and by grant No. 11371366 from the National Natural Science Foundation of China.

References

Anderson, K. M. (2007). Optimal spending functions for asymmetric group sequential designs. *Biom. J.* **49**, 337-345.

Anderson, H., Hopwood, P., Stephens, R. J., Thatcher, N., Cottier, B., Nicholson, M., Milroy, R., Maughan, T. S., Falk, S. J., Bond, M. G., Burt, P. A., Connolly, C. K., McIllmurray,

M. B. and Carmichael, J. (2000). Gemcitabine plus best supportive care (BSC) vs BSC in inoperable non-small cell lung cancer: A randomized trial with quality of life as the primary outcome. *British Journal of Cancer* **83**, 447-453.

Armitage, P. (1975). *Sequential Medical Trials*. Blackwell, Oxford.

Gridelli, C., Gallo, C., Shepherd, F. A., Illiano, A., Piantedosi, F., Robbiati, S. F., Manzione, L., Barbera, S., Frontini, L., Veltri, E., Findlay, B., Cigolari, S., Myers, R., Ianniello, G. P., Gebbia, V., Gasparini, G., Fava, S., Hirsh, V., Bezjak, A., Seymour, L. and Perrone, F. (2003). Gemcitabine plus vinorelbine compared with cisplatin plus vinorelbine or cisplatin plus gemcitabine for advanced non-small-cell lung cancer: A phase III trial of the Italian GEMVIN Investigators and the National Cancer Institute of Canada Clinical Trials Group. *Journal of Clinical Oncology* **21**, 3025-3034.

Gu, M. and Ying, Z. (1995). Group sequential methods for survival data using partial likelihood score processes with covariate adjustment. *Statistica Sinica.* **5**, 793-804.

Haybittle, J. L. (1971). Repeated assessment of results in clinical trials of cancer treatment. *Brit. J. Radiology* **44**, 793-797.

Heritier, S., Gebski, V. and Pillai, A. (2005). Dynamic balancing randomization in controlled clinical trials. *Stat. Med.* **24**, 3729-3741.

Hu, F. (2012). Statistical issues in trial design and personalized medicine. *Clinical Investigation* **2**, 121-124.

Hu, F., Hu, Y., Ma, W., Zhang L.X. and Zhu, H. (2015). Statistical inference of adaptive randomized clinical trials for personalized medicine. *Clinical Investigation* **5**, 415-425.

Hu, F. and Rosenberger, W.F. (2006). The theory of response-adaptive randomization in clinical trials (Vol. 525). John Wiley & Sons.

Hu, Y. and Hu, F. (2012). Asymptotic properties of covariate-adaptive randomization. *Annals of Statistics* **40**, 1794-1815.

Iacono, A. T., Johnson, B. A., Grgurich, W. F., Youssef, J. G., Corcoran, T. E., Seiler, D. A., Dauber, J. H., Smaldone, G. C., Zeevi, A., Yousem, S. A., Fung, J. J., Burckart, G. J., McCurry, K. R. and Griffith, B. P. (2006). A randomized trial of inhaled cyclosporine in lung-transplant recipients. *The New England Journal of Medicine* **354**, 2, 141-150.

Jakob, S. M., Ruokonen, E., Grounds, R. M., Sarapohja, T., Garratt, C., Pocock, S. J., Bratty, J. R. and Takala, J. (2012). Dexmedetomidine vs midazolam or propofol for sedation during prolonged mechanical ventilation: Two randomized controlled trials. *Journal of American Medical Association* **307**, 1151-1160.

Jennison, C. and Turnbull, B. W. (1997). Group-Sequential Analysis Incorporating Covariate Information. *J. Amer. Statist. Assoc.* **92**, 1330-1341.

Jennison, C. and Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials.* Chapman and Hall/CRC.

Krueger, G. G., Langley, R. G., Leonardi, C., Yeilding, N., Guzzo, C., Wang, Y., Dooley, L. T. and Lebwohl, M. (2007). A human interleukin-12/23 monoclonal antibody for the treatment of psoriasis. *The New England Journal of Medicine* **356**, 580-592.

Lai, H. L., Chen, C. J., Peng, T. C., Chang, F. M., Hsieh, M. L., Huang, H. Y., and Chang, S. C. (2006). Randomized controlled trial of music during kangaroo care on maternal state anxiety and preterm infants' responses. *Int. J. Nurs. Stud.* **43**, 139-146.

Lan, K. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-663.

Lehmann, E. L. (2004). *Elements of Large-Sample Theory.* Springer.

Ma, W., Hu, F. and Zhang, L. (2015). Testing hypotheses of covariate-adaptive randomized clinical trials. *J. Amer. Statist. Assoc.* **110**, 669-680.

Molander, A.,Warfvinge, J., Reit, C. and Kvist, T. (2007). Clinical and radiographic evaluation of one- and two-visit endodontic treatment of asymptomatic necrotic teeth with apical periodontitis: A randomized clinical trial. *Journal of Endodontics* **33**, 1145-1148.

Nordle, O. and Brantmark, B. (1977). A self-adjusting randomization plan for allocation of patients into two treatment groups. *Clin. Pharm. Therap.* **22**, 825-830.

O'Brien, P. C. and Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics* **35**, 549-556.

Ohtori, S., Miyagi, M., Eguchi, Y., Inoue, G., Orita, S., Ochiai, N., Kishida, S., Kuniyoshi, K., Nakamura, J., Aoki, Y., Ishikawa, T., Arai, G., Kamoda, H., Suzuki, M., Takaso, M., Furuya, T., Kubota, G., Sakuma, Y., Oikawa, Y., Toyone, T. and Takahashi, K. (2012). Efficacy of epidural administration of anti-interleukin-6 receptor antibody onto spinal nerve for treatment of sciatica. *European Spine Journal* **21**, 2079-2084.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J. and Smith, P. G. (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Brit. J. Cancer* **34**, 585-612.

Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 191-199.

Pocock, S. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31**, 103-115.

Proschan, M. A., Lan, K. and Wittes, J. T. (2006). *Statistical Monitoring of Clinical Trials, A Unified Approach.* Springer Science+Business Media, LLC.

Rosenberger, W. F. and Sverdlov, O. (2008). Handling covariates in the design of clinical trials. *Statistical Science* **23**, 404-419.

Shao, J., Yu, X. and Zhong, B. (2010). A theory for testing hypotheses under covariate-adaptive randomization. *Biometrika* **97**, 347-360.

Signorini, D. F., Leung, O., Simes, R. J., Beller, E. and Gebski, V. J. (1993). Dynamic balanced randomization for clinical trials. *Stat. Med.* **12**, 2343-2350.

Taves, D. R. (1974). Minimization: A new method of assigning patients to treatment and control groups. *J. Clin. Pharmacol. Therap.* **15**, 443-453.

Tilley, B. C., Alarcón, G. S., Heyse, S. P., Trentham, D. E., Neuner, R., Kaplan, D. A., Clegg,

D. O., Leisen, J. C., Buckley, L., Cooper, S. M., Duncan, H., Pillemer, S. R., Tuttleman, M., and Fowler, S. E. (1995). Minocycline in rheumatoid arthritis. A 48-week, double-blind, placebo-controlled trial. *Ann. Intern. Med.* **122**, 81-89.

Tsiatis, A. A., Rosner, G. L., and Tritchler D. L. (1985) Group sequential tests with censored survival data adjusting for covariates. *Biometrika* **72**, 365-373

Wald, A. (1947). *Sequential Analysis*. John Wiley and Sons Inc., New York.

Wang, S. K. and Tsiatis, A. A. (1987). Approximately optimal one-parameter boundaries for group sequential trials. *Biometrics* **43**, 193-200.

Wei, L. J. (1978). An application of an urn model to the design of sequential controlled clinical trials. *J. Amer. Statist. Assoc.* **73**, 559-563.

Zhu, H. and Hu, F. (2010). Sequential monitoring of response-adaptive randomized clinical trials. *The Annals of Statistics* **38**, 2218-2241.

Zhu, H. and Hu, F. (2012). Interim analysis of clinical trials based on urn models. *Canadian Journal of Statistics* **40**, 550-568.