Contemporary Clinical Trials Communications 7 (2017) 224–230

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Decision on performing interim analysis for comparative clinical trials

Kyongsun Pak[a,*], Susanna Jacobus[b], Hajime Uno[c]

[a] Kitasato University, School of Pharmacy, Department of Clinical Medicine (Biostatistics), 5-9-1 Shirokane, Minato-ku, Tokyo 108-0072, Japan
[b] Dana Farber Cancer Institute – ECOG-ACRIN Biostatistics Center, Boston, MA, USA
[c] Dana Farber Cancer Institute, Department of Medical Oncology, Division of Population Sciences, Boston, MA, USA

ARTICLE INFO

Keywords:
Interim analysis
Blinded analysis
Early termination for superiority
Early termination for futility
Data monitoring

ABSTRACT

In randomized-controlled trials, interim analyses are often planned for possible early trial termination to claim superiority or futility of a new therapy. While unblinding is necessary to conduct the formal interim analysis in blinded studies, blinded data also have information about the potential treatment difference between the groups. We developed a blinded data monitoring tool that enables investigators to predict whether they observe such an unblinded interim analysis results that supports early termination of the trial. Investigators may skip some of the planned interim analyses if an early termination is unlikely. We specifically focused on blinded, randomized-controlled studies to compare binary endpoints of a new treatment with a control. Assuming one interim analysis is planned for early termination for superiority or futility, we conducted extensive simulation studies to assess the impact of the implementation of our tool on the size, power, expected number of interim analyses, and bias in the treatment effect. The numerical study showed the proposed monitoring tool does not affect size or power, but dramatically reduces the expected number of interim analyses when the effect of the treatment difference is small. The tool serves as a useful reference when interpreting the summary of the blinded data throughout the course of the trial, without losing integrity of the study. This tool could potentially save the study resources and budget by avoiding unnecessary interim analyses.

1. Introduction

In randomized-controlled trials, interim analyses are often planned to review the efficacy or safety of the therapeutic interventions. Early termination of the trial may occur due to evidence of superiority or futility of the new therapy based on the interim analysis. To conduct interim analyses, we need to access the data prior to the completion of the trial. Particularly for blinded studies, interim analysis requires unblinding of the treatment allocation and conducting a formal between-group comparison [1,2]. Although unblinded data provide complete information of the observed data, blinded data also contain information about the treatment difference between the groups. For instance, when the observed response rate in the pooled sample is very low at the time of the interim analysis, we know the response rates in both groups are very low. Therefore, there is little chance a significant difference between the groups would be observed and, consequently, a formal comparison is a wasteful expenditure of alpha. Even when response rates are not that small, if the control rate can be reasonably estimated based on previous studies, the blinded data yields a decent estimate of the treatment difference.

There are several data monitoring tools [3–5] that use blinded data originating in the Bayesian approach for safety monitoring in single arm studies proposed by Thall and Simon [6]. For example, Ball [3] focused on the adverse event rate in the pooled sample and proposed a decision rule based on the posterior distribution of it using the Bayesian approach. On the other hand, our focus in this paper is a blinded data monitoring tool predicting the result of a formal unblinded interim analysis for superiority or futility of a new therapy. The proposed tool works with the hypothesis testing approach. Specifically, we assume that the alpha spending function approach [7] is used as a stopping guideline for superiority in the formal interim analysis. For futility, we assume that the result of stochastic curtailment method is used as a guideline of early stopping [8]. We performed extensive numerical studies to assess the impact of the implementation of the data monitoring tool on the type I error rate, power, expected sample size, expected number of interim analyses to be performed and bias in the treatment effect for both superiority and futility. We illustrated the practical application of our tool, using data from a clinical trial conducted by the ECOG-ACRIN Cancer Research Group. With our tool, investigators may skip some of the planned interim analyses when the result of an interim analysis at that time point is unlikely to support early termination of the trial for superiority or futility. Therefore, this

* Corresponding author.
E-mail address: pakk@pharm.kitasato-u.ac.jp (K. Pak).

http://dx.doi.org/10.1016/j.conctc.2017.08.001
Received 27 February 2017; Received in revised form 7 August 2017; Accepted 8 August 2017
Available online 16 August 2017

tool could ultimately avoid unnecessary spending of study resources while maintaining scientific integrity of the trial.

2. Methods

In this paper, we specifically focus on randomized controlled trials comparing binary endpoints, namely response rates, between a new therapy and a control. In the trial, interim analyses are planned for early termination for superiority or futility or both.

2.1. Typical procedure of interim analysis

Usually, the interim analysis is implemented at the time when the pre-planned information fraction is reached. For a binary outcome, the total information will be defined as the planned total sample size. Assume that, during the accumulating the preset sample size M, there are N $(\leq M)$ participants and T $(\leq N)$ responders in the two arms at the time of the interim analysis. Let (T_1, T_0) denote the numbers of responders in the arm of the new therapy and control respectively, and then $T = T_1 + T_0$. When unblinding the data, we can observe (T_1, T_0), and formal comparison would be implemented. Depending on the resulting test statistic, or the corresponding p-value or conditional power, we decide whether to stop or continue the trial.

2.2. Blinded data monitoring tool

Before breaking the blinded treatment assignment code, we may monitor (N, T) from the blinded data. Assume that each T_1 and T_0 follows a binomial distribution with a parameter p_1 for the new therapy and p_0 for the control therapy, respectively. The probability mass function of T, $Pr(T = t)$, can be expressed with a mixture of the aforementioned two binomials. Given the allocation ratio during the study q: $(1 - q)$ for the new therapy and control respectively, where $q \in (0,1)$, $Pr(T = t)$ is expressed that

$$Pr(T = t) = \binom{N}{t}\{qp_1 + (1-q)p_0\}^t\{q(1-p_1) + (1-q)(1-p_0)\}^{N-t}.$$

With the blinded treatment allocation, if we have enough certainty about p_0 and if the allocation ratio is close to q, we would be able to predict the response rate of the new therapy p_1. Specifically, if p_0 is a known value, the maximum likelihood estimator of p_1 is obtained by

$$\widehat{p_1} = \frac{T - N(1-q)p_0}{Nq}.$$

Then the standardized test statistics for testing the null hypothesis H_0: $p_1 = p_0$ is given by $Z_b = (\widehat{p_1} - p_0)/\sqrt{\widehat{\mathrm{Var}}(\widehat{p_1})}$, where $\widehat{\mathrm{Var}}(\widehat{p_1}) = N\hat{r}(1-\hat{r})/(Nq)^2$ and $\hat{r} = q\widehat{p_1} + (1-q)p_0$. Utilizing the observed Z_b at the interim analysis point, we can predict whether or not the unblinded interim analysis result will meet the stopping criteria for superiority or futility. For superiority, one can then obtain the threshold values of the total number of responders T with respect to each number of subjects N, with which the p-value of the test would meet the pre-specified stopping criteria corresponding to the information time at the interim analysis. For futility, one might use a conditional probability as criteria for stopping.

2.3. Illustrative example

To illustrate the aforementioned decision criteria, we consider a specific numerical example of a randomized controlled trial comparing

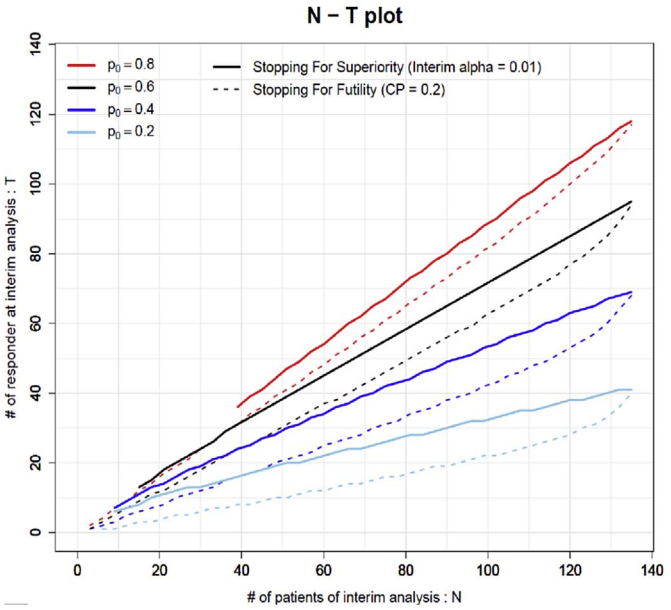

Fig. 1. N-T plot with $\mathbf{p}_0$= 0.8, 0.6, 0.4 and 0.2 for early stopping for superiority and futility.

response rates between the new and the control therapy. The accrual goal is 135 patients and the mixture proportion of allocation is q: $(1 - q) = \frac{2}{3}$: $\frac{1}{3}$ for the new therapy and the control, respectively.

First, we consider the case for interim analysis expecting early termination only for superiority and consuming type I error rate $\alpha = 0.01$ at the interim analysis. Under this scenario, the solid curves in Fig. 1 show the thresholds of N and T with various values of p_0. For example, the blue solid curve corresponds the case that $p_0 = 0.4$. Using the observed (N, T) with blindness maintained, these curves can be a reference to predict how likely the interim analysis result would meet the stopping criteria, if conducted. Specifically, in this example, when the observed (N, T) is above the blue curve, we can expect that the result of the interim analysis will support early stopping for superiority for the new therapy. Therefore, if we think that p_0 is very likely to be 0.4, we would conclude that an interim analysis should not be missed at this point.

Next, we consider the case of early termination for futility based on the conditional power less than 0.2. The dashed lines in Fig. 1 show the corresponding N-T curves for futility with various p_0's. Again, consider the case that $p_0 = 0.4$. The observed (N, T) below the blue dashed curve indicates the conditional power will likely be below 0.2. Thus, if we are confident with a $p_0 = 0.4$, we would determine the interim analysis should not be missed for potential futility stop. On the other hand, if the observed (N, T) is above the blue dashed curve, it may be an option to skip the scheduled interim analysis, if there are no other concerns on the study.

This tool can also be used for the cases that both superiority and futility stoppings are of interest. In those cases, we will use both solid and dashed curves in Fig. 1. When the observed (N, T) is in between solid and dashed lines, the interim analysis result will likely not meet stopping criteria for either superiority or futility. We may skip performing the interim analysis and continue the trial, unless there are other concerns in the study.

In this manner, the proposed blinded monitoring process is helpful for identifying whether it is a good time to conduct interim analysis, preserving the integrity of the study. Appendix 1 provides the computer programs to generate N-T plots with a documented example.