Phase 2 Selection Trials With a Prospective Control

Ying Kuen Cheung[1]

Department of Biostatistics, Columbia University,
722 West 168th Street, Room 641,
New York, NY 10032, USA
yc632@columbia.edu

Abstract. The paper reviews the SPRT-type selection procedures for phase 2 clinical trials, whose objective is to identify an experimental treatment that is more effective than a prospective control, or to declare futility if no such treatment exists. The SPRT-type procedures has a practical advantage in that the selection boundaries can be easily chosen with respect to a given set of error constraints. This paper illustrates the method using normal endpoints (Cheung, 2009, *Journal of Biopharmaceutical Statistics* in press), although the same principles has also been applied to binomial outcomes (Cheung, 2008, *Biometrics* **64**, 940–949).
Keywords. Sample size re-estimation; Sequential elimination; Sequential probability ratio test; Symmetric boundaries; Type I error.

1 Introduction

We address the same selection problem as in Cheung (2009) for phase 2 clinical trials with a set of treatments $\{0, 1, \ldots, K\}$, where 0 represents the control group. Precisely, let $X_{ij} \sim N(\mu_i, \sigma^2)$ denote a desirable outcome of the jth patient in arm i. We consider two particular scenarios (hypotheses). First, under the "global null" where $\mu_0 = \cdots = \mu_K$, we would like to control the selection probability for the control, denoted by P_0, to be at or above $1 - \alpha$. The value α in this context can be viewed as an extension of type I error rate from a hypothesis test to treatment selection. Second, when there is in truth a clinically superior treatment, i.e., $\mu_0 = \mu_1 = \cdots = \mu_{K-1}$ and $\mu_K = \mu_0 + \delta$ for some prespecified $\delta > 0$, our goal is to keep the selection probability for arm K, denoted by P_1, at about $1 - \beta$. The value β is analogous to type II error in a hypothesis test.

2 An SPRT-type selection procedure

We address the same selection problem as in Cheung (2009) for phase 2 clinical trials with a set of treatments $\{0, 1, \ldots, K\}$, where 0 represents the control group. Precisely, let $X_{ij} \sim N(\mu_i, \sigma^2)$ denote a desirable outcome of the jth patient in arm i. We consider two particular scenarios (hypotheses). First, under the "global null" where $\mu_0 = \cdots = \mu_K$, we would like to control the selection probability for the control, denoted by P_0, to be at or above $1 - \alpha$. The value α in this context can be viewed as an extension of type I error rate from a hypothesis test to treatment selection. Second, when there is in truth a clinically superior treatment, i.e., $\mu_0 = \mu_1 = \cdots = \mu_{K-1}$ and $\mu_K = \mu_0 + \delta$ for some prespecified $\delta > 0$, our goal is to keep the selection probability for arm K, denoted by P_1, at about $1 - \beta$. The value β is analogous to type II error in a hypothesis test.

THEOREM 1. Suppose the shifted outcome $Y_{ij} \sim N(\theta_i, \sigma^2)$ and $\theta_k \geq \theta_i$ for $i \neq k$. Then if the enrollment processes do not depend on θ_i's, an open-ended SPRT (i.e., $N_{\max} = \infty$) will correctly select arm k with probability bounded below by

$$\left\{ \sum_{i=0}^{K} \exp\left[\frac{2d}{\sigma^2}(\theta_i - \theta_k) \right] \right\}^{-1}. \tag{1}$$

If we run the SPRT with $a_0 = \cdots = a_K$ under the global null, then $\theta_0 = \cdots = \theta_K$ and the lower bound (1) is equal to $(1 + K)^{-1}$. By symmetry, we can deduce that the probability of selecting arm k, or any arm, by the SPRT is equal to $(1 + K)^{-1}$. Thus, the lower bound is exact in this case. Furthermore, we observe that the probability of selecting an experimental arm under the global null (i.e., type I error rate) will equal $K/(K + 1)$ which is apparently too large to be considered in practice. Therefore, we need to choose the shifts $a_0, a_1, \ldots, a_K$ differently so as to satisfy conventional error constraints.

COROLLARY 1. If the shifts are chosen such that $a_0 > a_i$ for $i = 1, \ldots, K$, then P_0 is bounded below by

$$\mathrm{LB}_0 = \left\{ \sum_{i=0}^{K} \exp\left[\frac{2d}{\sigma^2}(a_i - a_0) \right] \right\}^{-1}$$

which increases and converges to 1 as $d \to \infty$.

COROLLARY 2. If the shifts can be chosen such that $a_K + \delta > a_i$ for $i = 0, 1, \ldots, K-1$, then P_1 is bounded below by

$$\mathrm{LB}_1 = \left\{ \sum_{i=0}^{K-1} \exp\left[\frac{2d}{\sigma^2}(a_i - a_K - \delta) \right] + 1 \right\}^{-1}$$

which increases and converges to 1 as $d \to \infty$.

The corollaries guarantee we can always find a constant d for any given error constraints if the shifts are chosen to satisfy the conditions in the corollaries. Applications of these results are given in Section 3. The proof of Theorem 1 can be undertaken in the same manner as in Levin and Robbins (1981) who extend the SPRT to the multi-arm selection problem with binomial data and equal sample sizes, and is available from the author upon request.

3 Choosing the selection boundaries

3.1 Design parameters

In situations where the experimental regimens are exchangeable *a priori*, we may set $a_1 = \cdots = a_K$. Then the lower bounds become $\mathrm{LB}_0 = \left\{1 + K \exp\left[2d(a_1 - a_0)/\sigma^2\right]\right\}^{-1}$ and $\mathrm{LB}_1 = \left\{\exp\left[2d(a_0 - a_1 - \delta)/\sigma^2\right] + (K-1)\exp\left(-2d\delta/\sigma^2\right) + 1\right\}^{-1}$. We observe that LB_0 and LB_1 depend on a_0 and a_1 only through their difference, and therefore will set $a_1 = 0$ without loss of generality. As a result, we need $0 < a_0 < \delta$ in order to satisfy the conditions in Corollary 1 and Corollary 2, which then give

$$\mathrm{LB}_0 = \left[1 + K \exp\left(-\frac{2da_0}{\sigma^2}\right)\right]^{-1} \text{ and } \mathrm{LB}_1 = \left\{\exp\left[\frac{2d(a_0 - \delta)}{\sigma^2}\right] + (K-1)\exp\left(-\frac{2d\delta}{\sigma^2}\right) + 1\right\}^{-1}.$$

This constraint on the choice of a_0 is intuitive: it needs to be positive so that the control will look favorable under the global null with $\mu_0 + a_0 > \mu_i$, but smaller than δ so that a treatment with a clinically significant improvement remains superior in terms of the shifted mean. With the clinician-defined parameters δ, K and the error constraints α, β specified, the lower bounds LB_0 and LB_1 for the probability of correct selection depends on a_0, d, and the true variance σ^2.

3.2 d-minimal design

For a given a_0 that is between 0 and δ, LB_0 and LB_1 are increasing functions of d. This is expected because a larger value of d invokes trial termination or treatment elimination when more information has been accrued, and hence the decision is less likely to be error-prone. For the same reason, SPRT with a smaller d are expected to conclude a trial with fewer patients than when a larger d is used. We take the design approach in Cheung (2008, 2009) whereby the shift a_0 is chosen to be d-minimal: a shift a_0^* is d-minimal when it minimizes the required termination constant d^* for given error constraints $\mathrm{LB}_0 \geq 1 - \alpha$ and $\mathrm{LB}_1 \geq 1 - \beta$ and a given set of clinician-defined parameters δ and K. In particular, Cheung (2009) gives

$$a_0^* = \delta \left[\frac{\log K - \mathrm{logit}(\alpha)}{\log\left(\frac{K}{\alpha} - 1\right) - \mathrm{logit}(\beta)} \right] \text{ and } d^* = \frac{\sigma^2}{2\delta}\left[\log\left(\frac{K}{\alpha} - 1\right) - \mathrm{logit}(\beta) \right]. \tag{3.3}$$

Since a_0^* does not depend on the true variance σ^2, each observation in the control arm will be shifted by the same constant throughout the trial. Note that $0 < a_0^* < \delta$ when $\beta < 0.5$, or more precisely, $\beta < (K-\alpha)/\{K-\alpha+K(1-\alpha)\}$. In other words, the d-minimal criterion can be applied if the research team sets its goal to identify the superior experimental arm with a target probability greater than 0.5.

The termination constant d^*, on the other hand, does depend on σ^2. When implementing the method, we could repeatedly estimate σ^2 with the unbiased pooled sample variance $\hat{\sigma}^2$ throughout the trial. Therefore, the termination criteria will be slightly different at each interim; empirically, we find that the estimate of d^* becomes quite stable when σ^2 is estimated with at least 30 degrees of freedom.

3.3 Sample size determination

The lower bound formulae derived from (1) are based on the open-ended SPRT. While we expect the theoretical results will hold if the truncation $N_{\max}$ is sufficiently large, the choice of a sufficiently large $N_{\max}$ apparently depends also on σ^2. Since the motivation of a sequential design is to improve the enrollment feasibility on the single-stage design, we may initially set to truncate the sequential procedures at $N_{\max}$, which is defined according to a single-step procedure with respect to an assumed σ_0^2; see Cheung (2009). This guarantees the sequential designs adopted will always enroll fewer patients than the single-stage design. As seen in extensive simulations, the truncation $N_{\max}$ thus computed keeps the actual error rates at the target level, if the true variance is less than or equal to σ_0^2.

4 Discussion

This article reviews a SPRT-type design for treatment selection with a prospective control group using a normal endpoint. Extensive simulations have shown substantial gain in sample size over the traditional single-step design. Furthermore, Cheung (2009) suggest two practical modifications. First, a sequential elimination using the SPRT selection boundaries will likely further reduce sample size with comparable accuracies. Second, to anticipate the situations when the true σ^2 is larger than the assumed σ_0^2, we may re-estimate the truncation $N'_{\max}$ based on the pooled sample variance *if* the trial reaches $N_{\max}$ without reaching the termination criteria. That is, continue with an additional $N'_{\max} - N_{\max}$ subjects if $N'_{\max} > N_{\max}$; stop the trial and select the arm with the largest observed mean if otherwise. These two practical measures have been evaluated in Cheung (2009) and are recommended for practical use.

Finally, the proposed method is versatile and can be adapted for different outcome types. Cheung (2008) studies analogous procedures for binomial outcomes; we are currently working an extension for outcomes following a distribution from the exponential family.

5 Acknowledgement

This work was supported by NIH grant R01 NS055809-01.

References

Cheung, Y. K. (2008). Simple sequential boundaries for treatment selection in multi-armed randomized clinical trials with a control. *Biometrics* **64,** 940–949.

Cheung, Y. K. (2009). Selecting promising treatments in randomized phase II cancer trials with an active control. *J. Biopharm. Stat.* in press.

Levin, B., Robbins, H. (1981). Selecting the highest probability in binomial or multinomial trials. *Proceedings of the National Academy of Sciences of the United States of America* 78:4663-4666.

Robbins, H., Siegmund, D. O. (1974). Sequential tests involving two populations. *Journal of the American Statistical Association* 69:132-139.