

Finding Critical Values to Control Type I Error for a Biomarker Informed Two-Stage Winner Design

Jing Wang^{1,2*}, Sandeep Menon^{1,4} and Mark Chang^{1,3}

¹Department of Biostatistics, Boston University, 801 Massachusetts Avenue, Boston, MA 02118, USA

²Gilead Sciences, Inc. 303 Velocity Way, Foster City, CA 94404, USA

³AMAG Pharmaceuticals, Inc., 100 Hayden Avenue, Lexington, MA 02420, USA

⁴Pfizer, Inc., 35 Cambridge Park Drive, Cambridge, MA 02140, USA

Background

Adaptive clinical trial designs have been getting very popular in recent years. The PhRMA Working Group defines an adaptive design as a clinical study design that uses accumulating data to direct modification of aspects of the study as it continues, without undermining the validity and integrity of the [1]. These designs can assist in potentially accelerating clinical development and improving efficiency.

However, the multiple interim looks and adaptive adjustments with the design can lead to inflation of type I error. Over the past decade, several statistical approaches have been proposed to control the inflation, some of which have been widely applied in practice. Some of these approaches include: error spending approach for classical group sequential plans [2-4]; Combination of p-values, such as Fisher's combination test [5,6], Inverse Normal Method [7], sum of p-values approach [8]; conditional error function [9-11]; fixed weighting method [12]; variance spending method [13,14]; and multiple testing methodology such as closed test procedures [15-17].

In addition to the conventional adaptive designs which use the same endpoint at the interim and the final analysis of the study, with the surge in advanced technology especially in the "OMICS" space (eg. Genomics, proteomics, etc), interest has also been drawn towards the biomarker informed adaptive clinical trial designs recently. The biomarker informed adaptive designs make interim decisions based on inference on a potentially predictive biomarker, which may be a short-term endpoint that is indicative of the behavior of the primary endpoint.

Todd and Stallard [18] proposed a statistical approach to control type I error rate for group sequential trials where the interim treatment selections are based upon only on the biomarker. Stallard [19] later proposed a method for group sequential trials that use both the available biomarker and primary endpoint information for treatment selections. Their method controls the type I error rate in the strong sense. Friede et al. [20] considered a biomarker informed drop-the-losers design using combination tests for adaptive designs and closure principle for multiple testing to achieve strong control of the family-wise type I error rate. Scala and Glimm [21] studied the case with correlated time-to-event biomarker and primary endpoint where Bayesian predictive power combining evidence from both endpoints is used for interim selection, they investigated the precise conditions under which type I error control is attained. Jenkins et al. presented a type I error control approach for an enrichment design with survival biomarker and primary endpoint which allows both the subgroup and the full population as co-primary populations.

Shun et al. [22] studied a biomarker informed two-stage winner design with normal endpoints. In the design, the interim decisions are made by ranking of the observed effects of biomarker. They derived the unconditional distribution of the final test statistic for the design with two active treatment arms and one control arm, and proposed its

normal approximation for calculation of the critical value to preserve type I error rate. However, the proposed normal approximation procedure by Shun et al. [22] cannot be extended to designs with more active treatment groups. In this manuscript, we extend their work and propose a novel type I error control approach for biomarker informed two-stage winner design that can accommodate multiple active arms. Our approach preserves the type I error rate by adjusting critical rejection values of the final test statistic of the design. The exact distribution of the final test statistic is derived and R functions for calculating the adjusted critical rejection values from the skewed distribution are developed. The critical rejection values associated with one-sided type I error rate 0.025 for biomarker informed two-stage winner design with up to 7 active treatment groups are also tabulated for easy reference.

Since biomarker informed adaptive design have two endpoints i.e. the biomarker endpoint and clinical (or primary study endpoint) endpoint, it is important to define the robust model to describe the relationship between the two endpoints. Shun et al. [22] used the conventional approach to model the two normal endpoints in the biomarker informed two-stage winner design they considered, that is, a bivariate normal distribution with a correlation coefficient is used for modeling the two endpoints. However, the conventional approach is shown to be inappropriate when little historical knowledge is known about how the means of the two endpoints are related [23]. Wang et al. [23] proposed a two-level correlation model to describe the relationship between the two endpoints. Besides the correlation coefficient between the two endpoints, the uncertainty of the estimated mean level correlation between the two endpoints is also considered in their model. The two-level correlation model incorporates a new variable that describes the mean level correlation between the two endpoints. The new variable, together with its distribution, reflects the uncertainty about the mean-level relationship between the two endpoints due to a small sample size of historical data. It is shown in Wang et al. [23] that the two-level correlation model is a better choice for modeling the two endpoints than the conventional model. And in fact, the conventional model is a special case of the two-level correlation model. In this manuscript, we consider both (conventional and two-level correlation) models in our discussion for finding critical rejection values for final test statistic to preserve type I error rate.

*Corresponding author: Jing Wang, Gilead Sciences, Inc. 303 Velocity Way, Foster City, CA 94404, USA, Tel: 650-522-1167; E-mail: jwa222@bu.edu

Received May 06, 2014; Accepted July 04, 2014; Published July 10, 2014

Citation: Wang J, Menon S, Chang M (2014) Finding Critical Values to Control Type I Error for a Biomarker Informed Two-Stage Winner Design. J Biomet Biostat 5: 207. doi:10.472/2155-6180.1000207

Copyright: © 2014 Wang J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Finding Critical Rejection Values for Final Test Statistic in Biomarker Informed Two-Stage Winner Design

Biomarker informed two-stage winner design

In general, a biomarker informed adaptive design is a design that combines a phase II and a phase III study. It starts with several active treatment arms and a control arm with planned interim analyses on biomarker. At interim, the inferior arms will be terminated based on effects of biomarker (either by hypothesis testing or ranking of observations), and only the most promising treatment (“winner”) will be carried to the end of the study with the control arm. The final comparison between the winner arm and the control arm will be performed on data from both stages and on study primary endpoint. This design has the potential to shorten the duration of the trial for drug development and can be cost effective.

Shun et al. [22] studied a “biomarker informed two-stage winner design”. This design only has one interim look when each treatment group has n_1 patients (n_1 is the interim sample size), and uses ranking of biomarker observations for the interim selection. Additional n_2 patients will be recruited for the winner arm and the control arm, and the final comparison will be performed on the primary endpoint of the $2N$ ($N = n_1 + n_2$) patients.

Let K be number of treatment groups ($K-1$ active treatment groups, and 1 control group), and N be the maximum sample size for each treatment group. Assume the interim analysis is planned at the information time $I = \frac{n_1}{N}$, where $n_1 < N$. Two sets of measurements are obtained: $\{X_i^{(j)} | i=1, \dots, n_1\}$, the measurements of the biomarker at interim stage for i th person in j th treatment group; and $\{Y_i^{(j)} | i=1, \dots, N\}$, the measurements of the study primary endpoint at final stage for i th person in j th treatment group. $j=0, 1, \dots, K-1$. $j=0$ represents the control group while $j=0, 1, \dots, K-1$ the active treatment groups. Let $\bar{X}_n^{(j)} = \frac{1}{n_1} \sum X_i^{(j)}$ be the mean of the biomarker measurements for treatment group j at interim, and $\bar{Y}_N^{(j)} = \frac{1}{N} \sum Y_i^{(j)}$ be the mean of the primary endpoint measurements for treatment group j at final.

In the biomarker informed two-stage winner design, the interim decision rule is that if the interim biomarker observations $\bar{x}_n^{(j)} = \max(\bar{x}_n^{(1)}, \dots, \bar{x}_n^{(K-1)})$, select treatment j as the most effective treatment, and carry only treatment group j and the control group to the end of the study. When the interim biomarker outcomes are almost the same, the option that more than one treatment groups be carried to the end of the study is not considered in this design, because either treatment group can be selected in this case. The final assessment will be based on the study primary endpoint Y comparing the selected treatment group and the control group.

The two models for fitting the two endpoints

In this section, we briefly review two commonly used techniques for modeling the two endpoints (i.e., the biomarker and the study primary endpoint) in a biomarker informed two-stage winner design. We assume both endpoints are normally distributed in our context. For endpoints that are not normally distributed, a transformation could be considered.

The first is the conventional approach, which uses a bivariate normal distribution with a correlation coefficient to fit the two endpoints.

In this approach, the individual-level correlation coefficient ρ is the only variable to describe the relationship between the biomarker and the primary endpoint. The second is the two-level correlation model proposed by Wang, et al. [23], where a conditional bivariate normal distribution is used to model the two endpoints. This model considers both the individual-level and mean-level correlation between the biomarker and the primary endpoint.

More specifically, let u_j^x be the mean of biomarker for treatment group j , σ_x^2 be the variance. For a fixed j , assume $\{X_i^{(j)} | i=1, \dots, n_1\}$ i.i.d., and $X_i^{(j)} \sim N(u_j^x, \sigma_x^2)$. Denote the standardized mean of biomarker for each treatment group by $u_j^{x*}, u_j^{x**} = \frac{u_j^x}{\sigma_x}$.

Let u_j^y be the mean of study primary endpoint for treatment group j , σ_y^2 be the variance. For a fixed j , assume $\{Y_i^{(j)} | i=1, \dots, N\}$ i.i.d. and $Y_i^{(j)} \sim N(u_j^y, \sigma_y^2)$. Denote the standardized mean of study primary endpoint for treatment group j by $u_j^{y*} = \frac{u_j^y}{\sigma_y}$.

Assume also that $\{X_i^{(j)} | i=1, \dots, n_1\}$ and $\{Y_i^{(j)} | i=1, \dots, n_1\}$ are correlated with a correlation ρ for the same j and i , that is $\rho = \text{Corr}(X_i^{(j)}, Y_i^{(j)})$.

Since both endpoints are assumed to be normally distributed, the conventional approach uses a multivariate normal distribution with a correlation coefficient ρ for modeling the relationship between the biomarker and primary endpoint:

$$\begin{pmatrix} X_i^{(j)} \\ Y_i^{(j)} \end{pmatrix} \sim N \left(\begin{pmatrix} u_j^x \\ u_j^y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \right), j=0, 1, \dots, K-1.$$

This approach was used for modeling the two endpoints in the study of Shun et al. [22], etc. It has the limitation that the means for both endpoints have to be specified while running power simulation, which can be challenging especially in lieu of lack of solid historical knowledge about the relationship between the biomarker and the primary endpoint which is not uncommon. The two-level correlation model proposed by Wang et al. [23] define a new variable, R_j into the model, which was referred to the estimated mean level correlation between the biomarker and the primary endpoint. This new variable along with its distribution reflects the uncertainty of the mean-level relationship between the two endpoints especially due to a small sample size of historical data.

Assume R_j is normally distributed and centered at r_j ; $r_j = \frac{u_j^{y*}}{u_j^{x*}}$, which is the true mean-level correlation between the two endpoints. The two-level correlation model can be written as follows:

$$\begin{pmatrix} X_i^{(j)} \\ \sigma_x \\ Y_i^{(j)} | R_j \end{pmatrix} \sim N \left(\begin{pmatrix} u_j^{x*} \\ R_j u_j^{x*} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

$$R_j \sim N(r_j, \sigma_{rj}^2), j=0, 1, \dots, K-1.$$

And the unconditional distribution of the model could be expressed as follows:

$$\begin{pmatrix} X_i^{(j)} \\ \sigma_x \\ Y_i^{(j)} \end{pmatrix} \sim N \left(\begin{pmatrix} u_j^{x*} \\ u_j^{x*} \end{pmatrix}, \begin{pmatrix} 1 & \rho\sqrt{(u_j^{x*})^2\sigma_{rj}^2+1} \\ \rho\sqrt{(u_j^{x*})^2\sigma_{rj}^2+1} & 1 \end{pmatrix} \right), j=0, 1, \dots, K-1.$$

Notice that when $\sigma_{rj}^2 = 0$, the two-level correlation model could be specialized to

$$\begin{pmatrix} X_i^{(j)} \\ Y_i^{(j)} \end{pmatrix} \sim N \left(\begin{pmatrix} u_j^x \\ u_j^y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \right), j=0,1,\dots,K-1.$$

which is the conventional model.

As discussed in Wang et al. [23], the conventional approach is easy to overestimate the power of a biomarker informed two-stage winner design when historical knowledge about the two endpoints is not solid, while the two-level correlation model provides reasonable results as the uncertainty about the mean-level correlation is taken into account in the model. Both these models will be considered in the next two sections when we derive the distribution of the final test statistic for the biomarker informed two-stage winner design.

Test statistic and critical rejection region using conventional approach

To prevent type I error inflation of the biomarker informed two-stage winner design, we use the concept of adjusting critical rejection values of the final test statistic of the design. In this section, we derive the exact distribution and give the critical rejection region for the final test statistic of the biomarker informed two-stage winner design under the conventional model. As shown in the succeeding sections the proposed approach works well for biomarker informed two-stage winner designs with any number of active arms.

Consider the following hypotheses:

$$\begin{aligned} H_0 : u_1^y = \dots = u_{K-1}^y = u_0^y \\ H_1 : u_1^y > u_0^y \text{ or } \dots \text{ or } u_{K-1}^y > u_0^y \end{aligned} \quad (1)$$

It is reasonable to assume that $u_1^x = \dots = u_{K-1}^x = u_0^x$ when $u_1^y = \dots = u_{K-1}^y = u_0^y$ and $\rho \neq 0$. For simplicity, assume σ_y^2 is known.

Let G_j be the test statistic comparing the primary endpoint of the j th treatment group and the control group:

$$G_j = \sqrt{\frac{N}{2\sigma_y^2}} (\bar{Y}_N^{(j)} - \bar{Y}_N^{(0)})$$

where $\bar{Y}_N^{(j)} = \frac{1}{N} \sum Y_i^{(j)}$ is the mean of the primary endpoint measurements for treatment group j at final.

It could be shown that,

$$G_j \sim N \left(\frac{u_j^y - u_0^y}{\sqrt{\frac{2\sigma_y^2}{N}}}, 1 \right)$$

and under $H_0, G_j \sim N(0, 1)$.

The final test statistic of the biomarker informed two-stage winner design can then be expressed as:

$$W = G_j, \text{ if } \bar{X}_{n_1}^{(j)} = \max(\bar{X}_{n_1}^{(1)}, \dots, \bar{X}_{n_1}^{(K-1)}), j=1, \dots, K-1.$$

where, $\bar{X}_{n_1}^{(j)} = \frac{1}{n_1} \sum X_i^{(j)}$ is the mean of the biomarker measurements for treatment group j at interim.

That is, conditional on the interim selection, W takes on the value of the effect from the “winner” treatment group as the final test statistic.

For the very general case under H_1 , the distribution of the final test statistic W could be derived as:

$$\begin{aligned} F_w(w) &= P(W < w) = \sum_{j=1}^{K-1} P(G_j < w, \bar{X}_{n_1}^{(j)} = \max(\bar{X}_{n_1}^{(1)}, \dots, \bar{X}_{n_1}^{(K-1)})) \\ &= \sum_{j=1}^{K-1} P(G_j < w, \bar{X}_{n_1}^{(j)} - \bar{X}_{n_1}^{(m_1)} > 0, \dots, \bar{X}_{n_1}^{(j)} - \bar{X}_{n_1}^{(m_{K-2})} > 0) \\ &= \sum_{j=1}^{K-1} P(P_{j0} < a_{j0}, P_{j1} > a_{j1}, \dots, P_{j,K-2} > a_{j,K-2}) \end{aligned} \quad (2)$$

where $P_{j0} = G_j - \sqrt{\frac{N}{2\sigma_y^2}}(u_j^y - u_0^y)$

$$P_{j1} = \frac{\bar{X}_{n_1}^{(j)} - \bar{X}_{n_1}^{(m_1)}}{t} - \left(\frac{u_j^x}{t} - \frac{u_{m_1}^x}{t} \right)$$

$$P_{j,K-2} = \frac{\bar{X}_{n_1}^{(j)} - \bar{X}_{n_1}^{(m_{K-2})}}{t} - \left(\frac{u_j^x}{t} - \frac{u_{m_{K-2}}^x}{t} \right)$$

$$a_{j0} = w - \sqrt{\frac{N}{2\sigma_y^2}}(u_j^y - u_0^y)$$

$$a_{j1} = - \left(\frac{u_j^x}{t} - \frac{u_{m_1}^x}{t} \right)$$

$$a_{j,K-2} = - \left(\frac{u_j^x}{t} - \frac{u_{m_{K-2}}^x}{t} \right)$$

$$t = \sqrt{\frac{\sigma_x^2}{n_1}}$$

$$j \neq m_1 \neq \dots \neq m_{K-2}, j, m_1, \dots, m_{K-2} \in \{1, 2, \dots, K-1\}$$

$(P_{j0}, P_{j1}, \dots, P_{j,K-2})$ is from a multivariate normal distribution.

In general,

$$\begin{pmatrix} P_{j0} \\ P_{j1} \\ P_{j2} \\ \vdots \\ P_{j,K-2} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma & \gamma & \dots & \gamma \\ \gamma & 2 & 1 & \dots & 1 \\ \gamma & 1 & 2 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma & 1 & \dots & 1 & 2 \end{pmatrix} \right)$$

for any value of $K (K \geq 3)$.

Thus, for a biomarker informed two-stage winner design with $K (K \geq 3)$ treatments, the distribution of final test statistic under the conventional model is:

$$F_w(w) = \sum_{j=1}^{K-1} \int_{-\infty}^{w - \sqrt{\frac{N}{2\sigma_y^2}}(u_j^y - u_0^y)} \dots \int_{\frac{u_j^x}{\sqrt{\frac{\sigma_x^2}{n_1}}} - \frac{u_{m_1}^x}{\sqrt{\frac{\sigma_x^2}{n_1}}}}^{\dots} \int_{\frac{u_j^x}{\sqrt{\frac{\sigma_x^2}{n_1}}} - \frac{u_{m_{K-2}}^x}{\sqrt{\frac{\sigma_x^2}{n_1}}}}^{+\infty} f(p_{j0}, \dots, p_{j,K-2}) dp_{j,K-2} \dots dp_{j0} \quad (3)$$

where:

$$j \neq m_1 \neq \dots \neq m_{K-2}, j, m_1, \dots, m_{K-2} \in \{1, 2, \dots, K-1\},$$

$$f(p_{j0}, \dots, p_{j,K-2}) \text{ is p.d.f. of } N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma & \gamma & \dots & \gamma \\ \gamma & 2 & 1 & \dots & 1 \\ \gamma & 1 & 2 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma & 1 & \dots & 1 & 2 \end{pmatrix} \right), \text{ and}$$

$$\gamma = \sqrt{\frac{n_1}{2N}} \rho.$$

It can be seen that, the interim treatment selection of the design skewed the distribution of its final test statistic. Hence, appropriate statistical adjustment is necessary in order to preserve the type I error rate of the design. As the general distribution of the final test statistic is written, the type I error rate of the design can be preserved by adjusting the critical rejection value for the final test statistic.

Under H_0 , the distribution of the final test statistic (3) can be written as follows:

$$F_w(w) = \sum_{j=1}^{K-1} P(P_{j0} < w, P_{j1} > 0, \dots, P_{j,K-2} > 0) \\ = (K-1) \int_{-\infty}^w \int_0^{+\infty} \dots \int_0^{+\infty} f(p_{j0}, \dots, p_{j,K-2}) dp_{j,K-2} \dots dp_{j0} \quad (4)$$

where $f(p_{j0}, \dots, p_{j,K-2})$ is p.d.f. of

$$N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(K-1) \times 1}, \begin{pmatrix} 1 & \gamma & \gamma & \dots & \gamma \\ \gamma & 2 & 1 & \dots & 1 \\ \gamma & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ \gamma & 1 & \dots & 1 & 2 \end{pmatrix}_{(K-1) \times (K-1)} \right)$$

and $\gamma = \sqrt{\frac{n_1}{2N}} \rho$.

Denote the distribution of final test statistic W under H_0 by F_0 . Let w_α be the upper 100α percent quantile of F_0 ,

$$w_\alpha = F_w^{-1}(1 - \alpha | H_0) = F_0^{-1}(1 - \alpha).$$

The type I error rate of the design can be controlled at level α if the 1-sided rejection region is $\Omega = \{W : W > w_\alpha\}$.

Notice that, controlling type I error for the hypotheses (1) does not control the probability that the winning treatment will be deemed effective when it is in fact not effective. Even when the null hypothesis in (1) is rejected correctly, an error can still occur (ie. The wrong treatment can be selected). Hence, another interesting index for the performance of the biomarker informed two-stage winner design is “power with best treatment”, which is the probability that the null hypothesis will be rejected when the best treatment is selected at interim. Wang et al. [23] studied “power with best treatment”. In general, it is lower than power. However, the difference depends on the trend of the mean level relationship between biomarker and the primary endpoint. In this manuscript, we develop a method for controlling type I error of the design, “power with best treatment” will not be discussed.

Test statistic and critical rejection region using two-level correlation approach

In this section, we derive the approximate distribution and give the critical rejection region for the final test statistic of the biomarker informed two-stage winner design under the two level correlation model.

The test statistic of biomarker informed two-stage winner design comparing the primary endpoint of the j th treatment group and the control group under the two-level correlation model could be expressed as:

$$G_j = \sqrt{\frac{N}{\sigma_y^2 \left[\left(\widehat{u_j^{X^*}} \right)^2 \widehat{\sigma_j^2} + \left(\widehat{u_0^{X^*}} \right)^2 \widehat{\sigma_{r0}^2} + 2 \right]}} \left(\overline{Y_N^{(j)}} - \overline{Y_N^{(0)}} \right).$$

By law of large numbers, G_j is asymptotically normal,

$$G_j \sim N \left(\frac{N}{\sqrt{\left[\left(u_j^{X^*} \right)^2 \sigma_j^2 + \left(u_0^{X^*} \right)^2 \sigma_{r0}^2 + 2 \right]}} \left(u_j^{Y^*} - u_0^{Y^*} \right), 1 \right).$$

Hence under H_0 , G_j is asymptotically standard normal.

The final test statistic of the design could be expressed as:

$$W = G_j, \text{ if } \overline{X_{n_1}^{(j)}} = \max \left(\overline{X_{n_1}^{(1)}}, \dots, \overline{X_{n_1}^{(K-1)}} \right) j = 1, \dots, K-1.$$

and its distribution under the two-level correlation model could then be expressed as:

$$F_w(w) = P(W < w) = \sum_{j=1}^{K-1} P \left(G_j < w, \overline{X_{n_1}^{(j)}} = \max \left(\overline{X_{n_1}^{(1)}}, \dots, \overline{X_{n_1}^{(K-1)}} \right) \right) \\ = \sum_{j=1}^{K-1} P \left(G_j < w, \overline{X_{n_1}^{(j)}} - \overline{X_{n_1}^{(m_1)}} > 0, \dots, \overline{X_{n_1}^{(j)}} - \overline{X_{n_1}^{(m_{K-2})}} > 0 \right) \\ = \sum P \left(P_{j0} < a_{j0}, P_{j1} > a_{j1}, \dots, P_{j,K-2} > a_{j,K-2} \right) \quad (5)$$

where $P_{j0} = G_j - \sqrt{\frac{N}{\left[\left(u_j^{X^*} \right)^2 \sigma_j^2 + \left(u_0^{X^*} \right)^2 \sigma_{r0}^2 + 2 \right]}} \left(u_j^{Y^*} - u_0^{Y^*} \right)$

$$P_{j1} = \frac{\overline{X_{n_1}^{(j)}} - \overline{X_{n_1}^{(m_1)}}}{\sqrt{\frac{\sigma_X^2}{n_1}}} - \left(\sqrt{n_1} u_j^{X^*} - \sqrt{n_1} u_{m_1}^{X^*} \right)$$

$$P_{j,K-2} = \frac{\overline{X_{n_1}^{(j)}} - \overline{X_{n_1}^{(m_{K-2})}}}{\sqrt{\frac{\sigma_X^2}{n_1}}} - \left(\sqrt{n_1} u_j^{X^*} - \sqrt{n_1} u_{m_{K-2}}^{X^*} \right)$$

$$a_{j0} = w - \sqrt{\frac{N}{\left[\left(u_j^{X^*} \right)^2 \sigma_j^2 + \left(u_0^{X^*} \right)^2 \sigma_{r0}^2 + 2 \right]}} \left(u_j^{Y^*} - u_0^{Y^*} \right)$$

$$a_{j1} = - \left(\sqrt{n_1} u_j^{X^*} - \sqrt{n_1} u_{m_1}^{X^*} \right)$$

$$a_{j,K-2} = - \left(\sqrt{n_1} u_j^{X^*} - \sqrt{n_1} u_{m_{K-2}}^{X^*} \right)$$

$$j \neq m_1 \neq \dots \neq m_{K-2}, j, m_1, \dots, m_{K-2} \in \{1, 2, \dots, K-1\}.$$

$(P_{j0}, P_{j1}, \dots, P_{j,K-2})$ is approximately from a multivariate normal distribution.

$$\begin{pmatrix} P_{j0} \\ P_{j1} \\ P_{j2} \\ \vdots \\ P_{j,K-2} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma_j & \gamma_j & \dots & \gamma_j \\ \gamma_j & 2 & 1 & \dots & 1 \\ \gamma_j & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ \gamma_j & 1 & \dots & 1 & 2 \end{pmatrix} \right).$$

Hence,

$$F_w(w) = \sum_{j=1}^{K-1} \int_{-\infty}^w \int_{\left[\frac{\left(\widehat{u_j^{X^*}} \right)^2 \widehat{\sigma_j^2} + \left(\widehat{u_0^{X^*}} \right)^2 \widehat{\sigma_{r0}^2} + 2 \right]^{-1/2} \left(u_j^{Y^*} - u_0^{Y^*} \right)}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(p_{j0}, \dots, p_{j,K-2}) dp_{j,K-2} \dots dp_{j0} \quad (6)$$

where:

$$j \neq m_1 \neq \dots \neq m_{K-2}, j, m_1, \dots, m_{K-2} \in \{1, 2, \dots, K-1\};$$

$$f_j(p_{j0}, \dots, p_{j,K-2}) \text{ is p.d.f. of } N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma_j & \gamma_j & \dots & \gamma_j \\ \gamma_j & 2 & 1 & \dots & 1 \\ \gamma_j & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ \gamma_j & 1 & \dots & 1 & 2 \end{pmatrix} \right),$$

$$\gamma_j = \rho \sqrt{\frac{n_1}{N}} \sqrt{\frac{(\widehat{u}_j^{X^*})^2 \widehat{\sigma}_{rj}^2 + 1}{(\widehat{u}_j^{X^*})^2 \widehat{\sigma}_{rj}^2 + (\widehat{u}_0^{X^*})^2 \widehat{\sigma}_{r0}^2 + 2}}$$

Under H_0 , (6) could be simplified as:

$$F_w(w) = \sum_{j=1}^{K-1} P(P_{j0} < w, P_{j1} > 0, \dots, P_{j,K-2} > 0)$$

$$= \sum_{j=1}^{K-1} \int_{-\infty}^w \int_0^{+\infty} \dots \int_0^{+\infty} f_j(p_{j0}, \dots, p_{j,K-2}) dp_{j,K-2} \dots dp_{j0} \quad (7)$$

where $f_j(p_{j0}, \dots, p_{j,K-2})$ is p.d.f. of

$$N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(K-1) \times 1}, \begin{pmatrix} 1 & \gamma_j & \gamma_j & \dots & \gamma_j \\ \gamma_j & 2 & 1 & \dots & 1 \\ \gamma_j & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ \gamma_j & 1 & \dots & 1 & 2 \end{pmatrix}_{(K-1) \times (K-1)} \right)$$

$$\text{and } \gamma_j = \rho \sqrt{\frac{n_1}{N}} \sqrt{\frac{(\widehat{u}_0^{X^*})^2 \widehat{\sigma}_{rj}^2 + 1}{(\widehat{u}_0^{X^*})^2 \widehat{\sigma}_{rj}^2 + (\widehat{u}_0^{X^*})^2 \widehat{\sigma}_{r0}^2 + 2}}$$

Notice that, if $\sigma_{r0}^2 = \sigma_{r1}^2 = \dots = \sigma_{r,K-1}^2$, that is the variability of the estimated mean-level correlation for each treatment group is the same, the above distribution can be approximated by:

$$F_w(w) = \sum_{j=1}^{K-1} P(P_{j0} < w, P_{j1} > 0, \dots, P_{j,K-2} > 0)$$

$$= (K-1) \int_{-\infty}^w \int_0^{+\infty} \dots \int_0^{+\infty} f(p_{j0}, \dots, p_{j,K-2}) dp_{j,K-2} \dots dp_{j0}$$

where $f(p_{j0}, \dots, p_{j,K-2})$ is p.d.f. of

$$N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(K-1) \times 1}, \begin{pmatrix} 1 & \gamma & \gamma & \dots & \gamma \\ \gamma & 2 & 1 & \dots & 1 \\ \gamma & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ \gamma & 1 & \dots & 1 & 2 \end{pmatrix}_{(K-1) \times (K-1)} \right)$$

and $\gamma = \sqrt{\frac{n_1}{2N}} \rho$.

The type I error rate of the design can be controlled at level α if the 1-sided rejection region is $\Omega = \{W : W > w_\alpha\}$, where, $w_\alpha = F_w^{-1}(1 - \alpha | H_0) = F_0^{-1}(1 - \alpha)$.

R- functions to compute the critical value

We developed R functions for calculating the critical rejection values w_α for biomarker informed two-stage winner design with any active treatment arms K ($K \geq 3$) (Please refer to the appendix).

R function-*convention_cv_Kk* is for calculating the integration of $F_0(w)$ for biomarker informed two-stage winner design with $k-1$ active

treatment groups and a control group under conventional model and therefore can be used to find critical rejection values for the final test statistic. Notice that *convention_cv_Kk* is a function of α , K , $I = \frac{n_1}{N}$

and ρ . As it seems unlikely that in practice ρ will be known, the sample correlation coefficient $\hat{\rho}$ is suggested to be used for calculating an approximate critical rejection value. The question that how $\hat{\rho}$ affects the distribution of test statistic will be discussed in next section.

R function-*wang_cv_Kk* is for calculating values for biomarker informed two-stage winner design with $k-1$ active treatment groups and a control group under the two-level correlation model proposed by Wang et al. [23]. If there are unknown parameters incorporated in the functions, the parameter estimates can be used for calculating an approximate critical rejection value. As an additional check, simulations should be done to ensure type I error rate is preserved.

Critical values

Tables 1-5 provide the critical rejection values $w_{0.025}$ for biomarker informed two-stage winner design with up to 7 active treatment group under conventional model. As expected, the critical rejection value $w_{0.025}$ increases as ρ increases. It can also be seen that, critical value is a function of the information at the interim i.e. with more information at interim, the critical rejection value $w_{0.025}$ at the final analysis will be larger. Also the more active treatment groups, the larger the critical rejection value $w_{0.025}$ will be.

These tables also reflect partially how estimation of ρ affects the

ρ	0	0.2	0.5	0.8	1
$n_1/N=0.3$	1.96	2.023	2.108	2.182	2.225
$n_1/N=0.5$	1.96	2.041	2.146	2.232	2.279
$n_1/N=0.8$	1.96	2.061	2.186	2.281	2.328

Table 1: $w_{0.025}$ for biomarker informed two-stage winner design ($K=4$).

ρ	0	0.2	0.5	0.8	1
$n_1/N=0.3$	1.96	2.038	2.142	2.232	2.286
$n_1/N=0.5$	1.96	2.058	2.188	2.294	2.352
$n_1/N=0.8$	1.96	2.083	2.237	2.356	2.415

Table 2: $w_{0.025}$ for biomarker informed two-stage winner design ($K=5$).

ρ	0	0.2	0.5	0.8	1
$n_1/N=0.3$	1.96	2.048	2.165	2.27	2.33
$n_1/N=0.5$	1.96	2.072	2.218	2.34	2.408
$n_1/N=0.8$	1.96	2.1	2.275	2.411	2.48

Table 3: $w_{0.025}$ for biomarker informed two-stage winner design ($K=6$).

ρ	0	0.2	0.5	0.8	1
$n_1/N=0.3$	1.96	2.055	2.185	2.299	2.366
$n_1/N=0.5$	1.96	2.082	2.242	2.376	2.452
$n_1/N=0.8$	1.96	2.112	2.305	2.455	2.531

Table 4: $w_{0.025}$ for biomarker informed two-stage winner design ($K=7$).

ρ	0	0.2	0.5	0.8	1
$n_1/N=0.3$	1.96	2.063	2.202	2.323	2.396
$n_1/N=0.5$	1.96	2.091	2.262	2.407	2.488
$n_1/N=0.8$	1.96	2.123	2.33	2.492	2.575

Table 5: $w_{0.025}$ for biomarker informed two-stage winner design ($K=8$).

$\hat{\rho} (\rho)$	0 (0.8)	0.2 (0.8)	0.5 (0.8)	0.7 (0.8)	1 (0.8)
$n_1/N=0.3$					
$\widehat{W}_{0.025} (W_{0.025})$	1.96 (2.182)	2.023 (2.182)	2.108 (2.182)	2.159 (2.182)	2.225 (2.182)
Error ($\alpha=0.025$)	0.016	0.011	0.005	0.001	-0.002
$n_1/N = 0.5$					
$\widehat{W}_{0.025} (W_{0.025})$	1.96 (2.232)	2.041 (2.232)	2.146 (2.232)	2.205 (2.232)	2.279 (2.232)
Error ($\alpha=0.025$)	0.021	0.014	0.006	0.002	-0.003
$n_1/N = 0.8$					
$\widehat{W}_{0.025} (W_{0.025})$	1.96 (2.281)	2.061 (2.281)	2.186 (2.281)	2.253 (2.281)	2.328 (2.281)
Error ($\alpha=0.025$)	0.027	0.017	0.006	0.002	-0.003

Table 6: Error in Type I Error using $\hat{\rho}$ ($K=9$).

distribution of the test statistic. Table 6 lists the errors in Type I error using $\hat{\rho}$ instead of ρ for a biomarker informed two-stage winner design with 3 active treatment groups and 1 control group. If $\hat{\rho} = 0.5$ is used instead of $\rho=0.8$, the true Type I error of the design is around 0.03 when we thought the type I error rate is controlled at 0.025. If $\hat{\rho} = 1$ is used, the true Type I error of the design is around 0.023. Therefore, errors caused by misestimating ρ is in general in an acceptable region.

Discussion

In this manuscript, we have proposed a novel statistical approach for type I error control of the biomarker informed two-stage winner design. We leverage the concept of adjusting critical rejection values of the final test statistic of the design for preserving the type I error rate. The exact distribution of the final test statistic is derived under the conventional one-level correlation model, and the asymptotic distribution of the final test statistic is provided for Wang et al. [23] two-level correlation model. The critical rejection values w_a are computed through mathematical integrations. We developed R functions for calculating the adjusted critical rejection values from the skewed distribution of final test statistic. As shown, the critical rejection value $w_{0.025}$ increases if any of the following increases i.e. correlation (ρ), number of active treatment groups (k) and information at interim analysis $\left(\frac{n_1}{N}\right)$ increases.

Our proposed method circumvents the limitation of the normal approximation method proposed by Shun et al. [22], and works for designs with any number of treatment arms. However, it has the limitation that it works only for the biomarker informed two-stage winner design with normal interim and final endpoints. For the designs with non-normal endpoints, transformations might be used to convert the data to follow normal distribution. Developing novel approaches for type I error control for biomarker informed two-stage winner design with non-normal endpoints would be an interesting topic for future studies.

References

- Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, et al. (2006) Adaptive designs in clinical drug development—an Executive Summary of the PhRMA Working Group. J Biopharm Stat 16: 275-283.
- Pocock SJ (1977) Group sequential methods in the design and analysis of clinical trials. Biometrika 64: 191-199.
- O'Brien PC, Fleming TR (1979) A multiple testing procedure for clinical trials. Biometrics 35: 549-556.
- Lan KKG, DeMets DL (1983) Discrete sequential boundaries for clinical trials. Biometrika 70: 659-663.
- Bauer P (1989) Multistage testing with adaptive designs. Biometrie und Informatik in Medizin und Biologie 20: 130-148.
- Bauer P, Köhne K (1994) Evaluation of experiments with adaptive interim analyses. Biometrics 50: 1029-1041.
- Lehmacher W, Wassmer G (1999) Adaptive sample size calculations in group sequential trials. Biometrics 55: 1286-1290.
- Chang M (2008) Adaptive design theory and implementation using SAS and R. CRC Press.
- Proschan MA, Hunsberger SA (1995) Designed extension of studies based on conditional power. Biometrics 51: 1315-1324.
- Liu Q, Chi GY (2001) On sample size and inference for two-stage adaptive designs. Biometrics 57: 172-177.
- Müller HH, Schäfer H (2001) Adaptive group sequential designs for clinical trials: combining the advantages of adaptive and of classical group sequential approaches. Biometrics 57: 886-891.
- Cui L, Hung HM, Wang SJ (1999) Modification of sample size in group sequential clinical trials. Biometrics 55: 853-857.
- Fisher LD (1998) Self-designing clinical trials. Stat Med 17: 1551-1562.
- Shen Y, Fisher L (1999) Statistical inference for self-designing clinical trials with a one-sided hypothesis. Biometrics 55: 190-197.
- Bretz F, Schmidli H, Koenig F, Racine A, Maurer W (2006) Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: general concepts. Biometrical Journal 48: 623-634.
- Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Statist 6: 65-70.
- Hochberg Y (1988) A sharper Bonferroni's procedure for multiple tests of significance. Biometrika 75: 800-803.
- Hommel G (1988) A stagewise rejective multiple test procedure based on a modified Bonferroni test. Biometrika 75: 383-386.
- Todd S, Stallard N (2005) A new clinical trial design combining phases 2 and 3: sequential designs with treatment selection and a change of endpoint. Drug Information Journal 39: 109-118.
- Stallard N (2010) A confirmatory seamless phase II/III clinical trial design incorporating short-term endpoint information. Stat Med 29: 959-971.
- Friede T, Parsons N, Stallard N, Todd S, Valdes Marquez E, et al. (2011) Designing a seamless phase II/III clinical trial using early outcomes for treatment selection: an application in multiple sclerosis. Stat Med 30: 1528-1540.
- Di Scala L, Glimm E (2011) Time-to-event analysis with treatment arm selection at interim. Stat Med 30: 3067-3081.

23. Shun Z, Lan KK, Soo Y (2008) Interim treatment selection using the normal approximation approach in clinical trials. *Stat Med* 27: 597-618.
24. Sill MW, Sampson AR (2009) Drop-the-Losers Design: Binomial Case. *Comput Stat Data Anal* 53: 586-595.
25. Wang J, Chang M, Menon S (2013) Biomarker Informed Adaptive Design. Chapter 5. *Clinical and Statistical Considerations in Personalized Medicine*. CRC Press.