Short communication

A two-stage design with two correlated co-primary endpoints

Abstract

In phase II oncology trials, multiple binary outcomes are often of interest to evaluate the efficacy of a new treatment. Song[1] proposed an exact method of simultaneous evaluation of two independent co-primary endpoints in a two-stage design. The approach searches stage 1 and final stopping boundaries of both hypotheses based on binomial distribution under type I and II error constrains. Herein, I extend the design to a more common setting where the two co-primary endpoints are correlated.

Keywords: phase II oncology trials, co-primary, two-stage design

James Song

Suite 405, 2115 Linwood Avenue, Fort Lee, NJ 07450, USA

Correspondence: James Song, Suite 405, 2115 Linwood Avenue, Fort Lee, NJ 07450, USA, Tel 1(201)429-9918; Email jamesx.song@yahoo.com

Introduction

Simon two-stage design is often applied to phase II single-arm oncology trials where a binary outcome is of interest.[2] The trial design allows us to explore potential efficacy and safety signals quickly while limiting the number of patients exposed to an inefficacious new treatment. In the well-known Simon two stage design, hypotheses are typically set up as $H : P = P$ vs. $H_A : P = P_A$, where P, P_o and pA are the objective response rates (ORR). An interim futility analysis after stage 1 is implemented to stop the trial early in case of lack of efficacy. A numerical search is performed in identifying potential designs with sample sizes (n_1 and n) and boundaries (r_1 and r) at each stage via binomial distribution when P_o, pA, type I and II errors (α and β) are specified. The final design can be chosen among them based on different criteria.[3]

In oncology trials, besides ORR, other binary outcomes such as disease control rate (DCR, percentage of patients achieving stable disease or better); progression-free survival (PFS) or overall survival (OS) rate at a specific time point (e.g. 3-month PFS and 6-month OS) are often considered in selecting the primary endpoint. In some cases, it is important to include more than one measure of anticancer activities for better evaluation. To address the issue, Song[1] proposed an extension of Simon two-stage design by including a 2nd hypothesis, $H_o : P^` = P_o^`$ vs $H_A : P^` = P_A^`$ such that two binary outcomes can be tested simultaneously under independence assumption. The empirical type I and II errors for stopping boundaries (r_1, r, $r_1^`$, $r^`$) can be calculated via binomial distribution when sample sizes (n_1, n, $n_1^`$, $n^`$) are fixed. Subsequent selection of the stopping boundary is made based on an objective function that minimizes the type II errors (i.e. interim analysis and final type II errors within each endpoint and overall). Quite often, in practice, it is reasonable to assume a correlation between two binary efficacy outcomes. Therefore, it is desirable to extend this method to a setting with two correlated co-primary endpoints.

Method

The bivariate binomial distribution of X and Y proposed by Biswas and Hwang[4] is used. Let two binomial variables $(X, Y) \sim BVB(n, p, p, \tau)$, where numbers of trials in both X and Y are equal to n, the model defines dependence of X and Y via τ whose range is limited by the binomial parameters p and $p^`$. The correlation between $X$ and $Y$ ($\rho$) can be expressed as $\tau / (1+\tau)\sqrt{\{p(1-p)\} / \{p^`(1-p^`)\}}$. Its density function is expressed as

$$\Pr(x, y) = \binom{n}{x} p^x (1-p)^{n-x} f(y \mid x) \tag{1}$$

where

$$f(y \mid x) = (1+\tau)^{-n} \sum_{i+j=y} \binom{x}{i}\binom{n-x}{j} \{p^`+\tau(p^`-p)+\tau\}^i \{1-p^`-\tau(p^`-p)\}^{x-i} \{p^`+\tau(p^`-p)\}^j$$

$$\{1-p^`-\tau(p^`-p)+\tau\}^{n-x-j}$$

with $i = 0, \ldots, x$; $j = 0, \ldots, n - x$.

In the proposed two-stage design, the trial will move into stage 2 if number of successes in either endpoint $(x_1 \text{ or } y_1)$ passes its stage 1 stopping boundary $(r_1 \text{ or } r_1^`)$; the treatment will be deemed promising if at least one hypothesis is rejected when final number of success $(y \text{ or } y)$ crosses final boundary $(r \text{ or } r^`)$. Following these decision rules, the probability of accepting $H_0 s$ is

$$Pr\left(x_1 \le r_1 \cap y_1 \le r_1{}^`\right) \; + \; Pr\left(x \le r \cap y \le r`\,|\,x_1 > r_1 \cup y_1 > r_1{}^`\right) \text{ which}$$

is expressed as

$$\sum_{x_1 \le r_1 \cap y_1 \le r_1{}^`} Pr(x_1, y_1) + \sum_{x \le r \cap y \le r`} \sum_{x_1 > r_1 \cup y_1 > r_1{}^`} Pr(x, y) f(x_1, y_1) \qquad (2)$$

for $(x, y) \in \{x > r_1 \text{ or } y > r_1{}^`\}$. It can be calculated using equation (1) for a given $\tau$. The overall type I error is calculated when $p = p_0$ and $p` = p_0{}^`$; and the overall type II error is calculated when $p = pA$ and $p` = pA`$.

For a selected design, the probability of accepting H_0 of an individual endpoint can be calculated similarly. For example,

$$Pr\left(accept\ H_0 : p = p_0\right) = Pr\left(x_1 \le r_1 \cap y_1 \le r_1{}^`\right) \; + \; Pr\left(x \le r \,|\, x_1 > r_1\ U\ y_1 > r_1{}^`\right) =$$

$$\sum_{x_1 \le r_1 \cap y_1 \le r_1{}^`} Pr(x_1, y_1) + \sum_{x \le r} \sum_{x_1 > r_1 \cup y_1 > r_1{}^`} Pr(x) f(x_1, y_1) \qquad (3)$$

The type II error in H_0 can be obtained by setting $p = pA$.

An example

The proposed method is applied to a planned phase II trial in metastatic breast cancer. ORR and percentage of patients without deterioration in Global Health Status (GHS) of European Organization for Research and Treatment of Cancer Quality of Life Questionnaire – core 30 (EORTC QLC-C30) in the first two cycles of treatment were two key efficacy variables of interest. Hence, two hypotheses are set up as $H_0 : ORR = 5\%$ vs. $H_A : ORR = 15\%$ and $H_0 : GHS = 45\%$ vs. $H_A : GHS = 60\%$. Among 636 sets of boundaries satisfying overall $\alpha \le 0.05$ and $\beta \le 0.2$ assuming independence between ORR and GHS, the stopping boundaries $\left(r_1 = 0,\ r = 6,\ r_1` = 7\ and\ r` = 31\right)$ are chosen when $n_1 = 15$ and $n = 55$.

To show the effect of correlation between ORR and GHS on the type I and II errors, a range of τ is set such that the correlation under alternative hypothesis $\left(\rho_A\right)$ are -0.80, -0.50, -0.25, 0, 0.25, 0.5 and 0.8. Hence, the correlation under null hypothesis $\left(\rho_o\right)$ are -0.48, -0.30, -0.15, 0, 0.15, 0.30 and 0.48. Type I and II errors in the overall trial and within each endpoint are calculated using equations (2) and (3) (Table 1). In general, the type I errors decrease and the type II errors increase as correlation between two endpoints, ρ_0 and ρ_A, increases. The impact is more noticeable in the overall type I and II errors, which is consistent to the findings one would expect in a one-stage test involving multivariate normal variables. A similar, yet much smaller effect is observed in error rates of the individual endpoint; it is due to the fact that the correlation only matters at interim analysis in which only one boundary needs to be crossed in order to move both hypotheses tests to the final analysis. The results also show overall type II error is more sensitive to the level of correlation as compared to overall type I error, which is probably more specific to the design chosen and not to be generalized to others. Similar to the one-stage testing, the impact of correlation is not only influenced by the design characteristics within each test; but also the relative difference of type I or type II errors between the two endpoints.[5] In addition, the percentages of type I or II errors spent at interim analysis are also the factors in assessing correlation effect.

Table 1 The type I and II errors in the trial (overall) and within each endpoint: ORR and GHS

			Overall		ORR		GHS	
τ	ρ0	ρA	α	β	α1	β1	α2	β2
-0.5233	-0.48	-0.8	0.05	0.0124	0.0186	0.2632	0.0315	0.3377
-0.4069	-0.3	-0.5	0.0497	0.0379	0.0185	0.2644	0.0313	0.3389
-0.2554	-0.15	-0.25	0.0494	0.0668	0.0184	0.2666	0.0312	0.3409
0	0	0	0.0488	0.1001	0.0183	0.2701	0.0311	0.344
0.5221	0.15	0.25	0.0478	0.1366	0.0182	0.2749	0.0311	0.3478
2.1847	0.3	0.5	0.0466	0.1757	0.0181	0.2803	0.0311	0.3515
-11.2469	0.48	0.8	0.0445	0.2243	0.0181	0.2867	0.0312	0.3545

Conclusions

This short communication describes an extension of Song[1] to the two-stage design involving two correlated co-primary endpoints. The exact method using bivariate density function proposed by Biswas and Hwang[4] is implemented to calculate various probabilities under Simon two-stage design framework. The purpose of the current work is to explore the impact of correlation on the type I and II errors of the design chosen under independence assumption. First, under independence assumption, admissible designs are identified by exact binomial probability calculation which requires less computing resource. Desirable designs with comparable type I or II errors between two tests can then be selected among all admissible designs. In this selection process, designs with high type I errors at interim can

be screened out; an objective function $S(\beta)$ can also be used to select the designs minimizes the type II errors. Finally, the independence assumption is relaxed in the selected designs from earlier steps, type I and type II errors are recalculated to select the final desirable design. The method provides a useful tool for a more robust assessment of the design operating characteristics, especially when the independence assumption is questionable.

In case of overlapping between two endpoints, i.e., all responders in one endpoint are also the responders in the other, the correlation between two endpoints is defined when the marginal distribution of each variable is specified. For example, when setting $p_A = 15\%$ in ORR and $p_A\dot{} = 60\%$ in DCR, percentage of subjects in each cell of the $2x2$ ORR by DCR table is fixed since all subjects achieve objective response also have disease control (i.e. one is a subset of the other). Hence, correlation estimate such as phi coefficient can be obtained and be subsequently used in equations (2) and (3). Regardless correlation being implicitly specified or not, sensitivity analysis such as the one performed in the example is important to evaluate trial designs.

Acknowledgements

Conflicts of interest

None.

References

1. Song Jx. Sample size for simultaneous testing of rate differences in non-inferiority trials with multiple endpoints. *Computational Statistics and Data Analysis*. 2008;53(4):1201–1207.

2. Jung SH, Lee T, Kim K, et al. Admissible two-stage design for phase II cancer clinical trials. *Statistics in Medicine*. 2004;23(4):561–569.

3. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials*. 1989;10(1):1–10.

4. Biswas A, Hwang JS. A new bivariate binomial distribution. *Statistics & Probability Letter*. 2002;60(2):23–240.

5. Song JX. A two-stage design with two correlated co-primary endpoints. *Contemporary Clinical Trials Communications*. 2015;1:2–4.