

Clinical Trials

<http://ctj.sagepub.com>

Balanced two-stage designs for phase II clinical trials

Fei Ye and Yu Shyr

Clin Trials 2007; 4; 514

DOI: 10.1177/1740774507084102

The online version of this article can be found at:
<http://ctj.sagepub.com/cgi/content/abstract/4/5/514>

Published by:



<http://www.sagepublications.com>

On behalf of:



The Society for Clinical Trials

Additional services and information for *Clinical Trials* can be found at:

Email Alerts: <http://ctj.sagepub.com/cgi/alerts>

Subscriptions: <http://ctj.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.co.uk/journalsPermissions.nav>

Citations <http://ctj.sagepub.com/cgi/content/refs/4/5/514>

Balanced two-stage designs for phase II clinical trials

Fei Ye and Yu Shyr

Background Oncology phase II clinical trials are often designed using Simon's two-stage designs (optimal and 'minimax') for independent observations. Simon's designs do not include the use of correlated observations, and do not balance the sample sizes of the two stages. In these designs, the sample sizes of the two stages can be highly unequal. In certain circumstances, an alternative design option that balances the sample sizes is desirable.

Purpose To develop a two-stage phase II design that balances the sample sizes of the first and the second stages, while controlling for type I and type II error rates.

Methods We simulated designs based on response rates under various null and alternative hypotheses, type I and type II error constraints, and the degree of correlation in the case of correlated data. For correlated data, Sargent's method is adopted to account for the loss of information due to intra-person correlation.

Results Design characteristics for different parameter settings were generated using balanced design method, separately for independent and correlated data. Results were evaluated and compared to the optimal and minimax design.

Limitations For correlated data, designs were produced only for trials with half of the participants having one observation and the other half having two. Also, the degree of intra-person correlation was fixed at three levels.

Conclusion The balanced design provides an additional choice for two-stage phase II trials when the investigators would like to monitor the trial near a study's halfway point. Meanwhile, its total sample sizes are comparable with Simon's designs. *Clinical Trials* 2007; 4: 514–524. <http://ctj.sagepub.com>

Introduction

Statistical research on designs of phase II oncology clinical trials is focused on: (1) improvement on the estimation of treatment response rate and detection of the adverse effect of that treatment, and (2) determination of sample sizes. All else being equal, statistical power increases with increasing number of participants enrolled in the trial. However, the effort to increase power must be balanced with the greater costs associated with larger trials. Different approaches have been applied to phase II clinical trials: frequentist, Bayesian, and decision-theoretic [1]. In this article, we focus on those frequentist approaches that optimize the sample size with certain restrictions.

Currently, the most popular two-stage phase II oncology clinical trial designs with independent observations are Simon's optimal and minimax two-stage design [2]. Controlling for type I and type II error rates, the optimal design minimizes the expected sample size if the investigational treatment has a low response rate; the 'minimax' design minimizes the maximum sample size (i.e., the total sample size of the two stages). However, neither method balances the sample sizes of the two stages. Clinical investigators may not want to stop the trial based on a very small percentage of the target accrual, as the early-accrued participants in the first stage may not be sufficiently representative of the target population, or information might be lost on important secondary outcomes or correlational measures. Conversely, when the ratio of the

Division of Cancer Biostatistics, Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, TN, USA,
Author for correspondence: Yu Shyr, PhD, Division of Cancer Biostatistics Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, TN, USA

sample size of the first stage to the second stage ($n_1:n_2$) is significantly >1 , the potential savings in sample size from the two-stage design is minimal.

We propose an alternative method of two-stage phase II clinical trials design: balanced two-stage design. The balanced design seeks to equalize the sample size of the two stages, while maintaining total sample sizes that are comparable with Simon's designs. It provides an additional choice for two-stage phase II designs when investigators would like to monitor the trial near the halfway point of a study.

Some studies may have participants with multiple observations on the endpoint of interest. In that case, the assumption of independence between observations is violated and the standard two-stage phase II designs, such as the optimal design, cannot be directly applied. In this article, Sargent, Sloan and Cha's method [3] is incorporated into the balanced design for trials with correlated observations.

Two-stage phase II clinical trial designs with independent observations

Hypothesis testing and error rates

Two-stage phase II trials can be regarded as simple forms of a sequential phase II design, which is most often employed in oncology [4–6]. A typical two-stage phase II trial can be described as follows: at stage I, n_1 participants are treated. If the number of responses is not larger than the lower bound r_1 , the trial will be terminated due to lack of efficacy. Otherwise, an additional number (n_2) of participants will be enrolled in the second stage. If the total number of responses is not larger than the lower bound r by the end of the second stage, we conclude that the treatment does not have the desired effect [7–9].

Simon's optimal and minimax two-stage designs will be the comparators in this article. Among all feasible designs satisfying the type I and type II error constraints, the optimal design minimizes the expected sample size and the minimax design minimizes the maximum sample size.

Suppose the current standard treatment has the response rate 75% and the investigator wants to test whether a new treatment will improve this response rate by 20% ($p_0=0.75$, $p_1=0.95$), with 5% type I and 20% type II error constraints. For the optimal design, 3 participants are needed in stage I and 19 additional participants in stage II. The ratio of sample sizes of the first stage and the second stage ($n_1:n_2$) is 0.16. For the minimax design, when $p_0=0.45$ and $p_1=0.60$, with 5% type I and 10% type II error constraints, 93 participants need to be accrued in the first stage but only 2 participants in the second stage. The ratio of ($n_1:n_2$) is 46.5. The design settings and characteristics are shown in Table 1. In such cases, a design method seeking to balance the sample size between the two stages is desirable.

Balanced design for trials with independent observations

Four parameters, p_0 , p_1 , α , and β , are required for conducting a balanced design for trials with independent observations: p_0 is the control or standard response rate; p_1 is the desired response rate; type I error constraint (α) is the upper bound of the probability of falsely concluding that the new treatment has a higher response rate than the standard treatment when the response rates are equal; and type II error constraint (β) is the upper bound of the probability of falsely concluding that the response rate on the new treatment is equal to the response standard treatment, when it is in fact higher.

Table 1 Examples of optimal and minimax designs with unbalanced sample sizes

p_0	p_1	α	β	n_1^a	n_2^b	Ratio (n_1, n_2)
Optimal design						
0.75	0.95	0.05	0.20	3	19	0.16
0.11	0.31	0.1	0.2	5	20	0.25
0.70	0.90	0.05	0.20	6	21	0.29
Minimax design						
0.45	0.60	0.05	0.10	93	2	46.5
0.50	0.65	0.05	0.20	66	2	33.0
0.35	0.50	0.10	0.20	19	68	0.28

^aThe sample size of stage I.

^bThe sample size of stage II.

Next, let r_1 and r denote the lower bounds for the number of responses observed in stage I and the total number of responses observed at the end of stage II, respectively. Let n_1 and n_2 denote the sample size of stage I and stage II, respectively. Let n denote the maximum/total sample size ($n = n_1 + n_2$). The ratio of sample sizes between the first stage and second stage is $(n_1 : n_2)$.

Also, let $b(x; p, m)$ and $B(x; p, m)$ be the probability mass function and the cumulative distribution function for the binomial distribution with probability of success p and number of trials m . The probability of early termination (PET) is the probability of terminating the trial after the first stage when $p = p_0$:

$$PET = B(r_1; p, n_1) = \sum_{x=0}^{r_1} \binom{n_1}{x} p^x (1-p)^{n_1-x}$$

The expected sample size (EN) is estimated as:

$$EN = n_1 + (1 - PET)n_2.$$

Given parameters defined above, the algorithm for defining a balanced design can be described as the following: the method first searches for all feasible designs over n in the range of one (or the lower bound of the total sample size) and the upper bound (e.g., 120) that satisfy the type I and type II error constraints.

To minimize the increase in sample size over those of other designs while seeking better balance between the two stages, a feasible design will be defined as a candidate balanced design if it satisfies at least one of the following two conditions: (1) its expected sample size is not larger than that of minimax design, or (2) its maximum sample size is not larger than that of the optimal design.

Among these candidate designs, we then search for the design(s) with the ratio closest to one. If more than one design has the same ratio, the one with the smallest expected sample size will be defined as the balanced design.

Two-stage phase II clinical trial designs with correlated observations

All the phase II design methods discussed above are suited for trials having only independent observations. However, in many studies, an individual participant may have more than one observation on the endpoint of interest. For example, when a clinical trial is conducted to test a regimen for

sudden hearing loss, participants may have hearing loss in each ear measured separately, although the two readings would be related. Examples of correlated data also occur in otologic, ophthalmologic, and longitudinal studies [3,8]. In such cases, the assumption of independence is violated and therefore the standard phase II designs are no longer appropriate.

If they are treated as independent observations, the correlated observations cause an increase in type I error rate and a decrease in power; consequently, they cause an underestimation in required sample sizes. One solution proposed by Sargent *et al.* [3] is employed here to find two-stage designs for correlated data.

Balanced design for trials with correlated observations

Sargent *et al.* [3] proposed a method of determining sample size and design characteristics for phase II clinical trials with correlated observations. The algorithm can be described as follows:

- (1) Assuming single-observation per participant, a standard phase II design such as Fleming's single-stage design is used to calculate the required sample size N_T and the decision boundary S :

$$N_{PT} = \left(\frac{[Z_{1-\beta}\{p_1(1-p_1)\}^{1/2} + Z_{1-\alpha}\{p_0(1-p_0)\}^{1/2}]^2}{(p_1 - p_0)} \right)^2,$$

and

$$S = [N_{PT}P_0 + Z_{1-\alpha}(N_{PT} \times P_0(1 - P_0))^{1/2}]^* + 1$$

where x^* denotes the nearest integer to x .

- (2) For the correlated data, i.e., some participants have multiple observations, the actual type I and type II error rates can be calculated as follows:

$$\alpha_{\text{actual}} = 1 - \Phi\left(\frac{S + 1/2 - N_T p_0}{\sqrt{V}}\right)$$

$$\beta_{\text{actual}} = \Phi\left(\frac{S + 1/2 - N_T p_1}{\sqrt{V}}\right),$$

where

$$V = \text{Var}(X) = \sum_{i=1}^k iN_i p(1-p)(1+(i-1)\rho),$$

X is the number of responses, N_i is the number of participants with i observations, k is the maximum number of observations per participant, and ρ is the degree of correlation.

- (3) The type I and type II error rates from step 2 are inflated due to the intra-person correlation. The sample size N_T , i.e., number of participants, and the decision boundary S need to be increased to N_A and S_A so that the design meets the error constraints α and β . The N_A and S_A in this step refer to the number of participants who have one or multiple observations.
- (4) For per-observation analysis, the reduced type I and type II error rates are calculated based on the equations in step 2 with N_A , S_A , and $\rho=0$. The reduced error rates α_R and β_R are then used as parameter inputs (i.e., adjusted error constraints) in a standard two-stage design to ensure the desired error constraints α and β are approximately achieved.

For our study, balanced two-stage designs for trials with correlated observations were conducted using p_0 , p_1 , α_R , and β_R as parameter inputs. These designs optimally balance the sample sizes of the two stages, satisfy type I and type II error constraints, and adjust for the intra-person correlation effect.

Results

Table 2 describes two typical situations where a balanced design is preferable compared to

the optimal and the minimax design:

For case 1, the standard response rate is 0.63 and the desired response rate is 0.83. The type I error constraint is 0.05 and the type II error constraint is 0.20. After searching through a total number of 19,377 feasible designs satisfying the error rate constraints, i.e., $\alpha < 0.05$, $\beta < 0.2$ (with N in the range of 1 and the predetermined upper bound 120), 157 candidate designs were identified. These candidate designs satisfy the two conditions in defining a balanced design, i.e., $EN_i < EN_{\text{minimax}}$ or $N_i < N_{\text{optimal}}$, $i = 1, \dots, 157$. The ratio of the sample sizes ($n_1:n_2$) was calculated for all these candidate designs, two of which have the ratio one. The design with the smaller expected sample size ($EN_{\text{Balanced}} = 23.2$) was then defined as the balanced design.

Case 2 is an example of trials which have correlated endpoints: half of the participants has one observation, and the other half has two observations. The standard response rate and the desired response rate are 0.4 and 0.6, respectively. The type I error constraint is 0.05 and the type II error constraint is 0.1. Assuming a moderate intra-person correlation level ($\rho = 0.5$), the procedure for finding the balanced design can be described in the following steps:

- (1) Sargent's two-stage design method was used to calculate the reduced type I and type II error rates: $\alpha_R = 0.024$ and $\beta_R = 0.066$.
- (2) There are 19 867 feasible designs that satisfy the error constraints. A total number of 2957 candidate designs that satisfy the two conditions in defining a balanced design, i.e., $EN_i < EN_{\text{minimax}}$ or $N_i < N_{\text{optimal}}$, $i = 1, \dots, 2957$, 17 of which have the ratio one. Among these 17 designs, the balanced design was found to be the one with the smallest expected sample size ($EN_{\text{Balanced}} = 50.3$).

Table 2 Examples of optimal design, minimax design and balanced design

	n_1	n_2	n^a	EN ^b	PET*	Ratio ($n_1:n_2$)
Case 1. $p_0 = 0.63$, $p_1 = 0.83$, $\alpha = 0.05$, $\beta = 0.2$						
Optimal	8	30	38	19.4	0.62	0.27
Minimax	28	3	31	28.2	0.94	9.33
Balanced	18	18	36	23.2	0.71	1
Case 2. $p_0 = 0.4$, $p_1 = 0.6$, $\alpha = 0.05$, $\beta = 0.1$, $\rho = 0.5$						
Optimal	32	57	89	47.3	0.73	0.56
Minimax	75	1	76	75	0.98	75
Balanced	42	42	84	50.3	0.8	1

^aThe total sample size.

^bThe expected sample size. $EN = n_1 + (1 - \text{PET})n_2$.

*The probability of early termination. $\text{PET} = B(r_1; p, n_1) = \sum_{x=0}^{r_1} \binom{n_1}{x} p^x (1-p)^{n_1-x}$.

In both cases, the optimal design has a ratio significantly lower than one and the minimax design has a ratio significantly greater than one. However, the balanced design has the same sample size for the two stages, an expected sample size smaller than the minimax design, and a maximum sample size smaller than the optimal design. In these situations, the balanced design is recommended because the increase in the sample sizes is minor relative to the achieved balance between the two stages.

A total number of 588 design scenarios were considered for trials with independent observations. For testing a 15% improvement, 304 designs were generated with p_0 in the range of 0.05 and 0.80, for a step increment of 0.01. For testing a 20% improvement, 284 designs were generated with p_0 in the range of 0.05 and 0.75, for a step increment of 0.01. For each pair of p_0 and p_1 values, type I and type II error constraints used are: $(\alpha = 0.05, \beta = 0.1)$, $(\alpha = 0.05, \beta = 0.2)$, $(\alpha = 0.1, \beta = 0.1)$, and $(\alpha = 0.1, \beta = 0.2)$.

For trials with correlated observations, a total number of 153 design scenarios were considered. Degrees of intra-person correlation used are 0.2, 0.5, and 0.8. For each pair of p_0 and p_1 values, type I and type II error constraints used are: $(\alpha = 0.05, \beta = 0.1)$, $(\alpha = 0.05, \beta = 0.2)$, and $(\alpha = 0.1, \beta = 0.1)$. For testing a 15% improvement, 81 designs were generated with p_0 at 0.05, and throughout the range of 0.1 and 0.8 with a step increment of 0.1. For testing a 20% improvement, 72 designs were generated with p_0 at 0.05, and throughout the range of 0.1 and 0.7 with a step increment of 0.1.

Figure 1(a) and 1(b) show the density plots of ratio ($n_1:n_2$) for optimal designs with independent data and correlated data, respectively. The ratio ranges from 0.28 to 1.25 and the majority is below one. Figure 2(a) and 2(b) are the density plots for minimax designs with independent data and correlated data, respectively. The range of ratios extends from 0.28 to 148. More than 90% of the ratios ($n_1:n_2$) are > 1 for independent data, and 66% for correlated data. The density curves are skewed to the right with a long tail, truncated at the ratio of 20 for ease of interpretation (3% of the minimax designs have a ratio > 20). Figure 3(a) and 3(b) are the density plots for balanced design with independent data and correlated data, respectively. The ratio ranges from 0.29 to 1.11 for independent data and from 0.94 to 1.05 for correlated data.

In addition, balanced, optimal, and minimax designs were simulated to test 15 and 20% improvement on the response rate, for both independent and correlated data. Design parameters and characteristics are presented in Table 3–6. In each table, the first, second, and third rows are

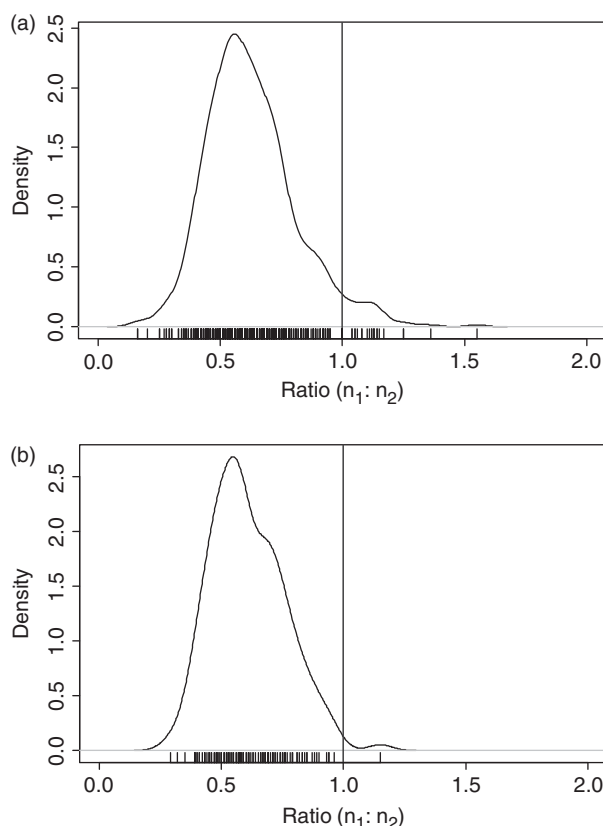


Figure 1 (a) Density of ratio for optimal design – independent data (b) Density of ratio for optimal design – correlated data

corresponding to the type I and type II error constraints (0.1, 0.1), (0.05, 0.2), and (0.05, 0.1), respectively. An asterisk sign (*) indicates that the corresponding balanced design has both expected sample size and maximum sample size between the optimal design and the minimax design. Otherwise, the design either has an expected sample size smaller than or equal to the minimax design, or has a maximum sample size smaller than or equal to the optimal design. Tables 3 and 4 present designs for testing 15 and 20% improvement on the response rate, respectively. Tables 5 and 6 present designs when half of the participants has one observation and the other half has two observations, for different levels of intra-person correlation (0.2, 0.5, and 0.8). Designs not shown on these tables may be found at <http://www.vicc.org/biostatistics/ts/freqapp.php>.

Discussion

As expected, larger sample sizes are required to detect smaller differences between p_0 and p_1 .

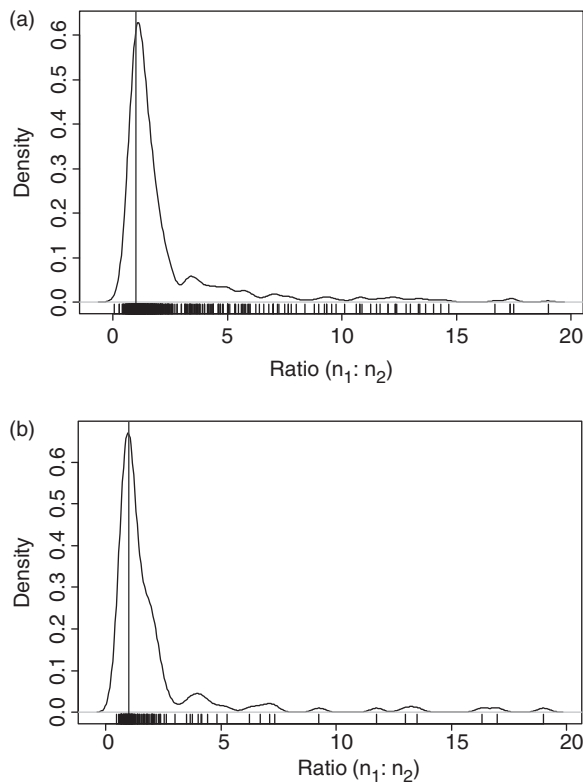


Figure 2 (a) Density of ratio for 'minimax' design – independent data (b) Density of ratio for 'minimax' design – correlated data

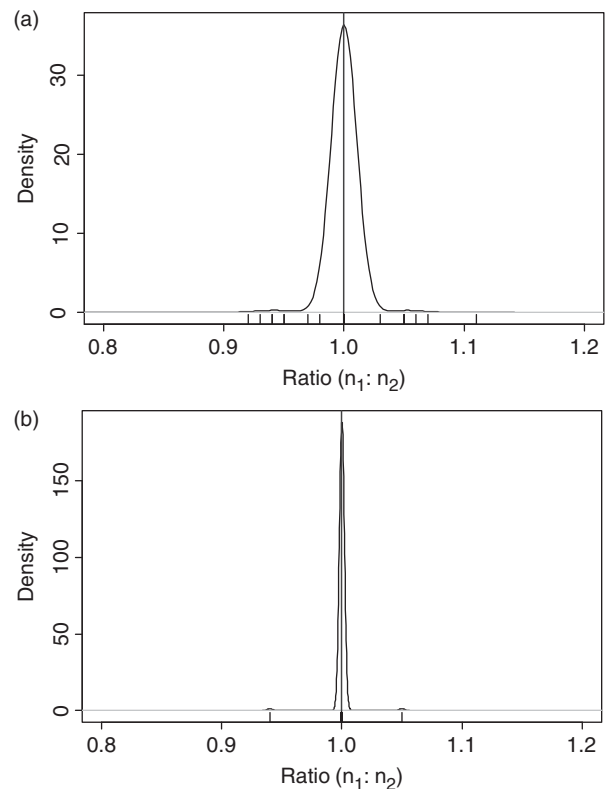


Figure 3 (a) Density of ratio for balanced design – independent data (b). Density of ratio for balanced design – correlated data

With type I and type II error constraints being equal, both the expected sample size and the maximum sample size tend to increase as p_0 and p_1 get closer to 0.5 since the number of responses follows a binomial distribution. Also, due to intra-patient correlation, designs for correlated data have larger sample sizes relative to independent data, with other parameters being equal. The sample sizes increase as the degree of correlation increases.

The purpose of this work is to provide additional choices to Simon's designs when the investigators prefer more balanced sample sizes between the two stages. Each of these three design methods, – optimal, minimax, and balanced – has advantages and disadvantages depending on the details of the study: if the optimal design and the minimax design have dramatically unbalanced sample sizes, the balanced design is more preferable. Otherwise, the optimal design is more attractive since it minimizes the sample size when the response rate is lower than desired. But if the participant accrual rate is low, as is the case for some rare diseases, minimax designs should be considered to minimize the duration of the trial.

Limitations and future research

For the two-stage phase II clinical trial designs with correlated data, only designs have half of the participants having one observation and the other half having two observations were produced. Future research might include designs with different percentages of participants having different numbers of observations.

Another possible direction for future research is the estimation of the degree of intra-person correlation. In these simulations, the degree of correlation is estimated roughly to be 0.2, 0.5, or 0.8, but the degree of correlation in particular clinical situations may differ, and this correlation is often not estimated or reported even when data are available.

Conclusion

The balanced design method provides an alternative choice for two-stage phase II clinical

Table 3 Optimal, minimax, and balanced designs for independent data, $p_1 - p_0 = 0.15$

p_0	p_1	Optimal design				Minimax design				Balanced design						
		Reject drug if response rate $\leq p^*/n$		EN (p_0)		Reject drug if response rate $\leq r_1/n_1$		EN (p_0)		Reject drug if response rate $\leq r_1/n_1$		EN (p_0)				
		$\leq r_1^a/n_1$	$\leq p^*/n$	PET (p_0)	$n_1 : n_2$	$\leq r_1/n_1$	$\leq p^*/n$	PET (p_0)	$n_1 : n_2$	$\leq r_1/n_1$	$\leq p^*/n$	PET (p_0)	$n_1 : n_2$			
0.05	0.20	0/12	3/37	23.5	0.54	0.48	0/18	3/32	26.4	0.4	1.29	1/19	3/38	23.7	0.75	1
		0/10	3/29	17.6	0.6	0.53	0/13	3/27	19.8	0.51	0.93	0/14	3/28	21.2	0.49	1
0.10	0.25	1/21	4/41	26.7	0.72	1.05	1/29	4/38	32.9	0.57	3.22	1/21	4/42	26.9	0.72	1
		2/18	7/43	24.7	0.73	0.72	2/22	7/40	28.8	0.62	1.22	2/21	7/42	28.4	0.65	1*
0.20	0.35	2/21	10/66	36.8	0.65	0.47	3/31	9/55	40.0	0.62	1.29	4/32	10/64	38.8	0.79	1*
		5/22	16/63	43.6	0.54	0.75	6/33	15/58	45.5	0.5	1.32	6/32	16/62	45.9	0.54	1.07
0.3	0.45	8/37	22/83	51.4	0.69	0.44	6/31	15/53	40.4	0.57	1.41	7/31	17/62	39.4	0.73	1*
		9/30	29/82	51.4	0.59	0.58	8/42	21/77	58.4	0.53	1.2	8/39	21/78	53.7	0.62	1*
0.4	0.55	9/27	30/81	41.7	0.73	0.5	16/46	25/65	49.6	0.81	2.42	11/34	26/68	44.4	0.69	1*
		13/40	40/110	60.8	0.7	0.57	27/77	33/88	78.5	0.86	7	18/53	39/106	64.4	0.78	1*
0.5	0.65	11/26	42/88	46.2	0.67	0.42	28/59	34/70	60.1	0.9	5.36	17/39	38/78	49.3	0.73	1*
		19/45	49/104	64	0.68	0.76	24/62	45/94	78.9	0.47	1.94	23/53	50/105	66.7	0.74	1
0.6	0.75	17/27	46/67	39.3	0.69	0.68	18/30	43/62	43.8	0.57	0.94	20/39	44/78	53.6	0.63	1*
		21/34	47/71	47.1	0.65	0.71	19/40	41/72	58	0.44	1.25	20/39	44/78	53.6	0.63	1*
0.7	0.85	22/42	60/105	62.3	0.68	0.67	28/57	54/93	75	0.5	1.58	25/50	58/100	72.2	0.56	1*
		21/34	47/71	47.1	0.65	0.92	25/43	43/64	54.4	0.46	2.05	24/38	50/76	49	0.71	1
0.80	0.95	17/27	46/67	39.3	0.69	0.68	18/30	43/62	43.8	0.57	0.94	22/34	46/67	41.7	0.77	1.03
		21/34	64/95	55.6	0.65	0.56	48/72	57/84	73.2	0.9	6	28/45	61/90	59.7	0.67	1*
0.85	0.95	14/20	45/59	36.2	0.58	0.51	15/22	40/52	36.8	0.51	0.73	18/26	40/52	38	0.54	1
		14/19	46/59	30.3	0.72	0.47	16/23	39/49	34.4	0.56	0.88	21/28	44/56	34.2	0.78	1*
0.95	0.95	18/25	61/79	43.4	0.66	0.46	33/44	53/68	48.5	0.81	1.83	28/38	59/76	47.7	0.74	1*
		5/7	27/31	20.8	0.42	0.29	5/7	27/31	20.8	0.42	0.29	5/7	27/31	20.8	0.42	0.29
		7/9	26/29	17.7	0.56	0.45	7/9	26/29	17.7	0.56	0.45	7/9	26/29	17.7	0.56	0.45
		16/19	37/42	24.4	0.76	0.83	31/35	35/40	35.3	0.94	7	18/22	39/44	29.3	0.67	1

*The balanced design has both the expected sample size and the maximum sample size in between of that of the optimal design and minimax design.
^aThe lower bound for the number of responses observed in stage I. ^bThe lower bound for the total number of responses observed at the end of stage II.

Table 4 Optimal, minimax, and balanced designs for independent data, $p_1 - p_0 = 0.20$

p_0	p_1	Optimal design				Minimax design				Balanced design						
		Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$	Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$	Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$			
0.05	0.25	0/9	2/24	14.5	0.63	0.6	0/13	2/20	16.4	0.51	1.86	0/11	2/22	15.7	0.57	1*
		0/9	2/17	12	0.63	1.13	0/12	2/16	13.8	0.54	3	1/12	3/25	13.5	0.88	0.92
		0/9	3/30	16.8	0.63	0.43	0/15	3/25	20.4	0.46	1.5	1/15	3/30	17.6	0.83	1
0.10	0.30	1/12	5/35	19.8	0.66	0.52	1/16	4/25	20.4	0.51	1.78	2/17	5/34	21	0.76	1
		1/10	5/29	15	0.74	0.53	1/15	5/25	19.5	0.55	1.5	1/13	5/26	17.9	0.62	1*
		2/18	6/35	22.5	0.73	1.06	2/22	6/33	26.2	0.62	2	3/22	8/44	25.8	0.83	1
0.20	0.40	3/17	10/37	26	0.55	0.85	3/19	10/36	28.3	0.46	1.12	4/20	11/40	27.4	0.63	1
		3/13	12/43	20.6	0.75	0.43	4/18	10/33	22.3	0.72	1.2	4/18	11/36	23.1	0.72	1
		4/19	15/54	30.4	0.67	0.54	5/24	13/45	31.2	0.66	1.14	5/24	14/48	32.3	0.66	1
0.3	0.5	7/22	17/46	29.9	0.67	0.92	7/28	15/39	35	0.36	2.55	6/21	16/42	30.4	0.55	1*
		5/15	18/46	23.6	0.72	0.48	6/19	16/39	25.7	0.67	0.95	7/21	17/42	26.8	0.72	1
		8/24	24/63	34.7	0.73	0.62	7/24	21/53	36.6	0.56	0.83	9/28	22/56	36.9	0.68	1
0.4	0.6	7/18	22/46	30.2	0.56	0.64	11/28	20/41	33.8	0.55	2.15	10/24	23/48	32.4	0.65	1
		7/16	23/46	24.5	0.72	0.53	17/34	20/39	34.4	0.91	6.8	12/25	25/50	28.8	0.85	1
		11/25	32/66	36	0.73	0.61	12/29	27/54	38.1	0.64	1.16	13/30	30/60	38.6	0.71	1
0.5	0.7	11/21	26/45	29	0.67	0.88	11/23	23/39	31	0.5	1.44	12/23	27/46	30.8	0.66	1
		8/15	26/43	23.5	0.7	0.54	12/23	23/37	27.7	0.66	1.64	12/21	26/43	25.2	0.81	0.95
		13/24	36/61	34	0.73	0.65	14/27	32/53	36.1	0.65	1.04	16/29	35/59	35.9	0.77	0.97*
0.6	0.8	6/11	26/38	25.4	0.47	0.41	18/27	24/35	28.5	0.82	3.38	10/18	25/36	28.1	0.44	1*
		7/11	30/43	20.5	0.7	0.34	8/13	25/35	20.8	0.65	0.59	11/18	26/36	24.7	0.63	1
		12/19	37/53	29.5	0.69	0.56	15/26	32/45	35.9	0.48	1.37	16/25	35/50	31.8	0.73	1*
0.7	0.9	6/9	22/28	17.8	0.54	0.47	11/16	20/25	20	0.55	1.78	11/15	24/30	19.5	0.7	1
		4/6	22/27	14.8	0.58	0.29	19/23	21/26	23.2	0.95	7.67	10/14	23/28	19	0.64	1
		11/15	29/36	21.2	0.7	0.71	13/18	26/32	22.7	0.67	1.29	13/17	31/39	21.4	0.8	0.77

*The balanced design has both the expected sample size and the maximum sample size in between of that of the optimal design and minimax design.

Table 5 Optimal, minimax, and balanced designs for correlated data^a ($\pi_1 = \pi_2 = 1/2$)^b with $p_1 - p_0 = 0.15$, $\rho = 0.5$

p_0	p_1	Optimal design						Minimax design						Balanced design							
		Reject drug if response rate			EN (p_0)			Reject drug if response rate			EN (p_0)			Reject drug if response rate			EN (p_0)				
		$\leq r_1/n_1$	$\leq r/n$	$n_1:n_2$	PET (p_0)	$n_1:n_2$	$\leq r_1/n_1$	$\leq r/n$	EN (p_0)	PET (p_0)	$n_1:n_2$	$\leq r_1/n_1$	$\leq r/n$	EN (p_0)	PET (p_0)	$n_1:n_2$	$\leq r_1/n_1$	$\leq r/n$	EN (p_0)	PET (p_0)	$n_1:n_2$
0.05	0.20	1/22	5/54	31.7	0.7	0.69	1/26	5/49	34.6	0.62	1.13	1/25	5/50	33.9	0.64	1*	1/25	5/50	33.9	0.64	1*
		1/17	6/53	24.5	0.79	0.47	1/19	6/49	26.4	0.75	0.63	2/25	6/50	28.2	0.87	1	2/25	6/50	28.2	0.87	1
		2/29	7/71	36.4	0.82	0.69	4/43	6/60	44.1	0.94	2.53	2/32	7/64	38.8	0.79	1*	2/32	7/64	38.8	0.79	1*
0.10	0.25	3/29	11/73	43.5	0.67	0.66	3/35	10/63	48.1	0.53	1.25	3/32	10/64	44.8	0.6	1*	3/32	10/64	44.8	0.6	1*
		3/25	10/59	33	0.76	0.74	3/29	10/56	37.9	0.67	1.07	4/30	10/60	35.3	0.82	1	4/30	10/60	35.3	0.82	1
		4/35	13/81	47.4	0.73	0.76	5/51	13/76	61.1	0.6	2.04	6/44	14/88	50.4	0.85	1	6/44	14/88	50.4	0.85	1
0.20	0.35	8/39	24/93	59.3	0.62	0.72	11/60	22/83	72.7	0.45	2.61	9/44	23/88	60.8	0.62	1*	9/44	23/88	60.8	0.62	1*
		7/30	26/96	45.8	0.76	0.45	7/35	22/77	51.8	0.6	0.83	12/46	25/92	51.3	0.89	1*	12/46	25/92	51.3	0.89	1*
		11/49	31/116	66.7	0.74	0.73	28/102	28/104	102.1	0.97	51	12/55	30/110	71.4	0.7	1*	12/55	30/110	71.4	0.7	1*
0.3	0.45	16/50	39/109	68.7	0.68	0.85	15/51	35/96	72	0.53	1.13	17/54	39/108	72.6	0.66	1	17/54	39/108	72.6	0.66	1
		12/36	38/100	52.9	0.74	0.56	13/44	34/87	63.4	0.55	1.02	16/47	36/94	57.3	0.78	1*	16/47	36/94	57.3	0.78	1*
		18/56	51/137	80.7	0.69	0.69	20/67	48/127	94	0.55	1.12	22/67	50/134	84.3	0.74	1*	22/67	50/134	84.3	0.74	1*
0.4	0.55	19/46	53/115	71.3	0.63	0.67	24/64	47/100	85.9	0.39	1.78	23/55	51/110	73.5	0.66	1*	23/55	51/110	73.5	0.66	1*
		17/38	61/128	58	0.78	0.42	20/51	48/97	73.3	0.52	1.11	26/56	54/112	63.4	0.87	1*	26/56	54/112	63.4	0.87	1*
		25/58	72/153	83.4	0.73	0.61	62/129	62/130	129	0.97	129	33/74	70/148	87.1	0.82	1*	33/74	70/148	87.1	0.82	1*
0.5	0.65	25/48	67/119	71.6	0.67	0.68	31/61	57/100	76.6	0.6	1.56	32/60	68/120	75.6	0.74	1	32/60	68/120	75.6	0.74	1
		17/32	67/115	56.8	0.7	0.39	33/66	56/94	78.6	0.55	2.36	25/48	57/96	64	0.67	1*	25/48	57/96	64	0.67	1*
		30/56	88/154	80.7	0.75	0.57	41/78	75/129	92.6	0.71	1.53	40/73	84/146	85.7	0.83	1*	40/73	84/146	85.7	0.83	1*
0.6	0.75	26/42	66/99	61.6	0.66	0.74	33/55	60/89	70.3	0.55	1.62	29/47	63/94	63.6	0.65	1*	29/47	63/94	63.6	0.65	1*
		22/34	71/104	50.3	0.77	0.49	36/59	60/86	69.5	0.61	2.19	34/51	70/102	57.7	0.87	1*	34/51	70/102	57.7	0.87	1*
		31/49	91/135	72.4	0.73	0.57	73/108	79/116	108.3	0.96	13.5	43/66	89/132	76.8	0.84	1*	43/66	89/132	76.8	0.84	1*
0.7	0.85	23/32	59/77	47.5	0.66	0.71	33/45	54/70	51.6	0.74	1.8	25/34	52/68	43.1	0.73	1*	25/34	52/68	43.1	0.73	1*
		20/27	60/76	39.6	0.74	0.55	19/27	55/69	44.3	0.59	0.64	29/38	60/76	43.8	0.85	1*	29/38	60/76	43.8	0.85	1*
		30/42	76/97	62	0.64	0.76	73/93	74/95	93.1	0.97	46.5	35/49	77/98	66.7	0.64	1	35/49	77/98	66.7	0.64	1
0.80	0.95	21/25	45/52	31.3	0.77	0.93	17/21	43/49	31.4	0.63	0.75	20/25	44/50	35.5	0.58	1	20/25	44/50	35.5	0.58	1
		15/18	48/54	27.8	0.73	0.5	42/47	45/51	47.1	0.97	11.75	23/27	48/54	31.9	0.82	1*	23/27	48/54	31.9	0.82	1*
		20/24	54/61	33.8	0.74	0.65	20/24	54/61	33.8	0.74	0.65	22/26	46/52	31.4	0.79	1*	22/26	46/52	31.4	0.79	1*

*The balance design has both the expected sample size and the maximum sample size in between of that of the optimal design and minimax design.

^aAll the numbers of responses and sample sizes are shown as absolute numbers of observations.

^b π_i : the proportion of participants with 'i' correlated outcomes.

Table 6 Optimal, minimax, and balanced designs for correlated data^a ($\pi_1 = \pi_2 = 1/2$)^b with $p_1 - p_0 = 0.20$, $\rho = 0.5$

p_0	p_1	Optimal design				Minimax design				Balanced design						
		Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$	Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$	Reject drug if response rate $\leq r_1/n_1$	EN(p_0)	PET (p_0)	$n_1 : n_2$			
0.05	0.25	1/17	4/38	21.4	0.79	0.81	1/19	4/35	22.9	0.75	1.19	1/18	4/36	22.1	0.77	1*
		1/13	5/38	16.4	0.86	0.52	1/14	5/35	17.2	0.85	0.67	2/19	5/38	20.3	0.93	1
		1/17	6/49	23.6	0.79	0.53	1/19	6/45	25.4	0.75	0.73	2/24	6/48	26.8	0.88	1
0.10	0.30	2/20	7/43	27.4	0.68	0.87	2/25	7/40	31.9	0.54	1.67	2/21	7/42	28.4	0.65	1*
		2/15	9/48	21.1	0.82	0.45	2/18	8/39	23.6	0.73	0.86	3/22	9/44	25.8	0.83	1
		3/24	10/56	30.9	0.79	0.75	2/25	10/52	37.5	0.54	0.93	3/27	10/54	34.6	0.72	1*
0.20	0.40	4/21	16/57	35.9	0.59	0.58	5/28	15/52	40	0.5	1.17	6/28	16/56	37	0.68	1*
		5/20	18/63	28.4	0.8	0.47	5/24	15/48	32.3	0.66	1	7/27	16/54	31.2	0.84	1*
		6/27	20/70	39.3	0.71	0.63	18/61	18/63	61.1	0.97	30.5	9/37	21/74	44.1	0.81	1
0.3	0.5	8/26	23/61	39	0.63	0.74	7/25	21/55	39.6	0.51	0.83	9/29	22/58	39.6	0.64	1*
		7/20	27/69	31.2	0.77	0.41	8/25	22/53	34	0.68	0.89	10/29	24/58	35.6	0.77	1
		11/33	34/87	47.4	0.73	0.61	12/38	30/75	50.6	0.66	1.03	13/39	31/78	49.2	0.74	1*
0.4	0.6	10/25	30/62	40.3	0.59	0.68	11/30	28/57	45.4	0.43	1.11	13/31	30/62	41.5	0.66	1*
		11/24	33/66	32.9	0.79	0.57	12/29	28/54	38.1	0.64	1.16	16/33	33/66	37	0.88	1*
		14/32	44/89	47.3	0.73	0.56	38/75	38/76	75	0.98	75	19/42	42/84	50.3	0.8	1*
0.5	0.7	13/25	37/63	38.1	0.65	0.66	16/31	32/54	39.3	0.64	1.35	15/28	33/56	36	0.71	1*
		11/20	42/70	32.6	0.75	0.4	15/28	34/55	35.7	0.71	1.04	19/33	40/66	37.9	0.85	1
		15/28	51/86	44.6	0.71	0.48	24/44	43/71	50.1	0.77	1.63	23/41	49/82	48.2	0.83	1*
0.6	0.8	14/23	39/57	36.2	0.61	0.68	18/30	36/52	39.5	0.57	1.36	20/31	42/62	38.6	0.75	1*
		11/17	42/59	28.1	0.74	0.4	22/32	36/50	34.1	0.88	1.78	18/27	39/54	32	0.82	1*
		15/24	50/71	39.4	0.67	0.51	39/56	46/65	56.5	0.95	6.22	26/39	55/78	45.1	0.84	1
0.7	0.9	9/13	34/43	25.6	0.58	0.43	12/17	32/40	25.9	0.61	0.74	14/20	32/40	28.3	0.58	1
		9/12	35/43	19.8	0.75	0.39	10/14	29/35	21.5	0.64	0.67	15/20	33/40	24.8	0.76	1
		18/24	45/56	31.3	0.77	0.75	20/27	43/53	33.7	0.74	1.04	21/28	45/56	34.2	0.78	1

*The balanced design has both the expected sample size and the maximum sample size in between of that of the optimal design and minimax design.
^aAll the numbers of responses and sample sizes are shown as absolute numbers of observations. ^b π_i : the proportion of participants with 'i' correlated outcomes.

trial designs when the investigators would like to monitor the trial near the halfway point of the study. The method balances the sample sizes of the two stages without causing a appreciable increase in total sample size over Simon's optimal and mini-max designs.

Acknowledgements

This work was supported by Vanderbilt University through grants CCSG (5P30 CA068485-11), Breast SPORE (5P50 CA098131-05), Lung SPORE (2P50 CA090949-06A1), and GI SPORE (2P50 CA095103-06). We would like to thank Charity G. Moore, Ph.D, for her helpful comments on this work. We also thank the editor and reviewers for their thoughtful review of this manuscript.

References

1. Heitjan DF. Bayesian interim analysis of phase II cancer clinical trials. *Statistical Medicine* 1997; **16**: 1791–802.
2. Simon R. Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 1989; **10**: 1–10.
3. Sargent DJ, Sloan JA, Cha SS. Sample size and design considerations for phase II clinical trials with correlated observations. *Controlled Clinical Trials* 1999; **20**: 242–52.
4. Bellissant E, Benichou J, Chastang C. Design and analysis of phase II trials in oncology with a group sequential method, the triangular test. *Therapie* 1991; **46**: 21–7.
5. Sebillé V, Bellissant E. Sequential methods and group sequential designs for comparative clinical trials. *Fundamental Clinical Pharmacology* 2003; **17**: 505–16.
6. Jennison C. Efficient group sequential tests with unpredictable group sizes. *Biometrika* 1987; **74**: 155–66.
7. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials* (3rd ed.) Springer Verlag, 1998.
8. Gehan EA. The determination of the number of participants required in a follow-up trial of a new chemotherapeutic agent. *J Chronic Disease* 1961; **13**: 346–53.
9. Fleming TR. One-sample multiple testing procedure for phase II clinical trials. *Biometrics* 1982; **38**: 143–51.