國立交通大學

統計學研究所

碩士論文

臨床試驗之連續性滿足點 Phase II/III 調適性設計

An Alternative Phase II/III Design in Clinical Trials for Continuous Endpoint

研究生：劉佳佩

指導教授：蕭金福　教授

中華民國九十九年六月

臨床試驗之連續性滿足點 Phase II/III 調適性設計

An Alternative Phase II/III Design in Clinical Trials for Continuous Endpoint

研 究 生：劉佳佩　　Student：Chia-Pei Liu

指導教授：蕭金福　　Advisor：Chin-Fu Hsiao

國 立 交 通 大 學

統計學研究所

碩 士 論 文

A Thesis

Submitted to Institute of Statistics

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

臨床試驗之連續性滿足點 Phase II/III 調適性設計

研究生：劉佳佩　　指導教授：蕭金福 博士

國立交通大學統計學研究所

中文摘要

醫藥研發是非常具有風險、複雜、高成本又費時的過程。其中大部分的時間和經費是用於臨床試驗。因此，我們需要更有效率且可靠的臨床試驗方法，來分析資料及評估藥物的風險和效用，以減少病人樣本數，縮短發展期間，進而降低藥物開發成本。在本文中，我們發展一個連續性滿足點的 phase II / III 試驗設計，在 phase II 試驗中評估不同劑量的試驗藥物與對照組兩兩之間比較。在 phase II 試驗，隨機分派病人接受一個試驗藥物劑量或對照組。如果一個或一些劑量有統計顯著的療效優於對照組，這些劑量將被選中進入 phase III 試驗。此外，病人在被選中的劑量組與對照組將繼續進入 phase III 試驗。同時也將招募新的病人，隨機分配至所選的藥物劑量組或對照組。我們會計算每階段的檢定統計量之臨界值，以確定各藥物劑量在試驗過程中是否應被淘汰或被選擇，並計算合適的樣本大小以方便招募病人。在我們的設計中，由於將傳統的 phase II 試驗及 phase III 試驗合併成單一的試驗，同時，在 phase II 試驗收集的資料也會進入 phase III 試驗做最後的分析，因此可以減少樣本數量和節省試驗時間。

關鍵詞：調適設計、臨床試驗、phase II/III 設計

An Alternative Phase II/III Design in Clinical Trials for Continuous Endpoint

Student：Chia-Pei Liu　　Advisor：Dr. Chin-Fu Hsiao

Institute of Statistics
National Chiao Tung University

Abstract

The pharmaceutical development is very risky, complex, costly, and time-consuming. Much of time and costs were spent in clinical trials. Hence, we need the methods which can be more efficient and reliable to minimize sample size, shorten the period of development duration, and thus reduce the cost for drug development. In this paper, we demonstrate the phase II/III design for continuous endpoint based on pariwise comparisons at the phase II stage when evaluating the efficacy of drugs. For the phase II stage, patients are randomly assigned to receive either one of the several doses of the test drug or to the control group. If one or some doses are declared to have a statistically significantly superior efficacy over control, these doses will be selected for the phase III trial. In addition, the patients in the selected doses and control groups will be continued to the phase III stage. Also new patients will be recruited and randomized to receive either the selected doses of the test drug or to the control group. We will find the critical value at each stage to determine whether the treatment should be dropped out or selected in the process of the trials, and compute the required sample size for facilitating recruitment of patients. In our design, since we integrate the traditional phase II and III trials into a single trial, and the data collected from phase II stage will also be included into the final analysis, sample size reduction and trial time saving may be possible.

Key Words: adaptive design, clinical trial, phase II/III design

誌謝

能完成這篇論文，我十分感謝指導教授蕭金福老師給予的細心教導，老師在百忙之中總是不厭其煩地給我寶貴的意見與方向，才使這篇論文能充實嚴謹的順利完成。同時謝謝鄒小蕙老師，在我發生困惑時，協助我解決問題；感謝翁賢學長提供的經驗和想法；在做此論文的過程中，謝謝敏琪陪伴我一起成長努力；以及感謝口試委員們的建議，使得論文更加完善。

在念研究所的這些日子，感謝所上所有老師們的認真教學，謝謝郭姊的幫忙還有處理所上的事務。最後謝謝家人和朋友們的鼓勵與支持，讓我更有信心完成這篇論文，謝謝同學們在這段期的照顧，不論是課業上或是生活上幫助我很多事情，雖然相處的時間不長，但我會永遠想念這兩年來和大家一起努力、一起相聚的美好回憶。

劉佳佩 謹誌于
國立交通大學統計研究所
中華民國九十九年六月

Content

1. Introduction

The pharmaceutical development is very risky, complex, costly, and time-consuming. Much of time and costs were spent in clinical trials. Even if there is a better understanding of disease etiology and higher technology in medical production, the success rate of drug development has been low. One of the probable reasons is that the current methods used for developing new drugs may not be practicable. Hence, we need the methods which can be more efficient and reliable to minimize sample size, shorten the period of development duration, and thus reduce the cost for drug development.

In the traditional phase II and phase III design, there exists a lead period between phase II and phase III trials. In addition, the data collected in the phase II trial are not used in the phase III trial. Adaptive seamless phase II/III designs have been regarded as a feasible way in which shortening the trial time may be possible. More specifically, an adaptive seamless phase II/III design combines the traditional separated trials (phase II and phase III trials) into a single trial, and use data from patients enrolled before and after the adaptation in the final analysis (Maca et al. 2006). In a seamless phase II/III design, the phase II stage corresponds to the learning phase to evaluate whether a new medication is effective and choose the best doses for confirmatory phase, and the phase III stage then corresponds to the confirmatory phase to confirm its effectiveness and evaluate its safety. Since the data collected from phase II stage will also be included into the final analysis, sample size reduction and trial time saving may be possible.

Some statistical methods related to the adaptive seamless phase II/III designs

have been proposed. Simon (1989) and Tsou et al. (2008) developed two-stage screening designs for phase II which can minimize the expected sample size and stop early if the new regimen has low activity subject to constraints on the size of the type I error rate and power for discrete and continuous efficacy endpoints respectively. Schaid et al. (1990) proposed the phase II/III designs based on pariwise comparisons for survival endpoints for cancer drugs. The design allows multiple treatments to be tested at the same time and bases the determination to proceed from the phase II trial on the same clinical endpoint evaluated in the same population as the phase III trial. A concurrent control group is also treated. This design may offer a substantial saving when the hazard rate is large relative to the patient accrual rate, a situation often encountered in clinical trials in advanced cancer. Scher et al. (2002) used this design for castrate metastatic prostate cancer for which multiple regimens appear to have similar activity at this time. Posch et al. (2005) described a general formulation of the adaptive testing procedure in the context of treatment selection. In addition, they proposed multiplicity adjusted p-values, introduced simultaneous confidence bounds, and investigated the statistical properties of point estimates. Bischoff et al. (2005) developed a two-stage adaptive design with a minimal number of patients that controls the type I error rate, and achieves a required power to detect a given clinically relevant difference in means, and controls the probability of wrong selection. The information received about the variance is used to determine the number of patients on the selected treatment and control for the second stage. Maca et al. (2006) have proposed the concept of adaptive seamless phase II/III designs and describe the statistical methodologies related to adaptive seamless designs. They also describe the decision process involved with seamless designs and present some illustrative examples.

An adaptive seamless phase II/III design is to conduct the learning phase trial (phase II stage) and the confirmatory phase trial (phase III stage) simultaneously under the same protocol and the same study population. Therefore, it can eliminate the lead time between the separated trials, and thus reduce the sample size required and shorten the drug development time. In this paper, we demonstrate the phase II/III design for continuous endpoint based on pariwise comparisons at the phase II stage when evaluating the efficacy of drugs. The goal is to find the critical value at each stage to determine whether the treatment should be dropped out or selected in the process of the trials, and compute the required sample size for facilitating recruitment of patients. In our design, both the data from the learning phase and the confirmatory phase will be used in the final analysis. We also compare our design with the traditional separated phase II and phase III design. In this paper, a seamless adaptive phase II/III design based on continuous efficacy endpoint is described in Section 2. Some numerical results of our proposed design are provided in Section 3. Final remarks and discussion are given in Section 4.

2. A Seamless Adaptive Phase II/III Design

Our main goal is to compare a test drug for several different doses with a control group based on some continuous efficacy endpoint. Here, we assume that the total number of doses is *K*. For the phase II stage, patients are randomly assigned to receive either one of the *K* doses of the test drug or to the control group. We aim to compare each dose of the test drug with the control group, and the Bonferroni correction is employed to adjust p-values for multiple comparisons. If one or some doses are declared to have a statistically significantly superior efficacy over control, these doses will be selected for the phase III trial. In addition, the patients in the selected doses and control groups will be continued to the phase III stage. Also new patients will be recruited and randomized to receive either the selected doses of the test drug or to the control group. The final analysis includes the data of the selected doses and control groups from both phase II and phase III stages.

Let n_2 be the number of patients recruited for each of K+1 groups at the phase II stage. Let Y_{0j} and Y_{ir} be the efficacy responses for patient j receiving control group and patient r receiving the i^{th} dose group respectively, $i = 1, 2, \ldots, K$, $j = 1, 2, \ldots, n_2$, $r = 1, 2, \ldots, n_2$. We assume that the efficacy responses are continuous and

$$Y_{0j} \overset{iid}{\sim} N(\mu_0, \sigma^2) \text{ and } Y_{ir} \overset{iid}{\sim} N(\mu_i, \sigma^2),$$

where $N(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2. Here we also assume σ^2 is known at the design stage. The selection of the doses is based on the following hypothesis:

$$\mathrm{H}_0 : \mu_i - \mu_0 \le 0 \text{ vs. } \mathrm{H}_1 : \mu_i - \mu_0 > 0, \quad i = 1, 2, \ldots, K. \tag{1}$$

The observed sample means for the control group and the i^{th} experimental group based on n_2 patients can be respectively derived by

$$\overline{Y}_0^{\text{II}} = \sum_{j=1}^{n_2} \frac{Y_{0j}}{n_2} \text{ and } \overline{Y}_i^{\text{II}} = \sum_{j=1}^{n_2} \frac{Y_{ij}}{n_2}, \quad i = 1, 2, \ldots, K .$$

Let $\Delta_i = \mu_i - \mu_0$ and $\hat{\Delta}_i^{\text{II}}$ be the estimate of Δ_i. Then $\hat{\Delta}_i^{\text{II}} = \overline{Y}_i^{\text{II}} - \overline{Y}_0^{\text{II}}$ and

$$\hat{\Delta}_i^{\text{II}} \sim N\left(\Delta_i, \frac{2\sigma^2}{n_2}\right), i = 1, 2, \ldots, K .$$

Let $T_i^{\text{II}} = \dfrac{\hat{\Delta}_i^{\text{II}}}{\sqrt{\dfrac{2\sigma^2}{n_2}}}$. It follows that

$$T_i^{\text{II}} \sim N\left(\frac{\Delta_i}{\sqrt{\dfrac{2\sigma^2}{n_2}}}, 1\right), i = 1, 2, \ldots, K . \tag{2}$$

At the phase II stage, if T_i^{II} is less than a given value C_1, for all $i = 1, 2, \ldots, K$, or any one of T_i^{II} is greater than C_2, we will stop the trial and further clinical development will not be considered. The former indicates that all of the doses for test product are futile, whereas the latter says that there exists at least one experimental treatment regimen to have overwhelming advantage. On the other hand, we will continue to recruit patients to the control group and all experimental groups for which $C_1 \le T_i^{\text{II}} \le C_2$ for the phase III stage.

Let n_3 be the number of patients required for each of groups selected for the phase III stage. Similarly, the observed sample means based on n_3 patients for the control group and the i^{th} experimental group selected for the phase III stage can

respectively be expressed as

$$\bar{Y}_0^{\text{III}} = \sum_{j=n_2+1}^{n_2+n_3} \frac{Y_{0j}}{n_3} \quad \text{and} \quad \bar{Y}_i^{\text{III}} = \sum_{j=n_2+1}^{n_2+n_3} \frac{Y_{ij}}{n_3} \quad \text{, for some } i,\ i = 1, 2, \ldots, K .$$

In addition, the overall sample means based on $n_2 + n_3$ patients for the control group and the i^{th} experimental group selected the phase III stage are respectively given by

$$\bar{Y}_0 = \sum_{j=1}^{n_2+n_3} \frac{Y_{0j}}{n_2+n_3} \quad \text{and} \quad \bar{Y}_i = \sum_{j=1}^{n_2+n_3} \frac{Y_{ij}}{n_2+n_3} \text{, for some } i,\ i = 1, 2, \ldots, K .$$

Let $T_i^{\text{III}} = \dfrac{\bar{Y}_i^{\text{III}} - \bar{Y}_0^{\text{III}}}{\sqrt{\dfrac{2\sigma^2}{n_3}}}$. Let T_i be the test statistic for the final analysis. Then

$$T_i = \frac{\bar{Y}_i - \bar{Y}_0}{\sqrt{\dfrac{2\sigma^2}{n_2+n_3}}},$$

and we have

$$T_i = \frac{\sqrt{n_2}T_i^{\text{II}} + \sqrt{n_3}T_i^{\text{III}}}{\sqrt{n_2+n_3}}.$$

The i^{th} dose is declared to be superior to the control group if $T_i > C_3$.

In our design, each dose for the test product is compared to the control group, and thus we define α to be the pairwise alpha-error for each comparisons and $K\alpha$ to be the Bonferroni approximation of the overall alpha-error. Consequently, the overall probability of "accepting" the i^{th} dose in both stages with the true parameters Δ_i is a function of $\sigma, \Delta_i, n_2, n_3, C_1, C_2,$ and C_3, and it is given as

$$\begin{aligned} &\phi(\sigma, \Delta_i, n_2, n_3, C_1, C_2, C_3) \\ &= P_{\Delta_i}\left(T_i^{\text{II}} > C_2\right) + \int_{C_1}^{C_2} f_{T_i^{\text{II}}}(x) P_{\Delta_i}\left(T_i > C_3 \middle| T_i^{\text{II}} = x\right) dx, \end{aligned} \tag{3}$$

where $f_{T_i^{\text{II}}}(\cdot)$ is the probability density function for T_i^{II}. Subsequently, the pairwise

power $1-\beta$ given $\Delta_i = \Delta$ and the pairwise alpha-error α can be written as

$$1-\beta = \phi(\sigma, \Delta, n_2, n_3, C_1, C_2, C_3), \tag{4}$$

and

$$\alpha = \phi(\sigma, 0, n_2, n_3, C_1, C_2, C_3) \tag{5}$$

respectively. The equation (3) can be re-expressed as

$$\begin{aligned}
&\phi(\sigma, \Delta_i, n_2, n_3, C_1, C_2, C_3) \\
&= 1 - \Phi\left(C_2 - \frac{\Delta_i}{\sqrt{\frac{2\sigma^2}{n_2}}} \right) \\
&+ \int_{C_1}^{C_2} \frac{1}{\sqrt{2\pi}} \exp\left(- \frac{\left(x - \frac{\Delta_i}{\sqrt{\frac{2\sigma^2}{n_2}}} \right)^2}{2} \right) \Phi\left(C_3 \sqrt{\frac{(n_2+n_3)}{n_3}} + \sqrt{\frac{n_2}{n_3}} x + \frac{\Delta_i}{\sqrt{\frac{2\sigma^2}{n_3}}} \right) dx.
\end{aligned}$$

Furthermore, the expected total sample size, $E(N)$, under the null hypothesis, can be evaluated by

$$E(N) = (K+1)n_2 p_0 + \sum_{j=1}^{K} \{(n_3+n_2)(j+1) + n_2(K-j)\} p_j, \tag{6}$$

where p_0 denotes the probability of stopping accrual at the phase II stage, and p_j ($j = 1, 2, \ldots, K$) is the probability that accrual will continue for the standard treatment and j of the experimental treatments. The derivations of p_0 and p_j can be found in the Appendix.

Under the specification of design parameters Δ, σ, α, β, and C_1, the proposed phase II/III design is to determine n_2, n_3, C_2, and C_3 numerically based on

constraints of pairwise type I and II error rates given in (4) and (5) and to minimize the expected total sample size (6). In our design, C_1 should be pre-specified. The determination of C_1 should meet the minimal clinically meaningful requirement that an investigator would need to observe before continuing accrual onto the phase III stage. Furthermore, for convention, C_3 is chosen as $C_3 = \Phi^{-1}(1-\alpha)$. In other words, we consider C_3 as if one-stage design was conducted and the Bonferroni procedure was applied to the K multiple treatment comparisons.

3. Results

In this section, we are at the position to give some examples. Given ($K\alpha, 1-\beta$) = (0.05, 0.8), ($K\alpha, 1-\beta$) = (0.025, 0.8), Tables 1－6 illustrate the seamless adaptive design for different combinations of design parameters with K=1, K=2, and K=3, respectively. For each K, we consider various combinations of values for C_1=0, C_1=0.5, σ=13, 15, and 17, Δ=5, 6, and 7. The tabulated results include the early stopping upper boundary for concluding efficacy for the test drug at the phase II stage (C_2), the critical value that would reject the test drug at the phase III stage (C_3), the required sample size per group for the phase II stage (n_2), the required sample size per group for the phase III stage (n_3), the expected total sample size ($E(N)$), numbers of sample sizes required for the traditional phase II and phase III trials ($n_2^{'}$, and $n_3^{'}$, respectively), and the ratios of the maximum of the total sample size for our phase II/III design vs. the maximum of the total sample size for the traditional designs (*ratio*). Here the total sample sizes per group required for traditional phase II and phase III design are derived by

$$n_2^{'} = \left[\frac{\sqrt{2\sigma^2}\left(Z_\alpha + Z_\beta\right)}{\Delta} \right]^2,$$

and

$$n_3^{'} = \left[\frac{\sqrt{2\sigma^2}\left(Z_\alpha + Z_\beta\right)}{\Delta} \right]^2$$

respectively.

For instances, Table 1 displays the results for $K\alpha$=0.05, $1-\beta$=0.8 and K=1. The first line considers the case of C_1=0, σ=13, and Δ=5. That is, only one dose

group for the test product is included in the drug development. In this case, the phase II stage needs to recruit 26 patients for the control group and the experimental group. When the study is completed at the phase II stage, if the observed value of T_1^{II} exceeds 2.37, the trial is terminated and we declare that the experimental treatment regimen to have overwhelming advantage. If T_1^{II} is less than 0, the trial is also terminated and we conclude the test product is futile. On the other hand, if $0 \le T_1^{II} \le 2.37$, the trial continues to the phase III stage to enroll additional 66 patients for the control group and the experimental group. After the recruitment of the patients at the phase III stage is completed and if the observed overall value of T_1 based on the cumulative data of $n_1 + n_2$ obtained at the end of the trial does not exceed 1.64, then the lack of efficacy of the test drug is concluded. However, on the other hand, if the observed overall value of T_1 is greater than 1.64, the test drug is declared to be superior to the control group. In addition, the numbers of sample sizes required for the traditional phase II trial and phase III trial are respectively 84 and 84 per group. Subsequently, it can be seen from Table 1 that the total sample size required for our phase II/III design can be reduced by around 45% compared with the traditional design.

As noted from Tables 1－6 that, for fixed K, C_1, and σ, as Δ increases, the corresponding required sample size decreases. This makes intuitive sense since the larger the effect size (Δ/σ), the smaller the required sample size. It can be seen that as C_1 increases, both C_2 and the expected total sample size decrease. This phenomenon occurs because we will spend more type I error rate and power for early stopping for efficacy, and thus the value of C_2 is reduced. Also the more stringent (large) the C_1 is, the larger sample size required for each stage. On the other hand, larger C_1 will cause that the ineffective doses will be quickly eliminated during the

phase II stage. Doing so will increase the probability of early termination for futility at the phase II stage and subsequently reduce the expected total sample size.

Given ($K\alpha, 1-\beta$) = (0.05, 0.8), Figures 1－3 show the simulations comparing our proposed design with the traditional design in terms of success rates for K=1, σ = 13, 15, and 17, $\Delta = 5, 6,$ and 7, $C_1 = 0$, and 0.5 respectively. For example, given $C_1 = 0$, $\sigma = 13$, and $\Delta = 5$, we can derive that $n_2 = 26$, $n_3 = 66$, $C_2 = 2.37$, $C_3 = 1.64$, $n_2^{'} = 84$, and $n_3^{'} = 84$. We assume that the true values of Δ are 0, 0.25, 0.5, 0.75, …, 9 respectively. The success rates are obtained from simulations of 10000 replicates, and the simulation results are presented in Figure 1. From Figure 1, when $\Delta = 5$, the success rate of our proposed design is close to the desired power 0.8. Also it can also be seen that our proposed design performs better than the traditional design. Other simulation results with different combinations of design parameters can be seen from Figures 2 and 3. All figures show the same phenomenon as Figure 1.

4. Discussion

In this paper, we developed a seamless phase II/III design based on continuous endpoint to evaluate the efficacy of a test drug. More specifically, we integrate the traditional phase II and III trials into a single trial. One attractive feature in our design is that the phase II and phase III trials are conducted in the same protocol with the same inclusion/exclusion criteria, the same study design, the same control, the same methods for evaluation, and the same efficacy/safety endpoints. This can avoid the difficulties arising from the current phase II and III paradigm including different patient populations recruited for phase II and phase III trials and possible different primary efficacy endpoints used in the phase II and III trials.

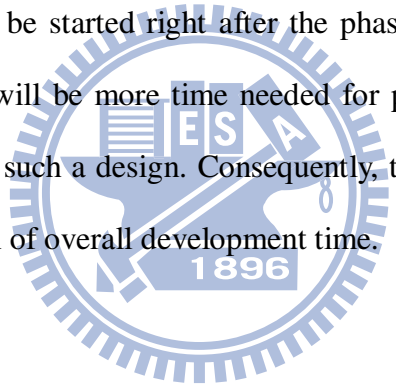
Another attractive feature is that our proposed design can shorten the time of clinical development because there is no lead time between the learning phase (phase II) and the confirmatory phase (phase III). Also data collected at the phase II stage are combined with those data obtained at the phase III stage for final analysis. As a result, the total sample size might be reduced and thus considerably valuable resource and cost can be saved.

Another point we wish to make is that choosing C_1 is rather critical. In fact, the determination of C_1 should meet the minimal clinically meaningful requirement that an investigator would need to observe before continuing accrual onto the phase III stage. Also the value of C_1 should be stringent enough that ineffective doses should be quickly eliminated during the phase II stage. It can be expected that larger value of C_1 will lead to higher probability of early stopping. Also larger value of C_1 can also increase the success probability of the phase III stage for the clinical development.

Note that when both powers in the traditional phase II and phase III designs are assumed to be 0.8, then the actual power is 0.64 (0.8×0.8) if phase II and III trials are conducted separately. However the power in our proposed phase II/III design is maintained at 0.8. Consequently, our proposed phase II/III design can derive more power than the traditional design. In the same way, suppose the type I error rates in the traditional phase II and phase III designs are both assumed to be 0.05, then the actual type I error rate is 0.0025 (0.05×0.05) if phase II and III trials are conducted separately. But the type I error rate in our proposed phase II/III design is controlled at 0.05 level. Hence, conducting phase II and III trials separately is more conservative than our proposed phase II/III design. We can clearly see the phenomenon in the Figures 1－3.

Although our proposed design has many advantages, not all clinical development can be conducted by such designs. Maca et al. (2006) outline some criteria for determining the feasibility for this type of design: endpoints and enrollment, and clinical development time. In our phase II/III design, prior to the interim analysis at which the dose to be continued will be chosen, there will be a period of the study during which some patients have been randomized but have not yet been followed long enough to reach the endpoint for evaluation. When the time needed to reach this endpoint is short relative to the total enrollment time of the study, enrollment can still continue uninterrupted with relatively few patients enrolled during this “transition” period. Even though patients enrolled during this period and randomized to doses that will not be continued will not be providing direct evidence for the comparison of the selected dose vs. the control at the phase III stage, they can be used to understand better the dose response and safety profile. If the endpoint

duration is too long, it could cause inefficiency because many patients will need to be randomized during this lead period between phase II stage and phase III stage. But if we suspend to recruit patients at the lead period, it may cause the trial to disrupt. In such case, a surrogate marker might be considered. They suggested using well-established and well-understood endpoints or surrogate marker when executing adaptive seamless designs. Also, an adaptive seamless design is not suitable if the target of the phase II is to decide a primary endpoint into phase III. If the seamless trial is the only pivotal trial, it is clear that the development time can be reduced. If the phase II/III trial is one of two required pivotal trials, the second pivotal trial should be completed close to the time the seamless study is completed. That is, the second pivotal trial should be started right after the phase II interim analysis in the phase II/III design. There will be more time needed for planning, development, and health authority review for such a design. Consequently, this additional time must be included into the evaluation of overall development time.



List of Tables

Table1. K=1, $(K\alpha, 1-\beta)$ = (0.05, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	2.37	1.64	26	66	116.82	84	84	0.55
		6	2.37	1.64	18	46	81.19	59	59	0.54
		7	2.37	1.64	14	34	61.40	43	43	0.56
	15	5	2.36	1.64	35	88	156.41	112	112	0.55
		6	2.37	1.64	24	61	107.91	78	78	0.54
		7	2.37	1.64	18	45	80.21	57	57	0.55
	17	5	2.36	1.64	45	113	200.93	143	143	0.55
		6	2.37	1.64	31	79	139.58	100	100	0.55
		7	2.37	1.64	23	58	102.97	73	73	0.55
0.5	13	5	2.11	1.64	35	70	110.77	84	84	0.63
		6	2.11	1.64	24	48	75.96	59	59	0.61
		7	2.11	1.64	18	36	56.97	43	43	0.63
	15	5	2.11	1.64	47	92	147.56	112	112	0.62
		6	2.11	1.64	32	64	101.27	78	78	0.62
		7	2.11	1.64	24	47	75.38	57	57	0.62
	17	5	2.11	1.64	60	119	189.25	143	143	0.63
		6	2.11	1.64	42	82	131.75	100	100	0.62
		7	2.11	1.64	31	61	97.53	73	73	0.63

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

Table2. K=2, $(K\alpha, 1-\beta)$ = (0.05, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	2.72	1.96	30	84	228.43	107	107	0.53
		6	2.72	1.96	21	58	158.60	74	74	0.53
		7	2.73	1.96	15	43	115.88	55	55	0.53
	15	5	2.71	1.96	40	111	302.88	142	142	0.53
		6	2.72	1.96	28	77	210.90	99	99	0.53
		7	2.72	1.96	21	57	156.95	73	73	0.53
	17	5	2.70	1.96	51	143	388.55	182	182	0.53
		6	2.71	1.96	36	99	271.13	127	127	0.53
		7	2.72	1.96	26	73	198.31	93	93	0.53
0.5	13	5	2.48	1.96	38	87	204.07	107	107	0.58
		6	2.48	1.96	27	61	144.17	74	74	0.59
		7	2.48	1.96	20	45	106.61	55	55	0.59
	15	5	2.47	1.96	51	116	273.05	142	142	0.59
		6	2.48	1.96	35	81	188.86	99	99	0.59
		7	2.48	1.96	26	59	139.10	73	73	0.58
	17	5	2.47	1.96	65	149	349.16	182	182	0.59
		6	2.47	1.96	45	104	242.65	127	127	0.59
		7	2.48	1.96	33	76	177.69	93	93	0.59

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

Table3. K=3, $(K\alpha, 1-\beta)$ = (0.05, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	2.88	2.13	31	96	337.89	120	120	0.53
		6	2.89	2.13	22	67	237.31	83	83	0.54
		7	2.89	2.13	16	49	173.21	61	61	0.53
	15	5	2.87	2.13	41	127	446.89	159	159	0.53
		6	2.88	2.13	29	89	314.31	111	111	0.53
		7	2.89	2.13	21	65	228.86	81	81	0.53
	17	5	2.86	2.13	53	163	574.98	204	204	0.53
		6	2.87	2.13	37	113	399.73	142	142	0.53
		7	2.88	2.13	27	84	295.17	105	105	0.53
0.5	13	5	2.65	2.13	39	100	298.64	120	120	0.58
		6	2.66	2.13	27	69	206.45	83	83	0.58
		7	2.66	2.13	20	51	152.78	61	61	0.58
	15	5	2.65	2.13	52	132	396.23	159	159	0.58
		6	2.65	2.13	36	92	275.24	111	111	0.58
		7	2.66	2.13	27	68	205.03	81	81	0.59
	17	5	2.65	2.13	67	170	510.33	204	204	0.58
		6	2.65	2.13	47	118	356.29	142	142	0.58
		7	2.66	2.13	34	87	260.12	105	105	0.58

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

Table4. K=1, $(K\alpha, 1-\beta)$ = (0.025, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	2.67	1.96	28	88	143.34	107	107	0.54
		6	2.68	1.96	20	61	100.56	74	74	0.55
		7	2.69	1.96	15	45	74.68	55	55	0.55
	15	5	2.67	1.96	38	117	192.10	142	142	0.55
		6	2.68	1.96	26	81	132.40	99	99	0.54
		7	2.68	1.96	19	60	97.56	73	73	0.54
	17	5	2.65	1.96	48	149	243.82	182	182	0.54
		6	2.67	1.96	34	104	171.21	127	127	0.54
		7	2.68	1.96	25	77	126.43	93	93	0.55
0.5	13	5	2.45	1.96	37	93	130.06	107	107	0.61
		6	2.45	1.96	26	64	90.59	74	74	0.61
		7	2.46	1.96	19	47	66.34	55	55	0.60
	15	5	2.44	1.96	49	123	172.12	142	142	0.61
		6	2.45	1.96	34	86	119.84	99	99	0.61
		7	2.45	1.96	25	63	87.99	73	73	0.60
	17	5	2.44	1.96	63	158	221.17	182	182	0.61
		6	2.45	1.96	44	110	154.29	127	127	0.61
		7	2.45	1.96	32	81	112.83	93	93	0.61

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

Table5. K=2, $(K\alpha, 1-\beta)$ = (0.025, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	2.97	2.24	32	105	270.09	129	129	0.53
		6	2.98	2.24	22	73	187.06	90	90	0.53
		7	2.99	2.24	16	54	137.56	66	66	0.53
	15	5	2.95	2.24	42	139	356.40	172	172	0.53
		6	2.97	2.24	30	97	250.83	119	119	0.53
		7	2.98	2.24	22	71	183.74	88	88	0.53
	17	5	2.94	2.24	54	178	456.97	220	220	0.53
		6	2.96	2.24	38	124	319.56	153	153	0.53
		7	2.97	2.24	28	91	234.89	113	113	0.53
0.5	13	5	2.76	2.24	40	110	236.03	129	129	0.58
		6	2.76	2.24	28	76	164.19	90	90	0.58
		7	2.77	2.24	20	56	119.10	66	66	0.58
	15	5	2.75	2.24	53	146	312.95	172	172	0.58
		6	2.76	2.24	37	101	217.54	119	119	0.58
		7	2.77	2.24	27	75	160.13	88	88	0.58
	17	5	2.74	2.24	68	186	400.06	220	220	0.58
		6	2.75	2.24	47	130	278.10	153	153	0.58
		7	2.76	2.24	35	96	206.27	113	113	0.58

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

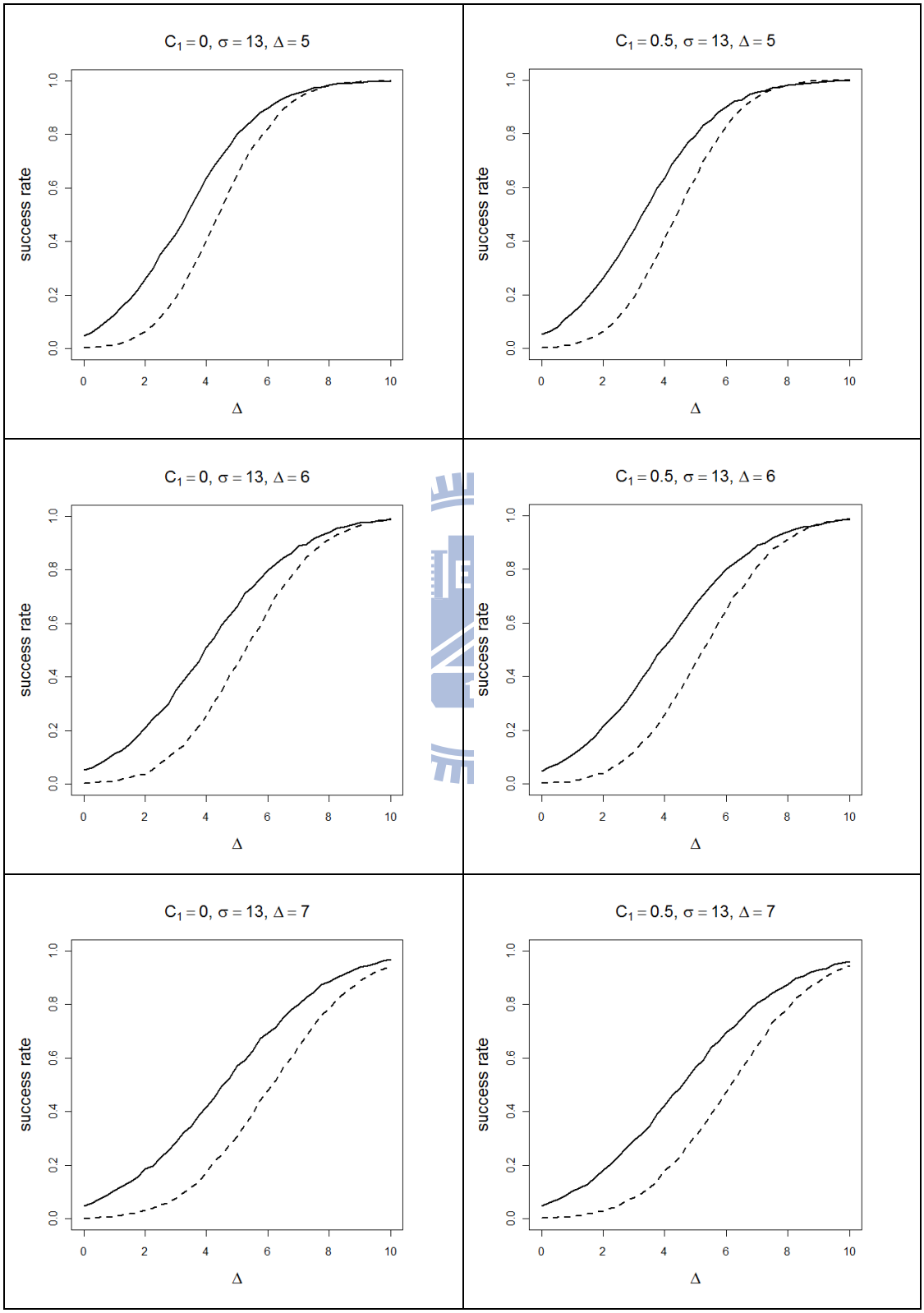
Table6. K=3, $(K\alpha, 1-\beta)$ = (0.025, 0.8)

C_1	σ	Δ	C_2	C_3	n_2	n_3	$E(N)$	$n_2^{'}$	$n_3^{'}$	ratio
0	13	5	3.10	2.39	33	117	393.99	142	142	0.53
		6	3.12	2.39	23	81	273.42	99	99	0.53
		7	3.13	2.39	17	60	202.41	73	73	0.53
	15	5	3.09	2.39	44	155	522.98	189	189	0.53
		6	3.11	2.39	30	108	361.85	131	131	0.53
		7	3.12	2.39	22	79	264.95	97	97	0.52
	17	5	3.07	2.39	56	198	667.10	243	243	0.52
		6	3.09	2.39	39	138	464.96	169	169	0.52
		7	3.11	2.39	29	102	344.42	124	124	0.53
0.5	13	5	2.91	2.39	41	121	339.19	142	142	0.57
		6	2.92	2.39	29	85	239.11	99	99	0.58
		7	2.92	2.39	21	62	173.82	73	73	0.57
	15	5	2.90	2.39	54	161	449.02	189	189	0.57
		6	2.91	2.39	38	112	314.17	131	131	0.57
		7	2.92	2.39	28	83	232.22	97	97	0.57
	17	5	2.89	2.39	69	207	575.48	243	243	0.57
		6	2.90	2.39	48	144	400.45	169	169	0.57
		7	2.91	2.39	36	106	297.50	124	124	0.57

$n_2^{'}$ and $n_3^{'}$ denotes the total sample sizes per group required for traditional phase II and phase III design respectively.

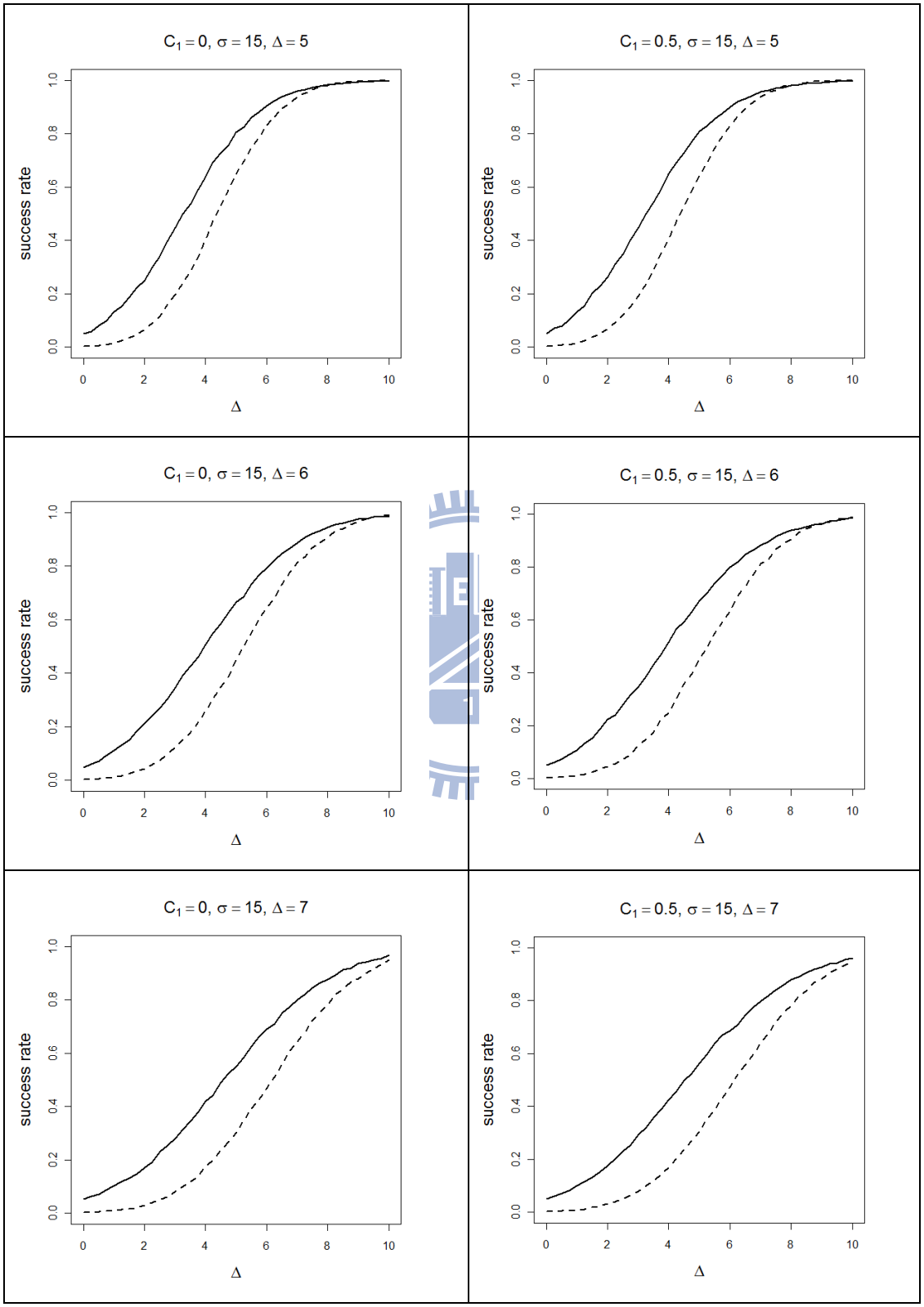
List of Figures

Figure1. Simulated success rates for the case of K=1, σ=13, $(K\alpha, 1-\beta)$ = (0.05, 0.8).



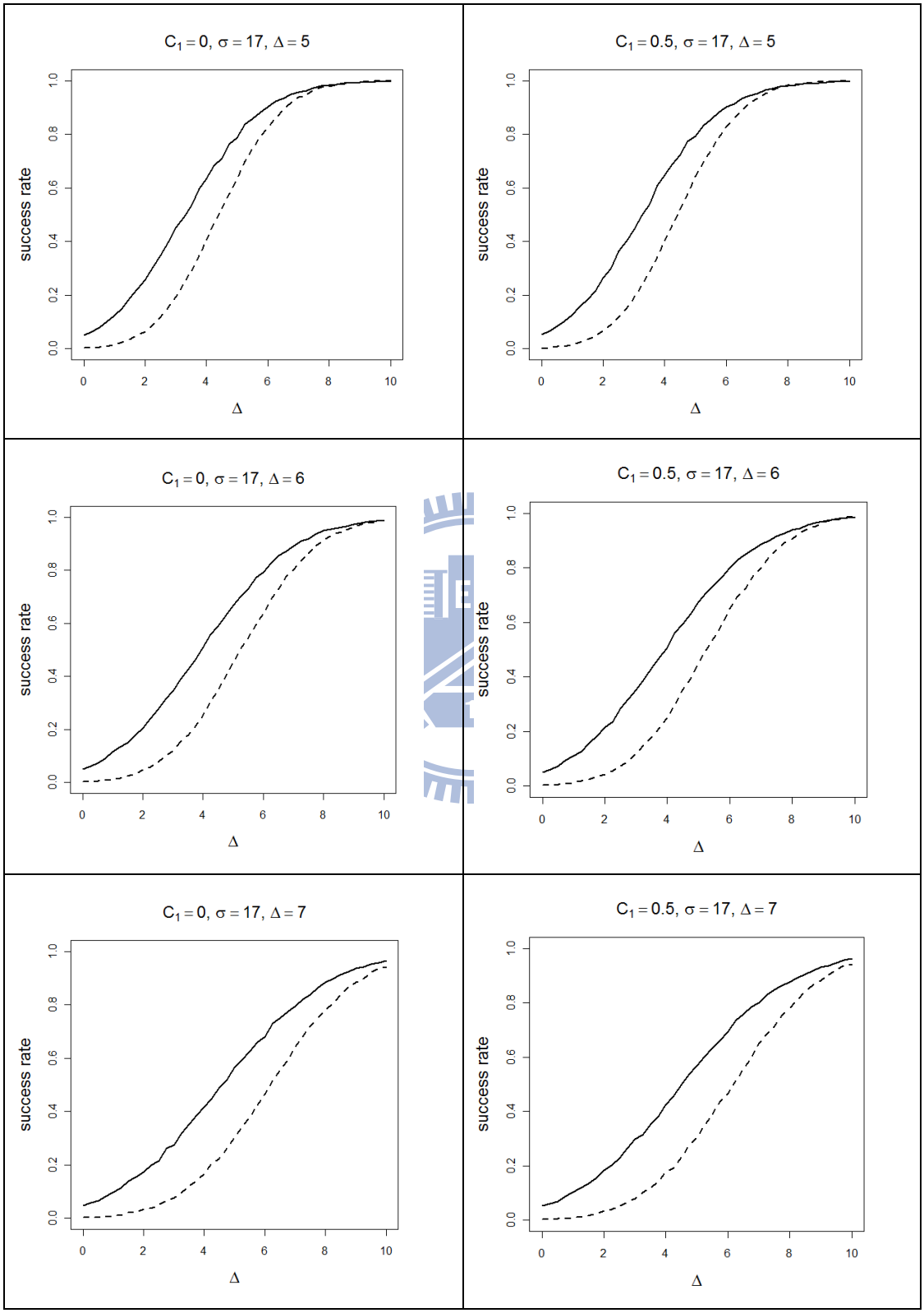
— proposed phase II/III design; --- traditional phase II and phase III trials

Figure2. Simulated success rates for the case of K=1, σ=15, $(K\alpha, 1-\beta)$ = (0.05, 0.8).



— proposed phase II/III design; --- traditional phase II and phase III trials

Figure3. Simulated success rates for the case of K=1, σ=17, $(K\alpha, 1-\beta)$ = (0.05, 0.8).



— proposed phase II/III design; --- traditional phase II and phase III trials

Appendix

Since

$$\overline{\mathbf{Y}}^{\mathrm{II}} = \begin{bmatrix} \overline{Y}_0^{\mathrm{II}} \\ \overline{Y}_1^{\mathrm{II}} \\ \vdots \\ \overline{Y}_K^{\mathrm{II}} \end{bmatrix} \sim N_K \left(\begin{bmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_3 \end{bmatrix}, \frac{\sigma^2}{n_2} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \right) \equiv N_K \left(\boldsymbol{\mu}, \frac{\sigma^2}{n_2} \mathbf{I}_K \right),$$

and from (2), we can derive the joint distribution of $\begin{bmatrix} T_1^{\mathrm{II}} & T_2^{\mathrm{II}} & \cdots & T_K^{\mathrm{II}} \end{bmatrix}^T$:

$$\begin{bmatrix} T_1^{\mathrm{II}} \\ T_2^{\mathrm{II}} \\ \vdots \\ T_K^{\mathrm{II}} \end{bmatrix} = \frac{1}{\sqrt{\frac{2\sigma^2}{n_2}}} \begin{bmatrix} \overline{Y}_1^{\mathrm{II}} - \overline{Y}_0^{\mathrm{II}} \\ \overline{Y}_2^{\mathrm{II}} - \overline{Y}_0^{\mathrm{II}} \\ \vdots \\ \overline{Y}_K^{\mathrm{II}} - \overline{Y}_0^{\mathrm{II}} \end{bmatrix} = \frac{1}{\sqrt{\frac{2\sigma^2}{n_2}}} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \overline{Y}_0^{\mathrm{II}} \\ \overline{Y}_1^{\mathrm{II}} \\ \vdots \\ \overline{Y}_K^{\mathrm{II}} \end{bmatrix} = \mathbf{A}\overline{\mathbf{Y}}^{\mathrm{II}} \sim N_K \left(\mathbf{A}\boldsymbol{\mu}, \frac{\sigma^2}{n_2} \mathbf{A}\mathbf{A}^{\mathrm{T}} \right).$$

We can write

$$N_K \left(\mathbf{A}\boldsymbol{\mu}, \frac{\sigma^2}{n_2} \mathbf{A}\mathbf{A}^{\mathrm{T}} \right)$$

$$\equiv N_K \left(\frac{1}{\sqrt{\frac{2\sigma^2}{n_2}}} \begin{bmatrix} \mu_1 - \mu_0 \\ \mu_2 - \mu_0 \\ \vdots \\ \mu_K - \mu_0 \end{bmatrix}, \frac{\sigma^2}{n_2} \frac{1}{\frac{2\sigma^2}{n_2}} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right)$$

$$\equiv N_K \left(\frac{1}{\sqrt{\frac{2\sigma^2}{n_2}}} \begin{bmatrix} \mu_1 - \mu_0 \\ \mu_2 - \mu_0 \\ \vdots \\ \mu_K - \mu_0 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ 1 & 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 2 \end{bmatrix} \right)$$

$$\equiv N_K \left(\frac{1}{\sqrt{\frac{2\sigma^2}{n_2}}} \begin{bmatrix} \mu_1 - \mu_0 \\ \mu_2 - \mu_0 \\ \vdots \\ \mu_K - \mu_0 \end{bmatrix}, \begin{bmatrix} 1 & 1/2 & 1/2 & \cdots & 1/2 \\ 1/2 & 1 & 1/2 & \cdots & 1/2 \\ 1/2 & 1/2 & 1 & \cdots & 1/2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/2 & 1/2 & 1/2 & \cdots & 1 \end{bmatrix} \right).$$

Under the null hypothesis, $\Delta_i = 0$,

$$\begin{bmatrix} T_1^{\text{II}} \\ T_2^{\text{II}} \\ \vdots \\ T_K^{\text{II}} \end{bmatrix} \overset{\text{H}_0}{\sim} N_K \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1/2 & 1/2 & \cdots & 1/2 \\ 1/2 & 1 & 1/2 & \cdots & 1/2 \\ 1/2 & 1/2 & 1 & \cdots & 1/2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/2 & 1/2 & 1/2 & \cdots & 1 \end{bmatrix} \right) \equiv N_K(\mathbf{0}, \boldsymbol{\Sigma}).$$

From above, the correlation coefficient ρ is $1/2$. The joint probability density function for $\begin{bmatrix} T_1^{\text{II}} & T_2^{\text{II}} & \cdots & T_K^{\text{II}} \end{bmatrix}^T$ is

$$f(x_1,\ldots,x_K) = \frac{1}{(2\pi)^{\frac{K}{2}}\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left\{ -\frac{\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}}{2} \right\}.$$

Under H_0, the probability of keeping on recruiting after the phase II stage for control group and j of the experimental group(s) is

$$p_j = \binom{K}{j} \underbrace{\int_{-\infty}^{C_1} \cdots \int_{-\infty}^{C_1}}_{K-j} \underbrace{\int_{C_1}^{C_2} \cdots \int_{C_1}^{C_2}}_{j} f(x_1,\ldots,x_K)\,dx_1 \ldots dx_K$$

$$= \binom{K}{j} \int_{-\infty}^{\infty} \left\{ \Phi\left(\frac{C_2 - \sqrt{\rho}\,y}{\sqrt{1-\rho}} \right) - \Phi\left(\frac{C_1 - \sqrt{\rho}\,y}{\sqrt{1-\rho}} \right) \right\}^j \Phi\left(\frac{C_1 - \sqrt{\rho}\,y}{\sqrt{1-\rho}} \right)^{K-j} \frac{\exp\left\{ -\frac{y^2}{2} \right\}}{\sqrt{2\pi}} dy,$$

where $j = 1,2,\ldots,K$. The probability of stopping to recruit after the phase II stage is

$$p_0 = \underbrace{\int_{-\infty}^{C_1} \cdots \int_{-\infty}^{C_1}}_{K} f(x_1,\ldots,x_K)\,dx_1 \ldots dx_K + \left(1 - \underbrace{\int_{-\infty}^{C_2} \cdots \int_{-\infty}^{C_2}}_{K} f(x_1,\ldots,x_K)\,dx_1 \ldots dx_K \right)$$

$$= \int_{-\infty}^{\infty} \Phi\left(\frac{C_1 - \sqrt{\rho}\,y}{\sqrt{1-\rho}} \right)^K \frac{\exp\left\{ -\frac{y^2}{2} \right\}}{\sqrt{2\pi}} dy + 1 - \int_{-\infty}^{\infty} \Phi\left(\frac{C_2 - \sqrt{\rho}\,y}{\sqrt{1-\rho}} \right)^K \frac{\exp\left\{ -\frac{y^2}{2} \right\}}{\sqrt{2\pi}} dy.$$

The results of p_0 and p_j are derived by Gupta (1963). And we have

$$p_0 + \sum_{j=1}^{K} p_j = 1.$$

The expected total sample size, $E(N)$, under the null hypothesis, is derived as

$$E(N) = (K+1)n_2 p_0 + \sum_{j=1}^{K} \{(n_3 + n_2)(j+1) + n_2(K-j)\} p_j$$

$$= (K+1)n_2 + \sum_{j=1}^{K} n_3 (j+1) p_j$$

References

1. Bischoff W, and Miller F. Adaptive two-stage test procedures to find the best treatment in clinical trials. *Biometrika.* 2005;92:197-212.
2. Gupta SS. Probability integrals of multivariate normal and multivariate t. *Ann. Math. Statist.* 1963;34:792-828.
3. Maca J, Bhattacharya S, Dragalin V, Gallo P, and Krams M. Adaptive seamless Phase II/III designs—Background, operational aspects, and examples. *Drug Information Journal.* 2006;40:463-473.
4. Posch M, Koenig F, Branson M, Brannath W, Dunger-Baldauf C, and Bauer P. Testing and estimation in flexible group sequential designs with adaptive treatment selection. *Stat Med.* 2005;24:3697-3714.
5. Schaid DJ, Wieand S, and Therneau TM. Optimal two-stage screening designs for survival comparisons. *Biometrika.* 1990;77:507-13.
6. Scher HI, and Heller G.. Picking the Winners in a Sea of Plenty. *Clinical Cancer Research.* 2002;8:400-404.
7. Simon R. Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials.* 1989;10:1-10.
8. Tsou HH, Hsiao CF, Chow SC, and Liu JP. A two-stage design for drug screening trials based on continuous endpoints. *Drug Information Journal.* 2008;42:253-262.