

行政院國家科學委員會專題研究計畫 期末報告

第二期臨床試驗下兩治療組之調適性二階段設計

計畫類別：個別型
計畫編號：NSC 101-2118-M-343-001-
執行期間：101年08月01日至102年07月31日
執行單位：南華大學自然醫學研究所

計畫主持人：陳嘉民

計畫參與人員：碩士班研究生-兼任助理人員：吳芳如
碩士班研究生-兼任助理人員：盧仲軒
大專生-兼任助理人員：李書馨

公開資訊：本計畫可公開查詢

中華民國 102年10月29日

中文摘要：在第二期臨床試驗中，兩治療組隨機試驗的主要目的是要瞭解新療法的有效性與標準療法相比是否具有較佳的療效。基於倫理的考量，設計上通常以兩階段的方式進行，以避免給太多的病人無療效的藥。在樣本數的計算上，研究者必須先給定目標療效反應率差。然而，在一個前導性研究中，由於研究人員對於新、舊治療之研究資訊往往相當缺乏，導致研究人員在決定新、舊治療之目標療效反應率差的時候，常常感到困擾。有鑑於此，本計畫根據 Lin 和 Shih (2004) 所提出之調適性設計概念，發展一個調適性二階段設計，其特點是允許研究人員可以根據第一階段所觀察到之有效反應人數的差異，彈性的選擇對立假設之目標療效反應率差。假如所觀察到呈有效反應之人數，差異夠大，則研究人員傾向於選擇一個樂觀的療效反應率差與較高的檢定力；反之，則選擇一個保守的療效反應率差與較低的檢定力。並在給定的顯著水準與檢定力下，根據大中取小與最佳化兩種準則，找出最合適的試驗設計參數。

中文關鍵詞：兩治療組隨機試驗、調適性二階段設計、試驗設計參數

英文摘要：In a drug development process, the primary goal of randomized phase II trials is to determine whether the new therapy has a higher efficacy than the standard one. For ethical considerations, the design of two-stage clinical trial is commonly used to avoid giving patient an ineffective therapy. The sample size required in the trial is determined by the expected difference in response rates between the two therapies under the specified power. In practice, however, it may be difficult to set the expected difference in the alternative hypothesis, especially when there is no available information about the two therapies in the literature. In view of this, following the idea of Lin and Shih (2004), we develop an adaptive two-stage design which allows the investigators to choose the expected difference with some flexibility. According to the outcome observed at the first stage, the expected difference can be chosen either a larger value with higher power or a smaller value with a lower power at the second stage. Moreover, for the selected significance level and power, the design parameters are derived based on the binomial distribution under Simon's minimax and

optimal criteria.

英文關鍵詞： randomized phase II trials, adaptive two-stage design ; expected sample size ; design parameters

1. 研究背景及目的

一般而言，在新藥研發的過程中，第二期臨床試驗（Phase II clinical trials）主要的目的是要瞭解新療法的有效性與安全性，以決定新療法是否具有較佳的療效，其結果可作為是否進一步進行第三期臨床試驗（Phase III clinical trials）的依據。通常在試驗開始前，研究者應先明確定義何謂有效的治療，而每位受試者在接受新療法後，依據受試者治療後的療效反應，可分為有效與無效兩種，為一個二元反應變數（binary response）。

除此之外，基於倫理與道德議題的考量，針對具生命威脅性之臨床試驗，如癌症臨床試驗，為避免提供無療效的藥給太多受試者，臨床試驗的設計上大多以多階段設計（multi-stage designs）或是群集逐次設計（group sequential design）的方式來進行試驗。例如，Gehan（1961）首先提出兩階段臨床試驗設計的概念；Fleming（1982）則針對多階段設計提供了一個簡易的檢定方法；針對二元反應變數，Simon（1989）建構了一個二階段設計，並直接利用二項分配推導出檢定力函數（power function），根據其所提出的大中取小設計（minimax design）與最佳化設計（optimal design）兩種準則，在給定的顯著水準（significant level）與檢定力（power）下，可決定所需的樣本個數與相對應的檢定臨界值。其後有多位學者根據賽門二階段設計（Simon two-stage design, 1989）進行推廣，例如 Chen（1997）將其推廣至三階段設計（three-stage design），藉以減少試驗過程中所需的受試者；Shuster（2002）則修正賽門二階段設計，使其在第一階段允許停止臨床試驗並宣稱新藥有療效；此外，針對受試者招募不易的情況，Chi 和 Chen（2008）根據賽門二階段設計，結合縮減抽樣程序（curtailed sampling procedure）發展出一種縮減兩階段設計（curtailed two-stage design），此設計一旦有足夠的訊息判定新藥療效與安全性時，即可立即停止招募病人，做成決策，進而有效減少受試者數目，縮短臨床試驗時間，加速進入第三期臨床試驗。上述臨床試驗設計在施行上皆是以單組（single arm）的方式進行，試驗過程中對所有參與臨床試驗的受試者均施予新療法的治療，並藉由參與試驗之受試者的整體療效反應率（response rate）與目前標準治療方式的療效反應率相比較，若整體有效反應的人數超過預設的標準，則表示此新療法比目前標準治療佳。

近年來，在舊藥開發新適應症之初期研究中，隨著治療方式與使用對象的改變，當研究人員欲比較兩種不同的治療方式在新適應症上何者較佳時，往往就直接將其視為兩個獨立的單組試驗進行研究。以 Vansteenkiste et al.（2007）針對易莎平（Ixabepilone）注射劑所進行的研究為例，此注射劑原先是用於乳癌患者，而後將其改用在非小細胞肺癌病人（non-small cell lung cancer），並進行第二期臨床試驗研究中，研究中以曾經使用含鉑化療藥品（Platinum-based Chemotherapy）失敗後的非小細胞肺癌病人為研究對象，在 21 天的治療週期中，研究人員考慮兩種不同的注射方式，其中治療組 A 以單次 3 小時靜脈輸注方式給予 $32\text{mg}/\text{m}^2$ 的易莎平注射劑；治療組 B 則是連續 5 天，每次 1 小時靜脈輸注方式給予 $6\text{mg}/\text{m}^2$

的易莎平注射劑。在病人個數的估算上，便直接以賽門最佳化二階段設計 (Simon optimal two-stage design) 來估算每個治療組所需的樣本個數。在上述易莎平注射劑的研究中，每個治療組在第一階段分別招募 27 位病人，若單組觀察到之有效反應人數大於等於 2 個，即客觀反應率 (objective response rate) 大於 7%，則進入第二階段，再招募 43 位病人。由於賽門最佳二階段設計僅能應用於單組資料與預先設定的療效反應率進行比較，進而在所要求的檢定力下判定此組資料之反應率是否較預先設定的反應率高。然而根據研究目的，研究人員欲瞭解治療組 A 與治療組 B 療效反應率高低時，此種決定樣本個數的方式，可能導致試驗無法達到所要求之檢定力。

因此，當研究人員欲探討兩種不同的治療方式何者較佳時，前述多種單組臨床試驗設計，如賽門二階段設計，便不適用。有鑒於此，針對上述比較兩種新的治療方式之檢定問題，為使其可以滿足臨床試驗中對檢定力的要求，Steinberg 和 Venzon (2002) 提出一個二階段設計，在第一階段，若兩個治療組有效反應人數差異少於預期的標準，則臨床試驗停止在第一階段；反之，則臨床試驗進入第二階段並額外再招募病人，若兩個治療組累積的有效反應人數差距超過臨界值，則拒絕虛無假設，並選擇 A 組為最佳治療組。然而在 Steinberg 和 Venzon (2002) 的研究中，並無控制整體的型一誤差率 (type I error rate) 與型二誤差率 (type II error rate)，而是在對立假設下，控制錯誤挑選最佳治療組的機率，作為挑選臨床設計參數的準則。Jung (2008) 沿用 Steinberg 和 Venzon (2002) 臨床試驗設計方式，並根據兩個獨立的二項分佈推導出檢定力函數，在給定目標療效反應率差下，利用 Simon (1989) 所提出之大中取小設計與最佳化設計兩種準則來計算臨床試驗所需的樣本數與檢定的臨界值。Kepner (2010) 則針對兩個治療組療效反應率的檢定問題，提出了三種不同的二階段檢定程序，考慮三種臨床試驗允許停止在第一階段的規則，並根據第一階段所觀察到的有反應人數差異來決定是否進行第二階段。

除此之外，由於第二期臨床試驗多為前導性研究，對於適當的目標療效反應率 (target response rate) 之選擇，臨床研究人員往往會有不同的看法，造成目標療效反應率設定上的困擾。然而，當研究者根據臨床試驗所要求的檢定力進行樣本數的估算時，目標療效反應率的設定卻是不可或缺的。因此，當研究人員無法明確指定目標療效反應率時，Lin 和 Shih (2004) 提出一個調適性二階段設計 (adaptive two-stage design)，其允許研究人員同時指定一個較高的目標療效反應率與一個較低的目標療效反應率，並依據第一階段的結果來選擇較適合的目標療效反應率，且仍能控制整體的誤判率在一個可容忍的標準下。然而，在一個包含兩個治療組的臨床試驗中，若研究人員無法明確決定兩個治療組臨床上有意義之差異時，Lin 和 Shih (2004) 所提出的調適性二階段設計便不適用，而目前亦尚無合適的臨床試驗設計可供研究者使用，所以針對包含兩個治療組的臨床試驗，發展一個可彈性選擇療效反應率差異的臨床試驗設計，使研究人員對於療效反應

率差的選擇可以更為彈性，確實有其必要性。調適性設計 (adaptive design) 的概念，目前普遍存在於各類型的臨床試驗中，其較為彈性的試驗進行方式，有助於解決試驗中所遭遇之問題，使臨床研究的進行更為順利。

因此，根據 Jung (2008) 所提出兩治療組隨機試驗，本研究擬結合 Lin 和 Shih (2004) 所提出的調適性二階段設計，發展出一個具調適性概念的兩組臨床試驗設計，其特點是允許研究人員可以根據第一階段的試驗結果，彈性的選擇兩個治療組之療效反應率差，並據此推導出相對應的檢定力函數，用以估算試驗所需的病人個數與檢定的臨界值，以確保可以控制整體型一誤差率同時擁有足夠的檢定力。在本文中，第二節將針對 Lin 和 Shih (2004) 所提出的調適性二階段設計與 Jung (2008) 所提出兩治療組隨機試驗作一說明；接下來，第三節將詳述本研究所提出的兩治療組調適性二階段設計 (Adaptive two-stage design for two arms)；最後，對於給定的顯著水準與檢定力，適用於兩治療組調適性二階段設計的樣本個數與臨界值將被呈現於第四節，並與 Jung (2008) 的結果做一比較。

2. 文獻探討

2.1 調適性二階段設計 (Lin 和 Shih, 2004)

由於第二期臨床試驗多為前導性研究，導致研究人員往往很難準確的訂定目標療效反應率。因此，當研究人員無法明確指定目標療效反應率時，Lin 和 Shih (2004) 提出一個調適性二階段設計 (adaptive two-stage design)。調適性二階段設計允許研究人員在試驗前的對立假說設定上，指定一個較低的目標療效反應率 p_1 與一個較高的目標療效反應率 p_2 ，並針對兩個不同的目標療效反應率，分別給予 $1 - \beta_1$ 與 $1 - \beta_2$ 兩種不同的檢定力要求，其中 $\beta_1 > \beta_2$ 。其後，在試驗進行的過程中，研究人員可以依據在第一階段中所觀察到的有效反應人數來決定適合的目標療效反應率與相對應的檢定力，且仍能控制整體的誤判率在一個可容忍的標準下。首先，其檢定之虛無假設與對立假設如下

$$H_0 : p = p_0 \text{ 和 } H_1 : p = p^*$$

其中， p^* 可能為 p_1 或是 p_2 ，其選擇由第一階段所觀察到的有效反應人數來決定。在研究進行的過程中，第一階段將招募 n_1 個病人，且令 x_1 表示 n_1 個病人中呈現有效反應的人數。接下來，依照第一階段所觀察到的有效反應人數，其結果可分為下列三種情況，簡述如下，

- (i) 若 $x_1 \leq s_1$ ，則在第一階段停止試驗，且試驗結果為不拒絕虛無假設。
- (ii) 若 $s_1 < x_1 \leq r_1$ ，則試驗允許進入第二階段，並再繼續招募 $m_2 = m - n_1$ 個病人，並選擇較低的目標反應率，即 $p^* = p_1$ ，與較小的檢定力 $1 - \beta_1$ 。合併兩個階段有效反應的人數，若有效反應的總人數大於臨界值 s ，則拒絕虛無假設。
- (iii) 若 $r_1 < x_1$ ，則試驗亦允許進入第二階段，並再繼續招募 $n_2 = n - n_1$ 個病人，並選擇較高的目標反應率，即 $p^* = p_2$ ，與較大的檢定力 $1 - \beta_2$ 。合

併兩個階段有效反應的人數，若有效反應的總人數大於臨界值 r ，則拒絕虛無假設。

為找出最佳的試驗設計參數， s_1 、 r_1 、 n_1 、 s 、 m 、 r 和 n ，Lin 和 Shih (2004) 推導其不拒絕虛無假設的機率為

$$G(s_1, r_1, n_1, s, m, r, n|p) = B(s_1, n_1|p) + \sum_{x=s_1+1}^{\min(r_1, s)} b(x, n_1|p)B(s-x, m_2|p) + \sum_{x=r_1+1}^{\min(r, n_1)} b(x, n_1|p)B(r-x, n_2|p)$$

此處， $b(x, n|p)$ 與 $B(x, n|p)$ 分別表示試驗次數為 n ，成功機率為 p 的二項分配之機率質量函數與與累積機率質量函數。在給定型一誤差率 α 與檢定力 $1 - \beta_1$ 和 $1 - \beta_2$ 下，試驗設計參數必須滿足下列三條限制式，

$$\begin{aligned} G(s_1, r_1, n_1, s, m, r, n|p = p_0) &\leq \alpha \\ G(s_1, r_1, n_1, s, m, r, n|p = p_1) &\geq 1 - \beta_1 \\ G(s_1, r_1, n_1, s, m, r, n|p = p_2) &\geq 1 - \beta_2 \end{aligned}$$

最後，進一步利用大中取小設計與最佳化設計兩種準則來決定最佳的試驗設計參數 s_1 、 r_1 、 n_1 、 s 、 m 、 r 和 n 。由於，此臨床試驗方式乃是針對單一治療組的情況進行設計，因此並不適合用於兩個治療組療效反應率的比較。

2.2 兩治療組隨機試驗 (Jung, 2008)

針對兩個治療組療效反應率比較的問題，Jung (2008) 提出了一個隨機對照的兩階段設計，簡述如下：首先，假定治療組 A 為預期結果較佳的治療方式 (superior arm A)，其療效反應率以 p_A 表示；治療組 B 為預期結果較差的治療方式 (inferior arm B)，其療效反應率以 p_B 表示。由於研究人員欲了解治療組 A 與治療組 B 相比是否有較佳的療效，因此其檢定假設可寫成

$$H_0 : p_A = p_B = p_0 \text{ 和 } H_1 : (p_A, p_B) = (p_0 + \Delta, p_0)$$

其中， p_0 表示在虛無假設下兩治療組的療效反應率， Δ 表示治療組 A 與治療組 B 之目標療效反應率的差。假設兩治療組在第一階段均招募 n_1 個受試者，令 x_1 和 y_1 分別表示 A 治療組與 B 治療組於第一階段觀察到有效反應的人數；若第二階段允許進行，兩治療組各再招募 n_2 個受試者，令 x_2 和 y_2 分別表示兩組於第二階段所各自觀察到的有效反應人數。試驗過程如下，

- (i) 在第一階段招募的 n_1 個受試者中，若 $x_1 - y_1 \geq a_1$ ，則試驗進入第二階段，並各再增加 n_2 個受試者，否則試驗停止試驗在第一階段並宣稱兩治療組療效無顯著差異；
- (ii) 當第二階段允許進行時，合併兩個階段所觀察到的有效反應人數，若 $x_1 + x_2 - y_1 - y_2 \geq a$ ，則拒絕虛無假設並宣稱治療組 A 優於治療組 B；反之，則不拒絕虛無假設並宣稱兩治療組療效無顯著差異。

根據上述的試驗過程，Jung (2008) 根據兩個獨立的二項分配，建構出精確的檢定力函數 (power function)，用以推導試驗過程中所需的最佳化試驗設計參數， a_1 、 n_1 、 a 和 n 。其檢定力函數如下，

$$G(a_1, n_1, a, n) = \sum_{k_1=a_1}^{n_1} \sum_{y_1=\max(0, -k_1)}^{n_1-\max(0, k_1)} \sum_{k_2=a-k_1}^{n_2} \sum_{y_2=\max(0, -k_2)}^{n_2-\max(0, k_2)} b(y_1, n_1; p_B) b(k_1 + y_1, n_1; p_A) \\ \times b(y_2, n_2; p_B) b(k_2 + y_2, n_2; p_A)$$

此處， $b(x, n; p)$ 表示試驗次數為 n ，成功機率為 p 的二項分配。在給定的型一誤差率與檢定力下，試驗設計須滿足

$$G(a_1, n_1, a, n | p_A = p_0, p_B = p_0) \leq \alpha$$

與

$$G(a_1, n_1, a, n | p_A = p_0 + \Delta, p_B = p_0) \geq 1 - \beta$$

的要求。由於滿足上述限制是要求的試驗設計參數並不唯一，因此 Jung (2008) 進一步利用 Simon (1989) 所提出之大中取小設計與最佳化設計兩種準則來決定最佳化的受試者個數與檢定的臨界值。

3. 研究方法

針對兩治療組療效比較的問題，本計畫擬將 Lin 和 Shih (2004) 所提出的調適性設計之概念應用到 Jung (2008) 的二階段隨機對照臨床試驗設計上，發展適用於兩個治療組之調適性二階段設計。首先，假定治療組 A 為預期結果較佳的治療方式，其療效反應率以 p_A 表示；治療組 B 為預期結果較差的治療方式，其療效反應率以 p_B 表示。在檢定之虛無假設與對立假設的設性上，沿用 Jung (2008) 設定方式，將之寫成下列形式

$$H_0 : p_A = p_B = p_0 \text{ 和 } H_1 : (p_A, p_B) = (p_0 + \Delta^*, p_0)$$

其中，在虛無假設下兩治療組的療效反應率相同，以 p_0 表示； Δ^* 則表示對立假設下兩治療組之目標療效反應率的差。首先，令 x_A 與 x_B 分別表示 A 治療組與 B 治療組於第一階段觀察到有效反應的人數，而 $D_1 = x_A - x_B$ 表示第一階段兩組呈有效反應人數的差異。若第二階段允許進行，令 y_A 與 y_B 則分別表示 A 治療組與 B 治療組於第二階段觀察到有效反應的人數，而 $D_2 = y_A - y_B$ 表示第二階段兩組呈有效反應人數的差異。此外，依據調適性設計之概念，為使研究者在目標療效反應率的設定上更有彈性，本計畫所提出之試驗設計允許研究人員同時設定兩個不同的目標反應率差 Δ_1 和 Δ_2 並搭配不同的檢定力 $1 - \beta_1$ 和 $1 - \beta_2$ 。試驗流程簡述如下：首先，假設在第一階段每組招募的病人個數為 n_1 ，

- (i) 若 $D_1 < r_1$ ，則試驗停止在第一階段，且不拒絕虛無假設；
- (ii) 若 $r_1 \leq D_1 < r_2$ ，則試驗允許進入第二階段。此時，每個治療組再持續招募 $m_2 = m - n_1$ 個病人，並選擇較低的目標反應率差，即 $\Delta^* = \Delta_1$ ，與設定較低之檢定力 $1 - \beta_1$ 。在第二階段結束後，合併兩個階段有效反應的人數，若 $D_1 + D_2 \geq r_m$ ，則拒絕虛無假設，並宣稱治療組 A 優於治療組 B；反之，則宣稱兩治療組療效無顯著差異。
- (iii) 若 $D_1 \geq r_2$ ，則試驗亦允許進入第二階段。此時，每個治療組則再持續招募 $n_2 = n - n_1$ 個病人，並選擇較高的目標反應率差，即 $\Delta^* = \Delta_2$ ，與設定較高之檢定力 $1 - \beta_2$ 。在第二階段結束後，合併兩個階段有效反應

的人數，若 $D_1 + D_2 \geq r_n$ ，則拒絕虛無假設，並宣稱治療組 A 優於治療組 B；反之，則宣稱兩治療組療效無顯著差異。

此處， $\Delta_1 < \Delta_2$ 且 $\beta_1 > \beta_2$ 。值得注意的是，當第二階段允許進行時，每組增加的人數將由第一階段所觀察到之兩組有效反應人數差來決定，而對立假設的選擇與檢定力的決定亦是根據第一階段的結果。

為計算臨床試驗中所需之樣本數與檢定之臨界值， $(r_1, r_2, n_1, r_m, m, r_n, n)$ ，首先推導出 D_1 的機率質量函數為

$$P(D_1 = d_1) = \sum_{x_A = \max(0, d_1)}^{\min(n_1, n_1 + d_1)} b(x_A, n_1; p) b(x_A - d_1, n_1; p)$$

此處， $d_1 = -n_1, \dots, n_1$ 。而在給定 D_1 下， D_2 的條件機率質量函數則為

$$P(D_2 = d_2 | D_1 = d_1) = \begin{cases} \sum_{y_A = \max(0, d_2)}^{\min(m_2, m_2 + d_2)} b(y_A, m_2; p) b(y_A - d_2, m_2; p), & r_1 \leq d_1 < r_2 \\ \sum_{y_A = \max(0, d_2)}^{\min(n_2, n_2 + d_2)} b(y_A, n_2; p) b(y_A - d_2, n_2; p), & r_2 \leq d_1 \end{cases}$$

此處，若 $r_1 \leq d_1 < r_2$ ，則 $d_1 = -m_2, \dots, m_2$ ；反之，若 $r_2 \leq d_1$ ，則 $d_1 = -n_2, \dots, n_2$ 。進一步依據上述檢定程序，可推導出檢定力函數為

$$\Theta(r_1, r_2, n_1, r_m, m, r_n, n | p_A, p_B) = \Theta(r_1, r_2, n_1, r_m, m | p_A, p_B) + \Theta(r_1, r_2, n_1, r_n, n | p_A, p_B)$$

其中

$$\Theta_m(r_1, r_2, n_1, r_m, m_2) = \sum_{d_1 = r_1}^{r_2 - 1} P(D_1 = d_1) \left\{ \sum_{d_2 = \max(-m_2, r_m - d_1)}^{m_2} P(D_2 = d_2 | D_1 = d_1) \right\}$$

表示當第二階段允許進行時，每組增加的人數為 m_2 且最後拒絕虛無假設的機率；而

$$\Theta_n(r_1, r_2, n_1, r_n, n_2) = \sum_{d_1 = r_2}^{n_1} P(D_1 = d_1) \left\{ \sum_{d_2 = \max(-n_2, r_n - d_1)}^{n_2} P(D_2 = d_2 | D_1 = d_1) \right\}$$

表示當第二階段允許進行時，每組增加的人數為 n_2 且最後拒絕虛無假設的機率。

在給定的型一誤差率 α 與檢定力 $1 - \beta_1$ 和 $1 - \beta_2$ 下，試驗設計須滿足下列三條限制式

$$\Theta(r_1, r_2, n_1, r_m, m_2, r_n, n_2 | p_A = p_0, p_B = p_0) \leq \alpha$$

$$\Theta(r_1, r_2, n_1, r_m, m_2, r_n, n_2 | p_A = p_0 + \Delta_1, p_B = p_0) \geq 1 - \beta_1$$

$$\Theta(r_1, r_2, n_1, r_m, m_2, r_n, n_2 | p_A = p_0 + \Delta_2, p_B = p_0) \geq 1 - \beta_2$$

由於滿足上述限制是要求的試驗設計參數並不唯一，因此本計畫根據 Simon (1989) 所提出之大中取小設計 (minimax design) 與最佳化設計 (optimal design) 兩種選擇最佳化試驗設計參數的準則，進一步確定臨床試驗中所需之樣本數與檢定之臨界值。大中取小設計的概念即是由所有滿足上述三個限制式的試驗設計參數中，挑選 $\max(m, n)$ 最小的試驗設計參數，然後再從中挑選使得期望樣本個數

為最小的試驗設計參數作為大中取小設計下的最佳解；反之，最佳化設計的概念則是直接由所有滿足上述三個限制式的試驗設計參數中，挑選使得期望樣本個數為最小的試驗設計參數作為最佳化設計下的最佳解。不論是採用大中取小設計或是最佳化設計的挑選參數準則，在給定試驗設計參數下，每個治療組的期望樣本個數為

$$E(N|p) = n_1 + \Phi_m \times m_2 + \Phi_n \times n_2$$

其中 $\Phi_m = \sum_{d_1=r_1}^{r_2-1} P(D_1 = d_1)$ 表示試驗進入第二階段且每組在增加 m_2 人的機率，而 $\Phi_n = \sum_{d_1=r_2}^{n_1} P(D_1 = d_1)$ 則表示試驗進入第二階段且每組在增加 n_2 人的機率。根據上述兩種挑選參數的準則所獲得的最佳解，在試驗過程中均可控制試驗整體的型一誤差率在 α 之下，同時在兩個不同的目標反應率差的設定下亦可以達到所要求的檢定力，其結果將於下一節呈現。

4. 結果與討論

為了與 Jung (2008) 的結果相比較，因此將整體的型一誤差率設定為 $\alpha = 0.2$ ，而目標療效反應率差則設定為 $\Delta_1 = 0.15$ 與 $\Delta_2 = 0.2$ 。當對立假設之 $\Delta^* = \Delta_1$ 時，相對應的型二誤差率則設定為 $\beta_1 = 0.2$ ；當對立假設之 $\Delta^* = \Delta_2$ 時，相對應的型二誤差率則設定為 $\beta_2 = 0.15$ 。Table I 和 Table II 分別表示以大中取小設計準則和最佳化設計準則所挑選出來的最佳解。由 Table I 可以發現，在中取小的參數挑選方法下， m 和 n 的差異並不大，在七種不同的 p_0 設定下，僅在 $p_0 = 0.1$ (此時 $m = 30, n = 29$) 與 $p_0 = 0.6$ (此時 $m = 55, n = 58$) 時，出現差異，其餘情況 m 和 n 均相同，此結果與 Lin 和 Shih (2004) 在單組調適性二階段設計下，利用大中取小準則決定參數所得到的結果類似；反之，如果是採用最佳化設計準則來挑選最佳解，則 m 和 n 的差異較大。此外，由 Table I 和 Table II 亦可以發現，當對立假設下之目標療效反應率差設定為 $\Delta^* = \Delta_2$ 時，雖然給定的相對應檢定力為 0.85，然而實際的檢定力卻都達到 0.9 以上，明顯高出 0.85 許多。Table III 為 Jung (2008) 與本計畫所提出雙組調適性二階段設計在中取小的參數挑選方法下之比較結果；Table IV 則為在最佳化設計的參數挑選方法下之比較結果。以 Table III 來說，在中取小準則下，調適性二階段設計中所需的最大樣本個數會高於 Jung (2008) 設定 $p_A - p_B = 0.2$ 與檢定力 0.85 下所需的樣本數，會略小於設定 $p_A - p_B = 0.15$ 與檢定力 0.80 下所計算的樣本數。

本質上，調適性二階段設計的概念其實就是在兩個不同的目標療效反應率差的設定下，使其均可以達到預設之檢定力。例如在中取小設計中，當 $p_0 = 0.3$ 時，試驗設計參數為 $r_1 = 0, r_2 = 4, n_1 = 43, r_m = 5, m = 56, r_n = 4$ 和 $n = 56$ 。不論 $\Delta^* = \Delta_1$ (即 $p_A = 0.45$ 與 $p_B = 0.3$) 或是 $\Delta^* = \Delta_2$ (即 $p_A = 0.5$ 與 $p_B = 0.3$) 時，調適性二階段設計的檢定力分別為 0.8005 與 0.9167，兩者實際之檢定力均可達到預設的 0.80 與 0.85 的水準。反觀在 Jung (2008) 的兩階段設計中，若採用目標療效反應率差為 0.2、檢定力為 0.85 的設定，則試驗設計參數為 $a_1 = -1$ 、

$n_1 = 26$ 、 $a = 4$ 與 $n = 40$ ，雖然樣本總數 40 小於調適性二階段設計之最大樣本個數 56，但若真正的目標療效反應率差為 0.15 而不是一開始所選擇之 0.2 時，其實際之檢定力就僅剩 0.7185。反之，若採用目標療效反應率差為 0.15、檢定力為 0.80 的設定，試驗設計參數為 $a_1 = -1$ 、 $n_1 = 37$ 、 $a = 5$ 與 $n = 60$ ，雖然不論真正的目標療效反應率差是 0.15 還是 0.2，此組試驗設計參數的實際檢定力為 0.8003 與 0.9194，均可達到預設的 0.80 與 0.85 的水準，但是樣本數卻較調適性二階段設計之最大樣本個數 56 大。因此，整體來說，本計畫所提出的調適性二階段設計可以同時在不同的對立假設下，取得一個平衡點，此特點將有助於執行臨床試驗者在目標反應率差異的選擇更有彈性。

Table I. Adaptive minimax two-stage design for two arms with $\Delta_1 = 0.15$, $\Delta_2 = 0.2$, $\alpha = 0.2$, $\beta_1 = 0.2$ and $\beta_2 = 0.15$.

p_0	n_1	r_1	r_2	m	r_m	n	r_n	$\max(m, n)$	α^*	$1 - \beta_1^*$	$1 - \beta_2^*$	$p_A = p_B = p_0$		
												$E(N)$	Φ_m	Φ_n
0.1	22	0	2	30	3	29	2	30	0.1985	0.8010	0.9040	26.5960	0.3820	0.2200
0.2	35	-1	2	45	3	45	4	45	0.1995	0.8013	0.9126	41.7369	0.3474	0.3263
0.3	43	0	4	56	5	56	4	56	0.1999	0.8005	0.9167	50.1091	0.3420	0.2049
0.4	48	-3	1	61	6	61	5	61	0.2000	0.8000	0.9200	57.9712	0.3085	0.4585
0.5	45	-1	2	62	6	62	5	62	0.1997	0.8004	0.9239	55.6077	0.2480	0.3760
0.6	35	-1	1	55	4	58	5	58	0.1998	0.8004	0.9286	49.2087	0.1912	0.4515
0.7	29	0	5	48	4	48	5	48	0.1980	0.8006	0.9377	39.5831	0.4589	0.0981

Table II. Adaptive optimal two-stage design for two arms with $\Delta_1 = 0.15$, $\Delta_2 = 0.2$, $\alpha = 0.2$, $\beta_1 = 0.2$ and $\beta_2 = 0.15$.

p_0	n_1	r_1	r_2	m	r_m	n	r_n	$\max(m, n)$	α^*	$1 - \beta_1^*$	$1 - \beta_2^*$	$p_A = p_B = p_0$		
												$E(N)$	Φ_m	Φ_n
0.1	14	0	2	28	2	42	3	42	0.1985	0.8005	0.9001	25.1091	0.3703	0.4658
0.2	22	0	3	51	4	46	3	51	0.1998	0.8000	0.9078	37.8253	0.4247	0.4038
0.3	28	0	2	49	3	68	5	68	0.1996	0.8006	0.9137	46.0025	0.4420	0.2273
0.4	29	0	2	56	4	73	5	73	0.1996	0.8001	0.9156	49.7849	0.4467	0.2093
0.5	28	0	1	60	4	69	5	69	0.1999	0.8002	0.9195	49.7207	0.4469	0.1061
0.6	25	0	3	54	4	73	5	73	0.1990	0.8001	0.9238	45.6334	0.4427	0.3220
0.7	20	0	1	34	3	55	4	55	0.1999	0.8001	0.9323	37.0201	0.4314	0.1371

Table III. Comparisons of required sample size and expected sample size for adaptive minimax two-stage design and Jung's minimax two-stage design under $\Delta_1 = 0.15$, $\Delta_2 = 0.2$, $\alpha = 0.2$, $\beta_1 = 0.2$ and $\beta_2 = 0.15$.

p_0	Jung's minimax design								adaptive minimax design				
	$p_A - p_B = 0.15$				$p_A - p_B = 0.2$				$\max(m, n)$	$E(N)$	α^*	$1 - \beta_1^*$	$1 - \beta_2^*$
	N	$E(N)$	α^*	$1 - \beta^*$	N	$E(N)$	α^*	$1 - \beta^*$					
0.1	34	30.75	0.1524	0.8004	27	19.50	0.1981	0.8500	30	26.60	0.1985	0.8010	0.9040
0.2	48	42.10	0.1835	0.8005	36	30.46	0.1990	0.8557	45	41.74	0.1995	0.8013	0.9126
0.3	60	51.91	0.1825	0.8002	40	35.45	0.1949	0.8503	56	50.11	0.1999	0.8005	0.9167
0.4	62	52.35	0.1996	0.8001	47	37.27	0.1968	0.8504	61	57.97	0.2000	0.8000	0.9200
0.5	64	50.88	0.1998	0.8006	47	36.63	0.1973	0.8515	62	55.61	0.1997	0.8004	0.9239
0.6	58	52.42	0.1961	0.8001	39	32.90	0.1991	0.8511	58	49.21	0.1998	0.8004	0.9286
0.7	49	38.44	0.1987	0.8001	34	30.54	0.1763	0.8500	48	39.58	0.1980	0.8006	0.9377

Table IV. Comparisons of required sample size and expected sample size for adaptive optimal two-stage design and Jung's optimal two-stage design under $\Delta_1 = 0.15$, $\Delta_2 = 0.2$, $\alpha = 0.2$, $\beta_1 = 0.2$ and $\beta_2 = 0.15$.

p_0	Jung's optimal design								adaptive optimal design				
	$p_A - p_B = 0.15$				$p_A - p_B = 0.2$				$\max(m, n)$	$E(N)$	α^*	$1 - \beta_1^*$	$1 - \beta_2^*$
	N	$E(N)$	α^*	$1 - \beta^*$	N	$E(N)$	α^*	$1 - \beta^*$					
0.1	35	28.95	0.1980	0.8073	27	19.50	0.1981	0.8500	42	25.11	0.1985	0.8005	0.9001
0.2	53	40.21	0.1811	0.8002	40	28.42	0.1988	0.8523	51	37.83	0.1998	0.8000	0.9078
0.3	68	48.61	0.1976	0.8010	45	33.83	0.1909	0.8504	68	46.00	0.1996	0.8006	0.9137
0.4	69	51.13	0.1942	0.8004	49	35.62	0.1995	0.8516	73	49.78	0.1996	0.8001	0.9156
0.5	67	50.40	0.1973	0.8000	50	34.92	0.1984	0.8503	69	49.72	0.1999	0.8002	0.9195
0.6	63	47.81	0.1887	0.8003	41	31.05	0.1959	0.8500	73	45.63	0.1990	0.8001	0.9238
0.7	51	37.25	0.1984	0.8005	42	25.71	0.1978	0.8500	55	37.02	0.1999	0.8001	0.9323

Reference

1. Chen, T. T. (1997). Optimal three-stage designs for phase II cancer clinical trials. *Statistics in Medicine*, 16:2701-2711.
2. Chi, Y. C. and Chen, C. M. (2008). Curtailed two-stage designs in phase II clinical trials. *Statistics in Medicine*, 27:6175-6189.
3. Fleming, T. R. (1982). One-sample multiple testing procedure for phase II clinical trials. *Biometrics*, 38:143-151.
4. Gehan, E. A. (1961). The determination of the number of patients required in a preliminary and follow-up trial of a new chemotherapeutic agent. *Journal of Chronic Diseases*, 13:346-353.
5. Jung, S. H. (2008). Randomized phase II trials with a prospective control. *Statistics in Medicine*, 27:568-583.
6. Kepner, J. L. (2010). On group sequential designs comparing two binomial proportions. *Journal of Biopharmaceutical Statistics*, 20:145-159.
7. Lin, Y. and Shih W. J. (2004). Adaptive two-stage designs for single-arm phase IIA cancer clinical trials. *Biometrics*, 60:482-490.
8. Shuster, J. (2002). Optimal two-stage designs for single arm phase II cancer trials. *Journal of Biopharmaceutical Statistics*, 12:39-51.
9. Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials*, 10:1-10.
10. Steinberg, S. M. and Venzon, D. J. (2002). Early selection in a randomized phase II clinical trial. *Statistics in Medicine*, 21:1711-1726.
11. Vansteenkiste, J., Lara, PN. Jr., Le Chevalier, T., et al. (2007). Phase II clinical trial of the Epothilone B Analog, Ixabepilone, in patients with Non-Small-Cell Lung Cancer whose tumors have failed first-line Platinum-Based chemotherapy. *Journal of Clinical Oncology*, 25:3448-3455.

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/28

國科會補助計畫	計畫名稱: 第二期臨床試驗下兩治療組之調適性二階段設計
	計畫主持人: 陳嘉民
	計畫編號: 101-2118-M-343-001- 學門領域: 臨床試驗與製藥統計
無研發成果推廣資料	

101 年度專題研究計畫研究成果彙整表

計畫主持人：陳嘉民		計畫編號：101-2118-M-343-001-					
計畫名稱：第二期臨床試驗下兩治療組之調適性二階段設計							
成果項目		量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p>其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	「無」
--	-----

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

調適性設計（adaptive design）的概念，目前普遍存在於各類型的臨床試驗中，其較為彈性的試驗進行方式，有助於解決臨床試驗中所遭遇之問題，使臨床研究的進行更為順利。本計畫所發展之調適性二階段設計乃根據 Lin 和 Shih (2004) 所提出之調適性二階段設計的概念，將其從原本僅適用於單一治療組的臨床試驗設計，推廣至可用於兩治療組療效比較之臨床試驗設計。本計畫所發展之調適性二階段設計，其特點是允許研究人員可以根據第一階段的試驗結果，彈性的選擇兩個治療組之目標療效反應率差，並據此推導出相對應的檢定力函數，用以估算試驗所需的病人個數與檢定的臨界值，同時確保整體型一誤差率可以控制在一定的水準同時擁有足夠的檢定力。