

# 科技部補助專題研究計畫成果報告 期末報告

## 兩相依二元反應變項之縮減二階段隨機設計

計畫類別：個別型計畫  
計畫編號：MOST 105-2118-M-343-001-  
執行期間：105年08月01日至106年07月31日  
執行單位：南華大學自然生物科技學系

計畫主持人：陳嘉民

計畫參與人員：碩士班研究生-兼任助理：李怡慧  
大專生-兼任助理：廖惠萱  
大專生-兼任助理：鄭如妤

中華民國 106 年 10 月 26 日

中文摘要：在藥物研發的過程中，第二期臨床試驗主要目的是進行新藥的有效性與安全性之評估，以確認是否具有足夠的證據來支持進行第三期臨床試驗。然而，由於在第一期臨床試驗中，新藥的最大耐受劑量的估計可能並不十分準確，使得在第二期臨床試驗中所使用的劑量並不恰當，導致出現過度的毒性反應。因此，在同時考慮新藥有效性與安全性的情況下，本研究以兩相依二元反應變項發展一個雙變量二階段隨機對照設計，並以縮減抽樣程序來建構雙變量縮減二階段隨機對照設計，以加速臨床試驗的進行。相較於雙變量二階段隨機對照設計，結果顯示縮減二階段隨機對照設計可以有效減少期望樣本個數。

中文關鍵詞：兩相依二元反應變項、二階段隨機設計、縮減抽樣程序、試驗設計參數、縮減二階段隨機設計

英文摘要：In a drug development process, an important goal of phase II clinical trial is to evaluate the efficacy and safety of a new drug and confirm whether the evidence is enough to enter the phase III clinical trial. However, the maximum tolerated dose of the new drug might not always be accurately estimated in the phase I clinical trial. As a result, the recommended dose could lead to excessive toxicity in the subsequent phase II clinical trial. In order to simultaneously evaluate the efficacy and safety of new drugs, this study proposed a two-stage randomized design with two dependent binary endpoints. In order to speed up the trial process, the curtailed sampling procedure is applied in the bivariate two-stage randomized design to develop the bivariate curtailed two-stage randomized design. To compare the expected sample sizes of the bivariate two-stage randomized design with and without the implementation of curtailment, the expected sample sizes of bivariate curtailed two-stage randomized design are consistently smaller than that of bivariate two-stage randomized design under each setting.

英文關鍵詞：two dependent binary endpoints; two-stage randomized design; curtailed sampling procedure; curtailed two-stage randomized design

## 一、 前言

在新藥研發的漫長過程中，在第一期臨床試驗(Phase I clinical trials)中，主要目的是尋找新藥的最大耐受劑量(maximum tolerated dose, MTD)，以利於後續第二期臨床試驗(Phase II clinical trials)的進行。而當新藥研發允許進入第二期臨床試驗時，其主要目的則是依據第一期臨床試驗的結果，針對新藥的有效性進行初步評估，以確認是否具有足夠的證據來支持進行第三期臨床試驗(Phase III clinical trials)。此外，由於第二期臨床試驗是屬於早期臨床試驗(early phase clinical trial)，基於倫理與道德議題的考量，減少不必要的風險，往往希望可以較少的受試者來完成試驗，因此試驗進行的方式通常是以多階段設計(multi-stage designs)或是群集逐次設計(group sequential design)的方式進行，以達到受試者保護的原則。典型的第二期癌症臨床試驗，可區分為單組設計與雙組設計兩大類。例如，Gehan (1961)、Fleming (1982)、Simon (1989)、Chen (1997)和 Shuster (2002)等學者提出了多種的單組試驗設計，其中賽門二階段設計(Simon's two-stage design, 1989)更是被廣泛的應用在第二期癌症臨床試驗，其特點是當試驗結果不如預期時，試驗可提前停止在第一階段，同時宣告新藥無療效。此外，為加速臨床試驗的進行，亦有學者開始嘗試應用縮減設計(curtailed design)來減少試驗所需的受試者平均人數，例如 Ayanlowo 和 Redden (2007)和 Kunz 和 Kieser (2012)分別提出兩種不同的隨機縮減二階段設計，然而使用上其檢定之臨界值會隨著樣本數的累積而改變；Chi 和 Chen (2008) 則是根據縮減抽樣程序(curtailed sampling procedure)的概念，提出另一種非隨機縮減二階段設計，其特點是檢定臨界值並不隨著受試者人數的增加而改變，在使用上較為方便。然而此類單組設計的試驗，需與一歷史對照組 (historical control) 為比較，往往會面臨兩者所招募之受試者存在著明顯的異質性，亦或是出現療效指標評估標準不一致等問題，進而造成比較上的偏差。

近年來，為解決歷史對照組之受試者與目前的臨床試驗所招募受試者之間可能存在著異質性(heterogeneous)，許多學者紛紛針對兩組療效反應率比較的問題，提出不同的二階段設計。例如 Jung (2008)以賽門二階段設計為基礎，提出了隨機分派的二階段設計，並根據兩個獨立的二項分配(binomial distribution)建構檢定力函數(power function)，並利用大中取小準則(minimax criteria)與最佳化準則(optimal criteria) 來決定試驗設計參數。Kepner (2010)則提出三種不同的二階段設計，其停止在第一階段的規則分別是僅考慮早期有效、僅考慮早期無效與同時考慮早期有效和無效三種。Hou et al. (2013)在不考慮對照組的情況下，針對兩個治療組的比較問題，提出一個隨機分派的二階段設計。Shan et al. (2013)與 Jung 和 Sargent (2014)則是分別利用 Barnard's exact test 與 Fisher's exact test 的方法來建構雙組比較的二階段設計，並進一步計算在此隨機分派二階段設計中所需的樣本數。Cellamare 和 Sambucini (2015) 則是利用貝氏預測法 (Bayesian predictive approach) 發展出一個隨機分派的二階段設計。在前述的臨床試驗方法中，不論是單組或是雙組設計，在臨床試驗上均是採二階段設計的概念，而在療效反應的評估上，僅關注新藥的療效反應，而忽略安全性的問題，亦或是直接將新藥有效性與安全性同時評估後採單變量二元反應變項來進行檢定。

然而，由於在第一期臨床試驗中，通常是小規模的研究，試驗中所使用的樣本數通常很少，因

此新藥的最大耐受劑量可能無法很清楚的確定，導致推薦用於第二期臨床試驗的劑量，可能具有過度的毒性，使得受試者出現毒性反應的比率比預期高得多。因此，近年來有部分學者認為在第二期臨床試驗中，除了進行新藥的有效性探索之外，亦應同時進行安全性的評估，對於出現毒性反應的受試者人數亦應該同時進行監控與評估，以避免出現過多毒性反應的受試者人數，進而發展多種可同時檢測新藥有效性與安全性的試驗。例如，在單組設計的第二期臨床試驗中，Bryant 和 Day (1995) 在賽門二階段設計中的決策規則中合併毒性反應的影響，同時監控試驗過程中毒性反應人數的結果，並探討療效反應率與毒性反應率在不同相關程度下對試驗結果的影響。Conaway 和 Petroni (1995) 則是用一個雙變量二元反應變項來同時描述每一個接受治療的受試者的療效反應與毒性反應的結果，並使用交集-聯集檢定 (Intersection-Union tests, IUT) 來同時檢定療效反應率與毒性反應率，然而對於檢定型 I 誤差的設定則是採取療效反應率與毒性反應率聯合設定的方式，無法依照療效指標的重要性給予不同的權重。因此，為了讓研究者在檢定過程中，對於療效反應率與毒性反應率的檢定型 I 誤差設定上更具彈性，Jin (2007) 採用與 Conaway 和 Petroni (1995) 相同的試驗設計，但針對型 I 誤差的設定則改採療效反應率與毒性反應率分離的邊際設定法，允許研究者在療效反應率與毒性反應率的檢定上，可依照需求分開設定不同的療效反應率與毒性反應率之檢定型 I 誤差，使其在使用上更具彈性。Ray and Rai (2011, 2012) 則是以賽門二階段設計為基礎，同時合併連續毒性監測 (continuous toxicity monitoring) 的方式，在每個病人在接受治療後去監測毒性反應人數，當出現毒性反應的人數過高，則立即停止試驗。Chen 和 Chi (2012) 將縮減抽樣設計進一步推廣，提出一個適用於兩相依二元反應變項的縮減二階段設計。Carsten 和 Chen (2015) 則是根據 Chi 和 Chen (2008) 的結果，將其應用兩組的比較上，然而其結論僅適用於配對隨機分派的雙組臨床試驗。反之，在雙組設計的第二期臨床試驗中，在同時考慮兩組療效反應率差異與毒性反應率差異的試驗設計，相關研究目前並不多，Bersimis et al. (2015) 考慮以雙變量的二元反應變項，提出一個合併安全性監測的單階段設計，並根據馬可夫鏈 (Markov's chain) 來計算所需的樣本數與檢定的臨界點，其特點是在試驗過程中，一旦研究者觀察到足夠的事件發生，研究者可提前停止試驗，雖然此設計同時考慮有效性與安全性，且具備縮減設計 (curtailed design) 提前停止試驗的優點，然而卻僅能在兩邊個數一樣的情況下方可比較，且僅在兩組呈現顯著差異時才有較小的期望樣本個數。

綜合上述相關文獻的探討，雖然已有多位學者提出在試驗過程中應合併考慮安全性的問題，並提出多種臨床試驗設計或是有關安全性的評估方法，然而多是針對單組設計的第二期臨床試驗。近年來，為解決受試者間的異質性的問題，雖然有多位學者提出以雙組設計來進行試驗，然而這些二階段隨機對照設計多僅考慮新藥有效性的評估，對於如何同時進行新藥安全性的評估，著墨甚少。此外，當研究者考慮以二階段隨機對照設計的方式進行臨床試驗時，另一個衍生出的問題是試驗過程中所需的受試者人數將比單組設計下的臨床試驗還要多。因此，本計畫以雙變量的二元反應變項為基礎，在同時考慮有效性與安全性的情況下，發展出一個適用於隨機對照臨床試驗之雙變量二階段設計，並進一步將縮減抽樣程序應用於此雙變量二階段設計，此種作法除了可以同時評估新藥的有效性與安全性外，亦可在未完成預定招募之受試者人數之前，便能根據早期試驗之結果做成決策，加快臨床試驗的

進行。

## 二、 研究目的

在第二期臨床試驗下的二階段隨機對照設計中，研究者為了控制型一誤差率同時具備足夠的檢定力，因此在試驗進行過程，基於檢定力的考量，對於受試者人數有一定的要求，往往必須先計算並招募到預定的受試者人數後，才能決定檢定結果。然而，一方面，在隨機對照的臨床試驗設計下，由於同時需要實驗組與對照組，使得在試驗過程中所需招募的受試者人數比單組臨床試驗設計還要多；另一方面，在第二期臨床試驗中為避免受試者異質性，對於受試者的納入條件設定較為嚴格，容易發生受試者招募不易的情況，以致必須花費更長的時間來完成臨床試驗。此外，由於在癌症新藥的第二期臨床試驗中，新藥劑量的決定是根據第一期臨床試驗中所決定的最大耐受劑量。然而，在第一期臨床試驗中，由於試驗過程中所使用的受試者人數通常不多，導致最大耐受劑量的估算容易產生偏差，進而使得第二期臨床試驗進行時，容易產生受試者安全性的疑慮。因此基於受試者安全性的考量，於第二期臨床試驗設計中增加安全性的評估，確有其必要性。所以，本計畫提出一個較具彈性的雙變量二階段隨機對照設計，此試驗設計一方面除了考慮新藥有效性的評估外，亦將同時合併新藥安全性的監控；另一方面為了加快臨床試驗的進行，在確保檢定力不變的條件下，本計畫亦將縮減抽樣程序應用於此雙變量二階段隨機對照設計中，允許研究者可以根據當前試驗結果來決定是否要提前停止試驗，以達到縮短試驗時間的目的。

## 三、 文獻探討

### 3.1 單組雙變量兩階段設計 (Conaway and Petroni, 1995; Jin, 2007)

在癌症新藥的第一期臨床試驗中，主要目的之一是決定新藥的最大耐受劑量，然而由於試驗過程中所使用的病人數目通常很少，導致對於最大耐受劑量的估計並不是很精準，因此當此劑量進一步用於第二期臨床試驗時，基於安全性的考慮，除了考慮新藥的療效反應率外，合併考慮新藥的毒性反應率確有其必要。Conaway 和 Petroni (1995) 提出一個單組的雙變量二階段設計，針對每一個進入研究的受試者，在其接受處理後，同時紀錄其是否呈現有效反應與是否出現毒性反應，為一雙變量二元反應變項，考慮虛無假設與對立假設分別為

$$H_0 : \{(p_r, p_t) | p_r \leq p_{r_0} \text{ OR } p_t \geq p_{t_0}\} \text{ vs. } H_1 : \{(p_r, p_t) | p_r > p_{r_0} \text{ OR } p_t < p_{t_0}\}$$

利用交集-聯集檢定 (Intersection-Union tests, IUT)，來同時檢定療效反應率與毒性反應率。其臨床試驗程序如下：在試驗第一階段所招募的 $n_1$ 位受試者中，假設共有 $X_{r_1}$ 位受試者呈現有效反應與 $X_{t_1}$ 位受試者出現毒性反應，

- 若 $X_{r_1} < c_{r_1}$ 或 $X_{t_1} > c_{t_1}$ ，則試驗停止在第一階段，且不拒絕虛無假設；
- 反之，若 $X_{r_1} \geq c_{r_1}$ 且 $X_{t_1} \leq c_{t_1}$ ，則試驗允許進入第二階段。

當第二階段允許進行時，則持續招募 $n_2$ 位受試者，假設共有 $X_{r_2}$ 位受試者呈現有效反應與 $X_{t_2}$ 位受試者出現毒性反應。合併兩個階段的結果，在 $n = n_1 + n_2$ 位受試者中，令 $X_r = X_{r_1} + X_{r_2}$ 與 $X_t = X_{t_1} + X_{t_2}$

- 若  $X_r < c_{r_2}$  或  $X_t > c_{t_2}$ ，則不拒絕虛無假設；
- 反之，若  $X_r \geq c_{r_2}$  且  $X_t \leq c_{t_2}$ ，則拒絕虛無假設。

上述之檢定臨界值  $c_{r_1}$ ， $c_{t_1}$ ， $c_{r_2}$  與  $c_{t_2}$ ，在虛無假設成立下，須滿足下列兩個限制式

$$P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_{r_0}, p_{t_0}), \theta) \leq \alpha$$

$$\sup_{p_r \leq p_{r_0} \text{ OR } p_t \geq p_{t_0}} P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_{r_0}, p_{t_0}), \theta) \leq \gamma$$

反之，在對立假設成立下，給定療效反應率  $p_{r_a}$  ( $> p_{r_0}$ ) 與毒性反應率  $p_{t_a}$  ( $< p_{t_0}$ )，須滿足下列限制式

$$P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_{r_a}, p_{t_a}), \theta) \geq 1 - \beta$$

此處  $\theta$  用於描述療效反應率  $p_r$  與毒性反應率  $p_t$  的相關強度，而  $\alpha$  與  $\gamma$  為給定的型一誤差率且  $\alpha \leq \gamma$ ， $\beta$  為給定的型二誤差率。最後，以賽門二階段設計(1989)之大中取小準則來決定最佳的試驗設計參數。值得一提的是，對於任意的  $\theta$  而言，因為

$$P(X_{r_1} \geq c_{r_1}, X_r \geq c_{r_2} | p_r) = P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_r, p_t) = (p_r, 0))$$

且

$$P(X_{t_1} \leq c_{t_1}, X_t \leq c_{t_2} | p_t) = P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_r, p_t) = (1, p_t))$$

因此在虛無假設下的最大型一誤差率可以簡化為單變量的型式，即

$$P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_r, p_t), \theta)$$

$$\leq \min(P(X_{r_1} \geq c_{r_1}, X_r \geq c_{r_2} | p_r), P(X_{t_1} \leq c_{t_1}, X_t \leq c_{t_2} | p_t))$$

因此，Jin (2007) 沿用 Conaway 和 Petroni (1995) 所提出的臨床試驗程序，提出彈性調整療效反應率與毒性反應率之型一誤差率的二階段設計。在給定的型一誤差率  $\alpha_r$  和  $\alpha_t$  與型二誤差率  $\beta$  下，檢定臨界值  $c_{r_1}$ ， $c_{t_1}$ ， $c_{r_2}$  與  $c_{t_2}$  須滿足下列限制式

$$P(X_{r_1} \geq c_{r_1}, X_r \geq c_{r_2} | p_{r_0}) \leq \alpha_r$$

$$P(X_{t_1} \leq c_{t_1}, X_t \leq c_{t_2} | p_{t_0}) \leq \alpha_t$$

與

$$P(X_{r_1} \geq c_{r_1}, X_{t_1} \leq c_{t_1}, X_r \geq c_{r_2}, X_t \leq c_{t_2} | (p_{r_a}, p_{t_a}), \theta) \geq 1 - \beta$$

而在最佳試驗設計參數的選擇上，則是採用賽門二階段設計 (1989) 之最佳化準則。比較 Conaway 和 Petroni (1995) 與 Jin (2007) 所提出的型一誤差限制式不難發現，若研究者想要控制整體型一誤差率，則可以將其設定為  $\gamma = \alpha$  或是  $\alpha_r = \alpha_t = \alpha$ ，此時兩者是一致的。

### 3.2 單組雙變量縮減兩階段設計 (Chi and Chen, 2012)

為了為加速臨床試驗的進行，縮短臨床試驗所需的時間，Chi 和 Chen (2012) 將縮減抽樣設計的概念應用至 Conaway 和 Petroni (1995) 的臨床試驗程序中，提出一個單組雙變量縮減二階段設計，其

臨床試驗程序如下：在試驗第一階段，最多招募 $n_1$ 位受試者，根據目前已觀察到結果的受試者

- 若無效反應的人數已經累積到 $n_1 - c_{r_1} + 1$ 或是毒性反應的人數已經累積到 $c_{t_1} + 1$ ，則試驗立即停止，且不拒絕虛無假設；
- 反之，若有效反應的人數已經累積到 $c_{r_1}$ 且無毒性反應的人數已經累積到 $n_1 - c_{t_1}$ ，則試驗允許提前進入第二階段。

當第二階段允許進行時，則持續招募受試者，至多到 $n = n_1 + n_2$ 位，合併第一階段的結果，

- 若無效反應的總人數已經累積到 $n - c_{r_2} + 1$ 或是毒性反應的人數已經累積到 $c_{t_2} + 1$ ，則試驗立即停止，且不拒絕虛無假設；
- 反之，若有效反應的人數已經累積到 $c_{r_2}$ 且無毒性反應的人數已經累積到 $n - c_{t_2}$ ，則試驗立即停止，且拒絕虛無假設。

藉由縮減抽樣設計程序的應用，當目前受試者的結果已足夠確認最後的檢定結果時，允許是臨床試驗立即停止，此方法可以有效減少期望受試者數目，加快臨床試驗的進行，然而此縮減二階段設計無法直接應用於雙組的隨機對照設計試驗中。

### 3.3 單變量二階段隨機對照設計 (Jung, 2008)

Jung (2008) 針對單變量二元反應變項提出了一個二階段隨機對照設計，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， $\Delta$ 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，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。

綜合前述文獻，不論是單變量二元反應變項或是雙變量二元反應變項，若為單組二階段設計，則研究人員均有相當多的試驗設計可供選擇；在雙組試驗下，若是考慮單變量二元反應變項，則可採用Jung (2008)所提出的二階段隨機對照設計。如果研究者考慮比較兩個二元反應變項，則可採用Bersimis et al. (2015)所提出的單階段設計進行，然而此方法僅能在兩邊個數一樣的情況下才可比較，使用上大受限制。因此，當受試者招募不易時，若研究人員希望可以有機會節省樣本量，提前結束試驗，目前尚無可用之臨床試驗設計。因此本計畫將Jung (2008)的二階段隨機對照設計推廣至適用於雙變量二元反應變項，且運用縮減抽樣程序來節省期望樣本個數，以達到加速試驗進行的目地。

## 四、 研究方法

### 4.1 雙變量單階段設計

在同時考量有效性與安全性的情況下，假設實驗組之療效反應率與安全應率分別為 $p_{A,r}$ 與 $p_{A,s}$ ，對照組之療效反應率與安全率分別為 $p_{B,r}$ 與 $p_{B,s}$ ，其檢定之虛無假設為實驗組與對照組比較沒有較佳的療效或是安全性，而對立假設則為實驗組與對照組比較有較佳的療效與安全性，可以寫成下列形式

$$H_0 : \{(p_{A,r}, p_{A,s}, p_{B,r}, p_{B,s}) | p_{A,r} \leq p_{B,r} \text{ or } p_{A,s} \leq p_{B,s}\} \text{ vs.}$$

$$H_1 : \{(p_{A,r}, p_{A,s}, p_{B,r}, p_{B,s}) | p_{A,r} > p_{B,r} \text{ and } p_{A,s} > p_{B,s}\}$$

假設實驗組與對照組預計分別招募 $n$ 個受試者，而每位受試者在接受治療後均可同時觀察其療效與安全性兩個結果。令 $X_{A,r}$ 與 $X_{A,s}$ 分別表示實驗組 $n$ 個受試者中有效反應人數與安全反應人數，分別為 $Bin(n, p_{A,r})$ 與 $Bin(n, p_{A,s})$ 之二項分配，則 $n$ 位受試者中出現 $X_{A,r} = x_{A,r}$ 與 $X_{A,s} = x_{A,s}$ 的機率為

$$P(X_{Ar} = x_{Ar}, X_{As} = x_{As})$$

$$= \sum_{i=\max(0, x_{Ar}+x_{As}-n)}^{\min(x_{Ar}, x_{As})} \binom{n}{i \quad x_{Ar}-i \quad x_{As}-i \quad n-x_{Ar}-x_{As}+i} p_{A,11}^i p_{A,10}^{x_{Ar}-i} p_{A,01}^{x_{As}-i} p_{A,00}^{n-x_{Ar}-x_{As}+i}$$

此處 $p_{A,11}$ ， $p_{A,10}$ ， $p_{A,01}$ 與 $p_{A,00}$ 分別表示每位受試者接受治療後呈現(有效，安全)、(有效，不安全)、(無效，安全)與(無效，不安全)四種反應的機率，且療效反應率 $p_{A,1} = p_{A,11} + p_{A,10}$ 與安全率 $p_{A,2} = p_{A,11} + p_{A,01}$ 之間的相關性可以 $\phi_A = p_{A,11}p_{A,00}/p_{A,10}p_{A,01}$ 來表示。同樣的，令 $X_{B,r}$ 與 $X_{B,s}$ 分別表示對照組 $n$ 個受試者中有效反應人數與安全反應人數，分別為二項分配 $Bin(n, p_{B,r})$ 與 $Bin(n, p_{B,s})$ 之二項分配， $n$ 位受試者中出現 $X_{B,r} = x_{B,r}$ 與 $X_{B,s} = x_{B,s}$ 的機率為

$$P(X_{Br} = x_{Br}, X_{Bs} = x_{Bs})$$

$$= \sum_{i=\max(0, x_{Br}+x_{Bs}-n)}^{\min(x_{Br}, x_{Bs})} \binom{n}{i \ x_{Br} - i \ x_{Bs} - i \ n - x_{Br} - x_{Bs} + i} p_{B,11}^i p_{B,10}^{x_{Br}-i} p_{B,01}^{x_{Bs}-i} p_{B,00}^{n-x_{Br}-x_{Bs}+i}$$

此處  $p_{B,11}$ ,  $p_{B,10}$ ,  $p_{B,01}$  與  $p_{B,00}$  分別表示每位受試者接受治療後呈現(有效, 安全)、(有效, 不安全)、(無效, 安全) 與 (無效, 不安全) 四種反應的機率, 且療效反應率  $p_{B,1} = p_{B,11} + p_{B,10}$  與安全率  $p_{B,2} = p_{B,11} + p_{B,01}$  之間的相關性可以  $\phi_B = p_{B,11}p_{B,00}/p_{B,10}p_{B,01}$  來表示。利用交集-聯集檢定的概念, 來同時檢定療效反應率與安全率, 在試驗所招募的  $n$  位受試者中, 若  $X_{A,r} - X_{B,r} \geq r$  且  $X_{A,s} - X_{B,s} \geq s$ , 則檢定結果拒絕虛無假設。為應用縮減抽樣程序的方便, 令  $D_r = X_{A,r} + (n - X_{B,r})$  表示實驗組有效反應人數與對照組無效反應人數的加總, 而  $D_s = X_{A,s} + (n - X_{B,s})$  表示實驗組安全反應人數與對照組不安全反應人數的加總, 則拒絕域可以進一步改寫成為  $\mathfrak{R} = \{(D_r, D_s) | D_r \geq c_r, D_s \geq c_s\}$ , 其中  $c_r = n + r$ ,  $c_s = n + s$ 。而  $D_r$  與  $D_s$  的聯合機率密度函數可以寫成

$$P(D_r = d_r, D_s = d_s) = \sum_{i=\max(0, d_r-n)}^{\min(n, d_r)} \sum_{j=\max(0, d_s-n)}^{\min(n, d_s)} P(X_{A,r} = i, X_{A,s} = j) \times P(X_{B,r} = i + n - d_r, X_{B,s} = j + n - d_s)$$

此處,  $0 \leq d_r \leq 2n$  和  $0 \leq d_s \leq 2n$ 。因此, 當試驗完成招募後, 拒絕虛無假設的機率即為

$$P(D_r \geq c_r, D_s \geq c_s) = \sum_{d_r=c_r}^{2n} \sum_{d_s=c_s}^{2n} P(D_r = d_r, D_s = d_s)$$

對於給定的型一誤差率  $\alpha$  與檢定力  $1 - \beta$ , 樣本數  $n$  與臨界點  $c_r$  和  $c_s$  必須滿足

$$\sup_{H_0} P(D_r \geq c_r, D_s \geq c_s) \leq \alpha \quad (1)$$

與

$$\inf_{H_1} P(D_r \geq c_r, D_s \geq c_s) \geq 1 - \beta \quad (2)$$

令  $\Delta_r = p_{A,r} - p_{B,r}$  和  $\Delta_s = p_{A,s} - p_{B,s}$ , 則前述假設可以簡化成

$$H_0 : \Delta_r \leq 0 \text{ or } \Delta_s \leq 0 \quad \text{vs.} \quad H_1 : \Delta_r > 0 \text{ and } \Delta_s > 0$$

其中虛無假設可以改寫成  $\{\Delta_r \leq 0 \text{ and } -1 \leq \Delta_s \leq 1\} \cup \{-1 \leq \Delta_r \leq 1 \text{ and } \Delta_s \leq 0\}$ 。因為當  $\Delta_s = 1$  時, 則表示  $p_{A,s} = 1$  與  $p_{B,s} = 0$ , 即代表  $P(X_{A,s} = n) = 1$  與  $P(X_{B,s} = 0) = 1$ , 此時  $P(D_s = 2n) = 1$ , 所以

$$\sup_{\Delta_r \leq 0, -1 \leq \Delta_s \leq 1} P(D_r \geq c_r, D_s \geq c_s | \Delta_r, \Delta_s) = P(D_r \geq c_r, D_s \geq c_s | \Delta_r = 0, \Delta_s = 1) = P(D_r \geq c_r | \Delta_r = 0)$$

同理,

$$\sup_{-1 \leq \Delta_r \leq 1, \Delta_s \leq 0} P(D_r \geq c_r, D_s \geq c_s | \Delta_r, \Delta_s) = P(D_r \geq c_r, D_s \geq c_s | \Delta_r = 1, \Delta_s = 0) = P(D_s \geq c_s | \Delta_s = 0)$$

因此,

$$\sup_{H_0} P(D_r \geq c_r, D_s \geq c_s) \leq \min\{P(D_r \geq c_r | \Delta_r = 0), P(D_s \geq c_s | \Delta_s = 0)\}$$

所以, 可將(1)進一步分開設定為  $P(D_r \geq c_r | \Delta_r = 0) \leq \alpha_r$  與  $P(D_s \geq c_s | \Delta_s = 0) \leq \alpha_s$ , (2)則可寫成

$$P(D_r \geq c_r, D_s \geq c_s | \Delta_r = \delta_r, \Delta_s = \delta_s) \geq 1 - \beta$$

對於給定的型一誤差率 $\alpha_r$ 、 $\alpha_s$ 與檢定力 $1 - \beta$ ，在虛無假設下，當給定 $p_{A,r} = p_{B,r} = p_{r0}$ 與 $p_{A,s} = p_{B,s} = p_{s0}$ 時，試驗設計參數 $(n, c_r, c_s)$ 須滿足

$$P(D_r \geq c_r | p_{A,r} = p_{B,r} = p_{r0}) \leq \alpha_r \quad (3)$$

與

$$P(D_s \geq c_s | p_{A,s} = p_{B,s} = p_{s0}) \leq \alpha_s \quad (4)$$

而在對立假設下，當給定 $(p_{A,r}, p_{B,r}) = (p_{r0} + \delta_r, p_{r0})$ 與 $(p_{A,s}, p_{B,s}) = (p_{s0} + \delta_s, p_{s0})$ 時，試驗設計參數 $(n, c_r, c_s)$ 則須滿足

$$P(D_r \geq c_r, D_s \geq c_s | (p_{A,r}, p_{B,r}) = (p_{r0} + \delta_r, p_{r0}), (p_{A,s}, p_{B,s}) = (p_{s0} + \delta_s, p_{s0})) \geq 1 - \beta \quad (5)$$

接下來，從滿足(3)、(4)與(5)的所有試驗設計參數 $(n, c_r, c_s)$ 中，尋找樣本數最小的參數組合即為最佳的試驗設計參數，若同時存在兩組以上的試驗設計參數滿足(3)、(4)與(5)，則以檢定力較高者為最佳的試驗設計參數。

## 4.2 雙變量二階段設計

根據 4.1 結果，進一步將其推廣至二階段設計，首先假設實驗組與對照組預計在第一階段分別招募  $n_1$  個受試者，令  $X_{A1,r}$  與  $X_{A1,s}$  分別表示實驗組於第一階段觀察到的療效反應人數與安全反應人數，同時令  $X_{B1,r}$  與  $X_{B1,s}$  分別表示對照組於第一階段觀察到的療效反應人數與安全反應人數，假設  $D_{1,r} = X_{A1,r} + (n_1 - X_{B1,r})$  為第一階段完成後實驗組有效反應人數與對照組無效反應人數的總和，而  $D_{1,s} = X_{A1,s} + (n_1 - X_{B1,s})$  則表示兩組於第一階段完成後實驗組安全反應人數與對照組不安全反應人數總和。

- 若  $D_{1,r} < c_{r1}$  或  $D_{1,s} < c_{s1}$ ，則試驗停止在第一階段，且不拒絕虛無假設；
- 反之，若  $D_{1,r} \geq c_{r1}$  且  $D_{1,s} \geq c_{s1}$ ，則試驗允許進入第二階段。

當第二階段允許進行時，兩組分別再持續招募  $n_2 = n - n_1$  個病人。令  $X_{A2,r}$  與  $X_{A2,s}$  分別表示實驗組於第二階段觀察到的有效反應人數與安全反應人數，同時令  $X_{B2,r}$  與  $X_{B2,s}$  分別表示對照組於第二階段觀察到的有效反應人數與安全反應人數。令  $D_r = (X_{A1,r} + X_{A2,r}) + (n - X_{B1,r} - X_{B2,r})$  表示完成兩個階段後實驗組有效反應人數與對照組無效反應人數的總和，而  $D_s = (X_{A1,s} + X_{A2,s}) + (n - X_{B1,s} - X_{B2,s})$  表示完成兩個階段後實驗組安全反應人數與對照組不安全反應人數的總和。

- 若  $D_r < c_r$  或  $D_s < c_s$ ，則試驗結果不拒絕虛無假設；
- 反之，若  $D_r \geq c_r$  且  $D_s \geq c_s$ ，則試驗結果拒絕虛無假設。

根據上述試驗方式，可定義檢定拒絕域為

$$\mathfrak{R} = \{(D_{1,r}, D_{1,s}, D_r, D_s) | D_{1,r} \geq c_{r1}, D_{1,s} \geq c_{s1}, D_r \geq c_r, D_s \geq c_s\}$$

並可推導出檢定力函數為

$$P(\mathfrak{R}) = \sum_{d_{1,r}=c_{r1}}^{2n_1} \sum_{d_{1,s}=c_{s1}}^{2n_1} P(D_{1,r} = d_{1,r}, D_{1,s} = d_{1,s}) P(D_{2,r} \geq c_r - d_{1,r}, D_{2,s} \geq c_s - d_{1,s})$$

此處  $D_{2,r} = X_{A2,r} + (n - X_{B2,r})$ ， $D_{2,s} = X_{A2,s} + (n - X_{B2,s})$ 。對於給定的型一誤差率 $\alpha$ 與檢定力 $1 - \beta$ ，試

驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 則必須滿足 $\sup_{H_0} P(\mathfrak{R}) \leq \alpha$ 與 $\inf_{H_1} P(\mathfrak{R}) \geq 1 - \beta$ 兩條限制式。根據 4.1 的結果，在給定的型一誤差率 $\alpha_r$ 、 $\alpha_s$ 與檢定力 $1 - \beta$ 下，設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 須滿足

$$\sum_{d_{1,r}=c_{r1}}^{2n_1} P(D_{1,r} = d_{1,r} | p_{A,r} = p_{B,r} = p_{r0}) P(D_{2,r} \geq c_r - d_{1,r} | p_{A,r} = p_{B,r} = p_{r0}) \leq \alpha_r \quad (6)$$

與

$$\sum_{d_{1,s}=c_{s1}}^{2n_1} P(D_{1,s} = d_{1,s} | p_{A,s} = p_{B,s} = p_{s0}) P(D_{2,s} \geq c_s - d_{1,s} | p_{A,s} = p_{B,s} = p_{s0}) \leq \alpha_s \quad (7)$$

而在對立假設下，當給定 $(p_{A,r}, p_{B,r}) = (p_{r0} + \delta_r, p_{r0})$ 與 $(p_{A,s}, p_{B,s}) = (p_{s0} + \delta_s, p_{s0})$ 時，試驗設計參數 $(n, c_r, c_s)$ 則須滿足

$$P(\mathfrak{R} | (p_{A,r}, p_{B,r}) = (p_{r0} + \delta_r, p_{r0}), (p_{A,s}, p_{B,s}) = (p_{s0} + \delta_s, p_{s0})) \geq 1 - \beta \quad (8)$$

接下來，針對所有滿足(6)、(7)與(8)的所有試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 中，尋找總樣本數最小的參數組合，若同時存在兩組以上的試驗設計參數，則以虛無假設下期望樣本個數 $E(N|H_0)$ 最小者為最佳的試驗設計參數，即大中取小設計(minimax criterion)。反之，針對所有滿足(6)、(7)與(8)的所有試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 中，直接尋找虛無假設下期望樣本個數 $E(N|H_0)$ 最小者為最佳的試驗設計參數，即最佳化準則(optimal criterion)。此處，期望樣本個數可以透過

$$E(N | p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s}) = 2 * (n_1 + (1 - PET)n_2)$$

來計算，表示在虛無假設下所需的平均受試者人數，而

$$PET = 1 - \sum_{d_{1,r}=c_{r1}}^{2n_1} \sum_{d_{1,s}=c_{s1}}^{2n_1} P(D_{1,r} = d_{1,r}, D_{1,s} = d_{1,s})$$

表示試驗停止在第一階段的機率。

### 4.3 雙變量縮減二階段設計

為加速臨床試驗的進行，接下來將依據 4.2 所得到的試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ ，應用縮減設計程序來減少試驗的期望樣本個數。其臨床試驗程序如下：在試驗第一階段，實驗組與對照組最多招募 $n_1$ 位受試者，根據目前兩組已觀察到結果的 $Y_{A1}(\leq n_1)$ 與 $Y_{B1}(\leq n_1)$ 位受試者，假設 $D_{1,r}(Y_{A,1}, Y_{B,1})$ 表示第一階段已觀察到試驗結果的 $Y_{A1}$ 位受試者中有效反應人數與 $Y_{B1}$ 位受試者中無效反應人數的總和， $D_{1,s}(Y_{A,1}, Y_{B,1})$ 表示第一階段已觀察到試驗結果的 $Y_{A1}$ 位受試者中安全反應人數與 $Y_{B1}$ 位受試者中不安全反應人數的總和，令 $D_{1,r}^c(Y_{A,1}, Y_{B,1}) = Y_{A,1} + Y_{B,1} - D_{1,r}(Y_{A,1}, Y_{B,1})$ 表示第一階段已觀察到試驗結果的 $Y_{A1}$ 位受試者中無效反應人數與 $Y_{B1}$ 位受試者中有效反應人數的總和， $D_{1,s}^c(Y_{A,1}, Y_{B,1}) = Y_{A,1} + Y_{B,1} - D_{1,s}(Y_{A,1}, Y_{B,1})$ 表示第一階段已觀察到試驗結果的 $Y_{A1}$ 位受試者中不安全反應人數與 $Y_{B1}$ 位受試者中安全反應人數的總和，

- 在目前已觀察到結果的 $Y_{A1}$ 與 $Y_{B1}$ 位受試者中，若觀察到 $D_{1,r}^c(Y_{A,1}, Y_{B,1}) = n_1 - c_{r1} + 1$ 或是 $D_{1,s}^c(Y_{A,1}, Y_{B,1}) = n_1 - c_{s1} + 1$ ，則試驗立即停止，且不拒絕虛無假設；
- 反之，在目前已觀察到結果的 $Y_{A1}$ 與 $Y_{B1}$ 位受試者中，若觀察到 $D_{1,r}(Y_{A,1}, Y_{B,1}) = c_{r1}$ 且

$D_{1,s}(Y_{A,1}, Y_{B,1}) = c_{s1}$ ，則試驗允許提前進入第二階段；

- 若上述兩種情況均未發生，則持續招募受試者至多 $n_1$ 。

當第二階段允許進行時，則持續招募受試者，至多到 $n = n_1 + n_2$ 位，合併第一階段的結果，假設 $D_r(Y_A, Y_B)$ 表示目前已觀察到試驗結果的 $Y_A (= Y_{A,1} + Y_{A,2})$ 位受試者中有效反應人數與 $Y_B (= Y_{B,1} + Y_{B,2})$ 位受試者中無效反應人數的總和， $D_s(Y_A, Y_B)$ 表示目前已觀察到試驗結果的 $Y_A$ 位受試者中安全反應人數與 $Y_B$ 位受試者中不安全反應人數的總和，令 $D_r^c(Y_A, Y_B) = Y_A + Y_B - D_r(Y_A, Y_B)$ 表示目前已觀察到試驗結果的 $Y_A$ 位受試者中無效反應人數與 $Y_B$ 位受試者中有效反應人數的總和， $D_s^c(Y_A, Y_B) = Y_A + Y_B - D_s(Y_A, Y_B)$ 表示目前已觀察到試驗結果的 $Y_A$ 位受試者中不安全反應人數與 $Y_B$ 位受試者中安全反應人數的總和，

- 在目前已觀察到結果的 $Y_A$ 與 $Y_B$ 位受試者中，若觀察到 $D_r^c(Y_A, Y_B) = n - c_r + 1$ 或是 $D_s^c(Y_A, Y_B) = n - c_s + 1$ ，則試驗立即停止，且不拒絕虛無假設；
- 反之，在目前已觀察到結果的 $Y_A$ 與 $Y_B$ 位受試者中，若觀察到 $D_r(Y_A, Y_B) = c_r$ 且 $D_s(Y_A, Y_B) = c_s$ ，則試驗立即停止，並拒絕虛無假設；
- 若上述兩種情況均未發生，則持續招募受試者至多 $n$ 。

根據上述雙變量縮減二階段設計之試驗流程，第一階段每組至多招募 $n_1$ 位受試者，然而試驗有機會提前停止試驗，反之雙變量二階段在第一階段則必定招募 $n_1$ 位受試者；同理，當第二階段允許進行時，雙變量縮減二階段設計之試驗流程，每組至多招募 $n$ 位受試者，然而試驗有機會提前停止試驗，反之雙變量二階段則必定招募 $n$ 位受試者，因此縮減二階段設計之期望樣本數 $E_c(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})$ 並定會小於等於雙變量二階段設計之期望樣本數 $E(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})$ ，為探討兩種試驗設計之差異，將以模擬方式進行比較。

## 五、 結果與討論（含結論與建議）

相較於固定每階段樣本的雙變量二階段設計，在雙變量縮減二階段設計中，除了考量反應人數外，尚須考慮兩階段目前所實際招募的受試者人數，導致在機率的計算較為困難且複雜，進而使得在試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 的尋找上需要耗費大量的時間。然而根據 Chi and Chen (2008, 2012)的結果，在單雙變量單組二階段設計中，在相同的型一誤差率與檢定力的設定下，有無採用縮減設計，兩種設計的試驗設計參數大多都是相同的。因此，在本研究中採用以雙變量二階段設計的試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 直接套用在雙變量縮減二階段設計中，並比較在相同的試驗設計參數下，兩種試驗設計在期望樣本數的差異。為了比較固定每階段樣本的雙變量二階段設計與雙變量縮減二階段設計，在虛無假設下進一步以模擬的方式比較兩種設計期望樣本個數的差異，並計算其減少的相對百分比

$$RS(\%) = \frac{E(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s}) - E_c(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})}{E_c(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})} \times 100\%。$$

此處， $E(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})$ 和 $E_c(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})$ 分別表示固定每階段樣本的雙變量二階段設計與雙變量縮減二階段設計的期望樣本個數。值得一提的是，當直接以固定每階段樣本的雙變量二階段設計之試驗設計參數 $(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 來作為雙變量縮減二階段設計之試驗設計參數時，此試驗設計參數

$(n_1, c_{r1}, c_{s1}, n, c_r, c_s)$ 不一定是雙變量縮減二階段設計中的最佳試驗設計參數。

為方便計算試驗設計參數，設定每組樣本數搜尋上限為 $n \leq 100$ ，且為避免第一階段或第二階段的樣本數過小，設定 $n_1, n_2 \geq 10$ 的條件。簡化虛無假設為 $p_{A,r} = p_{B,r} = p_{r0}$ 與 $p_{A,s} = p_{B,s} = p_{s0}$ ，對立假設為 $(p_{A,r}, p_{B,r}) = (p_{r0} + \delta_r, p_{r0})$ 與 $(p_{A,s}, p_{B,s}) = (p_{s0} + \delta_s, p_{s0})$ 時，若 $(\delta_r, \delta_s) = (0.2, 0.2)$ 時，在給定的型一誤差率 $\alpha_r = 0.15$ 、 $\alpha_s = 0.05$ 與檢定力 $1 - \beta = 0.8$ 下，假設不論是虛無假設或對立假設下，實驗組兩個反應率 $(p_{A,r}, p_{A,s})$ 與對照組兩個反應率 $(p_{B,r}, p_{B,s})$ 之相關性均定為 $\phi$ 。考慮四種 $\phi=0.1, 0.5, 2$ 與 $10$ 四種不同的相關強度，根據大中取小準則與最佳化準則所決定之試驗設計參數結果如 Table 1.和 Table 2.所示。由 Table 1. 可以發現，在大中取小準則下，當給定相同的 $p_{r,0}$ 時，隨著 $\phi$ 增加，在雙變量二階段設計中所需的樣本總數會略為下降，但期望樣本個數則無明顯趨勢；反之，在最佳化準則下，由 Table 2. 可以發現隨著 $\phi$ 增加，雙變量二階段設計中所需的期望樣本總數會逐漸上升。

為了解縮減抽樣程序應用於雙變量二階段設計的效益，以雙變量二階段設計在大中取小準則與最佳化準則中所決定的試驗設計參數進行模擬，步驟如下，

1. 根據給定的試驗設計參數來產生受試者人數資料與設定臨界點，在第一階段，隨機排列 $1, 2, \dots, 2n_1$ 的序列代表受試者依序進入臨床試驗的順序，若為 $1, 2, \dots, n_1$ ，表示實驗組；若為 $n_1 + 1, 2, \dots, 2n_1$ 表示對照組。若第二階段允許進行，則隨機排列 $1, 2, \dots, 2n - y_1 - y_2$ 的序列代表受試者依序進入臨床試驗的順序，若為 $1, 2, \dots, n - y_1$ ，表示實驗組；若為 $n - y_1 + 1, 2, \dots, 2n - y_1 - y_2$ 表示對照組，此處 $y_1$ 與 $y_2$ 分別表示試驗在第一階段兩組的收案人數。
2. 根據給定的 $(p_{A,r}, p_{A,s}, \phi_A)$ 可以求出 $p_{A,11}$ ， $p_{A,10}$ ， $p_{A,01}$ 與 $p_{A,00}$ ，並依照下列方式可產生每一位實驗組受試者的試驗結果。由均勻分配 $p \sim U(0, 1)$ 生成發生機率，若 $0 \leq p < p_{A,11}$ ，表示該名受試者有效且安全；若 $p_{A,11} \leq p < p_{A,11} + p_{A,10}$ ，表示該名受試者有效但不安全；若 $p_{A,11} + p_{A,10} \leq p < p_{A,11} + p_{A,10} + p_{A,01}$ ，表示該名受試者無效但安全；若 $p_{A,11} + p_{A,10} + p_{A,01} \leq p$ ，表示該名受試者無效且不安全。
3. 根據給定的 $(p_{B,r}, p_{B,s}, \phi_B)$ 可以求出 $p_{B,11}$ ， $p_{B,10}$ ， $p_{B,01}$ 與 $p_{B,00}$ ，依照步驟一的方式可產生每一位對照組受試者的試驗結果。
4. 依據 4.3 的試驗程序進行並計算每次完成試驗所需的受試者人數，在完成 1000000 次模擬後計算其平均人數即為雙變量縮減二階段設計的期望樣本個數 $E_c(N|p_{A,r}, p_{B,r}, p_{A,s}, p_{B,s})$ 。

根據前述的模擬方式，在虛無假設下所得到的期望樣本個數 $E_c(N|H_0)$ 如表 Table 1.和 Table 2.所示，與雙變量二階段設計相比，當採用縮減抽樣程序後，期望樣本個數的縮減幅度 $RS(\%)$ 大約在 10% ~ 15%之間。由此可知，當以固定每階段樣本的雙變量二階段設計之試驗設計參數進行臨床試驗時，若改以雙變量縮減二階段設計的試驗流程進行時，確實可以有效減少試驗的期望樣本個數。

## Reference

1. Ayanlowo, A. O., & Redden, D. T. (2007). Stochastically curtailed phase II clinical trials. *Statistics in medicine*, 26(7), 1462-1472.
2. Bersimis, S., Sachlas, A., & Papaioannou, T. (2015). Flexible designs for phase II comparative clinical trials involving two response variables. *Statistics in medicine*, 34(2), 197-214.
3. Bryant, J., & Day, R. (1995). Incorporating toxicity considerations into the design of two-stage phase II clinical trials. *Biometrics*, 1372-1383.
4. Carsten, C., & Chen, P. (2015). Curtailed Two-Stage Matched Pairs Design in Double-Arm Phase II Clinical Trials. *Journal of Biopharmaceutical Statistics*, (just-accepted).
5. Cellamare, M., & Sambucini, V. (2015). A randomized two-stage design for phase II clinical trials based on a Bayesian predictive approach. *Statistics in medicine*, 34(6), 1059-1078.
6. Chen, C. M., & Chi, Y. (2012). Curtailed two-stage designs with two dependent binary endpoints. *Pharmaceutical statistics*, 11(1), 57-62.
7. Chen, T. T. (1997). Optimal three-stage designs for phase II cancer clinical trials. *Statistics in medicine*, 16(23), 2701-2711.
8. Chi, Y., & Chen, C. M. (2008). Curtailed two-stage designs in Phase II clinical trials. *Statistics in medicine*, 27(29), 6175-6189.
9. Conaway, M. R., & Petroni, G. R. (1995). Bivariate sequential designs for phase II trials. *Biometrics*, 656-664.
10. Dale, J. R. (1986). Global cross-ratio models for bivariate, discrete, ordered responses. *Biometrics*, 909-917.
11. Fleming, T. R. (1982). One-sample multiple testing procedure for phase II clinical trials. *Biometrics*, 143-151.
12. Gehan, E. A. (1961). The determination of the number of patients required in a preliminary and a follow-up trial of a new chemotherapeutic agent. *Journal of chronic diseases*, 13(4), 346-353.
13. Hou, W., Chang, M. N., Jung, S. H., & Li, Y. (2013). Designs for randomized phase II clinical trials with two treatment arms. *Statistics in medicine*, 32(25), 4367-4379.
14. Jin, H. (2007). Alternative designs of phase II trials considering response and toxicity. *Contemporary clinical trials*, 28(4), 525-531.
15. Jung, S. H. (2008). Randomized phase II trials with a prospective control. *Statistics in medicine*, 27(4), 568-583.
16. Jung, S. H., & Sargent, D. J. (2014). Randomized Phase II clinical trials. *Journal of biopharmaceutical statistics*, 24(4), 802-816.

17. Kepner, J. L. (2010). On group sequential designs comparing two binomial proportions. *Journal of biopharmaceutical statistics*, 20(1), 145-159.
18. Kunz, C. U., & Kieser, M. (2012). Curtailment in single-arm two-stage phase II oncology trials. *Biometrical Journal*, 54(4), 445-456.
19. Ray, H. E., & Rai, S. N. (2011). An evaluation of a Simon 2-Stage phase II clinical trial design incorporating toxicity monitoring. *Contemporary clinical trials*, 32(3), 428-436.
20. Ray, H. E., & Rai, S. N. (2012). Operating characteristics of a Simon two-stage phase II clinical trial design incorporating continuous toxicity monitoring. *Pharmaceutical statistics*, 11(2), 170-176.
21. Shan, G., Ma, C., Hutson, A. D., & Wilding, G. E. (2013). Randomized two-stage phase II clinical trial designs based on Barnard's exact test. *Journal of biopharmaceutical statistics*, 23(5), 1081-1090.
22. Shuster, J. (2002). Optimal two-stage designs for single arm phase II cancer trials. *Journal of Biopharmaceutical Statistics*, 12(1), 39-51.
23. Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled clinical trials*, 10(1), 1-10.

Table 1. Comparisons of expected sample size between the bivariate two-stage randomized designs with fixed sample size and the bivariate two-stage randomized designs with curtailment based on minimax criterion,  $\delta_r = 0.2$ ,  $p_{A,s0} = p_{B,s0} = 0.7$  under  $H_0$  and  $p_{A,s1} = 0.9$ ,  $p_{B,s1} = 0.7$  under  $H_1$ .

$p_{r,0}$	$\phi$	$n$	$n_1$	$r_1$	$s_1$	$r$	$s$	$\alpha_r$	$\alpha_s$	$1 - \beta$	$E(N H_0)$	$E_c(N H_0)$	$RS(\%)$
0.1	0.1	63	25	24	25	67	72	0.145	0.047	0.800	79.5	71.62	11.0
	0.5	62	25	22	23	66	71	0.146	0.047	0.800	104.76	92.49	13.3
	2.0	62	24	22	23	66	71	0.145	0.047	0.800	94.54	84.70	11.6
	10	62	26	23	26	66	71	0.146	0.046	0.800	90.84	81.65	11.3
0.2	0.1	68	59	63	66	73	77	0.146	0.050	0.800	118.06	102.22	15.5
	0.5	68	59	63	66	73	77	0.146	0.050	0.804	118.26	102.96	14.9
	2.0	68	44	46	48	73	77	0.147	0.050	0.802	92.02	82.43	11.6
	10	66	51	54	56	71	75	0.145	0.049	0.800	104.04	92.80	12.1
0.3	0.1	74	50	52	52	80	84	0.149	0.044	0.800	103.24	93.76	10.1
	0.5	73	58	61	65	79	82	0.150	0.049	0.801	116.6	101.63	14.7
	2.0	70	54	56	60	76	79	0.150	0.050	0.801	109.84	97.24	13.0
	10	68	63	67	71	74	77	0.148	0.049	0.800	126.36	110.32	14.5
0.4	0.1	75	32	26	30	82	85	0.139	0.045	0.800	124.7	109.19	14.2
	0.5	75	32	29	32	82	85	0.138	0.044	0.800	101.38	91.42	10.9
	2.0	74	32	27	32	81	84	0.138	0.043	0.800	107.58	96.73	11.2
	10	72	59	62	66	79	81	0.131	0.050	0.801	119.5	105.41	13.4
0.5	0.1	75	31	28	31	82	85	0.143	0.044	0.800	97.58	87.56	11.4
	0.5	74	59	60	66	81	83	0.142	0.050	0.800	119.02	104.67	13.7
	2.0	73	59	61	66	80	82	0.139	0.050	0.802	119.34	105.13	13.5
	10	71	59	64	66	77	80	0.145	0.050	0.800	119.12	104.17	14.4
0.6	0.1	72	64	67	72	79	81	0.134	0.050	0.800	128.12	110.97	15.5
	0.5	72	59	62	66	79	81	0.131	0.050	0.800	118.56	103.48	14.6
	2.0	71	59	64	66	77	80	0.140	0.050	0.801	118.66	103.03	15.2
	10	67	59	64	66	73	76	0.144	0.049	0.800	118.78	104.03	14.2
0.7	0.1	65	49	50	53	71	74	0.144	0.050	0.800	99.94	89.80	11.3
	0.5	65	25	23	25	71	74	0.144	0.050	0.800	83.76	74.97	11.7
	2.0	64	27	23	26	70	73	0.144	0.050	0.800	99.94	89.05	12.2
	10	63	27	25	26	69	72	0.141	0.048	0.800	95.66	86.07	11.1

1. For each  $p_{r,0}$ , values in the first row to fourth row are calculated by  $\phi = 0.1, 0.5, 2, 10$ , respectively, under  $(\alpha_r, \alpha_s, \beta) = (0.15, 0.05, 0.2)$ .
2. The response rates for two group are  $(p_{A,r0}, p_{B,r0}) = (p_{r,0}, p_{r,0})$  under  $H_0$  and  $(p_{A,r1}, p_{B,r1}) = (p_{r,0} + \delta_r, p_{r,0})$  under  $H_1$ .

Table 2 Comparisons of expected sample size between the bivariate two-stage randomized designs with fixed sample size and the bivariate two-stage randomized designs with curtailment based on optimal criterion,  $\delta_r = 0.2$ ,  $p_{A,s0} = p_{B,s0} = 0.7$  under  $H_0$  and  $p_{A,s1} = 0.9$ ,  $p_{B,s1} = 0.7$  under  $H_1$ .

$p_{r,0}$	$\phi$	$n$	$n_1$	$r_1$	$s_1$	$r$	$s$	$\alpha_r$	$\alpha_s$	$1 - \beta$	$E(N H_0)$	$E_c(N H_0)$	RS(%)
0.1	0.1	89	19	20	20	93	99	0.126	0.048	0.802	54.1	48.65	11.2
	0.5	93	18	19	19	97	103	0.126	0.050	0.801	58.44	52.56	11.2
	2.0	85	24	26	25	87	95	0.148	0.048	0.801	61.52	55.35	11.1
	10	88	23	25	24	90	98	0.143	0.049	0.801	61.66	55.49	11.1
0.2	0.1	89	22	23	23	94	99	0.143	0.050	0.800	59.1	52.8	11.9
	0.5	88	22	23	23	93	98	0.142	0.049	0.802	65.58	58.93	11.3
	2.0	89	21	22	22	94	99	0.141	0.049	0.801	69.34	62.37	11.2
	10	87	21	22	22	92	97	0.140	0.048	0.801	71.78	64.74	10.9
0.3	0.1	100	23	24	24	106	111	0.142	0.043	0.802	63	56.27	12.0
	0.5	93	27	28	29	99	103	0.144	0.049	0.802	70.22	62.79	11.8
	2.0	94	26	27	28	100	104	0.143	0.049	0.801	74.08	66.46	11.5
	10	90	26	27	28	96	100	0.141	0.046	0.801	76.66	69.09	11.0
0.4	0.1	99	24	24	26	106	109	0.145	0.049	0.800	63.72	56.71	12.4
	0.5	97	24	24	26	104	107	0.143	0.048	0.800	70.64	63.23	11.7
	2.0	94	24	24	26	101	104	0.141	0.047	0.801	75.84	68.18	11.2
	10	90	24	24	26	97	100	0.138	0.045	0.801	79	71.34	10.7
0.5	0.1	91	20	20	21	98	101	0.139	0.050	0.801	64	57.12	12.0
	0.5	95	24	24	26	102	105	0.145	0.047	0.801	69.98	62.6	11.8
	2.0	92	24	24	26	99	102	0.143	0.046	0.801	75.14	67.53	11.3
	10	91	23	23	25	98	101	0.142	0.045	0.801	78.64	70.96	10.8
0.6	0.1	100	21	22	22	106	111	0.149	0.042	0.800	61.82	54.89	12.6
	0.5	89	22	23	23	95	99	0.146	0.050	0.801	66.42	59.32	12.0
	2.0	90	21	22	22	96	100	0.144	0.050	0.801	71.58	64.19	11.5
	10	89	24	25	26	95	99	0.149	0.044	0.801	75.72	68.33	10.8
0.7	0.1	89	19	20	20	95	99	0.130	0.048	0.801	56.36	49.97	12.8
	0.5	88	19	20	20	94	98	0.130	0.047	0.801	60.56	53.92	12.3
	2.0	86	19	20	20	92	96	0.128	0.046	0.801	65.92	58.94	11.8
	10	78	21	21	23	84	87	0.144	0.049	0.800	69.68	62.72	11.1

1. For each  $p_{r,0}$ , values in the first row to fourth row are calculated by  $\phi = 0.1, 0.5, 2, 10$ , respectively, under  $(\alpha_r, \alpha_s, \beta) = (0.15, 0.05, 0.2)$ .
2. The response rates for two group are  $(p_{A,r0}, p_{B,r0}) = (p_{r,0}, p_{r,0})$  under  $H_0$  and  $(p_{A,r1}, p_{B,r1}) = (p_{r,0} + \delta_r, p_{r,0})$  under  $H_1$ .

105年度專題研究計畫成果彙整表

計畫主持人：陳嘉民			計畫編號：105-2118-M-343-001-				
計畫名稱：兩相依二元反應變項之縮減二階段隨機設計							
成果項目			量化	單位	質化 (說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等)		
國內	學術性論文	期刊論文		0	篇		
		研討會論文		0			
		專書		0	本		
		專書論文		0	章		
		技術報告		0	篇		
		其他		0	篇		
	智慧財產權及成果	專利權	發明專利	申請中	0	件	
				已獲得	0		
			新型/設計專利		0		
		商標權		0			
		營業秘密		0			
		積體電路電路布局權		0			
		著作權		0			
		品種權		0			
		其他		0			
	技術移轉	件數		0	件		
		收入		0	千元		
	國外	學術性論文	期刊論文		0	篇	
			研討會論文		0		
			專書		0	本	
專書論文			0	章			
技術報告			0	篇			
其他			0	篇			
智慧財產權及成果		專利權	發明專利	申請中	0	件	
				已獲得	0		
			新型/設計專利		0		
		商標權		0			
		營業秘密		0			
		積體電路電路布局權		0			
		著作權		0			
		品種權		0			
其他		0					

	技術移轉	件數	0	件	
		收入	0	千元	
參與計畫人力	本國籍	大專生	2	人次	訓練學生文獻搜尋的能力，定期蒐集資料，查閱期刊，以獲取最新臨床研究的相關知識。
		碩士生	1		在計畫執行的過程中，透過相關臨床試驗文獻的研讀與部分計畫內容的執行，培養學生具備臨床試驗設計的概念，未來可從事臨床試驗相關研究與工作。
		博士生	0		
		博士後研究員	0		
		專任助理	0		
	非本國籍	大專生	0		
		碩士生	0		
		博士生	0		
		博士後研究員	0		
		專任助理	0		
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。)					

## 科技部補助專題研究計畫成果自評表

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（簡要敘述成果是否具有政策應用參考價值及具影響公共利益之重大發現）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（請於其他欄註明專利及技轉之證號、合約、申請及洽談等詳細資訊）

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200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）

在癌症新藥的第二期臨床試驗中，使用劑量的決定通常是根據第一期臨床試驗的結果。然而，在第一期臨床試驗中，由於試驗過程中所使用的受試者人數通常不多，導致劑量的估算容易產生偏差，因此基於受試者安全性的考量，於第二期臨床試驗設計中增加安全性的評估，有其必要性。本研究針對第二期隨機對照臨床試驗，在合併安全性的考量下，發展一個可適用於兩相依二元反應變項的縮減二階段隨機對照設計，其特點是研究者可根據目前的試驗結果，提前終止試驗並做出決策，而毋須等到所有的受試者均完成試驗，進而加快臨床試驗的進行，縮短臨床試驗所需的時間，進而有助於新藥臨床試驗的發展。

4. 主要發現

本研究具有政策應用參考價值： 否  是，建議提供機關

（勾選「是」者，請列舉建議可提供施政參考之業務主管機關）

本研究具影響公共利益之重大發現： 否  是

說明：（以150字為限）