



癌症藥品第一期臨床試驗設計與分析

林資荃¹

前言

新藥的發展過程中，第一期臨床試驗主要目的是評估試驗藥品的安全性、藥理及在人體的耐受性，然第一期臨床試驗通常假設藥品劑量越高、療效越好；但相對劑量越高、毒性也越強；因此，我們希望在可以接受的耐受性前提下，找出最高劑量，此劑量即稱為最大耐受劑量(maximum tolerated dose, MTD)。此劑量也將作為第二期臨床試驗的參考劑量。

研究人員執行第一期臨床試驗可能面臨到一些操作上問題，如劑量限制毒性(Dose-Limiting Toxicity, DLT)如何定義、起始劑量如何決定、劑量增量的間隔如何決定、每個劑量所納入的受試者數目與劑量增量(dose escalation)所採用的方法等。

有關劑量限制毒性之定義可以參考美國 National Cancer Institute (NCI)所頒布的 Common Terminology Criteria for Adverse Events (CTCAE)準則^[1]，此準則將 AE 分為 5 個等級，等級越高代表越嚴重，對於非血液性毒性(non-hematological toxicities) DLT 通常定義為大於或等於 3 的等級，而對於血液性毒性(hematological toxicities) DLT 通常定義為大於或等於 4 的等級。

起始劑量選取上通常會依據與人類生理反應相似的物種，未造成任何不良反應(no-observed-adverse-effect level, NOAEL)的劑量，來估算人體暴露量。若無暴露量可計算時，一般可以採用最合適的物種重複劑量試驗中 NOAEL 的 1/10 作為起始劑量的預估值。對於起始劑量選取更完整的介紹可以參考 US FDA guidance (2005)「Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers」^[2]與 ICH S9「Nonclinical Evaluation for Anticancer Pharmaceuticals」準則^[3]。

常見劑量增量的間隔是採用改良式費式數列(modified Fibonacci sequence)，藥品增量的比例依次為 100%、67%、50%、40%、33%、33%、33%...的方式。舉例來

¹ 財團法人醫藥品查驗中心新藥科技組



說，若起始劑量為 5 mg、下一個劑量則為 5×2 mg (10 mg)、再下一個劑量為 10×1.67 mg (16.7 mg)，依此數列增量，計算出所有劑量。對於不同試驗藥品，劑量增量的間隔可能有所不同，如癌症疫苗相關臨床試驗，劑量增量可能為指數倍數方式增加(Kaufman et al., 2006)^[4]，因此研究者應依據所要研究的藥品特性，來決定劑量增量的間隔。每個劑量所納入的病人數目通常為 1-3 名，宜採用依序收案的設計，即同一劑量之試驗藥品投予受試者，經觀察安全無慮後，再提高劑量投予其他受試者，依此類推。

劑量增量的試驗設計主要為規則設計 (rule-based designs) 與模型設計 (model-based designs) 這兩大類。常見的規則設計為 3+3 設計，而常見模型設計包括連續重評估方法 (continual reassessment method, CRM, O'Quigley et al., 1990)^[5]、連續重評估方法合併藥品過量控制 (CRM with overdose control, Babb et al., 1998)^[6] 與劑量增量基於毒性區間 (Escalation based on toxicity intervals, Neuenschwander et al., 2008)^[7]。上述所介紹模型分析方法皆可以透過統計 R 軟體套裝程式 Bayesian Continual Reassessment Method (bcrm) 來完成，網址如下：<http://CRAN.R-project.org/package=bcrm>。本文章第二章節將針對上述提及的劑量增量設計作完整介紹，並比較其優缺點。

劑量增量設計方法

(一) 3+3 設計

此設計是常見的劑量增量設計，其方法為從最低劑量開始收納 3 位受試者，每個劑量至多收納 6 位受試者，MTD 定義為，6 位受試者中有 1 名或 0 名經歷 DLT 之最高劑量，3+3 設計操作如下(圖一)。

- I. 若 3 位受試者中，0 位受試者產生 DLT，則提升一個劑量水準。
- II. 若 3 位受試者中，只有 1 位受試者出現 DLT，則在此劑量多增加 3 位受試者。
 - (i) 若 6 位受試者中，只有 1 位受試者出現 DLT，則提升一個劑量水準。
 - (ii) 若 6 位受試者中， ≥ 2 位受試者出現 DLT，則降低一個劑量水準。
- III. 若 3 位受試者中， ≥ 2 位受試者產生 DLT，則降低一個劑量水準。



IV. 降低一個劑量水準時：

- (i) 若此劑量已納入 6 位受試者，則此劑量為 MTD。
- (ii) 若此劑量已納入 3 位受試者，則再納入 3 位受試者。
 - 6 位受試者中， ≤ 1 位受試者出現 DLT，則此劑量為 MTD。
 - 6 位受試者中， ≥ 2 位受試者出現 DLT，則降低一個劑量水準。

V. 若最低劑量 3 位受試者中 ≥ 2 位受試者產生 DLT，則此劑量視為無法忍受 (intolerable) 且停止試驗，宣稱 MTD 不存在。

VI. 若最高劑量 3 位受試者中皆無受試者產生 DLT，則可以再收納 3 位受試者作為確認此劑量，若 6 位受試者中 0 或 1 個受試者產生 DLT，則此試驗最高劑量為 MTD。

3+3 設計的好處為容易操作與實施，設計上亦不需要了解統計方法學，然缺點為：
 (1) 估計出來 MTD 並不精確，因為每個劑量最多僅有 6 位受試者，所以計算出來信賴區間相當寬，變異程度相當大，如此大的變異將造成 MTD 決定上有很大的不精確性(表一)；
 (2) 此設計所找到的 MTD 多會低於真正的 MTD，因此會讓大部分受試者接受到一個無效的劑量(ineffective dose level)，若受試者為嚴重疾病族群，則更耽誤到受試者治療時程；及(3) 此設計僅利用最近的世代(cohort)結果來決定 MTD，此概念與歐盟 2017 所頒布準則「Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products」^[8] 建議採用所有世代結果來決定 MTD 概念有所不同。

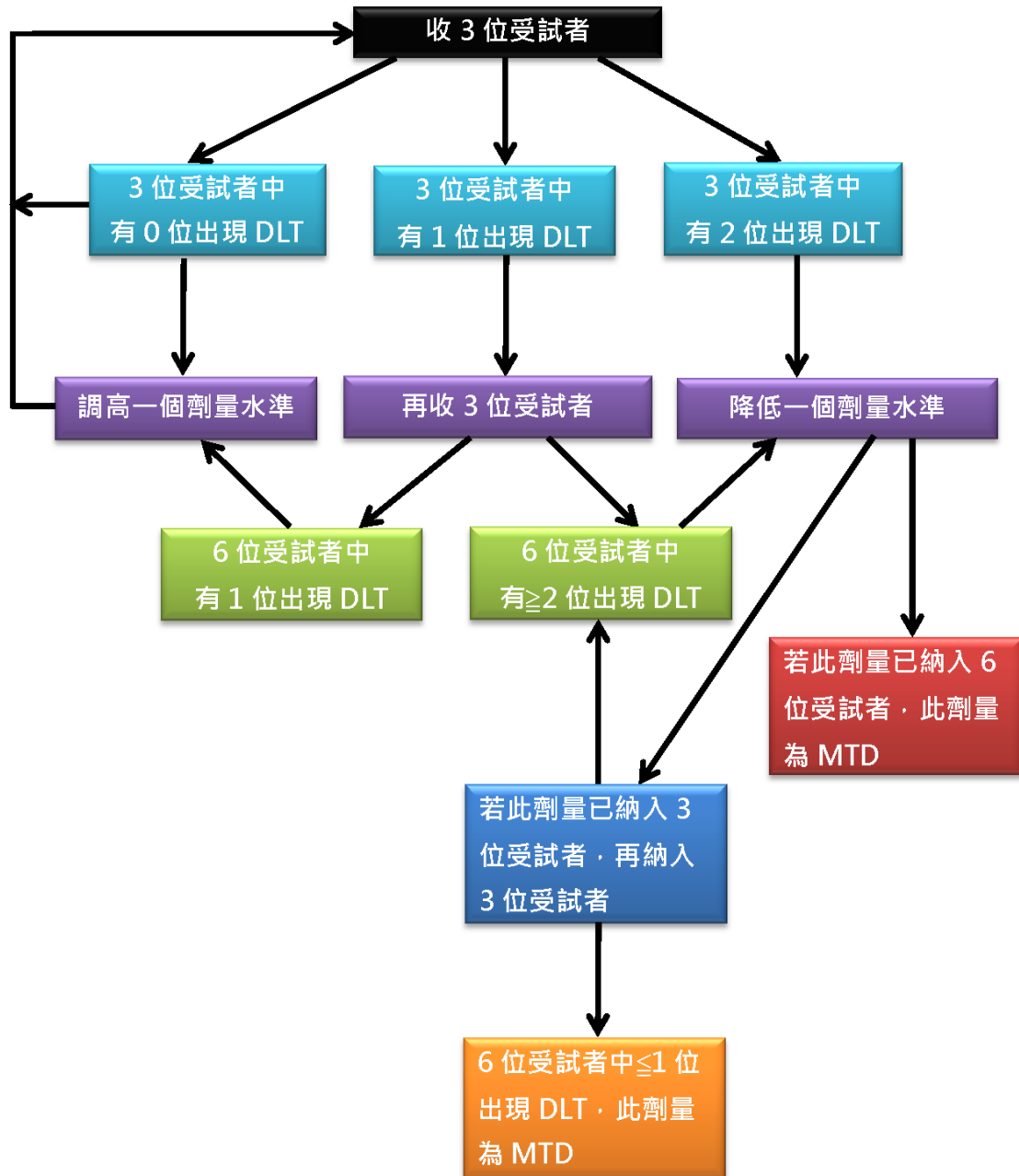
表一、DLT 之 95% 確切(exact)信賴區間

| 產生 DLT 人數 | 受試者人數 | 95% CI |
|-----------|-------|--------------|
| 0 | 6 | 0.000, 0.459 |
| 1 | 6 | 0.004, 0.641 |
| 2 | 6 | 0.043, 0.777 |
| 3 | 6 | 0.118, 0.882 |
| 4 | 6 | 0.223, 0.957 |
| 5 | 6 | 0.359, 0.996 |
| 6 | 6 | 0.541, 1.000 |



致力法規科學
守護生命健康

Regulatory Science, Service for Life



圖一、3+3 設計流程圖



由於 3+3 設計有上述的缺點，因此發展出藥品增量模型設計，此類設計是利用貝氏統計(Bayesian statistic)方法，藉由事前分配(prior distribution)的決定和受試者是否觀察到 DLT，計算出事後分配(posterior distribution)，事後分配將作為下一個受試者劑量選取之依據。目前常被使用模型設計為：(1)連續重評估方法^[5]；(2)連續重評估方法合併藥品過量控制^[6]；及(3)劑量增量基於毒性區間^[7]，分別介紹如下。

(二) 連續重評估方法(continual reassessment method, CRM)^[5]

連續重評估方法^[5]包括六個重要的步驟，茲整理如下：

步驟一：

決定劑量毒性曲線(dose-toxicity curve)模型：連續重評估方法^[5]利用非遞減函數曲線來描述劑量與毒性間相關性，目前常被使用的函數曲線有：(1)幂曲線(Power curve)；(2)雙曲正切函數曲線(hyperbolic tangent curve)；(3)單參數邏輯斯曲線(one parameter logistic curve)；及(4)雙參數邏輯斯曲線(two-parameter logistic curve)。四種不同曲線圖型如圖二所示，其中幂曲線、雙曲正切函數曲線與單參數邏輯斯曲線皆屬於單參數曲線(one-parameter curve)模型，而雙參數邏輯斯曲線屬於雙參數曲線(two-parameter curve)模型。雙參數曲線模型較單參數曲線模型多增加一個參數，可以更精確描述劑量與毒性間關係，進而提高找到正確 MTD 機率。

一旦決定劑量毒性曲線後，仍需要給定曲線中參數的分配，常用的分配為伽瑪(Gamma)、均勻(Uniform)或對數常態(Lognormal)分配。一般而言，研究人員可以依據臨床前動物試驗結果數據，來決定所欲採用的曲線與曲線中參數的分配。

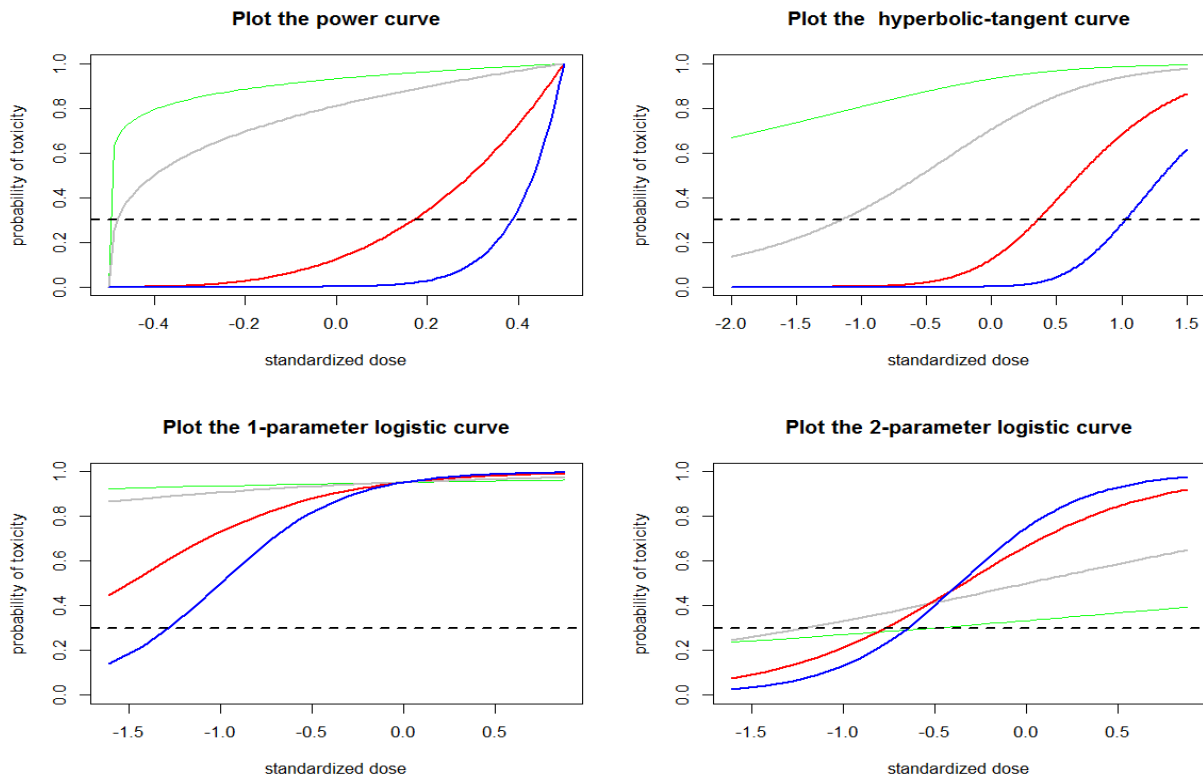
步驟二：

決定劑量增量的間隔：例如採用改良式費式數列或依據試驗藥品的特性，來決定所採用的藥品增量的間隔。



致力法規科學
守護生命健康

Regulatory Science, Service for Life



圖二、四種常用劑量毒性曲線圖

步驟三：

決定每個世代所需要的受試者數目：連續重評估方法^[5]，每個世代僅有 1 名受試者，但若要縮短整個試驗時程與降低試驗計算上的次數，改良式連續重評估方法(modified continual reassessment method, mCRM, Goodman et al. 1995)^[9]，可以允許每個世代至多收納到 3 名受試者。

步驟四：

決定第一個世代受試者所使用劑量與劑量增加法則：連續重評估方法^[5]讓第一個世代受試者接受先前猜測 MTD 的劑量，且允許跳躍式劑量增量。改良式連續重評估方法^[9]基於安全性考量，第一個世代受試者接受最低的試驗劑量，且不允許跳躍式劑量增量。



步驟五：

決定目標毒性率(target toxicity rate)：試驗進行前，研究人員須在計畫書中載明目標毒性率，一般來說，若試驗藥品毒性較強，則目標毒性率可以訂較高；而如果試驗藥品毒性較弱，或是需要長期服用，則目標毒性率可以訂較低。如抗高血壓藥品的試驗，其目標毒性率可能訂為 0.2；而抗癌藥品的試驗，其目標毒性率可能訂為 0.3。

步驟六：

決定試驗終止法則：連續重評估方法^[5]的試驗終止法則，為當病人數已達到事先設定的數目時，則試驗終止。改良式連續重評估方法^[9]的試驗終止法則設定較具彈性，如可以訂為至少要有 9 位受試者，接受到 MTD 劑量或是考量 MTD 之精確度，終止法則可以訂為尋找到的 MTD 劑量，其 95%信賴區間須要落在某個範圍內，如：0.2 到 0.4 間。不同的試驗終止法則，可能找到的 MTD 有所不同，一般來說，若試驗要求的精確度越高，則所需要的樣本數也相對較多，研究人員可以依據不同的模擬(simulation)方案，與經費預算考量來決定所採用的試驗終止法則。

(三) 連續重評估方法合併藥物過量控制(CRM with overdose control)^[6]

由於連續重評估方法^[5]允許跳躍式劑量增量，造成受試者接受到過高劑量的機率提高，於是 Babb 等人^[6]於 1998 提出，連續重評估方法合併藥物過量控制方法。在每一個世代的受試者接受藥物劑量後，都會計算出下一個世代的受試者接受劑量過量的機率，一旦機率超過可行性界限(feasibility bound)後，則藥物將不會再提高劑量，一般可行性界限都定在 25%。此方法所找到的 MTD 相較於連續重評估方法^[5]更為保守。

(四) 劑量增量基於毒性區間(Escalation based on toxicity intervals)^[7]

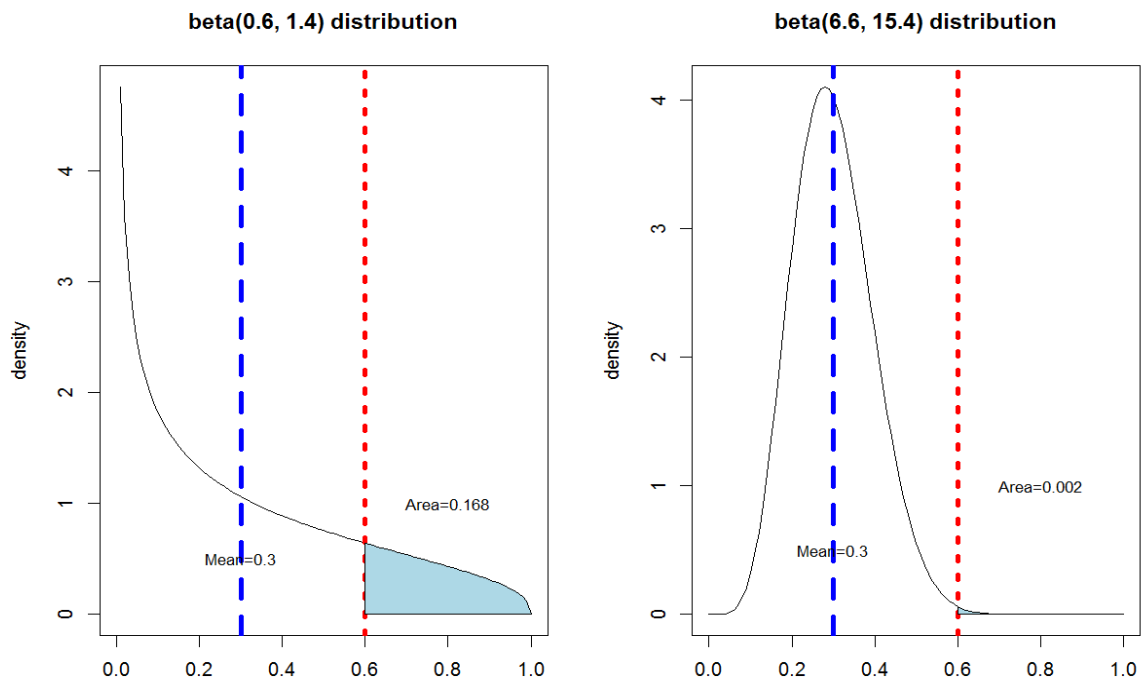
上述所提及連續重評估方法^[5]與連續重評估方法合併藥物過量控制^[6]，都是藉由事前分配的決定和受試者是否發生劑量毒性，計算出每個劑量的事後分配，透過每個劑量事後分配平均值的比較，來決定下一位受試者所接受劑量。然透過平均值來決定下一個病人所接受劑量，有時並不是一個最好的方法，舉例來說，假設某一個劑量之事後分配服從 beta 分配，由圖三可以看出 beta (0.6,1.4)與 beta (6.6,15.4)分配平均值皆為 0.3，但 beta (0.6, 1.4)分配產生 DLT rate 大於 0.6 機率為 0.168，然 beta (6.6,15.4)分配產



致力法規科學
守護生命健康

Regulatory Science, Service for Life

生 DLT rate 大於 0.6 機率為 0.002。如果採用事後分配平均值作為選取標準，可能選到 beta (0.6,1.4)或 beta (6.6,15.4)，但若採用劑量增量基於毒性區間方法，可能就只會選到 beta (6.6,15.4)分配，因此，採用劑量增量基於毒性區間方法，可以大大降低讓受試者暴露到過高毒性的機率。



圖三、beta (6.6,15.4)與 beta (0.6,1.4)

劑量增量基於毒性區間方法，由 Neuenschwander 等人於 2008 所提出^[7]，其作法如下，首先將(0,1)區間分成 4 等分，

- 過低毒性區間 (Under toxicity interval) : (0, 0.20]
- 目標毒性區間 (Target toxicity interval) : (0.20, 0.35]
- 過多毒性區間 (Excessive toxicity interval) : (0.35, 0.60]
- 無法接受毒性區間 (Unacceptable dosing interval) : (0.60, 1.00]

給予每個區間一個損失值，如落於過低毒性區間，則給予損失值為1；落於目標毒性區間，給予損失值為0；落於過多毒性區間，給予損失值為1；落於無法接受毒性區間，給予損失值為2。對於每個劑量，其所對應貝氏風險(Bayes risk)為：



$$1 \times P(\pi(d) \in (0, 0.20]) + 0 \times P(\pi(d) \in (0.20, 0.35]) \\ + 1 \times P(\pi(d) \in (0.35, 0.60]) + 2 \times P(\pi(d) \in (0.60, 1.00])$$

其中 $\pi(d)$ 為劑量 d 的事後分配機率密度函數。此方法是經由計算出每個劑量之貝氏風險，找出具最小貝氏風險的劑量，此劑量即為下一個世代受試者的使用劑量。採用此方法須要事先定義好每個區間的範圍與損失值，損失值定義：原則上為距離目標毒性區間越遠，所給予的損失值越大，如：落在無法接受毒性區間之損失值，應大於落在過多毒性區間之損失值。另外，若不想讓受試者曝露到過高的機率，則可以提高落在過多毒性區間與無法接受毒性區間的損失值。

除了上述介紹的方法外，仍有許多模型設計方法如Extended CRM (Moller, 1995)^[10]、Practical CRM (Piantadosi 與 Liu, 1996)^[11]與 modified toxicity probability interval 2 (mTPI2) (Guo et al., 2017)^[12]等方法，這些方法都可以供研究者試驗設計上的參考。然，模型設計方法並無法證明方法之間的優越性，因此研究人員在方法選取上，應考量臨床前的試驗數據及針對不同劑量毒性曲線、不同劑量增量的間隔、是否合併藥品過量控制、是否納入毒性區間考量等，執行統計模擬分析，找出最適當的方案，並作為試驗劑量增量的最佳模型。

結語

癌症藥品第一期臨床試驗主要目的，為找出人體可以接受MTD，此劑量也將作為第二期臨床試驗的參考劑量。然第一期臨床試驗所收納受試者人數通常不多，此情況下若採用傳統3+3設計較無法精確找到MTD，因此，近年來國際大廠紛紛採用模型設計方法，如本文所提及連續重評估方法^[5]、連續重評估方法合併藥品過量控制^[6]與劑量增量基於毒性區間^[7]等設計，這些設計都需要事先決定劑量毒性曲線與曲線中參數的分配，至於如何選取適當曲線與參數的分配，可以透過藥品的臨床前試驗及執行統計模擬的方式，選取出最適當的模式，進而更精準找到MTD。

最後，希望透過本文章介紹，期能協助廠商及研究人員針對預設之目標毒性率，透過模型設計，精確及有效地找到試驗藥品之MTD，作為後續臨床試驗劑量選取之依據，如此才更有機會，讓試驗繼續往下個階段進行，且提昇臨床試驗的成功率，以讓我國的



生醫產業能夠與國際接軌。

參考文獻

1. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0; Published: May 28, 2009. (v4.03: June 14, 2010) by National Cancer Institute.
2. US FDA guidance: Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers, July 2005.
3. International Conference on Harmonisation: "ICH Topic S9: Nonclinical Evaluation For Anticancer Pharmaceuticals" , June 2018.
4. Kaufman HL, Cohen S, Cheung K, DeRaffele G, Mithcam J, Moroziewicz D, Schlom J, Hesdorffer C (2006). Local delivery of vaccinia virus expressing multiple costimulatory molecules for the treatment of established tumors. *Human Gene Therapy*, 17:239–244.
5. O'Quigley J, Pepe M, Fisher L (1990). Continual Reassessment Method: a practical design for Phase I clinical trials in cancer. *Biometrics* 46, 33-48.
6. Babb J, Rogatko A, Zacks S (1998). Cancer phase I clinical trials: Efficient dose escalation with overdose control. *Statistics in Medicine*, 17: 1103-1120.
7. Neuenschwander B, Branson M, Gsponer T (2008). Critical Aspects of the Bayesian Approach to Phase I Cancer Trials. *Statistics in Medicine*, 27(13), 2420-2439.
8. EMA guidance: Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products, July 2017.
9. Goodman SN, Zahurak ML, Piantadosi S (1995). Some Practical Improvements in the Continual Reassessment Method for Phase I Studies. *Statistics in Medicine*, 14:1149-1161.



10. Moller S (1995). An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Statistics in Medicine*. May 15-30;14: 911-22; discussion 923.
11. Piantadosi S, Fisher JD, Grossman S (1998). Practical implementation of a modified continual reassessment method for dose-finding trials. *Cancer Chemother Pharmacol*. 41: 429-36.
12. Guo W, Wang SJ, Yang S, Lynn H, Ji Y (2017). A Bayesian interval dose-finding design addressing Ockham's razor: mTPI-2. *Contemporary Clinical Trials*. 58: 23-33.