

**ICH E10：臨床試驗中對照組的選擇與相關事項**

**CHOICE OF CONTROL GROUP AND  
RELATED ISSUES IN CLINICAL TRIALS**

**衛生福利部食品藥物管理署**

**中華民國 113 年 12 月**

# 目錄

第一章、緒論.....	7
1.1、指引概述與目標.....	7
1.2、對照組之目的.....	8
1.2.1、隨機分派 .....	8
1.2.2、盲性 .....	8
1.3、對照組的種類.....	9
1.3.1、安慰劑同步對照 .....	9
1.3.2、無治療同步對照 .....	10
1.3.3、劑量反應同步對照 .....	10
1.3.4、活性（陽性）同步對照 .....	10
1.3.5、外部對照（含歷史對照） .....	10
1.3.6、多重對照組 .....	11
1.4、臨床試驗的目標與相關事項.....	11
1.4.1、療效證明 .....	11
1.4.2、療效與安全性的比較 .....	11
1.4.3、比較之公平性 .....	12

1.4.3.1、劑量 .....	12
1.4.3.2、病人族群 .....	12
1.4.3.3、試驗指標的選擇與時機 .....	12
1.5、檢測靈敏度 .....	13
1.5.1、不劣性或相等性試驗之檢測靈敏度 .....	13
1.5.1.1、對藥品療效具檢測靈敏度之歷史證據，與不劣性臨界值之選擇 .....	14
1.5.1.2、適當的試驗進行方式 .....	15
1.5.2、以展現較優性為目的之試驗的檢測靈敏度 .....	16
第二章、關於對照類型的詳細考量 .....	17
2.1、安慰劑對照 .....	17
2.1.1、說明（見第 1.3.1 節） .....	17
2.1.2、將偏差最小化的能力 .....	17
2.1.3、倫理道德考量 .....	17
2.1.4、安慰劑對照試驗之適用性與在特定情況下之推論強度 .....	18
2.1.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題 .....	18

2.1.5.1、說明（見第 1.3.1 節） .....	19
2.1.5.1.1、三臂試驗：安慰劑與活性對照 .....	19
2.1.5.1.2、額外劑量 .....	19
2.1.5.1.3、因子設計(Factorial design) .....	19
2.1.5.2、試驗設計其他方面的調整 .....	19
2.1.5.2.1、安慰劑對照附加性試驗；替代試驗 .....	19
2.1.5.2.2、「及早退出」；救援性治療 .....	20
2.1.5.2.3、有限的安慰劑治療期 .....	20
2.1.5.2.4、隨機退出 .....	20
2.1.5.2.5、其他設計上的考量 .....	21
2.1.6、安慰劑對照試驗的優點 .....	21
2.1.6.1、展現療效的能力 .....	21
2.1.6.2、測量「絕對」療效與安全性 .....	21
2.1.6.3、效率 .....	22
2.1.6.4、將受試者與試驗主持人的期望的影響最小化 .....	22
2.1.7、安慰劑對照試驗的缺點 .....	22
2.1.7.1、倫理道德考量（見第 2.1.3、2.1.4 章） .....	22

2.1.7.2、病人與醫師的實務考量 .....	22
2.1.7.3、概推性(Generalizability) .....	22
2.1.7.4、無比較的資料 .....	23
2.2、無治療同步對照（見第 1.3.2 節） .....	23
2.3、劑量反應同步對照（見第 1.3.3 節） .....	23
2.3.1、說明 .....	23
2.3.2、將偏差最小化的能力 .....	24
2.3.3、倫理道德考量 .....	24
2.3.4、安慰劑對照試驗之適用性與在特定情況下之推論強度.....	24
2.3.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題.....	24
2.3.6、劑量反應試驗的優點 .....	24
2.3.6.1、效率 .....	24
2.3.6.2、可能存在的倫理道德優勢 .....	24
2.3.7、劑量反應試驗的缺點 .....	25
2.4、活性對照（見第 1.3.4 節） .....	25
2.4.1、說明 .....	25

2.4.2、將偏差最小化的能力 .....	25
2.4.3、倫理道德考量 .....	26
2.4.4、安慰劑對照試驗之適用性與在特定情況下之推論強度 .....	26
2.4.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題.....	26
2.4.6、劑量反應試驗的優點 .....	26
2.4.6.1、倫理道德與實務上的優點 .....	26
2.4.6.2、資訊的內容 .....	26
2.4.7、劑量反應試驗的缺點 .....	27
2.4.7.1、資訊的內容 .....	27
2.4.7.2、龐大樣本數.....	27
2.5、外部對照（含歷史對照，見第 1.3.5 節） .....	27
2.5.1、說明 .....	27
2.5.2、將偏差最小化的能力 .....	28
2.5.3、倫理道德考量 .....	28
2.5.4、安慰劑對照試驗之適用性與在特定情況下之推論強度 .....	29

2.5.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面 的問題.....	29
2.5.6、劑量反應試驗的優點 .....	29
2.5.7、劑量反應試驗的缺點 .....	30
第三章、同步對照組的選擇.....	30
表 1、特定同步對照類型在各種情形下之適用性 .....	31
圖 1、選擇同步對照以展現療效.....	32

# ICH E10：臨床試驗中對照組的選擇與相關事項

## 第一章、緒論

設計臨床試驗時，選擇適當的對照組至關重要。對照組的選擇會直接影響到根據試驗結果所能做出的推論，其影響所及還包括該試驗在倫理道德上的可接受性、在執行試驗與分析過程中偏差可降低的程度、收案族群以及收案程度、試驗適用的指標（endpoint），以及其他執行和結果解讀等。最終試驗結果的公信力與科學上的可靠性、法規主管機關對試驗結果的接受度，都與對照組的選擇有關。

### 1.1、指引概述與目標

本指引的目標主要是說明在探討治療療效的臨床試驗中，選擇對照組時所應依循的一般原則，並說明試驗設計與執行方面的相關問題。本指引並未討論任一區域的法規要求，而主要是說明試驗在依循建議原則時可呈現出的意義。本指引所述之一般原則可適用於任何對照性的試驗，但對照組選擇對於在藥品開發過程中探討藥品療效的臨床試驗尤為重要。選擇對照組時，須將現有的標準治療方法、是否有充分證據可支持所選試驗設計、倫理道德等面向均納入考量。

本指引首先說明設置對照組的目的，以及在呈現療效時常見的對照組類型。其後，對於使用活性對照組(active control)以不劣性（non-inferiority）或相等性（equivalence）來呈現療效的臨床試驗，討論試驗設計與解讀時相關之重要議題（第 1.5 章）。在某些情況下，即便試驗結果呈現不劣性，亦不足以作為藥品具療效的證據。明確而言，若欲將呈現不劣性的試驗結果作為藥品具療效的證據，該試驗本身必須具備能區別有療效和無療效治療之間差異的能力。

本指引接著進一步說明每一種對照組類型（見第 2-2.5.7 節），並針對各類型討論下列面向：

- 減少偏差的方法
- 試驗所涉及之道德上與實務上的議題
- 試驗的適用性以及在該情境下所作出之推論的強度
- 可以解決倫理道德、實務、推論等方面疑慮的方法，包括調整試驗設計或納入其他的對照組
- 試驗整體之優缺點

下列 ICH 指引與本指引具高度相關性：

- E3：藥品臨床試驗報告之格式及內容基準（Structure and Content of Clinical Study Reports）
- E4：支持藥品上市之劑量反應試驗資料（Dose-Response Information to Support Drug Registration）
- E5：接受國外臨床資料之族群因素（Ethnic Factors）
- E6：藥品優良臨床試驗準則（Good Clinical Practice: Consolidated Guideline）
- E8：臨床試驗一般性原則（General Considerations for Clinical Trials）
- E9：臨床試驗統計指導原則（Statistical Principles for Clinical Trials）

雖然採用任何本指引中提及的對照組試驗設計，至少能在某些情境下適用於查驗登記為目的之臨床試驗，但這些設計並非在所有情境下均同等適用或有用。選擇對照組類型時的考量方式請見第3節、圖1、表1。

雖然本指引主要聚焦在用於評估療效的臨床試驗，但文中所討論的許多考量亦可適用於以評估安全性為主要目的之臨床試驗，以及適用於比較兩種治療方法的安全性與療效之試驗。

## 1.2、對照組之目的

設置對照組的主要目的，是為了能夠觀察接受所研究的治療相對於非研究治療之下的療效結果（例如症狀、病徵、其他病症的變化）是否有差別，非研究治療包括其他治療、疾病的自然進程、觀察者或病人期望的心理因素所產生的影響等。對照組的經驗可以告訴我們，若病人並未接受試驗所研究的治療，抑或是接受另外一種已知有效的治療方法時，將會發生什麼情形。

若一項疾病在特定病人族群中，進程十分一致，或是可透過病人特徵準確預測受試者族群的病程，則試驗的療效結果便可與過去已知的結果相比較。舉例而言，我們可以假設疼痛會持續一段特定的時間、血壓不會發生變化、憂鬱症將會持續一段特定的時間、腫瘤將會持續惡化、急性心肌梗塞的死亡率將會與先前所見相同等等。在罕見的情況，特定病人族群的病程其實可以預測，因此可以將先前已研究過的相似病人族群作為試驗的比較對象，即歷史性對照組 (historical control) (見第1.3.5節)。然而，在多數情形中，因為不可能對病程作出足夠準確或可靠的預測，故仍必須設置同步對照組。

所謂同步對照組係指受試者與試驗組來自同一族群，在同一時期內接受該試驗所規定的治療方式。試驗組與對照組在基期 (baseline) 和治療期間，所有可能影響結果的因子都應盡可能相近，僅在測試藥物有所差異。兩組的因子若不相近，試驗可能產生偏差。此處偏差 (bias，用法同於《ICH E9》指引) 係指臨床試驗在設計、執行、分析、結果解讀等方面，系統性地導致治療效果估計值偏離真實情況的現象。隨機分派與盲性是降低偏差的常用方法，可確保試驗組與對照組於試驗開始時立足點相近，且於試驗過程中接受相似的治療方法 (見《ICH E9》)。試驗設計中是否有納入隨機分派與盲性的作法，是影響試驗品質與說服力 (persuasiveness) 的關鍵。

### 1.2.1、隨機分派

確保試驗組與對照組族群相近的最佳做法，便是從同一個族群中隨機將受試者分配至試驗組與對照組。隨機分派可以避免各組之間於可能影響結果的已知或未知因子有系統性差異。對於未設置隨機分配同步對照組的試驗，無法消除各組之間的系統性差異，便是一大問題 (見第1.3.5節：外部對照試驗)。此外，隨機分派亦能使統計結論更加可靠。

### 1.2.2、盲性

各組不僅須在基期時相近，在試驗期間所接受的觀察和治療亦應當相近，僅在主要治療藥物上有測試藥物與對照藥物之分。臨床試驗常有雙盲 (double-blind，亦作 double-masked) 的設計，意即試驗受試者與試驗主持人，以及試驗委託者和參與給予治療和療效評估的試驗相關人員，均不知受試者分派到何種治療藥物。當受試者或試驗主持人知道主要治療用藥時，病

人處置、治療方法、評估病情以及療效解讀等方面均可能因此有所差異，盲性的目的便在於將於此偏差降至最低。茲舉以下數例：病人

- 受試者若知道自己正接受被認為有活性的藥物 (active drug)，可能因為預期有效益而在療效評估時報告較佳的結果，抑或是因為知道自己正在接受活性藥而較有意願繼續留在試驗。
- 療效評估者有可能比較不會注意到無治療 (no treatment) 組的療效反應，而對於接受活性藥物的病人，則較容易注意到有利結果或不良事件。病人
- 了解治療分配情況可能影響收集試驗中數據或追蹤數據的努力程度。
- 了解治療方法可能影響是否讓受試者繼續接受治療或是否給予合併藥物 (concomitant medication) 或其他輔助療法的決定。
- 了解治療分配情況可能影響是否將某些受試者結果納入分析的決定。
- 了解治療分配情況可能影響統計分析方法的選擇。

盲性可以確保主觀的評估與決策不會受到了解治療分配情況所影響。

### 1.3、對照組的種類

臨床試驗的對照組可依兩種特性作分類：(1) 治療類型、(2) 決定如何納入對照組的方法。對照組的治療類型可以是以下四類當中的任一類：(1) 安慰劑、(2) 無治療、(3) 與試驗組不同的治療劑量或給藥方案、(4) 以另一不同活性成分的藥物進行治療。決定納入對照組主要有兩種方法，一是隨機分派，一是從不同於試驗族群的另一族群來選擇（外部對照或歷史對照）。本指引將對照組分為五類。第一至四類均屬同步對照（對照組與試驗組均來自同一族群且同時接受治療），通常採隨機分派。各組間以治療類型（如前所列）作區別。外部（歷史）對照組不論所採之治療類型，統一歸於第五類，因為此類的試驗有嚴重疑慮，包括無法確保試驗組與對照組之間的可比性，以及難以降低重要偏差，使得它們唯有在特殊情形下方能採用。

目前試驗內包含多種對照組的設計逐漸增加。每一類的對照組都適用於某種特定情況，但沒有一類對照組可適用於所有情況。茲將此五類對照組分述如下：

#### 1.3.1、安慰劑同步對照

在安慰劑對照試驗中，受試者被隨機分派至接受試驗所研究的藥物，或是接受外觀相同但不含測試藥物的治療方式。測試藥物的治療可以隨治療效果或耐受性進行調整，也可以給予一個或多個固定劑量。此類試驗幾乎都是雙盲試驗。所謂“安慰劑對照組”目的便是控制“安慰劑”效應，即受試者因認為正在接受治療而感到改善，但這並不是唯一的好處。事實上，安慰劑對照試驗由於有盲性和隨機分派，有接受非活性治療 (inert treatment) 的對照組，恰可反映出沒有測試藥物藥理作用時、所有其他潛在因子對病程的影響。這些因子包括自發性改變（病程的自然進程及回歸至平均表現）、受試者或試驗主持人的期望、參加試驗的效應、使用其他療法、診斷與評估時主觀因素的影響等。在研究治療效果時，安慰劑對照試驗主要是期

望觀察到組別之間療效的差異，但在評估安全性時，安慰劑對照試驗也可以試圖證明組別之間安全性並無差異。在此情況下，關鍵在於如果存在差異試驗是否能夠檢測出來（見第 1.5 節）。

安慰劑對照組有時並不代表對照組完全未接受治療。許多安慰劑對照試驗中，安慰劑或是試驗所研究的藥物是附加(add-on)在一種共同標準療法之上合併使用的（即所謂的附加性試驗，見第 2.1.5.2.1 節）。

### **1.3.2、無治療同步對照**

在無治療對照試驗中，受試者被隨機分配到接受所研究的治療或不接受治療（即無治療）。這種設計與安慰劑對照試驗有一個主要區別，即受試者和試驗主持人了解分配情況，不進行雙盲。由於雙盲設計有其優點(第 1.2.2 節)，故僅在難以或不可能雙盲（例如，治療有容易於辨識出的毒性）以及試驗結果屬客觀性觀察，不易受到第 1.2.2 節中列出的因素影響時，才需要或可採用不雙盲的設計。須注意的是，即便試驗採用不雙盲設計，通常仍然可以由一位未知治療分配情況的評估者進行盲性的療效指標評估。這種做法具有價值，在無法進行雙盲的試驗中都應考慮此種作法，惟此一作法並無法解決知曉治療分配時造成的其他偏差影響。（見第 1.2.2 節）。

### **1.3.3、劑量反應同步對照**

在隨機分配、多個固定劑量階層的劑量反應試驗中，受試者被隨機分配到數個不同固定劑量組中的其中一個。受試者可以一開始就接受該固定劑量，也可以逐漸增加到該劑量，但主要的比較是在各組達到最終劑量時進行。劑量反應試驗通常採用雙盲設計，可能包括安慰劑（零劑量）和/或活性對照。在濃度對照試驗中，治療組為數個固定濃度階層；概念上與多個固定劑量的劑量反應試驗相似。而在給藥方案對照試驗中，受試者被隨機分配到測試藥物的兩種或多種給藥方案中，例如每天一次或每天兩次，3 天給藥或 7 天給藥。

### **1.3.4、活性（陽性）同步對照**

在活性對照（或陽性對照）試驗中，受試者被隨機分派至接受試驗所研究藥物或其他活性藥物治療。此類試驗通常是雙盲的，但並非總是可行；例如許多腫瘤領域的臨床試驗中，由於使用不同的給藥方案、給藥途徑（見第 1.3.2 節）以及有明顯不同的毒性反應，此時便難以採用雙盲。活性對照試驗可以具有兩種不同方式顯示測試藥物的療效：(1) 藉由顯示試驗測試藥物的療效與已知有效治療(活性對照)一樣好，以證明測試藥物的療效；或 (2) 藉由顯示試驗測試藥物的療效優於 (superiority) 活性對照，以證明其療效。此外，活性對照試驗也可以用來比較兩種治療方法的療效和/或安全性（見第 1.4 節）。無論試驗的目的是顯示新治療的療效還是比較兩種治療方法，關鍵仍在於，如果存在差異試驗是否能夠檢測出來。

### **1.3.5、外部對照（含歷史對照）**

外部對照試驗指的是將接受測試治療的受試者與試驗之外的病人進行比較，而不是與試驗內部來自相同族群但接受不同治療的對照組進行比較。外部對照包括早期接受治療的病人（歷

史對照)，或者在相同時期但在不同條件情況下接受治療的病人。外部對照可以被明確定義(特定病人族體)，亦可不作明確定義(以一般醫學知識所認定的一群病人)。然而由於一般醫學知識的認定通常不夠精確，所以使用無明確定義的外部對照較不適當。此外，所謂的基期對照(baseline-controlled)試驗是指比較受試者在治療前後的狀態，例如血壓、腫瘤大小等，這種試驗通常缺乏內部對照，因此被視為無對照或外部對照(見第 2.5 節)。

### 1.3.6、複數對照組

正如將於下文進一步說明的(請參閱第 1.5.1 節)，一項試驗也可能同時使用多種不同類型的對照組，將有助於顯示更多資訊；例如同時採用活性對照與安慰劑，或同時包含多個測試藥物的劑量和多個活性對照藥物的劑量，是否包含安慰劑則視情況而定。這種設計可以直接比較兩種活性藥物的表現，尤其當兩種藥物的相對效力(potency)尚不明確時，或者試驗的主要目的便是確定它們之間的相對效力，便可能採取多重對照的方式。

## 1.4、臨床試驗的目標與相關事項

臨床試驗的主要可區分為兩方面：(1) 評估所研究治療方法的療效及/或安全性，以及(2)評估兩種治療方法的相對療效、安全性、效益/風險關係或實用性。

### 1.4.1、療效證明

使用任何類型對照組的試驗，都可以藉由顯示試驗組優於對照組(例如，安慰劑、無治療、測試藥物的低劑量、其他有效藥物)來證明其療效。此外，活性對照試驗還可以藉由顯示新治療與已知有效治療的療效相似來證明其療效。但唯有在可以合理認定試驗中活性對照藥物是確實有效的情況下，才能確立測試藥物的療效，因為有可能兩種治療在試驗中都無法發揮療效但也同樣呈現相似的結果(見第 1.5 節)。

藉由顯示測試藥物在療效上與標準治療相似而證明療效的臨床試驗，通常被稱為"相等性"試驗，此類試驗大多屬於不劣性試驗，亦即試驗試圖證實新藥物的療效不會比對照藥物低於一特定範圍，這個事先定義的範圍通常稱為“臨界值”(margin)。

### 1.4.2、療效與安全性的比較

在某些情況下，試驗目標是將一項治療方法與另一項相比較，而不僅僅是評估試驗藥物本身的療效。視其治療領域的不同，這些試驗可以為不同治療方法的相對風險/效益評估提供重要資訊。所選擇的活性對照藥物應當符合該區域的接受程度。不一定需要證明新治療方法優於活性對照藥物，而且視情況也可能不必證明不劣性。比如說，一個效果較弱的治療方法或許會因為安全性較佳，因此仍然有其實用價值。

儘管這類試驗的主要焦點是治療方法的比較，而非僅僅是療效的證明，但有關不劣性試驗在執行和解讀時須注意的事項(參見第 1.5 節)，在此類試驗也必須仔細考慮。具體而言，比較性試驗需要具備檢測兩種治療方法是否存在差異的能力，無法區分差異的試驗無法提供有用的比較資訊。

### 1.4.3、比較之公平性

為了確保比較性試驗能提供安全性和/或療效的有用資訊，試驗必須有公平的設計；換句話說，試驗情況不應該不恰當地偏袒其中一種治療方法。實際上，活性對照的相等性或不劣性試驗，都應盡可能提供公平的比較，因為任何對於試驗中對照組是否有發揮正常療效的疑慮，都將削弱試驗的檢測靈敏度 (assay sensitivity(見第 1.5 節))。試驗設計中，可能會不公平地偏袒某一治療方法的因素，包括：劑量選擇、病人族群，以及試驗指標的選擇和時間點。

#### 1.4.3.1、劑量

在比較試驗藥物與活性對照時，選擇適當的劑量和給藥方案十分重要。在評估比較結果時，對於療效明顯較弱的治療方式，需要仔細考慮是否劑量過低，對於明顯不耐受的治療方式，需要考慮是否劑量過高。在某些情況下，為使試驗的比較結果更具說服力，可能需要將對照組分為數個劑量作檢視，或是將測試藥物分數個劑量作檢視。

#### 1.4.3.2、病人族群

在活性對照試驗中，受試者的選擇可能會影響結果；在評估試驗結果時，應仔細考慮受試者族群。舉例而言，若試驗中有許多受試者先前對活性對照藥物無效，這樣便可能存在偏袒新治療的偏差。這種試驗的結果不能外推到未曾接受治療的整個病人族群。然而，對於該試驗族群 (之前對活性對照藥物無效)而言，發現新治療較佳的結果仍然是新治療療效的證據。實際上，在明顯對另一種治療方法無效的受試者身上試驗一種新療法，並將這些受試者隨機分派至接受新療法或失敗療法 (以對病人不具風險為前提)，若顯示出新療法對這些病人的療效，仍是十分具有臨床價值的觀察結果。

同樣的，有時知道某群病人對特定藥物較可能或是或較不可能產生有效反應或不良反應。舉例而言，黑人通常對  $\beta$  受體阻斷劑與血管收縮素抑制劑的降血壓效果反應較差，因此在此族群身上將新的抗高血壓藥與這些藥物比較往往會得到新藥效果較佳的結果。然而就此結果而判定新藥對整體族群較為優秀，並不適當。無論如何，在某一子族群中進行試驗時，若能清楚認知其局限性以及可以得出的適當結論，這樣的試驗仍可以提供有用的資訊。

#### 1.4.3.3、試驗指標的選擇與時機

當兩種治療方法分別用於同一疾病或病況時，需要特別注意它們可能會對該疾病的各種結果有不同程度的影響，尤其是兩種療法分別屬於不同的作用機轉或治療類別時。因此，臨床試驗中治療指標的選擇和時間安排，可能會偏袒其中一種治療方法。舉例而言，對於急性心肌梗塞病人，溶血栓藥可以降低死亡率，但是也會提高出血性中風的風險。如果將一種新的、藥理作用較強的溶血栓藥物與舊式溶血栓藥相比較，當治療指標訂為死亡率時，則藥理作用更強的治療可能看起來效果較好，但如果治療指標定為包含死亡率和嚴重中風的複合型指標時，則新藥物便可能看起來效果較差。同理，當比較兩種止痛

劑對牙痛的效果時，若在較早的時間點觀察，則對藥效較快發作的止痛劑較為有利，若於較晚的時間點觀察，則對藥效較持久的藥物較為有利。

## 1.5、檢測靈敏度

檢測靈敏度是臨床試驗的一項特性，其定義為區分有效治療與不太有效或無效治療的能力。檢測靈敏度對任何試驗皆具重要性，但對於不同試驗，例如旨在證明不同藥物療效差異的較優性試驗（superiority trial）和旨在證明不同藥物療效相似的不劣性試驗，其影響有所不同。對於較優性試驗而言，當試驗缺乏檢測靈敏度，則試驗便可能無法顯示出測試藥物的療效，而無法得到測試藥物較優性的結論。反之，對於不劣性試驗，當試驗缺乏檢測靈敏度，則試驗可能會錯誤地顯示一種無效的新治療不劣於已知有效的舊治療，而得出錯誤的結論。

當試驗已顯示兩種治療有不同的療效（即其中一種優於另一種）時，這一結果本身就證明了該試驗具有檢測靈敏度。然而，一個成功的不劣性試驗（即已顯示出不劣性），或者一個未成功的較優性試驗，通常不帶有可直接證明檢測靈敏度的證據。

### 1.5.1、不劣性或相等性試驗之檢測靈敏度

不劣性或相等性試驗是否具檢測靈敏度可由兩方面推論：

- 1) 有對藥物效果具檢測靈敏度的歷史證據，亦即過去具有相似設計的試驗經常能區分有效治療與較不有效或無效的治療。
- 2) 適當的試驗進行，亦即試驗的執行方式不會損及其區分有效治療與較不有效或無效治療的能力。

檢測靈敏度的歷史證據應當在開始進行不劣性試驗之前作評估。具體而言，應確認在某一治療領域內，經適當設計與執行的試驗曾使用一特定活性藥物，或是其他有同等療效的藥物，這些試驗能可靠地顯示治療效果。理想情況是擬納入試驗作為活性對照組的活性治療藥物，在過去試驗已經由可靠方式證明其較優於安慰劑。若是如此，這就為有相同設計的活性對照試驗提供了檢測靈敏度的歷史證據（見第 1.5.1.1 節）。

活性對照不劣性試驗的進行方式是否恰當，惟有在完成試驗之後才能作出完整評估。活性對照不劣性試驗不僅必須在設計上與先前用於確定檢測靈敏度歷史證據的試驗相近（例如收案標準、可允許的合併療法等），另外亦須對實際納入試驗的受試者族群、實際使用的合併療法等項目作評估，確保試驗在執行上確實等同於先前的試驗。試驗在執行上亦須有高品質，例如良好的遵從性（compliance）以及僅有少數失去追蹤（loss to follow-up）的個案。連同檢測靈敏度的歷史證據，適當的試驗進行方式（第 1.5.1.2 節）同樣確保了活性對照試驗的檢測靈敏度。

因此，不劣性試驗在設計與執行上可分為四個關鍵步驟：

1. 確認是否有對藥物效果具檢測靈敏度的歷史證據。缺乏此證據便無法藉由不劣性設計顯示療效，亦不應採用此策略。
2. 設計試驗。試驗設計的重要細節，例如受試者族群、合併療法、試驗指標、導入期(run-in period)等，須嚴格遵循檢測靈敏度歷史證據的試驗設計。
3. 設定臨界值。訂定可接受的臨界值，而在此一方面須將歷史數據、相關臨床與統計考量等均納入考量。
4. 執行試驗。試驗在執行上須嚴格遵循歷史試驗的做法且維持高品質。

### **1.5.1.1、藥物效果具檢測靈敏度之歷史證據，與不劣性臨界值之選擇**

如前所述，多數活性對照相等性試驗大多是以建立新治療方法療效為目的之不劣性試驗。不劣性試驗的結果分析方法已於 ICH 的 E9、E3 兩部指引中討論。簡言之，此類試驗中是將所研究的治療與已知有效的治療作比較。執行試驗之前，須先選擇相等性或不劣性的臨界值，有時亦稱為 delta 值。此臨界值代表的是新治療劣於其活性對照組到何等程度時，在統計分析時會予以排除(視為未達到不劣性)。若新治療與活性對照組之間差異的信賴區間尚未等於或大於臨界值，可宣稱新治療具不劣性。反之，若信賴區間等於或大於臨界值，則新治療方法無法宣稱不劣性。

不劣性試驗所選的臨界值不應大於在該試驗中活性藥物相較於安慰劑的最小預期療效差異(即療效幅度，effect size)。若新藥的療效數值低於活性對照藥物，且其差異等於或大於臨界值，則該新藥有可能完全不具療效。唯有存在藥物效果檢測靈敏度之歷史證據時，才可能合理訂定一個有活性的藥物至少應具備的療效表現(不超過臨界值)；事實上，臨界值的決定便是以歷史證據為基礎。一般是以先前安慰劑對照試驗的結果來訂定臨界值，此處的安慰劑對照試驗的設計須適當且與所規劃的試驗相近。但亦可以從劑量反應或活性對照較優性試驗的結果來訂定臨界值。無論先前試驗中使用何種對照類型，臨界值是基於活性治療相較於對照組的較優性數值表現來訂定，而非諸如自基期的變化值 (change from baseline) 之類無對照的測量值。本指引並未詳述臨界值的計算方式，文獻亦少有提及。

在不劣性試驗中，決定臨界值的過程包含統計推論與臨床判斷，故臨界值的訂定應當反映出所依據證據的不確定性，且應適當保守。若訂定過程執行得當，則當試驗結果顯示新藥與活性對照藥物的療效差異的信賴區間未超過臨界值時，即代表新藥的療效值大於零。在實務上，臨界值通常會訂定小於活性對照藥物的最小預期療效幅度，以確保新藥在臨牀上仍有可接受的療效表現。

舉例而言，在評估死亡率的不劣性試驗中，僅確保新藥的療效幅度大於零是不夠的，通常會希望新藥能達到活性對照藥物對死亡率療效的一定比例。這一點對於探討新藥和活性對照藥物相對療效的試驗(見第 1.4.2 節)同樣適用，而此類試驗通常會希望試驗結果能確保新藥與對照藥極為相似，而不光只是保證新藥能產生任何療效。

由於不劣性試驗的臨界值是以歷史證據為基礎，故不劣性試驗與歷史對照（外部對照）試驗具有一定程度的共通性。唯有在對照藥物療效大小的估算有先前充分的研究結果支持時，不劣性試驗的設計才被認為適當且可靠。先前研究結果須指向共同的結論，即活性對照藥物在適當規模且與所規畫試驗設計相近的試驗中，持續表現出與安慰劑有區別，且得出一個合理的最小療效幅度值。若先前有相似設計的安慰劑對照試驗不時顯示出活性對照藥物與安慰劑並無區別，而且此現象無法由先前試驗的某些特性來解釋，那麼在所規劃試驗中，新藥的療效僅能藉由較優性的比較來證明。

是否有適當的檢測靈敏度歷史證據，在某種程度上是常理判斷。某些情況下，檢測靈敏度從先前安慰劑對照試驗在結果上所展現的一致性便顯而易見，抑或是因為接受與未接受治療的結果有很大區別而不證自明。舉例而言，短期試驗中，許多傳染性疾病在接受有效的治療後，其治癒率通常會遠高於自發性痊癒率。然而，也有許多情況是在有良好對照的試驗中，被認為有效藥物的表現並非總是優於安慰劑，因而難以對該藥物的最小療效幅度作出可靠的決定。這樣的情況通常出現在安慰劑組也可能有可觀改善與變化的情況，及/或治療方法的效果較微小或變異較大，例如憂鬱症、焦慮、失眠、心絞痛、症狀性鬱血性心臟衰竭、季節性過敏、症狀性胃食道逆流疾病等。

以上列舉的例子裡，標準療法毫無疑問是有效的，因為有許多良好對照的試驗顯示這些藥品有效。然而，根據現有經驗，要敘述出讓藥物能可靠發揮出至少最低效果的試驗條件（即歷史證據的條件），從而制定適當的臨界值，十分困難。在某些情況下，具檢測靈敏度的歷史證據，其相關性也有疑慮，比如治療與診斷的標準已經隨時間經過而有很大的變化（參閱第 2.1.7.1 節）。若活性對照或不劣性試驗所訂定的不劣性臨界值，無法有充分的藥品效果靈敏度歷史證據支持，則即使試驗結果顯示具有不劣性，該結果亦難以作為療效的證據。

如前所述，具檢測靈敏度的歷史證據僅適用於特定的試驗設計。欲讓所規劃的不劣性試驗對藥物效果具同等的檢測靈敏度，該試驗的設計重點便必須跟歷史試驗相似，包括：收案條件（疾病嚴重程度、共病症、診斷方法）、對照藥的劑量與給藥方案、合併療法、指標和評估指標的時機點、以及用來排除某些病人的廓清期（washout period）等。當某些設計的差異是無可避免或必要時（例如因為科技或療法上的進展），則須審慎考量這些差異存在時，歷史證據是否仍適用，以及是否對臨界值的訂定造成影響。

### 1.5.1.2、恰當的試驗進行方式

即便有檢測靈敏度的歷史證據，而且新試驗與過去試驗有相同的設計，試驗實際執行情況仍有可能削弱檢測靈敏度。為確保檢測靈敏度，試驗在執行上必須有高度的品質，而且實際上納入試驗的病人、實際上施予的治療方法（除了試驗所測試的治療方法以外）、實際上進行的評估都應該與有檢測靈敏度的歷史證據試驗相同。

試驗執行中有許多因素可能降低有效治療與較不有效或無效治療之間的差異，進而削弱試驗的檢測靈敏度。以下列舉數例：

1. 對療法遵從性不佳
2. 納入試驗的受試者族群對藥物反應不佳
3. 合併使用試驗計畫之外的藥物或其他治療方法，因而干擾測試藥物或降低原本可能具有的療效
4. 紳入試驗的族群有自發性改善的傾向，使得藥效沒有發揮的空間
5. 診斷標準應用不當（納入試驗的受試者並無欲研究的疾病）
6. 因為知曉所有受試者都接受可能具療效的藥物，使得指標評估有所偏差，例如傾向於將血壓的反應都視為正常，從而降低測試藥物與對照藥之間的差異

試驗主持人和委託者都希望試驗執行上具有高品質，而藥品優良臨床試驗規範（《ICH E6》）的指引能持續提升試驗品質。雖然如此，仍須注意以呈現治療方法之間的差異為目的之試驗，亟需有良好的試驗設計並降低試驗執行錯誤，因為試驗的瑕疵越多越可能導致差異難以表現出來。在安慰劑對照試驗中，已經有許多措施可以改善遵從性，並增進受試者族群對藥效的反應，從而確保可能有療效的藥物能與安慰劑有所區分。雖然如此，在許多臨床環境中，即便有強大的動機與多方努力來保障試驗品質與檢測靈敏度，臨床試驗仍然常常無法將可能有療效的藥物與安慰劑區分開來。

反之，在以呈現兩個治療方法的差異僅在一定範圍內（不劣性）為目的之試驗中，可能不會像欲證明差異存在的情況那樣耗費多方努力來保障試驗品質，從而確保能發現差異——亦即確保檢測靈敏度。導致差異降低的試驗錯誤（例如遵從性不佳、安慰劑反應大、某些合併療法、指標評估錯誤等），是維持檢測靈敏度的一大疑慮。然而，如果相信新藥療效優於其對照藥物時，便可能有極強的動機進行高品質的試驗，以利排除不劣性臨界值。此外，值得注意的是，某些試驗錯誤有可能提高變異性，從而擴大信賴區間，無法排除治療和對照之間的差異超過邊界，導致顯示不劣性的可能性降低。因此，執行不劣性試驗時，也可能有強大動機來減少這些可能源自於評估方式不佳的變異性。

如前所述，為判斷不劣性試驗是否執行恰當，不僅需要審查可能使治療方法之間差異縮小的因素，還需要檢視該試驗與確定不劣性臨界值的試驗之間不同的因素。尤須注意的是，在收案族群、合併療法、療法遵從性、退出試驗的程度與理由等方面所觀察到的任何差異是否有可能對檢測靈敏度造成不良影響。此外，即使該試驗在設計與執行上看似與確定不劣性臨界值的試驗極為相似，但若活性對照藥物出現明顯非典型的結果（例如抗生素試驗中治癒率異常高或異常低），即可能代表該試驗與過去試驗有重要差異存在。

### 1.5.2、以展現較優性為目的之試驗的檢測靈敏度

檢測靈敏度的議題在不劣性試驗中亟具關鍵性，但實際上對於各種試驗都十分重要，包括安慰劑對照試驗與劑量反應試驗。舉例而言，若未能顯示某治療方法優於安慰劑，代表該治療方法不具效果，抑或是該試驗在設計與執行上無法區別出有效的治療方法與安慰劑。

在評估檢測靈敏度時，三臂試驗(three-arm trial)是有效的評估方法。三臂試驗包含安慰劑與一種已知有活性治療方法，這種試驗設計有許多好處，可以評估測試藥物的療效幅度（測試藥

物相較於安慰劑），並在活性對照藥與安慰劑比較之下且建立了檢測靈敏度後，可以比較測試藥物與活性對照藥的表現（見第 2.1.5.1.1 節）。

## 第二章、關於對照類型的詳細考量

### 2.1、安慰劑對照

#### 2.1.1、說明（見第 1.3.1 節）

在安慰劑對照試驗中，受試者通常透過隨機分派的方式而被分配至試驗組或安慰劑組。所謂安慰劑是一個「假性」的治療方法，其外觀特色如色彩、重量、味道、口感等均盡可能與所研究的藥物盡可能相同，但其內容並未包含測試藥物。有些試驗可能研究測試藥物的多個劑量，或是同時具有活性對照與安慰劑對照。在這樣的情形下，採用多個安慰劑（雙虛擬〔double-dummy〕設計）可能會比試圖使所有藥物看起來相同更容易。採用安慰劑促使雙盲的進行，也幾乎總是會與雙盲設計搭配。在試驗條件之下對治療結果的測量，可以顯示具活性藥物與安慰劑組的差異。一般而言，有各種設計可以使用：平行(parallel)或交叉(cross-over)設計（見《ICH E9》指引）、活性藥物的單一固定劑量、漸增劑量或多個固定劑量。下文將說明一些需特別留意的設計。須注意並非每一個包含安慰劑的試驗都是安慰劑對照試驗。舉例而言，活性對照試驗有可能每一種藥品都搭配安慰劑以達成盲性目的（雙盲）。這樣的試驗仍然是活性對照試驗，而非安慰劑對照試驗。安慰劑對照試驗，是把測試藥物作為治療方法以及把安慰劑作為治療方法，兩者加以比較的試驗。

另外亦須注意的是並非所有的安慰劑都完全不具活性成分。舉例而言，某些皮膚局部製劑試驗所使用的基劑對照(vehicle controls)可能具某種有效作用，但這並不會影響試驗本身檢測藥物療效的能力。然而若載體對照會造成有害的作用，則會產生某些問題。在這種情形下，設計上若有一組別是「無治療」，便可以評估測試藥物加上載體對照的整體效果。

#### 2.1.2、將偏差最小化的能力

安慰劑對照試驗通常可以藉由隨機分派與盲性將受試者與研究人員的偏差降至最低。然而，這一類的試驗會因為注意到其中一種治療方法的藥理效果而導致盲性被破壞。這種情形下需要有盲性人員評估結果才能減少偏差。這項考量在交叉試驗中尤其需要注意。

#### 2.1.3、倫理道德考量

當以新治療方法測試於尚無已知有效療法的疾病時，通常採新治療方法與安慰劑對照作比較的設計，這樣一般不會牽涉到倫理道德問題。但是，若該疾病其實有已知有效的治療方法，則使用安慰劑對照可能引發倫理道德、可接受性、可行性等方面的問題。若已知有治療方法可以防止嚴重傷害，比如死亡或不可逆的病況，則採用安慰劑並不適宜。然而，偶爾也有例外情況，比如標準療法有嚴重毒性以致於許多病人拒絕接受的情況。

於其他情況，在不會造成嚴重傷害的前提之下，邀請病人參與安慰劑對照試驗一般認為是符合倫理道德的。即便試驗將令病人感到不適，只要是非強迫性，且已充分告知病人現有的治療方法以及延後接受治療的可能後果等前提即可。然而，即便此類試驗符合倫理道德，仍然可能在實務上涉及重要的問題。舉例而言，延遲對疼痛或其他症狀的治療有可能對病人或醫師而言是不可接受的作法，因此使他們無意願參與這類試驗。至於存在已知有效治療方法的情況下，是否應該讓病人與醫師參與新藥的安慰劑對照試驗，取決於試驗主持人、病人、人體試驗委員會 (institutional review board, IRB) / 獨立倫理委員會 (independent ethics committee, IEC) 的判斷，且其可接受性可能在各地區不同。試驗的可接受性取決於各別的試驗設計與所選之病人族群，並將於隨後章節 (第 2.1.5 節) 作討論。

在評估某個安慰劑對照試驗是否符合倫理道德時，有可能取決於該試驗的特定情況，以及研究人員認為試驗結果具有何臨床意涵。舉例而言，對有輕度原發性高血壓但無器官損害的病人進行新式抗高血壓藥的短期安慰劑對照試驗，一般可能可接受，但較長的試驗或收納病情較嚴重的病人則可能不行。

此處需強調的是，採用安慰劑或無治療對照並不一定代表病人完全不會受到任何治療。舉例而言，在腫瘤試驗中，若無已核准的藥品，則試驗組以及安慰劑或無治療對照組中的病人將接受必要的緩解療法(palliative treatment)，比如止痛劑，以及最佳的支持性療法。許多安慰劑對照試驗都是以附加性的形式進行的，亦即所有病人都接受一種指定的標準療法或由主治醫師或醫療機構自行選擇 (見第 2.1.5.2.1 節)。

#### 2.1.4、安慰劑對照試驗之適用性與在特定情況下之推論強度

當用於呈現療效時，安慰劑對照試驗應盡排除各種假設或和對外部 (試驗之外) 資訊的倚賴。試驗在設計與執行上的問題大多會增加無法呈現治療差異的可能性，因此試驗本身便會促使試驗執行有良好的品質。即便試驗以比較兩種活性藥或評估劑量反應為主要目的，增加安慰劑對照可以提供一道內部標準，強化了從其他比較獲得結論的可信度。

安慰劑對照試驗亦提供了可區分由藥物引發或是由疾病或併發事件引發不良事件的能力。然而，須注意的是，當呈現兩項治療方法的相似性時，例如藉由接受新藥的病人與接受安慰劑的病人之間不良事件發生率相近，來證明新藥並不特別具有不良影響時，此時安慰劑對照試驗與任何相等性或不劣性試驗一樣，都需要考慮檢測靈敏度的問題 (見第 1.5.1 節)。解讀試驗結果時，首先應確認的是，若測試藥物會引發不良事件，該事件應該能被實際觀察到，故通常這樣的試驗應包括一個實際上會引發該不良事件的活性對照藥物來做比較。但在某些情況下，可藉由過去相似設計「對藥物不良效果具有檢測靈敏度」的歷史證據，來支持本試驗具有檢測靈敏度。

#### 2.1.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題

安慰劑對照試驗的倫理道德問題或在實務方面的限制，往往可以透過修改試驗設計來解決，但同時保留此類試驗的推論優勢。此外，加入額外的試驗組，例如測試藥物數個劑量的試驗組，或加入一個已知有效的活性對照組，可以讓安慰劑對照試驗提供更多的資訊。

### 2.1.5.1、額外的對照組

#### 2.1.5.1.1、三臂試驗：安慰劑與活性對照

如第 1.5.1 節所述，包含一個活性對照組與一個安慰劑對照組的三臂試驗，可在未能區分測試藥物與安慰劑的差異時，評估此一失敗是否代表測試藥物不具療效，抑或是因為該試驗缺乏辨識出藥物療效的能力。在此類試驗中，將安慰劑與標準藥物比較，可以作為檢測靈敏度的內部證據。如果必須與活性對照藥物進行精準的比較，亦可令活性對照藥組的規模比安慰劑組為大，以提高比較結果的精確性。此一作法亦可能使得病人與研究人員更願意接受該試驗，因為被隨機分派至安慰劑組的機會降低。

#### 2.1.5.1.2、額外劑量

隨機分派到試驗用藥數個固定劑量的試驗組加上安慰劑組，這樣的做法，使試驗可以分析劑量反應關係，而且因為較能公平地比較各治療方法，故對於比較性試驗特別有用（見《ICH E4：供藥品登記佐證之劑量反應資訊》指引）。

#### 2.1.5.1.3、因子設計(Factorial design)

因子設計可用於探索測試藥物的多個劑量，不管是作為單方治療 (monotherapy) 或是與另一種藥以及另一種藥的多個劑量合併使用。此類試驗可以分析多種組合的結果。此類試驗常用於評估高血壓的新治療方法，或是同時進行超過一種治療方法的多種試驗中亦可考慮使用此作法。舉例而言，此類試驗可以展現乙醯水楊酸(aspirin)與鏈激酶(streptokinase)在心臟病發後預防死亡率的獨立加成作用。

### 2.1.5.2、試驗設計其他方面的調整

#### 2.1.5.2.1、安慰劑對照附加性試驗；替代試驗

附加性試驗是新藥品的安慰劑對照試驗，試驗中受試者同時也接受標準療法治療。這樣的試驗在現有的治療方法已知可有效降低死亡率或不可回復病況時，以及當標準療法做為活性對照的不劣性試驗是無法執行或難以詮釋時（見第 1.5 節），都顯得格外重要。此類試驗在抗癌、抗癲癇、心臟衰竭藥物的臨床研發中十分常見。此設計唯有在標準療法無法完全發揮療效時（但幾乎總是如此）才能派上用場，並且具有可提供臨床結果改善（而非「僅是」不劣性）的證據這樣的優點。當然，此類試驗所建立的療效證據僅適用於合併療法，而單藥治療情況下的劑量可能與有效的合併療法下的劑量不同。整體而言，此一作法僅有在新藥品與標準療法的藥理機轉不同時才有機會成功，雖然的確有例外。舉例而言，愛滋病人的合併療法有可能因為延遲抗藥性的出現而顯示在藥理上相關藥品的合併使用有正面效果。

替代試驗是此設計的一個變化，可以偶爾提供單藥治療的資訊，尤其適合在慢性疾病的情形。在替代試驗中，新藥品或安慰劑隨機加入有效劑量下的傳統治療中，然後將傳統治療撤除（通常是逐步減量）。接著，便使用事先制定的成功條件觀察新藥品與安慰劑將

受試者維持在基期狀態的能力。此作法曾被用以研究類固醇依賴型病人的類固醇節用(steroiod-sparing)替代品，藉以避免初期類固醇戒斷與廓清期的症狀復發。此作法亦曾用以研究抗癲癇藥物的單藥治療。

#### 2.1.5.2.2、「及早退出」；救援性治療

在試驗設計中規劃如何由無效的治療及早退出(early escape)是有可能的。及早退出一詞係指受試者一旦發生以下情形即馬上令其退出試驗：臨床狀況惡化或未能改善至特定標準（在事先決定的時限內血壓未能獲得控制、癲癇發作率高於預先決定的數值、血壓升高至特定標準、心絞痛頻率升高至特定標準、肝炎病人的肝酵素值未能在一定時間內正常化）、發生一次該治療欲避免的不良事件（首次復發不穩定型心絞痛、強直陣攣發作、陣發性室上性心率不整）、因其他原因而必須採取救援性治療等等。在這些狀況中，需要改變治療方法的事件即為試驗之指標。決定這些指標是否發生的標準須有詳盡規定，而且測量時機必須能確保不會讓病人在病況控制不佳的處境下仍未能接受活性藥物的治療。此試驗設計的主要困境在於或許只能提供短期的有效性資料。不過，透過在隨機退出試驗（見第 2.1.5.2.4 節）設計中加入及早退出機制，可提供長期的有效性資料。

#### 2.1.5.2.3、有限的安慰劑治療期

在長期安慰劑治療不被接受的情形下，在活性對照試驗初期的一段時間內採用一個安慰劑對照組可以評估試驗的檢測靈敏度（至少可適用於短期效果）。隨後該試驗可在撤除安慰劑組的情形下持續進行。

#### 2.1.5.2.4、隨機退出

在隨機退出試驗中，受試者接受試驗治療一段特定時間後，隨機分派為繼續接受試驗治療或改接受安慰劑治療（亦即撤除活性治療）。此類試驗的受試者可來自於開放式單臂試驗、來自既有的臨床病人族群（但是通常須在試驗計畫書中制定一「洗入期」[wash-in] 階段以建立接受治療初始時的基期狀態）、來自對照試驗的活性組、來自活性對照試驗的單組或雙組等。在持續接受試驗治療與改接受安慰劑治療的分組間所出現的任何差異即展現該活性治療的效果。在隨機分派之前接受試驗治療的觀察期長短並無限制，因此這個方法可用於在無法進行長期安慰劑治療的情況下，研究試驗治療長期持續的有效性(effectiveness)。隨機退出後的觀察期可以是一段固定的时间，也可以採用及早退出或事件發生時間 (time to event；例如憂鬱症復發) 等設計方式。如同及早退出設計一般，須多加注意病人的監測與試驗指標的評估，以確保能迅速發現治療失敗的病人。

隨機退出的方式在數種情形下十分有用。首先，它或許能適用於用以解決復發性疾病發作的藥品（例如抗憂鬱藥），在這種情形下該隨機退出試驗實際上算是一種預防復發的試驗。其次，它或許能適用於可抑制症狀或徵候（慢性疼痛、高血壓、心絞痛），但長期安慰劑對照試驗難以執行的藥品。在這種情況下，該試驗可確立藥品的長期療效。第三，它對於決定一項治療應持續多久特別有用（例如採用  $\beta$  受體阻斷劑的心肌梗塞後治療）。

隨機退出設計的整體優勢在於，當配合及早退出的試驗指標（例如症狀復發）使用時，可讓病人須接受反應不佳的安慰劑治療的時間極短。

這種類型的設計也可以評估劑量方面的問題。在所有病人都接受初步的固定劑量治療後，他們可以在退出階段被隨機分派到數個不同的劑量（以及安慰劑）分組。這在有特定理由認為初始劑量與維持劑量有可能不同時特別有用，不論其理由是基於藥效學考量抑或是因為活性藥因原型藥(parent drug)或活性代謝物的半衰期較長而出現可觀的累積量。須注意隨機退出設計可在初步安慰劑對照劑量調升試驗(titration study，見《ICH E4》指引)後用以評估劑量反應關係。劑量調升試驗對於確立有效性是頗有效率的設計，但在許多情形中不能妥善提供劑量反應關係的資訊。在隨機退出階段，發生反應者隨機被分派至數個固定劑量與安慰劑組，使得劑量反應關係得以進行嚴謹的探討，同時又可在試驗初期使用劑量調升(dose titration)設計的高效率。

在隨機退出設計的應用上，留意是否可能發生退出現象相當重要，這顯示須明智地採用相對較為緩慢的逐步遞減藥量方式。病人有可能對藥品產生耐受性而無法受益，但停藥可能又導致病況加劇，從而對藥效持續性作出錯誤結論。此外十分重要的是，這些試驗中所觀察到的治療效果可能高於未經篩檢的病人族群，因為隨機退出試驗富含(enriched)反應者，同時又排除無法耐受該藥品的受試者。此現象發生於試驗明顯僅收納對藥品有出現反應的受試者，或僅收納已完成前一階段試驗的受試者（此常代表受試者對藥品有良好反應，並總是代表受試者對藥品有耐受力）。在目標在決定治療應持續多久的試驗中，這樣的收案條件可以提供所需的受試者族群以及所欲研究的比較。

#### **2.1.5.2.5、其他設計上的考量**

在任何安慰劑對照試驗中，不平衡的隨機分派（例如試驗組對安慰劑組比例 2:1）有可能使安全性資料庫充實，並使得試驗對病人及／或試驗主持人更具吸引力。

#### **2.1.6、安慰劑對照試驗的優點**

##### **2.1.6.1、展現療效的能力**

如同其他的較優性試驗，安慰劑對照試驗本身即帶有檢測靈敏度的內部證據。在顯現出差異時，無需參考外部發現即可對差異進行詮釋。

##### **2.1.6.2、測量「絕對」療效與安全性**

安慰劑對照試驗測量一個治療透過藥理機轉達成的完整效應。與之相對，活性對照試驗或劑量比較試驗則是測量相對於另一種治療的效應。另外，安慰劑對照試驗還可區分不良事件是由試驗用藥所引發，或由潛在疾病或「背景雜訊」所引發。絕對的療效幅度的訊息對於三臂試驗（試驗組、安慰劑、活性對照）極具價值，即便試驗的主要目的係在比較試驗用藥與活性對照。

### 2.1.6.3、效率

安慰劑對照試驗極具效率，可以用比其他類型的同步對照試驗都小的樣本數量，來測得治療效果。

### 2.1.6.4、將受試者與試驗主持人的期望所致的影響最小化

使用經過盲性的安慰劑對照可以抑制因受試者與試驗主持人的期望所致的病況改善，因為雙方都理解到有一部份的受試者的治療並未用到活性藥。如此一來，將可提升試驗偵測藥品真實效果的能力。

## 2.1.7、安慰劑對照試驗的缺點

### 2.1.7.1、倫理道德考量（見第 2.1.3、2.1.4 節）

若已知有可在特定族群降低死亡率或不可回復病況的有效治療時，則通常由倫理道德觀點來看，該族群不應進行安慰劑對照試驗。但是，關於實際適用這一點的特定狀況與族群為何仍有爭議。倫理道德方面的考量也有可能促使試驗採用病況較不嚴重的受試者，或是即便長期試驗指標較為重要卻仍僅研究短期試驗指標。當安慰劑對照試驗有悖倫理道德，而活性對照試驗又不足採信時，新藥的研究將會窒礙難行。舉例而言，對急性心肌梗塞病人進行血栓溶解劑的安慰劑對照試驗會被視為有悖倫理道德之舉。然而，在當前的環境下也難以依靠歷史數據建立確鑿的不劣性臨界值，因為急性血管重建術 (revascularization)的出現或許已使血栓溶解劑的效益值發生改變。第 2.1.5 節所敘述的設計或許能適用於一部分的情形。

### 2.1.7.2、病人與醫師的實務考量

即便在暫緩或延遲治療不致造成傷害已達成共識的狀況下，醫師及／或病人或許仍然會不願意接受病人被分派至安慰劑治療組的可能性。自覺未有改善的受試者有可能將此歸咎於安慰劑而缺乏效果，並退出治療從而對試驗分析造成困難。然而，若謹慎處理，因為不見療效而退出試驗有時反而能成為試驗的指標。雖然這一點能提供一些關於藥品有效性的資料，但其精確性仍低於受試者接受所分配治療後的實際臨床狀況資料。

### 2.1.7.3、概推性(Generalizability)

偶爾有主張任何對照試驗均代表一種人工環境，尤其安慰劑對照試驗更是如此，其所得試驗結果有別於在「真實世界」的有效性。若安慰劑對照試驗的受試者因為倫理道德或實務上考量而不具足夠的代表性，試驗結果的概推性便有待商榷。舉例而言，罹患較嚴重疾病的病人有可能因為計畫書、試驗主持人或病人的選擇而被排除在安慰劑對照試驗之外。在某些情形下，願意參與試驗的病人或試驗中心可能數量頗為有限。這些考量是否實際上（而非理論上）對試驗的概推性造成限制尚無定論。

#### 2.1.7.4、無比較的資料

缺乏活性對照的安慰劑對照試驗幾乎無法提供有用的有效性比較資訊，而這卻正是在許多情形下亟需得知且極重要的資訊。這樣的資訊無法可靠地透過跨試驗比對取得，因為各試驗的情況有可能存在巨大差異。

### 2.2、無治療同步對照（見第 1.3.2 節）

隨機分派無治療對照在整體特性與優缺點方面與安慰劑對照試驗相似。但是，有別於安慰劑對照試驗，此方法無法做到完全盲性，而這一點對試驗的各方面都有影響，包括受試者存留、病人管理以及觀察工作的各個面向（見第 1.2.2 節）。此設計適用於執行安慰劑對照試驗是恰當的情形下，只不過盲性並不可行或不切實際。採用此設計時，最好能將試驗的關鍵決策，例如合適性判定(eligibility determination)、終試驗指標判定或管理變更等，交由對治療分派加以盲性的觀察人員來負責。與數據分析有關的決策，例如分析群體 (analysis sets) 所包含的病人，亦應交由無法得知治療分派的人員來負責。進一步討論可參閱《ICH E9》指引。

### 2.3、劑量反應同步對照（見第 1.3.3 節）

#### 2.3.1、說明

在劑量反應試驗中，受試者被隨機分派至兩個或多個劑量組，可能有或沒有包括安慰劑組。執行劑量反應試驗的目的在於建立劑量與療效和不良效果之間的關係，及／或展現療效。第一個目的已於《ICH E4》指引中討論，本指引則討論展現療效的部分。建立療效的證據可以基於劑量組間成對比較，或是劑量組與安慰劑組間的比較，抑或是隨劑量增加所展現的顯著趨勢之上，即便任兩組間均無顯著差異亦可。然而，在後者的情形中或許有必要進行額外試驗，以評估低劑量的有效性。如同《ICH E9》指引所述，須事先定義主要療效分析的處理方法。

包括不同給藥方案治療組的試驗具有許多跟劑量反應試驗相同的考量。由於透過給藥方案對照試驗來確定療效屬罕見作法，目前討論將聚焦於劑量反應試驗之上。

在劑量反應試驗中加入一個安慰劑（零劑量）組有數項優勢。首先，可避免試驗因各組結果相似而無法判讀結果的情形，畢竟若無安慰劑組將難以評斷究竟是所有劑量都同樣有效抑或是同樣無效。再者，有安慰劑組存在則可估算該治療透過藥理機轉達成的總效應，雖然在劑量組樣本數較小的情形下估算將無法相當精確。第三，由於藥品與安慰劑之間的差異通常大於個別劑量之間的差異，使用安慰劑組可能縮小所需樣本數。各劑量組的樣本數並不需要均等，比如可以用較大的樣本數來提供低劑量組更精確的效應資料，或以此提升試驗的檢定力 (power)，以展現預期的最佳劑量具有明確療效。劑量反應試驗可以包含活性對照治療的一個或多個劑量，亦可在隨機退出試驗設計將受試者分派至多個劑量組。

### 2.3.2、將偏差最小化的能力

如果劑量反應試驗有盲性，則跟其他隨機分派盲性設計一樣具備將受試者與試驗主持人偏差最小化的能力。當一項藥品具有致使部分受試者或試驗主持人破壞盲性的藥理效果時，劑量反應試驗維持盲性或許比安慰劑對照試驗來得容易。為維持治療盲性，有可能需要準備多個假治療（dummies）或使數個不同劑量的外觀看似相同。

### 2.3.3、倫理道德考量

劑量反應試驗所涉及的倫理道德與實務考量與安慰劑對照試驗相似。若已知能有效防止死亡或不可回復病況的治療，則刻意將受試者隨機分派到治療效果較差的控制組，在倫理道德上並不比隨機分派到安慰劑組來得可接受。在嚴重性較低的病症，或當治療的毒性與效益相當時，病人與試驗主持人有可能會認為使用較低劑量，可能效果較弱、毒性較低，或安慰劑的劑量反應試驗是可以接受的選擇。

### 2.3.4、劑量反應試驗之適用性與在特定情況下之推論強度

大體而言，在可以利用安慰劑對照試驗確立成效與安全性的情況下，有盲性的劑量反應試驗也具備相似的功用（見第 2.1.4 節）。

### 2.3.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題

大體而言，可以減少安慰劑對照試驗在倫理道德、實務、推論方面問題的試驗調整亦可適用於劑量反應試驗（見第 2.1.5 節）。

### 2.3.6、劑量反應試驗的優點

#### 2.3.6.1、效率

雖然將一個足夠且完全有效的劑量與安慰劑作比較可以最有效率地展現療效，這樣的設計卻有可能產生難以接受的毒性且無法提供劑量反應關係的資訊。當劑量反應關係是單調的情況下，劑量反應試驗在展現療效方面具有合理的效率，而且可以提供劑量反應關係的資訊。若最適當的有效劑量尚屬未知，則研究特定範圍內的數個劑量將會比逕直挑選單一劑量為佳，因為所選單一劑量有可能效果不佳或產生令人無法接受的不良反應。

#### 2.3.6.2、可能存在的倫理道德優勢

在某些情形中，尤其在可能存在劑量相關的療效以及劑量相關的重要毒性反應的情形下，劑量反應試驗有可能代表一種符合倫理道德或務實，且可展現差異的試驗，即使是在安慰劑對照試驗不符此條件的情形，因為病人與試驗主持人有理由接受較弱的效果以換取較高的安全性。

### 2.3.7、劑量反應試驗的缺點

劑量反應試驗有一項須注意的潛在問題，即當具有一個正向的劑量反應趨勢（意即劑量與療效結果間呈現顯著相關），卻無顯著的成對差異時，則雖可以確立療效（見第 2.3.1 節）但卻可能難以確定哪些劑量（除最高劑量外）實際上有效。當然，單劑量試驗亦有相似的問題，因完全無法產生關於比該測試劑量更低的劑量的任何資訊。

另外，亦須注意劑量反應試驗顯示各個劑量之間沒有差異的情形並不少見。在沒有安慰劑組的情況下，這樣的結果通常無法提供任何資訊。

若療效範圍完全未知，此設計亦可能效率不佳，因為有許多受試者可能會被分派到療效範圍以下或以上的劑量組中。

劑量反應試驗在顯現藥物效果方面的效率有可能比安慰劑對照調升試驗來得低，但在多數情形下，劑量反應試驗能提供較佳的劑量反應資訊（見《ICH E4》指引）。

## 2.4、活性對照（見第 1.3.4 節）

### 2.4.1、說明

活性對照（陽性對照）試驗是將試驗用藥跟一已知的活性藥作比較的試驗。此類試驗為隨機分派且通常採雙盲設計。活性對照試驗在設計上最關鍵的問題是試驗究竟係以展現兩種治療的差異，抑或展現不劣性或相等性為目的。若欲藉由在試驗中展現試驗用藥相對於標準治療的不劣性來證實試驗用藥的有效性，須留意該試驗的檢測靈敏度，如第 1.5 節所述。在不劣性或相等性試驗中，作為活性對照的治療必須在所用的劑量以及該試驗狀況下已確立療效（見《ICH E9：臨床試驗統計指導原則》）。一般而言，這意味著此活性對照必須在提交試驗報告的地區屬於合乎規範的藥物，且在適應症與劑量皆與試驗相同的狀況下。反之，以顯示試驗用藥較優性的試驗則可以直接作為療效證據進行詮釋，即便活性對照的劑量過低或活性對照的效益未知（但不得有害）。然而，這樣的試驗結果，即試驗用藥優於其對照，唯有在活性對照係依適當劑量與時程施用於適當病人族群時才可以解釋為試驗用藥真正優於其對照（見第 1.4.3 節）。對照治療的使用不適當亦會造成該試驗無法判別不劣性，因若試驗用藥無法顯示具較優性的情況下，則無從確認該試驗的檢測靈敏度（見第 1.5.2 節）。

### 2.4.2、將偏差最小化的能力

隨機分派與盲性活性對照試驗通常能將受試者與試驗主持人的偏差降至最低，但仍須有應注意之處。在不劣性試驗中，試驗主持人與受試者均清楚知曉所有的受試者都將接受活性藥治療，雖然並不清楚所分派的藥品。這一點有可能導致在部分主觀的評估中將邊緣案例認定為成功案例的傾向。例如，在抗憂鬱藥的試驗中，可能會降低治療方法間能觀察到的差異，並且提高即使觀察到不劣性亦不足以作為藥品有效性證據的可能性。

### 2.4.3、倫理道德考量

一般認為活性對照試驗在倫理道德與實務上的問題較安慰劑對照試驗來得少，因為所有的受試者都能接受活性藥治療。然而，仍然有必要認清受試者接受的新治療並非標準療法（就跟安慰劑對照組也不是標準療法一樣），而且可能接受不具效力，甚或是有害的藥品。這一點在作為活性對照的療法已知可提升存活率或降低不可回復病況的發生率，亦即令安慰劑或無治療對照成為不可接受的選項時，將變得極為重要。因此，試驗治療須有完整的合理性論述。如果缺乏強烈的理由預期新藥品至少與標準療法一樣有效，採取附加性試驗（見第 2.1.5.2.1 節）或許會更為恰當，在可採用此設計的狀況下。

### 2.4.4、活性對照試驗之適用性與在特定情況下之推論強度

當新治療顯得比活性對照更有優勢時，假設該活性對照實際不具傷害性，便可如同其他任何較優性試驗一般逕直將試驗認定為顯現出療效。若活性對照試驗透過展現不劣性的方式展現療效，則須考量其檢測靈敏度，如第 1.5 節所述。若可確立檢測靈敏度，則活性對照試驗亦可用於評估相對療效。

### 2.4.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題

如前文所述（第 2.1.5 節），活性對照試驗中可加入一個安慰劑組、試驗用藥的數個劑量組及／或活性對照的其他劑量組。比較性的劑量反應試驗，亦即試驗用藥與活性對照都有數個劑量的試驗，在止痛劑試驗中屬於典型設計。活性對照試驗中的劑量可以是固定劑量或採劑量調升，而且試驗可以採用平行或交叉設計。不劣性試驗的檢測靈敏度有時可以透過試驗後期的安慰劑對照隨機退出階段（見第 2.1.5.2.4 節），或在試驗初期與安慰劑作短期比較（見第 2.1.5.2.3 節）來加以證明。活性對照較優性試驗對特定族群（對其他療法或活性對照無反應者）往往頗為有用而且一般而言相當易於詮釋，雖然所獲得的結果未必具概推性。

### 2.4.6、活性對照試驗的優點

#### 2.4.6.1、倫理道德與實務上的優點

活性對照設計，不論是以展現不劣性或相等性或較優性為目的，都可以減少因為不使用經文獻證實有重大健康效益的藥品所引起的倫理道德疑慮。另外，此設計還可以解決病人與醫師對於不使用經文獻證實之有效療法的疑慮。此設計有助於招募受試者以及取得 IRB／IEC 的許可，而且也有可能允許使用較大的樣本數進行試驗。由於缺乏效果而退出試驗的人數亦有可能降低。

#### 2.4.6.2、資訊的內容

在展現出相對於活性治療的較優性時，活性對照試驗便可直接作為療效證據進行詮釋。在活性對照試驗中採用較大樣本數，有時比較容易且比較容易被接受，因而提供更多關於安全性的資訊。若經適當設計，活性對照試驗還可提供相對療效的資訊。

## 2.4.7、活性對照試驗的缺點

### 2.4.7.1、資訊的內容

參閱第 1.5 節有關檢測靈敏度，以及不劣性或相等性試驗能否支持對療效下定結論的討論。即便確立檢測靈敏度且試驗證實適合偵測療效，亦無法直接評估療效幅度，在安全性結果的量化方面也更為困難。

### 2.4.7.2、龐大樣本數

一般而言，選擇須排除的不劣性臨界值時須要保守，以期能合理確保臨界值不大於能可靠預期的活性對照最小療效幅度。另外，由於通常須排除對照藥效果具超過一定合理限度損失的情形（見第 1.5.1 節），往往會採用一個更小的不劣性臨界值。因此臨界值的選擇會趨於保守，樣本數可能相當龐大。在活性對照較優性試驗中，藥品間的預期差異永遠小於試驗用藥與安慰劑間的預期差異，同樣導致樣本數變得相當龐大。

## 2.5、外部對照（含歷史對照，見第 1.3.5 節）

### 2.5.1、說明

在外部對照試驗中，構成對照組的病人族群並不屬於試驗用藥族群所屬的隨機分派試驗中的一部份，亦即該試驗並不具有隨機分派同步對照組。因此，其對照組與接受治療的族群並非來自於完全相同的族群。通常，其對照組係在較早時間點觀察並留下完善紀錄的病人族群（歷史對照），但也可以是在其他機構同時進行觀察的族群，或甚至在同一機構但不屬於該試驗的族群。外部對照試驗可以是較優性試驗（例如與未治療族群作比較）或是不劣性試驗。有時候，較大外部族群的特定病人會被選為對照組，依據某些特質挑選使其與試驗族群相似，甚至可能試圖將對照組與試驗組的特定病人進行配對。

在所謂的基期對照試驗中，會將病人隨時間推移的狀態與其基期狀態作比較。雖然這些試驗有時候會被視為「將病人作為自身之對照」，但實際上並無內部對照。反之，是以自基期以迄的變化與病人在未接受試驗用藥治療的預估狀態相比。基期對照試驗與使用較為複雜的依序啟-閉-啟設計（試驗用藥、安慰劑、試驗用藥）的試驗皆屬此類，但並不包括隨機分派同步對照組。如前所述，這些試驗所觀察到自基期以迄或試驗不同時間間的變化，始終都會與未接受介入的預估狀態相比較（至少含蓄地比較）。這些預估整體上都是以常識為基礎，並未參考特定的對照族群。雖然在某些情況中這樣做的合理性顯而易見，比如效果極為戲劇性、於接受治療後隨即發生、自發性出現的可能性極低（例如全身麻痺、心律調整、腫瘤明顯縮小），在多數情形中卻並不特別明顯且須參考特定歷史經驗。設計與分析此類試驗者必須明瞭其限制並應對採用的正當理由有所準備。

### 2.5.2、將偏差最小化的能力

無法控制偏差是外部對照試驗廣為人知的主要缺點，而且使得該設計在許多情況下均不適用。確立試驗組與對照組的可比性，進而發揮對照組的主要功能始終是一道難題，而且在許多情形下不可能達成的工作（見第 1.2 節）。除了使用試驗治療之外，各分組之間還有各種足以影響結果的因素，包括人口特徵、診斷標準、疾病分期或嚴重程度、合併療法、進行觀察的條件（例如臨床結果的評估方法、試驗主持人的期望等），彼此不相似。這樣不相似的情況可能包括重要但未被察覺且未進行測量的預後因素。採用外部對照時無法進行盲性與隨機分派以降低偏差。相較於隨機分派試驗中以明顯相似方法挑選的對照組，未接受治療的歷史性對照組結果往往較差，這在文獻中已有充分記錄，或許反映出選樣偏差(selection bias)。隨機分派試驗中的對照組必須符合特定標準以進入試驗，而這些標準與外部對照組的典型標準相比，通常較為嚴格而納入病況較輕的病人族群。外部對照組常常以回溯的方式篩選受試者，導致其存在潛在的選擇偏差。未能控制偏差的後果是，相較於同步對照試驗所應展現的結果，外部對照試驗的發現必須達到遠遠更為極端的統計顯著性，且在治療方法間出現大出許多的預估差異，才會被視為具說服力。

由於不能控制偏差，外部對照設計的使用被限制在治療效果極為戲劇性且病程通常可預期性極高的情形。此外，外部對照的應用須限制在試驗指標為客觀指標而且基期狀態與治療變數對試驗指標的影響已被詳細分析的情形。

如前所述，缺乏隨機分派與盲性，以及因為不能保證試驗組與對照組具可比性而衍生的問題，使得此試驗設計天生帶有大量潛在偏差且無法定量。即便如此，仍可依一些方法來設計試驗並執行可以使外部對照試驗變得更具說服力，且可能減少偏差。對照組的選擇必須能提供詳細的資訊，視情形包括個別病人在人口統計、基期狀態、合併療法、在試驗中的歷程。作為對照的病人必須與試驗中預期接受試驗用藥的族群盡可能相同，另外除了試驗用藥的差異以外，還必須以相同的條件和方式下接受治療。試驗中的觀察時機與觀察方法，須採用與對照病人所接受的觀察相同。為降低採樣偏差，對照組的挑選必須在執行比較分析之前即完成。這一點並非永遠可行，因為這些對照組的結果可能早已公開發表。任何對挑選標準的配對或為因應受試族群間差異而作的調整，都應在挑選對照組與執行試驗之前即已訂定。在沒有明顯的單一最佳外部對照存在的情況下，最好能採用複數的外部對照，且在分析計畫中預先保守的規劃如何將各個外部對照作為參考（例如試驗組須遠優於結果最佳的對照組以確立其療效）。在某些情況下，由一群經盲性的獨立評審依通用的標準重新評估對照組與試驗組的試驗指標或許有所幫助。

### 2.5.3、倫理道德考量

若新藥品欲治療尚無令人滿意治療方法的嚴重疾病，尤其是依據期理論上的考量、動物數據、或早期人體經驗判斷該藥品似乎頗具潛力的情形下，可能因具有沒機會接受該新藥治療的病人同步對照組，而使執行對比性試驗的意願合理的偏低。同時，執行實際上不可能證實療效的試驗亦是不負責且不合倫理。須知許多有潛力的療法在對照試驗中的效果遠不如預期或完全未能顯出療效。在此類情況下，試驗主持人可能將面臨極為困難的抉擇。在某些特例

中，展開外部對照試驗，希望發生具說服力的戲劇性效果，一但不如預期則迅速切換至隨機分派試驗可能可行。

或者，在面臨尚無令人滿意的治療方法的嚴重疾病，而病程亦無法準確預測時，大體而言較可取的做法是即便在最初期的臨床試驗亦採用隨機分派設計。若能在產生該療法有效的印象之前即展開試驗，這一點通常是可以達成的。試驗可由獨立的資料監測委員會進行監測，以期能及早發現戲劇性的效益。同步對照試驗可以極快的速度發現極端效果，並且也可發現中等但仍具價值的效果，而此中等效果無法由外部對照試驗以令人信服方式的方式展現。

#### **2.5.4、外部對照試驗之適用性與在特定情況下之推論強度**

一般而言，唯有在對試驗治療優於一切現有其他治療的看法強大到令其他試驗設計顯得難以接受，而且欲治療的疾病或病況有紀錄完善且高度可預測的病程時，才可以考慮採用外部對照試驗。即便在這種情形下，往往還是能改用隨機分派同步對照設計（見第 2.1.5 節）。

外部對照試驗最可能具說服力的狀況是，當試驗指標為客觀指標、試驗治療的臨床結果與外部對照的臨床結果明顯不同且試驗治療與對照的比較達到極高的統計顯著性、影響疾病結果的共變量已有完善分析、並且對照組在所有已知的相關基期、療法（除試驗用藥外）、觀察變項都與試驗組高度相似時。然而，即使在這樣的情形下，仍然有此類試驗做出錯誤結論的案例在文獻中記載。

考慮採用外部對照試驗時，採用適當的試驗設計與執行方法或許有助於減少偏差（見第 2.5.2 節）。

#### **2.5.5、試驗設計的調整以及合併其他對照以解決倫理道德、實務、推論方面的問題**

外部對照設計可以透過採用安慰劑對照的隨機退出階段來整合隨機分派與盲性等元素，通常可加入及早退出的規劃，如同先前所述（見第 2.1.5.2.4 節）。在治療初期對治療產生反應的受試者可維持繼續接受治療，使得治療初期的結果可以被嚴謹且大致上並無假設(assumption)與偏差的試驗證實。

#### **2.5.6、外部對照試驗的優點**

外部對照試驗的主要優點是所有病人都能接受有潛力的藥品治療，使得該試驗對病人與醫師更有吸引力。

外部對照試驗的設計可能頗具效率，因為所有的病人都會接受試驗用藥治療，這點對罕見疾病特別重要。然而，即便在外部對照試驗中僅有一個試驗組，對外部對照組結果的預估始終必須保守，這樣有可能使樣本數比安慰劑對照試驗所需要的程度更大。外部對照試驗須格外謹慎（例如採用較嚴格的統計顯著水準），因為試驗組與對照組之間可能存在各種已知與未知或不能測量的差異，而且這些差異往往對試驗組較為有利。

### 2.5.7、外部對照試驗的缺點

外部對照試驗不能盲性，而且受制於病人、觀察者、分析者偏差。以上為其主要缺點。雖然有可能在一定程度上減輕這些缺點，但即使是第 2.5.2 節所建議的方法也無法完全解決這些問題，因為治療方法並非隨機分派，而對照組與試驗組在試驗開始時的可比較性、病人在試驗期間所接受治療的可比較性也無法保證或作出完善的評估。外部對照試驗會高估試驗治療方法療效的傾向在文獻中已有許多記載。須注意，在此類試驗中對統計顯著性的檢定並不如隨機分派試驗來得可靠。

## 第三章、同步對照組的選擇

表 1 說明特定對照組類型的適用性，而圖 1 決策樹協助在不同的對照組類型中作選擇。雖然表 1 與圖 1 聚焦於展現療效以選擇對照組，有些設計亦能用以比較試驗藥與對照藥。對照組的選擇會受到療法的可取得性與特定區域的醫療所影響。

在特定情形與特定目的下，各主要對照組類型（安慰劑、活性、劑量反應）的潛在適用性如表 1 所示。該表格須搭配內文對於特定狀況的說明，以瞭解對照組的潛在適用性。所有情形都是在試驗有適當設計的前提下。外部對照組太過特殊，所以並未納入表中。

在多數情形下，展現相對於同步對照治療的較優性是最有說服力的療效證據。若較優性試驗無法實行或基於臨床或實務因素顯得不恰當，而活性對照通常具有明確的治療效果時（例如抗生素在多數情況下），可採用不劣性或相等性試驗，且可具說服力。

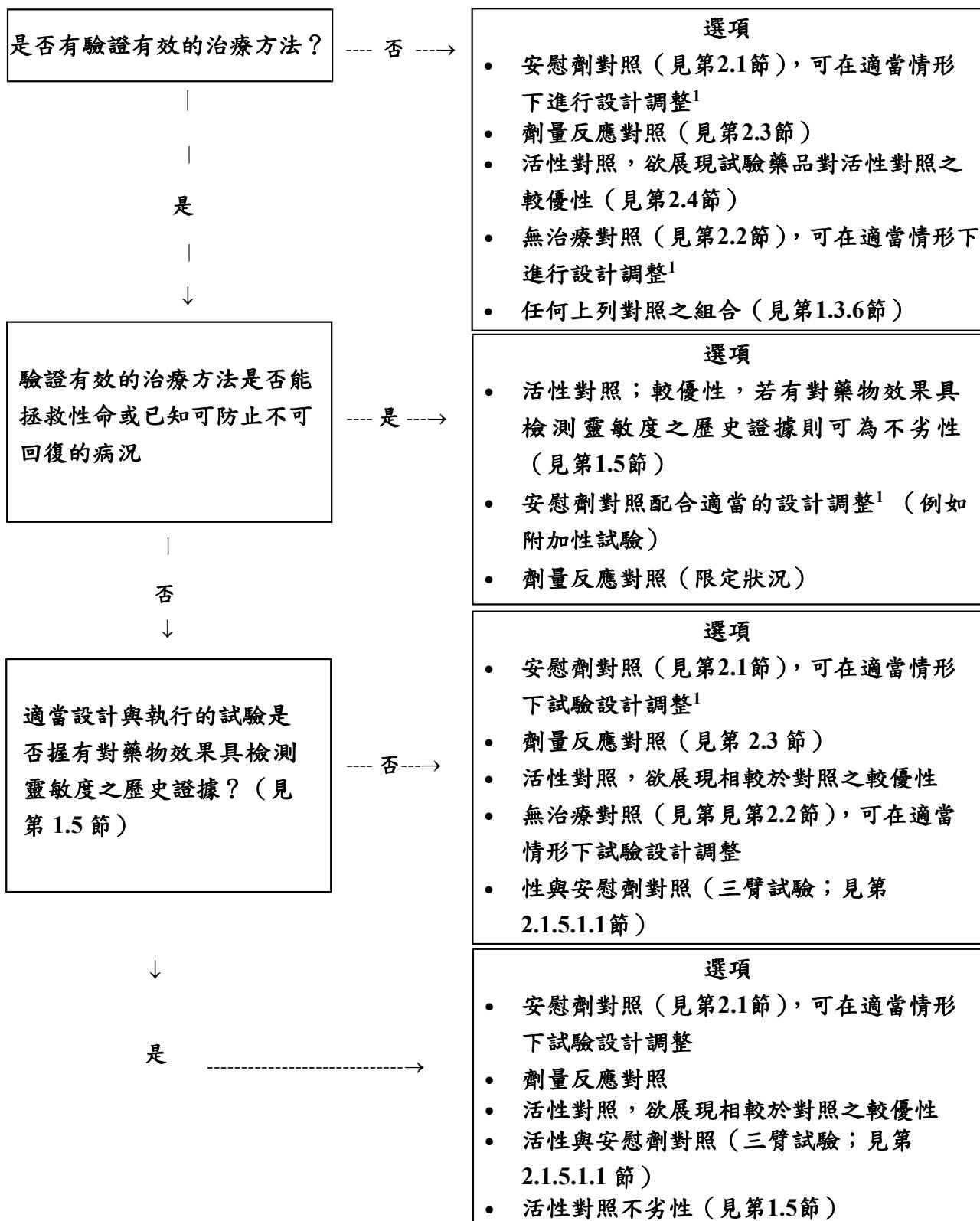
表 1、特定同步對照類型在各種情形下之適用性

		對照類型						
試驗目的	安慰劑 對照	活性 對照 不劣性	活性 對照 較優性	劑量 反應 (D/R)	安慰劑 + 活性	安慰劑 + 劑量 反應	活性 + 劑量 反應	安慰劑 + 活性 + 劑量 反應
測量 「絕對」 療效幅度	Y	N	N	N	Y	Y	N	Y
展現效果	Y	P	Y	Y	Y	Y	Y	Y
展現劑量 反應關係	N	N	N	Y	N	Y	Y	Y
比較療法	N	P	Y	N	Y	N	P	Y

Y=是，N=否，P=有可能，視是否有對藥物效果具檢測靈敏度之歷史證據而定

## 圖1、選擇同步對照以展現療效

此圖所顯示為選擇對照組之基本邏輯。所作決策視可取得之藥品與特定區域之醫療實務情況而定。



<sup>1</sup>附加性試驗、替代、及早退出、短暫的安慰劑時期、隨機退出（見第2.1.5.2節）