

出國報告（出國類別：進修）

多重檢定方法於臨床試驗上之應用

服務機關：國立臺灣大學醫學院附設醫院

姓名職稱：張晉豪

派赴國家：澳洲

出國期間：106年12月11日至107年12月10日

報告日期：108年2月10日

摘要

臨床試驗可採多重指標 (co-primary endpoints) 的研究設計，並假設任一個主要指標達統計檢定之顯著意義時，即宣稱該藥有效。然而這類方法會面臨統計上的問題，亦即多重比較時將會導致第一類誤差的增加。Benjamini 于 1995 年提出 FDR(False Discovery Rate)的方法以避免第一類誤差的增加。面對臨床試驗的多重指標或是多次比較的研究設計，本次前往雪梨大學臨床試驗中心進修報告將探討避免第一類誤差增加的統計方法，以避免增加無效藥上市的機率。此外，雪梨大學臨床試驗中心於 1988 年成立，迄今已達 30 年，具備有完整的臨床試驗人力組織架構。除了能夠提供全方位服務以利臨床醫師進行臨床試驗案，中心內部之研究人員亦有能力與臨床醫師合作，自行發起臨床試驗。在相關事務進行穩定之後，近年亦成立臨床試驗學位學程，以培養更多的臨床試驗人才。台大醫院臨床試驗中心未來除了可透過其發展經驗及人力組織架構作為借鏡進行調整，亦可尋求在臨床試驗各方面事務進行合作的可能性，藉以增加國際競爭力。

目次

壹、目的.....	4
貳、重要過程紀要.....	5
參、心得及建議.....	10

壹、 目的

一個嚴謹設計的臨床試驗，主要指標若達統計上之顯著意義，即表示此新藥或新醫材可宣稱達到疾病治療效果。在臨床試驗的研究設計中，主要指標只能有一個。如果臨床試驗有超過一個以上的主要指標，並假設任一個主要指標達統計檢定之顯著意義時，即宣稱該藥有效，此時接受一個無效藥上市的機率將會增加，亦即檢定時所犯的第一類誤差 (type I error) 會增加。此外，常見導致第一類誤差增加的原因，包括：多組治療組、期中分析、及次族群分析(在精準醫療的目的下尤為常見)。

許多新藥在發展後期時，針對單一主要指標的設定，往往需要花費許多的時間及金錢來加以摸索，而後得到最佳主要指標。這類傳統上的作法往往損及生技產業的經濟效益，進而延緩有效藥的上市，因而不利於民眾對於疾病治癒的期望。

為能加速新藥上市流程，有學者建議臨床試驗的研究設計，除了單一的主要指標外，可採多重指標 (co-primary endpoints) 的研究設計。然而這類方法需要校正第一類誤差以避免第一類誤差的增加，因此也面臨過於保守的問題。

多重比較所導致第一類誤差增加的問題，在全基因檢定的研究中，亦常常被提起。Benjamini 于 1995 年提出 FDR(False Discovery Rate)的方法以避免第一類誤差的增加。面對臨床試驗的多重指標研究設計，本次進修將探討減少第一類誤差增加的相關統計方法，以避免增加無效藥上市的機率，進而縮短主要指標之摸索過程。進而可增加未來藥物上市之時效性，有利於未來新藥之開發。

臺大醫院臨床試驗中心在台灣已屬於頂尖的臨床試驗中心之一，但為能更進一步與國際臨床試驗中心接軌並增加國際競爭力，學習國際頂尖臨床試驗中心之運作經驗，亦相當重要。本次進修目的亦將學習雪梨大學臨床試驗中心的運作經驗，以供臺大醫院臨床試驗中心之參考。

貳、 重要過程紀要

一、雪梨大學臨床試驗中心簡介

雪梨大學臨床試驗中心，係由澳洲國家衛生及醫療研究委員會 (National Health and Medical Research Council, NHMRC) 於 1988 年出資成立，迄今已達 30 年。NHMRC 並每年補助經費協助運作，所以其全名為 NHMRC-CTC。此外，該中心也從政府或其他研究贊助單位申請經費。其任務包括主動進行多中心臨床試驗、為研究醫師提供臨床試驗全方位之服務、與國內或國際的臨床試驗案進行合作、協助其他單位的臨床試驗、將臨床試驗研究結果進行統合分析、對已完成的臨床試驗研究進行轉譯醫學的研究等。

目前該中心有四位負責人，其中以 Prof. John Simes 為主要負責人。John Simes 為醫學博士，從 CTC 成立時即被延攬，將其大半生貢獻於此單位。目前他正於倫敦的臨床試驗中心訪問一年，在 CTC 成立 30 周年的慶祝茶會上，特地回來進行感性而簡短的演講。NHMRC-CTC 具有完整的臨床試驗人力組織架構，所以能夠提供臨床試驗全方位之服務或進行臨床試驗研究案。這些單位包括:

- (1) Oncology program management
- (2) Site management team
- (3) Quality assurance team
- (4) Data management team
- (5) Publication team
- (6) Systemic review team
- (7) Business director
- (8) Clinical trial program director

關於雪梨大學臨床試驗中心的更多詳細資料，可上網查詢 <http://www.ctc.usyd.edu.au/index.htm>

NHMRC-CTC 成功的要件除了有卓越的臨床試驗領導人才、研究團

隊、組織架構、運作系統、教育訓練以外，IT 人員的參與抑是不可或缺，因此平日即讓 IT 人員一起參與討論，形成共榮的策略團隊。成立 30 年來，在全體團隊持續的努力下，已經創造許多醫療及學術上的成就。

二、雪梨大學臨床試驗中心學習內容

(一) 進修單位

本次於雪梨大學臨床試驗中心進修，主要學習單位為該中心項下之生物統計部門。該單位負責人為 **Professor Val GebSKI**，目前有十一位統計分析師。由其中一位統計分析師，**David Espinoza**，負責協助我介紹及認識團隊成員，並解決學習期間所產生之各式問題。我參與並協助目前正進行且涉及資料處理，隨機分派及統計分析等相關事務之臨床試驗案。因此與許多統計分析師與臨床醫師皆有學術及實務上的交流，例如，**Andrew Martin**，**Chris Brown**，**Kristy Robledo**，**Rachel O'Connell** 及 **Chee Lee** 等人。

在參與臨床試驗案過程中，由於業務需要，亦曾經參與不同單位的會議，例如，資料管理部門，成本效益學術團隊，跨國臨床試驗案之線上會議等。

(二) 臨床試驗實務學習內容

在參與多個臨床試驗案後，除了涉獵多個重要主題，並充分了解這些主題在臨床試驗實務上的重要性，以避免違反科學精神以及產生臨床試驗案的無效結果。分述如下：

1、隨機分派

1-1 協助臨床試驗團隊進行隨機分派

1-2 建立隨機分派系統並使主持人無法接觸

2、病例報告表 (CRF) 的重要性

它不僅應該用於回答臨床問題的問題，還應該可用於回答：

2-1 臨床醫生的潛在問題

2-2 統計人員的統計問題

2-3 流行病學家的問題

3、統計分析計劃（SAP）

3-1 了解臨床試驗的基本生物或醫療概念，而不僅是統計分析

3-2 資料分析前需撰寫統計分析計劃書

4、臨床資料管理系統（CDMS）的必要性

臨床試驗務必使用CDMS，它的最低要求，包括：

4-1 資料管理及協調

4-2 資料庫設計

4-3 資料品質掌控

5、成立資料管理團隊的必要性

5-1 確保資料準確無誤

5-2 確保資料來源可信

6、分析結果輸出

6-1 分析結果顯示在螢幕，再手動輸入分析結果，會增加錯誤的可能性

6-2 無需手寫且直接從輸出中製作表格是必要的

7、詳細記錄統計過程

7-1 每個結果表應包括分析程式碼所存放的位置。

7-2 資料管理程式碼和資料分析程式碼應分別撰寫，避免混淆

7-3 應在每個資料管理的程式碼中，明確記錄每個資料的說明

(三) 多重比較學術研究學習內容

除了臨床實驗相關事務的學習之外，針對臨床試驗多重比較之學術研究問題，與 Prof. Val GebSKI 與臨床醫師 Chee Lee 進行了討論研究及實證資料分析。概述如下。

1、背景

一個嚴謹的臨床試驗，往往只能有一個主要目標(primary end point)。然而許多病因複雜的疾病，例如，癌症，自體免疫，心血管疾病等等，通常需要多個主要目標，或是多個次要目標(secondary end point)才有辦法準確描述新治療方法所得到之治療效益。在這種情況下，就會產生多重檢定的問題。此外，臨床試驗常採用的設計如，多組治療組、期中分析、及次群體分析(在精準醫療的目的下尤為常見)等等，同樣會產生多重檢定的問題。

多重檢定會增加某次檢定為顯著卻是假陽性的可能性。研究中進行的比較或檢定越多次，則達到統計上顯著差異的結果，就越有可能是錯誤的。因此為了能夠增加臨床試驗結果成功的可能性及推論的可信度，在多重性比較時所選擇的統計方法，將會是重要的影響因素。選擇正確的統計方法的使用，也意味著減少無效藥物上市的可能性，亦能減少無效持續研究的機會。

世界各國的藥物主管機關亦已經注意到，臨床試驗中解決多重比較的重要性。美國食品和藥物管理局(FDA)及歐洲藥品管理局(EMA)皆於 2017 年分別發布了關於處理臨床試驗多重比較的綜合指南。對於臨床試驗的要求，須能預先控制 I 型誤差的適當統計方法。新英格蘭醫學期刊(NEJM)亦於 2018 年刊載相關論文，尤可見此議題於臨床試驗上之重要性。

2、適當的統計方法

為避免統計方法之細節過於瑣碎，將不會描述統計方法使用之細節，僅提供統計方法名稱，供讀者參考。

2-1、安全性評估臨床試驗

對於安全性評估的二期臨床試驗，多重檢定不常發生。然若需進行多重檢定時，可採用二次式的 FDR(False Discovery Rate)，對於安全性可產生更嚴謹的評估。

2-2、型 I 誤差的調整

目前許多統計方法用在 III 期臨床試驗上，都是用來校正多次檢定下的型 I 誤差，使其能維持在常用的 5% 而不被放大。

2-2-1、 Fixed-sequence test: 若是多個 endpoints 的重要性能夠事先瞭解，也就是多次的檢定是有重要性上的先後順序，且試驗前即被清楚定義，此時可採用此統計方法。

2-2-2、 無母數檢定法: 例如，Bonferroni, Holm, 和 fallback 等方法。這些方法可適用於各種臨床試驗類型下所面臨的多重檢定，然而其過於保守。整體而言，可能會降低試驗成功的機率。

2-2-3、 半母數檢定法: Hochberg 和 Hommel 等方法。相較於無母數檢定法，半母數檢定法將可獲得較高之檢力，然而需有多次檢定彼此需要是獨立的假設。臨床試驗中之多組劑量比較或是多個次群體病人比較，多採用此法。

2-2-4、 母數檢定法: Regular Dunnett 和 step-down Dunnett 等方法。這是檢力最高的統計方法，然而統計假設也更嚴謹，需要先事先給定各個檢定之間彼此的相關性。臨床試驗中之多組劑量比較或是多個次群體病人比較，在知道統計假設的前提下，多採用此法。

2-3、多種因素組成的多重比較

由於多種因素組成的多重比較，例如，幾個主要目標，幾種劑量比較，所組成的多重比較。或者可簡化為，不同劑量下，數個主要目標的檢定，以及不同劑量的比較。這種具挑戰性的臨床試驗研究設計，可採用 Advanced gatekeeping 檢定法，以解決此類多重比較的問題。

題。

2-4、次群體病人相互比較

在臨床試驗中進行次群體分析的方法，可以是驗證性也可以是探索性的分析策略。驗證性次群體分析適用於晚期臨床試驗，亦可評估幾個預先定義的次群體病人之治療效果。探索性次群體分析則多以臨時方式發現治療效果或安全性的新現象。

隨著越多的標靶治療藥物正在開發，許多 3 期試驗已經進行次群體分析的研究設計。在這類試驗中，除了全部病人之外，另外已經預先指定兩個或更多個次群體，分別研究新治療方法在次群體中的療效。同樣會產生多重比較的問題。

次群體病人的分群策略，亦會影響統計方法的選擇。若病人分群因子是依照個人特性，分群因子之間並無相關，例如病人的基因型或性別。各個子群體病人可分別做一次檢定，全部病人亦可做一次檢定。前面的檢定彼此是獨立的，但若包含全部病人的檢定，則會有相關性的產生。另外，若是病人是依嚴重程度來分群，亦有不同的考量。例如若新藥物在嚴重的病人有效，即可假設在輕微病人以及全部病人中，皆為有效。這些考量皆可透過以上的統計方法中，選取適當的方法，以解決多重比較的問題。

3、實證資料分析

由於研究資料尚未發表，所以僅以簡述來形容多重檢定的於實證資料上的重要性。由於雪梨大學臨床試驗中心有一跨國型的肺癌臨床試驗，台大醫院亦是其中的一個合作中心，因此我們以此臨床試驗所收集的上千位肺癌個案，進行實證分析。我們嘗試瞭解接受臨床試驗的三組病人，於參加試驗後六個禮拜的疾病概況進行分組。探討復發，部分反應，完全反應的病人，探討其於六個禮拜之後的存活概況，是否受到復發前的治療藥物，復發時的腫瘤的惡化情況，以及不同型

別肺癌病人，進行無病存活比較及總存活比較。這樣的研究主題，將產生多重比較的問題。透過多重比較的校正後，將原本很多的顯著結果減少成個位數。未來準備再透過其他肺癌資料來驗證這些存在顯著差異的因子是否仍然顯著，及被刪去的顯著因子是否真的不顯著。

參、心得及建議

臺大醫院臨床試驗中心在台灣已屬於頂尖的臨床試驗中心之一，但為能成為國際頂尖的臨床試驗中心，學習雪梨大學臨床試驗中心之運作經驗，亦相當重要。雪梨大學臨床試驗中心具備有完整的臨床試驗人力組織架構，可提供全方位服務以利臨床醫師進行臨床試驗案，甚至內部研究人員亦有能力與臨床醫師合作，自行發起臨床試驗。

目前台大醫院臨床試驗中心亦提供相當多的服務給予臨床醫師，例如提供研究護理師人力，協助檢視合約、提供研究經費、建置後端付款平台、臨床試驗稽核、統計諮詢、統計分析、臨床試驗教育訓練、論文寫作諮詢等等。然而受限於經費問題，仍有許多執行層面值得借鏡雪梨大學臨床試驗中心的經驗並進行調整。例如：

1、提供臨床試驗專案管理師。

減少研究醫師需要負擔的諸多行政流程，可讓其專心且願意參與臨床試驗案的進行。

2、提供資料管理團隊。

臨床試驗之研究資料，不論是勘誤或是輸入錯誤，都需要有專人進行管理及修正，而非由個人自行修改，以建立資料的可靠性及公平性。目前臺大醫院執行臨床試驗之醫師，對這一方面的認識大多缺乏。即使有這方面的認知，受限於經費及人力，也大多是自行修改及管理。

3、協助建置電子資料庫系統。

雪梨大學臨床試驗中心以往都使用InForm商業軟體進行修改，而建置成屬於自己專用的資料庫管理管體，以管理臨床試驗案所收集的病例報告表(case report form, CRF)。然這套商業軟體收費昂貴，因此去年已經決議改採另外一套較為便宜且專業的商業軟體。目前臺大醫院臨床試驗中心，去年已經開始推廣使用專業的電子資料庫系統，

REDCap。為了降低研究者使用電子資料庫的門檻，應有專門人力協助臨床醫師建立系統，並協助指導資料輸入人員。

4、協助進行隨機分派

隨機分派的進行，不同臨床試驗在執行層面上，有各自的不同的需求及工具，所以無法完全一致。應協助及檢視研究者是否達成正確的隨機分派程序，以避免產生研究偏差。

5、提供正確的統計分析流程

臨床試驗在資料分析時，所採用的統計方法，不僅是使用是否正確的問題，還有使用的時機點應該要能夠受到公開審查的。目前多數研究醫生大多自行分析資料，且沒有注意時機點，會產生多次檢定或是球員兼裁判的問題。

6、建立使用者付費的觀念

雪梨大學臨床試驗中心的所有服務都是需要付費的，有收費標準但是會有折扣優惠。唯有收費的服務，才能讓使用者及提供者雙方更加重視服務內容。目前臺大醫院臨床試驗中心所提供之服務，大多是免費的，容易被濫用而造成研究資源的浪費及分散。應建立收費標準以讓使用者珍惜這些服務，台大臨床試驗中心也才有更多的資源以提供更好的服務。