



生物相似性藥品之臨床研發策略-以trastuzumab為例

許巧縈、林重賢、陳紀勳¹

前言

近年來隨著多項生物藥品的專利到期，生物相似性藥品的研發與製造相應而生，但由於生物藥品的結構龐大複雜，且具有製程專一(process specificity)的特性，即當製程改變時，產品就會有所不同。因此生物相似性藥品無法在科學上達到與參考藥品完全相同(identical)，無法比照傳統小分子學名藥之規範，以生體相等性(bioequivalence)結果來連結參考藥品之療效與安全性。目前各國對於生物相似性藥品的查驗登記要求，皆採取「減免部分法規科學要求(abridged)的生物藥品」的概念^[1,2]，透過高度的特性鑑定(characterization)與比較性分析研究(comparability studies，亦有以Comparative Analytical Assessments稱之)，證實生物相似性藥品與參考藥品，具有相似的品質與功能活性(有時須體外生物活性)，對於非臨床與臨床資料的要求，則依個案適度的減免。

為了建立與參考藥品的生物相似性(biosimilarity)，生物相似性藥品的技術性資料主軸為與參考藥品的比較性研究，參考藥品為該成分原開發廠於國內核准的藥品，比較性研究執行項目基本包括結構與功能分析測試、生物效價分析、動物藥動藥效學研究、動物毒理研究、人體藥動藥效學試驗、人體免疫原性(immunogenicity)評估以及人體療效和安全性比較試驗等，採逐步漸進的方式(stepwise approach)進行，形成生物相似性的整體證據(totality of evidence)。

給藥途徑及用法用量須比照參考藥品，所採容器封蓋系統設計須配合用法用量；欲宣稱之適應症不得超出我國參考藥品之適應症範圍。

本文章旨在探討生物相似性藥品在進行各項非臨床比較性研究，以及臨床藥動藥效學比較性試驗後，規劃臨床療效安全性比較性試驗設計的考量重點。以下將此臨床療效安全性比較性試驗，稱為臨床比較性試驗，並以成分為trastuzumab的生物相似性藥品為例進行說明。

¹財團法人醫藥品查驗中心新藥科技組



本文章所指的比較性試驗，其目的並非證實生物相似性藥品與參考藥品之間的可互換性(interchangeability)與可替代性(substitution)。

生物相似性藥品的臨床比較性試驗

生物相似性藥品臨床比較性試驗的目的，在於探討該藥品與參考藥品之間，是否具有明顯的療效和安全性的差異；不同於傳統藥品臨床試驗目的在證明該藥品的療效安全性。

研發一個生物相似性藥品，臨床比較性試驗所須的規模或件數，其決定因素包括：參考藥品結構上的特性和複雜程度、比較性研究的規模及限制、比較性研究所發現的差異、各項差異造成臨床差異的可能性、目前對於參考藥品作用機轉和病理學的了解程度、藥動藥效指標可預測臨床預後的程度、以及參考藥品各項療效安全性的過往臨床經驗等。研發單位應依據上述各項因素的分析，合乎科學的規劃臨床比較性試驗的試驗設計，例如在決定包括臨床指標(endpoints)、試驗族群、制定臨界值等試驗架構時，如何將上述因素納入考量。另外，若參考藥品及所屬分子類別具有特殊的療效安全性顧慮時，法規單位將可能要求額外的比較性試驗。

臨床比較性試驗：試驗設計

比較性試驗須以直接比較來證明生物相似性藥品與參考藥品具有相似的療效，原則上採用隨機分派、雙盲、活性對照組的試驗設計，且具有足夠的檢定力。單以安慰劑做為對照組的試驗無法用以認定與參考藥品之相似性，不應該採用。

比較性試驗原則上採用相等性(equivalence)試驗設計。若可提供科學性及作用機轉證據，明確排除該生物相似藥品成分在該適應症具有療效增強的可能性，亦可採用不劣性(non-inferiority)試驗設計，但建議事先與法規單位諮詢採用不劣性(non-inferiority)試驗設計的合適性。

試驗期間(study duration)及受試者人數的設計，應足以偵測兩藥品間具臨床意義的差異。若採用較靈敏且短期可評估的臨床療效指標做為主要療效指標時，可能需要較短的試驗期間以及較少的受試者人數，但須注意在此情境下試驗期間與受試者人數，可能不足以偵測重要的安全性訊息，如此可能需要額外的安全性及免疫原性的評估資



訊。

對照藥品之選擇

比較性試驗的活性對照藥品原則上，應為生物相似性藥品的參考藥品，參考藥品(以代號R表示)為該成分原開發廠於國內核准的藥品。若欲使用原開發廠其他未經我國核准之製造廠產品(以代號R'表示)作為對照藥品，R'必須已在中央衛生主管機關認可的國家核准上市，且申請者應確認所使用的R'可以代表我國已核准上市的R，應提供適當的數據或資訊並合乎科學地說明R'與R之間的關聯性。此外，生物相似性藥品不得作為參考藥品，因此不可作為臨床比較性試驗中的活性對照藥品。

例如trastuzumab原開發廠藥品之商品名為Herceptin®，上市後分別於美洲和歐洲設廠供應全球市場需求，我國核准的Herceptin®來自歐洲之製造廠，因此Herceptin®美洲製造廠之產品屬於R'。當選擇以R'作為臨床比較性試驗之活性對照藥物時，須確定R'原則上經十大醫藥先進國核准上市，且應在研發前期的比較性研究以及/或臨床藥動藥效學比較性試驗中，同時以生物相似性藥品、R和R'產品執行三方比較性研究，建立R'與R之間關聯性的銜接資料。基於市場需求，目前trastuzumab生物相似性藥品皆有三方比對性資料。

試驗族群之選擇

相對於參考藥品已取得適應症的目標族群，比較性試驗的試驗族群應具代表性，當參考藥品有多個適應症時，生物相似性藥品臨床比較性試驗的試驗族群，應選擇具有足夠的靈敏度，足以偵測生物相似性藥品與參考藥品間可能具臨床意義的差異的適應症族群。一般而言，均質性較高的試驗族群，可以減少對於試驗結果的干擾，從而顯示二組藥品比較的結果主要來自兩組試驗藥品的差異。

以trastuzumab為例，Herceptin®取得的適應症範圍包括早期乳癌、轉移性乳癌和轉移性胃癌。其中，以早期乳癌患者的疾病狀況較一致、體能狀態較佳且沒有接受過先前的治療，被認為較合適做為trastuzumab生物相似性藥品臨床比較性試驗的試驗族群。轉移性乳癌的病人族群，雖然可能在腫瘤負擔、腫瘤轉移部位、身體狀態或是先前接受的治療等因素導致異質性較高，但藉由納入排除條件適當地規範，以及良好的隨機分派機制等設計，亦有trastuzumab生物相似性藥品選擇以轉移性乳癌作為



試驗族群，可參考(表一)US FDA與EMA皆已核准的trastuzumab生物相似性藥品所執行之臨床比較性試驗。

有時為了符合臨床治療實務上的改變，比較性試驗收納的試驗族群可能可以在併用藥物、先前的治療、或疾病嚴重度上，與參考藥品核准適應症的目標族群略有不同，針對此項差異，建議研發單位於試驗設計階段提出合理性說明並與法規單位討論。

療效指標之選擇

生物相似性藥品的臨床比較性試驗，主要療效指標應該採用具靈敏度的療效指標，以偵測生物相似性藥品與參考藥品間可能具意義的臨床差異，與參考藥品之療效確認性試驗所用之主要療效指標可能不同。

例如trastuzumab原開發廠的Herceptin，當年的療效確認性試驗的主要療效指標多採用無惡化存活期(progression-free survival, PFS)或整體存活期(overall survival, OS)，但因為無惡化存活期與整體存活期所需的追蹤時間較長，容易受到非藥品相關因素的影響，例如腫瘤負荷量、病人的體能狀態、先前或後續其他抗腫瘤治療等，因此對於偵測兩組藥品的差異以及建立生物相似性來說，顯得靈敏度不足。雖然如此，原則上比較性試驗仍建議評估無惡化存活期與整體存活期，以支持二藥品間長期療效不具明顯差異。

反應率(response rate)評估藥品短期間的抗腫瘤活性，較不易受到其他因素的干擾，因此在癌症比較性試驗中，常用於偵測藥品間可能存在的具臨床意義的差異。在以轉移性乳癌做為試驗族群的trastuzumab生物相似性藥品臨床比較性試驗中(表一)，已有多項產品選擇以腫瘤的客觀反應率(objective response rate, ORR)做為的主要療效指標。客觀反應為腫瘤縮小達預先定義的標準並維持一段時間，為完全反應(complete response, CR)和部分反應(partial response, PR)兩者的總和，客觀反應率的定義為有客觀反應之病人比率。

但在早期乳癌患者執行的臨床比較性試驗中，則是選擇完全反應率做為主要療效指標(表一)，可能原因為腫瘤切除後達完全反應者相較於整體反應者，與整體存活期的相關性較高。值得注意的是，早期乳癌試驗中常用的完全反應的定義可區分為二種，其中一種為Breast pathological complete response (bpCR)，代表手術切除的乳房



組織中沒有殘留侵襲性腫瘤細胞的存在；以及Total pathological complete response (tpCR)，代表除了乳房組織以外，手術切除的腋下淋巴組織也都沒有殘留侵襲性腫瘤細胞的存在。二種完全反應的定義中，以tpCR與整體存活期的相關性較高，且更嚴謹地代表藥品的抗腫瘤活性。在早期乳癌術前輔助療法(neoadjuvant)新藥臨床試驗中，tpCR為美國FDA用於新藥加速核准(accelerated approval)所要求之替代療效指標。在trastuzumab生物相似性藥品之早期乳癌患者的臨床比較性試驗中，使用tpCR做為主要療效指標為目前之趨勢，因此建議選擇tpCR為trastuzumab生物相似性藥品早期乳癌患者的臨床比較性試驗之主要療效指標。

除上述建議之指標以外，若比較性試驗的主要療效指標欲採用其他指標，例如腫瘤反應的時間(time to response, TTR)，建議研發單位提供該療效指標對於偵測療效安全性差異的敏感性和合適性，以及該療效指標與臨床預後的相關性等科學性論述與法規單位討論。

臨界值(margin)的制定與樣本數的計算

生物相似性藥品之臨床比較性試驗統計設計，不論為相似性試驗或不劣性試驗，皆須事先定義臨界值大小，臨界值不僅關係執行試驗樣本數的大小，也是試驗結果的決策法則，故須審慎評估臨界值的合理性。決定臨界值的考量要點依序如下。

(1) 恆定性假設(constancy assumption)原則

活性對照藥的療效證據通常來自較優性比較試驗，由顯著優於安慰劑組，來確認其療效；然而不論是相似性或不劣性試驗，試驗藥品療效的確認，若缺乏安慰劑組，則須由活性對照藥的療效，間接證明試驗藥品療效。因此，若無法納入安慰劑組，須假設活性對照藥目前執行的臨床比較性試驗中的療效，與之前較優性比較試驗中該活性對照藥所得的療效證據是相近的；此即為「恆定性假設」。因此設計生物相似性之臨床比較性試驗時，須考量該活性對照藥(即參考藥品)是否可能不符合「恆定性假設」，例如觀察到參考藥品在臨床研發階段的療效極小，或參考藥品從前的三期試驗療效結果變化很大等。

(2) 療效指標之組間比較方式



以trastuzumab臨床比較性試驗中常見之療效指標「客觀反應率(ORR)」為例，為群體水平的總結(population-level summary)，屬於二元資料(binary data)中的比率(rate)，以下泛稱 p ，其組間比較方式，可考慮採用風險相對值(Risk Ratio, RR)或風險差異值(Relative Difference, RD)。選用何種比較方式，從統計的觀點，端視其是否符合恆定性假設而定。舉例來說，假設活性對照藥過去執行兩個第三期試驗，試驗1結果兩組 p_1 與 p_2 分別為0.1與0.05；試驗2結果則分別為0.06與0.03，經整理後可得二試驗之RR點估計值皆為0.5，但RD點估計值則分別0.05與0.03，可發現二試驗之RR值較能符合恆定性假設^[1]；反之，若過去之試驗結果顯示RD值較為恆定，則可考慮使用RD值。選用何種比較方式，仍要考量該估計值是否能代表實際臨床的經驗並具意義，經綜合考量後，再謹慎選取，並宜事先與法規單位討論。

(3) 臨界值的考量

假若試驗藥品與活性對照藥品之差異值為 θ (試驗藥品-活性對照藥)，而可接受的臨界值為 $\pm\delta$ (以上下臨界值皆相同為例)，則試驗虛無假說(H_0)與對立假說(H_1)分別為：

$$H_0 : |\theta| \geq \delta \text{ vs } H_1 : |\theta| < \delta \dots \dots \dots (1)$$

可拆解為

$$H_{01} : \theta \leq -\delta \text{ vs } H_{11} : \theta > -\delta \text{ 與 } H_{02} : \theta \geq \delta \text{ vs } H_{12} : \theta < \delta \dots \dots \dots (2)$$

決定相似性試驗的臨界值，可參考不劣性試驗相關法規指引^[1,2]。在決定臨界值的過程，會面臨兩種不同考量，即M1(統計臨界值)與M2(臨床臨界值)兩者。M1代表活性對照藥品於目前試驗中預期的療效，一個完整且具備足夠證據的M1估計方法，須經過系統性文獻搜尋，再進行統合分析後，以適當且較保守之估計值作為M1；M2代表試驗藥品與活性對照藥臨床上可接受的最大療效差異(失去的療效)^[5,8]，值得注意的是，雖然最終臨界值以M2(較M1小)為主，但最終臨界值的選擇須有合理的科學證據^[12]，而經完整評估且證據力強的M1，對M2的選取與試驗結果的詮釋，扮演相當重要的角色^[5]。

若從藥物機轉或其它科學證據上，已充分顯示試驗劑量與副作用頻率多寡之關係甚低，且可明確排除療效增加的可能性，則可與法規單位討論，考慮在相似性試驗中，使用較寬鬆的臨界值^[3]。或採用不劣性試驗設計^[13]的可能性。



(4) 統計分析方法與樣本數考量

依據預先決定的臨界值，且透過公式(2)可知，假說檢定共拆解為二個，在控制型一誤差5%的原則下，若依循生體相等性(bioequivalence)試驗原則，建議可採用二次單尾檢定 (two one-sided test, TOST)方法^[3,14]，透過二個單尾(0.05)檢定的方式，分別檢定公式(2)的兩個假說，以確認與活性對照藥的相似性。若欲比照藥品三期療效確認性試驗，亦可接受型一誤差取較保守的雙尾0.05。此外，若療效指標屬於二元資料(如 ORR)，則估計療效差異之信賴區間時，若採用概似法(approximate methods)，結果判讀須謹慎，因須考慮樣本數是否過小或ORR值接近0或1，前述情況建議可採用精確法(exact methods)估計信賴區間^[15,16]。

相似性試驗樣本數的估計，所須參數除了臨界值與型一誤差採用單尾或雙尾(0.05)外，須特別注意型二誤差(β)或檢定力($1-\beta$)。在計算樣本數時，由於相似性試驗包含二個假說檢定，因此，型二誤差須取 $\beta/2$ ；故一般而言，其它參數不變條件下，進行相似性試驗所需樣本數往往較不劣性試驗多^[14,17]。

安全性評估

生物相似性藥品的臨床療效即使與參考藥品相似，仍可能具有不同的安全特性。在臨床比較性試驗中，除了療效的相似性以外，同樣需展現生物相似性藥品與參考藥品的安全性相當(comparable)，尤其針對該生物藥品已知的風險以及免疫原性，謹慎比較生物相似性藥品與參考藥品所產生的不良反應之類型、嚴重度與頻率。

此外，應依據在結構與功能分析和動物的比較性研究中所觀察到的差異，評估生物相似性藥品可能特有的安全性疑慮，尤其是輸注反應與免疫原性，並在臨床試驗中進行特別監測與評估。

雖然trastuzumab分子在單株抗體藥品中屬於免疫原性較低，且非屬終身長期施打的生物藥品，臨床上發生需要轉換(switch)藥品的可能性較低，仍建議針對免疫原性進行完整的評估，研發單位應以最先進之技術為基礎，研發出生物相似性藥品與參考藥品之免疫原性檢測方法，執行生體含量分析方法確效，於比較性試驗中以盲性方式進行抗藥品抗體(antidrug antibody, ADA)檢測，分析ADA的發生率、效價及特性，提供免疫原性與療效、安全性之關聯性分析比較。



生物相似性藥品與參考藥品之間較微小的差異，對於臨床預後所造成的影響，往往必須等到較多的臨床使用經驗後才能獲知，因此生物相似性藥品須有完善的上市後安全監測計畫及上市後免疫原性研究。

結語

雖然生物相似性藥品在我國及歐美日等國皆有類似之研發基本準則，但由於參考品在不同國家可能有不同的核准狀況，且不同法規單位可能會有不同的要求，生物相似性藥品的臨床研發策略須審慎考量各國不同的狀況，並建議於研發階段儘早諮詢中央衛生主管機關及/或醫藥品查驗中心。

表一、US FDA 與 EU EMA 皆已核准的 trastuzumab 生物相似性藥品所執行之臨床比較性試驗

| Biosimilar | Company | N | Patient setting | Primary endpoint | Equivalence margin (E) | Results |
|-------------------------------|--------------------|-----|---------------------|------------------|--|---|
| Kanjinti ABP 980 | Amgen/ Allergen | 725 | Neo-adjuvant EBC | tpCR | -13%, +13% with 90 % CI for RD; 0.759, 1.318 with 90% CI for RR | RD: 7.3% (1.2, 13.4) RR: 1.19 (1.033, 1.366) |
| Herzuma CT-P6 | Celltrion | 475 | MBC | ORR | -0.15, +0.15 with 95 % CI for RD; | RD: 5% (-0.14, 0.04) |
| | | 549 | Neo-adjuvant EBC | tpCR | -0.15, +0.15 with 95 % CI for RD; 0.74, 1.35 with 95% CI for RR | RD: -0.04 (-0.12, 0.05) RR: 0.93 (0.78, 1.11) |
| Ogivri MYL-14010 | Mylan/ Biocon | 500 | MBC | ORR | -15%, +15% with 95 % CI for RD; 0.81, 1.24 with 90% CI for RR | RD: 5.53 (-3.08, 14.04) RR: 1.09 (0.974, 1.211) |
| Trazimera PF-0528014 | Pfizer | 707 | MBC | ORR | 0.8, 1.25 with 95% CI for RR | RR: 1.09 (0.842, 1.049) |
| Ontruzant ⁺ SB3 | Samsung | 800 | Neo-adjuvant EBC | bpCR | -13%, +13% with 95 % CI for RD; 0.785, 1.546 with 95% CI for RR | RD: 10.70 (4.13, 17.26) RR: 1.259 (1.085, 1.460) |

EBC: early breast cancer, MBC: metastatic breast cancer, ORR: objective response rate, tpCR: total pathological complete response, bpCR : breast pathological complete response, RD: risk difference, RR : relative risk or risk ratio, CI: confidence interval.



參考文獻

1. U.S. Food and Drug Administration. Guidance for industry: Scientific Considerations in Demonstrating Biosimilarity to a Reference Product; 2015.
<https://www.fda.gov/media/82647/download> assessed 2020/11/17
2. European Medicines Agency. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues; 2015.
https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-similar-biological-medicinal-products-containing-biotechnology-derived-proteins-active_en-2.pdf assessed 2020/11/17
3. 生物相似性藥品查驗登記基準; 2015.
http://regulation.cde.org.tw/doc_data_display?sid=2349&doctype2= assessed 2020/11/17
4. 生物相似性單株抗體藥品查驗登記基準; 2015.
http://regulation.cde.org.tw/doc_data_display?sid=2409&doctype2= assessed 2020/11/17
5. U.S. Food and Drug Administration. Guidance for industry: non-inferiority clinical trials; 2016.
<https://www.fda.gov/media/78504/download> assessed 2020/11/17
6. European Medicines Agency. Guideline on the choice of the non-inferiority margin; 2006.
https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-choice-non-inferiority-margin_en.pdf assessed 2020/11/17
7. 吳雅琪. 不劣性試驗統計審查重點. 2012; RegMed, 22.
8. 吳雅琪. 淺談美國FDA不劣性試驗指引最終版之修訂. 2017; RegMed, 77.
9. Isakov L, Jin B, Jacobs IA. Statistical primer on biosimilar clinical development. Amer J Ther. 2016; 0:1–8.
10. 林資荃. 統合分析(Meta-analysis)簡介. 2014; RegMed, 48.



11. PRISMA 2009 Checklist; 2009.
<http://www.prisma-statement.org/documents/PRISMA%202009%20checklist.pdf> assessed 2020/11/17
12. He K, Chen H, Gwise T et al. Statistical considerations in evaluating a biosimilar product in an oncology clinical study. Clin. Cancer Res. 2016; 22(21), 5167–5170.
13. 陳紀勳,詹明曉.生物相似性藥品之臨床考量. 2018; RegMed, 92.
14. World Health Organization. Guidelines on Evaluation of Similar Biotherapeutic Products (SBPs); 2015.
http://www.who.int/biologicals/areas/biological_therapeutics/BIOTHERAPEUTICS_FOR_WEB_22APRIL2010.pdf. assessed 2020/11/17
15. Chan, I. S. F. and Zhang, Z. "Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions" . Biometrics. 1999; 55: 1202–1209.
16. Santner, T. J., Pradhan, V., Senchaudhuri, P, Mehta, C. R. and Tamhane, A. C.. Small-sample comparisons of confidence intervals for the difference of two independent binomial proportions, Computational Statistics & Data Analysis. 2007; 51, 5791–5799.
17. Chow S-C, Shao Jun, Wang Hansheng. Sample Size Calculations in Clinical Research. Second edition. Chapman & Hall/CRC, Taylor & Francis Group, 2008.