



台灣藥物法規
資訊網法規公告



台灣藥品
臨床試驗資訊



TFDA 藥物
食品安全週報



致力法規科學
守護生命健康

Regulatory Science, Service for Life

統計工具於藥品品質比對之應用

張凱琳¹

前言

統計工具雖已廣泛應用於藥品臨床試驗及生體相等性試驗等療效相關之評估，不過，在藥品製造品質管制領域方面，雖有國際規範如 ICH Q6A 建議以信賴界限 (confidence limit) 訂定品質屬性之允收標準^[1]，ICH Q1E 建議以 95% 信賴界限推定架儲期^[2]，但並未詳盡探討或制訂統計工具之應用範圍。近年來，由於生物相似性藥品之興起，歐盟 EMA 於 2017 年 3 月發布「Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development」草案，其中探討比較各項品質屬性差異時會採用之統計工具，並說明各項工具之優劣點^[3]。美國 FDA 亦於 2017 年 9 月發布「Statistical approaches to evaluate analytical similarity」草案，其中直接規定生物相似性藥品與對照藥品比較各項品質屬性之統計方法及可接受之範圍^[4]。該草案雖因收到各方意見認為其中規定可能影響生物相似性藥品研發之成本與效益，由美國 FDA 於 2018 年 6 月撤回^[5]，但仍可一窺美國 FDA 對各種統計工具之考量。

針對藥品製造方面之品質屬性比對，常用之統計工具有屬於敘述統計 (descriptive statistics) 之最大值最小值範圍 (Min-Max range)，以常態分布為基本假設之品質範圍 (quality range)，可用於推論統計 (inferential statistics) 之區間估計 (interval estimation)，包含信賴區間 (confidence intervals)、容忍區間 (tolerance intervals) 以及預測區間 (prediction intervals)。以下針對上述各項工具，依歐盟 EMA 及美國 FDA 之考量，整理其計算方式、應用範圍及優缺點探討。

最大值與最小值範圍

當品質屬性具有 3 個以上的數值時，即可觀測出一組數據中最大值與最小值。表 1 為一組虛擬數據，使用 Microsoft Excel 2010 的 MAX 與 MIN 函數功能，可快速列出

¹ 財團法人醫藥品查驗中心藥劑科技組



表 1 數據中最大值與最小值。

表 1、最大值與最小值計算方法

Lot	Percent
01	100.23
02	104.52
03	98.57
04	99.61
05	101.24
06	100.95
07	100.55
08	99.33
09	102.41
10	100.22

最大值 (=Max [B2:B11]) =	104.52
最小值 (=Min [B2:B11]) =	98.57

最大值與最小值為大部分藥品製造廠最常使用之比對工具，除了顯示數值之外，亦常以圖表方式呈現，無須複雜計算，僅須觀測數值分布範圍，並以所觀測範圍為準，比對製程變更後品質屬性是否仍在變更前觀測範圍之內，或者比對學名藥或生物相似藥品是否落在對照藥品檢測範圍之內。雖然使用上簡便，深受藥品製造廠青睞，但不論是歐盟 EMA 或是美國 FDA，均不認為使用最大值與最小值範圍比對相似性具有足夠可信度。歐盟 EMA 認為，當用於比對之批次數目較少時，此法反而有較高的機會達成具有相似性之推論，可能造成偽陽性機率增加(兩者實際上不相似，卻做出兩者相似之推論)^[3]。美國 FDA 則直接在「Statistical approaches to evaluate analytical similarity」說明，並不建議使用最大值與最小值範圍作為 tier 2 之比較方法^[4]。由美國 FDA 統計審查員所發表的學術論文「Statistical considerations in setting product specifications」^[6]則更進一步指出，此法在少量樣本數目下變異性極大，當樣本數目增加時，又無法收束在固定範圍內，無論樣本數目多寡，涵蓋範圍均與樣本數目不相關。換言之，樣本數少時，此法無法清楚界定使用者的風險(實際上不相似卻判定為相似，假設檢定之 type I error)



及製造廠的風險(實際上相似卻判定為不相似, 假設檢定之 type II error); 樣本數多時, 使用者的風險又可能因所界定的範圍接近樣本群體 100%範圍而增加^[6]。結論與歐盟 EMA 看法相同。

品質範圍

品質範圍乃美國 FDA 於「Statistical approaches to evaluate analytical similarity」所使用辭彙^[4], 於歐盟 EMA 「Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development」則稱為 x -sigma 方法, 直接指出計算中最重要標準差參數^[3]。Eq. (1)為品質範圍之計算式, 表 2 引述表 1 的虛擬數據, 使用 Microsoft Excel 2010 的 AVERAGE 與 STDEV 函數功能, 計算三倍(表 2 中 x 之數值)標準差之品質範圍。

$$\text{品質範圍} = \text{樣本平均值} \pm x \times \text{樣本標準差} \quad \text{Eq. (1)}$$

表 2、品質範圍計算方法

Lot	Percent		
01	100.23	最大值 (=Max [B2:B11]) =	104.52
02	104.52	最小值 (=Min [B2:B11]) =	98.57
03	98.57	平均值 (=AVERAGE [B2:B11]) =	100.763
04	99.61	標準差 (=STDEV [B2:B11]) =	1.696618
05	101.24	$x =$	3
06	100.95	品質範圍下限 (=Eq. [1]) =	95.67315
07	100.55	品質範圍上限 (=Eq. [1]) =	105.8529
08	99.33		
09	102.41		
10	100.22		

歐盟 EMA 在「Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development」提出 x 可為 3、



4 或 6，但並未特別建議各數字適用之狀況^[3]。美國 FDA 之「Statistical approaches to evaluate analytical similarity」則說明 x (在撤回的指引草案裡代號為 X) 應與主管機關討論合適性^[4]。若假設品質屬性母群體為常態分布， $\pm 3 \text{ sigma}$ (母群體之標準差) 的範圍即可涵蓋母群體 99.9%，因此多數藥品製造廠常選擇 $x = 3$ 作為範圍評估參數。同時，在 ICH Q6A 「Decision tree #1: Establishing acceptance criterion for a specified impurity in a new drug substance」所提到信賴界限亦為「three times the standard deviation of batch analysis data」^[1]，即品質範圍的計算方式，由於 ICH 相關指引之規定，亦擴大了品質範圍的應用。但歐盟 EMA 認為，母群體為常態分布之假設已有不確定性，並非所有品質屬性檢測數值均可確認母群體分布為常態分布。其次，所選擇涵蓋範圍大小並無可靠的科學論述^[3]。舉例而言， $x = 6$ 論述基礎主要為製程效能研究 (process capability study)，該研究允許製程平均值在長期製造過程中，偏離最初製造時平均值至 1.5 倍 sigma。為了涵蓋偏離後的平均值以及分布，最初合格範圍會設定至 6 倍 sigma^[7]。6 倍 sigma 方法 (six sigma methodology) 藉由擴大合格範圍，以減少缺陷品之比率，但卻未考量使用者之風險。允許平均值偏離 1.5 倍 sigma 之依據，主要為汽車工業製程常見之設定值，尚無嚴謹科學證據支持其可適用於製藥工業。

美國 FDA 統計審查員所發表的學術論文「Statistical considerations in setting product specifications」將品質範圍以「參考區間(reference intervals)」表示^[6] (為歐盟 EMA 2011 年發布之「Report on the expert workshop on setting specifications for biotech products」所使用辭彙^[8])，在論文中亦如歐盟 EMA 觀點，提出對品質範圍的疑問。美國 FDA 統計審查員從 Monte Carlo simulation 執行結果發現，少量抽樣樣本所得到品質範圍，會廣於實際母群體的範圍^[6]。實際運用品質範圍計算時，在品質屬性評估中通常僅有少量批次可執行檢測之情況，使用品質範圍建立對照藥品或變更前藥品之母群體預測值，可能與實際母群體範圍有落差。

信賴區間

信賴區間為常用於區間估計之統計工具。Eq. (2) 為抽樣樣本數目較少時 (通常為樣本數，或批次數少於 30 之情況) 信賴區間計算式^[9]。再引述表 1 之虛擬數據，使用 Microsoft Excel 2010 的 TINV 函數，計算 95% 雙尾信賴區間之上下限如表 3 所示。

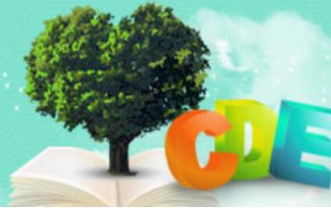


$$\text{信賴區間} = \text{樣本平均值} \pm t_{(\alpha/2)} \times \text{樣本標準差} / \sqrt{\text{樣本數}} \quad \text{Eq. (2)}$$

表 3、信賴區間計算方法

Lot	Percent		
01	100.23	最大值 (=Max [B2:B11]) =	104.52
02	104.52	最小值 (=Min [B2:B11]) =	98.57
03	98.57	平均值 (=AVERAGE [B2:B11]) =	100.763
04	99.61	標準差 (=STDEV [B2:B11]) =	1.696618
05	101.24	$x =$	3
06	100.95	品質範圍下限 =	95.67315
07	100.55	品質範圍上限 =	105.8529
08	99.33	$\alpha =$	0.05
09	102.41	$t(\alpha/2)$ (=TINV [E8, 10-1]) =	2.262157
10	100.22	信賴區間下限 =	99.54931
		信賴區間上限 =	101.9767

樣本群體的信賴區間乃用於估計母群體平均值，以 95%信賴區間為例，所代表意義為執行 100 次抽樣中，有 95 次抽樣群體所計算得到的信賴區間，可以涵蓋母群體的平均值^[9]。信賴區間在藥品開發上應用廣泛，以生體相等性評估為例，為確認對照藥品與試驗藥品是否具有生體相等性，即以兩抽樣樣本群體平均值之差為評估參數，確認其 90%信賴區間是否落在法規規定為相等之範圍內。但在品質領域中，信賴區間的應用並不普及。由表 3 可知，與品質範圍與最大值最小值範圍相較，信賴區間範圍最狹窄，雖然對使用者有利(不易使用到缺陷品)，但卻可能造成產品合格率過低，整體產率不足之狀況。即使排除在信賴區間範圍之外，仍落在品質範圍或最大值最小值範圍內之產品，是否就會影響產品之安全性與療效，亦有爭議。歐盟 EMA 認為信賴區間主要應用在臨床試驗之相等性評估或非劣性(non-inferiority)評估之中^[3]，美國 FDA 在「Statistical approaches to evaluate analytical similarity」亦應用了信賴區間評估屬於 tier 1 品質屬性之相等性^[4]。兩大法規單位均認為信賴區間主要應用於相等性評估。



致力法規科學
守護生命健康
Regulatory Science, Service for Life

容忍區間

近年來生物藥品的品質屬性允收範圍評估，除了品質範圍之外，亦常見容忍區間之應用。Eq. (3)為抽樣樣本數目較少時容忍區間之計算式^[6]。其中，參數 k 之計算涉及非中心 t 分布(non-central t distribution)，Microsoft Excel 2010 無直接對應之函數，改採用自由軟體 R 3.4.1 版本之 qt 函數計算 k 值。表 4 列出涵蓋範圍 $p = 99.5\%$ ，信心水準 $g = 95\%$ 之下，以 R 軟體計算 k 值之方式，表 5 則應用表 1 之虛擬數據計算容忍區間。

$$\text{容忍區間} = \text{樣本平均值} \pm k \times \text{樣本標準差} \quad \text{Eq. (3)}$$

表 4、k 值計算方式

> n = 10	#設定樣本數 n 為 10 批
> p = 0.995	#設定涵蓋範圍 p 為 99.5%
> g = 0.95	#設定信心水準 g 為 95%
> f = n-1	#設定自由度 f 為 n-1
> delta = qnorm(p)*sqrt(n)	#設定非中心參數 delta 為涵蓋範圍 p 之下，z 分布(常態分布)之反函數值乘以樣本數開根號： $\delta = z_p \times \sqrt{n}$
> k = qt(g,f,delta)/sqrt(n)	#設定 k 值為非中心 t 分布反函數除以樣本數開根號： $k = t_{(g,f,\delta)}/\sqrt{n}$
> k	#列出 k 計算後結果
[1] 4.378378	#k 值



表 5、容忍區間計算結果

Lot	Percent		
01	100.23	最大值 (=Max [B2:B11]) =	104.52
02	104.52	最小值 (=Min [B2:B11]) =	98.57
03	98.57	平均值 (=AVERAGE [B2:B11]) =	100.763
04	99.61	標準差 (=STDEV [B2:B11]) =	1.696618
05	101.24	$x =$	3
06	100.95	品質範圍下限 =	95.67315
07	100.55	品質範圍上限 =	105.8529
08	99.33	$\alpha =$	0.05
09	102.41	$t_{(\alpha/2)} (=TINV [E8, 10-1]) =$	2.262157
10	100.22	信賴區間下限 =	99.54931
		信賴區間上限 =	101.9767
		k (以 R 計算) =	4.378378
		容忍區間 (TI 99%/95%) 下限 =	93.33457
		容忍區間 (TI 99%/95%) 上限 =	108.1914

容忍區間的表示方法，舉例如 TI 99%/95%，具有兩個數字，前者為涵蓋範圍，後者為信心水準，與信賴區間僅有信心水準不同。依容忍區間之定義，若表示為 TI 99%/95%，代表抽樣 100 次所得容忍區間，有 95 次可以涵蓋母群體 99%之範圍^[9]。與信賴區間推估母群體單點平均值相較，容忍區間推估的是母群體的範圍，因此容忍區間會比信賴區間寬廣。對變異性較大的生物藥品來說，容忍區間不易產生超出區間範圍外的狀況，可以減少不良品的比率。但一旦擴大了合格範圍，同時就會增加使用者的風險(產品實際上不合格，但仍被放行使用)。歐盟 EMA 對容忍區間之見解與品質範圍類似，認為大多數容忍區間的評估均未確認母群體是否為常態分布。另外，涵蓋範圍與信心水準的訂定方式也欠缺科學依據^[3]。舉例而言，個別調整表 5 所訂定之容忍區間涵蓋範圍 p 以及信心水準 g，將計算結果列於表 6，可看出隨著涵蓋範圍或信心水準增大，容忍區間亦隨之增大。依表 6 所示之現象，歐盟 EMA 認為，雖然製造廠可將涵蓋範圍



與信心水準訂定得很高，如 TI 99%/99%，但卻會使人誤會依 TI 99%/99%所證明的相似性具有很高的可信度，其實只是因為擴大了整體區間範圍，使得所有數值都落在推測範圍內而已，仍無法準確預測母群體所在範圍^[3]。美國 FDA 統計審查員以 Monte Carlo simulation 評估結果顯示，容忍區間涵蓋範圍會較母群體真實範圍寬廣，不適於抽樣數較少之評估，否則可能會提高使用者之風險^[6]。即使增加樣本數量可使容忍區間範圍趨近母群體真實範圍，在美國 FDA 已撤銷之指引草案中，tier 2 之評估仍不建議使用容忍區間評估生物相似性藥品與對照藥品之相似性^[4]。

表 6、容忍區間依涵蓋範圍或信心水準變化之趨勢

涵蓋範圍 (p) 固定為 99%			
信心水準 g	95%	97.5%	99%
k	4.378378	4.878817	5.571023
TI 99%/g	[93.33457, 108.1914]	[92.48551, 109.0405]	[91.3111, 110.2149]
信心水準 (g) 固定為 95%			
涵蓋範圍 p	95%	97.5%	99%
k	3.402453	3.846398	4.378378
TI p/95%	[94.99034, 106.5357]	[94.23713, 107.2889]	[93.33457, 108.1914]

預測區間

生物藥品亦有使用預測區間界定允收範圍之案例。Eq. (4)介紹單一參數下預測區間的計算式^[10] (雖亦有利用線性回歸模式計算預測區間之計算式，但本文暫不討論此算法



[9]。表 7 則利用 Eq. (4) 以及表 1 之虛擬數據計算 95% 信心水準之預測區間。

$$\text{預測區間} = \text{樣本平均值} \pm t_{(\alpha/2)} \times \text{樣本標準差} \times \sqrt{[1 + (1/\text{樣本數})]} \quad \text{Eq. (4)}$$

表 7、單一參數預測區間計算方法

Lot	Percent		
		最大值 (=Max [B2:B11]) =	104.52
01	100.23	最小值 (=Min [B2:B11]) =	98.57
02	104.52	平均值 (=AVERAGE [B2:B11]) =	100.763
03	98.57	標準差 (=STDEV [B2:B11]) =	1.696618
04	99.61	$x =$	3
05	101.24	品質範圍下限 =	95.67315
06	100.95	品質範圍上限 =	105.8529
07	100.55	$\alpha =$	0.05
08	99.33	$t_{(\alpha/2)} (=TINV [E8, 10-1]) =$	2.262157
09	102.41	信賴區間下限 =	99.54931
10	100.22	信賴區間上限 =	101.9767
		k (以 R 計算) =	4.378378
		容忍區間 (TI 99%/95%) 下限 =	93.33457
		容忍區間 (TI 99%/95%) 上限 =	108.1914
		預測區間下限 =	96.73765
		預測區間上限 =	104.7883

預測區間的概念為預測一母群體之中，未來抽樣或觀測到的一個數值可能出現的範圍^[3]。計算式 Eq (4) 與計算式 Eq. (2) 雖類似，但計算式 Eq. (2) 的信賴區間，所推測的數值為母群體之平均值；計算式 Eq. (4) 的預測區間，所預測的數值為未來的單一觀測值，兩者所代表意義不同。美國 FDA 並未對預測區間有任何建議，歐盟 EMA 有簡略介紹預測區間的意義，但並未說明預測區間應用於品質領域之優缺點^[3]。預測區間的概念，主要是認為製造廠可在一個穩定的製程下，以過去歷史數據預測未來觀測值。不過，所



有區間估計統計工具可能有的問題，即過去歷史數據的母群體是否為常態分布，一樣會出現在預測區間的估計之中，而且 Eq (4) 已簡化預測區間之計算，通常預測區間需龐大數據建立模式後(通常具有兩個以上的參數，例如：安定性試驗的時間與含量測定數值^[9])，再預測未來單一觀測值。若無龐大數據與模式支持，以 Eq (4) 所計算出的預測區間可信賴度並不高。相較於信賴區間，預測區間的計算中考量了所有實測值個別之誤差，因此一如表 7 所示，預測區間範圍較信賴區間寬廣。即使預測區間的寬廣對製造廠有利，但考量預測區間的概念，「預測未來」並不同等於「比對不同群體」。當學名藥或生物相似性藥品欲與對照藥品比較，甚至是進入藥品製程確效的新藥批次欲與第三期臨床試驗所使用的新藥批次比較，用於比對兩種可能不同的母群體時，並不適用於同一母群體下預測區間的概念。

選擇統計工具之考量

在品質領域選擇統計工具時，最應優先考量的參數為樣本數目以及代表性。與臨床試驗等動輒上百或上千樣本數之狀況不同。樣本數以批次數目為單位時，即使是中央極限定理中觀察到最低可趨近標準常態分布之 30 個樣本數，在品質領域中也是非常龐大的研發成本，對中小型製藥企業而言，每項產品均投入 30 批進行研究實有難度。在樣本數目少，甚至可能僅有 1 批的極端狀況下，如何選擇統計工具進行分析，主要仍須倚賴紮實的物化性質研究、配方研究、製程研究與過去經驗，以說明投入統計分析中樣本的代表性。歐盟 EMA 的概念文章草案，大部分的內容均在描述如何確認樣本(批次)的代表性，統計工具的探討僅占一小部分。畢竟，若抽樣樣本不具代表性，無論使用何種統計工具，都無法得到明確的統計推論^[3]。美國 FDA 雖在已撤銷之指引草案中建議生物相似性藥品之相似性比對須各採樣對照藥品及生物相似性藥品至少 10 批^[4]，但各國法規單位對批數之標準仍未有定論。在藥廠針對美國 FDA 指引草案所公開的意見之中，亦可發現生物相似性藥品製造廠針對對照生物藥品批次抽樣代表性的疑慮，由於生物藥品變異性較大，又可能隨著時間而改變品質屬性之範圍，生物相似性藥品製造廠無法逐一掌握對照生物藥品變更履歷，如何確認抽樣所得之對照生物藥品代表性，成為生物相似性藥品製造廠的一大難題^[11]。

比較前述各項統計工具得到的範圍，相同信心水準下，最狹窄的範圍為信賴區間，最寬廣的範圍為容忍區間(表 7)。雖然有不少生物藥品採用容忍區間訂定品質屬性允收



範圍，但仍不應忽視提高使用者風險的可能性，尤其在沒有足夠臨床試驗數據支持的生物相似性藥品或學名藥，歐盟 EMA 及美國 FDA 均不建議以對照藥品的容忍區間訂定學名藥或生物相似性藥品的允收標準^[3,4]。品質範圍雖仍有容忍區間的缺陷，但因計算簡便，國際規範如 ICH Q6A 已有建議之 x 值(3 倍樣本標準差)^[1]，並有多數案例被歐盟 EMA 或美國 FDA 所接受採認，成為近期製造廠最常使用之評估工具。

除了上述各項統計工具所得到的範圍管制藥品品質之外，運用兩個抽樣範圍確認兩種藥品是否具有相似性/相等性，屬於品質比對最嚴謹之等級，但亦為最難以評斷合適性之範疇。推論統計之假設檢定雖可得到兩組數據是否具有顯著差異之結論，但顯著差異的有無並不同於相似性/相等性，相似性/相等性的結論必須倚賴國際法規單位達成的共識，訂定明確的相似性/相等性標準。舉例而言，國際公認之溶離曲線相似性標準訂定為 $f_2 \geq 50$ ，該標準並非統計假設檢定所判斷的結果，而為溶離試驗與體內生體相等性試驗建立相關性(correlation)後，發現 $f_2 \geq 50$ 之範圍與 $C_{\max, \text{test}}/C_{\max, \text{reference}}$ (試驗藥品與對照藥品最大血中濃度之比例) 80%至 125%之範圍重疊，因此認定 $f_2 \geq 50$ 具有相似性^[12]。除了溶離曲線比對相似性之外，目前並無其他品質屬性具有國際上公認之相似性/相等性標準。美國 FDA 偶爾會在個別產品生體相等性指引(product-specific guidances for generic drug development) 中制訂品質屬性的相似性/相等性標準，如 budesonide suspension/inhalation，美國 FDA 即針對品質屬性，如每噴所含藥量，訂定了一整套 population bioequivalence 計算流程以及相等性標準($\theta_p = 2.0891$)^[13]；在已撤銷的指引草案中，亦訂定了 $1.5\sigma_R$ (1.5 倍對照藥品品質屬性數據所得標準差) 之相似性/相等性標準^[4]，但前述兩項美國 FDA 所訂定之相似性/相等性標準並未廣泛被國際接納， $1.5\sigma_R$ 之規定甚至已被撤銷^[5]。其餘國際醫藥法規組織均未明確揭露相似性/相等性標準，有鑑於近期生物藥品專利逐漸到期，生物相似性藥品之興起，應可促進國際法規單位針對生物相似性藥品之相似性/相等性標準進行討論，進而達成國際共識。



結語

表 8 彙整本文所提到所有統計工具常見應用範圍及法規單位之建議。

表 8、統計工具常見應用範圍與法規單位建議彙整表

統計工具	常見應用範圍	US FDA/EMA 意見
最大值與最小值範圍	<ul style="list-style-type: none"> 製程變更前後品質屬性比較 學名藥/生物相似性藥品與對照藥品品質屬性比較 	<ul style="list-style-type: none"> 無法清楚界定使用者風險與製造廠風險 不建議用於兩群體比較
品質範圍	<ul style="list-style-type: none"> 訂定品質屬性允收範圍 學名藥/生物相似性藥品與對照藥品品質屬性比較 	<ul style="list-style-type: none"> 母群體不見得是常態分布 涵蓋範圍大小並無一致評估標準 樣本數/批次數過少可能無法準確預測母群體範圍
信賴區間	<ul style="list-style-type: none"> 學名藥/生物相似性藥品與對照藥品品質屬性比較 臨床試驗相等性或非劣性評估 生體相等性試驗相等性評估 	<ul style="list-style-type: none"> 母群體不見得是常態分布 常用於相等性/相似性評估
容忍區間	<ul style="list-style-type: none"> 訂定品質屬性允收範圍 製程變更前後品質屬性比較 生物相似性藥品與對照藥品品質屬性比較 	<ul style="list-style-type: none"> 母群體不見得是常態分布 樣本數/批次數過少可能無法準確預測母群體範圍 信心水準與涵蓋範圍之訂定缺少科學依據
預測區間	<ul style="list-style-type: none"> 訂定品質屬性允收範圍 製程變更前後品質屬性比較 	<ul style="list-style-type: none"> 無具體建議



受限於統計工具需大量抽樣樣本之前提，在品質領域運用相關統計工具比對各項品質屬性時，須注意樣本(批次)代表性及品質屬性特性等化學製造管制相關議題。研發生物相似性藥品及學名藥之製造廠，更須注意在對照藥品公開資訊不足的狀況下，設計合適之檢測時機與項目，並嚴謹分析各項公開資訊，整合所有資料，再與法規單位充分溝通，共同討論品質比對計畫與統計工具之合適性。在完善的化學製造管制基礎下，方得以合適之統計工具得到合理之推論。

參考文獻

1. International Conference on Harmonisation of Technical Requirements for registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline: Specifications: Test procedures and acceptance criteria for new drug substances and new drug products: Chemical substances, Q6A. 6 October 1999.
2. International Conference on Harmonisation of Technical Requirements for registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline: Evaluation for stability data, Q1E. 6 February 2003.
3. European Medicines Agency. Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (draft). 3 March 2017.
4. U.S. Food & Drug Administration. Statistical approaches to evaluate analytical similarity (withdraw). September 2017.
5. U.S. Food & Drug Administration. FDA takes steps to foster greater efficiency in biosimilar development by reconsidering draft guidance on evaluating analytical studies. 21 June 2018. URL available at: https://www.fda.gov/NewsEvents/Newsroom/FDAInBrief/ucm611409.htm?utm_campaign=06212018_FIB_FDA%20reconsiders%20draft%20guidance%20on%20analytical%20studies%20for%20biosimilars&utm_medium=email&utm_source=Eloqua [accessed 12 July 2018]



6. Dong, X., Tsong, Y., Shen, Meiyu. Statistical considerations in setting product specifications. *Journal of Biopharmaceutical Statistics* 25: 280-294, 2015.
7. Şenvar, Ö., and Tozan, H. Process capability and six sigma methodology including fuzzy and lean approaches. 2 November 2010. Products and Services Igor Fuerstner, IntechOpen. URL available at: <https://www.intechopen.com/books/products-and-services--from-r-d-to-final-solutions/process-capability-and-six-sigma-methodology-including-fuzzy-and-lean-approaches> [accessed 12 July 2018]
8. European Medicines Agency. Report on the expert workshop on setting specifications for biotech products. 9 September 2011.
9. Bolton, S. and Bon, C. Pharmaceutical statistics: practical and clinical applications. Marcel Dekker, Inc. 2004.
10. Ramírez, J. G. Statistical intervals: confidence, prediction, enclosure. W. L. Gore and Associates Inc. URL available at: http://www.sascommunity.org/mwiki/images/3/3c/Statistical_Intervals-Confidence,_Prediction,_Enclosure.pdf [accessed 12 July 2018]
11. Biotechnology Innovation, Organization. Re: Docket No. FDA-2017-D-5525: Statistical approaches to evaluate analytical similarity. 21 November 2017. URL available at: <https://www.bio.org/letters-testimony-comments/biomilars-bio-comments-fda-draft-guidance-statistical-approaches-evaluate> [accessed 12 July 2018]
12. Sugano, K. Biopharmaceutics modeling and simulations: theory, practice, methods, and applications. John Wiley & Sons, Inc. 2012.
13. U.S. Food & Drug Administration. Draft guidance on budesonide (suspension/inhalation). September 2012.