

實證本位的語言治療介入 與成效評估實務

鄭靜宜

國立高雄師範大學聽力學與語言治療研究所

摘要

本文介紹臨床醫學中以證據本位 (Evidence-Based Practice, EBP) 的實證取向，討論如何運用於語言治療領域，並探討可能的困境與克服的方法。其中並介紹 EBP 中推薦研究證據的分級，提出一些臨床研究產生好證據的方法，探討如何進行治療成效的評估才能產生好證據的可行方法。文章的後半部即著眼於語言治療的成效評估，介紹損傷、失能和障礙三個層次的成效評估指標，提出運用溝通功能性的評估落實以實證為本的介入治療，最後並討論在臨床養成教育中需培養的 EBP 有關能力，期望在未來 EBP 能落實於語言治療評估與介入的實務方面。

關鍵字：本於證據的執業、EBP、證據等級、成效評估、功能性評估

* 通訊作者：鄭靜宜

通訊地址：80201 高雄市苓雅區和平一路116號

電子郵件：jjeng@nknuc.nknu.edu.tw

投稿日期：2010.11.15

接受日期：2010.12.28

壹、何謂 EBP (Evidence-Based Practice, EBP) ?

在傳統的醫學是以權威為本位的執業 (authority-based practice)，相信權威者的判斷，權威者是指領域中的權威者，或是在醫病關係中治療者的角色。實證本位的執業 (Evidence-Based Practice, EBP) 則是以科學實證為治療介入的基礎，乃源自於實證醫學 (Evidence-Based Medicine, EBM)。實證醫學乃起源於英國流行病學家 Archie Cochrane (1972) 在他所著的 *Effectiveness and efficiency: Random reflections on health services* 一書中特別地強調隨機化控制實驗 (Randomized Controlled Trial, RCT) 的重要性，並指出臨床執業者應使用經過嚴謹研究考驗後證實為有效的治療法去治療病患，才能為患者提供優質的臨床服務。自從 1980 年以來 EBP 在醫療領域已有一段不算短的歷史，目前在各健康保健 (healthcare) 領域也開始蓬勃發展。事實上，目前 EBP 正衝擊傳統的醫學典範 (traditional medical paradigm)，也廣泛影響著醫學教學訓練的模式，並擴及到其他和介入或教學有關的領域中，如臨床心理學、復健醫學、特殊教育等。

所謂的 EBP 是指臨床介入者在做臨床介入決定之前，需透過詳盡地辨認、評估和應用最適當的臨床相關證據去治療他們的個案，即治療者做介入時在意識上與行為上應該使用當前最新的研究證據，最受支持的療法去治療他們的患者。治療者的任何治療的決定是出自於目前具有最強證據支持的方法，並且也考量患者的最大利益 (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000)。也就是臨床介入的處遇都要能有憑有據，而且所憑藉的是當前最好的證據，是經過科學研究驗證過的事實證據。因此，臨床執業者在做介入決定前，必須先去蒐集當前的研究證據，並評估這些研究證據的可信度及相關性。由於 EBP 是以證據為基礎，而所謂的證據大多是指科學研究，因此又可稱為以研究為本的執業 (research based practice) 取向。

Research-based practice is health care practitioners considered use of research findings and collective research evidence to shape general approaches to care, specific courses of action, and recommendations made to individual patients (Brown, 1999).

EBP 是整合「治療者的個人專業」、「目前可得最佳研究的證據」以及「患者的最大利益」等三方面來決定處遇的方式。事實上，它是一套相信科學證據，致力於尋找及運用科學證據來做臨床介入的作法。EBP 主張介入決定需回歸於科學研究證據，盡量減少個人的主觀判斷。只有經由科學研究所得到的證據才值得信賴和採用，並憑藉這些證據作為介入決定的根本。傳統的醫療典範較依賴執業者個人經驗以及權威的意見，當然介入者在接受教育訓練過程所獲知的知識和技能及一般生理病理學知識也常是介入決定的基礎。介入者在臨床上常會遇到具有類似問題的個案，久而久之臨床執業者常會發展出對於某常見特定類型個案的例行性評估與處遇模式。這些例行性的做法來自長久以來經驗的累積以及介入者個人的專業知能，然而，如此一成不變的介入模式卻可能會阻礙治療專業的發展，無法由日新月異的醫療、科技研究發現中得到好處，因而也無法為個案謀得真正最佳的利益。EBP 並非要屏除一般的生理病理知識或是推翻領域中權威者的意見，而是認為介入的決定需要整合目前新近的

研究實證結果 (Friedland, 1998)，將之整合入治療者的專業知識體系之中，以便為個案做最適切的處遇決定。況且許多以往可能被認為正確、具權威的意見或作法，在時代的巨輪下有可能禁不起日新月異的科技和實證研究的考驗而逐漸被淘汰或推翻。

Friedland (1998) 指出 EBP 包含三個大面向，包括資訊的取得、對資訊的評估以及做決定的技巧。對於採用 EBP 的介入者而言，介入的過程就如同做科學研究一樣，是一個假設考驗和問題解決的過程。在介入之初，執業者需要蒐集及熟知當前相關的研究證據，經由論文期刊資料庫的搜尋與相關文獻的閱讀，評估各種可能介入方案研究證據的支持程度，並使用最適當、最好的證據來做介入決定的依據，定出一個最佳的介入方案。介入之後審慎客觀地評估介入的結果，盡量控制無關變項，使之成為支持該介入法的成效證據之一。這裡所謂「最好的證據」是來自最直接相關的高品質研究，而非一些未經證實的意見。介入的決定一定有科學證據支持，而不能僅只是憑藉個人的經驗或直覺。此外，介入者還需考量有關患者的一些個體內、外在因素，做明智的介入決定，以解決患者的臨床問題。

目前在網路上期刊論文的取得十分便捷，也使得科學研究結果的交流更為迅速，知識的累積與更新更為迅速。介入者需常常上網搜尋最近新的研究發現，以便更新自己的專業資訊。此外，另一個更新資訊的捷徑則是定期閱讀一些事先已將最新研究整理好的專業主題回顧性的文章，可知目前相關領域中最新的研究發現或介入處遇的新發展。醫療科學的發展是日新月異的，各個專業領域期刊每期都有新的研究發表。臨床介入者若能時常更新專業資訊，則較能應用目前發現最新、較佳的介入法，掌握各種介入法的優缺點，也較能瞭解自己的個案在介入過程的可能變化和預測可能的結果。

在以 EBP 為取向的介入中，介入者與個案的關係是一種互動平行式的關係，介入者可和個案討論介入的計畫，個案不僅有被告知的權利，更可以積極地參與治療照護的決定過程，介入者和個案討論幾種可能的介入選項，選擇出最佳的治療法，並以此為根據去設計最佳的介入方案。EBP 的實施是依據目前可得的最佳研究證據、治療者個人的專業訓練和病患的最大利益（包括個人期望、喜好），而傳統模式的介入實施則是較屬權威導向，大多憑介入者的專業知能決定介入方式，個案本身多是較處於被動以及較弱勢的一方。

貳、EBP與語言治療的成效評估

一、語言治療的成效證據

對於臨床實務者而言，最需要的研究證據無非是有關各種介入法或是針對某一特殊類型患者的介入成效之研究或評論。成效測量 (outcome measure) 研究可以為臨床介入者提供許多可參考的意見以及支持其介入理念。「成效」(outcome) 是指介入後個案發生的正向改變，乃介入後的結果。介入行動之後，科學化的介入成效評估也是 EBP 重要的一環 (Dollaghan, 2004)。EBP 的成效評量是以科學的方法系統地測量介入成效，並提供臨床介入者一個可信賴的回饋機制，使介入的成效成為其他的介入者可參考的證據。

近來國內語言治療介入服務也日趨普及化，語言治療師成為身心障礙者專業服務團隊的一員，主要目的在幫助有語言障礙者克服溝通的困難，並且提升語言溝通功能。語言治療的成效測量和服務品質的提升漸受重視，治療師常需提供介入成效評估的證據，因此成效測量的重要性與日俱增。Lubinski 和 Golper (2006) 指出早在 1960 年美國的語言治療者即被要求將介入的成效以日常功能性的方式呈現，但是直到目前功能性量表的發展仍十分緩慢，可廣泛性使用具有信度、效度的評估工具仍舊缺乏，如此長期缺乏成效證據的結果將導致在健康服務和教育領域中喪失競爭力，甚至導致職缺數量的減少。語言治療缺乏實證性的成效證據將使介入的結果無法被肯定，影響大眾的接納程度，並涉及醫療、教育資源分配的減少。因此，無論是語言治療領域內或領域外對於語言介入的成效皆有共同的期待，對於成效評估的重要性也有一致性的共識。每個介入者皆有義務回答這樣一個問題：「是否語言治療對個案有改善的效果？」畢竟治療品質的改進是健康服務團隊始終關心的課題。對這個問題的回答可由治療法功效 (treatment efficacy) 和治療效果 (treatment effectiveness) 兩個層面著手。

「治療法功效」是指整體而言某一種治療或介入法的效果，此種證據需要系統性的廣泛評估研究調查，妥善控制無關變項，進行抽離情境條件的推論，必要時還需和它種介入法做對照或比較去顯現其較大的功效。治療法功效的評估需要回答的是廣泛性的問題，需要採較大型的控制研究設計，如隨機化實驗設計。通常研究的進行需在控制的情境中以避免一些無關變項的干擾，以保障治療功效的因果推論的正確性。治療法功效的研究是多數介入者賴以決定的重要證據。

「治療效果」則是指一位典型的治療者對某位(群)個案介入的有效性，是某介入者在對某個案介入後的成果評估。由於環境的設定是在一真實的情境和時間之中，對於效果的推論也侷限於該治療者、個案、情境和時間。Frattali (1998) 指出情境常會限制治療成效研究的推論，由於真實的情境中有許多不可控因素會降低因果推論的信心，因此治療效果的評估只能使用相關或估計的陳述，無法作因果式的推論。因為該位治療者以該種方法治好了某一位患者，並不保證其他治療者用該種方法也會有效，或是用在其他患者也同樣地有效。在真實的情境下，除了治療者對個案的介入因素外，事實上還有許多因素會影響介入的成效，包括治療者本身和個案本身的兩方面的因素，個案本身的因素例如有個案的年齡、嚴重程度、動機、情緒、社經情況、家人支持度等因素。

在另一方面，臨床上常發現某些介入法的實際應用效果往往不如「治療法功效」研究中所呈現的效果，亦即某種治療法介入法在團體性的平均資料中呈現有效，並不能保證用在個別的個案身上皆是有效的。因此，在研究文獻中一個有「治療功效」的介入法其實並不能保證用在某一個案上一定會有預期的「治療成效」。研究中「治療法功效」的證據僅是提供介入者評估預期成效機率的資訊來源之一，介入者在閱讀有關「治療法功效」的研究報告時，也需要注意到該研究報告中實施該種治療的介入對象的性質以及其他可能影響的因素。雖說「治療法功效」的研究嘗試跳脫情境因素的影響以較大樣本體現介入治療的效果，然而不可避免地還是會受到一些因素限制，如介入對象的特質(如年齡、嚴重程度等)和一些實驗當時的情境因素(如所使用的儀器或施行者)。因此，臨床執業者在將治療法的實際應用於自

己的個案時，需考慮到一些個別化的因素，例如個案的性別、年齡、特質、病症的嚴重程度是否符合「治療法功效」文獻研究中施行對象的性質。另一方面，當然也還需考慮目前的一些實際情境因素，如介入自己對該介入技巧的純熟度、介入療程的時間長短或相關器材的同質性等，因為這些因素也都會影響到一個執業者對於一個個案的介入效果。

由以上討論可知有關治療法功效的問題需有較大樣本的實驗控制研究來解答，想要知道語言治療中的一些治療法是否有其功效同樣也是需要較大樣本的實驗控制研究來探討。然而，和其他醫療專業領域相較，在溝通障礙領域中大樣本的實驗研究實不多，其中較多是屬於有關群體特性的評估和分類探討的研究（如某種障礙有哪些次類型），在治療法功效方面的研究數量一向稀少。其實，直至目前無論國內、外有關語言治療功效的研究需求皆十分殷切，但是大樣本的實驗研究往往受限於經費、時間、人力，以及可參加合格受試者人數等因素而往往無法順利進行，因此若要累積相當可觀的治療效用研究證據，尚需長時間的努力和等待。另一方面，語言治療師對於回答自己語言治療的「治療效果」的問題則屬於個人的職責，介入者有義務對所介入的個案或個案的家屬呈現這個問題的答案，並提出相關的證據來說明它，這是語言治療師的專業倫理。

二、什麼是「好證據」？

EBP 的臨床介入對於科學研究「證據」的重視是無需贅言的，而所謂的「證據」是來自使用科學方式的實證探索與研究。治療者做臨床介入前，需要先蒐集目前已有的科學研究證據，結合這些證據和考量個案的特性設計適當的介入計畫，並於介入後提出介入成效的證據。

在這個資訊爆炸的時代，蒐集到的相關研究證據可能十分龐雜、良莠不齊，臨床執業者需要具有「慧眼」，能分辨證據品質的能力。評估研究證據的品質，首先需系統性地對該類研究歷史脈絡有完整的了解，蒐集相關研究，並進行批判性的閱讀。評估各個研究所提供證據的優劣，服膺於來自優良研究的好證據，亦即介入決定需取決於有好的證據支持的治療法。對於相關研究證據的蒐集，經由搜尋領域有關的期刊或論文資料庫，可以得到許多介入問題相關的研究證據。有些研究證據是來自大型、控制較周延的實驗設計；有些證據則來自較不完美的實驗設計。證據的品質主要來自證據產生的方式，如研究的設計與方法，一個具有好品質的證據是來自於好的研究設計與方法，在治療法理念和治療法成效之間有清楚的因果關係存在，因此證據說服性也較強。可見，一個研究的實驗設計和方法會影響產出證據的強度。

後設分析（meta-analysis）研究是系統性的回顧以往主題相關的研究，並運用一些統計的方法，檢驗各研究結果的一致性，整合這些研究的統計效果，將其合併計算，得到一個整體的統計效果，整合各研究的發現後得到一個整合性的結論。當然，做此種研究的先決條件是需要累積有相當數量的相似主題研究。後設分析研究統整許多主題相關的研究，尤其是統整臨床療效研究的好證據。在科學界，單一實驗常不能決定某一論點的成立或否決，需要參看一系列或許多相關主題研究的結果，若有一致性的發現，如此的結論證據則較讓人信服，

因此若是有 RCT 研究的後設分析研究，則是讓 EBPs 執業者在蒐集文獻證據的工作上事半功倍的「好證據」(Wertz, 2002)。

根據醫療界發展出的證據分級制，牛津實證醫學中心(Oxford Center for Evidence-based Medicine, 2009)，將研究證據分為五個等級的推薦水準(Level of Evidence)，並依據研究問題的性質，如治療助益、診斷正確性、預後、治療的傷害性、篩檢等問題而略有差異，現就治療成效分述如下。

(一) 等級一

是最具有說服力的證據，而最強的證據通常是來自於具一致性結果的 RCT 研究、RCT 研究的後設分析研究或系統性回顧(systematic review)或是具有大效果量的 RCT 的實驗研究。在 RCT 研究中，經隨機取樣的受試者被隨機分派到實驗組(或治療組)或是控制組(或安慰劑組)，且在介入者並不知道受試者的組別情況下進行實驗，之後比較實驗組或是安慰劑組之間結果的差異。一個設計良好的雙盲隨機控制實驗(double-blind randomized controlled trials)是所有研究證據中的「黃金標準」(gold standard)，因為在這種研究中具有最少的偏誤和最大的歸因性(attribution)。大量隨機化取樣可得到具有相當代表性的樣本(有足夠的數量)。由於隨機化取樣與隨機組別分派控制了許多可能的混淆變項，使得自變項明確地成為組間差異的歸因所在，因此實驗結果可得到明確的因果關係推論。隨機控制實驗在研究證據等級中是位於最頂端的。然而，有時要做到完全隨機並不可行，尤其是在實際臨床情境中常無法做到隨機選擇病人或做到雙盲控制，各組之間病人的嚴重度難控制一致，各組人數不均，或是召集的受試人數不夠多，這些皆會造成取樣或控制的偏誤而影響實驗結果的效度。

(二) 等級二

證據的說服性次強，證據來自實驗效果量較弱的 RCT 實驗，或同儕群組研究(cohort study)，或是這些研究的系統性回顧。

(三) 等級三

證據具中等說服性，證據來自於非隨機準實驗設計、個案控制研究(case-control studies)、個案研究(case study)、相關性研究、橫斷面研究(cross sectional studies)，或以上這些研究的系統性回顧等。

(四) 等級四

證據的說服性較弱，證據來自個案系列報告或設計不佳的個案控制研究，以及經發表的指導原則(guidelines)或同行共識等。

(五) 等級五

證據的說服性最弱，來自於未經考證的領域權威的意見、以往曾閱讀的教科書中或同行的意見或討論。

此外，也有三級制的證據分類法，各是 RCT 研究、非 RCT 研究、專業權威意見和描述性個案報告等三種等級（U.S. Preventive Services Task Force, 1996）。這些證據等級分類法只是對研究推薦度做大致地分類，分類標準以研究的設計為主。但牛津實證醫學中心（2009）後續修訂的等級劃分表中，也指出研究證據等級的劃分不僅是依據實驗設計，還需考慮研究的品質、效果量的大小以及和其他研究結果的一致性。換句話說，除了研究設計的因素之外，對於一個研究證據好壞還需考量研究中眾多因素，如受試者的性質、效果量（effect size）、研究的信度、效度、可應用程度（直接相關性、花費時間）、技術或方法的成熟度。若在許多的研究中某一治療法皆有一致性的有效結果，則該介入方式或論點有充分的證據支持，但若只有少數有關的研究呈現效果，則證據支持度不夠充分。若是有許多的研究，但是各研究的結論卻不一致，則需要小心分辨是否因研究設計的差異而導致不同的結果。要判斷證據的優劣和各研究論點的真實性，需參考支持證據數量的多寡、證據結果間的一致性、相關程度、以及適用性來作為判斷的根據。

研究領域的特性會造成研究的性質差異，在溝通障礙領域中的研究常由於經費獲得的限制或是研究問題的迫切性較低，始終缺乏大型的隨機臨床實驗研究，多數為單一受試者（single subject）研究，但此類研究的證據推薦等級不高。由於專業領域性質的差異或可考慮將證據標準略加修訂以符合專業領域的需求，以提升領域之研究層次。事實上，單個的單一受試嘗試（n-of-1 trials）的研究證據推薦等級不高，但是若有多個性質相近重複的單一受試嘗試研究在CEBM的證據等級（Level of Evidence #2）第二版中的證據推薦等級，甚至可達第一級。另一方面，專業領域中應可擬定一個提升研究等級的進程，分階段進行，希望最終可促進RCT或後設性研究的發展。

「效果量」是評估一個研究效果強弱的重要指標，而好的研究證據通常具有高的效果量。效果量通常也是後設分析研究中分析的重要項目之一。什麼是「效果量」呢？簡單來說，就是實驗結果和虛無假設之間的標準化的距離，或者是實驗組和控制組之間在依變項呈現差距量的多寡（經過正規化後），即所謂的 d 值，或者也可以說是依變項（如治療成效）的所有變異數值可被自變項（如介入法）解釋的比值。 d 值愈大，則代表兩者差距的量愈大，即實驗組和控制組之間有很大的不同，代表實驗或介入的效果愈強，效果量愈大。在實驗統計上，有時實驗效果儘管已達到顯著，如 t 或 F 值達顯著水準，但有時效果量卻不一定很大，此時即使實驗組和控制組呈現統計上顯著差異，但兩者差異實際並不大，此時小的效果量說服性往往不強，如此就不是個「好證據」。傳統的顯著性考驗統計只告訴我們是否應拒絕虛無假設，只是「是」或「否」的二分法，並無法得知到底效果有多大，可拒絕虛無假設的信心多強，若有效果量的參考則稍可解答這些問題。在成效研究中，介入變項對成效依變項的效果量當然是希望能愈大愈好。此外，各研究效果量的整合還需考慮研究樣本大小，再做加乘計分和求平均以得到一個統整的效果量 D 。此外，計算信賴區間（confidence interval）也是一種估計效果量的方法，信賴區間愈大，代表不確定性愈大，效果量也就愈小。了解效果量和信賴區間的原理可以讓我們對研究結果的可信度和證據的良莠加以評估。

一種介入法通常等到相關研究累積到一定的程度，就會有相關研究的後設分析研究出現，因此為便於未來從事後設分析的研究者於研究中提取所需的數據資料，原始的研究報告中就必須載明使用的統計方法和相關統計數據，若能有變項效果的效果量 (effect size) 分析更好。自從 APA 出版手冊第五版開始也載明研究報告中需包含效果量或是信賴區間的資料 (American Psychological Association, 2001)。目前許多專業期刊發表也漸漸要求提出效果量或是信賴區間的資料。

從事後設分析的研究者在閱讀所有相關的個別研究報告後，汲取報告中的提供的統計數據，並將個別的研究的效果量標準化，分析比較，將個別的效果量集結起來估計整體的效果量。Robey 和 Dalebout (1998) 指出後設分析需要有高內在效度和外在效度的優良研究作為材料來分析，若是分析的研究控制變項不佳，品質不良或是研究間方法的差異性過大，都會威脅後設分析的效度。後設分析應該要涵蓋所討論問題有關的所有「好證據」，而可作為後設分析材料的研究最好為 RCT。然而，若是相關研究過少，不可避免地就需要涵蓋一些準實驗設計 (quasi-experiments)，或是一些控制較差的研究，但這些無可避免的會威脅該後設分析的效度。由於目前單一樣本設計或單一受試的研究無法算得效果量，尚無法納入後設分析的量化資料中。

三、實證本位的語言治療成效證據

發源於臨床醫學上的 EBP，是否適用於溝通障礙的專業領域呢？EBP 其實並非是新觀念，而是強調研究和證據對介入的重要。溝通障礙領域是屬於現代科學的學門之一，科學原本就重視研究和證據。EBP 即是要求使用有研究證據支持的治療法去做介入（合併考量個案的損益後），使用科學研究的精神和方法進行介入。對於介入程序做有系統性的嚴格控制，而非隨性的改變。

EBP 的作法強調成功的介入有賴於採用具有好證據支持的介入法，因此介入前需要好好地尋找好證據，需要蒐尋和目標個案具有類似特質以及一些目前可用介入法之成效相關文獻資料，首先要先將想問的問題先想好以及把可能相關的關鍵詞列出來。文獻搜尋時，需要使用適當的關鍵詞來窄化搜尋的範圍，以求得相關度較高的核心文獻。

圖書館提供的電子期刊資料庫的全文檢索是個便利的工具。語言治療的領域主要涵蓋有構音音韻異常、嗓音異常、運動言語異常、語暢異常、兒童語言異常、成人語言異常、擴大輔助溝通、吞嚥異常等，在表 1 整理了這些領域常用的期刊以及有收錄全文的電子期刊資料庫。是否有收錄全文則要依圖書館是否有近一步訂閱該電子資料庫而定，若沒有訂閱完整的資料庫則通常只有呈現論文的一般出版的訊息和摘要而已。有時 Google 的學術搜尋也可找到一些全文上線的文獻。

將蒐集到的文章依照主題相關的程度做分類，由相關性最強的核心研究文獻開始進行批判式閱讀。注意該研究所採的實驗設計為何？結論為何？所能提供的證據推薦等級為何？是否有高的內在效度和外在效度？所謂的內在效度是指研究變項的效果是否可以擺脫其他混淆

變項的糾纏而成為單純可以解釋依變項的比例，而常見的混淆變項有取樣的偏誤、測量的偏誤、程序的偏誤等。外在效度是指研究的變項效果可推論到研究之外的情境、個人或群體的程度，而具有完全隨機取樣的大樣本實驗通常外在效度較高，因為樣本較具有代表性。

表1

語言治療領域常用參考期刊和所收錄的電子全文資料庫

期刊名	收錄的電子資料庫
Journal of Speech, Language & Hearing Research	EBSCOhost-ASP, CMMC, Medline
Clinical Linguistics & Phonetics	EBSCOhost-CMMC
Folia Phoniatica et Logopaedica	EBSCOhost-Medline
Journal of Communication Disorders	SDOS
Language, Speech, & Hearing Services in Schools	EBSCOhost-ASP, CMMC
American Journal of Speech-Language Pathology	EBSCOhost-CMMC, ASP
Aphasiology	EBSCOhost-Medline
Language & Speech	EBSCOhost-ASP, CMMC
Seminars in speech and language	EBSCOhost-CMMC
Journal of Child Language	EBSCOhost-CDAS
Journal of Neurolinguistics	SDOS
Topics of Language Disorders	EBSCOhost-CMMC
Brain and Language	SDOS
Journal of Experimental Child Psychology	SDOS
Journal of Phonetics	SDOS
Journal of Acoustical Society of America	EBSCOhost-EJS
Journal of Voice	SDOS, SDOL (ScienceDirect)
International Journal of Pediatric Otorhinolaryngology	SDOS
Journal of Laryngology and Otology	EBSCOhost-Medline
Journal of Otolaryngology	EBSCOhost-Medline
Journal of Otolaryngology -- Head & Neck Surgery	EBSCOhost-Medline
Laryngoscope	Wiley InterScience
Journal of Fluency Disorders	SDOS
International Journal of Speech, Language & the Law	EBSCOhost-CMMC
Dysphagia	EBSCOhost-Medline
Dyslexia	EBSCOhost-Medline
The Cleft Palate-Craniofacial Journal	www.cpcjournal.org
Augmentative and Alternative Communication	EBSCOhost-CMMC
International Journal of Speech Technology	EBSCOhost-CMMC
台灣聽力語言學會雜誌、聽語會刊	CEPS中文電子期刊
中華心理學刊	CEPS中文電子期刊

四、語言治療成效的評估方式

EBP的介入者需為自己的介入成效提出實際的證據來。作為一個介入者最關心的課題也莫過於介入的方法是否有效。介入成效可分為長程性成效 (long term outcome) 和短程性的成效 (short term outcome)。長程性成效是過了一段時間後顯現出可歸因於介入或治療的效果，而短程的成效為介入後測量到的立竿見影的效果。以實證為本的臨床介入方式重視治療成效和治療效果的評估。以 EBP 為取向的介入歷程就如同做研究的過程，因此介入成效的評估也是一個實證研究的歷程。

進行評估成效首先要明確指出需要評估的行為是什麼以及評估的目的為何。成效評估首先要做的工作就是定義評估變項，為評估變項下一個操作型定義。定義測量的性質，操作型定義即是量化測量變項或指標，例如不流暢的出現頻率、語音清晰度的高低、子音的正確率或是平均語句長度等。完整的評估向度除了包含損傷層面，也應該包含功能性的評估，如溝通效能的評估。

和一般研究類似，成效評估研究的設計大致可分為大樣本和單一受試設計兩大類。大樣本實驗包括有群體前測——後測比較、長期縱貫式比較、同儕回溯研究等準實驗設計、隨機實驗組和控制組設計。在語言治療的療效評估研究中，大樣本的臨床研究常不易進行，原因之一具有病理同質性的群體並不容易尋找，不僅要有同質的實驗組，還要有同質的控制組，而同質的控制組設置在專業倫理上可能會有困難。原因之二是在介入方面較難控制一致。不管是一對一，或一對多的行為介入模式，皆難以控制使之等同。因為介入互動過程很難標準化，其中涉及的混淆變項極多，尤其在各組介入者若是不同時，在互動過程中許多在實務面均難控制一致。因為介入活動本是在情境中的動態歷程，介入者常需隨個案的反應而有彈性變通，因此要保持實驗組和控制組之間除了自變項外，其餘保持一致，實有相當大的困難。因此，實驗的結果常較難以單純地歸因於實驗自變項的效果，而這種研究也常只能是準實驗設計 (quasi-experiment)，而無法為真正的雙盲 RCT。

單一受試研究或個案控制研究也是提供 EBP 證據的來源之一，在以行為介入為主的領域，如特殊教育、復健醫學等，單一受試研究的使用是受到肯定的 (Horner et al., 2005)。在溝通障礙領域，單一受試研究也是常使用的一種實驗研究設計。因為臨床介入模式的個案大多為一對一的形式，且由於單一受試研究不需要尋找大量同質性的受試者，施行的可能性相對較高，因此臨床研究（特別是屬行為方面的介入）採用單一受試設計的數量較多。單一受試研究的優點是受試者數量少，施行較為容易，可以允許較深入且詳盡分析研究對象的行為改變 (Horner et al.)，可以詳盡地觀察並呈現個案改變的歷程，並且容易和臨床工作做結合。但另一方面，雖說單一受試嘗試的研究中呈現受試者的人數不盡然是單一，也可能是多位，大於一的情況 (許天威, 2003)，但要和 RCT 相較人數還是較不足，而且由於只有單一受試者，使之為實驗組，並無控制組，因此無法做組間比較。有些實驗設計中是以個案自身為控制組，在不同的時間點上做比較，以顯現治療效果，但由於介入效果只顯現在單一個案上，無法推論至群體。由於只有單一個案，在樣本數少、無關混淆變項多的情況下，單一受試研究的內在效度和外在效度會出現較多的問題，較易為人所質疑。例如常用的單一受試

研究設計中的前測——後測（AB）設計（A 為基線測量，B 為介入程序），由於無法釐清其他混淆變項的影響，介入成效的因果關係就無法得到明確的推論。其它有關介入者和個案的變項也常是混淆變項的來源，影響成效結果的推論性。

儘管如此，單一受試研究的進行若能採控制較優的設計，其實也能得到一些明顯的成效歸因證據，例如一般認為較優、也常使用的研究設計有 ABA 設計、倒返設計（reversal, ABAB）、多基線跨行為（multiple baseline across behaviors）、多基線跨情境（multiple baseline across situations）等。其中多基線設計一般較被推崇，它適合使用在不可逆的學習行為研究，例如在做構音介入時，以多個錯誤音為行為單位觀察介入後的改變，即適合使用這種設計。在介入開始同時觀察多種行為建立基線，再陸續處理每個目標行為，並持續紀錄尚未介入的行為。這種實驗設計有些控制機制存在，如各個處理期和基線期時長保持一致，階段期的嘗試次數保持一致。McReynolds 和 Kearns（1983）指出單一受試研究的結果非常需要加以反覆地施行於不同的個案，以印證自變項（如介入方法）的實際效果。若能以多基線設計做跨個案的研究，個案數量較多且具反覆印證性質，雖然樣本數無法達到如大樣本群體的研究一樣，但如能呈現客觀量化實驗的結果，單一受試研究應也具有有一些 EBP 參考的實證價值。優質的單一受試研究可作為 EBP 證據，Horner 等人（2005）提出以茲判斷的研究品質的標準或指標，如具有受試者和實驗情境的完整描述、具有對自變項和依變項的操作型定義、具穩定的基線、實驗控制的內、外在效度等。另外，在許天威（2003）所著的《個案實驗研究法》一書中也提出了一些有關單一受試者研究設計的信、效度和統計量化的方法，有興趣者可進一步加以參考。

參、語言治療的成效評估

一、介入成效的評估指標

可以作為成效指標的變項相當地多，可能包括了個體的生理、心理、社會各層面，如生理症狀改善、能力表現、功能、活動、適應行為、社會適應情況、獨立性、健康狀況、學業表現、滿意度、就業情況、對社會貢獻程度、生活品質等皆可成為測量的變項。對溝通障礙個案介入的成效評估要用哪些指標才較客觀、合適呢？溝通障礙乃是個體無法與他人進行有效的溝通，對於溝通障礙者的溝通效能指標可因層次觀點的不同而有不同的取向。世界衛生組織（World Health Organization [WHO], 1980）的 ICDH（International Classification of Impairments, Disability, and Handicaps）對於個體健康不良的狀況由三個層次的觀點來描述，即由「損傷」（impairment）、「失能」（disability）和「障礙」（handicap）。「損傷」是指個人在生理、心理或解剖結構的層次上的異常（disorders），如癱瘓、吞嚥異常、聽覺損傷等，是屬於個體的某部分內在結構層面出現問題或病徵。「失能」則屬於個體功能層次上的問題，由於損傷使得個體失去行使某些日常活動的能力，例如無法自行穿衣、行走、如廁或是與人溝通等問題。身心方面的損傷常會導致日常生活的失能，使個體活動方面受限制。「障礙」個體因為損傷或失能，造成在社會、經濟或環境的不利，無法滿足個人社會參與的角色，例如社會對殘障者的負面態度常造成殘障者低就業率或社會隔離等不利情況。「障礙」屬於社會環境層面，有關個人的生活品質或人生目標等大面向的不利情況，是

個人參與上的限制。障礙的評估涉及較多的個人和情境因素。在各種情境中，個體的障礙狀況常常不是固定，而是視環境中支持的多寡而改變，例如在一情境下由於輔具的使用使個體可以克服身心的限制而達成某些社會參與的目的，在某些情境下卻因為無法運用輔具而有無法參與的障礙。介入成效的評估即可由這三層面的觀點做評估，例如在「損傷」層面上評估個體在身心方面的能力缺陷。在「失能」層面上評估功能限制是否有減少，或適應行為是否有增加。在「障礙」層面上評估個體生活方面的品質與社會參與的限制。

ICF (International Classification of Functioning, Disability and Health) 是 WHO (2001) 依據以上三個層次發展出的一套的國際通用的健康損傷障礙編碼系統，採用階層性編碼架構系統。ICF 中最基本四大面向的編碼，包括身體功能 (body function, 開頭碼 b)、身體結構 (body structure, 開頭碼 s)、活動參與 (activities and participation, 開頭碼 d)、環境因素 (environmental factors, 開頭碼 e)。身體功能包括心智功能感官功能、聲音和言語功能以及其他生理系統或器官的功能。活動參與包括日常生活活動、自我照顧、人際互動交流等活動。身體結構是指身體的生理層面狀況，包括細胞、組織、器官和系統等部分。前三種屬個體內狀況，包括生理、解剖上的缺陷、臨床症狀等。第四種為個體外環境因素，包括物理環境、輔具的獲得、周圍人的支持、態度、社會福利制度，這些對個體而言有可能是阻礙，或是助益。由於 ICF 屬於國際性的標準，目前已有澳洲、加拿大等一些國家使用此編碼系統，未來的發展值得重視。

澳洲的聽語領域在臨床和研究上已開始使用 ICF 系統架構來定義溝通障礙患者的溝通問題。ICF 系統使用一組編碼來定義個人各方面的失能狀況和環境因素，在治療評估上也可使用此編碼來顯示成效，例如在加拿大的 CANCHILD 失能研究中心發展一個成效測量的標記法 (outcome measures rating form)，即使用 ICF 系統做成效資料登錄，其中溝通困難的問題也可用 ICF 系統來定義。例如，ICF 中 b320.2 是指構音功能有中度的損傷，b167.4 是指語言功能完全異常，完全無語言功能，d330.1 是指個體口語訊息的製造有輕微的困難，d350.3 代表個案日常會話嚴重困難。小數點前的數字碼代表哪一個方面的問題，可查表得知。小數點後的數字稱為 *qualifier*，代表嚴重度，0 代表無異常，1 代表輕微，2 代表中度，4 重度。然而，由於 ICF 是一個整合性的編碼系統，包含所有的損傷失能障礙的面向，而其中有關溝通障礙的部分實不多，而且使用五點量表的方式評估略嫌粗糙。因為量表對於一些較小的改進可能不夠敏感而顯現不出來。各嚴重等級的標注如何對應於各種相關的評估？要如何達到客觀？未來可探討 ICF 應如何加以修改以便適用於各種類型溝通障礙狀況。

若以「損傷」、「失能」和「障礙」三層次來看溝通障礙的評估，這三層次可依次為內在語言／言語能力 (capability) 的評估、溝通功能 (function) 的評估和生活品質 (quality of life) 的評估。言語溝通障礙領域中傳統的評估大多屬於損傷性層面的評估，主要在分析個體內在語言或言語能力的異常，診斷語言／言語問題的根源所在。如果說語言治療的主要目的是在提升個案的語言溝通效能，在日常生活中能有清晰可被了解的語言或溝通，滿足個案的溝通需求。功能性的評估正是所需要做的評估，評估個案在日常生活中的溝通功能，藉此可以了解個案的日常生活的溝通功能和情況，有助於介入目標的訂定、計畫的決定以及成效

的顯現。在「障礙」的層次，溝通障礙的評量工具較少，因為評估需由障礙者個人的觀點出發，涉及個體本身的生活方式、溝通環境、周圍親友的支持、文化背景、年齡、個性、價值觀等因素，因此評估的問題或向度較為個別化，較無法統一。對於此一層次的評估，介入者通常在訪談中由家屬或個案得到一些狀況改善的大致情形做推論。我們通常會預期對於一個個案這三向度的評估結果呈現正相關，但實際上卻不一定。在臨床上常可見兩個具有相似程度損傷的病人往往具有不同的日常生活能力。內在能力和外顯表現之間沒有一對一的對應關係，損傷的結果也不一定導致日常表現功能不佳，而障礙的程度高低需要考量一些其他外在因素，如可補償性或環境支持等。

內在能力或損傷的評估目的在評估個體內在語言的能力或潛能，評估工具可以細緻地分析語言／言語各方面的能力缺陷或不足。內在語言／言語能力的評估工具包括一些正式與非正式的測驗、量表或指標分數。通常針對個案的溝通障礙類別，如構音／音韻異常、音聲異常、語言異常、語暢異常、吞嚥異常或混合等情況，選擇適合的測驗工具評估其異常的情況，診斷異常的存在或評估其嚴重度。然而，此類的評估通常只是評估個案某部分的語言能力或功能，對於推論個案在實際溝通情境的表現實有所限制，要評估一個個案實際的溝通問題需有功能性的評估，即屬於「失能」層次的評估。

二、功能性評估工具

評量工具的運用是蒐集客觀證據的方法，評量工具的運用對介入成效的判斷有直接的影響。語言治療的主要目的在於促進個案的溝通功能，溝通功能的評量在於評估個案在日常生活上的溝通效能，評估個案在自然情境中能接收和傳遞訊息的能力。Frattali (1998) 指出溝通功能的評量包含三部分：個體意義傳達的有效性、速度流暢性和適當性。由於功能性評量的重點是在於日常生活場合中溝通的實際表現，通常採用實際觀察的方式做評估，而觀察者通常是與個案有較長時間相處的家人或教師。

目前語言障礙領域發展出的一些功能性評估工具，例如有 NOMS 所使用的功能性溝通評量表 (Functional Communication Measures, FCMs; American Speech-Language-Hearing Association [ASHA], 1997)、測量口語溝通效能的語音清晰度評估、測量吞嚥功能的吞嚥失能指標 (dysphagia disability index; Silbergleit, Jakobson, & Sumlin, 1997)、測量失語症患者的日常生活溝通能力評量 (Communicative Abilities in Daily Living, CADL; Holland, 1980)、功能性溝通剖面分析 (Functional Communication Profile, FCP; Sarno, 1969)、溝通效能指標 (Communicative Effectiveness Index, CETI; Lomas et al., 1989)、Amsterdam-Nijmegen 日常語言測驗 (Amsterdam-Nijmegen Everyday Language Test; Blomer, Kean, Koster, & Schokker, 1994)、以及測量語言損傷者的溝通剖面：功能技巧調查表 (Communication Profile: A Functional Skills Survey; Payne, 1994)。可見溝通功能性的評量工具多數還是依照個案的溝通障礙類型各有其特殊性，但和評估損傷的工具相較起來卻較不那麼特殊化，各種溝通障礙類別間可以通用，例如 CADL 原設計是給失語症的病人使用，但老人失智症、智能不足者和聽障者也可以使用。目前此層面的評估工具在失語症部分較多，評估兒童的語言或言語功能的較少。溝通環境檢核表 (Communication Environment Checklist; Rainforth, York, & Macdonald,

1992) 和教師——兒童溝通量表 (Teacher-Child Communication Scale; Bailey & Roberts, 1987), 可評估兒童在教室情境下的溝通功能。在言語的功能評估方面, Goldstein 和 Gierut (1998) 認為語音清晰度評估和連續語音樣本的聲母構音正確率 (Percentage of consonants correct, PCC) 是構音障礙嚴重度指標, 是屬於功能性層次的評估。

進行功能性評估需比較個案在各種情境下的溝通效能和發掘其溝通困難的所在, 介入成效則可顯現在各情境中溝通效能的提升上。整合性的語言評估至少應包含內在語言/言語能力的評估和外在日常溝通功能的評估, 而語言介入的任務或目標也應涵蓋此兩層面能力的提升, 不應偏重其中之一。若可能的話, 還應包括在「障礙」層次上對生活品質提升和個人角色滿足的評量。

這些功能性的評量工具大多使用心理量尺的點量表法做估計, 量尺由正常到最差分幾種程度, 而通常「正常」是指個案沒有患病前的情況或是符合在正常群體的範圍值之內。例如 FCP 為九點量表, 由正常到最差。CADL 為三點量表, 分錯誤、適當和正確三種等級。Communication Profile: A Functional Skills Survey則是使用五點量表評量各種溝通技能。功能性溝通測量 (Functional Communication Measures, FCMs; ASHA, 2000) 採用七點量表的方式, 以一級為最差, 七級為最佳。FCMs 可用來量化介入成效, 分別在介入之初與結案時由介入者依照個案的溝通功能做評量, 之後再藉著比較個案在前、後側的量表分數來估計治療介入所帶來的溝通能力的變化。

FCMs 是一系列適用各種溝障類型專用性的功能性溝通能力評量表。個案的功能性溝通功能等級要如何決定呢? FCMs 的指標是隨著個案在學校中的 IEP 中的項目而各有不同, 語言治療師觀察個案在高、低要求情況下的溝通表現。高要求情況 (high verbal demand) 是指適齡的溝通內容與形式需求情境, 低要求情況 (low verbal demand) 是指低於其年齡的溝通內容與形式需求情境, 並且評估個案在有外界幫助時的表現。以兒童口語表達為例 (ASHA, 2000), 一級程度是個案在口語啟動與反應在高要求情況時表現從未適齡, 在低要求情況下只有極少 (rarely) 的適齡表現。二級程度是指口語活動在低要求情況有時 (occasionally) 會有適齡表現, 而在高要求情況表現從未適齡, 但是當外界有最強支持時, 有時有少許適齡表現。三級程度是指在低要求情況通常有適齡表現, 在高要求情況少有適齡表現, 但是當有最強支持時, 有時會有適齡表現。四級程度是指在低要求情況持續有適齡表現, 而在高要求情況有時表現適齡, 若當有最強支持時通常有適齡表現, 但當情境需要時, 個案卻極少會使用補償性溝通策略。五級程度是指在低要求情況時持續有適齡表現, 在高要求情況有時表現適齡, 且當有最少支持時也會有適齡表現。當情境需要時, 個案有時會使用補償性溝通策略。六級程度是在低要求情況持續有適齡表現, 在高要求情況也通常表現適齡, 當情境需要時通常會使用補償性溝通策略。七級程度是指口語活動通常有持續適齡表現, 不需要外在額外的幫助, 當情境需要時持續會使用補償性的溝通策略以促進溝通。可見, 功能性溝通功能等級的定義決定要素有年齡、溝通情境需求 (高、低)、所獲支持的強弱和使用策略能力。個案溝通的表現評估介入者需要在各情境中仔細的觀察與比較, 比較包括個體間的比較與個體內的比較。個體內的比較則是個案在不同溝通需求時的表現以及若有幫助時的表現改變, 需

要跨情境的仔細觀察，也需要有適合個案年齡的常模來做個體間的比較。此外，對於溝通的高、低要求情況也需要做較為明確的定義和區別。

在 1997 年美國 ASHA 的溝通障礙療效國家中心（National Center for Treatment Effectiveness in Communication Disorders, NCTECD）成立國家成效測量系統（National Outcome Measurement System, NOMS），為一個介入成效資料匯集系統，具有全國性的資料庫（national database），為醫療與教育系統的語言治療師與聽力師介入成效的量化系統。介入的個案包括在醫療系統與教育系統的成人與兒童，因此資料庫分為三類，分別是成人、學齡兒童和學前兒童。NOMS 系統使用 FCMs 來量化個案的溝通功能，但也聲明並非要取代一般的測驗工具，而是一種輔助性評量，用來量化個案的溝通功能。溝通的功能是展現於日常生活之中，如學校、家庭等日常社交場合，FCMs 適用於跨情境的評估。也鑒於許多測驗工具因有許多的專業術語讓一般家長、個案或教師較難以了解，功能性的評估用語較易被一般人所接受。然而，FCMs 的使用還是出現了一些為人所詬病的缺點，主要是客觀性問題，其次是缺乏信度、效度的考驗，也無評分者間信度的考驗。Mullen（2004）即指出 NOMS 的三點限制，第一是在受試者方面沒有對照組的控制，第二是治療的程序不得而知，第三是 FCMs 的信度缺乏信度考驗的問題。事實上，語言治療師若只憑功能性評量來評估治療效果難免會流於主觀。臨床上常出現所謂的安慰劑效果（placebo），不僅會影響個案本身，也會影響介入者。主觀的期望效果不免影響語言治療師和家屬的判斷。語言治療師因涉及自己業務的成效，難免為保護自己而產生正向的偏誤，有不夠客觀的疑慮。EBP 取向的介入成效評估需要有客觀、量化的證據支持，若能結合功能性評量和能力評量的證據，使之相佐相輔應是較佳的作法。

肆、EBP與臨床養成教育

一、以EBP為取向的介入歷程

EBP 取向的介入歷程就如同做研究是問問題和回答問題的歷程，介入成效就是其中一個需要不斷提問的重要問題。在圖 1 列出介入歷程中實施 EBP 的幾個重要步驟。根據 EBP，介入者本身也是知識的生產者，研究理論和實務間關係密切，介入者同時也是研究者，在介入過程中蒐集研究問題有關的證據。如此可增進專業領域中知識的累積，當然對於增進介入成效也是可加以預期的。

在介入初期，介入者須對個案進行全盤性評估，並根據評估結果、相關研究證據，選擇介入的取向與介入目標，據此設計介入方案，此時要問的問題通常是：個案的語言能力可否用某一介入法來提升？某些言語上的困難是否可以用某些方法克服？這些為「治療法功效」的問題，要回答這種性質的問題需要有大樣本的臨床實驗證據，而臨床實驗證據來自於通常來自大規模的臨床測試（clinical trial），它們通常是研究者在有充裕的經費、人力、資源的支持下從事的大規模臨床測試研究，並將實驗研究的結果公開發表並討論，形成專業共識。如讓大家知道某治療的方法或技巧的有效性為何，是否可用於某群／個案身上，或對某些類型個案的效果如何。一般而言，醫學上的 RCT 大都為大藥廠或大型醫療儀器公司廠商出資從事研究的。

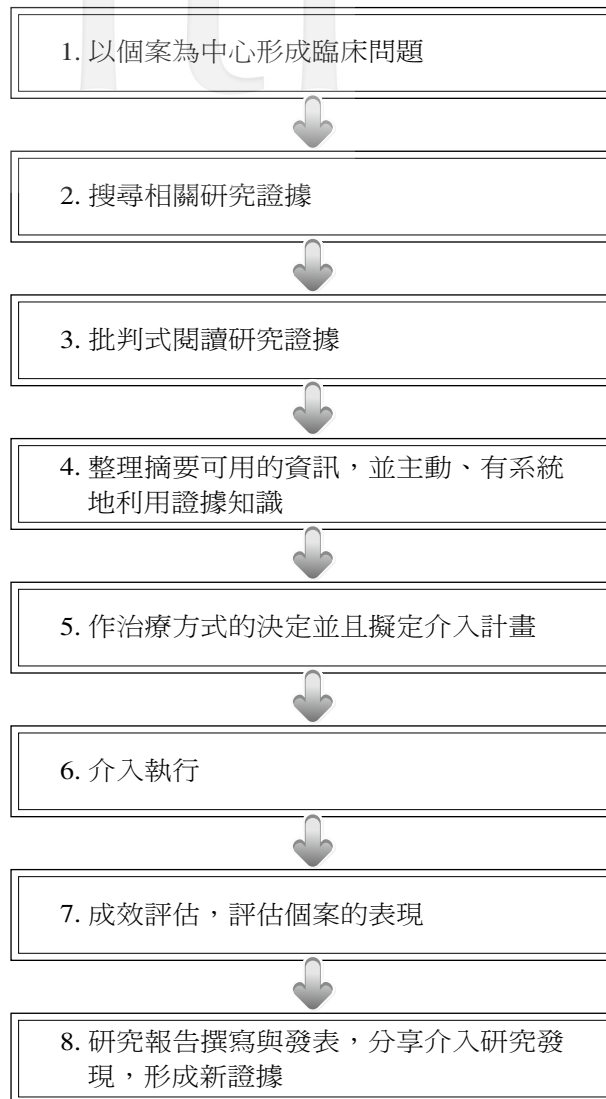


圖1 EBP 實施的步驟

一般的臨床介入者需要熟知領域內重要研究的結果，要知道這些研究的結果訊息就需要蒐集相關文獻的資料。藉由期刊論文網路資料庫的查詢，可以有效地在短時間內得到這些資料，經過批判性的閱讀和篩選後，找到最適切相關的證據，並在這些證據支持下設計合適的治療方案。在介入期間實施介入之後或固定期間，評估對個案介入實施的治療效果，療程結束後一段時期後有追蹤結果評估。可知在整個介入歷程中成效評估具有相當重要的角色，採取一般實證性成效評估的研究為證據，引導介入方案的設計，之後使用成效評估監控介入歷程，藉由動態性成效評估分析介入的短期效果，確認效果的存在與大小，並藉由成效評估的結果調整介入方式，計畫下一階段的介入。最後藉由成效評估顯示整體的介入的結果。如何能得知個案的改變是否來自於介入，或是其他因素導致的結果。這些以單一受試為對象的

研究有一些個體內控制法可以使用。其實，就單一個案的介入成果的評估本身也可成為一個研究。

EBP 的實施在醫療領域中常會一些有關於介入法的指導原則（guideline）的訂定與發佈，而這些指導原則的訂定是根據證據的等級強度和說服性製造專業指導原則將有助於臨床介入和介入效果的促進。Pinsky 和 Deyo（2000）指出 EBP 臨床介入者常面臨堆積如山的文獻研究和繁忙的臨床工作，而且閱讀研究文獻時，臨床介入者常會面臨一些看似互相矛盾的研究結果，而感到莫衷一是，此時運用專業領域的指導原則是最佳的選擇。但前提必須是，專業指導原則必須是根據經過多位專業者，經由審慎地整理和批判性地閱讀各個文獻後討論得到的結論，需有透明化的決策過程與實際研究的根據，也就是專業指導原則的訂定需要有良好的研究證據為基礎，經過多位專業者討論而得的共識。將研究的證據發展為專業領域的指導原則，於臨床執業者在做介入方式決定時有重要的幫助，並可促進專業領域內的共識形成與發展。當然，介入者在使用專業指導原則還是必須以批判性態度評估專業指導原則的品質，以決定是否使用於臨床介入中，而不是使之成為問題發生後推諉卸責的護身符。專業指導原則需要常常能定期回顧、檢討與修訂，整合最新近的研究發現與檢討，需要涵蓋最近一、二年的最新研究文獻（Frattali et al., 2003）。而目前國內在聽語介入專業指導原則的訂定方面尚有相當大的空間需要努力。

二、在臨床教育養成中培養的EBP能力

EBP 在聽語領域的實施目前可能會遇到許多的問題，其中最大的問題可能是大家對它的認識不足，其次是從業人員的訓練不足。例如臨床介入者可能會認為 EBP 費時費工，增加工作量。為克服這些問題美國聽語學會（ASHA）建議未來臨床聽語介入者在養成教育中需包含有 EBP 的介紹以及 EBP 的能力培養。要成為 EBP 專業執業者，在語言治療師的養成教育中需能培養以下的能力。

（一）蒐集研究證據的能力

認識並能使用圖書館或電子資料庫，能搜尋並運用網路資源。

（二）期刊論文的閱讀能力

在外語閱讀能力有限的情況下無法有效消化大量搜尋到的專業期刊論文是語言治療師常感到頭痛的問題，更遑論更進一步進行批判式的閱讀。若有專業文獻的整理與定期文獻回顧的出版以供閱讀，可稍加解決此問題。但治標不如治本，應想辦法對於語言介入實務者外語能力加以訓練，以提升專業期刊論文的閱讀能力。

（三）判斷證據良莠的能力

藉由期刊論文網路資料庫的查詢，可以在短時間內得到很多資料。然而，目前網路的資訊雖十分豐富，但也充斥著各種良莠不齊的資訊，亦即資訊的獲得往往不是問題，但對於資訊批判性的閱讀和有效地篩選則是需要經過不斷地學習與練習。因此，在教育訓練歷程中需培養此能力。

(四) 成效評估的能力

瞭解成效評估的實施方式，提出合適的介入成效支持證據。在損傷或是功能層次的評估能選擇合適個案的測驗工具或指標，能執行正確的施測、計分與分析解釋。介入者需認識目前可用的各種測驗評估工具的性質和其優、缺點。

(五) 做研究的能力

在以證據為本的介入中，一位介入者不僅是證據的使用者，也是證據的產生者，從事臨床研究產生科學研究證據以供大家參考。一個介入行動同時也是一個科學的研究。利用研究證據作為自我引導改善的依據。在介入者的養成教育中需普及並深化研究法的教導，以及提供練習的機會。在養成教育的過程中能使每個介入者皆有做研究的能力，了解 EBP，知道如何設計一個研究，操弄自變項，測量依變項以及控制無關變項，並且能寫出清楚、可讀的結果報告。

(六) 團隊合作的能力

積極推動語言治療專業和特殊教育的密切合作。在美國有超過一半比率的語言治療師是被學校聘請全日在學校工作的，而這些在學校服務的語言治療師服務的對象大多是身心障礙的學生，對於特殊需求學生的 IEP 極為重視，語言治療師需學習 IEP 的擬定，成為 IEP 中支持的一份子。語言治療師須以開放的態度，虛心面對他人對治療成效的提問，適時地提出成效評估的支持證據，合理地說明語言治療的成效。

伍、結論與建議

EBP 就是臨床介入需本於證據，而證據則是來自科學研究。科學研究產生新知識，知識即力量。同樣地，臨床介入的成效一樣需要證據的支持，缺乏證據支持的介入成效不免令人的質疑，而可信度堪慮。語言治療介入者需要學習臨床研究的方法，控制無關變項的方式，了解目前的最新治療取向，設計介入的方案，從事介入活動。成效報告中載明治療介入的方案、取向、治療的根據、治療歷程中的一些問題，治療介入的評估方式和結果，並使之刊登於相關的期刊中成為治療成效資料庫中的一筆資料。之後有研究者定期整理這些相關研究證據做系統性回顧，出版研究報告供大家參考，成為治療者的重要介入根據。

在過去，學術領域中的「研究」和「實務」之間常有一條跨不過去的鴻溝，實務人員較少去閱讀研究報告，搜尋追蹤證據，形成研究是研究，實務是實務，各不相干的情形。科學研究的目的是累積知識，但知識的力量在於運用，沒有被運用的知識是死的知識。EBP 以證據為基礎的作法可整合研究證據，使之應用於臨床介入，並在研究和實務之間搭一座橋樑，讓介入實務者跨越這條溝，實務者也同時也可為研究者。當實務者從事研究證據的搜尋時，會了解目前缺乏哪些重要的實務相關研究，提出研究需求，接下來或許也會本著實事求是的精神，在臨床介入過程中蒐集可用的證據資料，從事相關的小型研究，如個案研究，其實這同時也會無形間提升介入服務的品質，因為有可靠的成效評估證據的支持。

不可諱言地，臨床醫學和復健醫療的處遇介入性質是存在著一些差異的（Reilly, Oates, & Douglas, 2003），臨床醫學門診執業者的個案數量常較多，尤其是常可看到國內熱門醫生的門診，一個早上的掛號常有上百人之多，對一個病人的問診時間常不超過十分鐘。通常若非為慢性病，一個療程通常最多在二、三個禮拜內結束。而在復健醫療部門的患者，治療師處理一個個案所需的時間就需花上較長的時間，且療程通常較長，都在一個月以上，甚至一年以上的時間。語言介入治療的本質和多數臨床醫學治療的性質實有相當的差異性，而語言治療的性質則較接近復健醫學性質。語言治療對於個案一次介入時間通常需要 30 分鐘以上甚至 5、60 分鐘，療程也通常在一、兩個月以上。但是若整體介入的療程時間較長，由短期上來看相對的效果量則會較小。就介入性質而言，由於語言介入多屬於教導使其行為發生改變，多屬於非侵入式的介入，因此相對的風險較小，而醫學臨床治療通常需要用藥、手術或化學處理，多屬於侵入式的介入，相對地風險較大。Brown（1999）指出對於低風險的介入（low risk intervention）研究證據的評量標準較不需要如此嚴格。但也因為介入風險爭議較小，傷害性低（甚至無傷害性），就較不具急迫性，因而能得到的研究經費資源通常較為有限，因此領域中也就較缺乏較大規模的臨床隨機試驗，研究大多屬於準實驗設計、相關性調查研究或描述性的個案研究。這些實驗研究中樣本人數通常不多，因此長久下來研究證據的累積也較慢，整體研究證據的說服性會較弱。可見，一般的醫學臨床介入和語言介入的本質是有差異的，原本應用於臨床醫學的 EBP，若是將之推廣應用於聽語治療專業上可能需要略為修改以適應於語言治療專業的特性。例如證據的等級劃分，依據聽語溝通障礙領域內多數的研究在目前的情境可能就不適合分為五級。由於領域中有關介入的研究，大樣本的研究無法多得，多數為小樣本的研究。其實整合一些有相同主題的小樣本研究，將之集結起來，也可成為值得參考的證據。

筆者曾以非正式的方式訪問一些語言治療師和臨床督導，發現許多語言治療師們對於 EBP 的了解還是十分有限且模糊，多停留在研究文獻搜尋部分。事實上，目前不僅是在國內，就是國外的聽語治療師對於 EBP 也還是在認識不足的情況。然而，另一方面，目前整個醫療和復健領域，以及美國 ASHA 和各大學的聽語課程也都在朝著 EBP 的道路前進。相信國內未來聽語治療的執業模式也將朝這個方向去下走。但是目前在研究動力和資源普遍不足的情況下，走這條路可預期將是十分的艱辛與漫長。EBP 本身以證據為本，需要累積許多可成為供為證據的實證研究，目前在聽語溝通領域中治療成效的相關研究不多，況且連有關 EBP 本身的說服證據目前也付之闕如，因此有關 EBP 在聽語介入應用的研究和各種介入法的成效研究證據正是我們日後需要努力產生的，未來期待能出現更多與介入相關的研究，大家一起努力吧！

參考文獻

- 許天威（2003）。個案實驗研究法。台北市：五南。
- American Speech-Language-Hearing Association. (ASHA) (1997). *National treatment outcome data collection project*. Rockville, MD: Author.

- American Speech-Language-Hearing Association. (ASHA) (2000). National outcomes measurement systems (NOMS): K-6 speech-language pathology user's guild. In *Functional communication measures* (pp. 13-17). Rockville, MD: Author.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Bailey, D., & Roberts, J. E. (1987). *Teacher-child communication scale*. Chapel Hill, CA: University of North Carolina.
- Blomer, L., Kean, M.-L., Koster, C., & Schokker, J. (1994). Amsterdam-Nijmegen everyday language test: Construction, reliability, and validity. *Aphasiology*, 8, 381-407.
- Brown, S. J. (1999). *Knowledge for health care practice: A guide to using research evidence*. Philadelphia: W. B. Saunders.
- Cochrane, A. L. (1972). *Effectiveness and efficiency: Random reflections on health services*. London: Nuffield Provincial Hospitals Trust.
- Dollaghan, C. (2004). Evidence based practice in communication disorders: What do we know, and when do we know it. *Journal of Communication Disorders*, 37, 391-400.
- Friedland, D. J. (1998). *Evidence-based medicine: A framework for clinical practice*. Columbus, OH: McGraw-Hill.
- Frattali, C. M. (1998). *Measuring outcomes in speech-language pathology*. New York: Thieme.
- Frattali, C. M., Bayles, K., Beeson, P., Kennedy, M. R. T., Wambaugh, J., & Yorkston, K. (2003). Developing evidence-based practice guidelines: Applying science to the art of clinical care. *Journal of Medical Speech-Language Pathology*, 11(3), ix-xviii.
- Goldstein, H., & Gierut, J. (1998). Outcome measurements in child language and phonological disorders. In C. M. Frattali (Ed.), *Measuring outcomes in speech-language pathology* (pp. 406-436). New York: Thieme.
- Holland, A. L. (1980). *Communicative abilities in daily living -- CADL 2*. Austin, TX: Pro-Ed.
- Horner, R., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject design research to identify evidence-based practice in special education. *Council of Exceptional Children*, 71(2), 165-179.
- Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). Communicative effectiveness index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54, 113-124.
- Lubinski, R., & Golper, L. C. (2006). Professional issues: From roots to reality. In R. Lubinski & C. M. Frattali (Eds.), *Professional issues in Speech-Language pathology and audiology* (3rd ed., pp. 3-45). San Diego, CA: Singular Thomson Learning.
- McReynolds, L. V., & Kearns, K. P. (1983). *Single subject experimental designs in communicative disorders*. Baltimore: University Park Press.
- Mullen, R. (2004). Evidence for whom? ASHA's national outcomes measurement system. *Journal of Communication Disorders*, 37, 413-417.
- Oxford Centre for Evidence-based Medicine. (2009). Oxford centre for evidence-based medicine -- Levels

- of evidence. Retrieved October 10, 2010, from <http://www.cebm.net/index.aspx?o=1025>
- Reilly, S., Oates, J., & Douglas, J. (2003). *Evidence based practice in speech pathology*. London: Whurr.
- Robey, R., & Dalebout, S. (1998). A tutorial on conducting meta-analyses of clinical outcome research. *Journal of Speech Language and Hearing Research, 41*, 1227-1241.
- Payne, J. (1994). *Communication profile: A functional skills survey*. San Antonio, TX: Communication Skill Builders.
- Pinsky, L. E., & Deyo, R. A. (2000). Clinical guidelines: A strategy for translating evidence into practice. In J. P. Geyman, R. A. Deyo, & S. D. Ramsey (Eds.), *Evidence-based clinical practice, concept and approaches* (pp. 119-124). Boston: Butterworth-Heinemann.
- Rainforth, B., York, J., & Macdonald, C. (1992). *Collaborative teams for students with severe disabilities*. Baltimore: Paul Brookes.
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM* (2nd ed.). Edinburgh, UK: Churchill Livingstone.
- Sarno, M. T. (1969). *Functional communication profile: Manual of direction*. New York: Institute of Rehabilitation Medicine.
- Silbergleit, A. K., Jakobson, B. H., & Sumlin, T. (1997). *Dysphagia disability index: Development, design and validation*. Rockville, MD: American Speech Language Hearing Association.
- U.S. Preventive Services Task Force. (1996). *Guide to clinical preventive services* (2nd ed.). Baltimore: Williams & Wilkins.
- Wertz, R. T. (2002). Evidence-based practice guidelines: Not all evidence is created equal. *Journal of Medical Speech-Language Pathology, 10*(3), xi-xv.
- World Health Organization. (WHO) (1980). *International classification of impairment, disabilities, and handicaps*. Geneva, Switzerland: Author.
- World Health Organization. (WHO) (2001). International classification of functioning, disability and health. Retrieved September 1, 2010, from <http://www.who.int/classifications/icf/en/>

Evidence-Based Practice in Speech Therapy and Outcome Measurement

Jing-Yi Jeng

Graduate Institute of Audiology and Speech Therapy, National Kaohsiung Normal University

Abstract

Evidence-Based Practice is the present trend in modern medical health care. The aim of this article is to introduce the evidence-based practice and discuss the issues of EBP in the field of speech therapy. The level of evidence recommendation were also introduced and the methods to generate the good evidence for EBP. In addition to promoting the capability of searching evidence, this article provided a critical thinking about the quality of research evidence. In order to determine the state and quality of evidence on the effects of treatment, the outcome measurement is pretty important. In last part of this article is to introduce various indexes for outcome measurement from the perspective of impairment, disability and handicap. The importance of functional assessment was also emphasized. In order to promote EBP in speech therapy, six related ability are proposed to be incorporated in the training courses for the future speech pathologists for the expectation about the implication EBP in the field of speech therapy.

Keywords: Evidence-Based Practice; EBP; Level of evidence; Outcome measurement; Functional assessment