

「嬰幼兒綜合發展測驗」動作分測驗與「皮巴迪動作發展量表第二版」的診斷準確度

吳雪玉 廖華芳¹ 姚開屏² 李旺祚³ 王天苗⁴ 謝正宜⁵

目的：1)以臨床診斷為基準，探討「嬰幼兒綜合發展測驗」(Comprehensive Developmental Inventory for Infants and Toddlers, 簡稱 CDIIIT)動作分測驗(Motor subtest, 簡稱 CDIIITMS)與「皮巴迪動作發展量表第二版」(Peabody Developmental Motor Scales 2nd ed., 簡稱 PDMS-2)對動作發展障礙兒童之診斷準確率；2)求得兩量表不同分界點之敏感度、特異度、陽性概率比與陰性概率比，並由相關研究所提之盛行率提供各分界點之陽性預測率與陰性預測率，以為臨床使用此兩量表的參考。方法：本研究共收取動作障礙兒童 85 位，與一般兒童 137 名。由同一位測驗者對受試者同時施以兩種量表的測驗，再依各量表的評分標準計分。依臨床診斷為判別基準，以 Receiver operating characteristics (ROC)曲線的面積推估 CDIIITMS 與 PDMS-2 的診斷準確率。並以列聯表分析兩量表在各個不同分界分數的診斷估計值，並由估計盛行率計算陽性預測率與陰性預測率。結果：CDIIITMS ROC 曲線面積為 0.97，PDMS-2 則為 0.98，具顯著之高診斷準確率($P < 0.001$)。綜合判斷 CDIIITMS 之最佳分界點為動作發展商數 70，其敏感度為 87%，特異度 97%，陽性概率比 29.0，陰性概率比 0.13，以 6%盛行率推估得陽性預測率 65%，陰性預測率為 99%。而 PDMS-2 之最佳分界點為動作發展商數 85，其敏感度為 81%，特異度 99%，陽性概率比 81.0，陰性概率比 0.19，陽性預測率 84%，陰性預測率為 99%；兩量表與臨床診斷之 Kappa 係數分別為 0.86 與 0.82 ($P < 0.001$)，具顯著高相關。結論：CDIIITMS 與 PDMS-2 皆具高的診斷準確率，各項診斷估計值大致符合標準，為學前兒童可用之動作發展評估工具。

關鍵字：動作評估，敏感度，特異度，診斷準確率，ROC 曲線，發展障礙兒童
(台灣醫學 Formosan J Med 2005;9:312-22)

前 言

根據 2003 年修定之「兒童及少年福利法」、「身心障礙者保護法」與 2001 年修訂之「特殊教育法」，皆規定對發展遲緩兒童之早期療育，應由醫療主管機關召集，並結合醫療、教育、社政等主管機關，共同規劃及辦理早期療育工作。為能夠早期發現、早期通報，早期發展評量(early development evaluation)的規劃與執行，就變成推動早期療育的

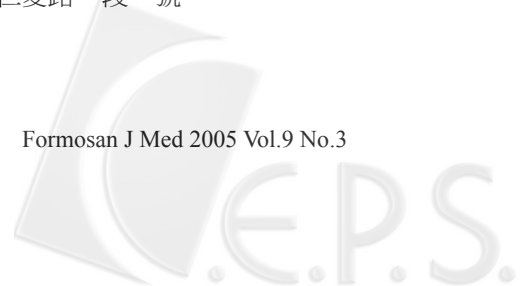
首要工作。完整的早期發展評量包括標準化發展診斷評量(standardized developmental diagnosis test)、臨床觀察、家長提供之訊息、相關專業人員之訊息與臨床判斷；其中標準化發展診斷評量尤其重要 [1]。由於動作發展遲緩常是發展遲緩兒童早期的表徵之一 [2]，因此有必要了解國內正在使用的動作發展診斷評量之診斷準確度。

「嬰幼兒綜合發展測驗」(Comprehensive Developmental Inventory for Infants and Toddlers, 簡稱 CDIIIT)是一套由國人自行編製，多向度、標

羅東博愛醫院復健技術科，¹台大醫學院物理治療學系暨研究所，²台大理學院心理學系暨研究所，³台大醫學院小兒科，⁴台灣師範大學特殊教育學系暨研究所，⁵台大醫院復健部

受文日期 民國 93 年 10 月 19 日 接受刊載 民國 94 年 1 月 18 日

通訊作者聯絡處：廖華芳，台大醫學院物理治療學系暨研究所，台北市仁愛路一段一號



準化、具本土常模的學前兒童個別施測發展診斷量表，具有信度、內容效度與建構效度[3]，於國內學前教育系統正普遍使用中。由於原 CDIIIT 作者僅提供總測量分數對嚴重發展遲緩兒童與輕度發展遲緩兒童之區別正確率分別為 82%與 55%[3]，因此有必要進一步對其診斷準確度作一探討。CDIIIT 的動作分測驗(motor subtest of CDIIIT, 簡稱 CDIIITMS)包含完整的動作評量內容，並經本研究小組其他研究證實，具有與「皮巴迪動作發展量表第二版」(Peabody Developmental Motor Scales 2nd ed., 簡稱 PDMS-2)及「貝萊氏嬰兒發展量表第二版」(Bayley Scales of Infant Development Inventory 2nd ed.)動作量表的同時效度[4-5]，因此可單獨成爲兒童動作發展的診斷評量[6]。

PDMS-2 與其第一版(Peabody Developmental Motor Scales, 簡稱 PDMS)爲目前最常被使用的兒童動作發展量表[7-8]，PDMS 被用來作爲發展診斷工具以區別嬰幼兒動作之正常與異常發展的效度已經獲得證實[7]。重新建立常模並根據專家意見修訂之 PDMS-2，爲 0-83 個月嬰幼兒動作發展之標準化、常模參照的個別施測量表，在兒童動作發展的評量上具有良好的內容效度、建構效度與效標預測效度[9]。然而 PDMS-2 爲新近完成之兒童動作發展量表，至今尚未有診斷準確度之報告。

運用發展診斷評量於發展遲緩的偵測時，要注意 4 個重要的診斷估計值：敏感度(sensitivity)、特異度(specificity)、陽性預測值(positive predictive value)與陰性預測值(negative predictive value)，這些估計值對利用測驗分數來作受試者的分類與區辨參考相當重要[10,11]。對於以連續分數來呈現測驗結果的診斷評量，常需探討哪個分界分數(cutoff score)可以區分發展正常與發展遲緩，而得最佳的診斷效率。通常不同之分界點可能導致敏感度上升，但相對會降低特異度，因此綜合考量常會加入陽性概率比(positive likelihood ratio)或陰性概率比(negative likelihood ratio)兩估計值。此外，ROC 曲線(Receiver Operating Characteristic Curve) 除可用以找出最佳分界點外，其下之面積常被用來綜合表示一個診斷評量之準確機率[12]。

爲增進 CDIIITMS 與 PDMS-2 兩量表在臨床運用上的效率，本研究目的爲：1)以臨床診斷爲基

準，探討兩量表之診斷準確率；2)得兩量表不同分界點之敏感度、特異度、陽性概率比與陰性概率比，並由相關研究所提之盛行率提供各分界點之陽性預測率與陰性預測率，以爲臨床使用此兩量表的參考。

研究對象與方法

受試者

本研究爲病例對照研究(case control study)。共計 85 位動作障礙兒童及 137 名一般兒童。動作障礙兒童之選取標準爲：1)經醫師診斷爲疑似或確定爲動作發展遲緩，並分別於台北市臺大、新光、國泰、羅東博愛與花蓮慈濟等教學醫院接受物理治療及/或職能治療者；2)年齡範圍 3 個月至 6 歲；3)家長願意接受評量，並填具同意書者。平均年齡 33.3 (標準差 19.6)個月，年齡範圍 4-70 個月，52 位男童與 33 位女童。動作障礙兒童所包含的臨床診斷類別包括：痙攣型雙邊腦性麻痺(spastic diplegic cerebral palsy)(n=27)、身心發展遲緩(psychomotor retardation)(n=23)、痙攣型四肢腦性麻痺(spastic quadriplegic cerebral palsy)(n=8)、染色體異常(chromosome abnormality)(n=7)、唐氏症(Down's syndrome)(n=7)、痙攣型半邊腦性麻痺(spastic hemiplegic cerebral palsy)(n=4)、脊柱裂(spinal bifida)(n=2)、動作發展遲緩(motor delay)(n=2)、協調不良型腦性麻痺(ataxic cerebral palsy) (n=1)、痙攣徐動型腦性麻痺(spastic athetoid cerebral palsy) (n=1)、徐動型腦性麻痺(athetoid cerebral palsy) (n=1)、裘馨氏肌失養症(Duchenne muscular dystrophy)(n=1)、脊肌肉萎縮症(spinal muscular atrophy)(n=1)與成骨不全(osteogenesis imperfecta)(n=1)。一般兒童選取標準爲 3 個月至 6 歲，足月生產，無神經系統、肌肉骨骼系統與呼吸循環系統之病史，且發展爲正常者。一般兒童就近在宜蘭地區與台北地區收取，共 137 名，平均年齡 29.4 (標準差 19.3)個月，年齡範圍 3-70 個月，69 位男童與 68 位女童。(表一)

研究工具

CDIIIT 於 1995 年由國內跨專業領域研究小組



表一：受試兒童基本資料與「嬰幼兒綜合發展測驗動作分測驗」(CDIITMS)及「皮巴迪動作發展量表第二版」(PDMS-2)之動作發展商數

	動作障礙兒童(n=85)	一般兒童(n=137)
月齡 (平均±標準差)	33.3 ±19.6	29.4 ± 19.3
月齡範圍 (月)	4-71	3-70
性別 (男/女)(人數)	52/33	69/68
動作發展商數 (平均±標準差)		
CDIITMS	58.3±9.4	94.5 ±14.5
PDMS-2	66.0 ±17.2	101.3 ±7.8

合力編製，包括診斷與篩檢兩種題本，適用於評估 3 至 71 個月嬰幼兒多個向度的發展情形。診斷測驗含認知、語言、動作、社會性能力與自理能力等 5 個分測驗與總測驗。CDIIT 以台灣北、中、南、東地區隨機取樣所得之 3703 名嬰幼兒的發展資料建立常模，並具有合宜的信效度[3]。

CDIITMS 內容包含粗動作領域的「肌肉控制」、「移位」、「身體協調」，及精細動作領域的「手操作」與「視動協調」等次領域。CDIITMS 的測驗題目採 0 與 1 之計分方法：受試嬰幼兒表現符合通過標準以 1 分計；不符合通過標準或沒有表現某種反應則計 0 分，若陪同的主要照顧者在平時看過該項反應則視同通過。施測由起始年齡組題項開始，施測者依一般兒童的實足年齡、身心障礙兒童則依其預估之發展年齡來決定施測的起始年齡組。由起始題項起，若通過則往困難的題目進行，否則則退回較簡易的題項，直到全部題目皆通過的年齡組，即達其基準層(basal level)；再往困難的題項進行，直到全部題目皆不通過的年齡組，即為上限層(ceiling level)。基準層以下的題目，全部以 1 分計，上限層以上的題目皆以 0 分計[13]。依兒童年齡與合作度，整個施測需時 15-20 分鐘。

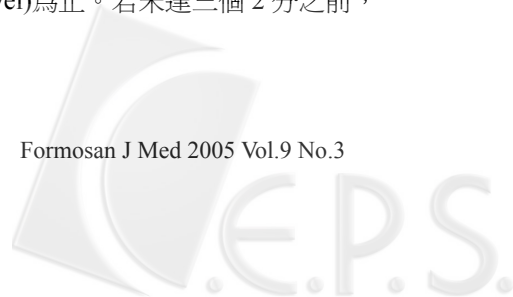
CDIIT 動作發展商數 (Total Motor Developmental Quotient, 簡稱 TMDQ) 為根據 CDIIT 常模表將施測所得之動作量表原始總分轉換而得。發展商數以 100 為平均值，15 為標準差。於兒童 24 個月前 CDIIT 提供個別月齡常模表，大於 24 個月後，則提供每三個月為一年齡組之常模，如 24-26 個月、27-29 個月等等。TMDQ 範圍為 54 至 146，兒童之得分若超出常模表範圍，視

其分數給與最高分(146)或最低分(54)。根據 CDIIT 指導手冊，兒童發展商數 70 以下為發展水準以下，需特別輔導。本研究群對 15 位足月兒與 15 位發展障礙兒童之信度研究顯示信度良好 [14]。

PDMS-2 測驗內容包含「反射」(reflex)、「靜態控制」(stationary control)、「移位」(locomotion)、「球類操控」(object manipulation)等粗動作之次領域，與「抓握」(grasping)、「視動協調」(visual-motor integration)之精細動作次領域。PDMS-2 信效度資料顯示良好[9]。過去本研究群對 15 位腦性麻痺兒童與 15 位健康足月兒使用 PDMS-2 得動作發展商數之研究顯示信度良好[15]。

PDMS-2 採三分法計分，各題目給分標準如下：2 分，表現完全符合通過標準；1 分，表現接近通過標準，但未完全符合；0 分，兒童不能或不願表現出欲測項目之技巧。每個題目各個分數的通過標準，皆明載於紀錄紙上。使用時，在該題目未得到滿分之前，可重複測試 3 次，以孩子所得到的最高分數來記錄[9]。

為節省施測時間，PDMS-2 有以下的施測規則：1). 起始點(entry points)：常模樣本中同一年齡層 75%孩子可通過的題目。對於一般兒童施測者可依孩子的實際年齡，障礙兒童則可依孩子的表現與專業的判斷來選擇適當的施測起始點。2). 基準層(basal level)：在項目中，不論其年齡層為何，只要往上測得連續 3 個 2 分，即達基準層。基準層之前的項目不必施測，全部以 2 分計算。施測時由起始點進入，往困難的項目進行，若得到連續三個 2 分，即已達基準層可繼續再往困難項目施測，直到上限層(ceiling level)為止。若未達三個 2 分之前，



有一個項目為 0 或 1，則需回到起始點之前的項目，直到找到基準層，再回到較困難的項目測試，直至上限層為止。3). 上限層：測試進行如前所述，當達連續 3 個 0 分即為上限層，上限層以上不必再施測，皆以 0 分計。此項規則，除了粗動作的反射次領域外皆適用，反射次領域由於要看的是孩子反射的整合情形，在一般兒童實足年齡、早產兒矯正年齡或障礙兒童粗動作發展年齡 12 個月以下者，皆須逐題測試[9]。依兒童年齡與合作度，整個施測需時 45-60 分鐘。

將測驗各領域的原始分數總合，查常模對照表算出各次領域的標準分數(standard score)，再將標準分數加總得動作總分的標準分數，依標準分數求得動作發展商數(TMDQ)。動作發展商數以 100 為平均值，15 為標準差。PDMS-2 於兒童 24 個月前提供個別月齡常模，於 24 個月至 53 個月間，則提供每三個月為一年齡組之常模，大於 54 個月後，則 6 個月一個年齡組，如 24-26 個月，54-59 個月等等。TMDQ 範圍為 38 至 165，兒童之得分若超出常模表範圍，視其分數給與最高分(165)或最低分(38)。PDMS-2 施測手冊建議總動作商數 70-79 為發展低下(poor)、總動作商數 80-89 為低於平均(below average) [9]。

施測程序

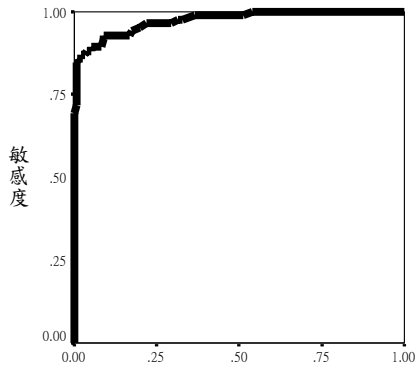
研究施測由同一位測驗者對受試者同時施以兩種量表的測驗，再依各量表的評分標準計分。為減少不必要的重覆以增進兒童的表現，本研究施測時以 PDMS-2 的測驗題目為主，將 CDIITMS 的項目加入其中。精細動作項目依所使用器材順序，粗動作項目則依不同的發展姿勢施測；對於兩量表共同的項目，只施測一次，再依各量表的計分標準分別評分。測驗的進行由精細動作項目開始，繼而測驗粗動作。粗動作以活動量較少的靜態控制項目為先，以避免受試者因體力耗盡而無法繼續受測的情形，並增進受試者在精細動作項目施測時的專注性 [9]。施測一個兒童所需時間，依年齡與合作度之不同，約需 45 至 60 分鐘。

統計分析

兩量表之診斷準確機率方面，以 ROC 曲線計

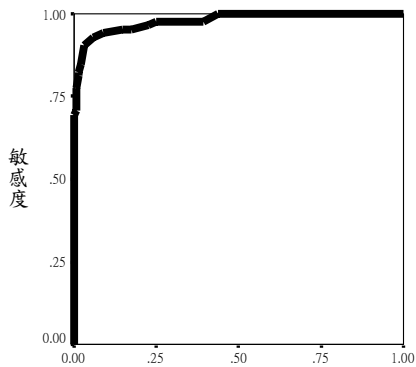
算，用所有受試兒童的臨床診斷為目標變數，畫出 CDIITMS 與 PDMS-2 總動作發展商數的 ROC 曲線，得 ROC 曲線的面積與 95%信賴區間(confidence interval, CI)，曲線面積越高，正確診斷機率越高 [12]。為得兩量表不同分界點之敏感度、特異度、陽性概率比與陰性概率比，依臨床診斷為效標，兒童分為發展障礙與發展正常組。在量表之分界分數方面，因 CDIIT 施測手冊建議 DQ70 [13]，PDMS-2 施測手冊建議 TMDQ80 或 DQ90 為分界點 [9]，兩量表並不一致；而根據發展診斷測驗常態曲線分佈，常建議的判別標準為分界點在 z 值 -1(DQ85)、-1.5(DQ77.5)與 -2(DQ70) [16]，因此本研究分別以總動作發展商數 70、77.5、80、85 與 90 為分界分數，來分析兩量表在不同分界點的診斷準確度各個指標：敏感度、特異度、陽性概率比、陰性概率比 [10-11,17-18]。所謂敏感度指在所有發展遲緩兒童中可由量表結果判斷為發展遲緩的比例；而特異度則為所有發展正常兒童中，可由量表結果判斷個案為發展正常的比例，兩者又稱為診斷準確度(diagnostic accuracy) [11,17]。陽性概率比即測驗為陽性之後疾病勝算比，評量後為陽性者，其為有病之概率為無病概率之倍數，以敏感度/(1-特異度)計算；陰性概率比則反之，意為評量後為陰性者，其為有病之概率為無病概率之倍數，以(1-敏感度)/特異度計算 [19]。並以 Kappa 值檢測不同分界點之分類與臨床診斷之相關係數值。一般定義分類相關強度如下，Kappa 值在 0 - 0.2 之間為「微」(slight)，0.21- 0.40 間為「尚可」(fair)，0.41-0.60 間為「中度」(moderate)，0.61-0.80 間為「顯著」(substantial)，0.81-1.0 間為「接近完美」相關(almost perfect) [10]。本研究 α 值定為 0.05。

由於目前對發展遲緩尚無全球公認之定義 [20]，也因此難得到一致的發展遲緩盛行率。但如相關研究所提及，Boyle 以美國國家衛生調查(National Health Interview Survey)中的資料統計發展障礙兒童的盛行率，17110 個案中成長與發展障礙兒童約占 4% [21]；Thompson 等人以美國佛羅里達州的活產統計(Birth Vital Statistics)與早期療育服務計畫(Children's Medical Services Early Intervention Program)比照研究出生體重對發展障礙的影響，在 170874 個案中發展障礙罹患率約為



1 - 特異度

圖一：以臨床診斷為目標變數，所畫出「嬰幼兒綜合發展測驗動作分測驗」總動作商數的 ROC 曲線。面積 0.97 (95%CI: 0.96-0.99) (P<0.001)



1 - 特異度

圖二：以臨床診斷為目標變數，所畫出「皮巴迪動作發展量表第二版」總動作商數的 ROC 曲線。面積 0.98 (95%CI: 0.96-0.99) (P<0.001)

5.3% [22]；而 Drillen 等人以篩檢測驗檢視英國 5334 名新生兒並追蹤至學齡前，發現至學前有 9% 被診斷為中度及中度以上的神經發展障礙[23]。本研究依上述研究假設盛行率 4%、6%、9%，再依下列貝氏定理的公式代入修正，分別計出各分界點之陽性預測率與陰性預測率[18,24]。

$$\text{陽性預測率} = \frac{\text{敏感度} \times \text{盛行率}}{\text{敏感度} \times \text{盛行率} + (1 - \text{特異度}) \times (1 - \text{盛行率})}$$

$$\text{陰性預測率} = \frac{(1 - \text{特異度}) \times (1 - \text{盛行率})}{(1 - \text{特異度}) \times (1 - \text{盛行率}) + \text{敏感度} \times \text{盛行率}}$$

結果

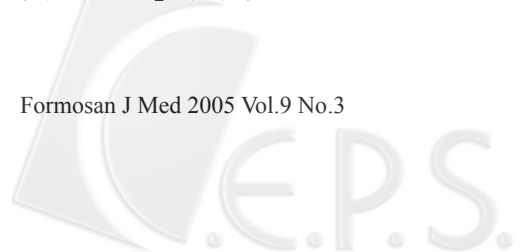
受試兒童基本資料與兩量表之動作發展商數如表一所示。動作障礙兒童組 CDIITMS 的 TMDQ 平均為 58.3(標準差 9.4, 範圍 54-100), PDMS-2 的 TMDQ 平均為 66.0(標準差 17.2, 範圍 37-100); 一般兒童組 CDIITMS 的 TMDQ 平均為 94.5(標準差 14.5, 範圍 63-132), PDMS-2 的 TMDQ 平均為 101.3(標準差 7.8, 範圍 79-124)。

圖一與圖二為以 ROC 曲線所得之兩動作量表之診斷準確率，ROC 曲線面積即為診斷準確機率。CDIITMS TMDQ 的診斷準確機率為 0.97 (95%CI 0.96-0.99, P< 0.001); PDMS-2 則為 0.98 (95%CI 0.96-0.99, P< 0.001), 因此兩個動作測驗之診斷準確機率皆接近 1, 且有高度顯著意義。

表二為 CDIITMS TMDQ 與 PDMS-2 TMDQ 在不同分界分數下(70, 77.5, 80, 85, 90)之敏感度、特異度、陽性概率比、陰性概率比與 Kappa 值的結果。於 CDIITMS 之診斷估計值方面，其敏感度之範圍為 87% - 99%，皆高於 85%，隨著分界分數之增加其敏感度也逐漸增加；特異度之範圍為 65%-97%，隨著分界分數之增加其特異度也逐漸減少，除分界分數 TMDQ 90 之特異度低於 70 以外，其餘分界分數之特異度皆高於 70；陽性概率比與陰性概率比皆隨分界分數之增加而逐漸減少，陽性概率比之範圍為 2.8- 29.0，除分界分數 TMDQ 70 之陽性概率比大於 10 以外，其餘分界分數皆小於 10；陰性概率比之範圍為 0.02- 0.13，除分界分數 TMDQ 70 之陰性概率比大於 0.1 以外，其餘皆小於 0.1。

PDMS-2 之敏感度範圍為 59% - 93%，較 CDIITMS 為低，隨著分界分數之增加其敏感度也逐漸增加，除分界分數為 TMDQ 70 或 77.5 外，其敏感度皆高於 70%；其特異度之範圍為 94%-100%，皆高於 90%，較 CDIITMS 為高。PDMS-2 之陽性概率比之範圍為 15.5-100，皆大於 10；陰性概率比之範圍為 0.07- 0.41，除 TMDQ 90 以外，皆高於 0.1。

以 Kappa 值檢測兩量表於不同分界點之分類與臨床診斷之相關係數值，得 CDIITMS 之 Kappa 值之範圍為 0.58- 0.86 (P< 0.001)，皆有統計顯著相關。於分界分數 TMDQ 70 及 77.5，其 Kappa 值分別為 0.86 與 0.81，屬「接近完美」相關程度；PDMS-2



表二：以臨床診斷為效標所計得「嬰幼兒綜合發展測驗」動作分測驗與「皮巴迪動作發展量表第二版」總動作發展商數(TMDQ)各分界點的診斷準確指標值

分界分數	敏感度(%)	特異度(%)	陽性概率比	陰性概率比	Kappa 值
嬰幼兒綜合發展測驗動作分測驗					
TMDQ 70	87	97	29.0	0.13	0.86***
TMDQ 77.5	94	89	8.5	0.07	0.81***
TMDQ 80	94	85	6.3	0.07	0.80***
TMDQ 85	96	77	4.2	0.05	0.70***
TMDQ 90	99	65	2.8	0.02	0.58***
皮巴迪動作發展量表第二版					
TMDQ 70	59	100	-	0.41	0.64***
TMDQ 77.5	68	100	-	0.32	0.73***
TMDQ 80	71	99	71.0	0.29	0.74***
TMDQ 85	81	99	81.0	0.19	0.82***
TMDQ 90	93	94	15.5	0.07	0.87***

***P < 0.001, by Kappa coefficient test

表三：「嬰幼兒綜合發展測驗」動作分測驗與「皮巴迪動作發展量表第二版」各分界點不同盛行率所得之陽性預測率(positive predictive rate, PPR)與陰性預測率(negative predictive rate, NPR)

盛行率	4%		6%		9%	
	PPR (%)	NPR (%)	PPR (%)	NPR (%)	PPR (%)	NPR (%)
嬰幼兒綜合發展測驗動作分測驗						
TMDQ 70	55	99	65	99	74	99
TMDQ 77.5	26	100	35	100	46	99
TMDQ 80	21	100	29	100	38	99
TMDQ 85	15	100	21	100	29	99
TMDQ 90	11	100	15	100	22	100
皮巴迪動作發展量表第二版						
TMDQ 70	100	98	100	97	100	96
TMDQ 77.5	100	99	100	98	100	97
TMDQ 80	75	99	82	98	88	97
TMDQ 85	77	99	84	99	89	98
TMDQ 90	39	100	50	100	61	99

Kappa 值之範圍為 0.64- 0.87 (P < 0.001), 也皆有統計顯著相關, 於分界分數 TMDQ 85 及 90, 其 Kappa 值分別為 0.82 與 0.87, 也屬「接近完美」相關程度。

表三為兩量表各分界點在不同盛行率下(4%, 6%, 9%)所得之陽性預測值與陰性預測值。CDITMS 之陽性預測值之範圍為 11% - 74%, 陰性預測值之範圍為 99%-100%; PDMS-2 之陽性預測值之範圍為 39% - 100%, 陰性預測值之範圍為 96%-100%。當盛行率越高, 其陽性預測值會增加, 相對地, 陰性預測值會略為下降。在盛行率為 6% 之情況下, CDITMS 僅分界分數 TMDQ 70, 其陽

性預測值接近 70%, 其餘分界分數之陽性預測值皆小於 70%; 而 PDMS-2 除分界分數 TMDQ 90 之陽性預測值小於 80% 以外, 其餘分界分數之陽性預測值皆大於 80%。CDITMS 之陰性預測值率皆大於或等於 99%, 而 PDMS-2 之陰性預測值皆大於或等於 97%。

討論

兩量表之診斷準確機率

ROC 曲線是用來描繪隨著測驗的陽性定義改變時, 敏感度與特異度估計值改變的情形, ROC



曲線下的面積即為測驗的診斷準確機率[17,18]。完美的測驗，其 ROC 曲線下的面積等於 1；面積若等於 0.5 則測驗幾無診斷功能[18]。本研究以所有受試兒童的臨床診斷為目標變數，所畫出 ROC 曲線的面積顯示，兩量表總動作發展商數的診斷準確率非常良好(皆接近 1)且相近(CDIITMS：0.97，PDMS-2：0.98)。(見圖一、二)由於病例對照組研究之病患通常較屬中重度，會高估診斷正確率[25]，檢視本研究動作障礙兒童由醫院復健科轉介進行物理治療和/或職能治療者，可能會較嚴重，然而進一步檢查其 CDIITMS 的 TMDQ 範圍為 54-100，其中有 3 位(4%)TMDQ \geq 85，9 位(11%)TMDQ 在 70-84 之間；而 PDMS-2 的 TMDQ 範圍 37-100，其中有 16 位(19%)TMDQ \geq 85，19 位(22%)TMDQ 70-84 之間，因此並不完全是中重度障礙兒童，因而推論本研究診斷正確率的高估程度應不是很高。然未來需再進一步進行前瞻性追蹤研究，以對此二工具之診斷正確率有更深入之了解。理論上接受何種治療與本研究主題發展評量工具之診斷效度無關，因此本研究未再進一步分析治療種類對此 2 種工具效度之影響。

兩量表各分界分數的診斷估計值

不同分界分數的選擇就會有不同的診斷估計值，為了解診斷評量的準確度，研究中常以效標評量或臨床診斷為黃金標準，依所定的分界分數來探討診斷評量的敏感度與特異度。如 Shoemaker 等人(1993)以 280 個 11-13 個月大(矯正年齡 8-11 個月)正在接受高危險群追蹤計畫的嬰幼兒，以「貝萊氏嬰兒發展量表」智力發展指數(mental development index)70 為標準，檢測「明尼蘇達兒童發展量表」(Minnesota Child Developmental Inventory, 簡稱 MCDI)的診斷準確度，結果顯示 MCDI 可有 56% 的敏感度與 92% 的特異度[26]。診斷分界點的選擇常是依據不同分界分數下的各診斷估計值來判斷。因敏感度與特異度不受研究樣本障礙盛行率(prevalence rate)的影響[10-11,17]，研究中在選擇較佳分界點時，常以此二估計值為依據，來選擇具有較佳診斷效率指標值的分界點。臨床上以敏感度 80%、特異度 90% 及 Kappa 值 0.75 以上為選擇標準[27-28]。根據此標準，CDIITMS 以 TMDQ 70

或 77.5 為分界點時，所得的敏感度分別為 87% 及 94%、特異度 97% 及 89%、及 Kappa 值 0.86 及 0.81；PDMS-2 則以 TMDQ 85 或 90 為分界點，所得的敏感度分別為 81% 及 93%、特異度 99% 及 94%、及 Kappa 值 0.82 及 0.87 為較佳(見表二)。

進一步以陽性概率比與陰性概率比值來審視這些分界點。陽性概率比以(敏感度/(1-特異度))計得，亦可以解釋為測驗結果真陽性(true positive)與偽陽性(false positive)的比值，換言之陽性概率比愈高，測驗愈利於用來診斷障礙的存在，一般以大於 10 為較佳數值[17]。陰性概率比以((1-敏感度)/特異度)計得，亦可以解釋為測驗結果偽陰性(false negative)與真陰性(true negative)的比值，陰性概率比值愈低，測驗愈利於用來確定障礙的不存在，一般以小於 0.1 為較佳數值 [17]。CDIITMS 以 TMDQ 70 為分界點時，陽性概率比為 29.0，陰性概率比 0.13；以 TMDQ 77.5 為分界點時，陽性概率比 8.5，陰性概率比 0.07。兩個分界點皆接近上述陽性概率比大於 10 與陰性概率比小於 0.1 之標準，但未完全符合。

而 PDMS-2 以 TMDQ 85 為分界點時，陽性概率比 81.0，陰性概率比 0.19，符合陽性概率比值大於 10 之標準，但不符陰性概率比值小於 0.1 之標準。以 TMDQ 90 為分界點時，陽性概率比 15.5，陰性概率比 0.07，完全符合陽性概率比值大於 10 與陰性概率比值小於 0.1 之標準。根據概率比之標準比較兩個分界點，PDMS-2 以 TMDQ90 為分界點較好。

分界點下的預測率

對於臨床工作者或受測兒童的家長而言，更想知道的是當測驗結果為陽性或陰性時，孩子為發展遲緩或發展正常的可能性會有多少，此即為預測值的內涵。陽性預測值的定義為測驗結果為陽性的孩子中真正為發展遲緩的比率，亦有學者稱之為測試後機率(post-test probability)，即當測驗結果為陽性時孩子可能罹患發展遲緩的機率。反之，陰性預測值為測驗結果為陰性的孩子其為發展正常的機率，這些值對於個人分數的解釋相當有用。但預測值的檢測結果會隨研究樣本障礙盛行率的不同而有不同的結果[10,11,17,29]。在評量敏感度與特異

度都固定之狀況下，障礙盛行率越高則其陽性預測值也越高，相對陰性預測值則越低 [29]。一般建議研究樣本的障礙盛行率必須與所欲使用測驗的族群障礙盛行率相同，研究結果所得的預測值方可運用於一般族群，否則需利用該障礙已知盛行率的統計加以修正應用[18,24]。

本研究引用其他文獻發展遲緩之盛行率，所得各分界點的預測率結果如表三顯示。於 CDIITMS 分界分數 TMDQ 70 時，其陽性預測值 55-74%、陰性預測值 99%；TMDQ 77.5 時其陽性預測值 26-46%、陰性預測率 99-100%。而 PDMS-2 於分界分數 TMDQ 85 時其陽性預測值 77-89%、陰性預測值 98-99%；TMDQ 90 時其陽性預測率 39-61%、陰性預測值 99-100%。Glascoe 等人認為，為免增加家長太多的焦慮，測驗的陽性預測值最好在 70%以上[28]，而 Palisano 也建議陰性預測值最好大於 95%，以減少低估轉介率，進而促使早期介入[30]。根據該原則，兩量表之分界分數相較之下，CDIITMS 之分界分數 TMDQ 70 優於 TMDQ 77.5，PDMS-2 之分界分數 TMDQ 85 優於 TMDQ 90。因此綜合敏感度、特異度、陽性概率比、陰性概率比、陽性預測值與陰性預測值各指標，本研究判斷 CDIITMS 之最佳分界點為動作發展商數 70，在此分界點下，其敏感度為 87%，特異度 97%，陽性概率比 29.0，陰性概率比 0.13，以 6%盛行率推估得陽性預測值 65%，陰性預測值為 99%。而 PDMS-2 之最佳分界點為動作發展商數 85，其敏感度為 81%，特異度 99%，陽性概率比 81.0，陰性概率比 0.19，陽性預測值 84%，陰性預測值為 99%。臨床評估者可根據上述指標作為臨床決策之參考。

然於評量表分界點之選擇上，除根據心理計量學之標準外，尚需牽涉法律與專業議題[27]。例如若所要診斷的障礙可以經由早期治療來增加改善程度，研究者可能選擇高敏感度低特異度的分界分數，降低偽陰性(false negative)的機率。反之，若障礙的介入時點對其治療成效之影響不大，則研究者可能選擇低敏感度與高特異度的分界分數，以減少偽陽性(false positive)的機率[10]。因此本研究提供各分界分數之診斷指標值，以為臨床施測之參考。

兩量表於台灣地區使用的成本效益

發展診斷評量施行的目的，除確定障礙診斷外，最大的目的更在於確定兒童接受治療的需求，分界點的選擇會影響接受治療者的多寡。通常以常模樣本為常態分佈，可得理論上小於各發展商數之百分比[19,27]，由此來當發展遲緩盛行率之估計值。根據 Portney 等人(2000)常態分佈曲線圖 z 分數，得 TMDQ70 其 z 值為-2，z 小於-2 之百分比為 2.3%，即盛行率理論值為 2.3% [31]。依此類推，得小於 TMDQ77.5、80、85 與 90 之盛行率理論值分別為 6.7%、9.2%、15.9%及 25.1%。依內政部兒童局公告全國 6 歲以下的人口數約為 193000 人，以每週接受兩種專業治療(物理、職能、語言治療)各一次，每次治療費用 600 元，並接受縣市政府社會局的交通補助費每人每月 3000 元估計。以 CDIITMS TMDQ 70 為分界點，一年所需早期療育健保費用每年為 28 億元(2,769,936,000 元)、交通補助費為 16 億元(1,598,040,000 元)；若以 TMDQ 77.5 為分界點則需增加 53 億的健保費用與 31 億的交通補助費。而 PDMS-2 在不同分界點間，經費需求也會相差 175 億之多。因此，評估工具分界點之選擇，除心理計量學考量外，社會資源與經費的多寡，也應考慮在其中。

另就評估工作本身的成本效益來看，CDIITMS 共有測試題目 97 題，施測時間只需 15 至 20 分鐘，而 PDMS-2 施測時間則須 45 至 60 分鐘，以專業人力月薪 50000 元計，PDMS-2 所需測試之人力成本為每人 234-313 元，CDIITMS 則僅有 1/3。而測驗工具的價格 CDIITMS 也只有 PDMS-2 的一半左右，不管是題數、施測時間或施測成本，CDIITMS 顯然更具經濟效益。然而如表二與表三所示，一般而言，CDIITMS 有較好的敏感度，PDMS-2 則有較佳的特異度，PDMS-2 之陽性預測值高於 CDIITMS，CDIITMS 之陰性預測值略高於 PDMS-2。因此，CDIITMS 與 PDMS-2 同樣具有良好的診斷準確率，也各有其優缺點，皆為可作為動作發展評估之診斷工具。

有關學前兒童動作發展量表，除本研究所探討之兩種量表外，尚有「布魯茵克斯-歐索瑞斯基動作精練度測試」(Bruininks-Oseretsky Test of Motor Proficiency)、「亞伯它嬰兒動作評估表」(Alberta

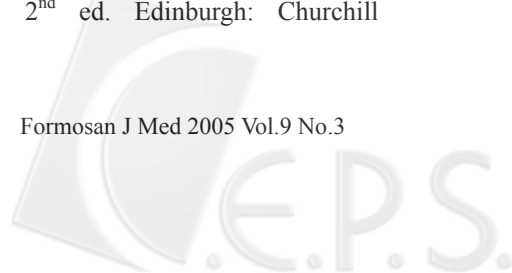
Infant Motor Scale)、「兒童動作 ABC 評量表」(Movement Assessment Battery for Children)，而如「兒童功能評估量表」(Pediatric Evaluation of Disability Inventory)與「貝萊氏嬰兒發展量表第二版」(Bayley Scales of Infant Development 2nd edition)等量表中也有動作分量表 [32]，後續可針對不同量表於台灣地區動作障礙兒童之診斷準確度進一步研究，以為臨床使用之參考。

誌謝

本研究承蒙衛生署研究經費補助(DOH92-TD-1016)，特此致謝。

參考文獻

1. Washington K, Scott DT, Johnson KA, et al: The Bayley Scale of Infant Development – II and children with developmental delays: a clinical perspective. *J Dev Behav Pediatr* 1998;19:346-9.
2. Shapiro BK, Palmaer FB, Capute AJ: The early detection of mental retardation. *Clin Pediatr* 1987; 26: 215-20.
3. 王天苗、蘇建文、廖華芳、林麗英、鄒國蘇、林世華：嬰幼兒綜合發展測驗之編製報告。測驗年刊 1998;45:19-46。
4. 吳雪玉：嬰幼兒綜合發展測驗動作分測驗的同時效度與診斷效率檢測。[碩士論文]。台北：台灣大學，2002;92。
5. Liao HF, Wang TM, Yao KP, Lee WY: Concurrent Validity of the Comprehensive Developmental Inventory for Infants and Toddlers with the Bayley Scales of Infant Development-II in Preterm Infants. *J Formos Med Assoc* 2005, in press.
6. Stengel TJ: Assessing motor development in children. In: Campbell SK ed. *Pediatric Neurological Physical Therapy*. 2nd ed. New York: Churchill Livingstone, 1991: 33-66.
7. Palisano RJ, Lydic JS: The Peabody Developmental Motor Scales: an analysis. *Phys Occup Ther Pediatr* 1984; 4: 69-75.
8. Hinderer KA, Richardson PK, Atwater SW: Clinical implication of the Peabody Developmental Motor Scales: a constructive review. *Phys Occup Ther Pediatr* 1989; 9: 81-106.
9. Folio MK, Fewell R: Peabody Developmental Motor Scales: Examiner's Manual. 2nd ed. Austin, Texas: PRO-ED Inc., 2000.
10. Domholdt E: *Physical Therapy Research: Principles and Applications*. 2nd ed. Philadelphia: W. B. Saunders Company; 2000: 180-4, 231-4.
11. Helewa A, Walker JM: *Critical Evaluation of Research in Physical Rehabilitation: Toward Evidence-Based Practice*. Philadelphia, Pennsylvania: W. B. Saunders Company, 2000:61-78.
12. Hanley JA, McNeil BJ: The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology* 1982;143:29-36.
13. 王天苗：嬰幼兒綜合發展測驗指導手冊。修正版。台北：教育部特殊教育小組，2003。
14. 潘懿玲、廖華芳：嬰幼兒綜合發展測驗使用於正常兒童、發展遲緩兒童與早產兒之信度[摘要]。中華民國物理治療學會第四十七次學術論文研討會，台大醫學院基礎醫學大樓，台北，2003。
15. 吳雪玉、廖華芳：皮巴迪動作發展量表第二版簡介。兒童福利 2002; 3:265-74。
16. Anatasi A, Urbina S: *Psychological Testing*. 7th ed. Upper Saddle River, NJ: Printice Hall. 1997: 48-145.
17. Mant J: Is this test effective? In: Dawes M, Davies P, Gray A, et al eds. *Evidence-based Practice*. London: Churchill Livingstone. 1999:133-58.
18. Rosner B: *Fundamentals of Biostatistics*. 5th ed. Pacific Grove, CA: Brooks/Cole. 2000:58-66.
19. Sackett DL, Straus SE, Richardson WS, et al: *Evidence-based Medicine, How to Practice and Teach EBM*. 2nd ed. Edinburgh: Churchill



- Livingstone, 2000:68-93.
20. Frankenburg WK: Developmental surveillance and screening of infants and young children. *Pediatrics* 2002;102:144-5.
 21. Boyle CA, Decoufle P, Yeargin-Allsopp M: Prevalence and health impact of developmental disabilities in US children. *Pediatrics* 1994;93:399-403.
 22. Thompson JR, Carter RL, Edwards AR, et al: A population-based study of the effects of birth weight on early developmental delay or disability in children. *Am J Perinatol* 2003; 20: 321-32.
 23. Drillen CM, Pickering RM, Drumond MB: Predictive value of screening for different areas of development. *Dev Med Child Neurol* 1988; 30:294-305.
 24. Irene GH, Hans CVH: The validity of sensitivity and specificity. *Statist Med* 2000; 19:1783-92.
 25. Lijmer JG, Mol BW, Heisterkamp S, et al: Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.
 26. Shoemaker OS, Saylor CF, Erickson MT: Concurrent validity of the Minnesota Child Developmental Inventory with high-risk infants. *J Pediatr Psychol* 1993;18: 377-88.
 27. Murphy KR, Davidshofer CO: *Psychological Testing, Principles and Applications*. 5th ed. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 2001:18-107, 169-91.
 28. Glascoe FP, Byme KE: The usefulness of the Developmental Profile-II in developmental screening. *Clin Paediatr* 1993;32:203-8.
 29. Glaros AG, Kline RB: Understanding the accuracy of tests with cutting scores: the sensitivity, specificity, and predictive value model. *J Clin Psychol* 1988; 4:1013-23.
 30. Palisano RJ: Neuromotor and developmental assessment. In: Wilhelm IJ ed. *Physical Therapy Assessment in Early Infancy*. New York, NY: Churchill Livingstone, 1993:173-224.
 31. Portney LG, Watkins MP: *Foundations of Clinical Research, Applications to Practice*. 2nd ed. Upper Saddle River, NJ: Prentice Hall Health, 2000:674-675.
 32. 廖華芳：小兒物理治療。台北：禾楓書局，2004;69-103。

Diagnostic Accuracy of the Motor Subtest of Comprehensive Developmental Inventory for Infants and Toddlers and the Peabody Developmental Motor Scales-Second Edition

Hsueh-Yu Wu, Hua-Fang Liao¹, Grace Yao², Wang-Tso Lee³, Tien-Miau Wang⁴, Jeng-Yi Hsieh⁵

Abstract: Purpose: 1) Using clinical diagnosis as the criterion to investigate the overall diagnostic accuracy of both the motor subtest of the Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) and the Peabody Developmental Motor Scales-Second Edition (PDMS-2). 2) To investigate the sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR-), and then to use the prevalence rates of 4%, 6%, 9% to calculate the corresponding positive and negative predictive value (PPV & NPV) for clinical application of both scales. Methods: Eighty-five children with motor disabilities and 137 non-disabled children were recruited for this study. Both the CDIIT and PDMS-2 were administered by the same tester. The area under the receiver operating characteristics (ROC) curve was used to investigate the overall diagnostic accuracy of both tests. The cross-tabs were used to calculate the diagnostic estimates, and then the estimated prevalence rate was used to calculate the PPR and NPR. Results: The overall diagnostic accuracy of the two scales was high, with ROC area 0.97 in the CDIIT, and 0.98 in PDMS-2 ($P < 0.001$). The best cutoff point of the CDIIT was developmental quotient 70, with sensitivity 87%, specificity 97%, LR+ 29.0, LR- 0.13. While a 6% prevalence rate was assumed, the PPV and NPV were 65% and 99% respectively. And the best cutoff point of PDMS-2 was developmental quotient 85 with sensitivity 81%, specificity 99%, LR+ 81.0, and LR- 0.19. The PPV of the PDMS-2 was 84%, and the NPV was 99%. The correlations between both scales and the clinical diagnoses were high with Kappa coefficients of 0.86 and 0.82 respectively ($P < 0.001$). Conclusions: Both scales had good and similar diagnostic accuracy. The diagnosis estimates of both scales were slightly different but within an acceptable range. Both scales are useful for motor development evaluation in preschool children.

Key Words: child, diagnostic accuracy, motor assessment, receiver operating characteristics curve, sensitivity, specificity

(Full text in Chinese: Formosan J Med 2005;9:312-22)

Rehabilitation Department, Poh-Ai Hospital, I-LAN; ¹School and Graduate Institute of Physical Therapy, College of Medicine, National Taiwan University; ²Department of Psychology; ³Department of Pediatrics, National Taiwan University Hospital; ⁴Department of Special Education, National Taiwan Normal University; ⁵Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital
Address correspondence to: Hua-Fang Liao, School and Graduate Institute of Physical Therapy, College of Medicine, National Taiwan University, 7 Chun-Shan S. Rd, Taipei, Taiwan

