

5 點量尺與 3 點量尺量表之反應性比較-- 在伯格式平衡量表之發現

王淳厚 . 薛漪平* 姚開屏** 謝清麟*

研究背景與目的：一般量表設計，若增加個別項目之度量 (scaling) 等級，應可提升量表之反應性，然而最近有學者以實際的數據挑戰這個「常理」，故本研究之目的在檢驗 5 點量尺的伯格式平衡量表 (Berg balance scale, BBS-5P) 是否比轉譯 (recode) 為 3 點量尺的 BBS (BBS-3P) 具備更高的反應性 (responsiveness)。**方法：**46 位首次發病住進台大醫院復健部之中風病人，以 5 點量尺及 3 點量尺的 BBS 評估其於住院三天內及住院後 14 天之平衡能力。作者將原本 5 點量尺之 BBS 的中間 3 個級距變成同一級距，意即轉譯為 3 點量尺，並以三種反應性指標檢驗二種度量等級之 BBS 評估中風病人復健住院兩週之平衡功能恢復程度。**結果：**本研究發現二種評分方式的反應性良好而且類似。BBS-5P 與 BBS-3P 於復健住院初期之三種反應性指標分別如下：1. 效應值 (effect size) d 為：0.59 (BBS-5P) 及 0.68 (BBS-3P)；2. 標準化之反應平均值 (standardized response mean)：1.39 (5P) 及 1.44 (3P)；3. 威爾卡遜檢定 (Wilcoxon's signed ranks test) 之 p 值： <0.001 (5P) 及 <0.001 (3P)。以上結果顯示 BBS-5P 與 BBS-3P 用於復健住院初期之中風病患皆具有良好的反應性，即 BBS-5P 之反應性未比 BBS-3P 高。本研究亦發現 BBS-3P 之內在一致性良好，且同時效度高。**結論：**以上結果支持最近有些學者的發現：部分量表的度量等級可能過於繁複。將 BBS 簡化為 3 點量尺並不損及量表的反應性。未來研究可繼續深入探討簡化後的 BBS 量表之心理計量特性，及嘗試簡化其他量表，使量表之使用更為簡單、快速。〈FJPT 2001; 26(2):91-96〉

關鍵詞：度量等級、反應性、伯格式平衡量表、腦中風

評估量表之反應性 (responsiveness or sensitivity to change) 係指量表偵測欲測量變項之些微變化的能力。¹ 對臨床及研究而言，一個良好的量表能適當地反映病患或臨床工作者感受到或觀察到病患之功能變化。因此 Wilkin, Hallam 與 Doggett 等人認為反應性是判斷評估量表良莠的重要標準之一，測量工具之反應性也是療效驗證研究是否成功的重要基礎。因此，如何挑選具備反應性的項目及選擇適當的度量等級以充分反映個案實際變化的情形，是測量工具發展的重大挑戰之一

。一般量表設計時，通常以增加個別項目之度量 (scaling) 等級，以便提升量表之反應性，² 如由 3 點量尺 (three point scale) 細分為 5 點量尺，或 4 點量尺細分為 7 點量尺。測量中風病患相關功能的量表中，較著名的例子，如日常生活活動 (activities of daily living, ADL) 評量工具的發展，Shah 等人³ 將巴氏量表 (Barthel index, BI) 由 2-3 點量尺提升為 5 點量尺的評分方式；生活功能獨立程度量表 (Functional Independence

中山醫學院物理治療學系

* 臺大醫學院職能治療學系

** 臺大理學院心理學系

通訊作者：謝清麟 臺灣大學醫學院職能治療學系 台北市中山南路七號

Measure, FIM) 的發展者將 ADL 個別項目的度量等級更提升為 7 點量尺。¹ 然而這些學者的作法與目的, 最近受到挑戰, 如 5 點量尺的 BI 並未比原來的 BI 具備較高的反應性,² 甚至 7 點量尺且項目較多的 FIM 與 3 點量尺為主的 BI 用於評估中風病人的反應性竟然不相上下。⁶ 所以, 量表項目之度量等級對量表反應性之影響亟待深入探討。本研究之目的在於檢驗量表項目之度量等級簡化後是否影響量表之反應性。以臨床及研究常用的伯格式平衡量表⁷(Berg balance scale, BBS) 為驗證對象, 檢驗其原始之度量等級設計 5 點量尺(BBS-5P) 是否比簡化後 3 點量尺的 BBS(BBS-3P) 對於復健住院初期之中風病人具備更高的反應性。

方 法

樣 本

樣本選自民國 87 年 9 月至 88 年 8 月連續住進台大醫院復健部的中風病患, 篩選標準如下: 一、診斷: 中風病人之診斷類別依國際疾病分類 (ICD-9-CM) 為參考依據,⁸ 凡被歸類為 ICD-9-CM430 (蜘蛛膜下腔出血), 431 (腦內出血), 432 (其他未明示之顱內出血), 433 (腦前動脈阻塞及狹窄), 434 (腦動脈阻塞), 436 (診斷欠明之腦血管疾病), 或 437 (其他及診斷欠明之腦血管疾病) 皆為本研究之對象; 二、首次中風; 三、發病二個月內; 四、能聽從口令完成評估者; 五、沒有伴隨其他重大疾病 (如: 惡性腫瘤、截肢等); 六、同意參與此研究。

評估工具

BBS⁷ 為一臨床及研究常用之平衡量表, 其主要評估日常生活中常見的平衡能力, 共 14 項, 主要分為三部份: 一、評估病人維持不同難度姿勢的能力, 例如: 坐在無靠背椅子上、維持站姿、兩腳並攏站立、兩腳前後站立、到最難的單腳站立等, 由於病人之支撐面積減少, 難度隨之增高; 二、評估病人在不同姿勢間的轉換能力, 例如: 由坐到站、轉位、轉身、在站姿下撿起地上的東西、由站到坐等; 三、評估病人調整身體重心之能力, 例如: 兩腳交替踏在階梯上、站立時將平舉的手臂盡量前伸等。BBS 之度量等級為 5 點量尺 (0-4 分), 0 分表示病人無法完成, 1-3 分代表從事活動的能力高低, 如需要不同程度的協助或完成活動的時間長短, 4 分表示病人能獨立達成或在合理的時間內完成, 總分為 0-56 分。每個項目皆有一定的評分標準, 評分手冊可向原作者索取。BBS 經驗證對中風病人之測量具有良好的內在一致性、施測者間信度及同時效度等。^{9,10}

福格邁爾運動量表¹¹ (Fugl-Meyer motor assessment scale, FM) 為國內外廣泛使用之評估中風病人動作控制能力的量表, FM 含六個次量表: 上、下肢動作功能、關節活動度 (range of motion)、疼痛、感覺缺失及平衡能力。評分方式為 0-2 分, 共三個等級: 病人完全無法從事者得 0 分, 能獨立完成者得 2 分。此量表具良好之信度及效度。^{12,13} FM 之上、下肢動作功能二次量表為本研究採用, 藉以呈現病人動作缺失之嚴重程度, FM 包含上肢 33 個項目、下肢 17 個項目, 共 50 個項目, 總分為 0-100 分。

資料蒐集

參與研究的病人於復健部住院後三天內及住院 14 天時, 分別接受上述所有的評估。BBS 與 FM 各由一位職能治療師對病人負責施測, 評估順序隨機安排。二位治療師對於評估結果互相不告知。病人的基本資料 (如年齡、性別等), 及病情資料 (如診斷、發病日期等) 皆由查閱病歷記錄獲得。

資料分析

資料分析步驟為三: 一、簡化度量等級: 將 BBS 原本 5 點量尺之每一項目 (0-4) (BBS-5P), 簡化為 3 點量尺 (0-2) (BBS-3P), 即將 BBS 各項目中原本評分為 1-3 分者合併為同一等級, 而各項目之最低分 0 分與最高分 4 分則保持不變 (即 0 轉譯為 0; 1、2、3 轉譯為 1; 4 轉譯為 2)。將各項中 1-3 分的情況合併為同一等級, 主要因為這些評分等級代表個案需要不同程度的協助或完成活動之時間較長, 且由於 BBS 每個項目中 1-3 分的評分標準並不一致, 若能結合為同一等級, 將可簡化評分的過程。二、信度及效度驗證: 量表之信度以 Cronbach's alpha 檢驗二種度量等級量表項目之內在一致性 (internal consistency)。量表之效度使用 Spearman rho 檢驗二種度量等級量表間的關連性, 以驗證 BBS-3P 之同時效度。三、以三種指標檢驗二種度量等級的 BBS 量表應用於中風復健住院二週的病人的反應性: 1. 效應值 (effect size) d : (平均變化值/起始時段分數之標準差), 本研究之起始時段為第一次 (病人於復健部住院後三天內) 之評估結果。效應值主要提供評估工具反應性大小的數值; 效應值越大, 反應性越高。2. 標準化之反應平均值 (standardized response mean): (平均變化值/前後變化分數之標準差), 是為另一種計算效應值的方式。Cohen 認為效應值若大於 0.2, 即已有小的效應 (small effect), 大於 0.5 代表中度效應 (medium effect), 大於 0.8 代表高度效應 (large effect)。¹⁴ 所以效應值高於 0.2 為判斷量表是否具備反應性的最低標準。3. 以 Wilcoxon's signed ranks test 檢驗二種度量等級的 BBS 量表評量到之病人平衡能力變化 (

反應性)是否具備統計上顯著 (statistical significance) 之水準。有關反應性指標的計算方式及其意義,讀者可參考Husted等¹⁵及Stratford等學者¹⁶之文獻。此外,將病人以嚴重程度分為二類:較輕微中風(FM分數30-100)及較嚴重中風(FM分數0-29),再以上述反應性指標驗證BBS-5P與BBS-3P之反應性。

結 果

共有46位中風病人符合樣本篩選標準,並於復健住院初期完成各項評估,病人之基本資料列於表1。

由病人住院三日內及二週後於BBS-5P項目得分計算出之內在一致性alpha值分別為:0.96及0.97;BBS-3P項目之alpha值分別為:0.95及0.96,這些結果代表BBS-5P與BBS-3P之內在一致性皆很高。病人住院三日內及二週後,BBS-5P與BBS-3P二者總分之rho係數分別為:0.98及0.99,二者之相關性極高,代表BBS-3P之同時效度高。

BBS-5P與BBS-3P對於復健住院初期中風病人的反應性,列於表2。結果顯示BBS-5P與BBS-3P之反應性良好、具備統計顯著程度($p < 0.001$)、而且反應性大小相當。二種度量等級的BBS量表之於不同嚴重程度的病人皆有良好的反應性、具備統計顯著程度($p < 0.001$)、而且反應性大小差異不大(表3)。

表1. 研究對象之基本資料與病情資料(n=46)

性別(男/女)	27/19
年齡(歲, Mean \pm SD)	64.5 \pm 10.1
診斷	
出血性中風	17
缺血性中風	29
患側(左/右)	24/22
發病至初次評估日期之天數(Median (range))	20 (7-50)
住院時之動作能力: FM分數 (Mean \pm SD)	
上肢動作分數	19.3 \pm 20.2
下肢動作分數	14 \pm 9.8
上、下肢動作分數	32.3 \pm 27.3
住院時之平衡能力: BBS分數 (Mean \pm SD)	14.5 \pm 14.3

FM: Fugl-Meyer motor assessment scale
BBS: Berg balance scale

表2. 二種度量等級BBS之反應性比較

	平均變化分數 (SD)	效應值 <i>d</i>	標準化之反應 平均均值	<i>p</i> *
BBS-5P	8.4 (6)	0.59	1.39	<0.001
BBS-3P	4.1 (2.8)	0.68	1.44	<0.001

BBS: Berg balance scale

5P: 5點量尺(0-1-2-3-4)

3P: 3點量尺(0-1-2)

* 使用Wilcoxon's signed ranks test

表3. 二種度量等級BBS於不同嚴重程度中風病人之反應性比較

	平均變化分數 (SD)	效應值 <i>d</i>	標準化之反應 平均均值	<i>p</i> *
較輕微者 (n=19)				
BBS-5P	8.2 (6.9)	0.85	1.19	<0.001
BBS-3P	4.1 (2.8)	1.08	1.3	<0.001
較嚴重者 (n=27)				
BBS-5P	8.7 (16)	0.55	1.83	<0.001
BBS-3P	4.5 (7.1)	0.64	1.65	<0.001

BBS: Berg balance scale

5P: 5點量尺(0-1-2-3-4)

3P: 3點量尺(0-1-2)

* 使用Wilcoxon's signed ranks test

討 論

在臨床上,量表之反應性代表量表偵測病人功能變化的能力;在療效研究上,反應性是呈現治療效果的指標,所以反應性是臨床及研究中用以選擇量表的重要標準。¹⁵本研究檢驗5點量尺的BBS量表是否比3點量尺的BBS具備更高的反應性。結果發現:對於復健住院初期的中風病人而言,二種評分方式的BBS反應性均良好,也就是說BBS-5P竟然

未比 BBS-3P 具備較高的反應性。研究者將中風病人依嚴重程度分成二大類，再分別驗證二種度量等級 BBS 的反應性，亦獲致同樣結果。這些結果支持最近有些學者的發現：量表之度量等級和項目之多寡與反應性高低未必成正比。³⁶ 以上研究結果代表部分量表的度量等級可能過於繁複，這些結果或有助於簡化臨床常用量表之度量等級，使量表之使用更為簡單、快速，而且仍能反應病人的功能變化，值得臨床及研究人員參考。

造成簡化量表度量等級卻不影響量表之反應性的原因尚未明瞭，可能是較多的度量等級，施測較複雜，尤其是本研究所採用的 BBS 量表，其每個項目之評分標準並不一致，可能因此使量表之穩定度（信度）偏低（如造成施測者內信度偏低），意即提高測量結果的變異量，進而影響量表之反應性。³⁷ 姚開屏與陳坤虎³⁸亦指出：過度提高量表項目之度量等級可能降低測量結果的準確度及可信度。因此後續研究可驗證度量等級對量表信度（如施測者間信度與施測者內信度）及效度的影響。

本研究以轉譯（recode）的方式，將 5 點量尺的 BBS 簡化為 3 點量尺，結果發現 BBS-3P 項目之內在一致性極高，與原始 BBS-5P 項目之內在一致性相當，此結果代表 BBS-3P 與 BBS-5P 量表項目間之同質性高。本研究亦發現病人於 BBS-3P 與 BBS-5P 所得總分之相關性極高，代表 BBS-3P 的同時效度良好。這些初步結果支持 BBS-3P 之信度與效度，且如上所述，二者之反應性亦相當，由於 3 點量尺的評分方式遠較 5 點量尺簡單，而且費時較少，因此 3 點量尺的 BBS 似乎可取代 5 點量尺的 BBS，然而此點尚須進一步驗證。

有關 5 點量尺 BBS 的轉譯方式可有許多種，如轉譯成不同點數（例如 2、4 點量尺）或轉譯方式不同（例：將 0-1 轉譯為 0；2 轉譯為 1；3-4 轉譯為 2 等）。然而若轉譯成 2 點量尺，測量結果將過於粗略；若轉譯成 4 點量尺則與 5 點量尺差異不大，較無簡化之意義。另外，若將 0-1 轉譯為 0，將提高評估結果之下限效應（floor effect）；若將 3-4 轉譯成同一點量尺，將提高評估結果之上限效應（ceiling effect）。故本研究僅採用一種轉譯方式，主要因為 3 點量尺之評量模式簡單、易於臨床及研究使用，而且將中間三點量尺簡化為一，較不影響評量結果之分數分佈。

本研究有部分限制值得讀者留意，如樣本數較少，使得病人的嚴重程度分類較為粗略。再者，選擇評估工具的標準有許多種，³⁹ 反應性僅是其中之一。若過度簡化量表之度量等級，可能影響資料的屬性及其相關的統計分析。另外，本研究直接由 BBS-5P 轉譯取得 BBS-3P，所得結果容易造成二者產生相關，因而高估 BBS-3P 的同時效度。所以研究者呼籲繼續深入探索評估量表之度量等級與項目多寡對於評估工具信度、效度及反應性的影響，這些研究結果將有助於臨床及

研究人員選擇較佳的評估量表。

誌 謝

本研究係由台大醫學院附設醫院醫療科技研究計畫 NTU H89S2007 資助。

參考文獻

1. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985;38:27-36.
2. Wilkin D, Hallam L, Doggett M. *Measures of Need and Outcome for Primary Health Care*. Oxford: Oxford University Press; 1992.
3. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol* 1989; 42:703-9.
4. Hamilton BB, Granger CV, Shervin FS, Zielezny FS, Tashman JS. A uniform national data system for medical rehabilitation. In: *Further MJ, ed. Rehabilitation Outcomes: Analysis and Measurements*. Baltimore: Paul H Brooks; 1987:137-47.
5. Hocking C, Williams M, Broad J, Baskett J. Sensitivity of Shah, Vanclay and Cooper's modified Barthel Index. *Clin Rehabil* 1999; 13:141-7.
6. Van der Putten JJ, Hobart JC, Freeman JA, Thompson AJ. Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel index and the Functional Independence Measure. *Journal of Neurology, Neurosurgery & Psychiatry* 1999;66: 480-4.
7. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada* 1989;41:304-11.
8. 中華民國醫院行政協會病例委員會編譯：ICD-9-CM 中英對照增訂版。台北市：巨流圖書公司；1993。
9. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehabil* 1992;73:1073-80.
10. 黃小玲、謝清麟、林伯欣、黃恢濤、楊怡君。伯格式平衡量表之信度及同時效度檢定。《職能治療學會雜誌》1995;13:21-8。
11. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient: a method for evaluation of physical performance. *Scand J Rehabil Med* 1975;7:13-31.
12. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys Ther* 1983;63:1607-10.
13. Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke. *Phys Ther* 1993;73:447-54.
14. Cohen J. *Statistical power analysis for the behavior sciences*. New Jersey: Lawrence Erlbaum Associates; 1983.
15. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epi* 2000;53: 459-468.

16. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther* 1996;76:1109-1123.
17. Rothwell PM. Responsiveness of outcome measures in randomised controlled trials in neurology. *J Neurol Neurosurg Psychiatry* 2000;68:274-5.
18. 姚開屏、陳坤虎。如何編製一份問卷：以「健康相關生活品質」問卷為例。《職能治療雜誌》1998；16：1-24。
19. Wade DT. *Measurement in Neurological Rehabilitation*. New York: Oxford University Press; 1992.

Comparison of Responsiveness of 5-point Scale and 3-point Scale: The Findings on the Berg Balance Scale

Chun-Hou Wang I-Ping Hsueh* Kai-Ping Grace Yao** Ching-Lin Hsieh*

Background and Purpose: Increasing the number of points in the scale of individual items is often used to improve the responsiveness of a scale. However, this principle has been challenged recently. The purpose of this study was to compare the responsiveness of the original 5-point Berg balance scale (BBS-5P) with a recoded 3-point BBS (BBS-3P) in patients after a stroke. **Methods:** Forty-six stroke inpatients receiving rehabilitation participated in this study. The BBS was used to assess the patients' balance at admission and 14 days after admission to the Physical Medicine and Rehabilitation Department of National Taiwan University Hospital. The three points in the center of the BBS were recoded as a single point, thus making a new three-point scale (BBS-3P). **Results:** Three indices of responsiveness obtained from the BBS-5P and BBS-3P are listed as follows: 1. effect size d : 0.59 for BBS-5P and 0.68 for BBS-3P; 2. standardized response mean: 1.39 for BBS-5P and 1.44 for BBS-3P; and 3. p value using the Wilcoxon's signed ranks test: <0.001 for BBS-5P and <0.001 for BBS-3P. These results indicate that both scales exhibit high and similar responsiveness on stroke inpatients receiving rehabilitation. Furthermore, internal consistency of the BBS-3P was high. The relationship between the BBS-5P and BBS-3P scores was extreme, indicating high concurrent validity of the BBS-3P. **Conclusions:** The study supports arguments made by previous researchers to the effect that a number of clinical measures are overcomplicated. The BBS could be simplified without sacrificing responsiveness. Further studies are needed to investigate the psychometric characteristics of the simplified BBS. All measures based on a scale larger than three might be carefully re-evaluated in the interest of simplifying measures. (FJPT 2001; 26(2):91-96)

Key words: Points of scaling, Responsiveness, Berg balance scale, Stroke

School of Physical Therapy, Chung Shan Medical & Dental College

* School of Occupational Therapy, College of Medicine, National Taiwan University

** School of Psychology, College of Science, National Taiwan University

Correspondence to: Ching-Lin Hsieh 7, Chun-Shan S. Rd, Taipei, Taiwan, ROC.

