行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※

中風病人平衡量表心理計量特性之分析與比較(2/2)

※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：■ 個別型計畫　　□整合型計畫

計畫編號：NSC 90-2314-B-002-328

執行期間：　90 年　8 月 1 日至　91 年　7 月　31 日

計畫主持人：毛慧芬 資深講師

共同主持人：謝清麟、湯佩芳

計畫參與人員：

本成果報告包括以下應繳交之附件：

□赴國外出差或研習心得報告一份

□赴大陸地區出差或研習心得報告一份

□出席國際學術會議心得報告及發表之論文各一份

□國際合作研究計畫國外研究報告書一份

執行單位：台大醫學院職能治療學系

中　華　民　國　　91　年　10　月　31　日

行政院國家科學委員會補助專題研究計畫成果報告

中風病人平衡量表心理計量特性之分析與比較(2/2)

編號：NSC 90-2314-B-002-328

執行單位：台大醫學院職能治療學系

主持人：毛慧芬 資深講師

一、中文摘要

背景與研究目的：

本研究目的為比較三種常用之中風病人平衡量表：柏格氏平衡量表(Berg Balance Scale，BBS)，富格梅爾動作測驗中的平衡次測驗(The balance subscale of Fugl-Meyer Motor Assessment Scale，BFM)，及中風病人姿勢評估量表(The Postural Assessment Scale for Stroke Patients，PASS) 之心理計量特性。本研究進行 180 天之追蹤，以了解各平衡量表對於不同階段及嚴重程度中風病人之適用性及反應性，提供臨床及研究上選擇評估工具時之參考依據。

方法：

共計有 123 位中風患者，分別於發病後 14、30、90 及 180 天接受三種量表之評量。並進行信度、效度及反應性之分析。

結果：

FM-B 與 BBS 具有明顯地板及天花板效應，而 PASS 無此現象；三量表在病患中風後不同恢復階段均呈現良好的信度及效度，而由效應值來檢測三量表之反應性,則顯示在發病後前 90 天具有中等到良好的反應性,但 90 天到 180 天則只有低程度的反應性。另外，PASS 對於發病後 14 天至 30 天的病患及嚴重之中風病患,相較於另二量表反應性佳。

結論：

三量表都是具可接受的心理計量特性而適用於臨床或研究。其中 PASS 又較優於其他二個量表。

關鍵詞：

平衡、信度、效度、反應性、腦中風

Abstract:

Background and Purpose—This study compared the psychometric properties of three clinical balance measures, the Berg Balance Scale (BBS), the Balance subscale of the Fugl-Meyer test (FM-B), and the Postural Assessment Scale for Stroke patients (PASS), in stroke patients with a broad range of neurological and functional impairment from the acute stage up to 180 days post onset.

Methods—One hundred and twenty-three stroke patients were followed up prospectively using the three balance measures on the 14th, 30th, 90th, and 180th days after stroke onset (DAS). Reliability (inter-rater reliability and internal consistency) and validity (concurrent validity, convergent validity, and predictive validity) of each measure were examined. A comparison of the responsiveness of each of the three measures was made based on the entire group of patients and three separate groups classified by degree of neurological severity.

Results— The FM-B and BBS showed a significant floor or ceiling effect at some DAS points, while the PASS did not show these effects. The BBS, FM-B, and PASS all had good reliability and validity for patients at different recovery stages after stroke. The results of effect size demonstrated fair to good responsiveness of all three measures within the first 90 DAS, but, as expected, only a low level of responsiveness at 90 to 180 DAS. The PASS was more responsive to changes in severe stroke patients at the earliest period after stroke onset, 14-30 DAS.

Conclusions—All three measures tested showed very acceptable levels of reliability, validity, and responsiveness for both clinicians and researchers. The PASS showed slightly better psychometric

characteristics than the other two measures.

Key Words—balance, cerebrovascular disorders, reliability, validity, responsiveness

二、緣由與目的

Balance training is an important component of stroke rehabilitation.[1,2] Several studies have found that changes in balance ability correlate significantly with changes in function.[3-7] Measuring balance can assist the clinician in diagnosis, selecting the most appropriate therapy, and outcome measurement.[3,8]

A variety of laboratory approaches to assess balance have been proposed,[9-16] but the functional scales of balance measures are most commonly applied to stroke patients in clinical settings.[14,17] To date, more than 15 different functional scales measuring balance have been developed and used in stroke research.[11,13,14,17-20] However, only a few are specifically designed for stroke patients. [17] The balance subscale of the Fugl-Meyer test (FM-B) [19] and the Berg Balance Scale (BBS) [18] are the most commonly used. Recently, Benaim et al[17] adapted items from the FM-B and developed a new scale, the Postural Assessment Scale for Stroke Patients (PASS) for measuring balance function in stroke patients.

To be clinically useful, a scale must be scientifically sound in terms of three basic psychometric properties: reliability, validity and responsiveness.[8,23,24] Although many researchers have examined the reliability and validity of each of the three balance measures described above,[10,17-22,25-29] some limitations were noted. First, these studies did not compare the properties of different balance measures on the same cohort of patients. Secondly, most of the subjects in previous studies were tested only up to three months after stroke onset. Furthermore, no previously reported studies have evaluated whether these balance measures have similar psychometric properties for patients with different degrees of neurological severity. The purpose of this prospective study was to compare the reliability, validity and responsiveness of these three balance measures concurrently in a cohort of patients on the 14th, 30th, 90th and 180th day after stroke onset (DAS).

三、方 法：

Subjects

Subjects were patients with stroke admitted to National Taiwan University Hospital. Patients were included in the study if they met the following criteria: (1) diagnosis (ICD-9-CM codes) of cerebral hemorrhage (431), cerebral infarction (434), or other (430, 432, 433, 436, 437); (2) first onset of CVA, without other major diseases; (3) stroke onset within 14 days prior to hospital admission; (4) ability to follow commands; and (5) ability to give informed consent personally or by proxy. The clinical diagnosis of stroke was confirmed by neuroimaging examination (CT/MRI). Subjects were excluded if they suffered from another stroke or other major diseases during the follow-up period.

Procedures

The three balance measures and related measures were administered to patients on the 14th, 30th, 90th and 180th DAS. The protocol of this study was divided into two parts. The first part was an inter-rater reliability study. The three balance measures were administered by two occupational therapists (A and B) individually, on the same patients on the 14th DAS, and within a 24 hour period. The second part of the protocol was a validity and responsiveness study. In this study, the Barthel Index (BI) [30] was used as the external criteria (for the examination of convergent validity) and was administered on the 14th, 30th, 90th and 180th DAS. The walking subscale of the Motor Assessment Scale (MAS)[31] was also used as an external criterion (for the examination of predictive validity) to evaluate the performance of ambulation on the 180th DAS. The degrees of responsiveness of the three balance measures were calculated based on the changes occurring between 14th to 30th DAS, 30th to 90th DAS, 90th to 180th DAS, 14th to 90th DAS, and 14th to 180th DAS.

Instruments

The Balance Measures

The FM-B is one of six subscales of the FM which was designed to evaluate impairment following stroke.[19] The FM-B contains 7 three-point items, three for sitting and four for standing. The total score ranges from 0 to 14. Results of previous studies investigating the reliability and validity of the FM-B have been controversial.[19-21,28,29]

The BBS[22] evaluates a person's performance on 14 items (one sitting and 13 standing items) related to balance function that are frequently encountered in everyday life. The scoring method is based on a five-point ordinal scale of 0 to 4, with the total score ranging from 0 to 56.[22]

The PASS[17] was developed to be applicable to all stroke patients, even for those with very poor postural performance. The PASS contains 12 four-point items which grade performance for situations of varying difficulty in maintaining or changing a given lying, sitting, or standing posture. It has a total score ranging from 0 to 36. The psychometric properties of the PASS have been reported to be satisfactory in stroke patients during the first 3 months after stroke.[17] However, no studies of the responsiveness of the PASS have been reported.

Statistical analysis

Distribution

The score range and distribution of each of the three measures were examined. Floor and ceiling effects exceeding 20% are considered to be significant.[36] The existence of the floor and ceiling effects are indicative of the limited ability of a measurement to discriminate between subjects.

Reliability

Item reliability. The inter-rater agreement on individual items of the three balance measures was analyzed using the weighted kappa statistic. The weighted kappa score measures the agreement among raters adjusted for the amount of agreement expected by chance and the magnitude of disagreements.[37] A kappa value greater than 0.75 indicates excellent agreement, 0.4-0.75 indicates fair to good agreement, and less than 0.4 indicates poor agreement.[38]

Total score reliability. The inter-rater reliability of the total score of the three balance measures was analyzed using the intra-class correlation coefficient (ICC) statistic. The fixed effect of ICC Model 3[39] was used to compute the ICC value for the degree of agreement between repeated measurements by the two raters on the same patient. An ICC value of more than 0.80 indicates high reliability.[40]

Internal consistency. The internal consistency of each balance measure was expressed using Cronbach's alpha coefficients. An alpha coefficient greater than 0.70 is considered adequate for group comparison.[41]

Validity

Concurrent validity. The inter-relationship between the three balance measures on the 14th, 30th, 90th, and 180th DAS were examined using the Spearman rho correlation coefficient.

Convergent validity. The relationships between the total score of the three balance measures with those of the BI at each DAS point were examined using the Spearman rho correlation coefficient.

Predictive validity: The predictive validity of the three balance measures was assessed by comparing the results of the three balance measures on the 14th, 30th, and 90th DAS with that of the MAS at 180 DAS using the Spearman rho correlation coefficient.

Responsiveness

Two approaches were employed in this study. First, effect size (ES) was calculated by dividing the mean change scores by the standard deviation of the change score in the same subjects, according to Cohen's criteria.[42] Furthermore, patients were stratified into one of the following three groups based on their FM scores: 0 to 35, severe; 36-79, moderate; and 80 or greater, mild.[43]

四、結 果

A total of 128 patients with a wide spectrum of balance deficits were recruited in the study. Table 1 presents detailed characteristics of the cohort of patients in the study.

Distribution

Table 2 shows the distributions of the three balance measures at 4 DAS points. For the entire sample, scores at the 4 DAS points spanned virtually the entire range; however, the FM-B and BBS showed notable floor effects at 14 DAS, and the BBS also showed significant ceiling effects at 90 and 180 DAS.

Reliability

One hundred and twelve patients participated in the reliability investigation. The medians (ranges) of weighted kappa statistics for each item of the PASS, FM-B and BBS were 0.88 (0.61 - 0.96), 0.79 (0.71 - 0.95) and 0.92 (0.59 - 0.94), respectively, indicating good individual item agreement. The ICCs (95% CI) for the total scores of the PASS, FM-B and BBS were 0.97 (0.95 - 0.98), 0.92 (0.88 - 0.95) and 0.95 (0.93 - 0.97), respectively, indicating excellent total score agreement. The Cronbach's alphas of the PASS, FM-B and BBS ranged from 0.94 to 0.96, 0.85 to 0.91, and 0.92 to 0.98, respectively, on all 4 DAS, indicating high internal consistency.

Validity

Concurrent validity: Table 3 shows the intercorrelations between three balance measures. The pair-wise correlations of the three balance measures were high at each stage of stroke.

Convergent validity: The scores of the three balance measures were highly correlated with those of the BI scores (Spearman rho ≥ 0.86, $p<0.0001$)(Table 4).

Predictive validity: The scores of the three balance measures at the earlier three DAS points were highly correlated with the MAS scores on evaluations on the 180th DAS (Spearman rho≥ 0.8, $p<0.0001$)(Table 4).

Responsiveness

The ES showed that the three balance measures were moderately to highly responsive in detecting changes before 90 DAS (14 - 30 DAS: ESs ≥ 0.8, 30 - 90 DAS: ESs ≥ 0.63) and that the levels of responsiveness of these measures were low, as expected, at 90 and 180 DAS ($0.31 \leq$ ESs ≤ 0.4) (Table 5). Table 5 also shows that the changes of the three scales at each stage were all significant ($p \leq 0.006$). Table 6 shows the responsiveness of the three balance measures at different stages for subjects with different levels of stroke severity (ESs ≥ 0.21). All of these results indicate that the three balance measures are generally sensitive to change over time after a stroke. In particular, the BBS was found to be less responsive than the FM-B and the PASS for severe stroke patients at 14 to 30 DAS (Table 6).

五、討 論

(FB-M, BBS, and PASS) for stroke patients were systematically compared for the first time. This study recruited stroke subjects with a wide range of degrees of neurological severity from a major academically based teaching hospital in Taiwan. In addition, this study followed subjects at 4 specific time points after stroke for an extended period (up to 180 DAS) in order to evaluate how appropriate these measures are for use at different recovery stages post stroke. Furthermore, the data of responsiveness were analyzed according to different degrees of neurological severity of the subjects.

Distribution

Distribution of balance measures is rarely reported in clinical studies. However, it is important to show the score distribution of the study sample to understand whether the scale is measuring a restricted range of stroke patients. In this study, analysis of the score distribution of the three balance measures at 14 DAS revealed some limitations of these measures. The BBS and FM-B showed significant floor effects at 14 DAS while the PASS did not. Almost one third of the subjects had the lowest scores on these two measures. The reason for this result might be that the least demanding test in the BBS and FM-B is to sit independently; however, some stroke patients may not regain their ability to sit independently in the very early stage[7,44] thus leading to significant floor effects. These results indicated that the PASS, including four items on bed mobility, was more appropriate than the BBS and FM-B in assessing patients in the early recovery stage. We also found that the BBS had significant ceiling effects on the 90th and 180th DAS. These results indicate that the BBS might not discriminate the patients' balance function after 90 DAS.

Reliability

In this study, the reliabilities of these three balance measures were examined in terms of inter-rater reliability and internal consistency. In agreement with many previousstudies,[10,18,21] our results indicated that both the item and total score inter-rater reliabilities and internal consistency of these three measures were equally high. In particular, the individual item and total score reliability of the FM-B after adjusting for the scoring criteria of the sitting items was much higher than those found in a previous study.[28] Thus, the reliability of FM-B with the modified sitting balance items was well supported.

The very high internal consistency of the three balance measures indicated that the items of each of these instruments measured the same concept – balance.[11,22] However, the extremely high internal consistency of the three measures might indicate the possibility of item redundancy, which needs further examination.

Validity

Results of concurrent, convergent, and predictive validity of all three measures were generally in accordance with the findings of previous studies.[5,10,17,19,26-28] For example, previous studies found that BBS scores were correlated to motor performance,[10] to ADL function,[10,27] and to walking ability.[27] The findings of this study further confirmed the validity of the three measures.

Responsiveness

The responsiveness of an instrument is of key importance in outcome studies. If the instrument is unable to detect change in balance function, an intervention that improves balance may indicate no significant differences between treated and untreated patients. Unfortunately, this property is often overlooked, and information about the responsiveness of the three measures is scarce.

The results of the ES indicate that these three measures had fair to good levels of responsiveness before 90 DAS and in the overall stages (14 - 90 DAS and 14 - 180 DAS) of recovery. In addition, at later stages (90 - 180days) of recovery the three measures had, as expected, only low levels of responsiveness. This might have been due to a plateau in the improvement of balance function after 90 DAS. The motor and ADL functions have also been reported to reach a plateau after 90 DAS.[43] The other possible reason might be that these three balance measures lack items sensitive enough to detect patients' improvement after 90 DAS.

Investigation of how disease severity affects the responsiveness of the three balance measures revealed that the BBS was less responsive than the PASS and FB-M in severe stroke patients at the initial stages

(14-30 days). The reason for this finding might be that the BBS was not originally designed for stroke patients,[22] and only one item of the scale assesses balance ability in the sitting position. Since the sitting balance is one of the first postures to be restored after a stroke, it seems that BBS is lacking items to detect change in patients who are unable to stand independently.

Level of scaling and number of items

The numbers of scoring levels on the FM-B, PASS and BBS are 3, 4, and 5, respectively. The number of items in these three scales also varies. However, as found in this study, the reliability and responsiveness of these three measures are generally similar if the subjects are considered as a whole. These results indicate that increasing the number of items or the grading levels does not improve the responsiveness or decrease the reliability of these three balance measures. Interestingly, a recent study found that the BI (10 items, mainly a 3-point scale) and the motor Functional Independence Measure (13 items, a 7-point scale) showed similar responsiveness in patients with stroke and multiple sclerosis.[45] The similar responsiveness of the three balance measures has important implications for both clinicians and researchers. The FM-B and the PASS are quicker and simpler to rate than the BBS. From this point of view, the BBS is the least suitable instrument for use in both clinical and research settings.

One of the limitations of this study was that the intra-rater reliability of the three measures was not examined. Some studies have reported excellent intra-rater reliability results for the BBS[18] and the PASS.[17] In addition, we found a high inter-rater reliability of the measures. Therefore, the intra-rater reliability of the three measures might not be an issue of important concern.

In summary, the BBS, FM-B, and PASS are clinical balance measures with good reliability, good validity, and accepted responsiveness at different post-stroke stages of recovery. The PASS showed slightly better psychometric characteristics than the other two measures and thus appears to be better suited for use by clinicians and researchers.

六、參考文獻：

1. Bobath B. *Adult Hemiplegia: Evaluation and Treatment.* 3rd ed. London: William Heinemann Medical Books; 1990.
2. Ryerson SD. Hemiplegia. In: Umphred DA, ed. *Neurological Rehabilitation.* 3rd ed. St Louis: CV Mosby Co; 1995:681-721.
3. Bohannon RW, Leary KM. Standing balance and function over the course of acute rehabilitation. *Arch Phys Med Rehabil.* 1995;76:994-996.
4. Dettmann MA, Linder MT, Sepic SB. Relationships among walking performance, postural stability, and functional assessments of the hemiplegic patient. *Am J Phys Med.* 1987;66:77-90.
5. Juneja G, Czyrny JJ, Linn RT. Admission balance and outcomes of patients admitted for acute inpatient rehabilitation. *Am J Phy Med Rehabil.* 1998;77:388-393.
6. Nichols DS, Miller L, Colby LA, Pease WS. Sitting balance: its relation to function in individuals with hemiparesis. *Arch Phys Med Rehabil.* 1996;77:865-869.
7. Sandin KJ, Smith BS. The measure of balance in sitting in stroke rehabilitation prognosis. *Stroke.* 1990;21:82-86.
8. Wade DT. *Measurement in Neurological Rehabilitation.* Oxford: Oxford University Press; 1992
9. Berg K. Balance and its measure in the elderly: a review. *Physiother Can.* 1989;41:240-245.
10. Berg K, Maki BE, Williams JI, Holliday PJ. Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehabil.* 1992;73:1073-1080.
11. Berg K, Norman KE. Functional assessment of balance and gait. *Clin Geriatr Med.* 1996;12:705-723.
12. Horak FB. Clinical measurement of

postural control in adults. *Phys Ther.* 1987;67:1881-1885.

13. Horak FB, Esselman P, Anderson ME, Lynch MK. The effects of movement velocity, mass displaced, and task certainty on associated postural adjustments made by normal and hemiplegic individuals. *J Neurol Neurosurg Psychiatry.* 1984;47:1020-1028.
14. Leonard E. Balance tests and balance responses: performance changes following a CVA: a review of the literature. *Physiother Can.* 1990;42:68-72.
15. Stevenson TJ, Garland J. Standing balance during internally produced perturbations in subjects with hemiplegia: validation of the balance scale. *Arch Phys Med Rehabil.* 1996;77:656-662.
16. Whitney SL, Poole JL, Cass SP. A review of balance instruments for older adults. *Am J Occup Ther.* 1998;52:666-671.
17. Benaim C, Pérennou DA, Villy J, Rousseaux M, Pelissier JY. Validation of a standardized assessment of postural control in stroke patients: the Postural Assessment Scale for Stroke Patients (PASS). *Stroke.* 1999;30:1862-1868.
18. Berg K, Wood-Dauphinee S, Williams JI. The balance scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med.* 1995;27:27-36.
19. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient I. a method for evaluation of physical performance. *Scand J Rehabil Med.* 1975;7:13-31.
20. Poole JL, Whitney SL. Motor assessment scale for stroke patients: concurrent validity and interrater reliability. *Arch Phys Med Rehabil.* 1988;69:195-197.
21. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys Ther.* 1983;63:1606-1610.
22. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can.* 1989;41:304-311.
23. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. *J Neurol Neurosurg Psychiatry.* 1996;60:127-130.
24. Streiner DL, Norman GR. *Health measurement scales.* 2nd ed. Oxford: Oxford University Press; 1995.
25. Berg K, Wood-Dauphinee S, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health. 1992;83(Suppl 2): S7-S11.
26. Wood-Dauphinee S, Berg K, Bravo G, Williams JI. The balance scale: responsiveness to clinically meaningful changes. *Can J Rehabil.* 1997;10:35-50.
27. Hsueh IP, Mao HF, Huang HL, Hsieh CL. Comparisons of responsiveness and predictive validity of two balance measures in stroke inpatients receiving rehabilitation. *Formosan J Med.* 2001 [in Chinese] (in press).
28. Malouin F, Pichard L, Bonneau C, Durand A, Corriveau D. Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer assessment and the motor assessment scale. *Arch Phys Med Rehabil.* 1994;75:1206-1212.
29. 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;73:447-454.
30. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J.* 1965;14:61-65.
31. Carr JH, Shepherd RB, Nordholm L, Lunne D. Investigation of a new motor assessment scale for stroke patients. *Phys Ther.* 1985;65:175-180.
32. Berglund D, Fugl-Meyer A. Upper extremity function in hemiplegia: a cross-validation study of two assessment methods. *Scand J Rehabil Med.* 1986;18:155-157.

33. Richards SH, Peters TJ, Coast J, Gunnell DJ, Darlow MA, Pounsford J. Inter-rater reliability of the Barthel ADL index: how does a researcher compare to a nurse? *Cli Rehabil.* 2000;14:72-78.
34. Gosman-Hedstrom G, Svensson E. Parallel reliability of the functional independence measure and the Barthel ADL index. *Disabil Rehabil.* 2000;22:702-715.
35. Hsueh IP, Lee MM, Hsieh CL. The Psychometric Characteristics of the Barthel ADL Index in Patients with Stroke. *J Formos Med Assoc.* 2001 (in press).
36. Holmes W, Shea J. Performance of a new, HIV/AIDS-targeted quality of life (HAT-QOL) instrument in asymptomatic sero-positive individuals. *Qual Life Res.* 1997;6:561-571.
37. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70:213-220.
38. McCluggage WG, Bharucha H, Caughley LM, Date A, Hamilton PW, Thornton CM, Walsh MY. Interobserver variation in the reporting of cervical colposcopic biopsy specimens: comparison of grading systems. *J Clin Pathol.* 1996;49:833-835.
39. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420-428.
40. Richman J, Makrides L, Prince B. Research methodology and applied statistics. Part 3: measurement procedures in research. *Physiother Can* 1980;32:253-257.
41. Ware JE Jr. SF-36 *Health survey: manual and interpretation guide*. Boston, Massachusetts: The Health Institute, New England Medical Centre; 1993.
42. Cohen J. *Statistical power analysis for the behavior sciences.* New Jersey: Lawrence Erlbaum Associates; 1983.
43. Feigin L, Sharon B, Czaczkes B, Rosin AJ. Sitting equilibrium 2 weeks after a stroke can predict the walking ability after 6 months. *Gerontology.* 1996;42:348-353.

附錄：

Table 1. Characteristics of the stroke patients participating in the study (n=123)

Characteristic		
Gender	male/female	66/57
Age	mean year (SD)	69.3 (11.2)
Diagnosis	cerebral hemorrhage	32 (26%)
	cerebral infarction	74 (60%)
	other	17 (14%)
Side of hemiplegia	right/left	56/67
PASS (14 DAS)	mean (SD)	17.6 (12.8)
FM-B (14 DAS)	mean (SD)	6 (4.8)
BBS (14 DAS)	mean (SD)	22.3 (22.2)
BI (14 DAS)	mean (SD)	39 (30.5)
FM (14 DAS)	mean (SD)	50.8 (35.8)

BI: Barthel Index PASS: Postural Assessment Scale for Stroke patients
FM-B: Balance subscale of the Fugl-Meyer test
BBS: Berg Balance Scale FM: motor subscale of the Fugl-Meyer test
DAS: days after stroke onset

Table 2. Distributions of the three balance measures at different stages of stroke recovery

	PASS		FM-B		BBS	
	Floor effect n (%)	Ceiling effect n (%)	Floor effect n (%)	Ceiling effect n (%)	Floor effect n (%)	Ceiling effect n (%)
14 DAS (n=123)	10 (8.1)	4 (3.3)	36 (29.3)	4 (3.3)	43 (35)	6 (4.9)
30 DAS (n=110)	5 (4.5)	11 (10)	16 (14.5)	11 (10)	19 (17.3)	13 (11.8)
90 DAS (n=93)	2 (2.2)	13 (14)	6 (6.5)	13 (14)	6 (6.5)	20 (21.5)
180 DAS (n=80)	3 (3.8)	14 (17.5)	5 (5)	14 (17.5)	4 (5)	23 (28.8)

Table 3. Concurrent validity of the three balance measures at different stages of stroke recovery

	PASS vs BBS	PASS vs FM-B	BBS vs FM-B
14 DAS (n=123)	0.95	0.97	0.92
30 DAS (n=110)	0.95	0.96	0.90
90 DAS (n=93)	0.92	0.95	0.90
180 DAS (n=80)	0.93	0.95	0.92

Table 4. Convergent validity and predictive validity of the three balance measures at different stages of stroke recovery

	Convergent validity (rho)			Predictive validity (rho)		
	PASS Vs BI	FM-B vs BI	BBS vs BI	PASS vs MAS*	FM-B vs MAS*	BBS vs MAS*
14 DAS (n=123)	0.90	0.86	0.89	0.86	0.80	0.82
30 DAS (n=110)	0.92	0.89	0.94	0.86	0.85	0.84
90 DAS (n=93)	0.88	0.88	0.90	0.90	0.87	0.91
180 DAS (n=80)	0.92	0.89	0.91	-	-	-

Table 5. Responsiveness of the three balance measures at different stages of stroke recovery

	Effect size			Wilcoxon Z		
	PASS	FM-B	BBS	PASS	FM-B	BBS
14-30 days (n=110)	0.89	0.82	0.80	8.22*	7.16*	8.46*
30-90 days (n=93)	0.64	0.63	0.69	6.05*	5.88*	6.7*
90-180 days (n=80)	0.31	0.33	0.40	2.75†	2.74†	3.63‡
14-90 days (n=93)	1.07	1.06	1.07	8.08*	7.95*	8.04*
14-180 days (n=80)	1.12	1.14	1.11	7.65*	7.55*	7.67*

PASS: Postural Assessment Scale for Stroke patients
FM-B: Balance subscale of the Fugl-Meyer test
BBS: Berg Balance Scale
* $p < .00001$
† $p = .006$
‡ $p = .0003$

Table 6. Responsiveness (effect size) of the three balance measures at different stages of stroke recovery for subjects with different levels of severity

	Mild stroke (n=44)			Moderate stroke (n=27)			Severe stroke (n=52)		
	PASS	FM-B	BBS	PASS	FM-B	BBS	PASS	FM-B	BBS
14-30 days	0.60	0.53	0.75	1.14	1.17	1.24	1.01	0.9	0.64
30-90 days	0.29	0.42	0.21	0.5	0.63	0.73	1.02	0.85	1.07
90-180 days	0.34	0.33	0.37	0.29	0.26	0.43	0.36	0.37	0.39
14-90 days	0.76	0.78	0.81	1.18	1.14	1.28	1.51	1.41	1.24
14-180 days	0.85	0.86	0.88	1.20	1.22	1.28	1.54	1.57	1.21

PASS: Postural Assessment Scale for Stroke patients
FM-B: Balance subscale of the Fugl-Meyer test
BBS: Berg Balance Scale