

## SHORT COMMUNICATION

# A COMPARISON OF RESPONSIVENESS AND PREDICTIVE VALIDITY OF TWO BALANCE MEASURES IN PATIENTS WITH STROKE

Wan-Hui Yu, MS<sup>1</sup>, I-Ping Hsueh, MA<sup>1</sup>, Wen-Hsuan Hou, MD, MSc<sup>2</sup>, Yen-Ho Wang, MD<sup>3\*</sup> and Ching-Lin Hsieh, PhD<sup>1\*</sup>

*From the <sup>1</sup>School of Occupational Therapy, College of Medicine, National Taiwan University and Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, <sup>2</sup>Department of Physical Medicine and Rehabilitation, E-Da Hospital and I-Shou University and <sup>3</sup>Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan.*

*\*Both authors contributed equally to this paper.*

**Objective:** To compare the responsiveness and predictive validity of the Balance Computerized Adaptive Test (Balance CAT) and the Postural Assessment Scale for Stroke patients (PASS) in inpatients with stroke receiving rehabilitation.

**Design:** A pre-post test design.

**Subjects:** Eighty-five inpatients after stroke.

**Methods:** Effect size  $d$  and Wilcoxon signed-rank test were used to assess the internal responsiveness of the Balance CAT and PASS. The changes in the Barthel Index (BI) and the mobility subscale of the Stroke Rehabilitation Assessment of Movement (MO-STREAM) scores were both chosen as the external criteria for examining external responsiveness. Moreover, to investigate the predictive validity, the admission scores of the two balance measures, and the discharge score of the BI/MO-STREAM, were examined by simple linear regression analysis.

**Results:** Both the Balance CAT and PASS had high internal responsiveness (effect size  $d \geq 0.87$ ) and fair external responsiveness ( $r^2 \geq 0.20$ ). The predictive validities of both measures were sufficient ( $r^2 \geq 0.33$ ). The Balance CAT took approximately 3 items (min–max = 2–4) to complete.

**Conclusion:** The Balance CAT and PASS have sufficient responsiveness and predictive validity in inpatients with stroke receiving rehabilitation. The Balance CAT is more efficient to administer and is thus recommended over the PASS.

**Key words:** responsiveness; predictive validity; computerized adaptive test; balance; stroke.

J Rehabil Med 2012; 44: 176–180

*Correspondence address: Ching-Lin Hsieh, School of Occupational Therapy, College of Medicine, National Taiwan University, 4th Floor, 17, Xuzhou Road, Taipei 100, Taiwan. E-mail: clhsieh@ntu.edu.tw*

Submitted February 11, 2011; accepted September 9, 2011

## INTRODUCTION

Balance deficit is common in patients with stroke and can seriously impair their function in activities of daily living (ADL). Measuring balance is important for clinicians in selecting an appropriate therapy and evaluating treatment outcomes (1,

2). Moreover, a short and precise balance measure can both enhance administration efficiency and reduce the assessment burden for raters and patients (3).

To date, several balance measures have been developed for stroke patients; however, only a few measures achieve both brevity and precision, which are needed in busy clinics (4, 5). Computerized adaptive testing (CAT) has been suggested to satisfy this need (6, 7). CAT chooses items tailored to an individual patient and skips items that are apparently too easy or too difficult for the patient (8). For instance, if a patient can stand independently, the computer “knows” not to ask whether he or she can sit without assistance. Instead, the computer asks whether he or she can pick up a pen from the floor while standing. Thus, CAT can achieve both efficient and precise assessments simultaneously. In recent years, CAT has been applied successfully to evaluate functional outcomes (e.g. lower extremity function, ADL function) in the rehabilitation field (9, 10).

Hsueh et al. (3) have developed a CAT system for assessing balance function (Balance CAT) in patients with stroke. The Balance CAT takes only approximately 4 items (83 s on average) to complete, which is only 18% of the average time of the Berg Balance Scale (BBS) (3). In addition, it has sufficient reliability and concurrent validity with the BBS (3). Therefore, given the efficiency and preliminary psychometric evidence, the Balance CAT demonstrates great potential for use in both clinical and research settings.

To improve the utility of the Balance CAT, evidence on the other psychometric properties (e.g. responsiveness and predictive validity) of this measure is needed. The purpose of this study was to compare the responsiveness and predictive validity of the Balance CAT with those of a traditional balance measure, the Postural Assessment Scale for Stroke patients (PASS), in inpatients with stroke receiving rehabilitation. In addition, we compared the efficiency of the Balance CAT and PASS in terms of the number of items needed to complete the assessment.

## METHODS

### *Participants*

A sample of patients with stroke undergoing inpatient rehabilitation at the National Taiwan University Hospital were recruited from 1

January 2009 to 31 July 2010. Inclusion criteria were: (i) diagnosis of cerebral haemorrhage or cerebral infarction; (ii) ability to follow simple instructions without severe cognitive deficits; and (iii) absence of comorbidities (e.g. brain tumour, fracture, amputation, or severe rheumatoid arthritis) that would reduce or limit a subject's ability to perform movements. Informed consent for participation was obtained from the participants personally or by proxy. Patients who did not stay in the rehabilitation ward for more than 7 days were excluded. The study was approved by the Institution Review Boards of the National Taiwan University Hospital.

#### Procedure

The Balance CAT and PASS were administered to patients at admission to the rehabilitation ward and at discharge from the hospital. Both measures were administered separately by two occupational therapists in a counterbalanced sequence. In addition, both the mobility subscale of the Stroke Rehabilitation Assessment of Movement measure (MO-STREAM) and the Barthel Index (BI) were administered to patients at admission as well as at discharge. The MO-STREAM was administered by a research assistant, and the BI was administered by the patient's attending physician. All measures were administered within 24 h. All raters were blind to both the purposes of the study and results of each other's assessments during the study period.

#### Measures

**Balance Computerized Adaptive Test.** The Balance CAT (3) is a computerized adaptive test that can be administered through a personal digital device via the internet. This measure contains 34 easily administered items and was developed to evaluate balance function in patients with stroke according to their ability. Therefore, patients with different levels of balance function were assessed by different numbers of items (less than 34), which were tailored to each patient's individual level of ability. Of the 34 items, 26 items have 2 response categories (able or unable to perform a balance-related task). The other 8 items have 3 response categories (i.e. 0: unable, 1: able to complete the task but not smoothly, and 2: able to complete the task smoothly; alternately, 0: unable, 1: able to maintain balance while performing a task for 1–5 s, and 2: able to maintain balance while performing a task for more than 5 s). The original item response theory estimates of the Balance CAT are standardized scores ranging from –2.4 to 2.3. For easier interpretation, we further linearly transformed these scores to 0 to 10 (i.e. 0: the patient is not able to pass the easiest item: sitting with trunk support for 10 s; 10: the patient is able to pass the most difficult item: hopping in place on the more affected foot for more than 5 times). The reliability and concurrent validity of the Balance CAT are sufficient in patients with stroke (3).

**Postural Assessment Scale for Stroke patients measure.** The PASS measure (11) was specially developed to assess postural control in all stroke patients, even those with very poor postural performance. The PASS contains 12 4-level (0–1–2–3) items of varying difficulty that grade performance while maintaining or changing a lying, sitting, or standing position. Its total score ranges from 0 to 36. The psychometric properties of the PASS are satisfactory in patients with stroke (11, 12).

**Stroke Rehabilitation Assessment of Movement instrument.** The STREAM instrument (13) evaluates the motor and basic mobility function of patients after stroke. It consists of 30 items equally distributed among 3 subscales: upper-limb movements, lower-limb movements, and mobility subscales. The psychometric properties of the STREAM are satisfactory in patients with stroke (13–15). In this study, only the mobility subscale (MO-STREAM) was used. This 4-point (0–1–2–3) subscale contains 10 mobility items, including rolling, bridging, supine to sitting, sitting to standing, standing for a count to 20, placing affected foot onto first step, 3 steps backward, 3 steps to affected side, 10-m walk, and walking down 3 stairs. The total score for the MO-STREAM

ranges from 0 to 30. The responsiveness of the MO-STREAM is sufficient for use with stroke patients (16). The MO-STREAM was used to examine the external responsiveness and predictive validity of the Balance CAT proposed in this study and the PASS.

**Barthel Index.** The BI, a measure of the basic ADL function, includes 10 fundamental items of ADL: feeding, grooming, bathing, dressing, bowel and bladder care, toilet use, ambulation, transfers, and stair climbing (17). The total score ranges from 0 to 100, with 3 categories of disability using the following cut-off values: severe (0–50), moderate (51–75), and mild to no disability (76–100) (18). The reliability, validity and responsiveness of the BI are sufficient in patients with stroke (19, 20). The BI was used to test the external responsiveness and predictive validity of the Balance CAT proposed in this study and the PASS.

#### Statistical analyses

**Score distribution.** The score distributions of the Balance CAT and PASS were examined. The floor effect is the percentage of the sample scoring the minimum possible score, whereas the ceiling effect represents the opposite extreme (21). Floor and ceiling effects exceeding 20% were considered notable (22).

**Internal responsiveness.** Internal responsiveness can be defined as the ability to detect change over a pre-specified time frame, in which the characteristic measured changes naturally over time or due to proven interventions (23). Two approaches were employed to examine the internal responsiveness of the Balance CAT and PASS in the periods between admission to the rehabilitation ward and discharge from the hospital. First, effect size  $d$  was defined as the observed mean change score divided by the standard deviation of the baseline score. An effect size  $d$  greater than 0.8 was large, 0.5–0.8 was moderate, and 0.2–0.5 was small (24). Secondly, we used the Wilcoxon signed-rank test to determine the statistical significance of the change in scores. In addition, a sample size of 35 was needed for an effect size  $d=0.5$  with statistical significance ( $p<0.05$ ) to achieve 80% statistical power.

**External responsiveness.** External responsiveness can be described as the relationship between change in a measure (i.e. the two balance measures in this study) and change in a reference measure of function indicator (i.e. basic ADL and mobility in this study) (23). If the relation is substantial, change in the measure can reflect a patient's functional change and support its external responsiveness. To examine the external responsiveness of the two measures, the changes in the BI and MO-STREAM scores during rehabilitation stay were both chosen as external criteria. A simple linear regression analysis was conducted to investigate the association between the change in score of the Balance CAT/PASS and the change in score of the BI/MO-STREAM. The values of  $\beta$  and  $r^2$  represent the extent of the external responsiveness (23). A value of an  $r^2$  between 0.25 and 0.64 was defined as moderate association between the changes in score of these measures, indicating sufficient external responsiveness of the Balance CAT and PASS (25).

**Predictive validity.** Predictive validity describes the ability of a measure to be a valid predictor of some future health-related criterion (26). An instrument with good predictive validity can help clinicians make a prognosis (26, 27). For example, scores on a sitting balance measure at an early stage can predict ADL function at a late stage (2, 28–31). We determined the predictive validity of the Balance CAT and PASS by the strength of the associations between the scores of the two measures at admission and those of the BI and MO-STREAM at discharge. The level of the associations, which was examined using the  $\beta$  and  $r^2$  in simple linear regression analysis, indicates the strength of the predictive validity. In addition, the  $r^2$  indicates the extent of the explanatory (predictive) power. Greater than or equal to 25% explanatory power ( $r^2\geq 0.25$ ) was considered to indicate sufficient predictive validity of the Balance CAT and PASS (25).

## RESULTS

*Participant demographics*

A total of 140 patients were originally recruited in this study. Fifty-five patients were lost to follow-up because they declined further participation or were discharged early, without notice. Eighty-five patients completed both assessments. These patients were not significantly different from those lost to follow-up in terms of demographic characteristics (i.e. age and gender) or balance-related functions (i.e. scores of the PASS, Balance CAT, BI, and MO-STREAM) ( $p > 0.07$ ). In addition, these 85 participants had a wide spectrum of balance deficits (PASS: min–max = 0–36). The BI median score at admission was 30 (min–max = 0–90), indicating that most of the patients had severe disability. The Balance CAT took approximately 3 items (min–max = 2–4) to estimate a patient's balance function at admission and discharge. Moreover, the median times for the Balance CAT were 61 s (min–max = 23–132) at admission, and 62 s (min–max = 15–110) at discharge. Further characteristics of the patients are shown in Table I.

Table I. Basic characteristics of the subjects in the study

Characteristic	Patients who completed the study ( $n=85$ )	Patients lost to follow-up ( $n=55$ )
Sex, $n$		
Male	59	34
Female	26	21
Age, years, mean (SD)	65.5 (11.6)	68.5 (14.3)
Stroke type, $n$		
Cerebral haemorrhage	29	20
Cerebral infarction	56	35
Side of hemiplegia, $n$		
Right	37	25
Left	47	29
Bilateral	1	1
Period of onset to initial evaluation, days, median (min–max)	19 (5–79)	19 (2–98)
Days of rehabilitation ward stay, median (min–max)	34 (8–78)	–
Admission BI score, median (min–max)	30 (0–90)	30 (0–85)
Discharge BI score ( $n=77$ ), median (min–max)	75 (5–100)	–
Admission MO-STREAM score, median (min–max)	9 (1–30)	8 (0–30)
Discharge MO-STREAM score ( $n=84$ ), median (min–max)	22 (2–30)	–
Admission PASS score, median (min–max)	16 (0–36)	15 (0–36)
Discharge PASS score, median (min–max)	31 (1–36)	–
Admission Balance CAT score, mean (SD)	4.0 (2.4)	3.6 (6.8)
Discharge Balance CAT score, mean (SD)	6.2 (2.0)	–

SD: standard deviation; BI: Barthel Index; MO-STREAM: mobility subscale of the Stroke Rehabilitation Assessment of Movement measure; PASS: Postural Assessment Scale for Stroke patients; Balance CAT: Balance Computerized Adaptive Test.

*Distribution of scores at admission and discharge*

In terms of score distribution, neither the Balance CAT nor the PASS showed a notable floor or ceiling effect at admission (< 15%). However, the percentage (12.9%) of the participants achieving the lowest score on the Balance CAT was higher than that (2.4%) on the PASS at admission. Neither measure displayed a notable floor or ceiling effect at discharge (< 10%).

*Internal responsiveness*

The changes in score between admission and discharge of the Balance CAT and PASS were generally large and similar (effect size  $d = 0.87$ – $0.90$ ; Table II). The changes in the two measures were all significant ( $p < 0.001$ ).

*External responsiveness*

Table II shows that change in score of the Balance CAT/PASS had fair association with that of the BI ( $r^2 = 0.20$ ). The association between change in score of the Balance CAT and that of the MO-STREAM was moderate ( $r^2 = 0.44$ ). Moreover, the association between change in score of the PASS and that of the MO-STREAM was also moderate ( $r^2 = 0.59$ ).

*Predictive validity*

Table II shows that the admission score of the Balance CAT/PASS had sufficient explanatory power to predict the discharge score of the BI (Balance CAT:  $r^2 = 0.33$ ; PASS:  $r^2 = 0.39$ ). The admission score of the Balance CAT/PASS also had sufficient explanatory power to predict the discharge score of the MO-STREAM (Balance CAT:  $r^2 = 0.57$ ; PASS:  $r^2 = 0.63$ ).

We conducted the power analyses of Wilcoxon signed-rank test and simple linear regression analysis for our sample size (85), and the statistical powers were all above 99%.

## DISCUSSION

We examined two types of responsiveness in this study. We found that the changes in score between admission and dis-

Table II. Comparison of responsiveness and predictive validity of the Balance Computerized Adaptive Test (CAT) and Postural Assessment Scale for Stroke patients (PASS)

Psychometric property	Balance CAT	PASS
Responsiveness		
Internal responsiveness		
Effect size $d$	0.90	0.87
Wilcoxon $Z$	7.3*	7.7*
External responsiveness, $\beta$ ( $r^2$ )		
Change in BI	0.44 (0.20)*	0.44 (0.20)*
Change in MO-STREAM	0.67 (0.44)*	0.77 (0.59)*
Predictive validity, $\beta$ ( $r^2$ )		
BI at discharge	0.57 (0.33)*	0.62 (0.39)*
MO-STREAM at discharge	0.76 (0.57)*	0.80 (0.63)*

\* $p < 0.001$ .

BI: Barthel Index; MO-STREAM: mobility subscale of the Stroke Rehabilitation Assessment of Movement measure.

charge (i.e. effect size) of the Balance CAT and PASS were both large. The results indicate that the Balance CAT and PASS had high internal responsiveness in inpatients with stroke receiving rehabilitation. Regarding the external responsiveness of both measures, the change in score of the Balance CAT/PASS exhibited a moderate association with that of the MO-STREAM. In addition, the change in score of the Balance CAT/PASS exhibited a fair and significant association with that of the BI. That is to say, improvement exhibited in the Balance CAT and PASS reflected a substantial functional change in mobility, and a significant functional change in ADL in stroke patients. These results support the external responsiveness of the Balance CAT and PASS. These findings demonstrate the value of the Balance CAT and PASS in measuring the recovery of balance function in stroke patients.

Our results showed that the Balance CAT and PASS had similar internal/external responsiveness. These findings are important and useful for clinicians and researchers in choosing between competing measures. In comparison with other traditional balance measures, the PASS has shown slightly better internal responsiveness than the BBS (12). However, the Balance CAT required completion of only 3 items (25% of the items of the PASS), demonstrating the efficient administration of CAT. Due to a similar responsiveness, using a more efficient measure (i.e. the Balance CAT) can not only reduce the assessment time, and thereby the burden on raters and patients, but also maintain adequate power to detect a statistically significant finding.

Early prediction of a patient's functional status is important for patient management (e.g. setting treatment goals and plans). We found that the Balance CAT and PASS at an early stage (i.e. within 3 months) could predict basic ADL and mobility at 1 month on average after the first assessment. Moreover, the predictive powers of the Balance CAT and PASS for basic ADL and mobility found in this study may have clinical utility because of their sufficient predictive powers ( $r^2 > 0.3$ ). According to previous findings, the PASS has already shown good predictive validity (11, 12). However, the Balance CAT, with fewer items needed, showed a predictive validity comparable to the PASS. These observations strongly support the predictive validity and clinical use of the Balance CAT.

Our results also showed that the predictive powers of the Balance CAT and PASS (assessing balance function) at admission were high for the MO-STREAM (assessing mobility) and moderate for the BI (assessing basic ADL function) at discharge. Similarly, Hsueh et al. (32) found that the balance function of patients with stroke at admission to the rehabilitation ward had high correlation with walking performance at discharge. Compared with the high predictive validity for mobility, the moderate predictive validity of both balance measures for basic ADL function might be expected. It could be that the basic ADL function of a patient depends on several factors (including balance, motor function, cognition, age and environmental context) (2, 33). Therefore, clinicians should consider conducting a comprehensive evaluation of the related factors at an early stage to provide better patient management in basic ADL training.

We found that neither the Balance CAT nor the PASS showed a notable floor or ceiling effect in the subjects at admission, indicating that both measures assess a wide spectrum of balance deficits. However, the Balance CAT showed a slight floor effect in the patients at admission. Therefore, the Balance CAT may have less discriminative ability for patients with extremely poor balance, who are more likely to be found at an early stage of stroke. Nevertheless, the Balance CAT did show sufficient responsiveness in our participants at a sub-acute stage. Thus, the slight floor effect of the Balance CAT seems not to affect its ability to detect balance improvement.

This study has 4 limitations. First, we examined the responsiveness of the Balance CAT and PASS only on inpatients with stroke receiving rehabilitation. In addition, most of the patients were severely disabled. Further studies are needed to examine the responsiveness of the two balance measures on stroke patients with higher levels of functioning to further validate our findings. Secondly, we investigated the predictive validity of the Balance CAT and PASS only on basic ADL function in stroke patients. As instrumental ADL (e.g. meal preparation, shopping) is suggested as another primary outcome after stroke (34), further studies are needed to examine the predictive validity of the two balance measures on stroke patients' instrumental ADL to further promote their utility. Thirdly, we did not record the time needed to complete the PASS, so we could not compare the average time of the Balance CAT with that of the PASS. Further studies should include the time information to further validate the efficiency of both balance measures. Fourthly, a large number of patients were lost to follow-up (i.e. 39% of the original recruitment), which reduced the size of our sample (i.e. from 140 to 85). Although we compared patients who completed the study with those who did not, the large number lost to follow-up may affect the generalization of our findings.

In conclusion, the results of this study provide strong evidence that both the Balance CAT and the PASS have sufficient responsiveness and predictive validity in inpatients with stroke receiving rehabilitation. Due to the short assessment time, the Balance CAT is thus suggested for use in patients with stroke in both clinical and research settings.

#### ACKNOWLEDGEMENTS

This study was supported by research grants from the National Science Council (NSC96-2628-B-002-034-MY3) and the National Health Research Institute (NHRI-EX97-9512PI & NHRI-EX98-9512PI) of Taiwan. These funding sources have not had any influence on the interpretation of data or the final conclusions drawn.

#### REFERENCES

1. Bohannon RW, Leary KM. Standing balance and function over the course of acute rehabilitation. *Arch Phys Med Rehabil* 1995; 76: 994–996.
2. Hsieh CL, Sheu CF, Hsueh IP, Wang CH. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. *Stroke* 2002; 33: 2626–2630.

3. Hsueh IP, Chen JH, Wang CH, Chen CT, Sheu CF, Wang WC, et al. Development of a computerized adaptive test for assessing balance function in patients with stroke. *Phys Ther* 2010; 90: 1336–1344.
4. Chou CY, Chien CW, Hsueh IP, Sheu CF, Wang CH, Hsieh CL. Developing a short form of the berg balance scale for people with stroke. *Phys Ther* 2006; 86: 195–204.
5. Chien CW, Lin JH, Wang CH, Hsueh IP, Sheu CF, Hsieh CL. Developing a short form of the postural assessment scale for people with stroke. *Neurorehabil Neural Repair* 2007; 21: 81–90.
6. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: Item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res* 2007; 16 Suppl 1: 133–141.
7. Jette AM, Haley SM. Contemporary measurement techniques for rehabilitation outcomes assessment. *J Rehabil Med* 2005; 37: 339–345.
8. Wainer H. Computerized adaptive testing: a primer. Mahwan: Lawrence Erlbaum Associates; 2000.
9. Hart DL, Mioduski JE, Stratford PW. Simulated computerized adaptive tests for measuring functional status were efficient with good discriminant validity in patients with hip, knee, or foot/ankle impairments. *J Clin Epidemiol* 2005; 58: 629–638.
10. Haley SM, Siebens H, Coster WJ, Tao W, Black-Schaffer RM, Gandek B, et al. Computerized adaptive testing for follow-up after discharge from inpatient rehabilitation: I. activity outcomes. *Arch Phys Med Rehabil* 2006; 87: 1033–1042.
11. Benaim C, Perennou 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.
12. Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke* 2002; 33: 1022–1027.
13. Daley K, Mayo N, Wood-Dauphinee S. Reliability of scores on the stroke rehabilitation assessment of movement (STREAM) measure. *Phys Ther* 1999; 79: 8–19.
14. Wang CH, Hsieh CL, Dai MH, Chen CH, Lai YF. Inter-rater reliability and validity of the stroke rehabilitation assessment of movement (STREAM) instrument. *J Rehabil Med* 2002; 34: 20–24.
15. Hsueh IP, Hsu MJ, Sheu CF, Lee S, Hsieh CL, Lin JH. Psychometric comparisons of 2 versions of the Fugl-Meyer motor scale and 2 versions of the stroke rehabilitation assessment of movement. *Neurorehabil Neural Repair* 2008; 22: 737–744.
16. Ahmed S, Mayo NE, Higgins J, Salbach NM, Finch L, Wood-Dauphinée SL. The stroke rehabilitation assessment of movement (STREAM): a comparison with other measures used to evaluate effects of stroke and rehabilitation. *Phys Ther* 2003; 83: 617–630.
17. Wade DT, Collin C. The Barthel ADL index: a standard measure of physical disability? *Int Disabil Stud* 1988; 10: 64–67.
18. Supervia A, Aranda D, Marquez MA, Aguirre A, Skaf E, Gutierrez J. Predicting length of hospitalisation of elderly patients, using the barthel index. *Age Ageing* 2008; 37: 339–342.
19. Hsueh IP, Lee MM, Hsieh CL. Psychometric characteristics of the barthel activities of daily living index in stroke patients. *J Formos Med Assoc* 2001; 100: 526–532.
20. Hsueh IP, Lin JH, Jeng JS, Hsieh CL. Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel Index, and 10 item Barthel Index in patients with stroke. *J Neurol Neurosurg Psychiatry* 2002; 73: 188–190.
21. 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. *J Neurol Neurosurg Psychiatry* 1999; 66: 480–484.
22. Holmes WC, Shea JA. Performance of a new, hiv/aids-targeted quality of life (HAT-QoL) instrument in asymptomatic seropositive individuals. *Qual Life Res* 1997; 6: 561–571.
23. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000; 53: 459–468.
24. Cohen J. Statistical power analysis for the behavior sciences. Hillsdale, NJ: Lawrence Erlbaum Assoc; 1988.
25. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. *Radiology* 2003; 227: 617–622.
26. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River: Pearson Prentice Hall; 2009.
27. Brown T. Construct validity: a unitary concept for occupational therapy assessment and measurement. *HK J Occupat Ther* 2010; 20: 30–42.
28. Franchignoni FP, Tesio L, Ricupero C, Martino MT. Trunk control test as an early predictor of stroke rehabilitation outcome. *Stroke* 1997; 28: 1382–1385.
29. Loewen SC, Anderson BA. Predictors of stroke outcome using objective measurement scales. *Stroke* 1990; 21: 78–81.
30. Sandin KJ, Smith BS. The measure of balance in sitting in stroke rehabilitation prognosis. *Stroke* 1990; 21: 82–86.
31. Wade DT, Skilbeck CE, Hewer RL. Predicting Barthel ADL score at 6 months after an acute stroke. *Arch Phys Med Rehabil* 1983; 64: 24–28.
32. Hsueh I, Mao H, Huang H, Hsieh C. Clinical applications of balance measures in stroke patients. *J Formos Med Assoc* 2001; 5: 261–268.
33. Veerbeek JM, Kwakkel G, van Wegen EE, Ket JC, Heymans MW. Early prediction of outcome of activities of daily living after stroke: a systematic review. *Stroke* 2011; 42: 1482–1488.
34. Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke* 2000; 31: 1429–1438.