Developing a Short Form of the Postural Assessment Scale for People With Stroke

Chi-Wen Chien, Jau-Hong Lin, PhD, Chun-Hou Wang, I-Ping Hsueh, MA, Ching-Fan Sheu, PhD, and Ching-Lin Hsieh, PhD

Objective. To develop a Short Form of Postural Assessment Scale for Stroke patients (SFPASS) with sound psychometric properties (including reliability, validity, and responsiveness). *Methods.* This study consisted of 2 parts: developing the SFPASS and cross-validation. In the 1st part, 287 people with stroke were evaluated with the PASS at 14- and 30-day post-stroke intervals. The authors reduced the number of test items that constitute the PASS by more than half (i.e., making 5-, 6-, and 7-item sets) and simplified the scoring system (i.e., collapsing the 4-level scale in the original PASS into a 3-level scale [PASS-3L]), making both 4-L and 3-L versions available. Thus, a total of 6 SFPASSs were generated. In addition, 2 external criteria, the Barthel activities of daily living index and the Fugl-Meyer motor test, were used to examine the validity of the 6 SFPASSs. The psychometric properties of the new 6 SFPASSs were compared with each other as well as with those of the original PASS to determine which scale outperformed the others. In the 2nd part of the study, the authors cross-validated the best SFPASS using another independent sample of 179 people with stroke. *Results.* All 6 SFPASSs demonstrated good reliability, validity, and responsiveness. However, the Bland-Altman plots showed that only the 5-item PASS-3L demonstrated no systematic trend between the difference and mean score of the 5-item PASS-3L and the original PASS. The 5-item PASS-3L also had psychometric properties similar to those of the original PASS, as demonstrated in a cross-validation sample. *Conclusion.* The authors' results provide strong evidence that the 5-item PASS-3L has sound psychometric properties in people with stroke. The 5-item PASS-3L is simple and fast to administer and is thus recommended.

Key Words: *Cerebrovascular disorders—Posture—Psychometrics.*

From the School of Occupational Therapy, College of Medicine, National Taiwan University, and Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taiwan (C-WC, C-LH, I-PH). School of Physical Therapy, College of Health Science, Kaohsiung Medical University, Kaohsiung, Taiwan (J-HL). School of Physical Therapy, College of Medical Technology, Chung-Shan Medical University, Taichung, Taiwan (C-HW). Institute of Cognitive Science, National Cheng Kung University, Tainan, Taiwan (C-FS).

Address correspondence to Ching-Lin Hsieh, School of Occupational Therapy, College of Medicine, National Taiwan University, 4F, No. 17, Shiujou Rd, Taipei 100, Taiwan. E-mail: mike26@ha.mc.ntu.edu.tw.

Chien C-W, Lin J-H, Wang C-H, Hsueh I-P, Sheu C-F, Hsieh C-L. Developing a short form of the postural assessment scale for people with stroke. *Neurorehabil Neural Repair* 2007;21:81–90.

DOI: 10.1177/1545968306289297

Measuring balance is important for clinicians in diagnosing the severity of a stroke, selecting the most appropriate therapy, and evaluating treatment outcome for people with stroke.[1-3] A variety of functional scales measuring balance are commonly used with people with stroke; however, only a few are specifically designed for this use.[4] Benaim et al[4] developed a new scale, the Postural Assessment Scale for Stroke patients (PASS), that directly addressed the need for an assessment tool that specifically measures balance in people with stroke. Mao et al[5] further compared the reliability, validity, and responsiveness of the PASS with 2 balance scales (i.e., the Berg Balance Scale and the balance subscale of the Fugl-Meyer test) in people with stroke and found the PASS to have superior psychometric characteristics among the 3 balance measures.[5] Thus, the PASS demonstrated great potential for use in both clinical and research settings.

However, in addition to having sound psychometric properties, a clinical measure should ideally require minimal time to administer in a clinical setting.[6,7] Clinicians are aware of the negative effects of time-consuming assessment (e.g., the PASS may take about 10 min to complete[4]); therefore, reducing the number of items of the PASS (12 items) could produce positive consequences for both clinicians and, most important, the people with stroke. Currently, each item of the PASS has 4 levels of scaling, but there are 5 different criteria for scoring these 12 items.[4] Such an inconsistency in scoring criteria highlights the need for qualified administrators and may lead to difficulties for less-trained raters when making judgments about the patients' conditions. Furthermore, the extremely high internal consistency of the PASS (i.e., Cronbach's α was found to be as high as 0.95[4] and 0.94~0.96[5]) indicates possible redundancy among the items.[8] Some recent studies have demonstrated that an increase in the number of items or scoring levels does not necessarily improve the reliability, validity, and responsiveness of mobility measures,[9] balance measures,[5] and activities of daily living measures.[6,10] These observations motivate us to simplify the PASS.

The purpose of this study was to develop a short form of the PASS (SFPASS) that has psychometric properties

similar to those of the original PASS. We hypothesized that 5 to 7 items would be adequate based on the previous studies' finding[6,11] and that a validated simplified scaling level (i.e., a 3-level PASS[12]) could be employed. Therefore, several SFPASSs were proposed, and the psychometric properties of each SFPASS were compared with those of the original PASS in people with stroke.

METHODS

Subjects

In this study, 2 cohorts of people with stroke were used. The purpose of the study was to develop several SFPASSs, examine their psychometric properties, and determine the best SFPASS. A calibration and validation cohort was retrieved from a prospective study (the Quality of Life after Stroke Study in Taiwan), initiated on 1 December 1999.[5] Each subject within this group was evaluated at 14 days after stroke onset and reassessed at other specific time points (e.g., 30 and 90 days) subsequent to the onset of stroke to characterize his or her recovery of neurological impairments (e.g., as measured by the Fugl-Meyer motor test [FM]), balance ability (e.g., as measured by the standard version of the PASS), and functional abilities (e.g., as measured by the Barthel Index [BI]). These measures (i.e., the PASS, the FM, and the BI) were administered by an occupational therapist who was not informed of the purpose of this study.[5]

The other cohort of patients, who were used to cross-validate the psychometric properties of the best SFPASS, were recruited from 4 medical centers throughout Taiwan. Each subject in this cohort was evaluated twice: once at the initial stage of admission for rehabilitation and, second, prior to hospital discharge. The standard version of the PASS and BI were administered by trained occupational therapists and physical therapists in each center.

Subjects of both groups met the following criteria: 1) first or recurrent onset of cerebrovascular accident without other major diseases (e.g., cancer, dementia, severe rheumatoid arthritis); 2) ability to follow verbal instructions to complete the PASS; and 3) ability to provide informed consent personally or by proxy. Subjects were excluded if they had another stroke or other major disease(s) during the follow-up period.

Measures

The PASS was specifically developed to measure balance function in people with stroke.[4] The PASS contains 12 four-level items that gauge a person's balance performance in situations of varying difficulty, that is, maintaining or changing a lying, sitting, or standing position. Its total score ranges from 0 to 36, and the psychometric properties of the PASS were found to be satisfactory when used to assess people with stroke.[4,5]

The current study employed both the original PASS and a simplified PASS with a 3-level (PASS-3L[12]) in the development of several SFPASSs. To produce the PASS-3L, the data retrieved for this study were recoded as 0-1.5-3 by collapsing the 2 middle levels of the original 4-level scale (0-1-2-3). The PASS-3L was found to feature similar psychometric properties to those of the original PASS.[12]

The BI was developed to measure the severity of a disability.[13] The BI evaluates 10 basic activities of daily living items with a total possible score range of 0 to 100.[13] The BI has been previously shown to reveal good interrater reliability and high convergent validity for people with stroke.[10,14] The BI was used to examine the convergent validity and predictive validity of the SFPASS proposed in this study.

The FM[15] has been used to measure motor impairment following stroke. The FM consists of 50 items of upper and lower extremities in motor functioning. Each item is graded on a 3-level scale. The total possible score ranges from 0 to 100 points, and it has been found to yield good interrater reliability and high concurrent validity for people with stroke.[16-18] The FM was used to test the convergent validity of the SFPASS proposed in this study.

Data Analysis

A method proposed by Hobart and Thompson[6] has been shown to be useful for the development of a short form of the Barthel Index[6] and a short form of the Berg Balance Scale[11] for people with stroke. Their method was based on classical test theory and was used to develop and validate the SFPASS in this study because of the ease of application. Several SFPASSs were developed by conducting an item analysis and selecting those items with the best measurement properties (i.e., higher internal consistency and greater responsiveness) in the calibration group of people with stroke. Moreover, the psychometric properties (including the acceptability, reliability, validity, and responsiveness) of these SFPASSs were examined to determine the best SFPASS. Thus, this method would appear to be especially useful for developing such a measure for monitoring recovery after stroke and measuring outcome after treatment and therefore was adopted in this study. Additionally, the main psychometric properties of the best SFPASS were further compared with those of the original PASS in the cross-validation group.

Development of the SFPASS. To develop the SFPASS, we selected items with the highest internal consistency (i.e., minimizing measurement error) and greatest responsiveness (i.e., maximizing the ability to detect change).[6] The best items were those with the lowest values from an overall item index.[6] The overall item index[6] of each item is the cross-product of the 2 rank orders (i.e., the rank of corrected item total correlation [indicating internal consistency] and the rank of effect size [indicating responsiveness] for an item). The corrected item total correlation for an item is the correlation between the scores of an individual item and the sum scores of all the items on the scale minus that item. The ranking of the corrected item total correlation allows the removal of test items that show the lowest correlation to the overall construct measured in the PASS. Furthermore, the effect size for an item is the mean change score (14–30 days after stroke) divided by the standardized deviation of the scores at 14 days after stroke. The ranking of the effect size is useful in removing test items that show little sensitivity to change. Finally, the corrected item total correlation and effect size for each item were respectively ranked, and then the cross-products of these rank orderings were computed, that is, the overall item index of each item. For example, if one item's total correlation ranking is 2, and its effect size ranking is 3, then its overall item index is $2 \times 3 = 6$. Smaller values of the overall item index indicated superior items.

We hypothesized that the use of 5 to 7 of the best items would be adequate for the SFPASS. Three sets of SFPASSs were generated (i.e., the 5-item PASS, 6-item PASS, and 7-item PASS). In addition, the current study also employed a technique for simplifying the scale by collapsing the 2 levels in the center of the PASS into a single level. Thus, we developed another 3 sets of SFPASSs (i.e., the 5-item PASS-3L, 6-item PASS-3L, and 7-item PASS-3L). Therefore, a total of 6 SFPASSs were generated.

Psychometric evaluation of the 6 SFPASSs. To test the psychometric properties between the 6 SFPASSs, the scores were linearly transformed into the same score range as that for the original PASS (0–36). Four psychometric properties were tested in this study, these properties being acceptability, reliability, validity, and responsiveness.

Acceptability is a determination of whether the score distributions of a measure can match the distribution corresponding to the individuals intended to be measured.[6] A measure exhibiting good acceptability should reveal observable scores spanning the entire range of the scale, with a mean score near the scale midpoint, and featuring small floor and ceiling effects, that is, less than 15% of the subjects achieving the lowest or the highest scores.[6]

Reliability reflects the degree of precision of a measure; that is, high reliability requires a low rate of errors to be generated.[19] To estimate test reliability, the study by Hobart and Thompson[6] recommended the examination of the internal consistency of a specific measure, using Cronbach's α coefficients to determine the intercorrelations between the items.[6] It has been suggested that reliability estimations exceed 0.8 for group comparison studies and 0.95 for individual patient clinical decision making.[20] Confidence intervals for the α coefficients were computed.[21] Furthermore, confidence intervals for individual stroke patient scores were computed by calculating the standard error of measurement (SEM),[20] indicating an estimate of the dispersion of scores that would be obtained if a measure was administered to a patient multiple times.[22] The following 2 formulae were used: 1) SEM = (standard deviation of sample scores) $\times \sqrt{(1 - \text{reliability})}$; 2) 95% confidence intervals for individual scores = $\pm 1.96 \times$ SEM.

Validity indicates whether a measure actually determines what it has been constructed to determine.[19] We examined the agreement between the results of the 6 SFPASSs and the original PASS at 14 days poststroke by using a random effects model intraclass correlation coefficient (ICC) and the method proposed by Bland and Altman.[23] Bland and Altman's method involves plotting the scores of the difference between the original PASS and 6 SFPASSs against those of the average between the original PASS and 6 SFPASSs. Ideally, there should be no systematic trend (i.e., the mean score of the short form and the original scale is not associated with the difference between the short form and the original scale) in a Bland-Altman plot.[23] These results are useful in determining whether or not the 6 SFPASSs and the original PASS can be used interchangeably.

In addition, 3 validity indicators were examined for the comparisons of the 6 SFPASSs and the original PASS. First, concurrent validity at 14 days subsequent to stroke was examined by computing the intercorrelations between the scores of the 6 SFPASSs and the original PASS. Second, convergent validity between the scores of the 6 revised SFPASSs, the FM, and the BI at 14 days poststroke was also examined. Third, predictive validity of the 6 SFPASSs at 14 days subsequent to stroke was examined by correlating the scores of the 6 SFPASSs to those of the BI at 90 days following stroke.

Responsiveness reflects a measure's effectiveness at detecting changes in the longitudinal follow-up of the participants.[24,25] The degree of responsiveness in the 6 SFPASSs was investigated by calculating the effect sizes.[25,26] Effect sizes were determined by computing the mean of the total score difference between 14 and 30 days following stroke for each patient, divided by the standard deviation of the total score at 14 days subsequent to stroke. Larger values suggested greater responsiveness.

Table 1. Characteristics of the People With Stroke in This Study

	Calibration and Validation Group ($n = 287$)	Cross-Validation Group ($n = 179$)
Age (mean [*SD*])	65.5 (11.3)	63.7 (13.7)[a]
Sex (male/female)	174/113	133/46[b]
PASS (mean [*SD*])	19.0 (11.8)	18.1 (11.6)[a]
BI (mean [*SD*])	41.0 (30.0)	40.9 (30.5)[a]
FM (mean [*SD*])	52.6 (34.6)	—

PASS = the Postural Assessment Scale for Stroke patients; BI = Barthel Index; FM = Fugl-Meyer motor test.
a. No significant difference between these 2 groups (P value > 0.05, t test).
b. Significant difference between these 2 groups (P value < 0.05, chi-square test).

Cross-validation of the best SFPASS. Further comparison of the main psychometric properties of the best SFPASS included acceptability, internal consistency, and concurrent validity with the original PASS, as well as the responsiveness and predictive validity (as measured by the BI at discharge) in another independent cohort of people with stroke served to cross-validate the best SFPASS. In addition, the best SFPASS was applied to a small group of people with stroke to record its time of administration.

RESULTS

In the calibration group, 287 subjects were evaluated at 14 days subsequent to stroke, and a total of 262 and 218 subjects were successfully followed up at 30 and 90 days following stroke, respectively. In the cross-validation group, 179 subjects were assessed twice: at admission to rehabilitation and at hospital discharge. There was no significant difference between balance function in these 2 groups of people with stroke (P value < 0.05), and the 2 cohorts had similar PASS scores as well as similar BI scores, indicating that the clinical characteristics of both cohorts were similar (Table 1).

Development of the SFPASS

Table 2 shows that corrected item total correlations ranged from 0.38 to 0.93 and that the effect sizes ranged from 0.16 to 0.53, for individual items. According to the overall item index listed in Table 2, the 7-item PASS and 7-item PASS-3L were developed by including the 7 best items (in a top-down order): standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg, standing without support, and standing with support. The 6-item and 5-item PASS and PASS-3L were developed by sequentially removing, respectively, the worst 1 or 2 items from those best 7 items. Therefore, a total of 6 SFPASSs were developed.

Psychometric Evaluation of the 6 SFPASSs

Acceptability. All 6 SFPASSs exhibited good variability, as the test scores spanned the full possible ranges of the scales. Mean scores (17.1~18.1) were very close to the midpoint (18), and ceiling effects were small (≤7% of the subjects). The floor effects of the 6 SFPASSs, however, were notable, that is, about 20% of the subjects achieving the lowest scores (Table 3).

Reliability. Additionally, the 6 SFPASSs revealed very high α coefficients (≥0.95), but the lower limit confidence intervals (0.78~0.64) of all the 6 SFPASSs did not meet the criterion of 0.8 (Table 3). The SEM of the 6 SFPASSs ranged from 2.8 to 3.4, which was less than 3.6 (i.e., 10% of the highest possible score of 36, where such a score indicates clinical importance[27]).

Validity. The agreement between the original PASS and 6 SFPASSs was excellent (ICCs ≥ 0.96; Table 3). The limits of agreement between the 6 SFPASSs and the original PASS were similar and within the appropriate range. Nevertheless, Figures 1 and 2 reveal that all the short forms, except for the 5-item PASS-3L, demonstrated a systematic trend with $r^2 \geq 0.18$, indicating that only the 5-item PASS-3L among the short forms can be used interchangeably with the original PASS.

Table 4 shows that all of the 6 SFPASSs demonstrated very high concurrent validity with the original PASS ($r \geq 0.96$). Moreover, all 6 SFPASSs exhibited extremely similar and high convergent validity with the BI ($r = 0.86$~0.87) and with the FM ($r = 0.74$~0.75). The extent to which each of the 6 SFPASSs was able to predict the score of the BI at 90 days subsequent to the patient's stroke ($r = 0.46$~0.49) was also very similar to the original PASS ($r = 0.49$).

Responsiveness. Table 4 shows that the 6 SFPASSs and the original PASS revealed similar and satisfactory effect sizes. The 6 SFPASSs revealed slightly larger effect sizes (between 0.43 and 0.44) than the original PASS (effect size = 0.42).

In brief, we found that the 5-item PASS-3L highly agreed with the original PASS and that it was slightly superior to the other 5 SFPASSs, as there was no systematic trend between the difference and mean of the

Table 2. Item Analysis of the Postural Assessment Scale for Stroke Patients Scores at 14 Days After Onset (Calibration and Validation Group)

	Item Analysis		
Item[a]	Item Total Correlation[b] (RO[c]) $N = 287$	Effect Size[d] (RO) $N = 262$	Overall Item Index[e] (RO)
11 Standing up to sitting down	0.91 (2)	0.37 (3)	6 (1)
8 Supine to sitting up on the edge of the table	0.90 (3)	0.37 (3)	9 (2)
9 Sitting on the edge of the table to supine	0.50 (3)	0.37 (3)	9 (2)
10 Sitting to standing up	0.93 (1)	0.34 (9)	9 (2)
4 Standing on nonparetic leg (no other constraints)	0.55 (11)	0.53 (1)	11 (5)
3 Standing without support	0.88 (6)	0.42 (2)	12 (6)
2 Standing with support	0.90 (3)	0.36 (6)	18 (7)
12 Standing, picking up a pencil from the floor	0.84 (7)	0.36 (6)	42 (8)
1 Sitting without support	0.78 (9)	0.34 (9)	81 (9)
7 Supine to nonaffected side lateral	0.81 (8)	0.24 (11)	88 (10)
5 Standing on paretic leg (no other constraints)	0.38 (12)	0.35 (8)	96 (11)
6 Supine to affected side lateral	0.63 (10)	0.16 (12)	120 (12)

a. The 1st 7 items were selected for developing the short forms of the Postural Assessment Scale for Stroke patients.
b. Calculated as the correlations between the score of each item and the total score of the remaining 11 items.
c. Ranking order: 1 = highest value, 12 = lowest value.
d. Calculated as the mean change score (the score at 14 days after onset minus the score at 30 days after onset) divided by the standard deviation of the score at 14 days after onset.
e. Cross-product of rank order for item total correlation and rank order of effect size; for example, standing up to sitting down: $2 \times 3 = 6$.

Table 3. Comparison of Acceptability, Reliability, and Agreement of the 6 Short Forms of Postural Assessment Scale for Stroke Patients at 14 Days After Onset (Calibration and Validation Group, $N = 287$)

Psychometric Properties	12-Item PASS[a]	7-Item[b] PASS [PASS-3L][c]	6-Item[d] PASS [PASS-3L]	5-Item[e] PASS [PASS-3L]
Acceptability				
Mean score (*SD*)	19.0 (11.8)	18.1 (13.4) [17.9 (13.1)]	17.2 (13.2) [17.1 (12.9)]	17.5 (12.9) [17.4 (12.6)]
% Floor/Ceiling effect[f]	6.3/2.8	19.5/6.3 [19.5/6.3]	20.2/6.3 [20.2/6.3]	20.2/7.0 [20.2/7.0]
Reliability				
α (LL 95% CI[g])	0.96 (0.85)	0.96 (0.77) [0.95 (0.78)]	0.95 (0.73) [0.94 (0.74)]	0.93 (0.64) [0.93 (0.66)]
SEM[h]	2.4	2.8 [2.8]	3.1 [3.1]	3.4 [3.4]
95% confidence interval[i]	±4.7	±5.5 [±5.5]	±6.0 [±6.0]	±6.7 [±6.7]
Agreement with 12-item PASS				
ICC[j]	—	0.98 [0.98]	0.97 [0.96]	0.98 [0.97]
Mean difference (*SD*)[k]	—	1.0 (2.3) [1.1 (2.4)]	1.8 (2.5) [1.9 (2.8)]	1.5 (2.3) [1.6 (2.7)]
Limits of agreement[l]	—	–3.7 to 5.6 [–3.7 to 5.9]	–3 to 6.6 [–3.5 to 7.3]	–3 to 6 [–3.7 to 6.8]

a. Original 4-level Postural Assessment Scale for Stroke patients.
b. 7-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg (no other constraints), standing without support, and standing with support.
c. The data in brackets are the results of the PASS simplified in scale from 0-1-2-3 to 0-1.5-3 (PASS-3L).
d. 6-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg (no other constraints), and standing without support.
e. 5-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, and standing on nonparetic leg (no other constraints).
f. Percentage of sample scoring 0 (floor effect) and 36 (ceiling effect).
g. Lower limit 95% confidence interval calculated as (α- 1.96 SE), where SE = $\sqrt{(\text{SD rii})}/\sqrt{(\text{k}/2\,(\text{k-1})\text{-1})}$. SD rii = standard deviation of item inter-correlations; k = number of items in scale.
h. Standard error of measurement calculated as $\text{SD}\sqrt{(1-\alpha)}$.
i. Calculated as 1.96 × SEM.
j. Intraclass correlation coefficient (random-effects model).
k. Calculated as the score on the 12-item PASS at 14 days after stroke onset minus the transformed score of the Short Form of Postural Assessment Scale for Stroke patients at 14 days after stroke onset.
l. Mean difference ± 1.96 *SD*.

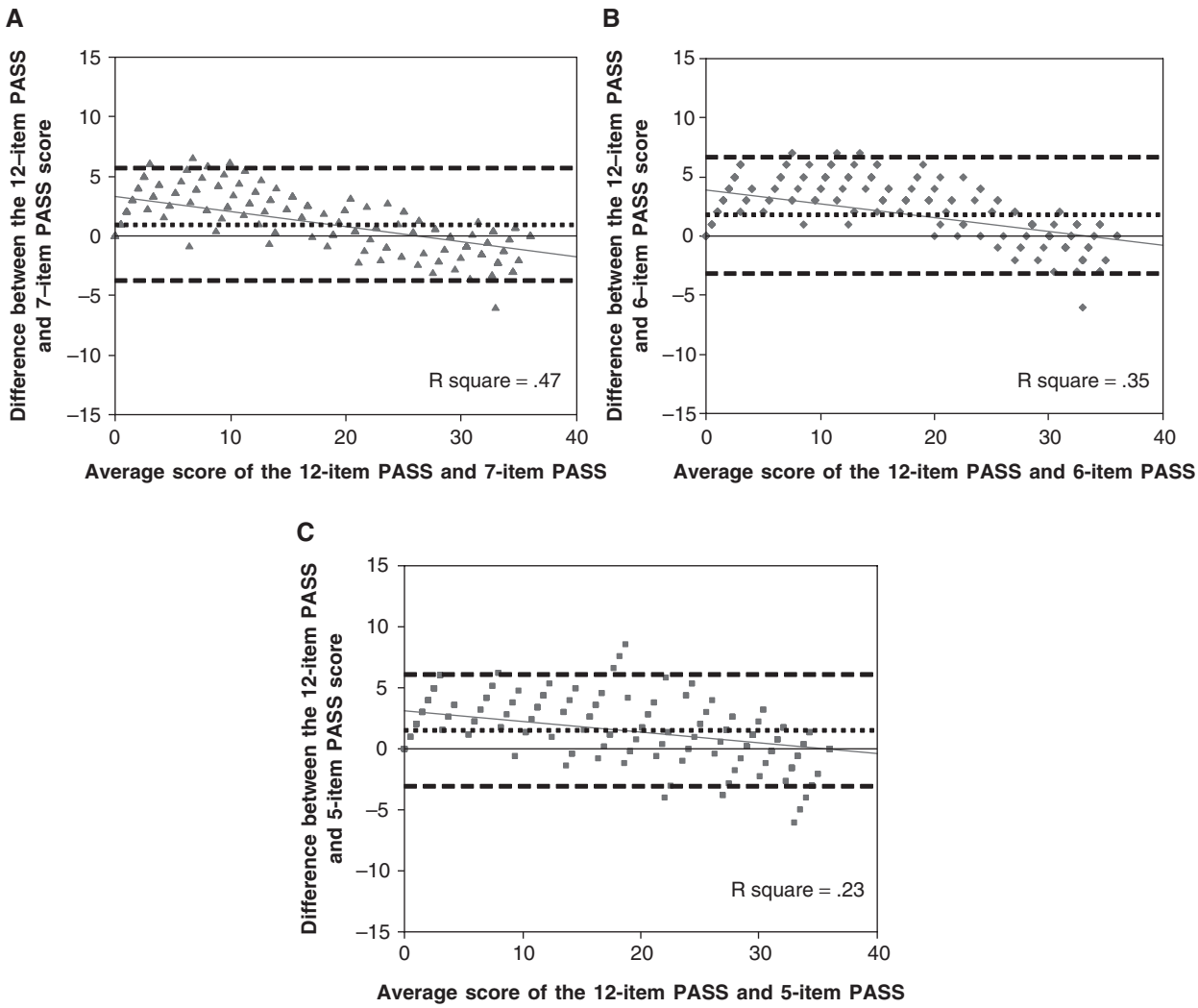

Figure 1. Bland-Altman method for plotting the difference of scores against mean scores of the original Postural Assessment Scale for Stroke patients (PASS) and 3 short forms of the PASS, including 7-, 6-, and 5-item PASS (A, B, and C, respectively). The 2 bold dashed lines define the limits of agreement (mean of difference ± 2 *SD*).

5-item PASS-3L and the original PASS in the Bland-Altman plots (Figures 1 and 2). The 5-item PASS-3L met the predefined psychometric criteria, with the exception of the floor effect (20.2%) and lower limit of the confidence interval for the α coefficient, which were also found in the other 5 SFPASSs. Furthermore, the 5-item PASS-3L is the simplest of the short forms. Therefore, the 5-item PASS-3L was determined to be the best SFPASS in the study.

Cross-Validation of the Best SFPASS

The findings of cross-validation also supported the requirement that the 5-item PASS-3L demonstrate satisfactory acceptability, internal consistency, concurrent validity, predictive validity, and responsiveness as compared with the original PASS (Table 5). Note that the floor effect of the 5-item PASS-3L was not obvious in this cohort (16.2% of the subjects).

In addition, the 5-item PASS-3L was applied to 10 people with stroke. All of the assessments were completed within 4 min.

DISCUSSION

A short, psychometrically sound measure offers clinicians a more efficient way to quantify patients' outcomes, given that it retains the reliability, validity, and responsiveness of the longer version. In this study, minimizing measurement error and maximizing the ability

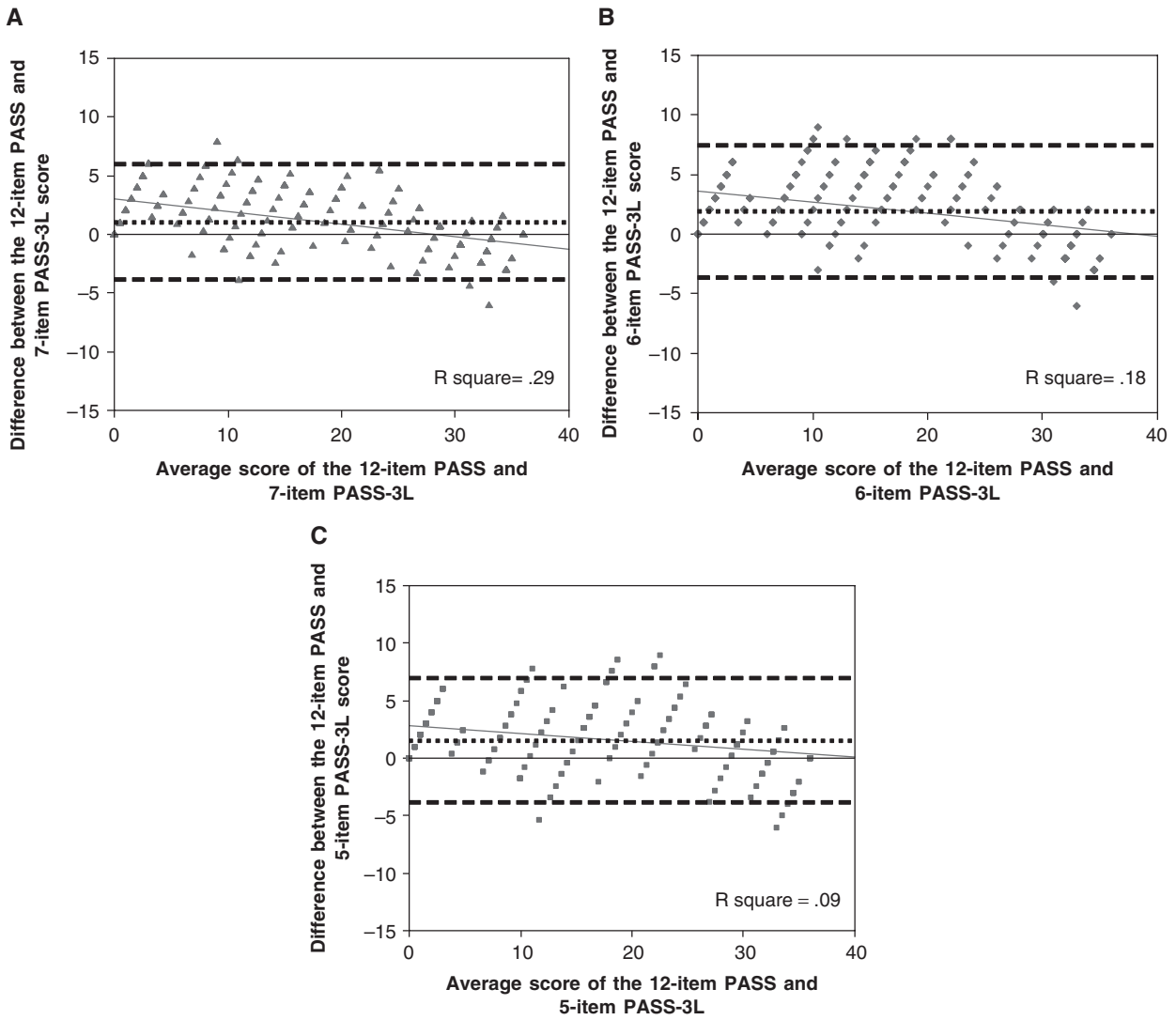

Figure 2. Bland-Altman method for plotting the difference of scores against mean scores of the original Postural Assessment Scale for Stroke patients (PASS) and 3 short forms of the PASS-3L, including 7-, 6-, and 5-item PASS-3L (A, B, and C, respectively). The 2 bold dashed lines define the limits of agreement (mean of difference ± 2 *SD*).

to detect change in the 12 items of the PASS directed the development of the 6 SFPASSs. We compared the psychometric properties of the 6 SFPASSs with the original PASS and found that the 5-item PASS-3L had the fewest items among the short forms. Psychometric properties of the 5-item PASS-3L were very similar to the original PASS. Furthermore, the psychometric properties of the 5-item PASS-3L were cross-validated and well supported in another sample. These results provided strong evidence that the 5-item PASS-3L was psychometrically similar (including internal consistency, concurrent validity, predictive validity, and responsiveness) to the original PASS for people with stroke. Results from the cross-validation testing suggested that we did not "overfit" the results of the 5-item PASS-3L to one single data set and that the findings of the study were well supported.

The 5-item PASS-3L improved in 3 significant aspects. First, the number of items is reduced from 12 items to only 5 items. Second, the number of scoring criteria is reduced from 5 to 2. Third, the scoring level is reduced from 4 to 3. Furthermore, in clinical or research settings, the 5-item PASS-3L can be administered in a flow sequence to further facilitate its administration, that is, beginning in "sitting to lying supine," then "changing supine to sitting up," "sitting to standing up," "standing on nonparetic leg," and "standing up to sitting down." More important, the contents of the 5-item PASS-3L are in concordance with the main concepts/contents of the original PASS, that is, gauging a person's balance

Table 4. Comparison of Validity and Responsiveness of the 6 Short Forms of Postural Assessment Scale for Stroke Patients (Calibration and Validation Group)

Psychometric Properties	12-Item PASS[a]	7-Item[b] PASS [PASS-3L][c]	6-Item[e] PASS [PASS-3L]	5-Item[f] PASS [PASS-3L]
Validity				
Concurrent validity ($N = 287$)				
12-item PASS (%V)[g]	—	0.99 (98) [0.99 (97)]	0.99 (97) [0.98 (96)]	0.99 (97) [0.98 (96)]
Convergent validity ($N = 287$)				
BI	0.87	0.86 [0.87]	0.87 [0.87]	0.86 [0.86]
FM	0.75	0.74 [0.75]	0.74 [0.75]	0.74 [0.75]
Predictive validity ($N = 218$)				
BI at 90 days after onset	0.49	0.46 [0.47]	0.48 [0.49]	0.46 [0.48]
Responsiveness ($N = 262$)				
Change score[h] (mean [*SD*])	4.9 (5.6)	5.8 (7.2) [5.7 (6.9)]	5.8 (7.1) [5.6 (6.8)]	5.6 (6.8) [5.4 (6.7)]
Effect size[i]	0.42	0.44 [0.44]	0.44 [0.44]	0.44 [0.43]

BI = Barthel Index; FM = Fugl-Meyer motor test.
a. Original 4-level Postural Assessment Scale for Stroke patients.
b. 7-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg (no other constraints), standing without support, standing with support.
c. The data in brackets are the result of the PASS simplified in scaling from 0-1-2-3 to 0-1.5-3 (PASS-3L).
d. 6-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg (no other constraints), standing without support.
e. 5-item: Standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, standing on nonparetic leg (no other constraints).
f. Product-moment correlations between the score at 14 days after onset.
g. Percent variance of 12-item PASS score explained.
h. The score at 30 days minus the score at 14 days. All change scores were statistically significant ($P < 0.01$).
i. Calculated as the mean change score divided by the standard deviation of the score at 14 days after stroke.

Table 5. Comparison of the Main Psychometric Properties of the Original Postural Assessment Scale for Stroke Patients and 5-Item PASS-3L (Cross-Validation Group, $N = 179$)

Psychometric Property	12-Item PASS	5-Item PASS-3L[a]
Mean score (*SD*)	18.1 (11.6)	17.5 (12.3)
% Floor/Ceiling effect[b]	6.1/1.7	16.2/8.4
Internal consistency, α[c]	0.96	0.93
Concurrent validity, ICC[d]	—	0.98
Predictive validity, r[e]	0.83	0.82
Responsiveness, effect size[f]	0.43	0.42

PASS = Postural Assessment Scale for Stroke patients.
a. 5-item PASS-3L simplified in scaling from 0-1-2-3 to 0-1.5-3. The 5 items are standing up to sitting down, supine to sitting up on the edge of the table, sitting on the edge of the table to supine, sitting to standing up, and standing on nonparetic leg (no other constraints).
b. Percentage of sample scoring 0 (floor effect) and 36 (ceiling effect).
c. Cronbach's α coefficients.
d. Intraclass correlation coefficient (random-effects model).
e. Product-moment correlations between the PASS score after admission to rehabilitation and the Barthel Index score before hospital discharge.
f. Calculated as mean change score divided by standard deviation of the score after admission to rehabilitation.

performance in situations of varying difficulty, that is, maintaining or changing from a lying, sitting, or standing position. That is, the 5-item PASS-3L not only retains clinically significant concepts but also contributes to efficiency in evaluation.

It is noted that some important aspects of the balance performance of individual patients (e.g., the abilities to roll in a lying position and pick up a pencil from the floor) are not recorded after the deletion of the items. Thus, the 5-item PASS-3L cannot entirely replace the original PASS in clinical settings, especially when the specific balance functions, originally measured by the items deleted from the original measure, are deemed to be treatment goals. Nevertheless, the high level of concurrence of the 5-item PASS-3L with the original PASS suggested that the 5-item PASS-3L and the original PASS can be used interchangeably or alternatively. For example, the original PASS can be used in the primary assessment (e.g., at admission or discharge) and the 5-item PASS-3L can be used for routinely monitoring patients' progress.

The 5-item PASS-3L was found to be psychometrically similar to the original PASS in this present study, except for the floor effect and the lower limit of the confidence interval for the reliability estimate found in the calibration group. The notable floor effect of the 5-item PASS-3L may have resulted from the removal of 3 lying and sitting items, which appeared to be the easiest tasks among the 12 original items. Removing these items from the original PASS could reduce the ability of the 5-item PASS-3L to detect changes in lying and sitting function and lead to a floor effect, which might dampen the responsiveness of the short form. However, the 5-item

PASS-3L still includes 2 lying and sitting items (i.e., "supine to sitting up" and "sitting to supine"), which could retain its ability to assess lying and sitting functions at a certain extent. Moreover, the reliability, validity, and in particular, responsiveness of the 5-item PASS-3L were very similar to those of the original PASS (Tables 3, 4, and 5). Thus, the floor effect may affect this measure's ability to discriminate some patients with severe stroke, but not worsen the psychometric merits of the 5-item PASS-3L, especially its use in detecting balance improvement. In addition, the reason why the lower limit of confidence intervals for each SFPASS's reliability estimates did not meet the set criterion may be the α coefficient, which is easily influenced by the number of items[8]; therefore, the lower limit of confidence interval for Cronbach's α coefficient is also related to the number of items. Fortunately, the SEM of the 5-item PASS-3L was within 10% of the highest possible score, indicating that the measurement error did not exceed clinical importance. The measurement precision of the 5-item PASS-3L is well supported.

The present study mainly employed the methodology presented in Hobart and Thompson's study for developing and cross-validating the 5-item PASS-3L. The 6 SFPASSs were developed by selecting items with the highest internal consistency and the greatest responsiveness. However, some psychometric indicators (e.g., the interrater reliability or test-retest reliability) could be taken into account in item-reduction criteria. Other item-reduction methods, such as the Item Response Theory modeling,[28,29] may also be used when developing a short form. Further studies may thus compare these different item-reduction methods with various kinds of item-reduction indicators on the development of multi-item measures.

In summary, our results provide strong evidence that the 5-item PASS-3L is psychometrically sound, except for the floor effect, and efficient to administer on patients who have had a stroke. The 5-item PASS-3L is thus suggested for use in people with stroke in both clinical and research settings.

ACKNOWLEDGMENTS

The study was supported by research grants from the National Science Council of Taiwan (NSC-93-2314-B-002-032 & NSC94-2314-B-002-078).

APPENDIX

5-Item PASS-3L Items and Criteria for Scoring

Note: Items 1 and 2 are to be performed with a 50-cm-high examination table, like a Bobath plane; item 3 is to be performed without any support; no other constraints.

Item	Scoring criteria
1. Sitting on the edge of the table to supine 2. Supine to sitting up on the edge of the table 3. Sitting to standing up 4. Standing up to sitting down	0 = cannot perform the activity 1.5* = can perform the activity with help 3 = can perform the activity without help
5. Standing on nonparetic leg (no other constraints)	0 = cannot stand on nonparetic leg for a few seconds 1.5* = can stand on nonparetic leg for a few seconds (but less than 10 seconds) 3 = can stand on nonparetic leg for more than 10 seconds

*The middle level of the 5-item PASS-3L is created by combining the middle 2 levels (1 and 2) of the original PASS-4L. PASS = Postural Assessment Scale for Stroke patients.

REFERENCES

1. Wade D. *Measurement in neurological rehabilitation*. Oxford (UK): Oxford University Press; 1992.
2. Bohannon RW, Leary KM. Standing balance and function over the course of acute rehabilitation. *Arch Phys Med Rehabil* 1995;76:994-6.
3. Duarte E, Marco E, Muniesa JM, et al. Trunk control test as a functional predictor in stroke patients. *J Rehabil Med* 2002;34:267-72.
4. 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;30:1862-8.
5. 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;33:1022-7.
6. Hobart JC, Thompson AJ. The five item Barthel Index. *J Neurol Neurosurg Psychiatry* 2001;71:225-30.
7. Talley NJ, Verlinden M, Jones M. Quality of life in functional dyspepsia: responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. *Aliment Pharmacol Ther* 2001;15:207-16.

8. Frisbie DA. Reliability of scores from teacher-made tests. *Educ Meas* 1988;7:25-35.
9. Rossier P, Wade DT. Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. *Arch Phys Med Rehabil* 2001;82:9-13.
10. 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;73:188-90.
11. 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.
12. Wang CH, Hsueh IP, Sheu CF, Yao G, Hsieh CL. Psychometric properties of 2 simplified 3-level balance scales used for patients with stroke. *Phys Ther* 2004;84:430-8.
13. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J* 1965;14:61-5.
14. Hsueh IP, Lee MM, Hsieh CL. Psychometric characteristics of the Barthel activities of daily living index in stroke patients. *J Formos Med Assoc* 2001;100:526-32.
15. Fugl-Meyer AR, Jaasko L, Norlin V. The post-stroke hemiplegic patient. II. Incidence, mortality, and vocational return in Goteborg, Sweden with a review of the literature. *Scand J Rehabil Med* 1975;7:73-83.
16. 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.
17. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys Ther* 1983;63:1606-10.
18. Lin JH, Hsueh IP, Sheu CF, Hsieh CL. Psychometric properties of the sensory scale of the Fugl-Meyer Assessment in stroke patients. *Clin Rehabil* 2004;18:391-7.
19. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. *J Neurol Neurosurg Psychiatry* 1996;60:127-30.
20. Nunnally J, Bernstein I. *Psychometric theory*. New York: McGraw-Hill; 1994.
21. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;52:281-302.
22. Guilford J. *Psychometric methods*. New York: McGraw-Hill; 1954.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
24. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171-8.
25. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985;38:27-36.
26. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27(Suppl 3):178-89.
27. Hebert R, Spiegelhalter DJ, Brayne C. Setting the minimal metrically detectable change on disability rating scales. *Arch Phys Med Rehabil* 1997;78:1305-8.
28. Lord FM, Novick MR. *Statistical theories of mental test scores*. Reading (MA): Addison-Wesley; 1968.
29. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press; 1960.