Analysis and Comparison of the Psychometric Properties of Three Balance Measures for Stroke Patients

Hui-Fen Mao, MS; I-Ping Hsueh, MA; Pei-Fang Tang, PhD; Ching-Fan Sheu, PhD; Ching-Lin Hsieh, PhD

Background and Purpose—This study compared the psychometric properties of 3 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 after onset.

Methods—One hundred twenty-three stroke patients were followed up prospectively with the 3 balance measures 14, 30, 90, and 180 days after stroke onset (DAS). Reliability (interrater 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 3 measures was made on the basis of the entire group of patients and 3 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, whereas 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 3 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 to 30 DAS.

Conclusions—All 3 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 2 measures.

Key Words: cerebrovascular disorders ■ reproducibility of results

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, selection of 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, >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 al17 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 3 basic psychometric properties: reliability, validity, and responsiveness.5,21,22 Although many researchers have examined the reliability and validity of each of the 3 balance measures described above,10,17–20,23–29 some limitations were noted. First, these studies did not compare the properties of different balance measures on the same cohort of patients. The measures in the previous studies were also rarely administered at specific times after onset. It is thus difficult for clinicians and researchers to compare the balance measures because of the limited results of previous studies. Second, most of the subjects in previous studies were tested only up to 3 months after stroke onset. Therefore, information is lacking on whether these balance measures are appropriate for assessing patients after that stage. Furthermore, no previously reported studies have evaluated whether these balance measures have similar psychometric properties for patients with different degrees of neurological severity. As a consequence, researchers and clinicians have found that they are faced with a greater range of choices but limited information on which to base their selection.

No reported studies have concurrently compared the psychometric properties of the 3 balance measures, the BBS,
FM-B, and PASS. The purpose of this prospective study was to compare the reliability, validity, and responsiveness of these 3 balance measures concurrently in a cohort of patients 14, 30, 90, and 180 days after stroke onset (DAS).

**Methods**

**Subjects**

Subjects were recruited from the registry of the Quality of Life After Stroke Study in Taiwan between December 1, 1999, and May 31, 2000. The Quality of Life After Stroke Study is an ongoing prospective cohort study of patients with stroke admitted to National Taiwan University Hospital. Individuals enrolled in the Quality of Life After Stroke Study were evaluated on 14 DAS and reassessed at other specific DAS times for up to 3 years after stroke to characterize recovery of neurological impairments, functional abilities, and health-related quality of life. Patients were included in the study if they met the following criteria: (1) diagnosis (International Classification of Diseases, Ninth revision, clinical modification codes) of cerebral hemorrhage (431), cerebral infarction (434), or other (430, 432, 433, 436, 437); (2) first onset of cerebrovascular accident without other major diseases; (3) stroke onset within 14 days before 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 or lived >40 miles from the hospital.

**Procedures**

The 3 balance measures and related measures were administered to patients 14, 30, 90, and 180 DAS. The protocol of this study was divided into 2 parts. The first part was an interrater reliability study. Only subjects at 14 DAS were used in this part of the study. The 3 balance measures were administered by 2 occupational therapists (A and B) individually on the same patients on 14 DAS. The therapists administered the assessments in a random order within a 24-hour period to minimize the effects of a possible spontaneous recovery. The therapists who administered these tests were blinded to each other's results.

The second part of the protocol was a validity and responsiveness study. Assessment of validity requires the use of standard instruments with which the scale is to be compared. In this study, the Barthel Index (BI) was used as the external criteria (for the examination of convergent validity) and was administered 14, 30, 90, and 180 DAS. The walking subscale of the Motor Assessment Scale (MAS) was used as an external criterion (for the examination of predictive validity) to evaluate the performance of ambulation on 180 DAS. The degrees of responsiveness of the 3 balance measures were calculated on the basis of changes occurring between 14 to 30, 30 to 90, 90 to 180, 14 to 90, and 14 to 180 DAS.

Because the testing protocol took ~30 to 40 minutes to administer, patients were allowed to rest during testing if necessary. All of the above assessments were made by an occupational therapist who was blinded to the purpose of this study. Data on demographic characteristics and comorbidity were collected from medical records.

**Instruments**

**Balance Measures**

The FM-B is 1 of 6 subscales of the FM, which was designed to evaluate impairment after stroke. The FM-B contains 7 three-point items, 3 for sitting and 4 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. Some studies found the sitting balance items, especially the 2 parachute reaction items, to be unreliable and invalid (item to total, \( r = -0.03 \)). Therefore, Hsueh et al revisited the scoring criteria of these 2 items but retained the original testing procedures. In this modified version, patients received a score of 0 if they lost balance easily, 1 if they partially lost balance, and 2 if they maintained sitting balance well when firmly pushed on the affected or nonaffected side. The validity of the modified FM-B was quite acceptable (\( r = 0.84 \)). The modified FM-B was used in this study.

The BBS evaluates a person's performance on 14 items (1 sitting and 13 standing items) related to balance function that are frequently encountered in everyday life. The scoring method is based on a 5-point ordinal scale of 0 to 4, with the total score ranging from 0 to 56. The BBS was originally developed for screening the elderly at risk of falling, but the psychometric properties of the BBS used in stroke patients have also been examined by various researchers with supportive results.

The PASS was reported to be able to detect changes of the same magnitude as the BI during the initial 12 weeks after stroke. The PASS was developed to be applicable to all stroke patients, even those with very poor postural performance. The PASS contains 12 four-point items that grade performance for situations of varying difficulty in maintaining or changing a given lying, sitting, or standing posture. Its total score ranges 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. The interrater reliability and intrarater reliability of the PASS have been shown to be very high. There are moderate correlations between scores of the PASS and the Functional Independent Measure, motor score of lower extremities, and postural stability. The PASS has fair predictive validity for later functional performance. However, no studies of the responsiveness of the PASS have been reported.

**Statistical Analysis**

**Distribution**

The score range and distribution of each of the 3 measures were examined. The floor and ceiling effects, the percentages of the sample scoring the minimum and maximum possible scores, respectively, reflect the extent that scores cluster at the bottom and top of the scale range. Floor and ceiling effects >20% are considered to be significant. The existence of the floor and ceiling effects is indicative of the limited ability of a measurement to discriminate between subjects.

**Reliability**

**Item Reliability**

The interrater agreement on individual items of the 3 balance measures was analyzed with 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. A \( \kappa \) value >0.75 indicates excellent agreement, 0.4 to 0.75 indicates fair to good agreement, and <0.4 indicates poor agreement.

**Total Score Reliability**

The intrarater reliability of the total score of the 3 balance measures was analyzed with the intraclass correlation coefficient (ICC) statistic. The fixed effect of ICC Model 3 was used to compute the ICC value for the degree of agreement between repeated measurements.
Responsiveness

Because there is no consensus regarding how best to assess the responsiveness of measurement instruments, 2 approaches were used 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 an ES > 0.8 is large, 0.5 to 0.8 is moderate, and 0.2 to 0.5 is small. In addition, by the 2 raters on the same patient. An ICC value of > 0.80 indicates high reliability.40

Internal Consistency

The internal consistency of each balance measure was expressed using Cronbach’s α coefficients. An α coefficient > 0.70 is considered adequate for group comparison.41

Validity

Concurrent Validity

Concurrent validity is usually established by demonstrating a high correlation between the scale and a gold standard. The interrelationship between the 3 balance measures on 14, 30, 90, and 180 DAS were examined with Spearman’s ρ correlation coefficient.

Convergent Validity

Convergent validity was determined by examining the relationships between the 3 balance measures and instruments measuring similar constructs. The relationships between the total score of the 3 balance measures and those of the BI at each DAS point were examined with Spearman’s ρ correlation coefficient.

Predictive Validity

The predictive validity of the 3 balance measures was assessed by comparing the results of the 3 balance measures at 14, 30, and 90 DAS with that of the MAS at 180 DAS by use of Spearman’s ρ correlation coefficient.

Responsiveness

Results

A total of 128 patients with a wide spectrum of balance deficits, ranging from asymptomatic patients to the bedridden, were originally recruited for the study. However, 5 patients declined to participate in the study; 13 patients were lost to follow-up at 30 DAS; and another 17 patients were lost to follow-up at 90 DAS. A total of 80 patients completed follow-up at 180 DAS. Table 1 presents detailed characteristics of the cohort of patients in the study.

Distribution

Table 2 shows the distributions of the 3 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 showed significant ceiling effects at 90 and 180 DAS.

Reliability

One hundred twelve patients participated in the reliability investigation. This study group consisted of 52 women and 60 men with a mean age of 69 years (SD, 11.3). Eleven patients were excluded from the analysis because they were not rated by both raters within 24 hours.

The medians (ranges) of weighted κ statistics for each item of the PASS, FM-B, and BBS were 0.88 (0.61 to 0.96), 0.79 (0.71 to 0.95), and 0.92 (0.59 to 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 to 0.98), 0.92 (0.88 to 0.95), and 0.95 (0.93 to 0.97), respectively, indicating excellent total score agreement. The Cronbach’s α 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.

TABLE 1. Characteristics of the Stroke Patients Participating in the Study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, M/F</td>
<td>66/57</td>
</tr>
<tr>
<td>Age (mean±SD), y</td>
<td>69.3±11.2</td>
</tr>
<tr>
<td>Diagnosis, n (%)</td>
<td></td>
</tr>
<tr>
<td>Cerebral hemorrhage</td>
<td>32 (26)</td>
</tr>
<tr>
<td>Cerebral infarction</td>
<td>74 (60)</td>
</tr>
<tr>
<td>Other</td>
<td>17 (14)</td>
</tr>
<tr>
<td>Side of hemiplegia, right/left</td>
<td>56/67</td>
</tr>
<tr>
<td>PASS at 14 DAS, mean±SD</td>
<td>17.6±12.8</td>
</tr>
<tr>
<td>FM-B at 14 DAS, mean±SD</td>
<td>6±4.8</td>
</tr>
<tr>
<td>BBS at 14 DAS, mean±SD</td>
<td>22.3±22.2</td>
</tr>
<tr>
<td>BI at 14 DAS, mean±SD</td>
<td>39±30.5</td>
</tr>
<tr>
<td>FM at 14 DAS, mean±SD</td>
<td>50.8±35.8</td>
</tr>
</tbody>
</table>

n=123.

TABLE 2. Distributions of the 3 Balance Measures at Different Stages of Stroke Recovery

<table>
<thead>
<tr>
<th></th>
<th>PASS, n (%)</th>
<th>FM-B, n (%)</th>
<th>BBS, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Floor Effect</td>
<td>Ceiling Effect</td>
<td>Floor Effect</td>
</tr>
<tr>
<td>14 DAS (n=123)</td>
<td>10 (8.1)</td>
<td>4 (3.3)</td>
<td>36 (29.3)</td>
</tr>
<tr>
<td>30 DAS (n=110)</td>
<td>5 (4.5)</td>
<td>11 (10)</td>
<td>16 (14.5)</td>
</tr>
<tr>
<td>90 DAS (n=93)</td>
<td>2 (2.2)</td>
<td>13 (14)</td>
<td>6 (6.5)</td>
</tr>
<tr>
<td>180 DAS (n=80)</td>
<td>3 (3.8)</td>
<td>14 (17.5)</td>
<td>5 (5)</td>
</tr>
</tbody>
</table>

Wilcoxon matched-pairs signed-rank tests were performed to determine the statistical significance of the change scores. Furthermore, to determine whether the responsiveness of the measures varied depending on the initial stroke-induced deficits, patients were stratified into 1 of the following 3 groups on the basis of their FM scores: 0 to 35, severe; 36 to 79, moderate; and ≥80, mild.43

Results

A total of 128 patients with a wide spectrum of balance deficits, ranging from asymptomatic patients to the bedridden, were originally recruited for the study. However, 5 patients declined to participate in the study; 13 patients were lost to follow-up at 30 DAS; and another 17 patients were lost to follow-up at 90 DAS. A total of 80 patients completed follow-up at 180 DAS. Table 1 presents detailed characteristics of the cohort of patients in the study.

Distribution

Table 2 shows the distributions of the 3 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 showed significant ceiling effects at 90 and 180 DAS.

Reliability

One hundred twelve patients participated in the reliability investigation. This study group consisted of 52 women and 60 men with a mean age of 69 years (SD, 11.3). Eleven patients were excluded from the analysis because they were not rated by both raters within 24 hours.

The medians (ranges) of weighted κ statistics for each item of the PASS, FM-B, and BBS were 0.88 (0.61 to 0.96), 0.79 (0.71 to 0.95), and 0.92 (0.59 to 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 to 0.98), 0.92 (0.88 to 0.95), and 0.95 (0.93 to 0.97), respectively, indicating excellent total score agreement. The Cronbach’s α 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.

TABLE 3. Concurrent Validity of the 3 Balance Measures at Different Stages of Stroke Recovery

<table>
<thead>
<tr>
<th>DAS</th>
<th>PASS vs BBS</th>
<th>PASS vs FM-B</th>
<th>BBS vs FM-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.95</td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td>30</td>
<td>0.95</td>
<td>0.96</td>
<td>0.90</td>
</tr>
<tr>
<td>90</td>
<td>0.92</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>180</td>
<td>0.93</td>
<td>0.95</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Downloaded from http://stroke.ahajournals.org/ by guest on April 7, 2017
Validity

Concurrent Validity
Table 3 shows the intercorrelations between the 3 balance measures. The pairwise correlations of the 3 balance measures were high at each stage of stroke, indicating high concurrent validity.

Convergent Validity
The scores of the 3 balance measures were highly correlated with those of the BI scores at the 4 selected time points after stroke (Spearman’s ρ ≥0.86, P<0.0001), indicating good convergent validity (Table 4).

Predictive Validity
The scores of the 3 balance measures at the earlier 3 DAS points were highly correlated with the MAS scores on evaluations on 180 DAS (Spearman’s ρ ≥0.8, P<0.0001), indicating good predictive validity (Table 4).

Responsiveness
The ES showed that the 3 balance measures were moderately to highly responsive in detecting changes before 90 DAS (14 to 30 DAS, ES ≥0.8; 30 to 90 DAS, ES ≥0.63) and that the levels of responsiveness of these measures were low, as expected, at 90 and 180 DAS (0.31≤ES≤0.4) (Table 5). Table 5 also shows that the changes in the 3 scales at each stage were all significant (P<0.006). Table 6 shows the responsiveness of the 3 balance measures at different stages for subjects with different levels of stroke severity (ES ≥0.21). All of these results indicate that the 3 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 PASS for severe stroke patients at 14 to 30 DAS (Table 6).

Discussion
In this study, the comprehensive psychometric properties of 3 balance measures (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 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) to evaluate how appropriate these measures are for use at different recovery stages after stroke. Furthermore, the responsiveness data 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 3 balance measures at 14 DAS revealed some limitations of these measures. The BBS and FM-B showed significant floor effects at 14 DAS, whereas the PASS did not. Almost one third of the subjects had the lowest scores on these 2 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 4 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 90 and 180 DAS. These results indicate that the BBS might not discriminate the patients’ balance function after 90 DAS.

### Table 4. Convergent Validity and Predictive Validity of the 3 Balance Measures at Different Stages of Stroke Recovery

<table>
<thead>
<tr>
<th>DAS</th>
<th>n</th>
<th>PASS vs BI</th>
<th>FM-B vs BI</th>
<th>BBS vs BI</th>
<th>PASS vs MAS*</th>
<th>FM-B vs MAS*</th>
<th>BBS vs MAS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>123</td>
<td>0.90</td>
<td>0.86</td>
<td>0.89</td>
<td>0.86</td>
<td>0.80</td>
<td>0.82</td>
</tr>
<tr>
<td>30</td>
<td>110</td>
<td>0.92</td>
<td>0.89</td>
<td>0.94</td>
<td>0.86</td>
<td>0.85</td>
<td>0.84</td>
</tr>
<tr>
<td>90</td>
<td>93</td>
<td>0.88</td>
<td>0.88</td>
<td>0.90</td>
<td>0.90</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td>180</td>
<td>80</td>
<td>0.92</td>
<td>0.89</td>
<td>0.91</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

*The MAS was administered at 180 DAS.

### Table 5. Responsiveness of the 3 Balance Measures at Different Stages of Stroke Recovery

<table>
<thead>
<tr>
<th>DAS</th>
<th>n</th>
<th>PASS</th>
<th>FM-B</th>
<th>BBS</th>
<th>PASS</th>
<th>FM-B</th>
<th>BBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>14–30</td>
<td>110</td>
<td>0.89</td>
<td>0.82</td>
<td>0.80</td>
<td>8.22*</td>
<td>7.16*</td>
<td>8.46*</td>
</tr>
<tr>
<td>30–90</td>
<td>93</td>
<td>0.64</td>
<td>0.63</td>
<td>0.69</td>
<td>6.05*</td>
<td>5.88*</td>
<td>6.7*</td>
</tr>
<tr>
<td>90–180</td>
<td>80</td>
<td>0.31</td>
<td>0.33</td>
<td>0.40</td>
<td>2.75†</td>
<td>2.74†</td>
<td>3.63‡</td>
</tr>
<tr>
<td>14–90</td>
<td>93</td>
<td>1.07</td>
<td>1.06</td>
<td>1.07</td>
<td>8.08*</td>
<td>7.95*</td>
<td>8.04*</td>
</tr>
<tr>
<td>14–180</td>
<td>80</td>
<td>1.12</td>
<td>1.14</td>
<td>1.11</td>
<td>7.65*</td>
<td>7.55*</td>
<td>7.67*</td>
</tr>
</tbody>
</table>

*P<0.00001; †P=0.006; ‡P=0.0003.
Table 6. Responsiveness (ES) of the 3 Balance Measures at Different Stages of Stroke Recovery for Subjects With Different Levels of Severity

<table>
<thead>
<tr>
<th>DAS</th>
<th>Mild Stroke (n=44)</th>
<th>Moderate Stroke (n=27)</th>
<th>Severe Stroke (n=52)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PASS</td>
<td>FM-B</td>
<td>BBS</td>
</tr>
<tr>
<td>14–30</td>
<td>0.60</td>
<td>0.53</td>
<td>0.75</td>
</tr>
<tr>
<td>30–90</td>
<td>0.29</td>
<td>0.42</td>
<td>0.21</td>
</tr>
<tr>
<td>90–180</td>
<td>0.34</td>
<td>0.33</td>
<td>0.37</td>
</tr>
<tr>
<td>14–90</td>
<td>0.76</td>
<td>0.78</td>
<td>0.81</td>
</tr>
<tr>
<td>14–180</td>
<td>0.85</td>
<td>0.86</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Reliability
In this study, the reliabilities of these 3 balance measures were examined in terms of interrater reliability and internal consistency. In agreement with many previous studies, our results indicated that both the item and total score interrater reliabilities and the internal consistency of these 3 measures were equally high. In particular, the individual item and total score reliabilities of the FM-B after adjustment for the scoring criteria of the sitting items were much higher than those found in a previous study. Thus, the reliability of FM-B with the modified sitting balance items was well supported. The very high internal consistency of the 3 balance measures indicated that the items of each instrument measured the same concept: balance. However, the extremely high internal consistency of the 3 measures might indicate the possibility of item redundancy, which needs further examination.

Validity
Results of concurrent, convergent, and predictive validity of all 3 measures were generally in accordance with the findings of previous studies. For example, previous studies found that BBS scores were correlated to motor performance, ADL function, and walking ability. The findings of this study further confirmed the validity of the 3 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 3 measures is scarce.

The results of the ES indicate that these 3 measures had fair to good levels of responsiveness before 90 DAS and in the overall stages (14 to 90 and 14 to 180 DAS) of recovery. In addition, at later stages (90 to 180 DAS) of recovery, the 3 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.

The other possible reason might be that these 3 balance measures lack items sensitive enough to detect patients’ improvement after 90 DAS.

Investigation of how disease severity affects the responsiveness of the 3 balance measures revealed that the BBS was less responsive than the PASS and FB-M in severe stroke patients at the initial stages (14 to 30 DAS). The reason for this finding might be that the BBS was not originally designed for stroke patients, and only 1 item of the scale assesses balance ability in the sitting position. Because sitting balance is 1 of the first postures to be restored after a stroke, it seems that the BBS is lacking items to detect change in patients who are unable to stand independently.

Level of Scaling and Number of Items
The FM-B, PASS, and BBS have 3, 4, and 5, scoring levels, respectively. The number of items in these 3 scales also varies. However, as found in this study, the reliability and responsiveness of these 3 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 3 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. The similar responsiveness of the 3 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 intrarater reliability of the 3 measures was not examined. Some studies have reported excellent intrarater reliability results for the BBS and PASS. In addition, we found a high interrater reliability of the measures. Therefore, the intrarater reliability of the 3 measures might not be an important issue.

In summary, the BBS, FM-B, and PASS are clinical balance measures with good reliability, good validity, and accepted responsiveness at different poststroke stages of recovery. The PASS showed slightly better psychometric characteristics than the other 2 measures and thus appears to be better suited for use by clinicians and researchers.

Acknowledgments
This study was supported by research grants from the National Taiwan University Hospital (89A015) and National Science Council (NSC-89-2314-B002-534 and NSC-89-2314-B002-468).
References

Analysis and Comparison of the Psychometric Properties of Three Balance Measures for Stroke Patients
Hui-Fen Mao, I-Ping Hsueh, Pei-Fang Tang, Ching-Fan Sheu and Ching-Lin Hsieh

Stroke. 2002;33:1022-1027
doi: 10.1161/01.STR.0000012516.63191.C5
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/33/4/1022