Relative and Absolute Reliability of a Vertical Numerical Pain Rating Scale Supplemented With a Faces Pain Scale After Stroke

Li-ling Chuang, Ching-yi Wu, Keh-chung Lin, Ching-ju Hsieh

**Background.** Pain is a serious adverse complication after stroke. The combination of a vertical numerical pain rating scale (NPRS) and a faces pain scale (FPS) has been advocated to measure pain after stroke.

**Objective.** This study was conducted to investigate whether an NPRS supplemented with an FPS (NPRS-FPS) would show good test-retest reliability in people with stroke. The relative and absolute reliability of the NPRS-FPS were examined.

**Design.** A test-retest design was used for this study.

**Methods.** Fifty people (>3 months after stroke) participating in an outpatient occupational therapy program were recruited through medical centers to rate current pain intensity twice, at a 1-week interval, with the NPRS-FPS (on a scale from 0 to 10). The relative reliability of the NPRS-FPS was analyzed with the intraclass correlation coefficient for determining the degree of consistency and agreement between 2 measures. The standard error of measurement, the smallest real difference, and Bland-Altman limits of agreement were the absolute reliability indexes used to quantify measurement errors and determine systematic biases of repeated measurements.

**Results.** The relative reliability of the NPRS-FPS was substantial (intraclass correlation coefficient = .82). The standard error of measurement and the smallest real difference at the 90% confidence interval of the NPRS-FPS were 0.81 and 1.87, respectively. The Bland-Altman analyses revealed no significant systematic bias between repeated measurements for the NPRS-FPS. The range of the limits of agreement for the NPRS-FPS was narrow (−2.50 to 1.90), indicating a high level of stability and little variation over time.

**Limitations.** The pain intensity of the participants ranged from no pain to a moderate level of pain.

**Conclusions.** These findings suggest that the NPRS-FPS is a reliable measure of pain in people with stroke, with good relative and absolute reliability.
Reliability of NPRS-FPS

Pain is a serious adverse complication after stroke and may negatively affect a patient’s rehabilitation process, psychological status, daily function, working capacity, and quality of life. The prevalence of pain after stroke varies from 19% to 74%, depending on inclusion criteria, pain definitions, follow-up interval, and measurement methods. Pain after stroke may be central (caused by thalamic lesions) or peripheral (from secondary impairments, abnormal muscle tone, or both). The upper extremity is the most common site of peripheral pain after stroke, and pain in the upper extremity is easier to specify than other types of pain after stroke. However, no scale has yet been developed specifically for pain after stroke. A patient’s self-report of pain is a reliable indicator for determining the presence and severity of pain. Because of speed and ease of administration and scoring, 4 self-report pain rating scales have been commonly applied to people with stroke: visual analog scale, numerical pain rating scale (NPRS), verbal rating scale, and faces pain scale (FPS). Assessment of pain after stroke must consider a patient’s ability to complete the scale correctly. Some people with right hemispheric stroke and hemispatial neglect may have a problem with a horizontal scale and tend to ignore numbers or faces located on their left side. Therefore, the visual analog scale is not recommended for people with stroke and cognitive or visuospatial impairments. However, the NPRS is more practical and understandable than the visual analog scale, is easy to administer and record, and does not require clear vision, dexterity, paper, or pen. Some people with left hemispheric stroke have speech disorders, may have comprehension deficiencies, and are not able to complete verbal assessments. The verbal rating scale is simple but lacks sensitivity in detecting changes in pain and is not suitable for people with severe aphasia. The NPRS has better sensitivity and provides a more accurate assessment of pain than the verbal rating scale. The NPRS has adequate reliability, validity, and responsiveness in people who are healthy; in elderly people with persistent pain; and in patients with chronic pain, shoulder pain, low back pain, postoperative pain, and cervical radiculopathy. However, the NPRS has not been specifically examined in people with stroke.

Although these scales are valid and reliable measures of pain intensity in the general population, only the vertical FPS has been found to be valid in the assessment of shoulder pain in people with stroke and reliable in people with left hemispheric stroke. The FPS is simple to administer and understand and does not require speaking, reading, or writing. It has established validity and reliability in adults; in elderly people with normal to moderately impaired cognitive functioning; and in people with illiteracy, chronic pain, acute burn, and shoulder pain after stroke.

Different versions of the FPS have been developed and 14 versions of the FPS were identified by a systematic review. The commonly used versions of the FPS have 7 faces or 6 faces that represent various degrees of pain. Among the various FPS instruments, the FPS, FPS-Revised, Oucher Pain Scale, and Wong-Baker FPS are adequately supported by psychometric data. The 6-face Wong-Baker FPS was preferred by parents and children and often used in older people. A major difference in these instruments is that the FPS and the FPS-Revised begin with a more neutral “no pain” face, and the Wong-Baker FPS begins with a smiling “no pain” face. The 6 cartoon-type faces of the Wong-Baker FPS range from a smiling “no pain” face at the lower anchor face to a crying “most pain” face at the upper anchor face. Previous FPS studies for assessing shoulder pain after stroke revealed that the vertical FPS, with 7 facial expressions, showed good correlations with the NPRS and the visual analog scale. Previous FPS studies for assessing shoulder pain after stroke revealed that the vertical FPS, with 7 facial expressions, showed good correlations with the NPRS and the visual analog scale.

The FPS and the NPRS were the preferred scales in Chinese adults after surgery, mature adults who were hospitalized, and older adults. It was recommended that the NPRS be supplemented with other measures in older people, and the combination of the NPRS and the Wong-Baker FPS was advocated for patients with acute burns. The evidence suggests that the 11-point NPRS is the most popular for patients to use to quantify their pain. The use of a vertical format for the scale was suggested to be more friendly, and the vertical NPRS is more sensitive and easier for patients to use. Therefore, a vertical scale incorporating the NPRS and the FPS theoretically could be used by all people with stroke as an a priori preferable measurement.

In the present study, the 6-face version of the Wong-Baker FPS was used to make it compatible with the NPRS in terms of scoring by use of a common metric (0–10). The aim of the present study was to determine the relative and absolute reliabilities of the vertical NPRS incorporated with the 6-face FPS (NPRS-FPS) for assessing pain in the affected arm in people with stroke.
Method
Participants
Fifty participants were selected from a sample of convenience referred to 3 medical centers according to the following inclusion criteria: enrollment in an outpatient occupational therapy program; first-ever stroke; subacute and chronic stroke (onset of ≥3 months); ability to follow study instructions and complete the scale (Mini-Mental State Examination scores of ≥22); and no participation in any experimental rehabilitation or drug studies during the study period. People who had stroke and no pain but who met the other criteria were recruited to ensure that the NPRS bottom anchor points were used in the study.

Because we expected that pain medications and intervention might complicate measurements of pain, exclusion criteria were as follows: irregular use of medications for pain or other pain-relieving treatment during the study period (and if people were on regular treatment, they had been receiving it for more than 1 week) and acute pain after surgery. Also excluded were people known to have physician-determined major medical problems or severe neuropsychological impairments that would interfere with participation (eg, global aphasia, severe attention deficits, and severe neglect).

Of the 50 participants, 29 had left-side brain lesions, and 21 had right-side brain lesions. The mean age of the participants was 52.63 years (SD = 11.09), and the mean time after stroke was 22.20 months (SD = 15.40).

Procedure
At outpatient rehabilitation, eligible patients who received routine occupational therapy focusing on functional training were invited by their rehabilitation physicians to participate in the study. Written informed consent was obtained from all of the participants. Before pain rating, participants were evaluated for the Brunnström stage and with the Fugl-Meyer Assessment to determine the motor recovery of the affected arm.

For determination of the test-retest reliability of the scale, all participants responded to the NPRS-FPS for the intensity of hemiplegic arm pain before therapy at 2 assessments, with a 1-week interval, at the same time of day to minimize any possible diurnal variation in pain. The 1-week interval was chosen to allow time to reduce the memory effect of the first assessment. Test and retest assessments were administered by the same research assistant. Participants received an average of 2 treatments (30 minutes per session) between the test and retest assessments. Of the participants with regular medication schedules, 3 (6% of all participants) received a muscle relaxant (tizanidine; Tonful, Yung Shin Pharmaceutical Industrial Co, Taichung, Taiwan), 2 (4%) had a prescription for chronic gout (allopurinol; Gouless, Standard Chemical and Pharmaceutical Co Ltd, Tainan, Taiwan), and 2 (4%) took medicine for heart disease (bokey; Tapal, Yung Shin Pharmaceutical Industrial Co).

Outcome Measures
Participants were provided a full explanation of the NPRS-FPS, which combined a vertical NPRS and the 6-face Wong-Baker FPS, and were instructed on how to complete the assessment. The NPRS-FPS was a combination of the vertical NPRS with word anchors on a scale from 0 to 10 and the 6 facial expressions of the FPS, facilitating scoring of the intensity of participants’ pain (Fig. 1).
Reliability of NPRS-FPS

The NPRS consisted of a 100-mm vertical line numbered with 0 as the bottom number to indicate “no pain” and 10 as the top number to indicate “worst possible pain” and containing small horizontal markers every 10 mm on the line. The 6 facial expressions of the FPS, suggesting various pain intensities, were used to supplement the vertical NPRS. The bottom face (a smiling face) was accompanied by the number 0 (“no pain”), and the top face (a sad, tearful face) was accompanied by the number 10 (“worst possible pain”).

Participants were asked to point only to a number, not a face, on the NPRS-FPS that best represented their current level of arm pain. The question posed to all participants rating their level of pain intensity was as follows: “How much pain do you feel today? Please point to a number that best reflects your current level of arm pain.”

Data Analysis

Statistical analyses were conducted with SPSS 16.0 software (SPSS Inc, Chicago, Illinois). The level of statistical significance was set at .05. Descriptive statistics are presented as the mean (standard deviation) or median (range) for continuous data and as numbers for discrete baseline characteristics. The relative and absolute reliabilities of the NPRS-FPS were estimated separately for participants with left-side brain lesions as well as for all participants with stroke. The relative reliability of the NPRS-FPS was determined through the calculation of a single measure, the mixed-model intraclass correlation coefficient (ICC); an ICC of greater than .75 indicated good reliability, and an ICC of less than .75 indicated poor to moderate reliability.

The absolute reliability of the NPRS-FPS was quantified through the standard error of measurement (SEM), the smallest real difference (SRD), and Bland-Altman analyses. The SEM represents the smallest change between 2 measures that indicates within-subject variability and a real change for a group. The SEM was calculated as follows:

\[ \text{SEM} = \text{SD}_{\text{pooled}} \times \sqrt{(1 - \text{ICC})} \]

In this equation, \( \text{SD}_{\text{pooled}} \) is the standard deviation for all observations from test sessions 1 and 2, and ICC is the test-retest reliability coefficient. The SRD represents the magnitude of change necessary to exceed the measurement error of 2 measures that indicates a true statistical change at a certain confidence interval (CI) for a single individual. The SRD90, which was used to determine whether the change score for an individual was real at the 90% CI, was calculated as follows:

\[ \text{SRD}_{90} = 1.65 \times \sqrt{2} \times \text{SEM} = 1.65 \times \sqrt{2} \times \text{SD}_{\text{pooled}} \times \sqrt{(1 - \text{ICC})} \]

In this equation, 1.65 is the 2-tailed tabled \( z \) value for the 90% CI, and \( \sqrt{2} \) represents the variance of 2 measures. Generally, a difference between 2 measurements that is larger than the SEM and the SRD can be attributed to a real change or one beyond measurement error at a specified confidence level. The smaller the SEM and the SRD90, the greater the reliability.

The Bland-Altman analyses and plots were established by a series of statistical procedures to indicate repeatability. The plots allowed us to visualize any possible relationship between the measurement error and the true values, systematic bias, and random error by examining the direction and magnitude of the scatter around the 0 line. The mean of all measurements, the mean and standard deviation of the differences (SDdiff) between the measurements, and the 95% CI of the mean difference were calculated. The 95% CI of the mean difference, which was used to determine the presence of systematic bias, was calculated as follows:

\[ 95\% \text{ CI of mean difference} = \text{mean difference} \pm 1.96 \times \text{SE} \]

\[ = \text{mean difference} \pm 1.96 \times \frac{\text{SD}_{\text{diff}}}{\sqrt{n}} \]

In this equation, SE is the standard error and \( n \) is the sample size. If 0 is included within the 95% CI, then no significant systematic bias between measurements can be inferred. The 95% limits of agreement (LOA)—that is, the mean difference \( \pm 1.96 \times \text{SD}_{\text{diff}} \)—were used to examine the natural variation over time; a narrow LOA indicated a higher level of stability. Each participant was represented in the Bland-Altman plots by assigning the mean of the 2 measurements as the abscissa (x-axis) value and the difference between the 2 values as the ordinate (y-axis) value.

Role of the Funding Source

This project was partly supported by the National Health Research Institutes (NHRI-EX101-9920PI and NHRI-EX101-10010PI), the National Science Council (NSC-102-2314-B-182-003, NSC-100-2314-B-002-008-MY3, and NSC-99-2314-B-182-014-MY3), and the Healthy Ageing Research Center at Chang Gung University (EMRPD1B0371) in Taoyuan, Taiwan.

Results

The detailed characteristics of the participants and the descriptive statistics for the NPRS-FPS in the 2 test sessions are shown in Table 1. There was no significant difference in pain intensity between the test and retest assessments for participants with right hemispheric stroke and partic-
Table 1.
Characteristics of the Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Participants With Stroke (N=50)</th>
<th>Participants With Left Hemispheric Stroke (n=29)</th>
<th>Participants With Right Hemispheric Stroke (n=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>36</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>Women</td>
<td>14</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Age, y, X (SD)</td>
<td>52.63 (11.09)</td>
<td>48.65 (10.42)</td>
<td>56.71 (10.48)</td>
</tr>
<tr>
<td>Side of brain lesion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>21</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Left</td>
<td>29</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>Mo after stroke onset, X (SD)</td>
<td>22.20 (15.40)</td>
<td>24.72 (17.01)</td>
<td>18.71 (12.40)</td>
</tr>
<tr>
<td>Aphasia</td>
<td>14</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Hemispatial neglect</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Brunnstrom stage of upper limb, median (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal part</td>
<td>4 (1–6)</td>
<td>4 (1–6)</td>
<td>4 (1–5)</td>
</tr>
<tr>
<td>Distal part</td>
<td>3 (1–6)</td>
<td>4 (1–6)</td>
<td>3 (1–6)</td>
</tr>
<tr>
<td>First assessment of pain intensity with NPRS-FPS, X (SD)</td>
<td>2.12 (1.90)</td>
<td>2.03 (1.74)</td>
<td>2.24 (2.13)</td>
</tr>
<tr>
<td>Participants with severe pain of 7–10</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Participants with moderate pain of 4–6</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Participants with mild pain of 1–3</td>
<td>24</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Participants with no pain</td>
<td>14</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Second assessment of pain intensity with NPRS-FPS, X (SD)</td>
<td>1.82 (1.95)</td>
<td>1.74 (1.92)</td>
<td>1.93 (2.03)</td>
</tr>
<tr>
<td>Participants with severe pain of 7–10</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Participants with moderate pain of 4–6</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Participants with mild pain of 1–3</td>
<td>22</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Participants with no pain</td>
<td>17</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Fugl-Meyer assessment of upper limb, X (SD)</td>
<td>34.50 (14.71)</td>
<td>38.9 (13.25)</td>
<td>28.43 (14.75)</td>
</tr>
<tr>
<td>Mini-Mental State Examination score, X (SD)</td>
<td>27.64 (2.36)</td>
<td>27.79 (1.86)</td>
<td>27.43 (2.96)</td>
</tr>
</tbody>
</table>

*Data are reported as number of participants, unless otherwise indicated. NPRS-FPS—numerical pain rating scale supplemented with faces pain scale.

Table 2.
Relative and Absolute Reliability of a Numerical Pain Rating Scale Supplemented With a Faces Pain Scale

<table>
<thead>
<tr>
<th>Participants</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>SRD</th>
<th>d</th>
<th>SD_{diff}</th>
<th>SE of d</th>
<th>95% CI of d</th>
<th>95% CI of d</th>
<th>LOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>.82 (.70–.90)</td>
<td>0.81</td>
<td>1.87</td>
<td>−0.30</td>
<td>1.12</td>
<td>0.16</td>
<td>−0.61 to 0.01</td>
<td>−2.50 to 1.90</td>
<td></td>
</tr>
<tr>
<td>Those with LHSP</td>
<td>.75 (.53–.87)</td>
<td>0.91</td>
<td>2.12</td>
<td>−0.29</td>
<td>1.29</td>
<td>0.24</td>
<td>−0.76 to 0.18</td>
<td>−2.82 to 2.24</td>
<td></td>
</tr>
<tr>
<td>Those with RHSP</td>
<td>.91 (.78–.96)</td>
<td>0.62</td>
<td>1.44</td>
<td>−0.31</td>
<td>0.87</td>
<td>0.19</td>
<td>−0.68 to 0.06</td>
<td>−2.02 to 1.40</td>
<td></td>
</tr>
</tbody>
</table>

*ICC—intraclass correlation coefficient, CI—confidence interval, SEM—standard error of measurement, SRD—smallest real difference, d—mean difference between the 2 test sessions (test session 2 minus test session 1), SD_{diff}—standard deviation of mean difference, LOA—limits of agreement (d±1.96×SD_{diff}), LHSP—left hemispheric stroke, RHSP—right hemispheric stroke.
Participants with left hemispheric stroke ($P > .05$).

**Relative Reliability**

Table 2 summarizes the results of our test-retest analyses of the NPRS-FPS. The ICCs for the NPRS-FPS in participants with right hemispheric stroke and participants with left hemispheric stroke were .91 and .75, respectively. For all participants with stroke, the ICC for the NPRS-FPS was .82, indicating that the NPRS-FPS had good relative reliability in participants with right or left hemispheric stroke.

**Absolute Reliability**

The SEM and SRD$_{90}$ of the NPRS-FPS were, respectively, 0.81 and 1.87 in all participants (Tab. 2), 0.62 and 1.44 in participants with right-side lesions, and 0.91 and 2.12 in participants with left-side lesions. The Bland-Altman statistics for the 2 test sessions with the NPRS-FPS are shown in Table 2. The mean difference between the testing sessions with the NPRS-FPS was close to 0 in participants with right ($-0.31$) or left ($-0.29$) hemispheric stroke. The 95% CI of the mean difference included 0, indicating that there was no significant systematic bias between the repeated measurements in participants with right-side or left-side brain lesions.

Three Bland-Altman plots (Fig. 2) that were representative of the NPRS-FPS showed the variability between the test and retest measurements in all participants with stroke and in participants with right or left hemispheric stroke. The graphs plot the differences between measurements from the 2 testing sessions for individual participants against the respective individual means of the test and retest measurements. The repeatability for most of the test and retest measurements was within the 95% CI. We noted 3 outliers for the NPRS-FPS in participants with left-
side lesions. The ranges of the LOA for the NPRS-FPS were −2.50 to 1.90 in all participants with stroke, −2.02 to 1.40 in participants with right hemispheric stroke, and −2.82 to 2.24 in participants with left hemispheric stroke.

Discussion

In the present study, we used standard methods to assess relative and absolute reliability in order to document the test-retest reliability of the NPRS-FPS in people with right or left hemispheric stroke. The NPRS-FPS showed good relative and absolute reliabilities, with high agreement, small measurement error, and no systematic bias for the assessment of pain after stroke.

Relative Reliability

Establishing the reliability of a tool for the adequate assessment of pain is an essential prerequisite before the tool is adopted as a standard measure of pain severity in patients with stroke. Test-retest reliability is the ability of an outcome measure to capture similar scores on 2 separate occasions of test administration, given that the patient’s condition has not changed.42 The ICC represents the degree of consistency between 2 measurements.50 The ICCs in all participants with stroke (.82) and in those with right (.91) or left (.75) hemispheric stroke indicated a high degree of agreement between test and retest measures and good reproducibility of the vertical NPRS-FPS. Our data were similar to values reported in previous studies, which indicated good reliability of the NPRS in Chinese adults after surgery (ICC=.82),25 elderly patients with persistent pain (ICCs=.75−.85),21 patients with chronic nociceptive or neurogenic pain (rank order agreement coefficient=.91),22 patients with shoulder pain (ICC=.74),23 patients with mechanical neck pain (ICC=.76),26 and acute pain in an emergency department (ICC=.88).31 Benaim et al13 used the FPS to assess shoulder pain after stroke and found that it was more reliable in people with left-side lesions (k coefficient=.74) than in those with right-side lesions (k coefficient=.53). In that study, no person with a right-side brain lesion and 42% of those with left-side lesions correctly ranked the 7 faces of the FPS from “least painful” to “most painful.” These data indicated that the 7 faces of the FPS might not discriminate different levels of pain depicted in the pictures.13

In the present study, we found that the relative reliability of the vertical NPRS-FPS appeared to be better in participants with right-side brain lesions than in those with left-side brain lesions (ICCs=.91 and .75, respectively). These results may have occurred because the expressions on the Wong-Baker FPS cartoon faces differed significantly from each other and may have been recognized by participants with stroke more easily than the neutral faces of the FPS. Another possible reason is that unilateral spatial neglect was present in 59% of the people with right hemispheric stroke in the study of Benaim et al,13 whereas none had this problem in the present study. The fact that no significant difference between participants with left hemispheric lesions and those with right hemispheric lesions was found for the NPRS-FPS at test and retest adds some validity to combining the NPRS and the FPS to assess pain after stroke. Our results indicated that the combination of NPRS and FPS was reliable in assessing pain in people with right or left hemispheric stroke.

Absolute Reliability

Determination of the absolute reliability of measures is critical to ensure repeated measurements with satisfactory stability and sensitivity to real changes over time.45 Reliable outcome measures demonstrate small measurement errors for a group of patients and small true changes for an individual patient. The SEM and SRD90 of the measures provide the absolute values of the measurement errors between repeated tests and determine whether changes in repeated measurements for a group of patients and for a single patient, respectively, are real.52,53 In the present study, the SEM and SRD90 of the NPRS-FPS for all participants were 0.81 and 1.87, respectively; that is, if the change in repeated measurements with the NPRS-FPS for a participant with stroke was more than 1.87, then the change exceeded the measurement error and was real at the 90% CI. The SEM and SRD90 of the NPRS-FPS were smaller than those in a previous study of NPRS (SEM=1.07, minimal detectable change=2.5).23 Possible reasons are different patient populations, study designs, and outcome measures.

The Bland-Altman statistics for the NPRS-FPS in all participants with stroke and in participants with right or left hemispheric stroke indicated no significant systematic bias between the repeated measurements because the mean differences between the 2 testing sessions were all close to 0 and the 95% CI of the mean differences included 0. From the Bland-Altman plots, the narrow range of the LOA and the 3 outliers in the NPRS-FPS indicated a high level of stability and little natural variation over time, perhaps because pain is easy to identify as a strong harmful feeling. The range of the LOA was slightly narrower (−2.82 to 2.24 versus −2.02 to 1.40) and the number of outliers was smaller (3 versus 0) in participants with right hemispheric stroke than in participants with left hemispheric stroke.

Patients with pain may have problems completing long questionnaires. The important aspects of selecting a pain scale for patients with stroke are that it should be short and easily understandable, place a low burden on patients, and
Reliability of NPRS-FPS

be reliable when replicated. The reliability results of the present study showed that a simple approach of combining the NPRS with the FPS created a reliable tool for assessing pain symptoms. The advantages of the NPRS-FPS are speed and ease of administration and scoring, similar to the strengths of the NPRS—reliability, known characteristics, time efficiency, ease of use, and not being an excessive burden on patients. In the absence of fully validated gold standards, the NPRS-FPS, with good test-retest reliability, could be used to monitor pain, facilitate faster communication between patients and clinicians regarding pain experience and response to treatment, and allow for future comparability across different studies.

Limitations

The main limitation of the present study was that it examined only a small number of participants with no pain to a moderate level of pain—not representative of the general population of people with stroke. In addition, the results cannot be generalized to people with stroke with cognitive-perceptual deficits due to the exclusion criteria. Thus, the findings should be interpreted with caution and are most generalizable to people without severe pain. Additional research with larger samples representing people with various levels of pain severity is needed to more fully examine the relative and absolute reliabilities of combined scales as clinically useful tools.

Another related argument is that a measure of pain assessed from a single measure of pain intensity may be at risk of being unreliable and invalid. The day-to-day fluctuations of and environmental contributions to pain intensity may lower the test-retest reliability of a single pain intensity rating in small groups of patients. Because chronic pain usually varies throughout the day and night, an assessment of pain should aggregate different time periods by averaging multiple measurements obtained at different times during a single day or across different days. All participants in the present study completed the NPRS-FPS at 2 assessments, with a 1-week interval, at the same time of day to minimize any diurnal variation in pain. Future studies could use multiple measures to control for the effects of environments and times to improve the psychometric properties of an assessment. Composite measures of pain intensity, consisting of 3 ratings per day across 4 days, have been recommended as being more valid and reliable in patients with chronic pain.

Another limitation is related to the design of the facial expressions on the FPS with respect to whether a smiling face and a tearful face were presented. Children’s pain ratings appear to be influenced by a smiling face, tending to be higher than those generated from face scales with a neutral face. Because the actual choice of the scale used clinically may be affected by the attractiveness of the scale, the Wong-Baker FPS, with its cartoon-like facial illustrations, was chosen to discriminate different levels of pain depicted in pictures because of its potential usefulness in a clinical setting with older adults. The use of smiling and tearful faces may present no problem for adults, who understand the underlying dimension from no pain to severe pain.

Although this research investigated the test-retest reliability of the NPRS-FPS in people with stroke, sensitivity to change and response to specific interventions have yet to be determined. At the beginning of the present study, participants were asked to maintain routine outpatient occupational therapy, consisting of an average of 2 treatments (30 minutes per session) between the test and retest assessments. Treatments were provided during the study period, but the treatments were not likely to have had substantial effects on the ultimate outcome. The sensitivity of the NPRS-FPS to change with disease progression or treatment needs to be proven if the NPRS-FPS is to be used as an outcome measure in a clinical trial. Future studies are needed to address the responsiveness and minimal clinically important changes observed for the NPRS-FPS in patients with stroke.

Conclusion

The present study has demonstrated empiric evidence for good test-retest reliability of the NPRS-FPS for documenting pain intensity in the clinical setting for patients with stroke. The NPRS-FPS is a reliable instrument that can provide the basis for assessing pain experience in clinical communications between people with stroke and health care providers. The use of a reliable patient-rated outcome measure across pain studies would enhance the comparability and clinical applicability of such studies.

All authors provided concept/idea and research design. Dr Chuang and Dr Lin provided writing. Dr Wu provided data collection. Dr Chuang provided data analysis. Dr Chuang, Dr Wu, and Dr Lin provided project management. Dr Wu and Dr Lin provided fund procurement. Dr Lin provided facilities/equipment. Dr Chuang and Dr Wu provided institutional liaisons. Dr Hsieh provided clerical support. Dr Chuang, Dr Wu, and Dr Hsieh provided consultation (including review of manuscript before submission). Dr Chuang and Dr Wu contributed equally to this article.

The institutional review boards of the participating sites approved this study.

This project was partly supported by the National Health Research Institutes (NHRI-EX101-9920PI and NHRI-EX101-10010PI), the National Science Council (NSC-102-2314-B-182-003, NSC-100-2314-B-002-008-MY3, and NSC-99-2314-B-182-014-MY3), and the Healthy Ageing Research Center at...
Reliability of NPRS-FPS

Chang Gung University (EMRPD1B0371) in Taoyuan, Taiwan.


References


44. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsive- ness comparisons using individual threshold criteria. *J Clin Epidemiol.* 2004;57:1008–1018.


January 2014

Volume 94 Number 1 Physical Therapy 137
Reliability of NPRS-FPS


