This is the published version of a paper published in *Physiotherapy Theory and Practice*.

Citation for the original published paper (version of record):

https://doi.org/10.1080/09593985.2018.1482511

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

© 2018 The Author(s). Published with license by Taylor & Francis. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-ncnd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-68074
Validity and test–retest reliability of the six-spot step test in persons after stroke

Mialinn Arvidsson Lindvall, Agneta Anderzén-Carlsson, Peter Appelros & Anette Forsberg

To cite this article: Mialinn Arvidsson Lindvall, Agneta Anderzén-Carlsson, Peter Appelros & Anette Forsberg (2020) Validity and test–retest reliability of the six-spot step test in persons after stroke, Physiotherapy Theory and Practice, 36:1, 211-218, DOI: 10.1080/09593985.2018.1482511

To link to this article: https://doi.org/10.1080/09593985.2018.1482511

© 2018 The Author(s). Published with license by Taylor & Francis.

Published online: 06 Jun 2018.

Article views: 652

View related articles

View Crossmark data

Citing articles: 2 View citing articles
Validity and test–retest reliability of the six-spot step test in persons after stroke

Mialinn Arvidsson Lindvall, RPT, MSc^a, Agneta Anderzén-Carlsson, RN^a,b, Peter Appelros, MD^a, and Anette Forsberg, RPT^c

^aDepartment of University Health Care Research Centre, Faculty of Medicine and Health, Örebro University, Örebro, Sweden; ^bFaculty of Health, Science, and Technology, Department of Health Sciences, Nursing, Karlstad University, Karlstad, Sweden; ^cDepartment of Physiotherapy, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

ABSTRACT
Background and Purpose: After stroke, asymmetric weight distribution is common with decreased balance control in standing and walking. The six-spot step test (SSST) includes a 5-m walk during which one leg shoves wooden blocks out of circles marked on the floor, thus assessing the ability to take load on each leg. The aim of the present study was to investigate the convergent and discriminant validity and test–retest reliability of the SSST in persons with stroke. Methods: Eighty-one participants were included. A cross-sectional study was performed, in which the SSST was conducted twice, 3–7 days apart. Validity was investigated using measures of dynamic balance and walking. Reliability was assessed using intraclass correlation coefficient, standard error of the measurement (SEM), and smallest real difference (SRD). Results: The convergent validity was strong to moderate, and the test–retest reliability was good. The SEM% was 14.7%, and the SRD% was 40.8% based on the mean of four walks shoving twice with the paretic and twice with the non-paretic leg. Conclusion: Values on random measurement error were high affecting the use of the SSST for follow-up evaluations but the SSST can be a complementary measure of gait and balance.

Introduction
Balance control is a complex function that coordinates the body’s center of gravity relative to the support surface and can be defined as the ability to maintain, achieve, or restore postural stability (Shumway-Cook and Woollacott, 2012). After a stroke, hemiparesis with muscle weakness, spasticity, and impaired sensory function may result in an asymmetric weight distribution in standing and walking, with more loads on the non-paretic limb (Langhorne, Bernhardt, and Kwakkel, 2011). Balance control may also be affected by impaired vision and cognition. A consequence of an asymmetric walking pattern and impaired balance control is high risk for falls (Hendrickson et al., 2014; Kamphuis, De Kam, Geurts, and Weerdesteyn, 2013). Several activities in daily living require the ability to perform multiple tasks simultaneously. Persons with stroke have been reported to have difficulties with performing another task while walking (An et al., 2014; Yang et al., 2007), as dual tasking can affect balance control.

Assessing gait performance after stroke requires measuring instruments that include various components of walking used in everyday life. Recommended instruments that measure balance control during walking activities are the Dynamic Gait Index (DGI) (Shumway-Cook and Woollacott, 1995), and the Functional Gait Assessment (FGA) (Lin et al., 2010; Wrisley, Marchetti, Kuharsky, and Whitney, 2004). The DGI includes eight items where a physiotherapist (PT) rates performance of different walking tasks such as changing walking speed, walking while turning the head, walking around or over obstacles, and walking with a pivot turn. The FGA is based on the DGI adding three more challenging items. A well-known instrument of functional balance is the Timed Up and Go (TUG) test (Podsiadlo and Richardson, 1991). The TUG test includes a rise from a chair, a 3-m walk, and a turn. However, in daily life, walking is often performed in complex situations; thus, there can be a need of measuring more advanced balance control during walking also capturing the unilateral limitations and asymmetric weight distribution.

The six-spot step test (SSST) was originally developed as a walking test for persons with multiple sclerosis (MS) and includes a motor dual task. The SSST comprises a short walk of 5 m involving shifting weight from one foot to the other, and single-leg standing while shoving wooden blocks (Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006).
The SSST with the complex combination of a forced weight shift from one foot to another, into a single-leg standing, during a simultaneous ambulation activity involving coordination on time, can give the opportunity to provide a more comprehensive picture of walking capacity after stroke than the DGI, the FGA, and the TUG test. However, psychometric properties of the SSST for use in persons with stroke need to be evaluated. In studies of people with MS, the SSST has had good test–retest reliability with high intraclass correlation coefficient (ICC) values of 0.95–0.98 (Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006; Pavan, Tilbery, Lianza, and Marangoni, 2010), and strong associations ($r = 0.73–0.92$) with the measuring instruments: timed 25-ft walk test, 6-min walk test, and the TUG test (Fritz, Jiang, Keller, and Zackowski, 2016; Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006; Pavan, Tilbery, Lianza, and Marangoni, 2010; Sandroff, Motl, Sosnoff, and Pula, 2015).

Psychometric evaluation includes establishing construct validity that reflects the ability of an instrument to measure a concept (Portney and Watkins, 2007). Construct validity can be gathered in different ways such as convergent validity that two measuring instruments are believed to reflect the same phenomenon. Discriminant validity indicates that lower correlations are expected from measuring instruments that are believed to assess different characteristics (Portney and Watkins, 2007). Instruments that are believed to measure other aspects of balance control, such as balance confidence and overall lower extremity function, are expected to have a low correlation to the SSST. The aim of the present study was to investigate the convergent and discriminant validity and test–retest reliability of the SSST in persons with stroke. The hypotheses were that the SSST would be strongly associated with measuring instruments of walking balance: the TUG test, the DGI, and the four square step test (FSST). Discriminant validity was investigated using the timed-stands test (TST) and the Activities-specific Balance Confidence (ABC) scale, and low-to-moderate associations with these were hypothesized.

## Methods

### Study design

This study had a cross-sectional design, and data were collected at five primary health-care centers in Sweden from June 2014 until November 2015. A convenience sample of community-dwelling persons was recruited. The inclusion criteria were (1) a minimum of 6 months poststroke; (2) medically stable with no other disease/condition that significantly influenced gait performance; (3) ability to understand both verbal and written information; and (4) self-reported walking impairment, but ability to walk at least 10 m with or without assistive device. Individuals meeting the inclusion criteria received written information about the study. Thereafter, they were contacted by telephone and invited to participate. Prior to data collection, they signed an informed consent. The study was approved by the Regional Ethics Committee of Uppsala/Örebro, Sweden (2014/058). Reliability studies are recommended to include at least 30–50 participants (Flansbjer, Blom, and Brogardh, 2012; Hopkins, 2000), with larger sample sizes giving more confident results. Therefore, we set out to recruit about 80 individuals.

### Procedures

The SSST is performed in a taped walkway (test field) that is 5 m long and 1 m wide (Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006). A circle with diameter of 20 cm is marked in the middle of each end-line of the test field. Two more circles of the same size are marked on each of the sidelines of the test field, with a distance from the end-line of 1 and 3 on one side, and 2 and 4 m on the other side. One of the end circles is empty and is used as the starting point. In the center of the other five circles, a cylindrical wooden block is placed, with a diameter of 8 cm and a height of 4 cm, and weighing 134 g (Figure 1).

The test starts when the participant first lifts one foot from the start circle and ends when the last block is shoved out of the end circle. The instruction is to do the test in fast walking, criss-crossing from circle to circle. The test is then repeated with the other leg shoving the blocks. The test is performed twice for each leg.

![Figure 1. Six-spot step test (SSST).](image)
circle and always using the same foot to shove the blocks out of the circles. In the present study, the SSST was performed in four trials, two trials for shoving with the paretic and two for shoving with the non-paretic limb. The total mean time for the four trials was calculated, as well as for shoving with the paretic and the non-paretic leg, respectively.

In the TUG test, time is measured from when a person rises from a chair, walks 3 m at comfortable speed, turns around, walks back, and sits down. The participants performed one trial. The TUG was developed as a measure of functional mobility (Podsiadlo and Richardson, 1991) and has been found valid and reliable in persons with stroke (Flansbjer et al., 2005).

The DGI consists of eight items assessing dynamic balance in different walking tasks such as changing walking speed, walking while turning the head, walking around or over obstacles, and walking with a pivot turn. Performance in each item was rated by the PT on a 4-point scale ranging from 0 (severe impairment) to 3 (normal walking ability without a walking aid). The maximum total score is 24 points (Shumway-Cook and Woollacott, 1995). The DGI has been found valid and reliable in persons with stroke (Jonsdottir and Cattaneo, 2007).

The FSST measures the ability to quickly step over four sticks on the floor marking four squares. Four sticks of 2.5-cm height were placed on the floor at 90-deg angles to each other. The person, steps forward, backward, and sideways first in a clockwise and an anticlockwise direction, putting most of the load on one leg at a time but making contact with the floor in each square with each foot (Dite and Temple, 2002). The participants were instructed to step in each square, trying not to touch the sticks. Time was noted as the participants completed the task. Two trials were performed and the mean value of these was calculated. If the sequence was not completed correctly, the person lost balance or moved the stick out of place, and it was possible to repeat (restart) the sequence one time. All participants completed the task. Two trials were performed, and the mean value of these was calculated, as well as for shoving with the paretic and the non-paretic leg, respectively.

The ABC scale is a patient-reported questionnaire that assesses self-rated balance confidence when performing daily activities (Myers, Fletcher, Myers, and Sherk, 1998). Using an 11-point scale, the participants stated how safe and secure they felt when performing 16 different indoor and outdoor activities, ranging from 0% (not at all safe and secure) to 100% (completely safe and secure). The item scores were summarized and then divided by 16 (the number of items). A validated Swedish version of the ABC scale was used (Forsberg and Nilsagard, 2013).

**Procedure**

Data collection was performed on two occasions, 3–7 days apart, by five PTs specialized in neurological physical therapy (including the first author, MAL). The first author performed 70% of the data collection. All conditions were kept as stable as possible; the same PTs performed both tests, at the same location in an ordinary hall-way at the primary health-care centers, and around the same time at the day. If walking aids were needed, it was used during both test occasions.

Before start of the data collection, a training session was performed where the PTs were introduced to the study protocol, all the measuring instruments and the procedure. Although inter-rater reliability of the DGI has been found good (Jonsdottir and Cattaneo, 2007), to achieve consistency of the ratings from the PTs, they watched a video recording of a person performing the DGI and then discussed the scoring. In the ABC scale, the PTs followed the associated instructions and were available for questions when the person rated their own balance confidence.

The first test session lasted approximately 45 min and started with an interview regarding demographic characteristics, use of walking aids indoors and outdoors, and self-reported number of falls during the last 3 months. To describe the participant’s severity of stroke, item 5 (arm) and 6 (leg) from the National Institutes of Health Stroke Scale (Lyden et al., 1994) were tested. The participant was in a supine position, asked to lift one arm to 45 deg, and hold it there for 10 s, then lift a leg to 30 deg and hold that position for 5 s. Performance was graded from 0 to 4 where 0 = able to hold the arm for 10 s and the leg for 5 s, and 4 = no movement in either arm or leg, see further information on the scoring in Table 1.

The balance measurements were then performed in a standardized order, starting with the ABC scale, followed by the TST, the FSST, the DGI, the TUG, and, finally, the SSST. On the second test occasion only, the SSST was performed. The participants were instructed not to change their activity habits between the two test occasions. During the tests, they wore comfortable shoes and if needed they could rest between the measures.
illustrate heteroscedasticity with a systematic variation around the zero line, and a wider spread with larger and lower mean values and also possible outliers. The possibility of heteroscedasticity can be addressed by performing the Pearson’s correlations coefficient test between the mean and the absolute difference between tests 2 and 1 and the means of the two sessions (Bland and Altman, 1999; Huang et al., 2011). An outlier was considered to be present when the difference between the two tests sessions was outside two SDs. All significance tests were two-sided and conducted at the 5% significance level. Statistical analyses were performed using SAS version 9 (SAS, Cary, NC, USA) and SPSS version 22.0 (SPSS Inc., Chicago, IL, USA).

Results

Eighty-one participants, 34 women and 47 men, with a mean age of 70 years were included. Of the total, 43% had left hemiparesis and 57% had right, and 59% used walking aids outdoors. In Table 1, further baseline characteristics are presented.

Correlation coefficients between the SSST and the instruments of walking balance (TUG, DGI and FSST) were $r = 0.832–0.918$, and the TST $r = 0.568$ and the ABC scale $r = 0.596$ (Table 2). The mean time for performing the TUG test (Table 2) was 19.0 s which indicates walking disability.

In Table 3, the mean values from the two test sessions of the SSST are presented. Participants walked significantly faster while shoving the blocks with their non-paretic leg. Thirty-one participants (38%) used one to two canes or crutches to perform all tests.

The mean change from the first to the second test was about 1 s (Figure 2) (Table 3). The Bland–Altman plot gives more values under the line than above, illustrating a better performance at the second test session.

Table 1. Baseline characteristics of the 81 participants.

<table>
<thead>
<tr>
<th>Mean (SD; range)</th>
<th>Mean (SD; range)</th>
<th>Mean (SD; range)</th>
<th>Mean (SD; range)</th>
<th>Mean (SD; range)</th>
<th>Mean (SD; range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean time since most recent stroke (years)</td>
<td>4.8 (5.0; 0.5–31)</td>
<td>4.8 (5.0; 0.5–31)</td>
<td>4.8 (5.0; 0.5–31)</td>
<td>4.8 (5.0; 0.5–31)</td>
<td>4.8 (5.0; 0.5–31)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>70 (7.9; 51–86)</td>
<td>70 (7.9; 51–86)</td>
<td>70 (7.9; 51–86)</td>
<td>70 (7.9; 51–86)</td>
<td>70 (7.9; 51–86)</td>
</tr>
<tr>
<td>Male/Female</td>
<td>47 (58)/34 (42)</td>
<td>47 (58)/34 (42)</td>
<td>47 (58)/34 (42)</td>
<td>47 (58)/34 (42)</td>
<td>47 (58)/34 (42)</td>
</tr>
<tr>
<td>Hemiparesis, left/right</td>
<td>35 (43)/46 (57)</td>
<td>35 (43)/46 (57)</td>
<td>35 (43)/46 (57)</td>
<td>35 (43)/46 (57)</td>
<td>35 (43)/46 (57)</td>
</tr>
<tr>
<td>Type of stroke</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracerebral infarction</td>
<td>60 (74)</td>
<td>60 (74)</td>
<td>60 (74)</td>
<td>60 (74)</td>
<td>60 (74)</td>
</tr>
<tr>
<td>Intracerebral hemorrhage</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
</tr>
<tr>
<td>Subarachnoid hemorrhage</td>
<td>3 (4)</td>
<td>3 (4)</td>
<td>3 (4)</td>
<td>3 (4)</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Previous strokes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>72 (89)</td>
<td>72 (89)</td>
<td>72 (89)</td>
<td>72 (89)</td>
<td>72 (89)</td>
</tr>
<tr>
<td>Yes</td>
<td>9 (11)</td>
<td>9 (11)</td>
<td>9 (11)</td>
<td>9 (11)</td>
<td>9 (11)</td>
</tr>
<tr>
<td>Falls in previous 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>47 (58)</td>
<td>47 (58)</td>
<td>47 (58)</td>
<td>47 (58)</td>
<td>47 (58)</td>
</tr>
<tr>
<td>One fall</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
<td>18 (22)</td>
</tr>
<tr>
<td>More than one fall</td>
<td>16 (20)</td>
<td>16 (20)</td>
<td>16 (20)</td>
<td>16 (20)</td>
<td>16 (20)</td>
</tr>
<tr>
<td>Walking aids outdoors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>27 (33)</td>
<td>27 (33)</td>
<td>27 (33)</td>
<td>27 (33)</td>
<td>27 (33)</td>
</tr>
<tr>
<td>Unilateral support</td>
<td>25 (31)</td>
<td>25 (31)</td>
<td>25 (31)</td>
<td>25 (31)</td>
<td>25 (31)</td>
</tr>
<tr>
<td>Walking cane/crutch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral support</td>
<td>23 (28)</td>
<td>23 (28)</td>
<td>23 (28)</td>
<td>23 (28)</td>
<td>23 (28)</td>
</tr>
<tr>
<td>Rolling walker</td>
<td>22 (27)</td>
<td>22 (27)</td>
<td>22 (27)</td>
<td>22 (27)</td>
<td>22 (27)</td>
</tr>
<tr>
<td>Two walking poles</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Wheelchair</td>
<td>6 (7)</td>
<td>6 (7)</td>
<td>6 (7)</td>
<td>6 (7)</td>
<td>6 (7)</td>
</tr>
</tbody>
</table>

NIHSS: National Institutes of Health Stroke Scale; SD: standard deviation.

Statistical analysis

The distribution of the variables is given as means, standard deviations (SDs), minimums and maximums for continuous variables, and numbers and percentages for categorical variables. To evaluate the external convergent validity between the first test session of the SSST and the other measures, Spearman’s and Pearson’s correlation coefficients were calculated with correlation coefficients <0.30 interpreted as weak, 0.30–0.59 as moderate, and ≥0.60 as strong (Andresen, 2000).

Test–retest reliability was evaluated based on the distribution of the change between the two measurements, using intra-individual SD (standard error of the measurements [SEM]), intra-individual coefficient of variation (SEM%), repeatability coefficient (smallest real difference, [SRD] = $\sqrt{2} \times 1.96 \times$ intra-individual SD), SRD%, and ICC (ICC2,1). Shrout–Fleiss reliability: random set (Bland and Altman, 1996; Shrout and Fleiss, 1979). Wilcoxon’s signed-rank test was used for analysis of systematic changes between the repeated measures.

A Bland–Altman plot was used to visually present the data (Bland and Altman, 1986). Bland–Altman plots are useful for determining whether the difference between the measurements is dependent on the mean of the measurements. The plots can also be used to
Table 3. Results of the six-spot step test (SSST) \( n = 81 \) performed on two test occasions and showing scores for shoving with the paretic versus the non-paretic limb, and the test–retest results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Occasion 1</th>
<th>Occasion 2</th>
<th>Difference test 2–1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean time (SD)</td>
<td>Within group p-value</td>
<td>Mean time (SD)</td>
</tr>
<tr>
<td>Mean of four SSST measures</td>
<td>23.7 (15.8)/17.1 (9.1; 85.6)</td>
<td>3.42</td>
<td>9.48</td>
</tr>
<tr>
<td>Mean of two measures shoving with the paretic limb</td>
<td>25.0 (16.8)/18.0 (8.9; 86.8)</td>
<td>4.90</td>
<td>13.59</td>
</tr>
<tr>
<td>Mean of two measures shoving with the non-paretic limb</td>
<td>22.4 (15.6)/16.6 (8.7; 98.7)</td>
<td>3.94</td>
<td>10.91</td>
</tr>
<tr>
<td>Difference of shoving between the paretic and the non-paretic limbs</td>
<td>2.6 (7.3)/0.7 (−26.2; 24.6)</td>
<td>0.0001</td>
<td>2.7 (6.9)/0.99 (−21.9; 31.5)</td>
</tr>
</tbody>
</table>

CI: Confidence interval; ICC2.1: intraclass correlation coefficient; SEM: standard error of measurement; SEM%: standard error of measurement, in per cent values; SRD: smallest real difference; SRD%: smallest real difference.

Discussion

This study showed strong convergent validity between the SSST and the TUG test, the DGI and the FSST. Discriminant validity was shown with moderate association with the ABC scale and the TST. Test–retest reliability for the SSST was high; however, the fairly high values on SEM and SRD indicate that sensitivity for repeated measures is limited for the SSST.
Most prior studies of the SSST have been performed in persons with MS where similar results on correlations with other walking tests have been found (Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006; Pavan, Tilbery, Lianza, and Marangoni, 2010; Sandroff, Motl, Sosnoff, and Pula, 2015). In the MS population, the SSST seems to be most suitable for persons with moderate MS (Fritz, Jiang, Keller, and Zackowski, 2016; Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006). In people with MS, a strong correlation between the SSST and the TUG has been shown (Fritz, Jiang, Keller, and Zackowski, 2016; Sandroff, Motl, Sosnoff, and Pula, 2015). These studies also found a strong association between the SSST and walking tests involving longer distances, such as the 2- and 6-min walking tests. The correlation coefficients between the SSST and the FSST and the DGI indicate that the SSST can capture dynamic balance also in persons with stroke. In earlier studies in persons with Parkinson disease, the combination of the DGI and the TUG is described as providing a complementary and comprehensive understanding of gait and balance control (Huang et al., 2011). The SSST, as the DGI, includes a more challenging balance activity, and together with the TUG, the SSST could be an option in assessing dynamic balance in persons with stroke.

Discriminant validity was here investigated with a measure of lower extremity function (the TST) and a questionnaire on self-rated balance confidence (the ABC scale). Sandroff, Motl, Sosnoff, and Pula (2015) presented moderate-to-strong correlation with balance confidence and highlighted the novelty of SSST as a measure of ambulatory function also including elements of balance confidence. The moderate associations in the present study suggest that the SSST is also related to strength and endurance in the lower limbs. Feeling confident to do activities in the home and community as assessed in the ABC scale was also found related to results on the SSST.

In the present study, high retest agreement (ICC 0.96) was found, which indicates consistency between the two measures and between the individuals. Test–retest reliability has also investigated in people with MS with good test–retest reliability when performing the second test about 2 h after the first, with about the same median time in both tests (Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg, 2006; Pavan, Tilbery, Lianza, and Marangoni, 2010). Callesen et al. (2018) also found high ICC values on test–retest reliability in MS when the SSST was performed 2 days apart. Reliability values were high for performing the SSST shoving with both the paretic and the non-paretic limbs. However, the participants walked significantly faster when standing on their paretic leg while shoving the block with their non-paretic limb, than vice versa. This could be due to the possibility to take a quick step with the non-paretic leg to restore balance after the kick. Another suggestion could be that lack of strength and coordination in the paretic limb makes shoving with that leg difficult and, therefore, takes longer time.

Comprehensive evaluations of reliability should include assessments of the variability of within-subject measurements. The Bland–Altman plot displays that a better performance was seen at the second test, where one explanation could be the participants being familiar with the test procedure. This finding is consistent with other studies (Callesen et al., 2018; Green, Forster, and Young, 2002) and to minimize this effect, a practice trial allowing participants to familiarize themselves with the test protocol is proposed. An increased intra-individual SD in individuals with slower walking was also seen in the Bland–Altman plot, also seen in earlier reliability gait performance studies in persons with stroke (Flansbjer et al., 2005).

Heteroscedasticity was present with a correlation coefficient of 0.7. In the Bland–Altman plot, one outlier was presented, a person with a slow performance on the SSST and with a better performance on the first test occasion. Even when this person was excluded, heteroscedasticity was still present with a correlation coefficient above the suggested value of 0.3 (Atkinson and Nevill, 1998). The existence of heteroscedasticity may easily occur in patient groups with variability in functioning (Flansbjer et al., 2005). In persons with stroke, it is vital to consider that persons with more challenging balance deficits may present variability in results on walking measures. This may depend on inconsistent level of motor capacity needed for ambulation in persons with a slower walking speed or less walking endurance. Persons with stroke struggle with residual symptoms such as weakness, stiffness, and fatigue, where the influence on balance and walking may vary. To minimize exterior influences, the conditions were kept as stable as possible, with the same PT at both test, in the same location, at the same time at the day, and if needed walking aids, it was the same in both test occasions. To decrease the within variability when evaluating the SSST, a shorter time period between assessments as performed by Nieuwenhuis, Van Tongeren, Sorensen, and Ravnborg (2006) may be needed. Using more defined inclusion criteria may also provide a more homogenous sample, thus probably reducing the existence of heteroscedasticity.

SEM and SRD have not previously been investigated for the SSST. There are no universally accepted limits
for SEM% but Flansbjer et al. (2005) suggest that values \(\leq 10\%\) are clinically relevant with respect to sensitivity. Huang et al. (2011) described limits for the percentage of the minimal detectable change (MDC%), a calculation similar to SRD%, where a MDC% score less than 30 is considered acceptable and less than 10 as excellent (Smidt et al., 2002). The fairly high SEM% and SRD% in the present study pointed to limited sensitivity for detecting changes beyond threshold of error. The mean time of four trials was 23.2 s, and in this group of individuals, the amount of change to reflect a true difference above measurement error must be larger than 3.4 s. The SRD reflects the minimal detectable difference and with a SRD% value of 40.8%, a person needs to walk almost 9.5 s faster to achieve true improvement. We suggest that for persons with stroke, the SSST is performed twice with each leg, thus calculating a score of mean of four trials. The high values on response stability affect the utility of the SSST as a follow-up measure. As suggested by Huang et al. (2011) and Flansbjer et al. (2005), using a percentage of minimal detectable change is more appropriate for interpreting a true change. To promote the utility of the SSST in stroke rehabilitation, it is necessary to investigate the responsiveness and improve the high SRD% values.

Another aspect of challenge of the forced load in the weight shifting task is the compulsory awareness of both the paretic to the non-paretic leg. In persons with stroke, the awareness of the body can be reduced (Guidetti, Asaba, and Tham, 2007) and in the SSST, the person might increase their limits of stability and their awareness of the body, by being forced to take load on each leg. The SSST was easily performed in all the clinical settings and the test takes only about 5–10 min.

Limitations

This study has several limitations. All participants were ambulatory; however, there was variation in balance performance and use of walking aids that are reflected in the large range of scores on the used tests, with relatively high heteroscedasticity. However, the sample can be regarded as representative of a population with mild-to-moderate disability after stroke seen in the primary health care. The inclusion criteria stated self-reported balance impairment where an objective measure of balance would have been appropriate, but the different test results did show that the participants had balance limitations. Furthermore, the results may have been affected by the order of the test items with the SSST as the last measure on the first test occasion. Validity was investigated using well-known measures of walking balance. In previous studies (Callesen et al., 2018; Fritz, Jiang, Keller, and Zackowski, 2016; Pavan, Tilbury, Lianza, and Marangoni, 2010) in people with MS, gait speed and walking distance have been used for validation, which also may be established in stroke. In further studies of the SSST, inter-rater reliability of the SSST also needs to be evaluated.

Conclusion

In summary, the SSST showed strong convergent validity with other instruments of gait and balance validated for stroke. Values on random measurement error were high affecting the use of the SSST as a measure for follow-up evaluations. Based on the SEM% and SRD% values, the most suitable way of applying the SSST is to calculate a person’s mean time in four trials, shoving twice with the paretic leg and twice with the non-paretic leg. The SSST includes a challenging weight shift from the paretic to the non-paretic limb, and the findings in this study suggest that the SSST can be a complementary measure of gait and balance in stroke rehabilitation.

Acknowledgments

We would like to thank the persons with stroke who volunteered to participate, and the physiotherapists who performed the data collection: Helena Tholin, Niclas Borgqvist, Ewa Petterson, and Anna Lövgren.

Declaration of interest

The authors report no declarations of interest.

Funding

This work was supported by the Swedish Stroke Association: [Grant Number 2011], Norrbacka-Eugenia Foundation: [Grant Number 804-14], and the Research Committee, Örebro County Council: [Grant Number OLL-457471].

References

