Intra-observer and interobserver reliability of One Leg Stand Test as a measure of postural balance in low back pain patients

Thomas Maribo1,2, Elena Iversen3, Niels Trolle Andersen4, Kristian Stengaard-Pedersen2, Berit Schiøttz-Christensen3

1Department of Physiotherapy, Aarhus University Hospital, Aarhus, Denmark
2Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark
3Aarhus Rheumatology Clinic, Aarhus, Denmark
4Department of Biostatistics, Institute of Public Health, Aarhus University, Aarhus, Denmark

Objective: To determine the absolute and relative reliability of intra-observer and interobserver measurements of postural balance using the One Leg Stand Test in patients with low back pain.

Patients and methods: Forty-eight patients with chronic low back pain advised to do exercise were included. A test protocol was established based on review of the literature and a pilot test. Patients were timed for their ability to stand on one leg with their eyes open and with eyes closed. After a 20-min break, the procedure was repeated to check for reproducibility.

Results: Eyes open: a ceiling effect was seen as more than 48% of patients were able to stand for the maximum time, and no further analysis was done. Eyes closed: intra-observer reliability was tested in 21 patients; absolute reliability showed a standard error of the measurement (SEM) of 2.48 s and a minimal detectable change (MDC) of 6.88. The relative reliability was acceptable with an intra class correlation coefficient (ICC) of 0.86. Interobserver reliability was tested in 27 patients; absolute reliability showed a SEM of 1.42 s and a MDC of 3.95. The relative reliability was acceptable with an ICC of 0.91.

Conclusions: The One Leg Stand Test can be used to test postural balance in patients with chronic low back pain, as the reliability of the test was acceptable.

Keywords: postural balance, chronic low back pain, One Leg Stand Test, reliability

Introduction

The assessment of postural balance is useful for a wide spectrum of diagnoses, for example, disorders in the lower extremities, neurological diseases, low back pain, or risk of falling.1-4 Postural balance is often defined as ‘the act of maintaining, achieving or restoring a state of balance during any posture or activity’.7 Maintenance of postural balance in a standing position is achieved by continuous afferent and efferent co-ordination between the somatosensory system and the neuromuscular system. The afferent information requires a process of integration of different sensory information from the proprioceptive, visual, and vestibular senses.7,8 The efferent system requires sufficient muscle strength and a co-ordinated neural system.9

An ideal test of postural balance would examine both static and dynamic balance. However, although some clinical tests like the Berg Balance Scale and the
Tinneti Balance Assessment Tool examine both static and dynamic balance, these tests are only relevant for patients with very poor postural balance. In patients with impaired postural balance, but no fall risk, these tests are not sensitive enough to show small deficits. A ceiling effect will appear, for example, patients score close to maximum or maximum points and a re-test will not be able to display improvement. To the best of our knowledge, there is no valid clinical test for both dynamic and static balance in patients with impaired, but not very poor, balance.

Computerised and sophisticated balance tests performed in movement laboratories are found to be valid tests of postural balance, but are not suited for daily routine.

The One Leg Stand Test, where the patient is observed and timed while standing on one leg, is an easy, quick and low-cost test for evaluating the postural balance; however, in the literature there is no consensus on how the test is performed. The reliability and validity of the One Leg Stand Test is not clear.

The purpose of this study was, therefore, to describe a detailed test protocol and determine the absolute and relative reliability (measurement error) of intra-observer and interobserver measurements of postural balance using the One Leg Stand Test on patients with low back pain.

Patients and methods

Patients
Forty-eight patients with chronic low back pain participated in the study. Inclusion criteria were: chronic low back pain (more than 6 months), between 18–75 years of age, and according to a specialist in rehabilitation advised to do exercise. Patients with severe lung and heart diseases, dementia or language problems were excluded.

Patients were randomly assigned to the intra-observer group, where the same physiotherapist tested the patients twice (n = 21) or the interobserver group, where two physiotherapists tested the patients independently (n = 27). All participants read and signed an informed consent. The Ethics Committee for Medical Research in Aarhus County approved the study protocol, and all procedures were according to the declaration of Helsinki II.

Procedure
The test protocol was established after a review of the literature and a pilot test conducted on 10 patients. Patients were tested barefoot with their eyes fixed on a point 1.5 m away, arms crossed on their chest, and one foot lifted approximately 10 cm and placed behind the weight-bearing leg (Fig. 1). Two trials were performed before timing. Patients were timed with their eyes open and eyes closed, standing on either leg. Timing was stopped at 60 s (eyes open) or 30 s (eyes closed) or if it was impossible to maintain the position. Patients were timed four times while standing on each leg; two times with their eyes open, and two times with eyes closed. For each leg, the best result with eyes closed and eyes open were recorded. Whenever patients maintained the position for a full 60 s or 30 s, no second timing was made. After a break of minimum 20 min, the entire procedure was repeated to check for interobserver or intra-observer reproducibility. Patients were assessed by two skilled physiotherapists; who were blinded to the previous test.

In clinical practice, intervention will be made on the basis of an observed poor balance; therefore, the result from the leg with the shortest stand time was used for analysis.
Statistical analysis

Descriptive statistics were used for age, gender, and stand time. Systematic bias was assessed using the \( t \)-test. Scatters of the between-procedure differences were plotted against the procedure means\(^9\) to indicate if the differences were related to time (heteroscedasticity).

The standard error of measurement (SEM) was calculated from the equation standard deviation (SD)/\( \sqrt{2} \), by using the SD of the stand time from all patients in the group.\(^{31} \) It represents the typical error in a single measurement. The 95% confidence interval (95% CI) for the SEM was calculated. The wider the scope of the 95% CI, the less accurate the estimation of the true score and, consequently, the less accurate the detection of a real change after intervention.

The minimal detectable change (MDC) is defined as the measure of statistically significant change between two measurements. The MDC is found by using the equation \( 1.96 \times \sqrt{2} \times \text{SEM} \).\(^{32} \) For a statistically significant change between two observations to be detected, the change must be at least the MDC.

Relative reliability was calculated using the intra class correlation coefficient (ICC) model 2.1. The ICC is a ratio of the variance between subjects over the total variance and ICC\(_{2.1} \) is a fixed model addressing both systematic and random error.\(^{33} \) An ICC > 0.75 was defined as acceptable reliability. Finally, the proportion of patients with differences between tests longer than 5 s were calculated. Ceiling effect were defined as more than 25% of patients scoring within the error margin of the highest score.\(^{34} \) The level of significance was set at \( P < 0.05 \).

Results

Forty-eight patients with chronic low back pain participated in the study; the time-span of low back pain ranged from 8 months to 20 years. All patients were advised to do exercise by an expert in rehabilitation. Twenty-one patients were assigned to the intra-observer group and 12 of these were males (57.1%). The mean age was 49.3 years (SD 9.98; range, 34.9–65.3 years).

Twenty-seven patients were assigned to the interobserver group and nine of these were males (33.3%). The mean age was 51.5 years (SD 12.35; range, 26.8–72.7 years).

Eyes open

Seventeen patients (81%) were able to keep the position at maximum time when intra-observer reliability was tested and 13 patients (48%) when interobserver reliability was tested. As this ceiling effect was seen, the results of the test with eyes open were of no interest and no further analysis was done.

Eyes closed

Intra-observer reliability

Intra-observer reliability was tested in 21 patients. One patient was able to keep the position for 30 s and was excluded from further analysis. For the remaining 20 patients, results of the differences of One Leg Stand Test were normally distributed. Time and differences are shown in Table 1; there were no statistical difference between the two tests (\( P = 0.18 \)). Scatters of the differences plotted against the means showed no heteroscedasticity. Measures of absolute reliability

<table>
<thead>
<tr>
<th>Seconds, mean (95% CI)</th>
<th>Mean difference (95% CI)</th>
<th>SEM (95% CI)</th>
<th>MDC</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>9.95 (6.98;12.92)</td>
<td>-1.10 (-2.74;0.54)</td>
<td>2.48 (1.89;3.62)</td>
<td>6.88</td>
</tr>
<tr>
<td>Test 2</td>
<td>11.05 (7.62;14.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The mean stand time in seconds, the difference between two tests, the standard error of measurement (SEM), the minimal detectable change (MDC), and the intra class correlation coefficient (ICC) for the intra-observer agreement are presented.
Intra-observer and interobserver reliability of One Leg Stand Test as a measure of postural balance in low back pain patients

Maribo et al.

Table 2 Interobserver reliability of One Leg Stand Test indicated by stand time in seconds with eyes closed obtained twice by two physiotherapists (PT A) and (PT B) in 27 patients with chronic low back pain

<table>
<thead>
<tr>
<th>Seconds, mean (95% CI)</th>
<th>Mean difference (95% CI)</th>
<th>SEM (95% CI)</th>
<th>MDC</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT A 6.81 (4.95; 6.68)</td>
<td>−0.30 (−1.09; 0.50)</td>
<td>1.42 (1.12; 1.95)</td>
<td>3.95</td>
<td>0.91</td>
</tr>
<tr>
<td>PT B 7.11 (5.21; 9.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The mean stand time in seconds, the difference between two tests, the standard error of measurement (SEM), the minimal detectable change (MDC), and the intra class correlation coefficient (ICC) for the interobserver agreement are presented.

Indicates measurement noise of 2.48 s and that a difference greater than 6.88 s is likely to be a real difference with 95% certainty (Table 1). Measures of relative reliability were acceptable (Table 1). Differences longer than 5 s were seen in 5% of the tests (Fig. 2).

**Interobserver reliability**

Interobserver reliability was tested in 27 patients. No patients were able to keep the position for 30 s. Results of the differences of One Leg Stand Test were normally distributed. Time and differences are shown in Table 2; there were no statistical difference between the two tests ($P = 0.45$). Scatters of the differences plotted against the means showed no heteroscedasticity. Measures of absolute reliability indicate measurement noise of 1.42 s and that a difference greater than 3.95 s is likely to be a real difference with 95% certainty (Table 2). Measures of relative reliability were acceptable (Table 2). Differences longer than 5 s were seen in none of the tests (Fig. 3).

**Discussion**

This study evaluated the absolute and the relative reliability of the One Leg Stand Test as a measure of postural balance in a population of patients with chronic low back pain. We found an acceptable reliability; with SEM of 2.48 and 1.42 and MDC of 6.88 and 3.95 for intra-observer and interobserver reliability, respectively. In daily practice, the MDC can visualise whether a patient has experienced a real change (with a 95% certainty). Our study showed that a change of more than 6.88 s should be regarded as a change in postural balance. The study indicated measurement noise of 2.48 s.

Previous studies revealed no consensus on how to perform the One Leg Stand Test, and thus we established a test protocol on the basis of the literature, and our own pilot test. No other study has used the same protocol as ours, making it difficult to compare our absolute reliability results with previous studies. So far, three studies have investigated the relative reliability of postural balance in populations of patients with low back pain, all using measures of body sway. Harringe et al. examined postural balance in a population of young female gymnasts with and without low back pain, finding the eyes-closed test more reliable than the eyes-open test, as well as a greater variability in patients with low back pain. They found an ICC of 0.57 with respect to path length of body sway tested on a computerised balance platform. Salavati et al. tested reliability of centre of pressure and presented an ICC between 0.44 and 0.91, but low back pain patients were mixed with subjects suffering from other musculoskeletal problems, making their results difficult to use in daily practice. Leitner et al. cited an ICC of 0.77 when testing sway with eyes closed on a firm surface. We found an ICC of 0.86 and 0.91, making our results comparable with findings from the three previous studies. This indicates that the One Leg Stand Test has the same reliability as more sophisticated measures of postural balance with respect to the problems inherent in the ICC.

When comparing the stand time in our study with the normative data presented by Vereeck et al. and Bohanon et al., we found a significantly poorer
performance. In the study by Vereeck et al.,¹⁴ more than 79\% of asymptomatic adults under 60 years of age were able to stand with their eyes closed for 10 s or more; in our study, only 28\% of patients under 60 years of age could match this. Bohanon et al.¹⁵ found a mean of 21 s or higher for persons under 60 years of age; in our study, the under-60-year-olds had a mean stand time of 8.9 s. These differences are most likely due to the fact that patients in our study had chronic low back pain whereas the studies of normative data only included persons who did not suffer from vertigo, balance problems, or neurological or orthopaedic disorders. The studies on normative data have shown that stand time during the One Leg Stand Test deteriorates markedly for persons over 60 years;¹⁴,¹⁶ we included relatively few subjects aged 60 years or older.

Previous studies have shown that the One Leg Stand Test with eyes open is not challenging enough for younger, healthy persons.¹⁴,²⁶ We, therefore, initially investigated a more challenging test position with 90º flexion in both the hip and the knee joint. This pilot test position triggered increased low back pain in some patients, causing us to change the test position.

Studies have shown impaired balance in patients with musculoskeletal problems such as low back pain or osteoarthritis in the knee or ankle.¹,²,⁵,⁶ When impaired balance is measured in patients having pain, it might give rise to impaired motor control, leading to changes in activities such as walking, running, and daily activities. As clinicians, we measure balance in order to form an impression of whether their motor control reflects any dysfunction in nerves or muscles. To be able to measure balance and effect changes in any dysfunctions, we need reliable and valid tests.

When using common balance tests such as the Berg Balance Scale or the Tinetti Balance Test,¹⁰,¹¹ most of these patients will score maximum or close to maximum scores. Using the protocol given in this study, the One Legged Stand Test provides reliable information on postural balance in patients who have impaired postural balance. The test has focused on relatively well-functioning patients, and the results are not representative for all patients; especially not for patients who are at risk of falling.

The test protocol developed here had no adverse effect on any of the 48 patients in this study and, as only one patient was able to maintain the eyes-closed position for the maximum time, the test protocol seems relevant. Although we found a very strong ceiling effect in testing with eyes open, we still recommend this as a part of the procedure to help patients get accustomed to the test.

The ICC was used to estimate the relative reliability. The result of the ICC is highly dependent on the variability in the population observed; one can get a higher ICC if some patients perform poorly and some patients perform very well. As shown in Figure 3, one patient was able to stand for 27 s observed by both physiotherapists. If this patient were excluded, the ICC would change from 0.91 to 0.70, illustrating how sensitive ICC is to outliers. Thus, one should take into consideration that the ICC must be interpreted with caution. We decided not to conclude anything on the basis of the ICC, but the results of the ICC are shown as it has been widely used by others in reliability studies.³¹–³³,³⁸

Furthermore, if the aim is to show changes in health status, results of absolute rather than relative reliability are preferred.³⁹ Such results describe the agreement of repeated tests by estimating the error in a single measurement applying the scale used for the measuring.³⁹ The error should be smaller than the changes one seeks to detect.

Kappa statistics are often used in reliability studies, providing data in the form of dichotomised answers (yes/no, pain/no pain) or on an ordinal scale (weighted kappa). However, when data are supplied on an interval scale, such as the One Leg Stand Test, other statistical analyses are preferred.

As described earlier, the relative reliability of the One Leg Stand Test seems to be as good as computerised balance tests performed on force platform. The One Leg Stand Test is far easier to use in daily routine as it requires no technical equipment. The question remains whether test of postural balance provides useful and valid information regarding low back pain patients. Future research should investigate how changes in low back pain are reflected in changes in postural balance.

**Conclusions**

A detailed One Leg Stand Test was planned based on literature and a pilot test; when using this protocol in a population of patients with chronic low back pain, the reliability is acceptable with a SEM of 2.48 and 1.42 and a MDC of 6.88 and 3.95 for the intra-observer and interobserver reliability, respectively. The study showed that a change of more than 6.88 s is regarded as real change. The study indicated measurement noise of 2.48 s, meaning that results of the One Leg Stand Test might differ up to 2.48 s without any real difference in balance.
Acknowledgements

This work was supported by a grant from The Association of Danish Physiotherapists’ Research Fund. The help with inclusion of patients by the Danish Rheumatism Association, CST Rehabilitation Centre Aarhus is highly appreciated.

References