A simplified version of the weight-bearing ankle lunge test: Description and test–retest reliability

Antonio Cejudo, Pilar Sainz de Baranda, Francisco Ayala, Fernando Santonja

**A R T I C L E   I N F O**

Article history:
Received 15 October 2013
Received in revised form 17 February 2014
Accepted 21 March 2014

Keywords:
Range of motion
Reproducibility
Sports therapy
Dorsiflexion

**A B S T R A C T**

The purpose of this study was twofold: (1) to describe a new version of the weight-bearing ankle lunge test (WBLT) that is simple to administer, that allows clinicians and sports medicine practitioners to directly assess (in degrees) the ankle dorsiflexion range of motion in a very short period of time while adopting a comfortable testing position; as well as (2) to determine the test–retest reliability of the ankle dorsiflexion range of motion measure obtained from the new version of the WBLT. A total of 50 active adults completed this study. All participants performed the new version of the WBLT on three different occasions, with a two-week interval between testing sessions. Reliability was examined through the change in the mean between consecutive pairs of testing sessions (ChM), standard error of measurement (SEM), minimal detectable change at 95% confidence interval (MDC95), and intracluster correlation coefficient (ICC2,k). The findings showed negligible or trivial ChM values for all the flexibility measures analysed (<1°). Furthermore, the SEM and MDC95 scores for the ankle dorsiflexion measure were 1.3 and 3.8 respectively, and the ICC2,k was 0.95. Therefore, this study demonstrated that the ankle dorsiflexion measure obtained from the new version of the WBLT has excellent test–retest reliability scores. Thus, an observed change larger than 3.8° from baseline scores after performing a treatment would indicate that a real change in ankle dorsiflexion range of motion was likely.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Sports therapy practitioners routinely assess ankle dorsiflexion range of motion (ROM) because inadequate levels of mobility have been proposed as a predisposing factor for increasing the likelihood of several foot and ankle pathologies (DiGiovanni et al., 2002). The weight-bearing ankle lunge test (WBLT) is probably the most widely used measurement method to assess ankle dorsiflexion ROM. The WBLT is based on the “knee-to-wall” principle and requires patients or athletes to perform, in front of a wall, a forward displacement of the pelvis and trunk with knee flexed while a lunge position is adopted (Bennell et al., 1998). The goal of the WBLT is to determine the maximum distance in centimetres between the big toe and the wall without the heel lifting from the ground while the knee is able to touch the wall (Fig. 1) (Bennell et al., 1998).

However, despite the fact that the WBLT has been considered valid enough by the most important American medical organisations (American Academy of Orthopaedic Surgeons, 1965; American Medical Association, 2001) as well as the fact that high reliability scores have been reported (Bennell et al., 1998; SimONSEN et al., 2012), when its testing procedure and equipment is carefully analysed, some important limitations are noted, which might question the applicability of this test in the clinical and sports therapy contexts.

Although simple, the testing procedure of the WBLT is usually repeated up to 5–6 times to enable the foot to be moved away or towards the wall until the “end range” is found (Houch and McKEON, 2011; O’Shea and GRAFFON, 2013). The duration of this testing procedure (approximately 3 min per patient) might represent a relevant limitation when a large number of athletes need to be tested in a very short period of time. Furthermore, during the
testing procedure of the WBLT, the administrator must firmly hold the patient’s heel down and read the distance in the ruler, thereby adopting a very uncomfortable position. Finally, the linear measure obtained from the WBLT should be converted to degrees to know the “real” ankle dorsiflexion ROM. This conversion process may reduce the accuracy of the measurement obtained from the WBLT due to the typical error of the estimation equation (Hopkins, 2000b).

Therefore, the aim of the current study was to describe a new version of the WBLT that is simple to administer, that allows clinicians and sports therapy practitioners to directly assess the ankle dorsiflexion ROM in degrees in a very short period of time while adopting a comfortable testing position. However, before these data can be used for clinical and training goals, values must be defined so as to provide guidance in deciding whether an observed change upon reassessment is within the boundaries of assessment error or whether there has been a true change (Hopkins, 2000a). Consequently, a secondary purpose of this study was to estimate the test–retest reliability of the ankle dorsiflexion ROM measure obtained from the new version of the WBLT that has been described.

2. Method

2.1. Participants

A convenience sample of 24 men (stature: 176.0 ± 6.1 cm; weight: 73.5 ± 8.5 kg; age: 20.3 ± 4.1 years) and 26 women (stature: 165.6 ± 4.3 cm; weight: 57.1 ± 6.9 kg; age: 19.7 ± 3.4 years) who were recreationally active young adults completed this study. The exclusion criteria were: (1) episodes of triceps surae injury over the past six months, (2) missing a testing session, (3) presence of self-reported delayed onset muscle soreness at any testing session, and (4) for women, being in the ovulation phase of their menstrual cycle during testing, as fluctuating concentrations of oestrogen throughout the menstrual cycle affect musculo-tendinous stiffness (Eiling et al., 2007). The participants were verbally informed about the study’s procedures before testing, and they provided written informed consent. This study was approved by the University of Murcia Research Ethics Committees (Spain).

2.2. Procedure

A standard box (30.5 cm high) was used to position the participants for the test, and an ISOMED inclinometer (Portland, Oregon) with a telescopic arm was used as the key measure for the new version of WBLT.

For a better understanding of the assessment method (i.e. instruments, clinician positioning, starting point, final point), a detailed description of the new version of the WBLT is displayed in Fig. 2.

The test–retest reliability of the ankle dorsiflexion measurements obtained from the new version of the WBLT was analysed using a repeated measures design. Thus, each participant underwent the testing procedure twice on three different occasions, with a two-week interval between testing sessions. The rationale for using 50 participants and three testing sessions to determine the reliability in our study (instead of the two testing sessions that have been typically used in previous reliability studies) was based on the simulations run by Hopkins (2000a), who stated that to achieve an accurate reliability score, a minimum of three testing sessions and 50 participants are needed. A physical therapist with greater than 10 years’ experience conducted each of the three testing sessions at the same time of the day under the same environmental conditions. The physical therapist was blinded to the purpose of the study and test results from previous testing sessions (Fig. 3).

Participants were instructed to perform 2 maximal trials of the new version of the WBLT for each limb in a randomised order, and the mean score for each test was used in the subsequent analysis. The mean of the two trials of the WBLT performed at each testing session was used for subsequent statistical analysis instead of the highest score because the magnitude of the error component decreasing when the scores are averaged (Vincent, 1999). Patients who do not tolerate the sensation of stretching or with low experience with it might set the endpoint of a trial of the test before achieving his/her peak ankle dorsiflexion ROM score due to a feeling of apprehension. To avoid the possible influence of this source of error on the stability of the measure, when a variation >5% was found in the range of motion values between the two trials, an extra trial was performed, and the two most closely related trials were used for the subsequent statistical analyses.

2.3. Statistical analyses

The distributions of raw data sets were checked using the Kolmogorov–Smirnov test and demonstrated that all data had a normal distribution (p > 0.05). Men and women were not analysed separately based on the fact that previous studies have reported that in both men and women, the baseline joint ROM respond in the same way whether or not specific and systematic flexibility training is performed (Akbari et al., 2006; Voigt et al., 2007). Furthermore, the possible influence of the different concentrations of oestrogen throughout the menstrual cycle on musculo-tendinous stiffness and joint laxity was controlled for the women participants.
Descriptive statistics were calculated for the ankle dorsiflexion measurements. Paired $t$ tests were used to test for differences between the scores of the dominant and non-dominant limbs. The test—retest reliability of the ankle dorsiflexion measure was measured through the change in the mean (ChM), standard error of measurement (SEM), the minimal detectable change at a 95% confidence interval (MDC95) and intraclass correlations (ICC) (Weir, 2005). The test—retest reliability for the ankle dorsiflexion measure was calculated separately for the consecutive pairs of trials (2–1, 3–2) to be consistent with the interval time between testing sessions (two weeks) (Hopkins, 2000a).

Thus, the ChM was estimated using a spreadsheet designed by Hopkins (2007) via the unequal-variances $t$ statistic computed for changes in scores between paired sessions. To make inferences about the true value of the effect, the uncertainty in the effect was expressed as 90% confidence intervals and as likelihoods that the true value of the effect represents substantial change (negative or positive) (Hopkins et al., 2009). The probability that the true value of the effect was positive or negative was inferred as follows: <0.5%, most unlikely; 1–5%, very unlikely; 6–25%, unlikely; 26–75%, possibly; 76–95%, likely; 96–99%, very likely; >99%, most likely (Hopkins et al., 2009). The SEM was calculated using the raw data via the following formula: $\sqrt{\text{MSE}}$, where MSE is the error mean square from the repeated measures analysis of variance. The MDC95 was calculated as $\text{SEM} \times \sqrt{2} \times 1.96$. The ICC$_{2,k}$ were calculated using the following formula:

**Fig. 2.** Description of the new version of the weight-bearing ankle lunge test.
3. Results

Descriptive statistics for the ankle dorsiflexion measure are displayed in Table 1. The paired t-test analysis reported no significant difference between the dominant and non-dominant legs (mean difference: 1.9°; 90% CI: from –2.1 to 3.8°); thus, the mean of the two legs was used for subsequent reliability analysis.

Test–retest reliability statistics (ChM, SEM, SWC and ICC) for the ankle dorsiflexion measure are presented in Table 2 for both consecutive pairs of testing sessions. The reliability scores obtained for each of the consecutive paired testing sessions (2–1 and 3–2) were almost identical, so that the mean of the two paired testing sessions for each measure might be used as a test–retest reliability criterion of reference (Hopkins, 2000a). The ChMs between consecutive pairs of testing sessions were “possibly trivial” (2–1) and “most likely trivial” (3–2). The SEM and MDC95 scores for the ankle dorsiflexion measure were 1.3 and 3.8 respectively, and the ICC2k was 0.95 (high).

4. Discussion

The main benefits of the new version of the WBLT developed in the current study when contrasted with the original version are the speed of the test, the simplicity of the measurement method, and its ease of execution. These benefits are allowed thanks to the use of an inexpensive inclinometer with a telescopic arm (100€, approximately) and a standard box. The use of an inclinometer as the key measure allows practitioners to achieve the ankle dorsiflexion ROM with just one trial and directly in degrees without any mathematic conversion. Consequently, it may reduce the predictive typical error of the measurement (Hopkins, 2000b). Furthermore, thanks to the telescopic arm of the dynamometer, marking a bony landmark is not required because the maximum ankle dorsiflexion ROM can be determined as the angle formed by the longitudinal axis of the leg (lateral bisector of the leg) with the vertical plane. Finally, the use of a box to position the patients/athletes for the test allows the sports therapy practitioners to adopt a comfortable position while monitoring the tested limb and heel and measuring the ankle dorsiflexion range of motion. However, it should be noted that for patients who have a limited knee flexion ROM, the use of a shorter box (10–15 cm) may be more appropriate because the degree of knee flexion necessary to achieve the end of the ankle dorsiflexion ROM is lower.

Another purpose of the current study was to determine the test–retest reliability of the ankle dorsiflexion ROM measure obtained from the new version of the WBLT. In this regard, the results of this study showed that the ankle dorsiflexion measure analysed had excellent test–retest reliability scores.

The ChM between consecutive testing sessions was negligible or trivial (ranging from 0.6° to 1.2°). This finding may support the idea that the testing procedure is simple to administer and easy for the patients/athletes to follow. Another aspect of reliability that was assessed was the precision of measurements, which was determined using the SEM (Weir, 2005). Admittedly, the clinical decision regarding the cut-off precision values of a measure is challenging, especially since there are no clear guidelines for reference value establishment, as well as the potential need to evaluate multiple factors (training status, sex, age) to reach a knowledgeable decision. However, it appears to be accepted that a variability of a measure lower than 10% could be considered appropriate for clinical and research purposes (Vincent, 1994). Based on this criterion, the ankle dorsiflexion measure obtained from the new version of the WBLT showed an excellent precision, since its percentage of variability (% SEM) was 2.7%. In terms of practical applications, it has been suggested that the MDC95 can be used to indicate the limit for the smallest change that indicates a real improvement in a single person (Weir, 2005). Therefore, clinicians can be 95% confident that an observed change between 2 measures larger than 3.8° for the ankle dorsiflexion measure obtained from the WBLT would likely indicate a real change in ankle ROM. Similar precision in measurement results have been reported in previous studies for the original WBLT (Bennell et al., 1998). Finally, the results of the current study reported high relative reliability scores for the ankle dorsiflexion measure (ICC > 0.9).

While the results of this study have provided information regarding the test–retest reliability of this new musculoskeletal screening test, limitations to the study must be acknowledged. The age distribution of participants was relatively narrow; thus, the ability to generalise to the broader population cannot be

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Descriptive values (mean ± standard deviation) for the modified version of the weight-bearing ankle lunge test measures.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Session 1</td>
</tr>
<tr>
<td></td>
<td>47.0 ± 6.3°</td>
</tr>
<tr>
<td></td>
<td>Degrees.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Test–retest reliability statistics for the modified version of the weight-bearing ankle lunge test measures.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ChM</td>
</tr>
<tr>
<td></td>
<td>TS 2–TS 1</td>
</tr>
<tr>
<td></td>
<td>1.2°</td>
</tr>
</tbody>
</table>

TS: testing session; ChM: change in the mean between consecutive testing sessions; SEM: standard error of measurement; MDC95: minimal detectable change at 95% confidence interval; ICC2k: intraclass correlation coefficients.
ascertained. Similarly, whether the test would be as reliable in a population of injured participants must also be considered, although similar methods have proven reliable with ankle fracture patients (Simondson et al., 2012). In addition, pre-season screening is generally performed in healthy, uninjured populations.

5. Conclusion

The current study provides important information to clinicians and sports therapy practitioners regarding the use of a new modification of the WBLT that can be viewed as a clinically simplified version. Specifically, this new version allows clinicians and sports therapy practitioners to directly assess, in degrees, the ankle dorsiflexion range of motion in a very short period of time (approximately 30s), while adopting a comfortable testing position. In addition, this study demonstrated that the ankle dorsiflexion measure obtained from the new version of the WBLT has excellent test–retest reliability scores.

References