



# Brazilian Journal of Physical Therapy

<https://www.journals.elsevier.com/brazilian-journal-of-physical-therapy>



## ORIGINAL RESEARCH

### Inter-day reliability of the Upper Body Test for shoulder and pelvic girdle stability in adults

Alan dos Santos Fontes<sup>a,\*</sup>, Marta Silva Santos<sup>a</sup>, Marcos Bezerra de Almeida<sup>a</sup>, Pedro J. Marín<sup>b</sup>, Danilo Rodrigues Pereira da Silva<sup>a</sup>, Marzo E. Da Silva-Grigoletto<sup>a</sup>

<sup>a</sup> Physical Education Department, Universidade Federal de Sergipe (UFS), São Cristóvão, Sergipe, Brazil

<sup>b</sup> CYMO Research Institute, Valladolid, Spain

Received 8 November 2018; accepted 14 February 2019

Available online 1 March 2019

#### KEYWORDS

Core stability;  
Physical tests;  
Reproducibility;  
Physical therapy

#### Abstract

**Background:** There are a limited number of tests for the assessment of shoulder and pelvic girdle stability. Reliable instruments are important to evaluate movement dysfunction at these joints in order to provide more objective parameters.

**Objective:** To evaluate the inter-day reliability of the *Upper Body Test* in young adults.

**Methods:** A reliability study was carried out with three assessments of the shoulder and pelvic girdle stability within 48-h intervals (Monday, Wednesday, and Friday). The OctoBalance<sup>®</sup> platform was used to perform the *Upper Body Test* in 31 active young adults ( $24.5 \pm 8$  years). Intraclass Correlation Coefficient (ICC<sub>2,2</sub>) two-way mixed model, Coefficient of Variation, and Bland–Altman plots were used to verify the reliability of the test. The standard error of measurement (SEM) and the minimum detectable difference (MDD<sub>95%</sub>) were calculated for clinical applicability.

**Results:** The Intraclass Correlation Coefficient ranged from 0.87 to 0.94 – Featuring a mean difference of 0.89 (95%CI =  $-0.19$ – $1.97$ ) to left and 0.95 (95%CI =  $-0.38$ – $2.27$ ) for the right side, with a low variation coefficient (3.31–5.91%) between the second and third days of assessment. There was a statistically significant difference between the first assessment day and the other test sessions. The Bland–Altman analyses revealed low bias with scores within the limits of agreement. Minimum detectable difference scores were between 4.02 and 5.10, and standard error of measurement between 1.75 and 2.72, depending on the movement side.

**Conclusion:** The *Upper Body Test* presented good inter-day reliability for assessing the stability of the shoulder and pelvic girdle in young active healthy adults.

© 2019 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier Editora Ltda. All rights reserved.

\* Corresponding author at: Av. Gonçalo Roemberg Leite, 1960, Ed. Gama, apt. 1201, CEP: 49050-370 Aracaju, Sergipe, Brazil.  
E-mail: [al.sfontes@gmail.com](mailto:al.sfontes@gmail.com) (A.S. Fontes).

## Introduction

Closed kinetic chain exercises for the upper extremity challenge the individuals in terms of strength, speed, mobility and stability.<sup>1,2</sup> Stability is the ability of the musculoskeletal structures to maintain a position or to move despite the presence of intrinsic or extrinsic disturbances to the body during the movement.<sup>3</sup>

Assessing the stability of the shoulder and pelvic girdle may help therapists to determine risk of injuries, and to prescribe exercise programs to individuals with shoulder dysfunction.<sup>1,4</sup> Few tests for the assessment of functional movements and stability of the shoulder girdle in closed kinetic chain have been described in literature.<sup>5,6</sup> The One-arm Hop test, the Closed-Kinetic Chain Upper Extremity Stability and the Upper Quarter Y-Balance tests have a base of support that limits joint mobility, challenging the assessment of trunk and shoulder girdle stability.<sup>1</sup> The Upper Quarter Y-Balance Test guides the trajectory of the movement with plastic tubes. These sets of characteristics limit the observation of compensatory movements, which might have a direct effect on the occurrence of injuries. Therefore, these tests could mislead interpretation. The One-arm Hop test and the Closed-Kinetic Chain Upper Extremity Stability require that the individual quickly reach both left and right body sides, combining closed and open kinetic chain activities. These characteristics are relevant when assessing the plyometric ability of the upper limbs.<sup>7</sup>

The *Upper Body Test* (UBT) has been created to meet the need for more accurate assessment of the shoulder girdle stability.<sup>2</sup> To apply the UBT, a base of support with two points is provided to the individual, increasing the difficulty of the test and allowing the evaluation of the lumbo-pelvic segment.<sup>8</sup>

However, the literature lacks studies on the reliability of the UBT for assessing stability of the shoulder and pelvic girdle. The objective of the present study was to evaluate the inter-day reliability of the UBT in the assessment of the shoulder and pelvic girdle stability. We hypothesized that the UBT is a reliable inter-day measurement test to evaluate shoulder and pelvic stability.

## Methods

### Experimental design

This is a methodological study carried out to test the reliability of the UBT. Measurements were taken in 3 days with intervals of 48 h.

### Subjects

Based on a pilot study, sample size was calculated using the G\*power 3.1.9.2 program and considering  $\alpha = 0.05$ ,  $\beta = 0.20$ , power correlation ratio to null hypothesis of 0.35 and power correlation ratio to alternative hypothesis of 0.80.<sup>9,10</sup> At least 29 participants were needed for the study. Considering a potential loss of 20%, 35 university students were recruited. Of those, four missed the second test day, due to personal problems. Therefore, 31 individuals (9 men and

22 women) composed the final sample and completed the three-day assessment.

Inclusion criteria were participants aged 18–45 years, and physically active measured by the International Physical Activity Questionnaire short form.<sup>11</sup> Exclusion criteria were: history of orthopedic, neurological or cardiac problems; upper or lower limbs amputation; history of vestibular or physical disorders that could influence joint mobility and muscle tone; under medication that could inhibit the central nervous system; and failure to attend one of the sessions and if the measurement was assessed in a time interval over 48 h. The present study was approved by the ethical committee of the Universidade Federal de Sergipe (HU-UFS), São Cristóvão, Sergipe, Brazil under the number 053820/2017 and participants signed a free and informed consent for participation in the study.

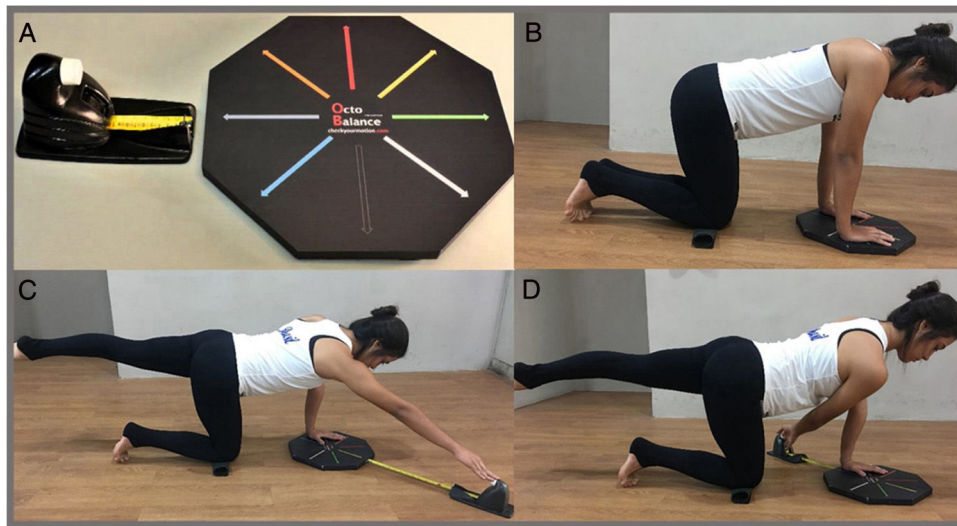
### Procedures

Body mass and height were obtained to calculate the body mass index (BMI,  $\text{kg/m}^2$ ) and characterize the sample. The OctoBalance<sup>®</sup> was used to assess the UBT. The device provides direction for performance of movements through colored arrows and a flexible metric scale. These features facilitate the performance of the movement as well as the range of motion achieved in each pattern of movement. The fixed platform is connected through magnets to a mobile part, which contains a scale that measures the distance in millimeters. These magnets are sensitive to changes in movement direction and get detached as it occurs. From the magnet detachment we can infer that the participant performed compensatory movements or changed movement direction. This feature helps understand the movement pattern.

Before the tests were carried out, the length of the upper limbs was obtained with a tape measure graded in millimeters and with the participant in the orthostatic position, shoulders flexed to 90°, elbows extended and wrists hyperextended. The measurement was taken from the acromion of the scapula to the first joint line, after the styloid process of the ulna. Upper limbs length was used to normalize the measurements obtained with the OctoBalance<sup>®</sup>, providing an assessment of the range achieved adjusted by limb length.<sup>1</sup> Following, the participants received instructions about the assessments.

In order to familiarize with the UBT procedures, the participant performed three movements in the superior-medial and inferior-lateral direction to the left and right sides, with a 30-s interval between movements. After one minute of rest, the individuals repeated the same movements but at this time the distance obtained in each movement pattern was recorded.<sup>7</sup> Thus, each participant performed 12 movements for each limb (24 total) per assessment day (Fig. 1).

A physical therapist, with three years of experience, conducted the experiment, and was blinded to scores that were recorded on a form by a research assistant. Prior to the start of each trial, the measuring tape was placed 20 cm from the platform of the OctoBalance<sup>®</sup>. The initial position of the test was standardized with the individual in a quadruped position (Fig. 1B) with the following recommendations: (a) individual in the quadruped position, supported in four points, with



**Figure 1** (A) Illustration of the OctoBalance®; (B) Initial position for assessment (quadruped); (C) Final position for assessment of the left superomedial pattern; (D) Final position for assessment of the left inferolateral pattern.

hands on the platform of the OctoBalance®; (b) metacarpophalangeal joints positioned on the arrows that were perpendicular to the body of the individual; (c) knees over a thin layer of foam, for protection; (d) hip, knees, ankles and shoulders flexed at 90°.

The participant performed specific movements supported in two points (knee and contralateral upper limb). To evaluate the left superior-medial pattern (Fig. 1C), the individual was asked to perform hip and knee extension on the left side, followed by movement of the right upper limb pushing the movable platform in the direction indicated by the yellow arrow. After this movement, the individual returned to the quadruped position and then, the assessment of the left inferior-lateral pattern (Fig. 1D) was conducted: elevation of the left lower limb followed by right upper limb pushing the platform in the direction indicated by the blue arrow. The assessments followed a specific order, starting with the left side followed by the right side, using the antero-medial and postero-lateral directions of the OctoBalance® (Fig. 1A).

The individuals were instructed to maintain their posture during the assessments. They were asked to avoid trunk rotation, flexion of the elevated lower limb, partial dislocation of the shoulder joint, flexion of the elbow of the supporting limb, and to maintain normal breathing and voluntary activation of the abdominal muscles to ensure trunk position.

Tests were discarded in the following cases: (A) abrupt pushing of the mobile platform; (B) intermittent displacement of the platform; (C) loss of balance during the test or failure to return to the initial supporting position; (D) diagonal displacement of the mobile platform in relation to the standard movement, (E) movement between the fixed base and the movable platforms of the OctoBalance®; (F) exaggerated elbow flexion (approximately 15°) of the supporting limb; (G) loss of hip extension; and (H) elevation of the ipsilateral lower limb to the side that moves the mobile platform.

The index of the UBT for each pattern of movement was calculated by dividing the mean value obtained in each

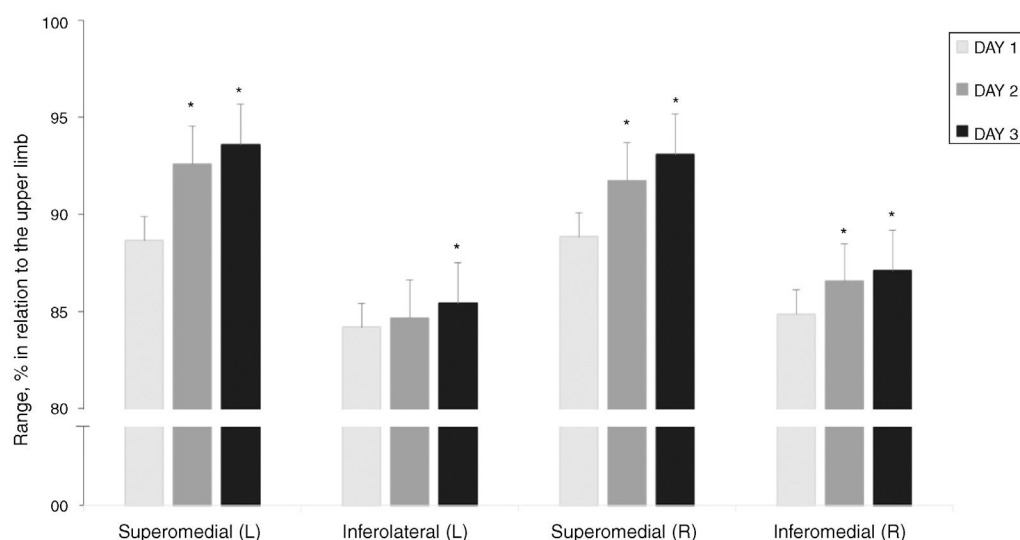
direction by the length of the corresponding limb, and then multiplied by 100.<sup>12</sup> The formula is showed below:

$$UBT = \left( \frac{(r1 + r2 + r3)/3}{UL} \right) \times 100$$

where  $r1$ ,  $r2$ ,  $r3$  are the scores obtained by the individuals in each direction; and  $UL$  the length of the upper limb. The index was obtained for the left and right sides of the body and expressed as percentages in relation to the length of the upper limb.

### Statistical analysis

The Shapiro-Wilk test was used to check the normality of the data. Descriptive statistics were expressed as mean and standard deviation (SD). After checking the assumptions of parametric statistics, the scores of the shoulder and pelvic girdle stability in the different assessment days were compared by repeated-measures Analysis of Variance (one-way ANOVA), followed by Bonferroni post hoc test. For the reliability, intraclass correlation coefficient (ICC<sub>2,2</sub>) two-way mixed model, with 95% confidence intervals and coefficient of variation (CV) were calculated. ICC values <0.25 were considered poor correlation, between 0.26–0.49 low correlation, 0.50–0.69 moderate correlation, 0.70–0.89 high correlation and >0.90 very high correlation.<sup>13</sup> Bland-Altman plots were used to verify the agreement between measurements.<sup>14</sup> Pearson correlation test was used for trend analyses. The standard error of measurement (SEM) was calculated using the equation:  $SEM = SD \times \sqrt{1 - ICC}$ , where SD corresponds to the standard deviation from day two. The minimum detectable difference (MDD<sub>95%</sub>) with 95% confidence interval was obtained with the equation  $MDD_{95\%} = 1.96 \times \sqrt{2 \times SEM}$ .<sup>15</sup> All analyses assumed statistical significance at  $p \leq 0.05$ . Statistical procedures were performed with SPSS® software version 22.0.



**Figure 2** Upper Body Test indexes in the three assessment days. Note. (\*) Statistical difference ( $p \leq 0.05$ ) in relation to day 1.

**Table 1** Values of the tests (day 2 and 3), intraclass correlation coefficient (ICC), coefficient of variation (CV), standard error of measurement (SEM) and minimum detectable difference (MDD<sub>95%</sub>) between the second and third assessment of the Upper Body Test.

| Movements       | Day 2           | Day 3           | Day 3–2           | Day 2 and 3      |      |      |                    |
|-----------------|-----------------|-----------------|-------------------|------------------|------|------|--------------------|
|                 | Mean $\pm$ SD   | Mean $\pm$ SD   | MD (CI 95%)       | ICC (CI 95%)     | CV   | SEM  | MDD <sub>95%</sub> |
| L superomedial  | 92.60 $\pm$ 7.7 | 93.62 $\pm$ 8.1 | 1.02 (–0.89–2.93) | 0.87 (0.73–0.94) | 5.6% | 2.59 | 7.18               |
| L inferolateral | 84.67 $\pm$ 8.3 | 85.43 $\pm$ 9.3 | 0.76 (–1.08–2.59) | 0.91 (0.82–0.96) | 5.9% | 2.46 | 6.82               |
| R superomedial  | 91.74 $\pm$ 7.8 | 93.09 $\pm$ 8.2 | 1.35 (–0.58–3.28) | 0.88 (0.75–0.94) | 5.7% | 2.70 | 7.48               |
| R inferolateral | 86.55 $\pm$ 7.6 | 87.10 $\pm$ 8.4 | 0.55 (–1.44–2.53) | 0.87 (0.74–0.94) | 6.2% | 2.72 | 7.54               |
| L side          | 88.63 $\pm$ 6.0 | 89.52 $\pm$ 6.5 | 0.89 (–0.19–1.97) | 0.94 (0.88–0.97) | 3.3% | 1.45 | 4.02               |
| R side          | 89.15 $\pm$ 6.4 | 90.09 $\pm$ 6.6 | 0.95 (–0.38–2.27) | 0.92 (0.83–0.96) | 4.0% | 1.84 | 5.10               |

Note: L, left; R, right; SD, standard deviation; MD, mean difference.

## Results

The mean age of the participants was  $24.8 \pm 5$  years, body mass of  $66.6 \pm 13.5$  kg, height of  $1.65 \pm 0.08$  m and body mass index of  $24.4 \pm 4$  kg/m<sup>2</sup>. Statistically significant differences in the superior medial pattern of the left and right sides were observed between the first and second days of assessment (Fig. 2). However, these differences were not observed between the second and third assessment days.

As there were significant differences between the first assessment compared to the second and third assessments, only the second and third measures are described in the following analysis. ICCs, CVs, SEM and MDD<sub>95%</sub> between the second and third assessments are presented on Table 1. High and very high correlation values (ICC 0.87–0.94) and low CV's were observed for all movements. The indicators of reliability and clinical relevance for the global indices of the left and right sides were: MDD<sub>95%</sub> sides Left 4.02 and Right 5.10.

The agreement between the measurements between the second and third days (Fig. 3) and body sides (Fig. 4) are presented below. A bias close to zero and a range of acceptable agreement between standards were observed. The

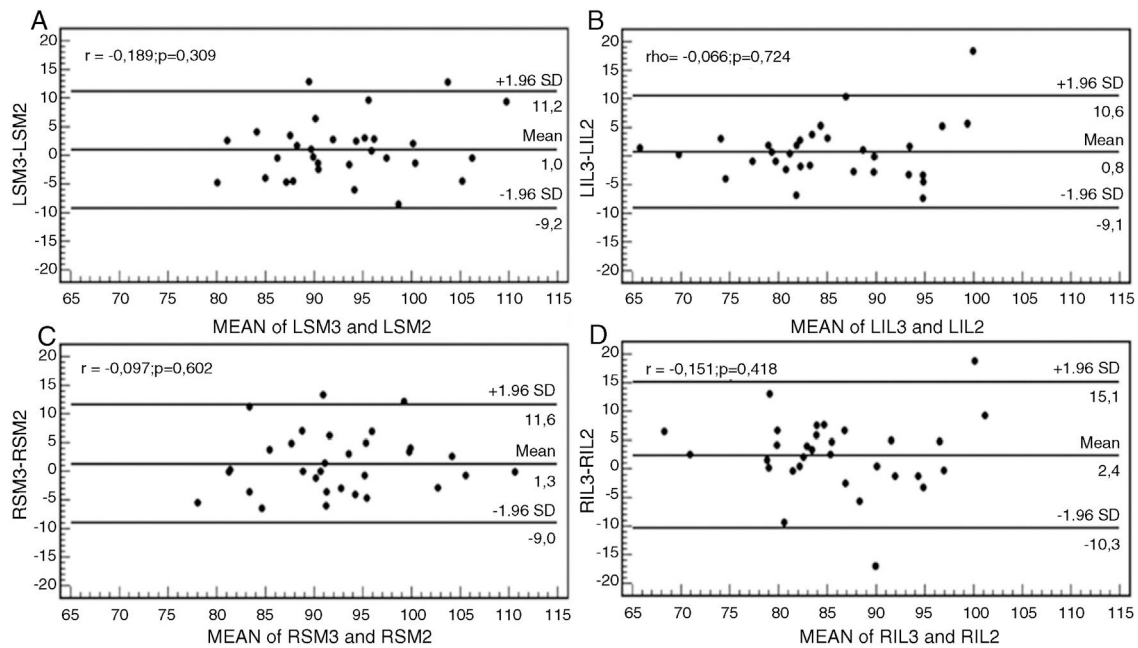
best indicators (low bias and values within the limits of agreement) were observed for the global analysis for the left and right body sides, with most individuals within the limits of agreement. Only one participant was above the limits of agreement of 95% for the left side (Fig. 4A).

## Discussion

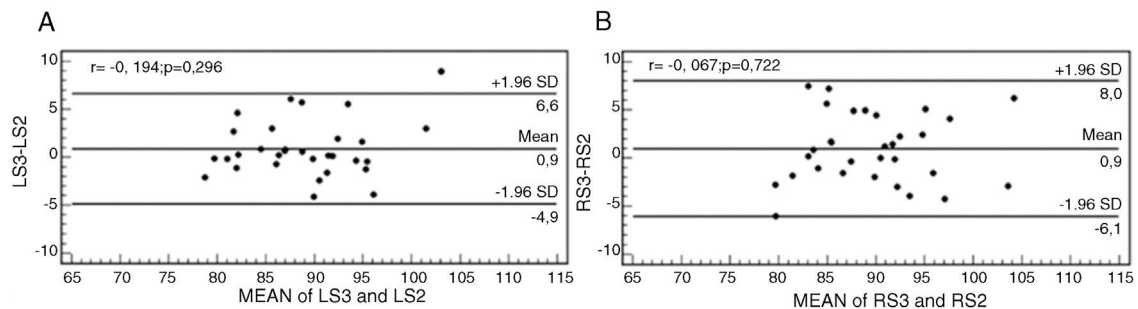
The results indicate that the UBT is a reliable instrument to assess stability of the shoulder girdle and lumbo-pelvic segment. However, familiarization to the UBT is essential before measures are taken.

High ICC values were found between the second and third assessments. The values were higher for the left and right sides of the body. Bland–Altman plots confirmed the reliability of the measures, showing a bias close to zero for most of the participants. Only one individual extrapolated the limits of agreement to the left side. The other individuals are within the limits of agreement. Thus, it is possible to suggest that there is good agreement between the attempts. However, we were unable to compare the values obtained in the





**Figure 3** Bland–Altman plots of differences and averages between the assessed patterns obtained in the UBT: (A) left superomedial (LSM); (B) left inferolateral (LIL); (C) right superomedial (RSM); (D) right inferolateral (RIL).



**Figure 4** Bland–Altman plots for visualization of the differences and means between the sides obtained by the UBT: (A) Left Side (LS); (B) Right Side (RS).

present study with other data, because reliability studies with the UBT are scarce.

Similar studies assessing the functional performance of the upper limbs, shoulder and pelvic girdle stability using the Upper Quarter Y-Balance Test in active adults presented reliable results.<sup>1,7</sup> However, the minimum differences detected in those studies (medial 8.1; superolateral 6.4; inferolateral 6.1)<sup>1</sup> were similar to those found with the UBT – measures lower than 7.54 for movement patterns, and lower for the body sides (4.02 for the left and 5.10 for the right side). These results indicate greater sensitivity of the UBT, permitting observation of slight changes between two assessments. It is important to highlight the limitations to visualize compensatory movements in the Upper Quarter Y-Balance Test due to the presence of a plastic tube.

MDD<sub>95%</sub> values differ from reproducibility tests but are important for clinical practice, since indicate the minimum difference detected between two assessments that are not attributable to measurement error.<sup>16,17</sup> Thus, if the UBT is used on a post-intervention and the scores found were below

the MDD<sub>95%</sub> presented on Table 1, the interpretation would be no clinical improvement because the value would be within the margin error of the instrument. On the other hand, scores higher than the MDD<sub>95%</sub> suggested in this study means functional improvement.<sup>17</sup>

The Closed-Kinetic Chain Upper Extremity Stability Test measures movements of the body sides simultaneously, limiting the comparison between sides.<sup>7,18</sup> The method used in the present study permits free and full range of motion, the detection of compensations by synergistic muscles and independent comparisons between the performance of the left and right sides.<sup>19</sup> These attributes associated with the results of the present study favor the selection of UBT.

Stability tests aim to differentiate levels of physical performance (school, university and at professional levels) among individuals, and identify thresholds for risk of injuries.<sup>20</sup> For health professionals, it may help in the prescription of exercises for stabilization and reducing the mechanical overload on the spine.<sup>20</sup> In addition, it has been shown that adolescent athletes who reported shoulder pain

had deficits in stability tests.<sup>21</sup> Therefore, upper limbs test in closed kinetic chain with the UBT characteristics may potentially function as a differential method between an injured limb and an asymptomatic limb. This type of comparison is necessary to evaluate rehabilitation programs, to provide indicators for the return to the activity, or to suggest biomechanical adjustments to improve performance.<sup>4</sup> Thus, future studies with UBT should include healthy and symptomatic individuals to provide parameters for treatment of injured upper limbs.

Finally, the familiarization measure is necessary to reduce the learning effects of the practice. Previous studies<sup>22,23</sup> have shown that this need is greater for exercises with high complexity, and high activation of the core muscles, as in the case of the movements used by UBT.

One limitation of the study was the impossibility to test the criterion validity of the instrument due to the lack of a gold standard measure. In addition, the results are limited to other populations, like older adults, children and people with injuries, emphasizing the need for studies in other populations. Lastly, future investigations should also use anchor methods to determine the clinical relevance to the UBT.

## Conclusion

The *Upper Body Test* presented satisfactory reliability for the assessment of shoulder and pelvic girdle stability in asymptomatic and physically active adults. However, we recommend at least two test sessions for good reliability results.

## Conflicts of interest

The authors declare no conflicts of interest. Except Pedro J. Marín who is the creator of the OctoBalance instrument.

## References

- Gorman PP, Butler RJ, Plisky PJ, Kiesel KB. Upper Quarter Y Balance Test: reliability and performance comparison between genders in active adults. *J Strength Cond Res*. 2012;26(11):3043–3048.
- Bullock GS, Brookreson N, Knab AM, Butler RJ. Examining fundamental movement competency and closed-chain upper-extremity dynamic balance in swimmers. *J Strength Cond Res*. 2017;31(6):1544–1551.
- Vera-García FJ, Barbado D, Moreno-Pérez V, Hernández-Sánchez S, Juan-Recio C, Elvira JLL. Core stability. Concepto y aportaciones al entrenamiento y la prevención de lesiones. *Rev Andal Med Deporte*. 2015;8(2):79–85.
- Contemori S, Biscarini A, Botti FM, Busti D, Panichi R, Pettorossi VE. Sensorimotor control of the shoulder in professional volleyball players with isolated infraspinatus muscle atrophy. *J Sport Rehabil*. 2017;12:1–29.
- Cramer J, Quintero M, Rhinehart A, et al. Exploration of score agreement on a modified upper quarter y-balance test kit as compared to the upper quarter y-balance test. *Int J Sports Phys Ther*. 2017;12(1):117–124.
- De Oliveira VM, Pitangui AC, Nascimento VY, Da Silva HA, Dos Passos MH, De Araújo RC. Test–retest reliability of the Closed Kinetic Chain Upper Extremity Stability Test (CKCUEST) in adolescents: reliability of CKCUEST in adolescents. *Int J Sports Phys Ther*. 2017;12(1):125–132.
- Westrick RB, Miller JM, Carow SD, Gerber JP. Exploration of the y-balance test for assessment of upper quarter closed kinetic chain performance. *Int J Sports Phys Ther*. 2012;7(2):139–147.
- García-Vaquero MP, Moreside JM, Brontons-Gil E, Peco-González N, Vera-García FJ. Trunk muscle activation during stabilization exercises with single and double leg support. *J Electromyogr Kinesiol*. 2012;22(June (3)):398–406.
- Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Med*. 2004;13(August (4)):251–271.
- Shieh G. Sample size requirements for the design of reliability studies: precision consideration. *Behav Res Methods*. 2014;46(September (3)):808–822.
- Matsudo SM, Araújo T, Matsudo VR, et al. Questionário Internacional de Atividade Física (IPAQ): estudo de validade e reprodutibilidade no Brasil. *Rev Bras Ativ Fis Saúde*. 2001;6(2):5–18.
- Myers H, Poletti M, Butler RJ. Difference in functional performance on the upper-quarter y-balance test between high school baseball players and wrestlers. *J Sport Rehabil*. 2017;26(May (3)):253–259.
- Domholdt E. *Physical Therapy Research*. Philadelphia, PA: W.B. Saunders Company; 1993.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327(8476):307–310.
- Terwee CB 1, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(January (1)):34–42.
- Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther*. 2006;86(5):735–743.
- Paço M, Cruz E. Fiabilidade intra-observador, erro de medida e mudança mínima detectável do weight-bearing lunge-test e do teste de deslizamento posterior do astrágalo em indivíduos com história de entorse do tornozelo. *ifisioonline*. 2011;2(2):25–31.
- Goldbeck TG, Davies GJ. Test–retest reliability of the closed kinetic chain upper extremity stability test. *J Sport Rehabil*. 2000;9:35–45.
- Gonzalo-Skok O, Serna J, Rhea MR, Marín PJ. Relationships between functional movement tests and performance tests in young elite male basketball players. *Int J Sports Phys Ther*. 2015;10(5):628–638.
- Butler RJ, Southers C, Gorman PP, Kiesel KB, Plisky PJ. Differences in soccer players' dynamic balance across levels of competition. *J Athl Train*. 2012;47(6):616–620.
- Oliveira VMA, Pitangui ACR, Gomes MRA, Silva HAD, Passos MHPD, Araújo RC. Shoulder pain in adolescent athletes: prevalence, associated factors and its influence on upper limb function. *Braz J Phys Ther*. 2017;21(2):107–113.
- Williams RM, Corvo MA, Lam KC, Williams TA, Gilmer LK, McLeod TC. Test–retest reliability and practice effects of the stability evaluation test. *J Sport Rehabil*. 2017;26(May (3)), <http://dx.doi.org/10.1123/jsr.2016-0163> (Epub 2017 January 17).
- Moya-Ramón M, Juan-Recio C, Lopez-Plaza D, Vera-García FJ. Dynamic trunk muscle endurance profile in adolescents aged 14–18: normative values for age and gender differences. *J Back Musculoskelet Rehabil*. 2018;31(February (1)):155–162.