Physical Therapy in Sport 38 (2019) 162–169

Contents lists available at ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Original Research

The validity and reliability of the Vail Sport Test™ as a measure of performance following anterior cruciate ligament reconstruction

Joseph Hannon [a,*], Sharon Wang-Price [b], Chad Swank [b], Kelli Brizzolara [b], Craig Garrison [c], James Bothwell [d], Curtis Bush [e]

[a] Texas Health Sports Medicine, 6301 Harris Pkwy, Suite 150, Fort Worth, TX, 76132, United States
[b] Texas Womans University, 5500 Southwestern Medical Ave, Dallas, TX, 75235, United States
[c] 800 5th Ave, Suite 150, Texas Health Sports Medicine, Fort Worth, TX, 76104, United States
[d] Texas Health Physician Group, 6301 Harris Pkwy, Suite 250, Fort Worth, TX, 76132, United States
[e] Texas Health Physician Group, 800 5th Ave, Suite 100, Fort Worth, TX, 76104, United States

A R T I C L E I N F O

Article history:
Received 15 March 2019
Received in revised form
17 May 2019
Accepted 18 May 2019

A B S T R A C T

Objectives: To determine the validity and inter-session reliability of the Vail Sport Test™.
Design: Cohort study-exploratory methodological research design.
Setting: Clinical Research Laboratory.
Participants: Forty-eight participants who underwent ACL-R.
Main outcome measure: Participants performance on the Vail Sport Test™ was graded by an experienced rater in real-time, and simultaneously recorded by a three-dimensional (3D) motion capture system. Construct validity was assessed using the reference standards of the camera system and the IKDC short form. To determine the between-day reliability, a subset of participants returned to repeat the test.
Results: There were no significant difference between the scores collected in real-time and from the kinematic data on the involved limb ($p = 0.222$). There was a significant difference for the uninvolved limb ($p = 0.015$). There was no significant difference between the scores collected in real time and those of the IKDC ($p = 0.885$). Good inter-session reliability (ICC = 0.787) was found for the involved limb.
Conclusion: The results of this study showed good reliability and partially support the validity of the Vail Sport Test as a measure of readiness to return to play.

1. Introduction

Currently, a number of return-to-sport tests can be found in literature and have been used in the decision to return athletes to sport, specifically for those following an anterior cruciate ligament (ACL) injury or an ACL reconstruction (ACL-R)(Barber-Westin & Noyes, 2011; Bien & Dubuque, 2015; Garrison et al., 2012; Harris et al., 2014). However, a systematic review examining return-to-sport testing showed that 65% of the studies reviewed did not report the criteria used to determine when an athlete is ready to return to sport (Harris et al., 2014). The tests that are implemented typically assess metrics of balance (e.g., Y-Balance Test), strength (e.g., isokinetic or hand-held dynamometric test), and power (e.g., hop test)(Ardern, Webster, Taylor, & Feller, 2011a; Augustsson, Roland, & Karlsson, 2004; Barber-Westin & Noyes, 2011; Butler, Lehr, Fink, Kiesel, & Plisky, 2013; Fitzgerald, Lephart, Hwang, & Wainner, 2001; Garrison et al., 2012; Gribble, Hertel, & Plisky, 2012; Grindem, Snyder-Mackler, Moksnes, Engebretsen, & Risberg, 2016; Kokmeyer, Wahoff, & Mymern, 2012). Although the above list may seem encompassing, many flaws continue to exist in the current return-to-sport testing regimen, such as a lack of evidence supporting the use of these tests to increase a safe return of injured athletes back to sports and lack of test components that challenge functional ability in planes other than the sagittal plane(Garrison et al., 2012).

The risk of secondary injury following ACL injury in young athletes is high. The total second ACL re-injury rate reportedly was 15%, with an ipsilateral re-injury rate of 7% and contralateral injury rate of 8%(Wiggins et al., 2016). Furthermore, the secondary ACL

* Corresponding author. Texas Health Sports Medicine, 6301 Harris Pkwy, Fort Worth, TX, 76132.
E-mail addresses: Josephhannon@texashealth.org (J. Hannon), Swang@twu.edu (S. Wang-Price), Cswank@bswrehab.com (C. Swank), Kbrizzolara@twu.edu (K. Brizzolara), Craiggarrison@texashealth.org (C. Garrison), Jamesbothwell@texashealth.org (J. Bothwell), Curtisbush@texashealth.org (C. Bush).

https://doi.org/10.1016/j.ptsp.2019.05.001
1466-853X/

injury rate for patients younger than 25 years was 21%, and for athletes who returned to sport was 20%. Combining these risk factors, athletes younger than 25 years who returned to sport had a secondary ACL injury rate of 23%(Wiggins et al., 2016). These high re-injury rates potentially indicate serious flaws in our return-to-sport criteria. Therefore, modifications of the current return-to-sport criteria and/or implementation of a better return-to-sport test is warranted to reduce re-injury risk after ACL-R(Ardern et al., 2011a; Ardern, Webster, Taylor, & Feller, 2011b; Wiggins et al., 2016; Xergia, Pappas, Zampeli, Georgiou, & Georgoulis, 2013).

As a variety of return-to-sport protocols have been reported in literature, it is likely that the current return-to-sport testing regimens may be inconsistent in both implementation and interpretation. Due to the inadequacy of current return-to-sport tests, the Vail Sport Test™ (Fig. 1a–e) was developed to assess an athlete's ability to perform four major sport-specific functional activities: (1) single leg squat, (2) lateral bounding, (3) forward running, and (4) backward running). For each of the test components, a clinician assesses the athlete's joint movements and subjectively judges how well the athlete performs during these four sport-specific functional activities and, therefore, identifies points of weakness.

Unlike the single-leg forward hop for distance test which is limited to the sagittal plane, the Vail Sport Test™ assesses an athlete's kinematics in multiple planes of motion. In addition, the Vail Sport Test™ requires the athlete to move through both the frontal and sagittal planes while continuing to go through vertical excursions. Lastly, the Vail Sport Test™ also incorporates external perturbation to the athlete during the testing procedure. A blue or black Sport Cord® resistive band (STI, Baton Rouge, LA) used during the test acts as resistance to further challenge the athlete to maintain appropriate trunk and lower extremity positioning (Garrison et al., 2012).

The Vail Sport Test™ has great potential to be an optimal test for safely returning athletes to sports. Although the within-session reliability of the test has been established, the validity and intersession reliability have not been established for the Vail Sport Test™ (Garrison et al., 2012). Therefore, the purposes of this study were to assess the reliability and validity of the Vail Sport Test™ as a measure of readiness to return to sports following ACL-R.

2. Methods

A power analysis using G*Power 3.1.3 (Faul, Erdfelder, Lang, & Buchner, 2007), with a medium effect size of 0.50 (Cohen, 1998) and an alpha level of 0.017, revealed that 43 participants were required to reach a power of 0.85 for the comparison between the involved and uninvolved limbs. A medium effect size was chosen based on the previous study, in which the reliability and validity of a jump-landing-rebound task was examined(Padua et al., 2009). Injured participants were recruited primarily from the XXXXXX. Potential participants were seen at their return-to-sport assessment once they were released by their orthopedic surgeon. Release by their surgeon was based on the results of clinical examination, including assessment of knee active range of motion (ROM), patellar and anterior interval mobility, swelling or pain, as well as assessments of passive hip and ankle ROM, hip and knee strength and lower extremity balance. A cut off score of 90% symmetry for all strength measures was considered acceptable; however, participants who scored below this were still considered for clearance if all other metrics exceeded the 90% threshold.

Injured athletes were considered for inclusion in this study if they: 1) were between 13 and 25 years of age, 2) had injured their ACL for the first time and underwent surgical reconstruction, 3) were involved in a level-1 sport (e.g., basketball, football, or soccer) or level-2 sport (e.g., baseball, racket sports, or skiing) which must include activities such as jumping, pivoting, or hard cutting for greater than 50 h a week (Daniel et al., 1994), 4) were in the return-to-sport rehabilitation stage of their treatment, which is typically five to eight months post-surgery depending on whether other structures (e.g., meniscus, articular cartilage, collateral ligaments, etc.) were involved. Injured athletes were excluded from this study if they: 1) injured their ACL more than once, 2) had a full-thickness chondral defect of 1 cm^2 or greater, 3) had a grade II or III medial or lateral collateral ligament sprain, 4) had a grade III posterior cruciate ligament tear, 5) had a simultaneous bony fracture with ACL tear, or 6) were not planning to return to sport after their ACL-R.

2.1. Instrumentation

An 8-camera 3-dimensional (3D) motion capture system (Qualisys AB, Göteborg, Sweden) with a capture rate of 120 Hz was used to capture joint motions in all three planes during the Vail Sport Test™. Two force plates (Advanced Mechanical Technology, Inc., Watertown, MA) were used during data collection to allow accurate time sequencing during data collection and processing, thus enabling identification of initial contact during the jumping tasks. The International Knee Documentation Committee short form (IKDC) was used to serve as a reference standard to examine the validity of the Vail Sport Test™. The IKDC is a standard measure used for assessing outcomes of patients following ACL-R surgery and has been shown to be reliable and to have acceptable psychometric properties(Collins, Misra, Felson, Crossley, & Roos, 2011; Crawford, Briggs, Rodkey, & Steadman, 2007; Grevnerts, Terwee, & Kvist, 2015; Irrgang, Ho, Harner, & Fu, 1998).

2.2. Procedures

Prior to the Vail Sport Test™, each participant performed a self-selected warm-up (e.g., stationary biking, elliptical, and gluteus muscle activation exercises) for about 10 min. Next, 33 reflective markers were affixed to their skin over the body landmarks in order for cameras to capture joint movements during the testing. For consistency makers were placed by the same two testers for all subjects. Marker locations included bilateral acromions, sternum, C7, T12, L5, bilateral anterior superior iliac crests, bilateral posterior superior iliac crests, bilateral superior sacral poles, inferior sacrum, bilateral greater trochanters, bilateral mid-thighs, bilateral medial and lateral femoral condyles, bilateral mid tibias, bilateral medial and lateral malleoli, bilateral first and fifth metatarsal heads, and bilateral calcanei (Mandengue et al., 2005).

Next, all participants were asked to complete each component of the Vail Sport Test™ in the following order: single leg squat, lateral bounding/agility, forward running and backward running. The Vail Sport Test was administered following the testing protocol described in previously published studies(Garrison et al., 2012). Briefly, a SportCord® resistance band (STI, Baton Rouge, LA) was used to provide resistance for the single leg squat test (Fig. 1a). Participants who weighed greater than 72 kg used a black (heavy resistance) SportCord® resistance band and those who weighed less than 72 kg used a blue (medium resistance) SportCord® resistance band. Participants held one end of the band and the other end wrapped around their foot to secure the band in a taut position. This starting position was standardized to ensure that the participants squatted against resistance. If necessary, participants could use two fingers of their hand on the uninvolved side to balance themselves. They performed 3 min of continuous squatting on their injured limb (Fig. 1a). Following a 2.5-min rest period they then completed the single leg squat test on their uninvolved limb (Garrison et al., 2012).

The lateral bounding component (Fig. 1band c) involved the

Fig. 1a. Single leg squat.

Fig. 1b. Lateral bounding starting position.

participant performing a lateral hopping motion against the resistance of a SportCord® resistance band. The injured leg was positioned as the inside leg or the leg closest to the fixation of the band. The participant was instructed to hop from one leg to the other in a leg-length distance, in which the leg length was measured from the participant's greater trochanter to the floor. The participant performed this lateral bounding test on the injured limb for 90 s and then on their uninvolved limb, with a 2.5-min rest period between limbs.

As with the lateral bounding, a SportCord® resistance band was used to provide resistance for both the forward and backward running (Fig. 1dand e). The participant was instructed to hop from one leg to the other in an up-and-down manner (similar to jogging in place) with the knees flexed between 30° and 60°. The participant ran in each direction for 2 min with a 2.5-min rest period between the two directions. A subset of participants were asked to return 2–7 days later to repeat the Vail Sport test™ to determine the between-day test-retest reliability.

The Vail Sport test™ was scored following the previously published criteria (Garrison et al., 2012). The grading criteria included assessment of technique for each component and was based on a binary scoring system (*yes* = 1, *no* = 0). One point was given for each standard completed with proper form during the set time intervals of each of the four testing components. The total possible number of points for the Vail Sport Test™ ranges from 0 to 54. A patient post ACL-R was required to score at least 46 out of 54 points in order to receive a passing score (Garrison et al., 2012). For each testing component, the participants received no points if they continued to perform with an incorrect movement pattern despite having received verbal feedback on three consecutive repetitions within the testing time interval (Garrison et al., 2012). Dueto the length of the testing period, 3D motion data for the final 10 s of each 30-s interval was collected and used for data processing.

Fig. 1c. Lateral bounding ending position.

Fig. 1d. Forward jogging.

2.3. Kinematic data processing and statistical analyses

For the kinematic data of each 10-s time period, the maximum value was the average of the peak value and the values extracted from two frames before and two frames after the peak value. Trunk flexion greater than 30° from the participant's starting position was considered excessive. Greater than 10 degrees of maximum knee frontal plane projection angle was considered excessive knee valgus. Greater than 0 degrees of knee extension from the starting position was considered excessive for the knee extension grading component. Sagittal plane knee kinematics for knee flexion was used to grade knee flexion during each test component. A knee flexion angle of less than 30° was considered a deduction for that test component. The tibial tuberosity marker was compared to the toe marker and this difference was used to assess for excessive anterior tibial translation. Any value in which the tibial tuberosity marker exceeded the toe marker coordinate by greater than 0.03 m was considered a failed test and a score of zero was given. When a joint motion was graded excessive, a "0" was given for that test component.

IBM SPSS statistics 19 (IBM Corp., Armonk, NY) was used for statistical analysis. Means and standard deviation were calculated for participants' characteristics (i.e., age, gender, limb dominance, operated limb, sport, and IKDC scores), as well as ROM and strength data at the time of the testing for this current study. Paired *t*-tests were used to assess the difference in the ROM and strength measurements between the involved and uninvolved limbs, with the α set at 0.05. Non-parametric signed rank tests were performed to analyze the collected data because the normality assumption was not met. The intraclass correlation coefficient ($ICC_{3,1}$) was used to determine the between-day test-rest reliability. To determine construct validity, two separate signed rank tests were used to compare the scores collected in real-time with those determined from the 3D kinematic data for the involved and uninvolved limbs, respectively. In order to further determine construct validity, one additional signed rank test was used to compare the scores collected in real-time with the IKDC scores for the involved limb. Because these two sets of scores were measured on different scales, Z-scores were computed from the raw scores for the comparison analysis. The alpha level was set at 0.0167 after correction as three

Fig. 1e. Backward jogging.

signed rank tests were performed. To determine the association between the three sets of data as mentioned above with $p < 0.01$, three separate Pearson correlation coefficients (r) were calculated.

3. Results

3.1. Participants

Sixty-six patients who had ACL-R and were referred by their surgeons for a return-to-sport assessment were screened for eligibility for the study from October 2016 to December 2017. Ten patients were excluded from the study because they previously underwent ACL-R. Seven eligible patients declined to participate in the study. Consequently, 49 eligible participants were enrolled in the study. The 49 subjects were referred by four different orthopedic surgeons; with one surgeon referring 40 participants, one surgeon referring 7, and two additional surgeons each referring 1 participant. Of the 49 enrolled participants, 48 completed the study. One participant was asked to discontinue the study because of an inability to keep markers attached to the participant's skin due to excessive sweating. The characteristics of the participants are summarized in Table 1, including age, gender, limb dominance, operated limb, sport, and IKDC scores. Table 2 includes ROM and strength measurements of the lower extremities of the participants. Significant differences were found between the involved and uninvolved limbs in the ROMs of knee flexion and extension, and strength of the quadriceps and hamstrings muscles.

3.2. Validity

Table 3 lists the scores determined in real-time and obtained from 3D kinematic data. The result showed no significant difference between these two scores on the involved limb ($p = 0.222$), but a significant difference on the uninvolved limb ($p = 0.015$). Pearson's correlation coefficients showed a significant moderate correlation for the two sets of scores (real-time vs 3D) on the involved limb, with $r = 0.55$ ($p < 0.001$) and on the uninvolved limb ($r = 0.46$, $p = 0.001$). The comparison analysis showed no significant difference between the scores collected in real-time and the IKDC scores, for the involved limb ($p = 0.885$), with a non-significant weak correlation ($r = 0.20$, $p = 0.174$) between the two sets of the scores.

3.3. Reliability

Fourteen participants returned 2–7 days later for a second testing session to determine the between-day test-retest reliability. The demographic information for this subset of patients as well as their Vail Sport Test™ scores are listed in Tables 4 and 5. Intraclass correlation coefficients ($ICC_{3,1}$) showed good inter-session reliability for the involved limb with the $ICC_{3,1}$ being 0.787 and 95% CI (confidence level) ranging from 0.459 to 0.926 Similarly, the results showed fair inter-session reliability for the uninvolved limb with $ICC_{3,1}$ being 0.485 and 95% CI ranging from −0.038 to 0.800. To further illustrate the reliability, a Bland-Altman plot was created to show the limits of agreement of the reliability data for both the involved limb (see Fig. 2) and the uninvolved limb (see Fig. 3).

4. Discussion

4.1. Construct validity

Using the reference standard of the 3D kinematic analysis, the results appeared to partially support the use of the Vail Sport Test™ as a measure for return-to-sports decision for athletes following ACL-R. This discrepancy of the findings between the limbs could be explained partially by inherent bias of the rater. The Vail Sport Test™ was designed to assess readiness to return to sports of the involved limbs (Garrison et al., 2012). However, the performance of both limbs was assessed in the current study with an assumption that the rater would grade both limbs at the same time equally in real-time. Therefore, the rater's past experiences of only grading the involved limb could have impacted the real-time scores of the study, specifically regarding the results of the uninvolved limb. In addition, considering that the majority of the participants (44 out of 48) passed the Vail Sport Test™ and scored between 46 and 54 on the test, this narrow range of scores may have contributed to the near-significant difference and the moderate correlation between the two sets of the scores. However, despite the findings that do not support the validity of the Vail Sport Test™ on the uninvolved limb, this should not diminish the overall validity of the test, as the tests

Table 1
Participant Characteristics of the Study.

	All ($n = 48$)	Pass Group ($n = 44$)	Fail Group ($n = 4$)
Age (years)	16.7 ± 1.5	16.9 ± 2.7	15.6 ± 0.8
Height (cm)	168.9 ± 10.4	168.8 ± 12.4	162.5 ± 7.8
Weight (kg)	68.0 ± 9.38	67.3 ± 12.8	68.1 ± 6.9
Sex	Women: 30	Women: 29	Women: 1
	Men: 18	Men: 15	Men: 3
Months post-surgery	7.0 ± 1.2	7.1 ± 1.8	7.0 ± 1.0
Concomitant injury	Meniscus Repair: 13	Meniscus Repair: 11	Meniscus Repair: 2
	Meniscectomy: 10	Meniscectomy: 9	Meniscectomy: 1
	None: 25	None: 24	None: 1
Mechanism of Injury	Direct: 10	Direct: 9	Direct: 1
	Indirect: 12	Indirect: 11	Indirect: 1
	Non-Contact: 26	Non-Contact: 24	Non-Contact: 2
Limb Dominance	Right: 45	Right: 43	Right: 2
	Left: 3	Left: 1	Left: 2
Injured Limb	Right: 21	Right: 21	Right: 0
	Left: 27	Left: 23	Left: 4
Sport			
Basketball	15	14	1
Football	10	9	1
Soccer	18	16	2
Volleyball	3	3	0
Softball	1	1	0
Cheerleading	1	1	0
IKDC	91.8 ± 8.2	91.3 ± 7.4	88.5 ± 6.4

Note: IKDC = International Knee Documentation Committee Subjective Knee Evaluation Form.

Table 2
Range of Motion and Strength Measurements of Lower Extremities of the Participants ($n = 48$).

	Involved	Uninvolved	*P* Value
AROM (°)			
Knee Flexion	139.5 ± 8.4	141.1 ± 8.5	0.040*
Knee Extension	1.6 ± 2.6	2.7 ± 1.9	0.003*
PROM (°)			
Hip Internal Rotation	44.0 ± 9.5	41.2 ± 8.0	0.060
Hip External Rotation	41.9 ± 7.5	41.5 ± 6.8	0.700
Ankle Dorsiflexion	40.4 ± 7.4	40.3 ± 7.4	0.340
Strength (kg)			
Hip Abduction	25.9 ± 5.2	24.6 ± 5.9	0.160
Hip External Rotation	20.5 ± 4.6	21.2 ± 4.8	0.360
Quadriceps(peak torque at 60°/sec)	72.7 ± 26.0	100.7 ± 34.9	<0.001*
Hamstring(peak torque at 60°/sec)	47.5 ± 18.1	49.8 ± 18.1	0.003*

Note: AROM = active range of motion; PROM = passive range of motion; *$p < 0.05$.

Table 3
Means and Standard Deviations of the Vail Sport Test™ Scores Collected Visually in Real-time and Obtained by Analyzing Post-capture 3D Kinematic Data.

	Real-time Data	Post-Capture kinematic data	*p* value
Involved Limb	49.3 ± 3.4	50.5 ± 2.8	0.013
Uninvolved Limb	48.2 ± 6.3	50.2 ± 3.3	0.006*

Note. 3D = 3-dimensional. *$p < 0.01$.

original intent was to assess the involved limb's readiness to return to sport.

3D motion analysis has been used previously as a reference standard to assess the validity of a visual movement screen (Padua et al., 2009). Padua et al. (2009) assessed the criterion validity of the Landing Error Scoring System (LESS) as a clinical assessment tool of jumping-landing biomechanics by comparing the scores graded by an expert to those obtained from 3D motion analysis. These authors found excellent agreements for the grading of ankle dorsiflexion, knee flexion ROM, trunk flexion at maximum knee flexion and foot position at initial contact, as well as moderate agreements for trunk flexion and knee valgus at initial contact and knee valgus ROM

Table 4
Characteristics of the Participants in the Reliability Part of the Study ($n = 14$).

Age (years)	15.8 ± 1.1
Height (cm)	166.3 ± 8.1
Weight (Kg)	64.2 ± 8.9
Sex	Women: 11
	Men: 3
Months post-surgery	7.1 ± 0.5
Days Between Testing Sessions	6.1 ± 0.5
Mechanism of Injury	Direct: 6
	Indirect: 1
	Non-Contact: 7
Limb Dominance	Right: 14
	Left: 0
Injured Limb	Right: 4
	Left: 10

Test™ Scores (Means and Standard Deviations) Collected Visually on Two Separate Sessions and Intraclass Coefficient Coefficients (ICC) for the Between-day Test-retest Reliability.

Table 5
Vail Sport Test™ scores for the inter-session reliability subjects.

	Session 1	Session 2	$ICC_{(3,1)}$	*p* value
Involved Limb	50.7 ± 1.8	49.7 ± 2.3	0.787	0.001
Uninvolved Limb	50.5 ± 1.9	49.4 ± 2.8	0.489	0.033

(Padua et al., 2009). Padua et al. (2009) also found that participants with poor jumping techniques (i.e., high LESS scores) displayed different kinematics and kinetics of lower extremities from those with excellent jumping techniques. Like the Padua et al. study (2009), a 3D motion analysis system was used in the current study to assess knee motions in the sagittal and frontal planes during jumping and landings tasks. The findings of the current study agreed with those of the Padua et al. study, therefore further supporting the use of 3D motion analysis as a reference standard for validating visual assessment of dynamic lower extremity movements.

The results of this study also showed no significant difference

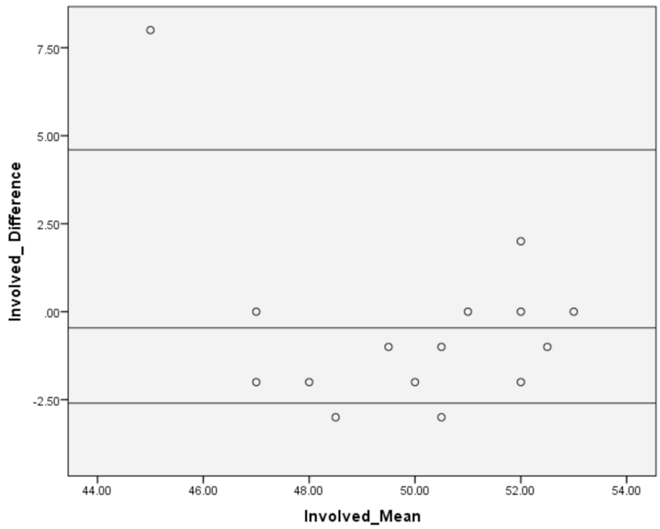

Fig. 2. Bland-Altman plot for limits of agreement of the inter-session reliability data for the involved limb.

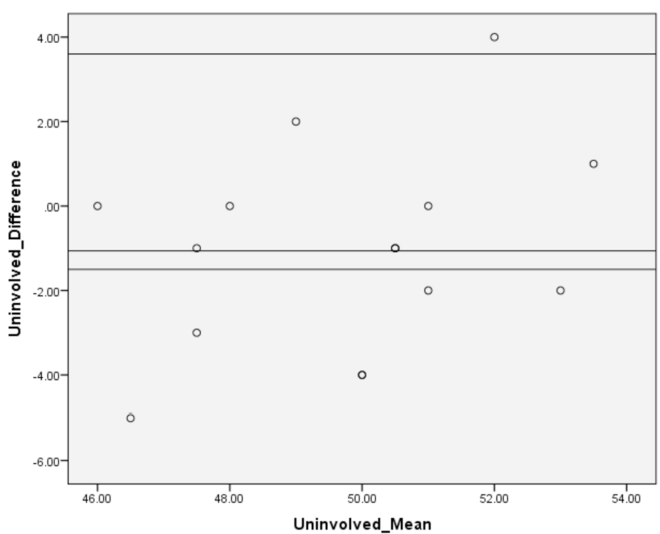

Fig. 3. Bland-Altman plot for limits of agreement of the inter-session reliability data for the uninvolved limb.

between the Z scores of the involved limbs Vail Sport Test™ scores and those of the IKDC scores. However, we found a weak correlation ($r = 0.20$) between the two sets of scores. The IKDC is a well-established outcome measure in this population and improved scores on the IKDC have been shown to be related to improved performance on clinical measures (Collins et al., 2011; Reinke et al., 2011). However, the IKDC is meant to assess the patients' perceived functional levels of the limbs of participants with knee disorders, whereas the Vail Sport Test™ was scored for involved and uninvolved limbs separately. This discrepancy may have impacted the results. As most of the participants had high passing scores on the Vail Sport Test indicating a high level of function, the fair correlation between these scores was not surprising. Further, it is also speculated that the relationship may not be linear in the upper quartile of the Vail Sport Test™ scores.

In addition, the high number of the participants who passed the test could be in part due to participant selection bias. The participants who were eligible to participate in this study were cleared by their surgeon and physical therapist to complete return to sport testing. At the time of the testing, the treating surgeon and physical therapist appeared to demonstrate confidence that the selected patients were ready for jumping and cutting tasks based on their overall clinical presentations, such as their pain, ROM, strength and balance. This indicated that the surgeon's and physical therapist's experience and clinical judgement for returning athletes to sports are intuitive and consistent with the results of the Vail Sport Test™. Nonetheless, failure of four participants to pass the Vail Sport Test™ suggests the inadequacy of relying solely on surgeon's and physical therapies' clinical judgement to determine an athletes' readiness to return to sports following ACL-R.

4.2. Reliability

The results showed good inter-session reliability for the involved limb and fair inter-session reliability for the uninvolved limb. These results are in agreement with those of previously published research examining both the inter-rater and intra-rater reliability of the Vail Sport Test™(Garrison et al., 2012). In the Garrison et al. (2012) study, excellent inter-rater and intra-rater reliability was reported, with the ICCs being 0.95 and 0.97, respectively. The between-day test-retest reliability value (ICC = 0.787 for the involved limb) in this study was expected to be lower than the ICC values reported by Garrison et al. (2012) because the reliability was established by the graders watching a video-taped test in the Garrison et al. study, whereas in this study, participants performed the test twice within a short period of time. It is worth noting that the participants in the reliability part of this study were those who self-selected to return for additional testing. This may have skewed the results because those participants who were not as challenged by the test may have been the participants who chose to return to test again. However, given the high overall pass rate in this study, it is likely that the reported ICC values in this study would remain high regardless of which participants returned to test.

Interestingly, the ICC value for the uninvolved limb was only 0.48. The poorer reliability on the uninvolved limb may be due to the grader's bias, as more detail is typically given to the involved limb during testing as discussed earlier. During the bilateral movement tasks (forward and backward jogging) the grader was asked to assess both limbs at the same time. This is not how the original test is described. However, given the difficulty of the test and the clinical nature of the study we could not ask the participants to complete the test a second time to create two separate scores. Additionally, the inter-session reliability has previously been established for the involved limb (Garrison et al., 2012) which we believe further supports the notion that the poorer reliability seen on the uninvolved limb is likely due to rater bias.

4.3. Limitations

The results of this study should be interpreted in light of several limitations. First, the criterion that the participants be cleared by the surgeon for return-to-sport testing, could have resulted in a non-normal (i.e., positively skewed) distribution of the Vail Sports Test™ scores, leading to notably unequal sizes of the "pass" and "fail" groups. Further, there may have been additional patients who would have passed but were not tested and vice versa. It is uncertain whether those patients who were referred to the return-to-sport test but declined to participate in the study would have passed the test. Another limitation of this study is that the rater was asked to grade both limbs simultaneously during the forward and backward jogging portion of the test. The Vail Sport Test™ was originally designed to grade only the involved limb. Grading of both limbs simultaneously is a challenging task for the rater. Given the good construct validity that was found between the real-time rating and post-capture grading of 3D kinematic data and the acceptable inter-session reliability for each limb, it is likely that the rater was able to successfully complete this task. However, this might explain why the inter-session reliability for the uninvolved limb was poorer than that of the involved limbs. If the rater was unable to appropriately perform a dual task, they may have focused their efforts on grading the involved limb rather than the uninvolved limb, as this is typically the limb of interest.

Lastly, only a portion of the collected 3D kinematic data was used for the post-capture grading. In an ideal situation, the complete time of the testing would have been captured and graded. However, due to the large amount of data that is captured, it was not feasible to collect the entire testing session. As such, a 10-s window was chosen to allow both appropriate grading and feasible data management. However, given that the Vail Sport Test™ scores determined by the 3D kinematic data were no different from the real-time test results, this data analysis approach seemed to be appropriate.

5. Conclusion

The results of this study established the construct validity and demonstrated good inter-session reliability of the Vail Sport Test™, supporting the use of the Vail Sport Test™ as a measure of readiness to return to sports for patients following ACL-R. The failure on the Vail Sport Test™ of some participants who were cleared to return to sports by surgeons suggested the need of an adequate return-to-sport test, such as Vail Sport Test™. A follow-up study to examine whether the participants who passed the Vail Sport Test™ successfully return to their sports and do not have recurrent ACL injury could further validate the Vail Sport Test™. Additionally, it may be worth examining the difference between expert and novice raters, as experience of the rater may influence the results of the test.

Conflicts of interest

None to declare.

Ethical approval

This study was approved by appropriate ethical committees related to the institution(s) in which it was performed and that subjects gave informed consent to the work.

Funding

This study was funded by the Texas Physical Therapy Foundation (PRO0004886).

Acknowledgements

We would like to acknowledge the research coordinators Angel Grondin, Kalyssa Creed, and Laura Giesler.

References

Ardern, C. L., Webster, K. E., Taylor, N. F., & Feller, J. A. (2011a). Return to sport following anterior cruciate ligament reconstruction surgery: A systematic review and meta-analysis of the state of play. *British Journal of Sports Medicine, 45*(7), 596–606. bjsports76364.

Ardern, C. L., Webster, K. E., Taylor, N. F., & Feller, J. A. (2011b). Return to the pre-injury level of competitive sport after anterior cruciate ligament reconstruction surgery two-thirds of patients have not returned by 12 months after surgery. *The American Journal of Sports Medicine, 39*(3), 538–543.

Augustsson, J., Roland, T., & Karlsson, J. (2004). Ability of a new hop test to determine functional deficits after anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy, 12*(15), 350–356.

Barber-Westin, S. D., & Noyes, F. R. (2011). Factors used to determine return to unrestricted sports activities after anterior cruciate ligament reconstruction. *Arthroscopy, 27*(12), 1697–1705.

Bien, D. P., & Dubuque, T. J. (2015). Considerations for late stage ACL rehabilitation and return to sport to limit Re-injury risk and maxamize athletci performance. *Internacional Journal Sports Physical Therapy, 10*(2), 256.

Butler, R. J., Lehr, M. E., Fink, M. L., Kiesel, K. B., & Plisky, P. J. (2013). Dynamic balance performance and noncontact lower extremity injury in college football players an initial study. *Sport Health: A Multidisciplinary Approach, 5*(5), 417–422.

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.

Collins, N. J., Misra, D., Felson, D. T., Crossley, K. M., & Roos, E. M. (2011). Measures of knee function: International knee documentation committee (IKDC) subjective knee evaluation form, knee injury and osteoarthritis outcome score (KOOS), knee injury and osteoarthritis outcome score physical function short form (KOOS-PS), knee outcome survey activities of daily living scale (KOS-adl), lysholm knee scoring scale, oxford knee score (OKS), western ontario and McMaster universities osteoarthritis index (WOMAC), activity rating scale (ARS), and tegner activity score (TAS). *Arthritis Care & Research, 63*(S11), S208–S228.

Crawford, K., Briggs, K. K., Rodkey, W. G., & Steadman, J. R. (2007). Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. *Arthroscopy, 23*(8), 839–844.

Daniel, D. M., Stone, M. L., Dobson, B. E., Fithian, D. C., Rossman, D. J., & Kaufman, K. R. (1994). Fate of the ACL-injured patient a prospective outcome study. *The American Journal of Sports Medicine, 22*(5), 632–644.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Fitzgerald, G. K., Lephart, S. M., Hwang, J. H., & Wainner, M. R. S. (2001). Hop tests as predictors of dynamic knee stability. *Journal of Orthopaedic & Sports Physical Therapy, 31*(10), 588–597.

Garrison, J. C., Shanley, E., Thigpen, C., Geary, R., Osler, M., & DelGiorno, J. (2012). The reliability of the Vail sport Test™ as a measure of physical performance following anterior cruciate ligament recosntruction. *Internacional Journal Sports Physical Therapy, 7*(1), 20.

Grevnerts, H. T., Terwee, C. B., & Kvist, J. (2015). The measurement properties of the IKDC-subjective knee form. *Knee Surgery, Sports Traumatology, Arthroscopy, 23*(12), 3698–3706.

Gribble, P. A., Hertel, J., & Plisky, P. (2012). Using the star excursion balance test to assess dynamic postural-control deficits and outcomes in lower extremity injury: A literature and systematic review. *Journal of Athletic Training, 47*(3), 339–357.

Grindem, H., Snyder-Mackler, L., Moksnes, H., Engebretsen, L., & Risberg, M. A. (2016). Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: The Delaware-oslo ACL cohort study. *British Journal of Sports Medicine, 50*(13), 804–808. bjsports-2016-096031.

Harris, J. D., Abrams, G. D., Bach, B. R., Williams, D., Heidloff, D., Bush-Joseph, C. A., & Cole, B. J. (2014). Return to sport after ACL reconstruction. *Orthopedics, 37*(2), e103–e108.

Irrgang, J. J., Ho, H., Harner, C. D., & Fu, F. H. (1998). Use of the International Knee Documentation Committee guidelines to assess outcome following anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy, 6*(2), 107–114.

Kokmeyer, D., Wahoff, M., & Mymern, M. (2012). Suggestions from the field for return-to-sport rehabilitation following anterior cruciate ligament reconstruction: Alpine skiing. *Journal of Orthopaedic & Sports Physical Therapy, 42*(4), 313–325.

Mandengue, S. H., Seck, D., Bishop, D., Cisse, F., Tsala-Mbala, P., & Ahmaidi, S. (2005). Are athletes able to self-select their optimal warm up? *Journal of Sports Science and Medicine, 8*(1), 26–34.

Padua, D. A., Marshall, S. W., Boling, M. C., Thigpen, C. A., Garrett, W. E., & Beutler, A. I. (2009). The landing error scoring system (LESS) is a valid and reliable clinical assessment tool of jump-landing biomechanics the JUMP-ACL study. *The American Journal of Sports Medicine, 37*(10), 1996–2002.

Reinke, E. K., Spindler, K. P., Lorring, D., Jones, M. H., Schmitz, L., Flanigan, D. C., & Martin, M. (2011). Hop tests correlate with IKDC and KOOS at minimum of 2 years after primary ACL reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy, 19*(11), 1806–1816.

Wiggins, A. J., Grandhi, R. K., Schneider, D. K., Stanfield, D., Webster, K. E., & Myer, G. D. (2016). Risk of secondary injury in younger athletes after anterior cruciate ligament reconstruction a systematic review and meta-analysis. *The American Journal of Sports Medicine, 44*(7), 1861–1876, 0363546515621554.

Xergia, S. A., Pappas, E., Zampeli, F., Georgiou, S., & Georgoulis, A. D. (2013). Asymmetries in functional hop tests, lower extremity kinematics, and isokinetic strength persist 6 to 9 months following anterior cruciate ligament reconstruction. *Journal of Orthopaedic & Sports Physical Therapy, 43*(3), 154–162.