IJSPT

ORIGINAL RESEARCH

INTER-RATER AND TEST-RETEST RELIABILITY OF TWO NEW SINGLE LEG SIT-TO-STAND TESTS

Andy Waldhelm,[1]
Coral Gubler,[1]
Katie Sullivan,[1]
Chris Witte,[1]
Devin Buchheister,[1]
Jacob Bartz-Broussard[1]

ABSTRACT

Background: The single leg sit-to-stand test (SLSTST) is a functional test that assesses quadriceps strength. The original SLSTST was used to diagnose lumbar nerve root impingement/radiculopathy – specifically at the L3 and L4 level. The original SLSTST used one repetition as the requirement for a successful test, therefore it may not identify quadriceps weakness in highly functional individuals with or recovering from an athletic injury.

Purpose/Hypothesis: The purpose of this study was to determine the interrater and test-retest reliability of two new SLSTSTs, one for maximum number of repetitions over 30 seconds and one for time to complete five repetitions.

Study Design: Cross-sectional, reliability study

Methods: Twenty healthy college-aged individuals (12 males, age: 22.5 years ± 1.37, height: 1.72 m ± 0.09; weight: 70.2 kg ± 11.0) participated in the study. Two testing sessions were held three to seven days apart, and two second-year physical therapy students served as examiners. The objective of the 30-second SLSTST was for the participant to perform as many single leg sit-to-stand repetitions they could in thirty seconds, while the five repetitions SLSTST measured how quickly the subjects could perform five single leg sit to stand repetitions. Both lower extremities were tested and Intraclass Correlation Coefficients (ICC) were calculated to determine reliability.

Results: Both SLSTSTs were found to have excellent interrater and good to excellent test-retest reliability. The 30-second SLSTSTs had inter-rater ICC = 0.99 on the right and 0.98 on the left while the test-retest ICCs ranged from 0.92 to 0.94. The five repetition SLSTSTs had an inter-rater ICC = .99 on both legs while the test-retest ICC ranged from 0.87 to 0.94.

Conclusions: The results of the current study indicate that the two new SLSTSTs had good to excellent test-retest and excellent inter-rater reliability. However, more research is needed to determine if SLSTSTs can be used to identify quadriceps weakness in individuals with recovering from an athletic injury or to be used as a return-to-sport (RTS) assessment.

Levels of Evidence: Level 2

Key Words: Functional test, Movement System, quadriceps weakness, return to sport assessment, single leg sit to stand

CORRESPONDING AUTHOR
Andy Waldhelm, PT, PhD, SCS, CSCS
Department of Physical Therapy
University of South Alabama
5721 USA Drive North
Mobile, AL 36688
Phone: 225-328-3892
Email: awaldhelm@southalabama.edu

[1] University of South Alabama, Department of Physical Therapy, Mobile, AL, USA

Conflicts of Interest: None.

DOI: 10.26603/ijspt20200388

INTRODUCTION

Anterior cruciate ligament (ACL) sprains are common among athletes with up to 250,000 injuries occurring per year in the United States.[1,2] ACL injuries typically occur in athletes who participate in sports which require running, jumping and cutting maneuvers and can occur with or without contact.[3,4] Typically athletes who undergo surgery to reconstruct the ACL return to sport six to 15 months post-surgery, with a mean return of 13 months.[3] Unfortunately, only about one-half of athletes who had an ACL reconstruction fully return to sport after one year.[3, 5] Furthermore, up to one-third athletes returning to sport after an ACL reconstruction have a second ACL injury within two years[6] which is up to six times higher than athletes who had not suffered an ACL injury.[7]

In order to reduce the risk of a second ACL injury in athletes, return-to-sport (RTS) criteria have been developed.[5,8] RTS assessments consist of an arsenal of tests designed to determine when an athlete is able to safely return to their sport with the ability to perform at or near pre-injury level.[5,8,9] RTS assessments calculate and use a symmetry indices utilizing the non-injured limb as a control to determine residual functional deficit of the injured side.[10] RTS criteria can include subjective patient reports as well as objective measures such as neuromuscular control, strength, and functional performance in sport related activities.[11-15] The assessments used in the RTS criteria can include strength tests of the quadriceps and hamstrings, single leg hop tests, single leg squat, in-line lunge, and various agility tests.[14] Each of these facets are equally important when the RTS decision is made, but there is disagreement on which assessments should be used and if these assessments are valid.[5] Furthermore, there is limited evidence that passing the RTS criteria can decrease the risk of a second ACL or knee injury.[5]

Quadriceps and other lower extremity muscle weakness is a common impairment following an ACL reconstruction and strength may never fully return to the pre-injury status.[16] Furthermore, quadriceps weakness is related to an increased risk of a secondary ACL injury.[16,17] Therefore, strength assessments that address the quadriceps are a key element in the RTS process.[10,18,19] Muscle strength is commonly measured through isometric, isotonic or isokinetic testing.[10,16]

Isokinetic assessments are considered the "gold standard" for strength testing.[20] The measurement is performed at a constant angular velocity and produces objective measurements of torque, work, and power.[20-22] Isokinetic machines make great tools to assess muscle strength as part of a RTS battery because there is currently no universally-accepted criteria, but they are often impractical for clinical use as they are expensive, occupy a large area, and they do not measure functional muscle performance.

The single leg sit-to-stand test (SLSTST) is a functional test used to assess quadriceps strength.[23,24] In individuals with lower extremity radicular pain, the SLSTST was used to diagnose mid-lumbar nerve root impingement and radiculopathy with a sensitivity of 0.48, a specificity of 0.90, a positive likelihood ratio of 4.60 and a negative likelihood ratio of 0.58.[24] The SLSTST requires a patient to lift their entire body weight with one leg from a chair.[23,24] Unlike isokinetic machines, this test can be easily performed with a standard 18 in (45.7 cm) dining room chair and a timing device being the only required equipment. Although the single leg squat tests do not isolate the quadriceps and muscle activation of the lower extremity muscles during the single leg sit to stand test has not been studied, several authors have found the quadriceps to be highly active during a single leg squat.[25-27] In a systematic review by Dedinsky et al,[27] percent of maximal voluntary isometric contraction of the quadriceps ranged from 33.9% to 116.2%, and a study by Zeller et al.[26] found the vastus lateralis to be most active when performing a single leg squat as compared to other muscles of the leg, including the hamstrings, gluteus maximus and gluteus medius.

The original test described by Rainville et al.[23] and Suri et al.[24] only called for the participant to perform one repetition of the single leg sit to stand movement, therefore, this version of the test may not be able to identify quadriceps weakness in higher functioning individuals such as athletes. A multiple repetition SLSTST may allow for assessment of power and muscular endurance in individuals with higher functional levels such as athletes. Therefore, the purpose of this study was to determine the interrater

and test-retest reliability of two new SLSTSTs one for maximum number of repetitions over 30 seconds and one for time to complete five repetitions.

METHODS

Design

A cross-sectional reliability research design determined the interrater and test-retest reliability of two novel SLSTSTs.

Participants

Twenty healthy participants (12 males, 8 females; age: 22.5 ± 1.37; height (cm): 172 ± 9.18; weight (kg); 70.2 ± 11.0) volunteered for the study. Inclusion criteria included individuals between the ages of 18 and 40, who exercised at least one time per week for at least 30 minutes. Potential participants were excluded if they had low back or lower extremity injury within the past three months which held them out of regular activity for more than three days. All participants wore their own exercise/athletic clothes and shoes during the study. All subjects provided verbal and written consent which was approved by the Institutional Review Board of the University of South Alabama before they started the study.

Procedures

Two testing sessions were used for this study. The first testing session included a five-minute warm up in which the participants walked at a self-selected speed. The order of tests and the initial leg tested were block randomized prior to data collection to decrease order effects. Before testing, the researcher demonstrated the desired single leg sit to stand movement and had the participant complete two practice repetitions to demonstrate understanding. For a repetition, the participant started with their knee in 90° of flexion while sitting on an 18 in (45.7 cm) chair and rose to full knee extension on the testing leg without the contralateral leg touching the ground. The participants were allowed to maintain their non-testing leg in a self-selected comfortable position which varied from knee fully extended to knee flexed at nearly 90°. Then the participant descended until their buttocks made contact with an 18 in (45.7 cm) chair (Figures 1 and 2).



Figure 1. *Starting position of the single leg sit to stand tests, with the knee flexed at 90 degrees.*

The participants were allowed to just touch the chair and move on to the next repetition, they did not have to bear weight through their buttocks. The patient performed the test with their arms crossed over their chest. If the participant did not correctly perform the movement (buttocks did not touch the chair, they did not come to full knee extension, or the non-testing leg touches ground, etc.), then the repetition was not counted.

The two SLSTSTs include a timed test to measure muscular endurance (SLST30sec) and a second test for speed over repetitions to assess power (SLST5rep). For the SLST30sec, the participants performed as many repetitions as possible in thirty seconds while the SLST5rep measured how quickly they were able to complete five repetitions. Each leg and test were performed twice with a one-minute rest between each trial and each test. The best score of the two trials for each test and leg was used in data collection. The participants performed a second session of testing three to seven days after the

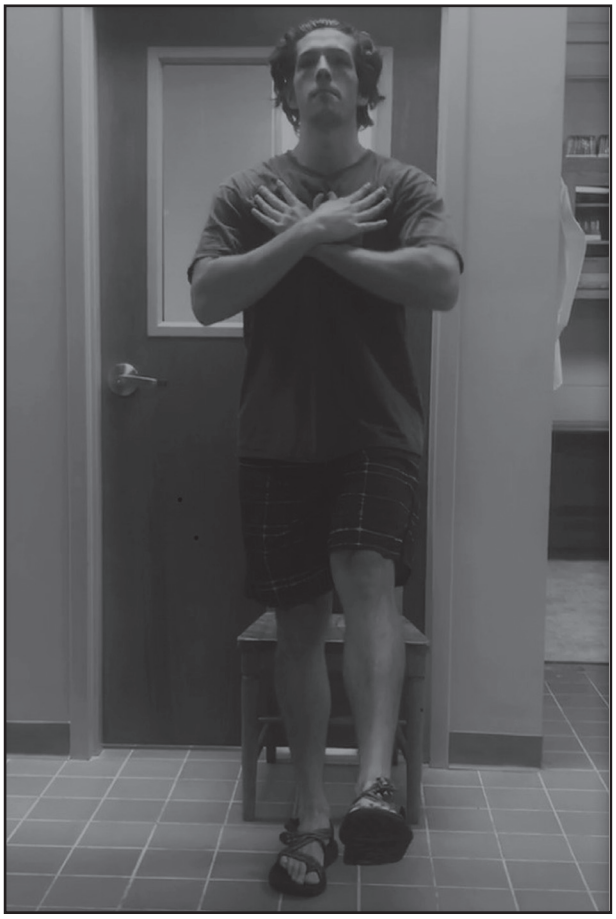


Figure 2. *Finishing position of the single leg sit to stand tests, with the knee fully extended.*

initial testing session following the same protocol and testing order as the first testing session.

Testing was performed in a human performance lab and the raters included two second year physical therapy students. The raters collected the data in real time and independently from each other since they were in different locations within the lab. Another researcher collected the data sheets from each of the two raters and entered the data on a spreadsheet. Inter-rater reliability was measured for both testing days while test-retest reliability was calculated with data obtained on Session 1 and Session 2 for each rater.

Statistical Analysis

SPSS Version 25 software (IBM SPSS Inc., Armonk, MY) was used for data analysis. Interrater and test-retest reliability was determined using Intraclass Correlation Coefficient (ICC) with a p value <0.05. Fisher's classification system for ICC values where: Poor = ICC < 0.5, Moderate = ICC between 0.50 and 0.75, Good = ICC between 0.75 and 0.90 and Excellent = ICC > 0.90, was used.[28]

RESULTS

Table 1 presents the results on each of the tests recorded by each rater on both days. Tables 2 and 3 display the interrater, ICC(3,2), and test-retest, ICC(3,1), reliabilities for each of the tests on both legs. The interrater reliability was much higher than the test-retest with all ICCs > 0.98. The test-retest reliability had ICC 's that ranged between 0.87-0.94. Rater 1 had a slightly higher ICCs on the SLST30sec tests. Rater 2 had higher scores on the SLST5rep tests.

Table 1. *Results (mean ± SD) of two single leg-sit-to-stand tests, maximum repetitions in 30 seconds (SLST30sec) and time to complete five repetitions (SLST5rep), for each test administrator and for both testing sessions, held three to seven days apart.*

	Rater 1	Rater 2
Session 1-SLST30sec, Right (reps)	12.6 ± 3.25	12.5 ± 2.99
Session 1- SLST30sec, Left (reps)	12.4 ± 3.53	12.1 ± 3.45
Session 2- SLST30sec, Right (reps)	14.0 ± 3.61	14.2 ± 3.58
Session 2- SLST30sec, Left (reps)	14.3 ± 3.23	14.1 ± 3.68
Session 1-SLST5rep, Right (s)	9.14 ± 3.79	9.04 ± 3.75
Session 1- SLST5rep, Left (s)	8.56 ± 2.41	8.47 ± 2.42
Session 2- SLST5rep, Right (s)	7.94 ± 2.09	7.86 ± 2.08
Session 2- SLST5rep, Left (s)	8.22 ± 2.49	8.15 ± 2.55

Table 2. *Interrater reliability (ICC $_{(2,1)}$) between test administrators for each Single Leg Sit to Stand Test for both testing sessions.*

	Session 1	Session 2
Right SLST30sec	0.99	0.99
Left SLST30sec	0.98	0.98
Right SLST5rep	0.99	0.99
Left SLST5rep	0.99	0.99

SLST30sec = single leg sit to stand 30 seconds test;
SLST5rep = single leg sit to stand five repetition test

Table 3. *Test-Retest reliability (ICC (2,1)) between testing session held three to seven days apart for each of the Single Leg Sit to Stand Tests for both test administrators.*

	Rater 1	Rater 2
Right SLST30sec	0.94	0.92
Left SLST30sec	0.93	0.92
Right SLST5rep	0.87	0.94
Left SLST5rep	0.87	0.93

SLST30sec = single leg sit to stand 30 seconds test; SLST5rep = single leg sit to stand five repetition test

DISCUSSION

Isokinetic strength testing is the gold standard for strength testing.[20] Isokinetic machines are cost prohibitive and are not readily available in physical therapy rehabilitation clinics. It may be beneficial to have a feasible and functional method to identify quadriceps weakness or imbalance. The purpose of this study was to determine the reliability of two new SLSTSTs which are purported to assesses muscular endurance and power. Both tests demonstrated excellent interrater and good to excellent test-retest reliability, thus, the results of the current study support the original hypothesis.

Previous research on the reliability of sit-to-stand tests has exhibited good to excellent reliability,[30,31] although only one study examined the interrater reliability of a single leg sit-to-stand test.[23] Similar to the Rainville et al. study, the examination of sit-to-stand assessment demonstrated excellent interrater reliability, but their testing methodology was different by only using only one repetition.

The purpose of the study was not to determine the validity of these new tests. Therefore, future research is required to assess if these tests are related to quadriceps muscle weakness, if they can identify athletes who are at risk for injury, and if they can be used as part as a return to sport criteria.

There are limitations to this study, however, an established standard protocol minimized the influence of inconsistent techniques. First, the study's interrater reliability was higher than test-retest reliability, and this was likely due to the timeline of testing as the interrater reliability was collected on the same day and time while test-retest reliability was collected three to seven days apart. Therefore, a learning effect may have decreased the test-retest reliability since all scores for both tests and both legs improved. Furthermore, the testers were of similar background knowledge and experience. If testers had varying levels of knowledge and experience, the test-retest reliability may have been different, but no change in the reliability is expected if experience clinicians had performed the study.

Next, individuals performing the single leg sit-to-stand tests used varying positions of the non-weight bearing limb when performing the test. Some

participants fully extended the non-weight bearing limb in front of them while others had their knee in a flexed position. The participants who had their knee flexed during the study, may have used less effort when performing the test since the non-testing foot is positioned closer to the center of mass. Fortunately, with good to excelled reliability the non-standardized positon of the non-testing leg did not have a major effect on the outcome of the study and therefore, standardizing the non-testing leg positon may not be required. Another limitation could be chair height. An 18 inch (45.7 cm) chair was used and participants started the test with their knee flexed to 90°, but the height of the chair was not standardized to participant's height or leg length or overall. Not observed it this study, but individuals who were extremely short or tall may have had more or less difficulty performing the assessment with a standard chair height, therefore the height of the chair can be standardized to half the individual's leg length.

Although the raters recorded their data individually, they were not in separate rooms, which may have had an influence on the reliability. Despite this, the testers were not allowed to speak to each other about test results and were not made aware of any results until after all testing was concluded to minimize unwanted influence. All information was processed by an experimenter not actively participating in the data recording process. Lastly, the participants used were all healthy and college-aged students; thus, the results can only be generalized to a similar population.

Future studies must include standardizing the non-testing limb position during the SLSTSTs to decrease momentum created by swinging the limb upon standing. Raters with more experience can examine the differences between raters of varying experience levels. A subject population of a wider age range with previous ACL or other lower extremity injuries will provide more generalizable results and will help determine if these new tests are used as an RTS assessment.

CONCLUSION

The results of the current study indicate that two new SLSTSTs, SLST30sec and SLST5rep, demonstrate excellent interrater and good to excellent test-retest reliability. More research is needed to determine if these tests are related to quadriceps weakness by assessing the relationships between the SLSTSTs and isokinetic knee extension strength tests. Lastly and more importantly, future studies must assess if these new SLSTSTs can effectively identify strength imbalance in athletes who have undergone ACL reconstruction surgery or other lower extremity injures and if they can be used as part of RTS criteria or testing in order to help clinicians with RTS decision making.

REFERENCES

1. Sorenson B. Kernozek TW, Willson JD, Ragan R. Hove J. Two-and three-dimensional relationships between knee and hip kinematic motion analysis: single-leg drop-jump landings. *J Sport Rehabil.* 2015;24(4):363-372.
2. Redler LH, Watling JP, Dennis ER, Swart E, Ahmad CS. Reliability of a field-based drop vertical jump screening test for ACL injury risk assessment. *Phys Sportsmed.* 2016;44(1):46-52.
3. Ardern CL, Webster KE, Taylor NF, Feller JA. Return to the preinjury level of competitive sport after anterior cruciate ligament reconstruction surgery: two-thirds of patients have not returned by 12 months after surgery. *Am J Sports Med.* 2011; 39(3):538-543.
4. Montalvo AM, Schneider DK, Webster KE, et al. Anterior cruciate ligament injury risk in sport: a systematic review and meta-analysis of injury incidence by sex and sport classification *J Athl Train*. 2019;54(5): 472-482.
5. Webster KE, Hewett TE. What is the evidence for and validity of return-to-sport testing after anterior cruciate ligament reconstruction surgery? a systematic review and meta-analysis. *Sports Med*. 2019;49(6):917-929.
6. Paterno MV, Schmitt LC, Thomas S, Duke N, Russo R. Quatman-Yates CC. Patient and parent perceptions of rehabilitation factors that influence outcomes after anterior cruciate ligament reconstruction and clearance to return to sport in adolescents and young adults. *J Orthop Sports Phys Ther.* 2019; 49(8):576-583.
7. Paterno MV, Rauh MJ, Schmitt LC, Ford KR, Hewett TE. Incidence of second ACL injuries 2 years after primary ACL reconstruction and return to sport. *Am J Sports Med.* 2014;42(7):1567-73.
8. Herbst E, Hoser C, Hildebrandt C, Raschner C, Hepperger C, Pointner H, Fink C. Functional assessments for decision-making regarding return to sports following ACL reconstruction. Part II: clinical

application of a new test battery. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(5):1283-91.

9. Hoog P, Warren M, Smith CA, Chimera NJ. Functional hop tests and tuck jump assessment scores between female Division I collegiate athletes participating in high versus low ACL injury prone sports: a cross sectional analysis. *Int J Sports Phys Ther.* 2016;11(6):945-953.
10. Beischer S, Hamrin Senorski E, Thomeé C, Samuelsson K, Thomeé R. Young athletes return too early to knee-strenuous sport, without acceptable knee function after anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2018;26(7):1966-1974.
11. May KH, Guccione AA, Edwards MC, Goldstein MS. The adolescent measure of confidence and musculoskeletal performance (AMCAMP): development and initial validation. *Int J Sports Phys Ther*. 2016;11(5):698-707.
12. Glazer DD. Development and preliminary validation of the injury-psychological readiness to return to sport (I-PRRS) scale. *J Athl Train*. 2009;44(2):185-189.
13. Dover G, Amar V. Development and validation of the athlete fear avoidance questionnaire. *J Athl Train*. 2015;50(6):634-642.
14. Narducci E, Waltz A, Gorski K, Leppla L, Donaldson M. The clinical utility of functional performance tests within one-year post-ACL reconstruction: a systematic review. *Int J Sports Phys Ther*. 2011;6(4): 333-342.
15. Soderman K, Alfredson H, Pietila T, Werner S. Risk factors for leg injuries in female soccer players: A prospective investigation during one out-door season. *Knee Surg Sports Traumatol Arthrosc*. 2001; 9(5):313-321.
16. Nagai T, Schilaty ND, Laskowski ER, Hewett TE. Hop tests can result in higher limb symmetry index values than isokinetic strength and leg press tests in patients following ACL reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2020;28(3):816-822.
17. Grindem H, Snyder-Mackler L, Moksnes H, Engebretsen L, Risberg MA. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: The Delaware-Oslo ACL cohort study. *Br J Sports Med.* 2016;50(13):804-808.
18. Keays SL, Bullock-Saxton JE, Newcombe P, Keays AC. The relationship between knee strength and functional stability before and after anterior cruciate ligament reconstruction. *J Orthop Res.* 2003;21(2):231-237.
19. Barfod KW, Feller JA, Hartwig T, Devitt BM, Webster KE. Knee extensor strength and hop test performance following anterior cruciate ligament reconstruction. *Knee*. 2019;26(1):149-154.
20. Undheim MB, Cosgrave C, King E, et al. Isokinetic muscle strength and readiness to return to sport following anterior cruciate ligament reconstruction: is there an association? A systematic review and a protocol recommendation. *Br J Sports Med*. 2015;49(20):1305-1310.
21. Feiring DC, Ellenbecker TS, Derscheid GL. Test-retest reliability of the biodex isokinetic dynamometer. *J Orthop Sports Phys Ther*. 1990;11(7):298-300.
22. Drouin JM, Valovich-McLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. *Eur J Appl Physiol*. 2004;91(1):22-29.
23. Rainville J, Jouve C, Finno M, Limke J. Comparison of four tests of quadricepss strength in l3 or l4 radiculopathies. *Spine*. 2003;28(21):2466-2471.
24. Suri P, Rainville J, Katz JN, et al. The accuracy of the physical examination for the diagnosis of midlumbar and low lumbar nerve root impingement. *Spine*. 2011;36(1):63-73.
25. Beutler AI, Cooper LW, Kirkendall DT, Garrett WE. Electromyographic analysis of single-leg, closed chain exercises: Implications of rehabilitation after anterior cruciate ligament reconstruction. *J Athl Train*. 2002;37(1):13-18.
26. Zeller BL, McCrory JL, Kibler B, Uhl TL. Differences in kinematics and electromyographic activity between men and women during the single-legged squat. *Amer J of Sports Med.* 2003; 31(3):449-456.
27. Dedinsky R, Baker L, Imbus S, Bowman M, Murray L. Exercises that facilitate optimal hamstring and quadriceps co-activation to help decrease ACL injury risk in healthy females: A systematic review of the literature. *Int J Sports Phys Ther*. 2017; 12(1):4-15.
28. Fisher RA. *Statistical methods of research workers*. 12th ed. Edinburgh: Oliver and Boyd; 1954.
29. Goldberg A, Chavis M, Watkins J, Wilson T. The five-times-sit-to-stand test: validity, reliability and detectable change in older females. *Aging Clin Exp Res*. 2012;24(4):339-344.
30. Regterschot GRH, Zhang W, Baldus H, Stevens M, Zijlstra W. Test–retest reliability of sensor-based sit-to-stand measures in young and older adults. *Gait & Posture*. 2014;40(1):220-224.