RELIABILITY OF THE TUCK JUMP INJURY RISK SCREENING ASSESSMENT IN ELITE MALE YOUTH SOCCER PLAYERS

PAUL READ, MSC, ASCC, CSCS*D, JON L. OLIVER, PhD, MARK B.A. DE STE CROIX, PhD, GREGORY D. MYER, PhD, CSCS*D, and RHODRI S. LLOYD, PhD, ASCC, CSCS*D

1School of Sport, Health and Applied Science, St Mary’s University, London, United Kingdom
2Youth Physical Development Unit, School of Sport, Cardiff Metropolitan University, United Kingdom
3Exercise and Sport Research Centre, School of Sport and Exercise, University of Gloucestershire, United Kingdom
4Division of Sports Medicine, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, USA
5Department of Pediatrics and Orthopaedic Surgery, College of Medicine, University of Cincinnati, Cincinnati, Ohio, USA
6The Micheli Center for Sports Injury Prevention, Boston, MA, USA

Abstract

Altered neuromuscular control has been suggested as a mechanism for injury in soccer players. Ligamentous injuries most often occur during dynamic movements, such as decelerations from jump-landing maneuvers where high risk movement patterns are present. The assessment of kinematic variables during jump-landing tasks as part of a pre-participation screen is useful in the identification of injury risk. An example of a field-based screening tool is the repeated tuck jump assessment. The purpose of this study was to analyze the within-subject variation of the tuck jump screening assessment in elite male youth soccer players. 25 pre and 25 post-peak height velocity (PHV) elite male youth soccer players from the academy of a professional English soccer club completed the assessment. A test, re-test design was used to explore the within-subject inter-session reliability. Technique was graded retrospectively against the 10-point criteria set out in the screening protocol using two-dimensional video cameras. The typical error range reported for tuck jump total score (0.90 – 1.01 in pre and post-PHV players respectively) was considered acceptable. When each criteria was analyzed individually, Kappa coefficient determined that knee valgus was the only criterion to reach substantial agreement across the two test sessions for both groups. The results of this study suggest that although tuck jump total score may be reliably assessed in elite male youth soccer players, caution should be applied in solely interpreting the composite score due to the high within-subject variation in a number of the individual criteria. Knee valgus may be reliably used to screen elite youth male soccer players for this plyometric technique error and for test, re-test comparison.

Keywords

Injury Screening; Landing Assessment; Youth; Soccer
INTRODUCTION

The sport of soccer imposes high physiological demand with aerobic endurance (24), speed, agility, strength and power key determinants of elite performance (22, 38). An inherent risk of injury is also present due to frequent repetitions of high intensity movements that expose players to high joint and muscular forces (5). Incidence data in elite male youth soccer indicates that injuries occur mainly in the lower extremities (71–80%), are largely non-contact in nature, and are predominantly comprised of ligament sprains at the ankle and knee (4, 20, 34). Specifically, knee injury incidence rates in elite male youth soccer have been reported as 0.71 injuries per player each year, equating to an average of 17 training days and 2 matches missed per season (25). Furthermore, 35% of these injuries were classified as severe (> 28 days absent) and other data is available to confirm the knee as the most frequent site of major injury in elite male youth soccer, including anterior cruciate ligament (ACL) and medial collateral ligament (MCL) sprains (40).

While limited evidence is available in male youths to confirm pertinent risk for injuries, altered neuromuscular control has been suggested as a mechanism in females and adults (1, 33, 36). Deficits in neuromuscular control can impose excessive stress to the passive ligamentous structures, exceeding their strength limit and increase the potential for compromised structural integrity and injury (21). Ligamentous injuries most often occur during dynamic movements, such as decelerations from jump-landing maneuvers where high risk movement patterns are present (17).

The assessment of kinematic variables during jump-landing tasks as part of a pre-participation screen is useful in the identification of injury risk (26, 27, 28, 30). A common diagnostic tool used within the available literature is a drop vertical jump (3, 7, 12, 28, 29, 32). These studies have confirmed this assessment provides both valid and reliable indicators of anterior cruciate ligament (ACL) injury risk factors including knee valgus and high knee abduction moments (12, 28, 29, 32). Due to the need for expensive laboratory equipment and labor intensive testing procedures, recent investigations have explored more clinician friendly alternatives using standard video cameras to analyze drop vertical jump tasks (28, 32). However, these clinic based methods still present limitations for youth soccer coaches working with large groups. There is a paucity of literature available in youth male populations and when assessing large groups comprising a variety of age ranges and anthropometric profiles, it is inherently difficult to select an appropriate drop height that accounts for individual jumping and landing abilities. Additionally, during drop landing tasks, participants begin in an elevated position; muscles that were previously at rest while standing on the box control impact forces at the point of ground contact eccentrically. In more functional soccer tasks (jumping, cutting etc.), the involved musculature are required to perform propulsive motions (i.e. the initial jump or running action) prior to ground contact and it could be inferred that drop jumping tasks may artificially induce feed-forward stabilization mechanisms. This could be considered a learned skill as preparatory muscle activation strategies are known to develop as a result of age and maturation (19). Further, despite the aforementioned methods (28) providing a more time efficient and ‘clinician friendly’ approach than laboratory techniques, there is still a requirement for digitization of
joint angles to determine knee valgus motion and maximal flexion angles which is time inefficient.

Repeated jumping tasks whereby athletes are required to respond to movement perturbations and forces may be more ecologically valid in sports such as soccer. An example of a field-based screening tool is the repeated tuck jump assessment (27). This method has been developed to identify neuromuscular deficits and identify lower extremity technique flaws during a plyometric exercise (27, 31, 35). It has been suggested that the presence of these neuromuscular deficits may be indicative of an increased risk of traumatic knee injury (26, 30). Assessed using a ten point rating scale, initial pilot studies have indicated moderate-to-strong inter-rater reliability (range = 0.72 – 0.97) (27). More recently, intra and inter-tester reliability have been reported demonstrating excellent agreement (kappa measurement, k = 0.88 [11]).

The tuck jump assessments was originally designed as a coach friendly plyometric landing assessment and training tool which targets risk factors for ACL injury mechanisms (27, 31, 35). However, less information is available to confirm its reliability and validity as a diagnostic screening tool, especially in male youth soccer players who have a high incidence of medial collateral ligament (MCL) injuries (25, 34) where a valgus mechanisms has been suggested as a high risk movement pattern (9, 16). It has been suggested that baseline levels and deficits can be compared following targeted training interventions to objectively track improvements in performance (31, 35). However, no data is currently available that report within-subject variation using test, re-test experimental design, especially in pediatric populations. Given the inherent risk of lower limb injury to male youth soccer players, especially around the time of peak height velocity (PHV) (39), and reported higher variability in jump performance in less mature individuals (10), it would seem pertinent to confirm the reliability of the tuck jump assessment with this particular population. The aim of this study is to determine the reliability of the tuck jump assessment tool in both pre and post-PHV elite male youth soccer players.

**METHODS**

**Experimental Approach to the Problem**

This study used a test, re-test design to explore the inter-session reliability of the tuck jump assessment. Participants were required to attend the club training ground on three occasions each separated by a period of seven days. The first session was used to familiarize participants to the test allowing both visual demonstrations and practice attempts until the principal researcher was satisfied that technique was appropriate and consistent. Data collected in the second and third sessions were subsequently analyzed to determine within-subject variation. In each of the two experimental test sessions, a 10-minute standardized warm up inclusive of a series of dynamic stretches was completed prior to commencement. Testing was completed at the same time on each day and participants were asked to wear the same clothing and footwear and refrain from strenuous exercise for at least 48 hours prior to testing. Participants were also asked to eat according to their normal diet and avoid eating and drinking substances other than water one hour prior to each test session.
Subjects

Twenty five pre-PHV (age 11.93 ± 0.43 yr; height 151.40 ± 4.84 cm; body mass 41.05 ± 5.62 kg; maturational offset −2.34 ± 0.41 yr;) and twenty five post-PHV (age 17.26 ± 0.69 stature 178.22 ± 5.47; body mass 72.27 ± 6.93 kg; maturity offset 2.91 ± 0.81 yr) youth soccer players from the academy of a professional English Championship soccer club volunteered to take part in the study. The research design required players to complete a familiarization session and two experimental test sessions separated by a period of one-week. None of the players reported injuries at the time of testing and all were participating regularly in football training and competitions. Parental consent and participant assent were collected prior to the commencement of testing. In addition, subjects completed a physical activity readiness questionnaire to determine their health status was appropriate and there was no significant physical reason why they should not partake in the research project. Ethical approval was granted by the institutional ethics committee in accordance with the declaration of Helsinki.

Procedures

During each recorded test session, tuck jumps were performed in place for a period of 10 seconds (27). Prior to the test, participants were required to place their feet 35 cm apart on two vertical strips of tape, which were connected by a horizontal line forming an H-Shape (35). Participants began by performing a countermovement before jumping in a vertical direction as high as possible while simultaneously pulling their knees up towards their chest. Upon ground contact, participants were encouraged to minimize ground contact time and utilize a toe to mid-foot rocking landing strategy. They were also instructed to land in the same footprint with each jump. Players were screened using a specified 10-point criteria as suggested previously (27, 31, 35). Deficiencies in plyometric technique were marked if it was not maintained for 10 seconds.

To increase assessment accuracy, two-dimensional video cameras were used to capture the test and each player’s technique was graded retrospectively. Kinematic data was collected at 50 Hz using two high-definition video cameras (Samsung, New Jersey, USA) positioned in the frontal and sagittal planes at a height 0.70 m, and a distance of five meters from the center of the capture area. To allow visible tracking of the knees, subjects were required to wear shorts with a line at approximately mid-thigh. When scoring each trial of the tuck jump assessment, each of the 10 criteria were analyzed and if a deficit was present on two or more repetitions it was marked as previously stated (27, 31, 35). Additionally, a composite score tallied from the recorded deficits was provided and higher scores were indicative of reduced performance, 10 being the maximum score attainable suggesting players should be targeted for further technique training (27). The same rater marked and recorded deficits to avoid the presence of inter-rater error. Further analysis was also completed to determine the frequency that each criteria was present and the total number of jumps in each session.

Statistical Analysis

Tests for normal distribution were completed on the recorded data and means and standard deviations for tuck jump total score were calculated across the two testing sessions. Twenty-five participants per group were selected to provide a sufficiently small confidence interval to the reliability statistics. This provided a confidence interval factor of approximately 1.35 to apply random variation. Doubling the sample size would only reduce the precision of the confidence interval factor to 1.22 (14), and thus have limited effects on the outcomes or interpretations of reliability.

Reliability statistics included: change in mean, intra-class correlation coefficient (ICC) to determine rank order repeatability and typical error estimate. Reliability data was computed through Microsoft Excel® 2010 using a freely available spreadsheet (15). To determine systematic bias between trials paired samples t-tests were used with a p value ≤0.05. The total number of jumps completed during each trial was calculated and a paired samples t-test was used to check for a significant difference between test sessions. Further, Pearson correlation coefficients were calculated to assess the relationship between tuck jump total score and the number of jumps performed. Paired samples t-tests were and Pearson correlation coefficients were processed using SPSS ® (V 21. Chicago Illinois).

To determine the accuracy of observations across the two test sessions for each individual criterion displayed in the tuck jump assessment checklist, agreement was assessed using Kappa coefficient with the following threshold classifications: >0.8 almost perfect agreement; 0.6–0.8 substantial agreement; 0.4–0.6 moderate agreement; 0.2–0.4 slight to fair agreement (18). Intra-rater reliability was calculated for the tuck jump total score using intra-class correlation coefficient (ICC) where the same rater scored a sample of the videos from experimental test session 1 on two occasions separated by a period of 7 days.

RESULTS

Descriptive statistics and all reliability values are displayed in table 1. Intra-rater reliability for the 10s repeated tuck jump total score was deemed strong (ICC = 0.88) based on previous classifications (13). A significant difference was observed between the total number of jumps performed in each test session only in the post pubertal group (P < 0.001) who completed less jumps in experimental test session 2. Relationships between tuck jump total score deficits and the numbers of jumps completed were significantly correlated for the pre-pubertal group only (r = 0.37; P < 0.001).

The kappa coefficient used to determine between-session agreement for each of the ten individual tuck jump criteria are displayed in table 2. Knee valgus was the only criterion that reached substantial agreement across the two test trials for both groups suggesting strong reliability. Other measures including unparalleled foot position during landing and a pause between jumps also reported substantial agreement for the pre-PHV players; however, this was not consistent across groups. Moderate agreement (k = 0.6 – 0.4) was present for foot positioning not being shoulder width during landing (pre-PHV group only), thigh motion not reaching parallel and high contact noise (post-PHV group only) suggesting these criteria could be measured with acceptable reliability. The agreement between all other criteria was
DISCUSSION

To the author’s knowledge, this is the first study to assess the reliability of the repeated tuck jump assessment in male youth soccer players which could be included as part of a field-based neuromuscular control screening battery. Research quantifying the effects of age, growth and maturation on the reliability of this test also does not currently exist. The total score obtained during the repeated tuck jump assessment demonstrated moderate ICC values. However, caution should be applied when using this approach for determining the reliability of a test. A re-test correlation measures how closely the values of two trials track each other specific to each individual and the reproducibility of the rank order of subjects during the re-test (14). ICC statistics are also affected by the heterogeneity of the values between participants, and the sample used for measurement (14). A sample which is of a homogenous nature will demonstrate a low value. The subjects in this study are reflective of a homogenous sample due to their status as elite male youth academy soccer players and thus, this provides a plausible explanation for the low ICC values. Examining the results of previous reliability studies using this approach and interpreting the findings is also difficult.

An alternative and more appropriate approach is to use statistical tests that measure the typical error to show variation in the participants score between experimental test sessions (2, 13). The typical error range reported (0.89 – 1.01 in pre and post-PHV players respectively) would be considered acceptable and indicates that changes in total score >1 following an intervention would be classified as a real change in injury risk reduction. Previous investigations have reported large percentage changes following augmented feedback interventions for various criteria in the tuck jump assessment (35), suggesting that a change in total score >1 is realistic. However, in-spite of the acceptable typical error values reported, when each of the 10 individual criterion measures were analyzed separately, a number of these demonstrated fair to low agreement. These findings may be explained in part by the greater number of criteria included in the total tuck jump score which may have provided an averaging effect where errors begin to cancel themselves out. This is representative of taking the mean of more trials. However, due to the paucity of literature which has investigated the reliability of this test, further research is warranted.

Criterion deficits of knee valgus, feet not parallel or shoulder width on landing, and a pause between jumps were the only measures to demonstrate acceptable reliability for pre-PHV players (kappa range = 0.52 – 0.78). Whereas, observed knee and thigh motion variables including: knee valgus on landing, and thighs not parallel in flight and high contact noise were the only criteria to report acceptable reliability in the post-pubertal group (range k = 0.43 – 0.67). This suggests that these plyometric technique flaws can be reliably assessed as part of a test-retetest screen in pre-PHV children. Thus, caution should be applied in purely interpreting the composite assessment score due to the high within-subject variation in a number of the individual criteria. Respective of a change in total score >1 which was identified as a reliable performance modification, it would be uncertain as to how injury risk has changed.
While this is the only study to analyze test, re-test within-subject reliability for each of the individual criteria during the tuck jump assessment, knee valgus on landing reported the strongest agreement for both groups. This suggests that screening for this criterion risk factor is reliable for test re-test comparison in male youth soccer players who are either pre- or post-PHV. Knee injuries comprise 17% of all injuries in elite male youth soccer players (25) and the valgus mechanism has been suggested as a high-risk movement pattern for incidences of both medial collateral ligament (MCL) (9, 16), and ACL injury (8, 12). MCL injuries occur frequently in male youth soccer accounting for up to 85% of knee ligament injuries (19, 33) and ACL lesions are the most frequently occurring serious injury (> 28 days absence) in this cohort (40). Knee valgus motion has been cited as a key contributing factor to high knee abduction moments and ACL injuries (29) and is a predictor of injury assessed during jumping and landing tasks (12). Therefore, practitioners using the tuck jump assessment to screen male youth soccer players may reliably track changes in knee valgus on landing following an intervention.

There is a paucity of available literature to compare the current findings with other studies and specifically no data is available in male youth of different ages or stages of maturation. In their initial review of this assessment, Myer et al. (27) stated that unpublished pilot studies indicated moderate-to-strong inter-rater reliability (range = 0.72 – 0.97). More recently, Herrington et al. (11) assessed the main outcome scores of males and females using two independent assessors demonstrating good intra and inter-tester reliability which is in agreement with the findings of this study (intra rater reliability = 0.88). This suggests that variation in test scores is most likely due to the participants and not the rater. However, other investigations have shown poor inter-rater and poor to moderate intra-rater reliability (6) and none of these studies investigated within-subject reliability. In the present study, high within-subject variance was present as indicated by low kappa agreement between trials on a number of the test criteria. This finding is not unexpected, as obtaining high reliability for repeated trials during tasks requiring dynamic jumping and landing activities that are reliant on reflexive muscle responses, proprioceptive and kinesthetic feedback will likely utilize a range of movement strategies, thus increasing variability (7, 23, 37). Previous literature has suggested greater variation in jump performance in younger subjects (10); lower levels of consistency in technique can be expected in youth players who may experience difficulty in consistently executing the correct jump-landing pattern.

A significant difference was noted between the numbers of jumps completed in each test in the post-pubertal group. This occurred despite the standardized 10-second protocol which has been previously suggested (27, 31, 35). Measures need to be reproducible if their purpose is for repeated screening. Intuitively, a player who performs more or less jumps during a re-test may heighten the chance of deficit changes (either higher or lower) being reported purely as a result of altering the number of jumps completed. Specifically, a greater number of jumps will typically increase the likelihood of more errors and this may partly explain the variation in jump performance between the two trials in the current study. This was likely evident in the pre-PHV group supported by a significant correlation between the number of jumps performed and the tuck jump total score. Also, fewer jumps during the 10 second test period may be representative of more powerful jump performance due to increases in flight time. Cumulatively, these findings suggest that for practitioners who wish
to use this assessment for regular athlete monitoring and comparison between groups, standardizing the number of jumps performed may be recommended.

**PRACTICAL APPLICATIONS**

This study is novel as data is now available to determine the within-subject reliability of the repeated tuck jump assessment in male youth soccer players. The total score comprised of 10 criterion measures reported acceptable within-subject variation and indicated that a change in score >1 is needed to reliably state that performance has improved. However, when each criterion was analyzed individually, moderate-substantial agreement between-trials were noted in some of the assessed criterion but these differed in each age group and other criteria reported low agreement. Therefore, caution should be applied when solely interpreting the composite score due to uncertainty of how their tuck jump technique has changed and subsequently which potential injury risk factors are present.

Knee valgus was the only measure to demonstrate substantial agreement in both groups and may be reliably used to screen male youth soccer players for this potential injury risk factor and for test, re-test comparison, on the provision that the total number of jumps performed is standardized. Thus, it is suggested that the tuck jump could be used to effectively assess jump-landing technique but only the presence of knee valgus should be evaluated. This provides a time-efficient tool for practitioners who screen large groups of athletes as part of regular field-based screening protocols. Also, lower total scores reported in the post-PHV players on both trials tentatively suggest that performance in the tuck jump assessment may improve as a result of growth and maturation.

**References**


Biographies

Paul Read, MSc, ASCC, CSCS*D

Contact Details: School of Sport, Health and Applied Sciences, St Mary’s University, Waldegrave Road, Twickenham, London, TW1 4SX, paul.read@stmarys.ac.uk Telephone Number: +44 20 8240 4000

Current Position: Paul Read Paul is a strength and conditioning coach and senior lecturer in strength and conditioning at the St Mary’s University
John L. Oliver, PhD

**Current Position:** Jon L. Oliver is a Reader in Applied Paediatric Exercise Science at Cardiff Metropolitan University and he regularly collaborates with international governing bodies and professional organizations on research in youth sports including soccer, rugby union, martial arts, and golf.

Mark De Ste Croix, PhD

**Current Position:** Professor Mark De Ste Croix is professor of paediatric sport and exercise and co-director of the Exercise and Sport Research Centre at the University of Gloucestershire. Mark works with a number of professional football clubs in the UK and Spain as well as National Governing bodies, focusing on injury risk in youth sport.

Gregory D. Myer, PhD

**Current Position:** Gregory D. Myer is director of Research at the Human Performance Laboratory for the Division of Sports Medicine at Cincinnati Children’s Hospital Medical Center and holds primary academic appointments of Associate Professor in the Departments of Pediatrics and Orthopaedic Surgery within the College of Medicine at University of Cincinnati.
Rhodri S. Lloyd, PhD

**Current Position:** Rhodri S. Lloyd is a Senior Lecturer in strength and conditioning at Cardiff Metropolitan University and serves as a Board Director for the UK Strength and Conditioning Association.
Figure 1. Tuck jump Assessment: Six common mistakes that clinicians should aim to correct for their athletes while they perform the tuck jump exercise: (1) athletes display unwanted medial knee collapse, (2) athletes do not achieve the desired knees parallel position at top of flight, (3) athletes do not displayed synchronized lower limb positions during flight, (4) athletes land with their feet too close together, (5) athletes land in undesirable staggered position, and (6) athletes do not land with both feet at the same time.
Table 1

Descriptive statistics and reliability values for Tuck Jump total score

<table>
<thead>
<tr>
<th>Test</th>
<th>Number of Jumps</th>
<th>TJ Total Score</th>
<th>Change in mean</th>
<th>ICC</th>
<th>Typical Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test 1</td>
<td>Test 2</td>
<td>Test 1</td>
<td>Test 2</td>
<td></td>
</tr>
<tr>
<td>TJ total score pre-PHV</td>
<td>14.4 ± 1.6</td>
<td>14.7 ± 1.48</td>
<td>6.88 ± 1.33</td>
<td>6.64 ± 1.50</td>
<td>−0.24 ± 1.27</td>
</tr>
<tr>
<td>TJ total score post-PHV</td>
<td>15.8 ± 0.8</td>
<td>14.4 ± 0.81 *</td>
<td>5.44 ± 1.29</td>
<td>5.40 ± 1.83</td>
<td>0.17 ± 1.43</td>
</tr>
</tbody>
</table>

TJ = Tuck Jump

* significantly different from test 1 (p < 0.05)
## Table 2

### Tuck Jump Kappa Statistics

<table>
<thead>
<tr>
<th>Phase of Jump</th>
<th>Criterion</th>
<th>Kappa Value (pre-PHV)</th>
<th>Kappa Value (post-PHV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee and thigh motion</td>
<td>knee valgus</td>
<td>k = 0.78</td>
<td>k = 0.67</td>
</tr>
<tr>
<td></td>
<td>not parallel</td>
<td>k = 0.34</td>
<td>k = 0.44</td>
</tr>
<tr>
<td></td>
<td>not equal</td>
<td>k = 0.11</td>
<td>k = 0.29</td>
</tr>
<tr>
<td>Foot positioning during landing</td>
<td>not shoulder width</td>
<td>k = 0.52</td>
<td>k = 0.25</td>
</tr>
<tr>
<td></td>
<td>not parallel</td>
<td>k = 0.62</td>
<td>k = 0.20</td>
</tr>
<tr>
<td></td>
<td>unequal contact time</td>
<td>k = 0.04</td>
<td>k = 0.05</td>
</tr>
<tr>
<td></td>
<td>high contact noise</td>
<td>k = 0.34</td>
<td>k = 0.43</td>
</tr>
<tr>
<td>Plyometric technique</td>
<td>pause between jumps</td>
<td>k = 0.65</td>
<td>k = 0.00</td>
</tr>
<tr>
<td></td>
<td>declines in 10s</td>
<td>k = 0.33</td>
<td>k = 0.31</td>
</tr>
<tr>
<td></td>
<td>excess flight motion</td>
<td>k = −0.06</td>
<td>k = 0.01</td>
</tr>
</tbody>
</table>