The reliability and criterion validity of 2D video assessment of single leg squat and hop landing

Authors
Lee Herrington PhD MCSP[1*]
Faisal Alenezi PhD[1]
Msaad Alzhrani MSc[2]
Hasan Alrayani MSc[1]
Richard Jones PhD[1]

1 Centre for Health, Sport and Rehabilitation Sciences, University of Salford, UK
2 Physiotherapy Department, Majmaah University, Kingdom of Saudi Arabia
*Corresponding author

The reliability and criterion validity of 2D video assessment of single leg squat and hop landing

Abstract

The objective was to assess the intra-tester, within and between day reliability of measurement of hip adduction (HADD) and frontal plane projection angles (FPPA) during single leg squat (SLS) and single leg landing (SLL) using 2D video and the validity of these measurements against those found during 3D motion capture. 15 healthy subjects had their SLS and SLL assessed using 3D motion capture and video analysis. Inter-tester reliability for both SLS and SLL when measuring FPPA and HADD show excellent correlations ($ICC_{2,1}$ 0.97-0.99). Within and between day assessment of SLS and SLL showed good to excellent correlations for both variables ($ICC_{3,1}$ 0.72-91). 2D FPPA measures were found to have good correlation with knee abduction angle in 3-D (r = 0.79, p = 0.008) during SLS, and also to knee abduction moment (r = 0.65, p =0.009). 2D HADD showed very good correlation with 3D HADD during SLS (r = 0.81, p = 0.001), and a good correlation during SLL (r = 0.62, p =0.013). All other associations were weak ($r<0.4$). This study suggests that 2D video kinematics have a reasonable association to what is being measured with 3D motion capture.

Keywords: biomechanics; landing; assessment

THE RELIABILITY AND CRITERION VALIDITY OF 2D VIDEO ASSESSMENT OF SINGLE LEG SQUAT AND HOP LANDING

INTRODUCTION

Three dimensional (3D) motion analysis has been used extensively to assess kinematic and kinetic variables during lower limb motion. It has been regarded as the 'gold standard' for the assessment of potentially high risk manoeuvres related to a variety of knee injuries (McLean et al., 2005). Although 3D motion capture is considered the gold standard for kinematic and kinetic analysis, it is frequently not used in the clinical environment or for pre-participation screening, possibly due to the time required to acquire and analyse the data, large cost of equipment, and the training needed to effectively use it. In the place of 3D motion capture, 2-dimensional (2D) video motion analysis has been used to quantify hip and knee kinematics (Munro et al., 2012). 2D motion capture though has an inherent limitation as it cannot measure kinematics that occurs in planes not perpendicular to the camera without potential for perspective error. As such, 2D motion capture may not be suitable for performance assessment of any motion that is not purely uniplanar such as the knee valgus motion at the knee, which in reality is a movement not only comprising of knee abduction and hip adduction in the frontal plane but also hip internal rotation and tibial external rotation in the coronal plane (Malfait et al., 2014). The work of McLean et al (2005) confirmed this noting that 2D knee valgus angles were inherently influenced by hip and knee joint rotations.

The extent to which non-uniplanar motions can be reflected in the uniplanar knee motion, measured with 2D video, has only been investigated in a limited number of studies. These studies have tested for a relationship between 2D measures of knee and hip motion and 3D hip and knee kinematics. For example, McLean et al (2005) reported the relationship between 2D and 3D motion capture in assessing frontal-plane knee kinematics during side-stepping, side-jumping, and shuttle run. They reported strong correlations of r = .76 and .80 between peak knee abduction angles during 2D and 3D motion capture for side-stepping and side-jumping, respectively; however, the shuttle run yielded a much lower relationship of just *r* = 0.20. Sorenson et al (2015) found a strong relationship between 2D frontal plane projection angle (knee abduction angle) and 3D knee abduction angle (R^2 = 0.72), and between 2D hip adduction and 3D hip adduction (R^2 = 0.52) during single leg hop landings. Gywnne and Curran (2014) found FPPA to correlate strongly with 3D knee abduction angle during single leg squat (r=0.78). The study of Willson and Davis (2008) found 2D knee abduction angle reflected 23%–30% of the variance of 3-D kinematic measurements during single leg squat, and also found knee abduction angle to be significantly correlated with hip adduction (r=0.32). However, none of these studies have looked at the relationship of 2D frontal plane measures to movements in other planes, or the external moments generated at the hip and knee.

There are currently only a limited number of publications which have reported reliability of 2D FPPA for both single leg squat and single leg landing (Gwynne and Curran, 2014; Munro et al 2012). There would appear to be no studies which have reported on reliability of the 2D video measurement of hip adduction angle during these tasks.

The overall aim of the study was to assess the reliability and validity of 2D kinematic video analysis of single leg squat and single leg landing, specifically: to assess the intra-tester and within and between day reliability of measurement of hip adduction and frontal plane projection angles during SLS and SLL using 2D video; and to assess the validity of these measurements against those found during 3D motion capture. The three hypotheses which will be tested by this study are: that 2D video parameters will have validity when compared to equivalent 3D parameters; 2D parameters measured will show strong between individual and within and between session reliability.

METHOD

Participants: Fifteen physically active healthy participants, after giving informed consent, volunteered to participate in this study which was approved by the university research ethics committee. Participants had to be free from lower limb or spinal injury or history of injury to participate in the study. Participant's details are to be found in table 1.

Procedures

Single leg Squat (SLS).

Participants were asked to stand on the test limb, facing the video camera. They were asked to squat down as far as possible, to at least 45° knee flexion, over a period of 5 seconds. Knee-flexion angle was checked during practice trials using a standard goniometer (Gaiam-Pro), and then observed by the same examiner throughout the trials. There was also a counter for each participant over this 5-second period, in which the first count initiates the movement, the third indicates the lowest point of the squat and the fifth indicates the end. This standardises the test for the participant, thereby reducing the effect of velocity on knee angles. Trials were only accepted if the particpant squatted to the minimum desired degree of knee flexion and maintained balance throughout.

Single-Leg Landing (SLL).

Participants dropped from a 28-cm step, leaning forward and dropping as vertically as possible. They were asked to take a unilateral stance on the ipsilateral limb and to hop forward to drop onto the force platform, ensuring that the contralateral leg made no contact with the ground on landing.

3D motion capture: The method is based on the procedure previously reported in Alenezi et al (2014). A ten-camera motion analysis system (Pro-Reflex, Qualisys, Sweden), sampling at 240 Hz, and a force platform embedded into the floor (AMTI, USA), sampling at 1200 Hz, were used to collect kinematic and kinetic variables during the support phase of single leg squat and landing tasks. Before testing, participants were fitted with the standard training shoes (New Balance, UK) to control shoe-surface interface. Reflective markers (14mm) were attached with self-adhesive tape to the participants' lower extremities over the following landmarks; anterior superior iliac spines, posterior superior iliac spines, iliac crest, greater trochanters, medial and lateral femoral condyles, medial and lateral malleoli, posterior calcanei, and the head of the first, second and fifth metatarsals (figure 1). The tracking markers were mounted on technical clusters on the thigh and shank with elastic bands.

The foot markers were placed on the shoes, and the same individual placed the markers for all participants. The calibration anatomical systems technique (CAST) was employed to determine the six-degree of freedom movement of each segment and anatomical significance during the movement trials. The static trial position was designated as the participants' neutral (anatomical zero) alignment, and subsequent kinematic measures were related back to this position. The markers were removed and replaced for the within-session trials and removed and replaced for the between-day trials. To orientate participants with the tasks, each participant was asked to perform 3–5 practice trials of each task before data collection. Participants were required to complete five successful trials for each task. Visual3D motion (Version 4.21, C-Motion Inc. USA) was used to calculate the joint kinematic and kinetic data. Motion and force plate data were filtered using a Butterworth 4th order bi-directional low-pass filter with cut-off frequencies of 12 Hz and 25 Hz, respectively, with the cut-off frequencies based on a residual analysis (Yu et al., 1999). All lower extremity segments were modelled as conical frustra, with inertial parameters estimated from anthropometric data (Dempster, 1959). Joint kinematic data was calculated using an X–Y–Z Euler rotation sequence. Joint kinetic data were calculated using three-dimensional inverse dynamics, and the joint moment data were normalised to body mass and presented as external moments referenced to the proximal segment. External moments were described in this study, for example, an external knee valgus load will lead to abduct the knee (valgus position), and an external knee flexion load will tend to flex the knee. The following discrete variables were calculated for each trial: peaks of hip internal rotation and adduction moments, knee abduction moment, and peaks of lower limb joint angles at frontal, sagittal, & transverse planes for the hip and knee.

2D video assessment: Before testing, markers were placed on the lower extremity of each participant to approximate the radiographic landmarks employed by Willson et al (2006) and Willson and Davis (2008). Markers were placed at the midpoint of the femoral condyles to approximate the centre of the knee joint, midpoint of the ankle malleoli for the centre of the ankle joint, and on the proximal thigh along a line from the anterior superior iliac spine to the knee marker, markers were also placed on the anterior superior iliac spines. The midpoints were determined using a standard tape measure, and all markers were placed by the same experimenter. These markers were used in order for FPPA of the knee and hip adduction angles to be determined from digital images using Quintic software package (9.03 version 17). Digital video footage was captured at the same time as 3D, recorded at a standard 10× optical zoom using a Sony HandycamDCR-HC37 camera positioned at a height of 60 cm, 10 m away from the participants and directly in front, after which the footage was downloaded to Quintic. A single experimenter digitised the markers placed on the participant across the three sessions, with another experimenter digitising the video from the first sessions to provide data for the inter-tester reliability element.

Outcome measures from 2D video

The frontal plane projection angle and the hip adduction angle were collected from the videos of all participants for all tasks. Frontal plane projection angle (FPPA) of the knee was measured as the angle subtended between the line from the markers on the proximal thigh to the knee joint and the line from the knee joint to the ankle at the frame that corresponded with the point of maximum knee flexion. Positive FPPA values reflected knee valgus, excursion of the knee toward the midline of the body so that the knee marker was medial to the line between the ankle and thigh markers, and negative FPPA values reflected knee varus. The hip adduction angle (HADD) was measured as an angle subtended between the line from the markers on the proximal thigh and the line between the two anterior superior iliac spines. In figure 1 the lines used are superimposed in the figure.

Participants were tested twice on day 1 (tests 1 and 2), with the tests separated by 1 hour, to assess within day reliability. Participants were then tested again exactly 1 week later (test 3) at the same time of day to assess between-days reliability. Participants were allowed practice trials before each test until they felt comfortable; this was typically 2 or 3 trials. After familiarisation each participant performed 3 trials of each test. Both legs were tested and analysed for all tests.

Analysis

Inter-tester reliability was assessed by two of the authors independently viewing the 2D videos of the first assessment session, and measuring FPPA and hip adduction angles from SLS and SLL for all participants, then the findings were compared using intra-class correlation (model 2,1). Within and between session reliability was assessed when a single examiner viewed the 2D videos of the first, second and third assessment sessions, and measuring FPPA and hip add angles from SLS and SLL for all participants, then comparing findings for within day and between day reliability using intra-class correlation (model 3,1). The standard error of measurement was also calculated for all variables using the formula SEM = SD (pooled) x (√1-ICC) (Thomas et al., 2005). The criterion validity of the 2D measures was then assessed by comparing the findings from the first assessment of SLS and SLL to the findings from 3D motion capture undertaken at the same time using a Pearson's product moment correlation. Paired T-tests were ran to ensure that there is no consistent between- or within-day difference in the means between rater one and rater two, between sessions and between 2D and 3D values.

RESULTS

Inter-tester 2D video measurement reliability

Inter-tester reliability between two testers of the video capture technique has been assessed in both tasks the ICC's and standard error of measurement (SEM) are shown in table 2. In SLS, the correlation was found to be very large for both variables. FPPA and HADD reported correlations were (r = 0.97, p = 0.001), (r = 0.96, p =0.001) respectively. Also, FPPA and HADD were found to be highly correlated (r = 0.99, p = 0.001), (r =0.99, p = 0.001) respectively in SLL.

Within and between day 2D video measurement reliability

Table 3 shows the within and between session ICC's and SEM's. In SLS, 2D FPPA measures demonstrated good within-session (ICC = 0.72, 95% CI = 0.09 to 0.92) reliability. The 2D HADD also demonstrated good within-session (ICC = 0.91, 95% CI = 0.71 to 0.97) reliability. The SLL, within-session reliabilities were (ICC = 0.87, 95% CI = 0.58 to 0.96), (ICC = 0.89, 95% CI = 0.65 to 0.97) for FPPA and HADD respectively. The SLS, 2D FPPA measures showed good between-session (ICC = 0.87, 95% CI = 0.57 to 0.96) reliability, along with 2D HADD demonstrating good between session (ICC = 0.79, 95% CI = 0.32 to 0.94) reliability. While during SLL, between-session reliabilities were (ICC = 0.87, 95% CI = 0.58 to 0.96), (ICC = 0.86, 95% CI = 0.54 to 0.96) for FPPA and HADD respectively.

Correlation 2D measurements to 3D motion capture kinematic and kinetic values

Table 4 shows the correlations between 2D and 3D variables. 2D FPPA measures were found to have good correlation with knee abduction angle in 3-D (r = 0.79, p = 0.008) during SLS. Also, 2D FPPA was found to be correlated with knee abduction moment (r = 0.65, p =0.009) during SLS. 2D HADD showed very good correlation with 3D HADD during SLS (r = 0.81, p = 0.001), and a good correlation during SLL (r = 0.62, p =0.013). All other associations were weak (r<0.4) and not significant (p>0.05).

Table 5 shows the mean values for hip adduction and knee abduction (FPPA) across sessions for both raters and for 3D motion capture. Paired t-tests were ran on this data the results are shown in table 5 and show no significant differences between values across all comparisons.

DISCUSSION

In line with previous work (Gwynne & Curran, 2014; Munro et al 2012) 2D assessment was shown to be reliable within and between days. In Munro et al (2012) SLS within day reliability was 0.59-0.86 and in Gwynne and Curran (2014) it was 0.86 (our study 0.72) and between day 0.72-0.82 (Munro et al 2012), 0.74 (Gwynne & Curran, 2014) our study 0.87. In Munro et al (2012) for SLL within day 0.75-0.79 versus 0.87 and between days 0.8-0.82 versus 0.87. Generally the standard error of measurement (SEM) in our study was slight less than those reported in Munro et al (2012) and Gwynne & Curran (2014) across the tasks. It is not clear why our study found between session reliability to be slightly better in terms of the ICC values than within day for SLS (table 3), it might be partially explained by the larger confidence intervals for within day, though there was no significant difference in mean scores (table 5). No previous work has attempted to assess inter-tester reliability for the assessment of FPPA or hip adduction angles from the videos during SLS or SLL so our findings cannot be compared to previous work. It would appear from our data that the method undertaken to assess FPPA and HADD for all tasks shows excellent inter-tester reliability. The findings of this paper along with those of Gwynne & Curran (2014) and Munro et al (2012) would indicate that the methods used are robust enough to provide reliable results across multiple testers and time points, this opens up the possibility of using these tests in multi-centre trials.

There are only a limited number of studies which have investigated the association between 2D and 3D knee and hip motion during functional tasks. These studies have generally, despite measuring the relationships during different tasks, found moderately strong correlations between certain parameters when undertaking 2D and 3D motion capture in line with the findings of this study (McLean et al. 2005). When assessing similar tasks, in for instance the study of Sorensen et al (2015) they found FPPA had a strong relationship to 3D knee abduction angle (r^2 =0.72) during SLL, similar to the current studies findings. Willson and Davis (2008) found a moderate relationship during SLS of FPPA to knee abduction angle assessed with 3D (r=0.48), though only a weak one with HADD (r=0.37), which differs slightly from our study. Gwynne & Curran (2014) found a strong relationship between FPPA during SLS and 3D knee abduction angle (r=0.78). Ageberg et al (2010) reported 3D hip internal rotation rather than knee abduction angle appeared to be related to a knee medial to the foot position during a single leg squatting task, this study could be regarded as showing partial agreement with their findings as there was a poor correlation between hip internal rotation angle and FPPA.

When the findings from the literature along with ours are taken on the whole, it would appear that both FPPA and HADD measured using 2D video bear a moderate to good relationship with the comparable measures produced with 3D motion capture, especially for less complex and dynamic tasks such as SLS. The difference in validity between the two tasks (SLS and SLL) might be due to the difference between the tasks and their impact on matching the exact moment of maximum knee flexion angle. FPPA was captured at the point of maximum flexion, because of the different capture speeds of the two systems being compared, during is the high speed task of SLL, the poor correlation could relate to an inability to be measuring at exactly the same knee flexion point. Whereas during the slower task of SLS it is more likely the two systems coalesce.

This study has limited generalisability as the relationships were only found in uninjured individuals and further study is required to identify if these or different relationships occur in individuals with pathology. In addition, as with other 3D motion capture studies using external marker sets, skin-movement artefact has the potential to influence such data. The intrinsic difference in sampling frequency between the two measurement systems used may influence the relationships found as there may have been temporal differences in the data extracted from each system leading to systematic error. Similarly different markers were used for measurement for the two systems which again could have led to some systematic error. Future research should perhaps consider using the same markers and taking care to measure at the same points. Further research is needed to identify whether the current findings extend to analyses performed using clinical techniques, as well as during other activities such as bilateral leg landings, cutting activities, and other dynamic tasks. Inter-tester reliability in this study only assessed the agreement between the two assessors in extracting the relevant angles from the 2D videos. The potential sources of inter-tester error could be: 1) placement of the markers; 2) digitization of the markers; 3) measure of the angles; the inter-tester reliability in this study only tested the latter two sources, it did not assess the first source.

This study adds to the growing body of evidence suggesting 2D video analysis of a variety of single leg tasks, have a reasonable association to what is being measured using 3D motion capture. The findings of this study also show the approach has good reliability, within and between sessions and also between examiners.

REFERENCES

1. Ageberg E, Bennell K, Hunt M, Simic M, Roos E, Creaby M. Validity and inter-rater reliability of medio-lateral knee motion observed during single limb mini squat BMC Musculoskel Dis 2010:11;265-273
2. Alenezi F, Jones R, Jones P, Herrington L. The reliability of lower limb biomechanical variables collected during single leg squat and landing tasks. J Electromyo Kinesio 2014:24;718-21
3. Dempster W. Space requirements of the seated operator. In: WADC Technical Report 0hio: L Wright-Patterson Air Force Base; 1959:55–159.
4. Ford K, Myer G, Hewett T. Reliability of landing 3D motion analysis: implications for longitudinal analyses. Med Sci Sports Exs. 2007:39;2021–8.
5. Gwynne C, Curran S. Quantifying frontal plane knee motion during single limb squats: reliability and validity of 2-dimensional measures. Int J Sports Phys Ther 2014:9;898-906
6. McLean S, Walker K, Ford K, Myer G, Hewett T, van den Bogert A. Evaluation of a two dimensional analysis method as a screening and evaluation tool for anterior cruciate ligament injury. Brit J Sports Med. 2005:39;355–362.
7. Malfait B, Sankey S, Azidin R, Deschamps K, Vanrenterghem J, Robinson M. How reliable are lower limb kinematics and kinetics during a drop vertical jump. Med Sci Sports Exs. 2014:46;678–85.
8. Munro A, Herrington L, Carolan M. Reliability of 2-dimensional video assessment of frontal-plane dynamic knee valgus during common screening tasks J Sports Rehab 2012:21;7-11
9. Sorenson B, Kernozek T, Willson J, Ragan R, Hove J. Two and three dimensional relationships between knee and hip kinematic motion analysis: single leg drop jump landings. J Sports Rehab 2015:24;363-372
10. Thomas J, Nelson J, Silverman S. 2005 *Research Methods in Physical Activity*. Champaign, IL: Human Kinetics.
11. Yu B, Gabriel D, Noble L, An K. Estimate of the optimum cut off frequency for the Butterworth low-pass digital filter. J Appl Biomech 1999:15;318–29.
12. Willson J, Davis I. Lower extremity mechanics of females with and without patellofemoral pain across activities with progressively greater task demands. Clin Biomech 2008:23;203–11.
13. Willson J, Ireland M, Davis I. Core strength and lower extremity alignment during single leg squats. Med Sci Sports Exs. 2006:38;945-950.

Figure1: Marker placements and lines for calculation of FPPA

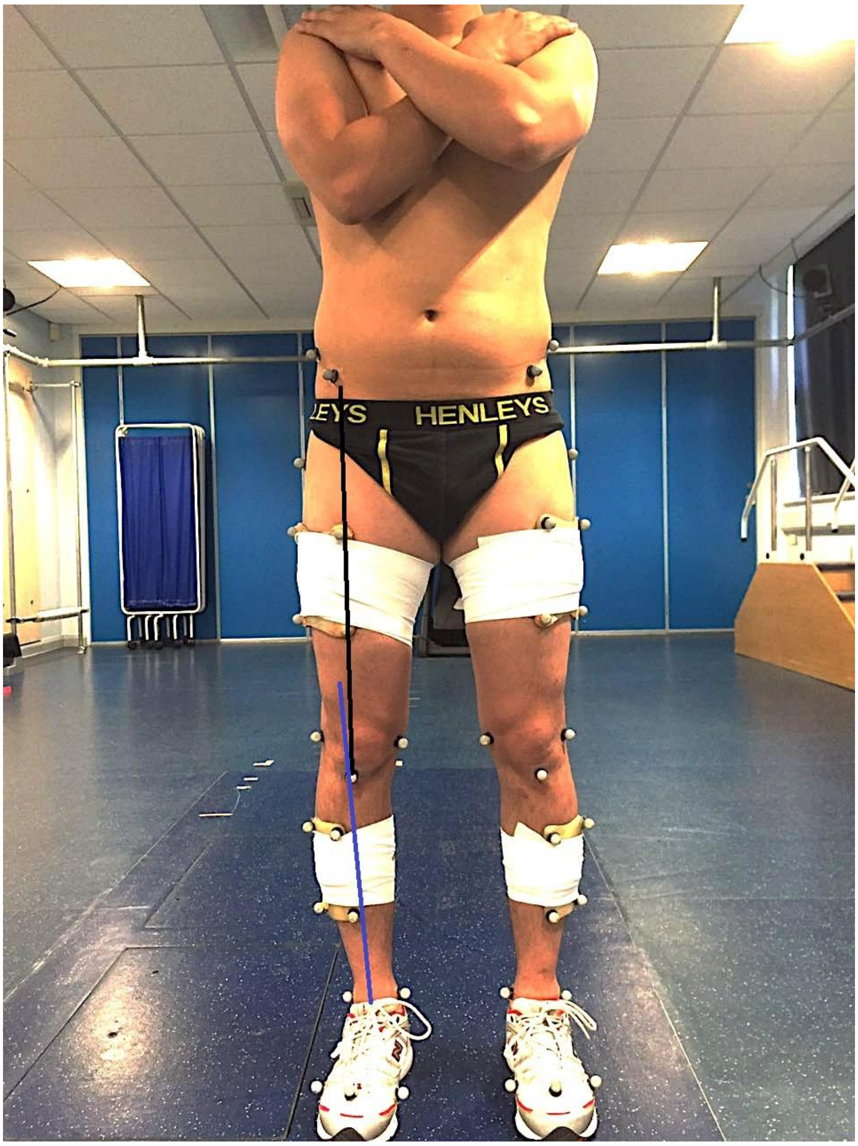

Table 1: Participant demographics

Characteristic	Gender	
	Males (N= 8)	Females (N= 7)
Age (years)	25.0 (±6.4)	26.6 (±3.5)
Height (cm)	171.0 (±6.7)	163.0 (±5.4)
Mass (kg)	69.7 (±10.7)	63.0 (±8.0)

Table 2: Inter-tester reliability

	Single leg Squat		Single leg Land	
	FPPA	Hip Add	FPPA	Hip Add
ICC (95% CI)	0.97 (0.91-0.99)	0.96 (0.89-0.99)	0.99 (0.97-1.0)	0.99 (0.97-1.0)
SEM(°)	1.97	1.96	1.99	1.99

Table 3: Within and between day reliability of 2D video

	Single leg Squat				Single leg land			
	Within session		Between session		Within session		Between session	
	FPPA	Hip Add	FPPA	Hip Add	FPPA	Hip Add	FPPA	Hip Add
ICC (95% CI)	0.72 (0.09-0.92)	0.91 (0.71-0.97)	0.87 (0.57-0.96)	0.79 (0.32-0.94)	0.87 (0.58-0.96)	0.89 (0.65-0.97)	0.87 (0.58-0.96)	0.86 (0.54-0.96)
SEM(°)	1.72	1.37	1.93	1.93	1.43	1.32	1.4	1.43

Table 4: Correlation 2D measurements to 3D motion capture kinematic and kinetic values

	2D video measurements			
3D motion capture measurements	Single leg Squat		Single leg Land	
	FPPA	Hip adduction angle	FPPA	Hip adduction angle
Hip adduction angle	r=0.45(p=0.1)	**r=0.81(p=0.001)**	r=0.13(p=0.65)	**r=0.62(p=0.013)**
Hip adduction moment	r=0.27(p=0.34)	r=0.29(p=0.41)	r=0.22(p=0.44)	r=0.1(p=0.9)
Hip internal rotation angle	r=0.16(p=0.56)	r=0.6(p=0.18)	r=0.33(p=0.23)	r=0.2(p=0.47)
Hip internal rotation moment	r=0.41(p=0.32)	**r=0.45(p=0.09)**	r=0.37(p=0.16)	r=0.1(p=0.78)
Knee Abduction angle	**r=0.79(p=0.008)**	r=0.25(p=0.37)	r=0.21(p=0.79)	r=0.36(p=0.19)
Knee abduction moment	**r=0.65(p=0.015)**	r=0.12(p=0.68)	r=0.36(p=0.19)	r=0.01(p=0.9)

Table 5: Mean (standard deviation) measurements for hip adduction and knee abduction angles (FPPA) across sessions and for second tester (2D video) and for 3D motion capture

Task	Single leg Squat					Single leg landing				
Session	1	1 (Rater 2)	1 (3D)	2	3	1	1 (Rater 2)	1 (3D)	2	3
FPPA (Hip abduction)	-9.1 (10.6)	-11.3 (11.3)[1]	-7.8 (4.7)[2]	-10.2 (7.8)[3]	-11.7 (9.8)[4]	-10.9 (6.4)	-11.2 (6.4)[1]	-8.4 (5.0)[2]	-10.3 (6.5)[3]	-10.9 (5.8)
Hip Adduction	70.6 (8.7)	70.6 (8.3)[1]	74.9 (6.3)[2]	71.7 (7.5)[3]	70.5 (8.2)[4]	79.1 (5.2)	78.8 (5.6)[1]	81.9 (5.7)[2]	80.2 (6.5)[3]	81.2 (6.6)[4]

[1] No significant difference between rater one and rater two (p>0.05)
[2] No significant difference between rater one and 3D motion capture (p>0.05)
[3] No significant difference between session 1 and session 2 (p>0.05)
[4] No significant difference between session 1 and session 3 (p>0.05)