Title

Inter-rater reliability, internal consistency and common technique flaws of the tuck jump assessment in elite female football players

Authors: Lawrence Mayhew (l.mayhew@leedsbeckett.ac.uk) [1], Mark I. Johnson (m.johnson@leedsbeckett.ac.uk) [1, 2], Peter Francis (peter.francis@leedsbeckett.ac.uk) [1], Nicky Snowdon (n.snowdon@shu.ac.uk) [3], Gareth Jones (g.j.jones@leedsbeckett.ac.uk) [1, 2]

Author affiliations

1. School of Clinical and Applied Sciences, Leeds Beckett University, Leeds, U.K.
2. Leeds Pallium Research Group, Leeds, U.K.
3. Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, U.K.

Corresponding author

Gareth Jones, School of Applied and Clinical Sciences, Leeds Beckett University, Leeds, LS1 3HE,
U.K., Telephone: +44 113 8123304, Fax: +44 113 2833124, e-mail: G.J.Jones@leedsbeckett.ac.uk

Article Type: Original Article

Running header: Inter-rater Reliability of Tuck Jump Assessment

Funding: This study was did not receive external funding

Ethical Approval: Ethical approval was granted by Sheffield Hallam University

Contribution of Authors:

- Study design: LM, GJ, NS
- Data collection: LM
- Data analysis: LM,GJ, MIJ
- Preparation of the manuscript: GJ, LM, MIJ, PF

Competing Interests: None

Acknowledgments: None

Accessing Research Materials: Underlying research materials related to your paper (for example data, trial master file documents etc.) can be accessed by contacting Mr Lawrence Mayhew

Key words: Tuck Jump Assessment; landing faults; female football; reliability, internal consistency

ABSTRACT

Injury rates between elite female and male players are comparable, although female players are more likely to sustain an anterior cruciate ligament (ACL) injury. The common mechanism of ACL injury is non-contact trauma sustained when landing from a jump. The Tuck Jump Assessment (TJA) is commonly used in football to assess jump landing technique. The aims of this study were to determine inter-rater agreement and internal consistency of the TJA and to identify commonly occurring technique flaws. Sixty elite female football players (mean (SD): age = 20.27 ± 3.44yrs) were video recorded whilst undertaking the TJA and independently assessed by four raters. Clinically acceptable levels of agreement were reached for ‘Lower extremity valgus at landing’ k = .83 (95% CI, .72 – .93); ‘Thighs do not reach parallel’ k = .84 (95% CI, .74 - .94); ‘Thighs not equal side to side’ k = .86 (95% CI, .75 - .96). The level of agreement for the composite score of all 10 criteria ranged from kw = .62 (95% CI, .48 – .76) to kw = .80 (95% CI, .70 – .90) suggesting a ‘fair-to-very good’ level of inter-rater agreement. The most common technique flaws were found for criterion 2 ‘Thighs do not reach parallel’ (N=147/665) and criterion 1 ‘Knee valgus on landing’ (N=80/665). However, internal consistency results suggest that the TJA is not unidimensional. We suggest ‘Knee valgus on landing’ may have clinical utility although further research is needed.

INTRODUCTION

The advent of professionalism and the exponential rise in the number of UEFA registered female football players has corresponded with a significant increase in the reported incidence of injury. The estimated incidence of injury for female players is between 12.6 and 24.0 injuries per 1000 hours of match play and between 1.2 and 7.0 injuries per 1000 hours of training (Giza, Mithofer, Farrell, Zarins, & Gill, 2005; Le Gall, Carling, & Reilly, 2008; Nilstad, Andersen, Bahr, Holme, & Steffen, 2014). Although injury rates between elite female and male players are comparable (Hagglund, Walden, & Atroshi, 2009), female players are more likely to sustain an anterior cruciate ligament (ACL) injury of the knee than their male counterparts (Walden, Hagglund, Magnusson, & Ekstrand, 2011). ACL injury in female players is more likely to occur at an earlier age (Renstrom et al., 2008) and a previous history of ACL injury is considered to be a significant risk factor for reinjury (Faude, Junge, Kindermann, & Dvorak, 2006). The most common mechanism of ACL injury is an acute non-contact trauma sustained during rapid decelerating movements, for example when landing from a jump (Walden et al., 2011). Reduced neuromuscular control during landing may result in increased knee valgus angles termed ligament dominance and increase the likelihood of an individual sustaining an ACL injury (Hewett, Myer, Ford, et al., 2005).

Observational screening tools are commonly used to assess jump landing tasks to identify faulty movement patterns. These screening tools include the Landing Error Scoring System (LESS) (Padua et al., 2009), the Drop Jump test (Barber-Westin, Smith, Campbell, & Noyes, 2010) and the Tuck Jump

Assessment (TJA) (Myer, Ford, & Hewett, 2008). There is a paucity of evidence to support the validity and reliability of these tools (Frohm, Heijne, Kowalski, Svensson, & Myklebust, 2012; Kiesel, Plisky, & Voight, 2007).

The TJA is widely used to assess jump and landing performance and forms part of the battery of physiological tests used to screen players within the English Women's Super League and the English Women's National Team. Performance during the TJA is scored using a 10 item observational tool that documents technique flaws associated with the jump landing action for knee and thigh motion, foot position during landing, and plyometric technique (Herrington, Myer, & Munro, 2013; Myer et al., 2008). An individual is identified as requiring interventions to correct technique flaws if the TJA composite score is ≥6. There is an ongoing debate about the clinical utility of the TJA and a lack of empirical evidence to support the choice of a cut-off point of ≥6 (Klugman, Brent, Myer, Ford, & Hewett, 2011; Myer et al., 2008; Myer, Ford, Khoury, Succop, & Hewett, 2011). Moreover, there is limited research on inter-rater reliability of the TJA (Dudley et al., 2013; Herrington et al., 2013; Read, Oliver, de Ste Croix, Myer, & Lloyd, 2016), although a recent study of 50 elite level male youth football players concluded that the TJA criterion for knee valgus was a reliable measure of landing performance (Read et al., 2016).

To date, there have been no studies that have investigated technique flaws associated with the TJA in elite female football players. The primary aim of this study was to determine the inter-rater reliability and degree of internal consistency of the TJA. The secondary aim was to identify the most commonly occurring technique flaws in elite female football players.

METHOD

Design

This study was designed to measure inter-rater reliability of two physiotherapists and two sports and conditioning coaches who independently scored elite female football players undertaking a tuck jump test.

Participants

Participants were 60 elite international female football players (mean ± SD: age = 20.27 ± 3.44yrs; height = 168.02 ± 5.26cm; mass = 62.54 ± 6.33kg) who were medically fit to complete mandatory physiological screening. Each participant completed a tuck jump test that was video recorded and subsequently assessed by each of the four raters. Assessment of the tuck jump test is routinely included in a battery of physiological tests used at the team's international training camps and therefore all players were familiar with the tuck jump procedure. Written consent was provided by all participants and raters, and ethical approval was granted by Sheffield Hallam University.

Raters

Two physiotherapy and two strength a conditioning coaching staff from the Women's English Football Association independently scored a video recording of each player undertaking the tuck jump test. Raters regularly used the TJA as a screening tool and were experienced with the TJA scoring process. Characteristics of the raters were:

- Rater 1: Physiotherapist - [redacted] Women's Football Association (FA) with 5 years of experience in elite female football; 10 years post qualifying experience (Health & Care Professions Council registered Physiotherapist)
- Rater 2: Physiotherapist - [redacted] for a women's super league team with 3 years of experience at an FA girls centre of excellence; 11 years post qualifying experience (Health & Care Professions Council registered Physiotherapist)
- Rater 3: Strength and Conditioning Coach - [redacted] Women's FA with 8 years of experience in elite football; 11 years post qualifying experience (United Kingdom Strength & Conditioning Association accredited)
- Rater 4: Strength and Conditioning Coach - [redacted] who had worked with multi-sport elite athletes and had 1 year post qualifying experience with football players of a national standard (United Kingdom Strength & Conditioning Association accredited)

The four raters had a total of 17 years' experience working with female football players at national and international standard.

Procedures

A video recording was taken of each player completing a tuck jump test on an indoor artificial 4G playing surface. All players wore 'astro turf' football shoes. Ambient temperature and humidity were not controlled during testing. The tuck jump test was facilitated by the Principal Investigator who provided standardised verbal instructions and a practical demonstration of tuck jumps to each participant immediately before they took the tuck jump test. A video recording of individual tuck jumps from the sagittal and coronal was made using two Sony PJ410 High Definition cameras on tripods. The tuck jump test was standardised in-line with previously published protocols (Dudley et al., 2013; Herrington et al., 2013; Myer et al., 2008). Two strips of 2.5cm tape were placed 20cm apart and aligned parallel to each other. Participants were instructed to stand with one foot on each tape strip and to perform repeated tuck jumps for 10 seconds, lifting their knees to be level with the hips in the horizontal plane, and to return to the start position. Participants were encouraged to use a high level of effort. No feedback was given to participants whilst they performed tuck jumps.

Each rater independently scored the tuck jump test of each participant by watching the video in real time. In order to standardise the test raters were instructed to view the video of each participant no more than 3 times prior to scoring their tuck jumps across the 10 criterion of the TJA screening tool (Dudley et al., 2013; Herrington et al., 2013; Myer et al., 2008). A score of 1 was assigned if the participant failed to meet an individual criterion on any occasion during the test (i.e. had a technique flaw) and a score of 0 was assigned if the participant did not exhibit a technique flaw. The total score was calculated for each participant with higher scores indicative of poorer performance.

Data Analysis

Raw data was screened for anomalies including data inputted incorrectly. A one-variable x^2 test was conducted to measure the association between the observed and expected frequencies of technique flaws. The minimum number of participants required to detect a kappa coefficient as statistically significant when the value of kappa (K) was set at k =.00 (with 80% power) was n=39 (Sim & Wright, 2005). Fleiss Kappa (an extension of the Cohen's kappa coefficient (k) for two raters) was utilised to assess multiple inter-rater agreement for each TJA criterion with standard error of measurement (SEM) and 95% confidence intervals (CI). The significance level was set at $p < 0.05$. Microsoft Office Excel 2010 was used to compute Fleiss Kappa. A weighted kappa (*Kw*) was performed on the total score to calculate the degree of disagreement. The interpretation of Cohen's kappa coefficient utilised arbitrary theoretical values set by Fleiss et al. (2003) as < 0.40 poor, 0.41 – 0.75 fair to good and 0.75 – 1.00 very good, with > 0.75 used as a cut off for clinically acceptable measure of inter-rater agreement (Sim & Wright, 2005).

Internal consistency of total scores was assessed by Cronbach's alpha reliability coefficient. There is no consensus for the lower limit of the coefficient so the following rules of thumb were applied: > .9 – Excellent, > .8 – Good, > .7 – Acceptable, > .6 – Questionable, > .5 – Poor, and < .5 – Unacceptable (George & Mallory, 2003). Cronbach's alpha analysis was performed using SPSS version 21.

RESULTS

Frequency of technique flaws

The sum of technique flaws scored by all four raters was 665 (Table 1). The most frequent technique flaw was Criterion 2 ‘Thighs do not reach parallel’ (N=147/665, 22%), the second most frequent technique flaw was criterion 1 ‘Knee valgus on landing’ (N=80/665, 12%) and the least frequent technique flaw was Criterion 9 ‘Pause between jumps’ (N=23/665, 4%).

[Insert Table 1 here]

The x^2 value of 152.1, DF=9 had an associated probability value of 0.0001. Thus we can accept that there was a significant difference between the observed and expected frequencies.

The frequency of technique flaws within each of the categories of the TJA (Knee and thigh motion, comprising 3 criterion; Foot position during landing, comprising 5 criterion; and Foot position during landing, comprising 2 criterion) were calculated relative to the maximum number of technique flaws possible was calculated (i.e. (60 participants x 4 raters) x the number of criterion included in each sub-category) There were 234/720 (32.5%) technique flaws for Knee and thigh motion 307/1200 (46%) technique flaws for Foot position during landing and 64/480 (13.3%) technique flaws for Plyometric technique.

Inter-rater agreement

The Fleiss kappa coefficient values used to determine inter-rater agreement ranged from ‘fair-to-good’, k = .46 (95% CI, .35 - .56) to ‘very good’ k = .86 (95% CI, .74 - .94). Raters reached substantial agreement for ‘Lower extremity valgus at landing’ k = .83 (95% CI, .72 – .93); ‘Thighs do not reach parallel (peak of jump)’ k = .84 (95% CI, .74 - .94); ‘Thighs not equal side to side’ k = .86 (95% CI, .75 - .96). A descending order of inter-rater agreement from criterion 1 to criterion 10 was observed in the results.

[Insert Table 2 here]

Weighted kappa (k_w) coefficient values used to determine inter-rater agreement of the composite score ranged from k_w = .62 (95% CI, .48 – .76) to k_w = .80 (95% CI, .70 – .90) suggesting a ‘fair-to-very good’ level of inter-rater agreement.

[Insert Table 3 here]

Internal Consistency

Low alpha values were detected across all four raters for the entire TJA scale. Internal consistency was reassessed with items 9 and 10 removed (Plyometric technique) as the repeated plyometric nature of the TJA over a 10 second period differentiates it from previous tests such as the Landing Error Scoring System (Padua et al 2015). As an 8 item scale there were negligible alterations in internal consistency (range α = .091 – .161, Table 4).

Internal consistency results suggest that the TJA scale and sub items are not unidimensional

[Insert Table 4 here]

DISCUSSION

Statement of principal findings

This is the first study to investigate technique flaws associated with the TJA in elite female football players. The TJA was designed for use with athletic populations to detect technique flaws in jump landing tasks (Myer et al., 2008). In our study four raters identified 665 technique flaws in 60 participants. The most frequent flaws were ‘Thighs do not reach parallel’ (criterion 2) and ‘Knee valgus on landing’ (criterion 1), which are part of the ‘Knee and thigh motion’ category of the TJA. The least frequent technique flaws were criterion 9 ‘Pause between jumps’ and criterion 10 ‘Technique declines prior to 10 seconds’, which form the ‘Plyometric technique’ category of the TJA. The inter-rater level of agreement for the total score of the TJA was ‘fair-to-very good’ with all criteria of the ‘Knee and thigh motion’ category reaching clinically acceptable levels of agreement. Low alpha values for internal consistency suggest the individual criteria contained within the TJA are not unidimensional therefore they are not measuring the same underlying construct (i.e. jump landing task).

Meaning of the study findings

The TJA is currently used by teams within the English Women’s Super League and has been used by the English Women’s National Teams. The TJA is utilised by coaches and medical staff as a screening tool to assess performance of jump landing tasks, and as an outcome measure in regard of neuromuscular retraining.

Previous studies investigating the TJA have not clearly identified the frequency of individual technique flaws and this limits our ability to compare between studies. In our study criterion 2 ‘Thighs do not reach parallel’ was the most frequently identified technique flaw and ‘Pause between jumps’ was the least frequently identified technique flaw. Dudley et al. (2013) also reported criterion 2 as the most frequently identified technique flaw but did not report the rank of other TJA criteria.

Herrington et al. (2013) reported the inter-rater level of agreement for the composite score of the TJA using 2 raters to be very good/excellent (K=0.88) in a sample of ten athletes. The inter-rater percentage of exact agreement between raters across all ten criteria was 93% (range 80%-100%, i.e. high). Interestingly, Dudley et al. (2013) reported the inter-rater level of agreement using 5 Raters to be poor in 40 recreationally active university students (ICC=0.47, 95% CI 0.33-0.62). Read et al. (2016) used a test-retest design to investigate intersession reliability of the TJA in 50 elite male youth football players. Although reliability was found to be strong (ICC=0.88) the authors suggested caution in interpreting the composite score of the TJA due to high within-subject variation in a number of individual criterion.

The difference in the reported levels of agreement may in part be explained by the statistical test selected by investigators. Sample sizes of at least 50 are recommended when using percentage of exact agreement (Birkimer & Brown, 1979). Therefore results from studies containing smaller sample sizes are quite probably the result of chance agreement and should be considered with caution. Each of the TJA criteria is scored in a dichotomous manner i.e. flaw occurred or no flaw occurred and the data is therefore characterised as nominal. Kappa coefficients are recommend for use as the preferred statistical test to determine the inter-rater level of agreement for nominal data (Hallgren, 2012). We utilised Fleiss Kappa to determine inter-rater agreement for individual TJA criteria and a weighted Kappa to determine inter-rater agreement for the composite score.

Cronbach's alpha is considered to infer the degree to which the criteria measures a single unidimensional construct. Our internal consistency statistics raise concern about the construct validity of the TJA suggesting redundancy of TJA criteria. Analysis with the 'Plyometric technique' category removed to determine if the psychometric properties of the test would be improved as an 8 item measure found that unidimensionality remained violated. However it is important to note that jump landing is a skill characterised by multiple factors.

Errors are also introduced into TJA by variability in the interpretation of what constitutes the occurrence of a technique flaw. Dudley et al. (2013) claims that instructions used to assess performance during the tuck jump test do not specify whether a technical flaw should be scored by the rater if observed only on a single occasion or whether it needs to occur repeatedly and consistently throughout the assessment, lead to inconsistency of scoring between assessors. Our raters were instructed to score the presence of a technique flaw when they identified a flaw within the time frame of the test i.e 10 seconds.

Myer et al. (2008) suggest that individuals with a total TJA score of ≥6 have an increased risk of sustaining an ACL injury and interventions to address landing errors should be employed. To our knowledge there is no empirical evidence to support the use of a cut point of ≥6. The results from the present study suggest that TJA criterion are not internally consistent and do not have a coherent empirical structure (i.e. are not interrelated). However if multiple items were highly interrelated then a case could be made that some items should be removed as they are measuring the same thing. It is important to note that although the items were not internally consistent it does necessarily mean that the composite score is not meaningful.

A recent study by Read et al. (2016) concluded that only the knee valgus criterion could be reliably used to screen elite youth male football players as a measure of landing performance. A prospective study by (Hewett, Myer, Ford, et al., 2005) found increased knee abduction angles (knee valgus) during a plyometric activity to be a significant predictor of ACL injury. In our study 'Knee valgus on landing' was the second commonest technique flaw reaching clinically acceptable levels of agreement. ACL strain from valgus knee loading has been confirmed through cadaver, in vivo and 3-dimensional motion analysis methods (Fukuda et al., 2003; Kanamori et al., 2000; Markolf et al., 1995). Increased internal hip rotation, coupled with increased external rotation of the tibia (dynamic knee valgus) has been found in female football players during jump landing and these have been used to predict ACL injury (Alentorn-Geli et al., 2009; Barber-Westin et al., 2010). Female athletes have been found to preferentially rely on increased quadriceps recruitment relative to hamstring recruitment during incremental vertical jump test using surface electromyography (Myer, Brent, Ford, & Hewett, 2011). In addition, a quadriceps dominant landing strategy may increase the risk of sustaining an ACL rupture (Alentorn-Geli et al., 2009; Hewett, Myer, & Ford, 2005).

Limitations of the study

A number of limitations need to be considered when interpreting our study findings. It is possible that the sequence of items in the TJA impacts on recall rates because we observed a trend of decreasing item frequencies and kappa scores through items 1 to 10. Furthermore, Cronbach's alpha is considered a crude measure of reliability (coefficient of reliability) and can be influenced by the number of scale items and redundant items (DeVellis, 2012). Exploratory factor analysis would have provided a more in-depth assessment of the factor structure and dimensionality of the TJA, although with such low internal consistency scores pursuing exploratory factor analysis at this stage may not have provided any further meaningful information. Thus, in future investigators may wish to consider the 'Knee valgus on landing' criterion during jump landing tasks as a predictor of ACL injury in female football.

Conclusion

There is a paucity of studies evaluating the psychometric properties of the TJA and those that exist have inconsistent findings (Dudley et al., 2013; Herrington et al., 2013; Read et al., 2016). Our study found that the criterion used in the TJA are not measuring the same underlying construct (i.e. jump landing task). This raises doubt about the clinical utility of the TJA in its current form. The TJA was intended for use in elite athletes, and assessors that were experienced in its use. Our study was concordant with these directives. Thus, we recommend that assessors should remain cautious when interpreting the composite score of the TJA. The three individual criterion that contribute to the 'Knee and thigh motion' category reached clinically acceptable levels of agreement and may be useful when assessing athletic performance of jump landing tasks. In addition the 'Knee valgus on landing' criterion may have clinical utility and contribute to the screening of elite female football players for potential ACL injury risk. We hope our study catalyses further research in this field.

Tables & Figures

TJA Criterion	Frequency response		
	Total flaws	% of total flaws available	% of relative flaws
Knee & thigh motion			
Lower extremity valgus at landing	80	33.3	12
Thighs do not reach parallel	147	61	21.1
Thighs not equal side to side	67	28	10.1
Foot position during landing			
Foot placement not shoulder width apart	67	28	10.1
Foot placement not parallel	68	28.2	10.2
Foot contact timing not equal	50	20.2	7.5
Excessive landing noise	44	18.3	6.6
Does not land in the same footprint	78	33	11.7
Plyometric technique			
Pause between jumps	23	9.5	3.5
Technique declines prior to 10seconds	41	17	6.2

Table 1: Frequency response of each TJA criterion listed within respective TJA categories; ‘knee & thigh motion’; ‘foot position during landing’; ‘plyometric technique’

Table 2: Fleiss Kappa Inter-rater agreement of TJA criterion

TJA criteria	Fleiss Kappa (κ) Inter-rater agreement
Lower extremity valgus at landing	κ = .83 (95% CI, .72 – .93), $p < .000$
Thighs do not reach parallel (peak of jump)	κ = .84 (95% CI, .74 - .94), $p < .000$
Thighs not equal side to side	κ = .86 (95% CI, .75 - .96), $p < .000$
Foot placement not shoulder width apart	κ = .75 (95% CI, .65 - .85), $p < .000$
Foot placement not parallel (front and back)	κ = .73 (95% CI, .62 - .82), $p < .000$
Foot contact timing not equal	κ = .70 (95% CI, .60 - .81), $p < .000$
Does not land in the same footprint	κ = .60 (95% CI, .50 - .71), $p < .000$
Excessive landing noise	κ = .63 (95% CI, .53 - .73), $p < .000$
Pause between jumps	κ = .60 (95% CI, .49 - .69), $p < .000$
Technique declines prior to 10seconds	κ = .46 (95% CI, .35 - .56), $p < .000$

Table 3: Weighted Kappa Inter-rater agreement of TJA criterion

Paired raters	Weighted Kappa *(K_W)* Inter-rater agreement (Sum score)
$PT_1 : PT_2$	κ_w = .65 (95% CI, .51 – .79)
$PT_1 : SC_1$	κ_w = .80 (95% CI, .70 – .90)
$PT_1 : SC_2$	κ_w = .67 (95% CI, .54 – .80)
$PT_2 : SC_1$	κ_w = .70 (95% CI, .54 – .84)
$PT_2 : SC_2$	κ_w = .79 (95% CI, .69 – .88)
$SC_1 : SC_2$	κ_w = .62 (95% CI, .48 – .76)

Abbreviations: TJA: tuck jump assessment, PT: physiotherapist, SC: strength & conditioning coach

Table 4: Internal consistency

Cronbach's Alpha (α)	Rater 1 (PT_1)	Rater 2 (PT_2)	Rater 3 (SC_1)	Rater 4 (SC_2)
Entire scale	.073	-.033	.018	.129

TJA categories	Rater 1 (PT_1)	Rater 2 (PT_2)	Rater 3 (SC_1)	Rater 4 (SC_2)
Knee & Thigh motion	-.397	-.720	-.653	-.509
Foot position during landing	.288	.163	.220	.191
Plyometric technique	.528	.306	.222	.339
With items 9 & 10 removed	.161	.091	.112	.154

Abbreviations: TJA: tuck jump assessment, PT: physiotherapist, SC: strength & conditioning coach

References

Alentorn-Geli, E., Myer, G. D., Silvers, H. J., Samitier, G., Romero, D., Lazaro-Haro, C., & Cugat, R. (2009). Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 1: Mechanisms of injury and underlying risk factors. *Knee Surg Sports Traumatol Arthrosc, 17*(7), 705-729. doi: 10.1007/s00167-009-0813-1

Barber-Westin, S. D., Smith, S. T., Campbell, T., & Noyes, F. R. (2010). The drop-jump video screening test: retention of improvement in neuromuscular control in female volleyball players. *J Strength Cond Res, 24*(11), 3055-3062. doi: 10.1519/JSC.0b013e3181d83516

Birkimer, J. C., & Brown, J. H. (1979). Back to basics: Percentage agreement measures are adequate, but there are easier ways. *J Appl Behav Anal, 12*(4), 535-543.

Boden, B. P., Torg, J. S., Knowles, S. B., & Hewett, T. E. (2009). Video analysis of anterior cruciate ligament injury: abnormalities in hip and ankle kinematics. *Am J Sports Med, 37*(2), 252-259. doi: 10.1177/0363546508328107

DeVellis, R. F. (2012). Scale Development: Theory and Applications (3rd edtion). Los Angeles, CA. Sage.

Dudley, L. A., Smith, C. A., Olson, B. K., Chimera, N. J., Schmitz, B., & Warren, M. (2013). Interrater and Intrarater Reliability of the Tuck Jump Assessment by Health Professionals of Varied Educational Backgrounds. *J Sports Med (Hindawi Publ Corp), 2013*, 483503. doi: 10.1155/2013/483503

Faude, O., Junge, A., Kindermann, W., & Dvorak, J. (2006). Risk factors for injuries in elite female soccer players. *Br J Sports Med, 40*(9), 785-790. doi: 10.1136/bjsm.2006.027540

Ford, K. R., Manson, N. A., Evans, B. J., Myer, G. D., Gwin, R. C., Heidt, R. S., Jr., & Hewett, T. E. (2006). Comparison of in-shoe foot loading patterns on natural grass and synthetic turf. *J Sci Med Sport, 9*(6), 433-440. doi: 10.1016/j.jsams.2006.03.019

Frohm, A., Heijne, A., Kowalski, J., Svensson, P., & Myklebust, G. (2012). A nine-test screening battery for athletes: a reliability study. *Scand J Med Sci Sports, 22*(3), 306-315. doi: 10.1111/j.1600-0838.2010.01267.x

Fukuda, Y., Woo, S. L., Loh, J. C., Tsuda, E., Tang, P., McMahon, P. J., & Debski, R. E. (2003). A quantitative analysis of valgus torque on the ACL: a human cadaveric study. *J Orthop Res, 21*(6), 1107-1112. doi: 10.1016/s0736-0266(03)00084-6

Giza, E., Mithofer, K., Farrell, L., Zarins, B., & Gill, T. (2005). Injuries in women's professional soccer. *Br J Sports Med, 39*(4), 212-216; discussion 212-216. doi: 10.1136/bjsm.2004.011973

Hagglund, M., Walden, M., & Atroshi, I. (2009). Preventing knee injuries in adolescent female football players - design of a cluster randomized controlled trial [NCT00894595]. *BMC Musculoskelet Disord, 10*, 75. doi: 10.1186/1471-2474-10-75

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol, 8*(1), 23-34.

Herrington, L., Myer, G. D., & Munro, A. (2013). Intra and inter-tester reliability of the tuck jump assessment. *Phys Ther Sport, 14*(3), 152-155. doi: 10.1016/j.ptsp.2012.05.005

Hewett, T. E., Myer, G. D., & Ford, K. R. (2005). Reducing knee and anterior cruciate ligament injuries among female athletes: a systematic review of neuromuscular training interventions. *J Knee Surg, 18*(1), 82-88.

Hewett, T. E., Myer, G. D., Ford, K. R., Heidt, R. S., Jr., Colosimo, A. J., McLean, S. G., . . . Succop, P. (2005). Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med, 33*(4), 492-501. doi: 10.1177/0363546504269591

Kanamori, A., Woo, S. L., Ma, C. B., Zeminski, J., Rudy, T. W., Li, G., & Livesay, G. A. (2000). The forces in the anterior cruciate ligament and knee kinematics during a simulated pivot shift test: A human cadaveric study using robotic technology. *Arthroscopy, 16*(6), 633-639. doi: 10.1053/jars.2000.7682

Kiesel, K., Plisky, P. J., & Voight, M. L. (2007). Can Serious Injury in Professional Football be Predicted by a Preseason Functional Movement Screen? *N Am J Sports Phys Ther, 2*(3), 147-158.

Klugman, M. F., Brent, J. L., Myer, G. D., Ford, K. R., & Hewett, T. E. (2011). Does an in-season only neuromuscular training protocol reduce deficits quantified by the tuck jump assessment? *Clin Sports Med, 30*(4), 825-840. doi: 10.1016/j.csm.2011.07.001

Le Gall, F., Carling, C., & Reilly, T. (2008). Injuries in young elite female soccer players: an 8-season prospective study. *Am J Sports Med, 36*(2), 276-284. doi: 10.1177/0363546507307866

Markolf, K. L., Burchfield, D. M., Shapiro, M. M., Shepard, M. F., Finerman, G. A., & Slauterbeck, J. L. (1995). Combined knee loading states that generate high anterior cruciate ligament forces. *J Orthop Res, 13*(6), 930-935. doi: 10.1002/jor.1100130618

Myer, G. D., Brent, J. L., Ford, K. R., & Hewett, T. E. (2011). Real-time assessment and neuromuscular training feedback techniques to prevent ACL injury in female athletes. *Strength Cond J, 33*(3), 21-35. doi: 10.1519/SSC.0b013e318213afa8

Myer, G. D., Ford, K. R., & Hewett, T. E. (2008). Tuck Jump Assessment for Reducing Anterior Cruciate Ligament Injury Risk. *Athl Ther Today, 13*(5), 39-44.

Myer, G. D., Ford, K. R., Khoury, J., Succop, P., & Hewett, T. E. (2011). Biomechanics laboratory-based prediction algorithm to identify female athletes with high knee loads that increase risk of ACL injury. *Br J Sports Med, 45*(4), 245-252. doi: 10.1136/bjsm.2009.069351

Nilstad, A., Andersen, T. E., Bahr, R., Holme, I., & Steffen, K. (2014). Risk factors for lower extremity injuries in elite female soccer players. *Am J Sports Med, 42*(4), 940-948. doi: 10.1177/0363546513518741

Padua, D. A., Marshall, S. W., Boling, M. C., Thigpen, C. A., Garrett, W. E., Jr., & 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. *Am J Sports Med, 37*(10), 1996-2002. doi: 10.1177/0363546509343200

Read, P. J., Oliver, J. L., de Ste Croix, M. B., Myer, G. D., & Lloyd, R. S. (2016). Reliability of the Tuck Jump Injury Risk Screening Assessment in Elite Male Youth Soccer Players. *J Strength Cond Res, 30*(6), 1510-1516. doi: 10.1519/JSC.0000000000001260

Renstrom, P., Ljungqvist, A., Arendt, E., Beynnon, B., Fukubayashi, T., Garrett, W., . . . Engebretsen, L. (2008). Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement. *Br J Sports Med, 42*(6), 394-412. doi: 10.1136/bjsm.2008.048934

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther, 85*(3), 257-268.

Walden, M., Hagglund, M., Magnusson, H., & Ekstrand, J. (2011). Anterior cruciate ligament injury in elite football: a prospective three-cohort study. *Knee Surg Sports Traumatol Arthrosc, 19*(1), 11-19. doi: 10.1007/s00167-010-1170-9

Zazulak, B. T., Hewett, T. E., Reeves, N. P., Goldberg, B., & Cholewicki, J. (2007). The effects of core proprioception on knee injury: a prospective biomechanical-epidemiological study. *Am J Sports Med, 35*(3), 368-373. doi: 10.1177/0363546506297909