



Contents lists available at ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Original Research

Between-session and inter-rater reliability of the modified tuck jump assessment in healthy adult athletes

Alli Gokeler ^{a, b, *}, Bart Dingenen ^c^a Luxembourg Institute of Research in Orthopedics, Sports Medicine and Science (LIROMS), Luxembourg^b Exercise Science & Neuroscience, Department Exercise & Health, Faculty of Science, Paderborn University, Paderborn, Germany^c Reval Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Agoralaan, 3590, Diepenbeek, Belgium

ARTICLE INFO

Article history:

Received 27 December 2018

Received in revised form

11 February 2019

Accepted 11 February 2019

Keywords:

Neuromuscular control

Jumping

Screening/Exercise test

Rehabilitation

Reliability

ABSTRACT

Objectives: This study aimed to determine the between-session and inter-rater reliability of the modified tuck jump assessment (TJA).**Design:** Reliability study.**Setting:** Human movement laboratory.**Participants:** Sixteen recreational participants (8 females, 8 males; mean age: 22.4 years) participated in the study. Participants performed the modified TJA in two sessions separated by one week.**Main outcome measures:** Between-session and inter-rater reliability were determined by two independent raters. The standard error of measurement (SEM) and smallest detectable difference (SDD) were calculated.**Results:** The between-session reliability for the total scores was excellent for rater 1 (ICC = 0.93, 95% CI = 0.78–0.98) and rater 2 (ICC = 0.96, 95% CI = 0.89–0.99), with SEM's and SDD's ranging between 0.5–0.7 and 1.3–2.1 respectively. The inter-rater reliability for the total scores was good for session 1 (ICC = 0.85, 95% CI = 0.58–0.95) and session 2 (ICC = 0.88, 95% CI = 0.66–0.96). In general, there was poor agreement for item analysis for both raters between sessions as well as between raters.**Conclusion:** The modified TJA revealed excellent between-session and good inter-rater reliability for the total score. Reliability for the between-session was poor for the individual items. The relatively poor reliability for the individual items imply that caution is warranted when solely interpreting the total score.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Analyzing athletes' motion during specific tasks may enable the identification of those at risk of injury (Hewett et al., 2005). In particular, athletes participating in pivoting sports may be at risk for lower extremity injuries such as ankle sprains, patellofemoral pain or anterior cruciate ligament (ACL) injuries (Kilic, Maas, Verhagen, Zwerver, & Gouttebauge, 2017; Myer et al., 2015). Traditionally, movement analysis has been assessed using laboratory based three-dimensional (3D) motion capture equipment. However, using this equipment is time consuming and very costly.

Several jumping and landing tests have been developed in the

clinical setting, including the landing error scoring system (LESS) (Padua et al., 2009), the drop vertical jump test (Myer, Ford, Khoury, Succop, & Hewett, 2010) and the tuck jump assessment (TJA) (Myer, Ford, & Hewett, 2008). The TJA may offer clinical advantages over the other tests. First, the TJA is a full effort plyometric jump task which may better reflect sport-specific jumping activities (Myer et al., 2008). Second, the TJA protocol requires athletes to perform consecutive jumps for 10 s, while the LESS (Padua et al., 2009) and drop vertical jump (DVJ) test (Myer et al., 2010) require only three jumps. Therefore, the TJA evaluates a measure of endurance, introducing a potential fatigue effect that might highlight landing flaws not observable in three non-consecutive jumps.

The TJA was developed as a practical “clinician-friendly” plyometric assessment, identifying jumping and landing technique flaws pertaining to risk of ACL injury (Myer et al., 2008). Using a double-leg jump task such as the TJA may allow clinicians to detect side-to-side asymmetries. In addition, the TJA may be used to

* Corresponding author. Luxembourg Institute of Research in Orthopedics, Sports Medicine and Science (LIROMS), 76 rue d'Eich, L-1460, Luxembourg.

E-mail address: alli.gokeler@liroms.org (A. Gokeler).

monitor rehabilitation of lower extremity injuries (Herrington, Myer, & Horsley, 2013b). However, the TJA could be also be considered in function of injury screening purposes for other lower extremity injuries such as patellofemoral pain. High knee valgus angles contributing to high knee abduction moments during DVJ landing have been linked prospectively to patellofemoral pain in adolescent girls (Myer et al., 2015). Similarly, the TJA has been shown to identify increased knee valgus angles in (female) athletes during landing (Myer et al., 2008), indicating the potential of the TJA for future injury screening purposes. A recent study demonstrated that both male and female youth athletes demonstrated an improvement in jump-landing technique with the modified TJA with maturation (Fort-Vanmeerhaeghe, Benet, Mirada, Montalvo, & Myer, 2018).

Originally, the jumps and landings during the TJA were analyzed with 10 items (Myer et al., 2008). These 10 items were used to assess the four neuromuscular imbalances that are suggested to be related to ACL injury risk in female athletes (ligament, quadriceps, leg, and trunk dominance) (Hewett, Ford, Hoogenboom, & Myer, 2010). While the TJA is clinically useful (Herrington, Myer, & Munro, 2013a), concerns have arisen related to the traditional dichotomous scoring system which was deemed too restrictive to assess complex movement patterns (Fort-Vanmeerhaeghe, Montalvo, Lloyd, Read, & Myer, 2017). Subsequently, a modified TJA has been developed (Fort-Vanmeerhaeghe et al., 2017) where the dichotomous scale (0 for no flaw or 1 for flaw) (Myer, Brent, Ford, & Hewett, 2011), has been changed into an ordinal scale (Fort-Vanmeerhaeghe et al., 2017). The 2-point dichotomous scale has been changed to a 3-point ordinal scale that may allow clinicians to score jump-landing performance with a slightly larger range given the subtleties often observed that may pose a problem to capture this in a yes/no configuration. Fort-Vanmeerhaeghe et al. (Fort-Vanmeerhaeghe et al., 2017) reported good to excellent intra- and interrater reliability for most of the items of the modified TJA. Although it is recognized that reliability of the modified TJA is an important measurement property, it leaves the question unanswered what the jump-landing variance of healthy athletes between sessions is. Within-session reliability has been established as good (Fort-Vanmeerhaeghe et al., 2017), but the variance of athletes between sessions remains unknown. In order to interpret differences during and following an intervention appropriately, an indication of the variability of the movement patterns between sessions is necessary. This could provide meaningful information to determine whether a change in the score of the modified TJA during and/or after an intervention program reflects a true improvement or lack thereof.

The primary aim of this study was therefore to determine the between-session reliability of the modified TJA in healthy adult athletes. It is common that a patient is under the care of several physiotherapists during a course of rehabilitation. Hence, within the team of physiotherapists, agreement between physiotherapists is important when using the modified TJA. So far inter-rater reliability for the modified TJA has been established for young athletes, but not for adult athletes (Fort-Vanmeerhaeghe et al., 2017). Therefore, the second purpose was to investigate the inter-rater reliability of 2 raters using the modified TJA in adult athletes.

2. Materials and methods

2.1. Participants

Sixteen recreational participants (8 females, 8 males; age: mean \pm SD = 22.4 \pm 1.9 years; weight: mean \pm SD = 68.1 \pm 10.3 kg; height: mean \pm SD = 176.3 \pm 9.3 cm; body mass index: mean \pm SD = 21.8 \pm 2.1 kg/m²) participated in the study.

Participants were recruited by advertisements in local sport clubs. All participants participated in sports involving multidirectional movements (11 football, 4 volleyball, 1 multisport including kick-boxing). Inclusion criteria for the study were (i) age 18–45 years, (ii) no musculoskeletal injuries of the lumbar spine or lower limb in the last 6 months and (iii) willing to sign the informed consent form. Exclusion criteria were (i) history of a major trauma in the previous 6 months prior to the measurements (ii) orthopaedic surgery of the lumbar spine, pelvis or lower limb, (iii) any presence of following conditions/constitutions: neurologic or vestibular disorders, (iiii) current pregnancy. Appropriate ethical approval was granted by the local ethical committee prior to the commencement of the study. Before participating in the study, all participants read and signed the informed consent form.

2.2. Procedure

Participants were instructed to place their feet in the middle of the rectangle marked (41 cm length, 35 cm width) on the floor. This square consisted of four smaller rectangles. In addition, basic instructions were given about how to carry out the test (information on lifting the knees to hip height and attempting to land on the same footprint with their feet shoulder width apart (Fort-Vanmeerhaeghe et al., 2017)). In order to fully understand the verbal instructions, study participants were shown a live demonstration of the correct TJA technique. Participants were then allowed to ask questions to clarify their understanding of what was required in order to correctly perform the test and were provided 3–5 practice trials to ensure accurate interpretation of the tuck jump. Following the warm-up, the TJA commenced on the designated location for 10 s after receiving again basic instructions about how to complete the test. The performance of each test session was recorded using two cameras (iPad Air, Apple, Inc., USA) sampling at 120 frames per second. The frontal and sagittal placed iPads (Apple Inc., USA) were positioned on a portable tripod perpendicular to respectively the frontal and sagittal plane at a height of 0.78 m, and a distance of 3.10 m. The participants were required to wear a short. For males, the upper body was bare whilst females were asked to wear a sports bra to enhance visible tracking of the trunk and lower extremities. The participants were required to wear sports shorts. All participants were tested twice with a one-week interval. Two sports physical therapists with respectively 26 and 9 years of experience independently assessed the videos of the tuck jumps. The scoring was based on a frontal and sagittal views. The videos were analyzed with freely available software (QuickTime 7.9 for Windows, Apple Inc., USA) that allowed for slow motion and frame by frame advancement. Each rater was allowed to play the videos as many times as needed to obtain a score for the tests. This procedure and scoring were according to the method by Fort-Vanmeerhaeghe et al. (Fort-Vanmeerhaeghe et al., 2017)]. Briefly, if participants failed to meet the criteria then they scored 1 or 2. Conversely, if participants met the criteria, then they scored 0 for the respective category. In order for a movement to be considered dysfunctional, the specific technique error had to appear two or more times during the 10-s test. The raters were blinded to the previous score as well as each other's scores, and both raters performed the second assessment one week after the first assessment to prevent recall bias.

2.3. Statistical analysis

To determine the between-session reliability of the total modified TJA scores, the first assessment was compared to the second assessment of both raters individually. To determine the inter-rater reliability, both assessments of rater 1 were compared to

both assessments of rater 2. A two-way random model was used with absolute agreement ($ICC_{2,2}$). The ICC values were interpreted as poor (<0.50), moderate ($0.50–0.74$), good ($0.75–0.89$) or excellent ($0.90–1.00$) (Portney & Watkins, 2000). The standard error of the measurement (SEM) was determined using the following equation, $SEM = SD \cdot \sqrt{1-ICC}$ (Weir, 2005). The smallest detectable difference (SDD) was estimated based on the 95% confidence interval (CI), where $SDD_{95} = SEM \cdot \sqrt{2} \cdot 1.96$ [18]. To determine the between-session and inter-rater reliability of each individual TJA item, Cohen's Kappa (k) values were calculated. The k -values were interpreted as follows: $k < 0$ less than chance agreement; $k \geq 0.01$ to ≤ 0.20 slight agreement; $k \geq 0.21$ to ≤ 0.40 fair agreement; $k \geq 0.41$ to ≤ 0.60 moderate agreement; ≥ 0.61 to ≤ 0.80 substantial agreement; ≥ 0.81 to 0.99 almost perfect agreement; $k = 1$ perfect agreement (Landis & Koch, 1977). All statistical analyses were performed using SPSS (SPSS Science, version 24 for Windows, USA).

3. Results

3.1. Total scores

The between-session reliability for the total scores was excellent for rater 1 ($ICC = 0.93$, 95% CI = $0.78–0.98$) and rater 2 ($ICC = 0.96$, 95% CI = $0.89–0.99$), with SEM's and SDD's ranging between $0.5–0.7$ and $1.3–2.1$ respectively (Table 1). The absolute differences between sessions were relatively small ($0.7–1.3$) (Table 1).

The inter-rater reliability for the total scores was good for session 1 ($ICC = 0.85$, 95% CI = $0.58–0.95$) and session 2 ($ICC = 0.88$, 95% CI = $0.66–0.96$), with SEM's and SDD's ranging between $0.8–1.0$ and $2.3–2.9$ respectively (Table 2). The absolute differences between raters were relatively small ($1.3–1.6$) (Table 2). The mean total scores were 10.9 ± 2.9 and 11.5 ± 2.7 for respectively the first and second session of rater 1, and 11.6 ± 2.4 and 11.6 ± 2.2 for respectively the first and second session of rater 2.

3.2. Item analysis

The Kappa measure of agreement between sessions differed between items and between raters. For rater 1, slight agreement was found for item 6, fair agreement for item 5, moderate agreement for items 2, 7, 9 and 10, substantial agreement for item 1 and almost perfect agreement for item 4. The 3rd item “thighs not equal side-to-side” had a less than chance agreement ($k < 0$) (Table 3). For rater 2, moderate agreement was found for item 9, substantial agreement for item 1, almost perfect agreement for items 2, 4, 6 and 10 and perfect agreement for items 5 and 7. Only the 3rd item “thighs not equal side-to-side” had a less than chance agreement ($k < 0$) (Table 3). Item 8 (“pause between jumps”) was scored 0 for all sessions. Therefore, kappa coefficients could not be calculated for this item.

The inter-rater Kappa coefficients differed largely between items. Slight to fair agreement was found for item 7, fair agreement for items 3 and 6, fair to moderate agreement for item 9, fair to substantial agreement for item 4, moderate agreement for item 2, moderate to substantial agreement for items 5 and 10, substantial to almost perfect agreement for item 1 (Table 4). Item 8 (“pause

between jumps”) was scored 0 for both sessions which precluded to calculate Kappa coefficients for this item.

4. Discussion

The main finding from the current study was that the modified TJA revealed excellent between-session and good inter-rater reliability for the total score. The relative lower reliability for the individual items in comparison with the total scores imply that caution is warranted when solely interpreting the total score. For example, the same total score could be obtained by different combinations of scores on individual items. Others have also cautioned to solely interpreting the total score due to the high within-subject variation in a number of the individual items (Mayhew, Johnson, Francis, Snowdon, & Jones, 2017; Read, Oliver, de Ste Croix, Myer, & Lloyd, 2016).

4.1. Between session reliability of the total scores

In the current study between session-reliability was determined based on analysis of participants performing two sessions of the modified TJA separated by one week.

In clinical practice, any improvement of the modified TJA following an intervention or rehabilitation programs needs to be checked against the clinical significance of the findings. Therefore, it is important to know the measurement properties of the tests being used. This the first study to the best of the author's knowledge to report between session reliability, including the SEM and SDD for the modified TJA. Using both the SEM and SDD data could help clinicians to determine whether a change (e.g. lower or higher score) in the modified TJA score over time, may reflect a true change rather than irrelevant fluctuations. The SEM provides an estimate of one standard deviation of the error associated with a single measurement (Davidson & Keating, 2014). Based on the current study, the SEM shows relative low variability of the modified TJA test caused by measurement error. Using the original TJA, Dudley et al. (Dudley et al., 2013) found that interrater reliability of three raters who completed two scoring sessions improved from 0.52 (95% CI $0.35–0.68$) for session one to 0.69 (95% CI $0.55–0.81$) for session two, one month later. The SDD is an indicator of the confidence that differences in performance are not due to measurement error or subject variability. The SDD was between 1.3 and 2.1 between sessions. Increases or decreases between sessions for the modified TJA that exceed this SDD can provide justification for continued care, indicate the need for additional evaluation or interventions, or identify appropriate benchmarks for care. Based on the results of the current study the SDD, the authors suggest that was at least 2 points decrease or increase on the total score represent a clinical meaningful change between sessions. Our findings suggest that the total scores of the modified TJA may have potential for making longitudinal comparisons of jump landing movement patterns for repeated measures. Although, the TJA has been shown to be reliable 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 (Read et al., 2016).

Table 1
Between-session reliability of the total score of the modified tuck jump assessment.

	Absolute difference between sessions *	$ICC_{2,2}$ (95% CI)	SEM	SDD
Rater 1	1.3 ± 0.8	0.93 ($0.78–0.98$)	0.7	2.1
Rater 2	0.7 ± 0.6	0.96 ($0.89–0.99$)	0.5	1.3

Abbreviations: ICC, intraclass correlation coefficients; CI, confidence interval; SEM, standard error of measurement; SDD, smallest detectable difference.

* Mean \pm standard deviation.

Table 2
Inter-rater reliability of the total score of the modified tuck jump assessment.

	Absolute difference between sessions *	ICC _{2,2} (95% CI)	SEM	SDD
Measurement 1	1.6 ± 1.2	0.85 (0.58–0.95)	1.0	2.9
Measurement 2	1.3 ± 1.0	0.88 (0.66–0.96)	0.8	2.3

Abbreviations: ICC, intraclass correlation coefficients; CI, confidence interval; SEM, standard error of measurement; SDD, smallest detectable difference.

* Mean ± standard deviation.

Table 3
Kappa coefficients for between-session reliability for each item of the modified tuck jump assessment.

Item	Rater 1	Rater 2
1. Lower extremity valgus at landing	.77	.63
2. Thighs do not reach parallel	.52	.80
3. Thighs not equal side-to-side	-.25	-.23
4. Foot placement not shoulder width apart	.86	.88
5. Foot placement not parallel	.29	1.0
6. Foot contact timing not equal	.19	.89
7. Excessive landing contact noise	.59	1.0
8. Pause between jumps	/	/
9. Technique declines prior to 10s	.44	.59
10. Does not land in same footprint	.43	.85

Table 4
Kappa coefficients for inter-rater reliability for each item of the modified tuck jump assessment.

Item	Measurement 1	Measurement 2
1. Lower extremity valgus at landing	.88	.77
2. Thighs do not reach parallel	.46	.56
3. Thighs not equal side-to-side	.25	.39
4. Foot placement not shoulder width apart	.40	.63
5. Foot placement not parallel	.77	.43
6. Foot contact timing not equal	.32	.36
7. Excessive landing contact noise	.24	.09
8. Pause between jumps	/	/
9. Technique declines prior to 10s	.41	.25
10. Does not land in same footprint	.67	.59

4.2. Between session reliability of the individual scores

The Kappa measure of agreement between sessions differed between items, and between raters. Rater 1 had slight to moderate agreement for six items, substantial agreement for one item and almost perfect agreement for one item. Rater 2 had substantial agreement for one item, almost perfect agreement for four items and perfect agreement for two items.

The substantial number of low Kappa scores for the individual items could indicate that the criteria for these items appear to be difficult to apply in practice when using the modified TJA. For example, both rater 1 and 2 in the current study scored 0 for item 8 “pause between jumps” for both sessions, indicating that none of the participants used a pause between jumps. In the instructions how to score this item, Fort- Vanmeerhaeghe et al. (Fort-Vanmeerhaeghe et al., 2017) provided the following criteria: (0) reactive and reflex jumps, (1) small pause between jumps, and (2) large pause between jumps. However, the authors do not explain the difference between small and large (Smith, Chimera, Lininger, & Warren, 2017).

4.3. Inter-rater reliability of the total scores

The inter-rater reliability for the total score of the modified TJA was good based on the ICC (0.85–0.88) values of the two raters involved in this study, which suggests that the modified TJA has the

capacity to be used by multiple examiners to document jump-landing movement patterns in athletes. Previously, the inter-reliability was determined but based on two raters who analyzed the same videos of a single session with a one-week interval and reported inter-rater ICC values of 0.96 (Fort-Vanmeerhaeghe et al., 2017).

4.4. Inter-rater reliability of the individual scores

The inter-rater Kappa coefficients of agreement revealed large differences for the various items. Specifically, scoring the items “thighs not equal side-to-side during flight”, “foot contact timing not equal” and “technique declines prior 10 s” yielded only slight to moderate agreement. Our findings are in line with previous expressed concerns that have arisen pertaining the clarity of some of the items used in the scoring of the modified TJA (Smith et al., 2017). Those authors proposed to score pauses longer than half a second as a large flaw, while those below constitute a small flaw (Smith et al., 2017).

4.5. Population and procedure

Several other factors need to be considered when interpreting the results of the modified TJA. First, the population being tested is essential to consider. Jump performance in adolescent athletes may be influenced by periods of rapid growth which were shown to have landing biomechanics associated with increased injury risk (Read et al., 2018). Fort- Vanmeerhaeghe et al. (Fort-Vanmeerhaeghe et al., 2018) found that knee valgus at landing had a significant interaction between sex and maturation. Youth female athletes demonstrated more deficits for knee valgus at landing and fatigue compared to males regardless of maturation status (Fort-Vanmeerhaeghe et al., 2018). If the modified TJA is used for screening purposes, e.g. yearly at pre-season, changes that are due to maturity should therefore be taken into account. Future research should investigate the reliability of the modified TJA scoring criteria in different types and level of athletes, various age and gender groups. As a consequence, the mean scores on the TJA reported in this study should not be considered as normative values while ignoring these specific individual constraints.

Second, if clinicians want to interpret an improvement in the TJA score, they need to reassure that the athlete performed the test with maximal effort, as an athlete could jump slower or lower while focusing on technique. To overcome this the number of jumps can be counted and recorded to provide reference, in case athletes focus primarily on the landing whilst decreasing performance (Smith et al., 2017).

4.6. Study limitations

Several study limitations need to be acknowledged. First, the findings of the current study should not be extrapolated to athletes of different age groups, athletic level and other sports. Second, not all participants were able to land in the same foot print and subsequently received a flaw. This is however not a weakness of the current study as the protocol of Fort- Vanmeerhaeghe et al. (A Fort-

Vanmeerhaeghe et al., 2017) was followed. They described a marked rectangle of the floor measuring 35 cm in width and 45 cm in length. While this may be feasible in youth athletes, for adult athletes the rectangle may be too small given their larger shoulder width and foot length.

5. Conclusion

The modified TJA revealed excellent between-session and good inter-rater reliability for the total scores, but considerable differences were found for the between-session and inter-rater reliability of the individual items.

Ethical approval

All participants signed written informed consent. The regional ethics committee approved the study (IRB-ID B371201731092).

Conflicts of interest

None declared.

Declarations of interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pts.2019.02.002>.

Funding

None declared.

References

- Davidson, M., & Keating, J. (2014). Patient-reported outcome measures (PROMs): how should I interpret reports of measurement properties? A practical guide for clinicians and researchers who are not biostatisticians. *British Journal of Sports Medicine*, 48(9), 792–796. <https://doi.org/10.1136/bjsports-2012-091704>.
- 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. *The Journal of Sports Medicine*, 1–5, 2013 <https://doi.org/10.1155/2013/483503>.
- Fort-Vanmeerhaeghe, A., Benet, A., Mirada, S., Montalvo, A. M., & Myer, G. D. (2018). Sex and maturation differences in performance of functional jumping and landing deficits in youth athletes. *Journal of Sport Rehabilitation*, 1–22. <https://doi.org/10.1123/jsr.2017-0292>.
- Fort-Vanmeerhaeghe, A., Montalvo, A. M., Lloyd, R. S., Read, P., & Myer, G. D. (2017). Intra- and inter-rater reliability of the modified tuck jump assessment. *Journal of Sports Science and Medicine*, 16(1), 117–124.
- Herrington, L., Myer, G. D., & Munro, A. (2013a). Intra and inter-tester reliability of the tuck jump assessment. *Physical Therapy in Sport*, 14(3), 152–155. <https://doi.org/10.1016/j.pts.2012.05.005>.
- Herrington, L., Myer, G., & Horsley, I. (2013b). Task based rehabilitation protocol for elite athletes following anterior cruciate ligament reconstruction: A clinical commentary. *Physical Therapy in Sport*, 14(4), 188–198. <https://doi.org/10.1016/j.pts.2013.08.001>.
- Hewett, T. E., Ford, K. R., Hoogenboom, B. J., & Myer, G. D. (2010). Understanding and preventing acl injuries: Current biomechanical and epidemiologic considerations - update 2010. *North American Journal of Sports Physical Therapy: NAJSPT*, 5(4), 234–251.
- Hewett, T. E., Myer, G. D., Ford, K. R., Heidt, R. S., Jr., Colosimo, A. J., McLean, S. G., et al. (2005). Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: A prospective study. *The American Journal of Sports Medicine*, 33(4), 492–501.
- Kilic, O., Maas, M., Verhagen, E., Zwerver, J., & Gouttebauge, V. (2017). Incidence, aetiology and prevention of musculoskeletal injuries in volleyball: A systematic review of the literature. *European Journal of Sport Science*, 1, 29. <https://doi.org/10.1080/17461391.2017.1306114>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Mayhew, L., Johnson, M. L., Francis, P., Snowdon, N., & Jones, G. (2017). Inter-rater reliability, internal consistency and common technique flaws of the Tuck Jump Assessment in elite female football players. *Science and Medicine in Football*, 1(2), 139–144. <https://doi.org/10.1080/24733938.2017.1282165>.
- 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 and Conditioning Journal*, 33(3), 21–35. <https://doi.org/10.1519/SSC.0b013e318213afa8>.
- Myer, G. D., Ford, K. R., Di Stasi, S. L., Foss, K. D. B., Micheli, L. J., & Hewett, T. E. (2015). High knee abduction moments are common risk factors for patellofemoral pain (PFP) and anterior cruciate ligament (ACL) injury in girls: Is PFP itself a predictor for subsequent ACL injury? *British Journal of Sports Medicine*, 49(2), 118–122. <https://doi.org/10.1136/bjsports-2013-092536>.
- Myer, G. D., Ford, K. R., & Hewett, T. E. (2008). Tuck jump assessment for reducing anterior cruciate ligament injury risk. *Athletic Therapy Today*, 13(5), 39–44. <https://doi.org/10.1123/att.13.5.39>.
- Myer, G. D., Ford, K. R., Khoury, J., Succop, P., & Hewett, T. E. (2010). Development of a clinic based prediction tool to identify high ACL injury risk female athletes. *Medicine & Science in Sports & Exercise*, 42(5), 168.
- 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. *The American Journal of Sports Medicine*, 37(10), 1996–2002. <https://doi.org/10.1177/0363546509343200>.
- Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research—applications to practice* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Read, P. J., Oliver, J. L., de Ste Croix, M. B. A., Myer, G. D., & Lloyd, R. S. (2016). Reliability of the tuck jump injury risk screening assessment in elite male youth soccer players. *The Journal of Strength & Conditioning Research*, 30(6), 1510–1516. <https://doi.org/10.1519/JSC.0000000000001260>.
- Read, P. J., Oliver, J. L., Myer, G. D., De Ste Croix, M. B. A., Belshaw, A., & Lloyd, R. S. (2018). Altered landing mechanics are shown by male youth soccer players at different stages of maturation. *Physical Therapy in Sport*, 33, 48–53. <https://doi.org/10.1016/j.pts.2018.07.001>.
- Smith, C. A., Chimera, N. J., Lininger, M. R., & Warren, M. (2017). Modified tuck jump assessment: Reliability and training of raters. *Journal of Sports Science and Medicine*, 16(3), 440–442. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/28912664>.
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, 19(1), 231. <https://doi.org/10.1519/15184.1>.