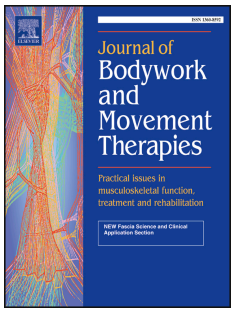
Journal Pre-proof



Intrarater and interrater reliability of the modified Thomas Test

K. Cady, M. Powis, K. Hopgood

PII: S1360-8592(21)00215-1

DOI: https://doi.org/10.1016/j.jbmt.2021.09.014

Reference: YJBMT 2339

To appear in: *Journal of Bodywork & Movement Therapies*

Received Date: 23 January 2021

Revised Date: 16 June 2021

Accepted Date: 19 September 2021

Please cite this article as: Cady, K., Powis, M., Hopgood, K., Intrarater and interrater reliability of the modified Thomas Test, *Journal of Bodywork & Movement Therapies* (2021), doi: https://doi.org/10.1016/j.jbmt.2021.09.014.

Credit Author Statement

Kate Cady : Supervision, writing review and editing, methodology

Molly Powis: Data collection and writing original draft, methodology.

Kirsty Hopgood: Writing, reviewing and editing.

Intrarater and Interrater Reliability of the Modified Thomas Test

Cady, K *1,2, Powis, M 2, Hopgood, K 2

1. University of Hertfordshire, College Lane Campus, Hatfield, Hertfordshire AL10 9AB, k.cady2@herts.ac.uk, +447810440728

2. University of Gloucestershire, Oxstalls Lane, Longlevens, Gloucestershire, GL2 9HW

- Corresponding author

INTRODUCTION

Range of movement (ROM) is an integral part of a clinical assessment to build a clinical picture to enable a diagnosis and/or identify predisposing factors to injury (Roach et al., 2015; van Dillon, McDonnell, Fleming, Sahrmann, 2000). Researchers have suggested that hip-flexor extensibility can influence postural dysfunction causing low back pain, hip pain, knee pain and may affect human performance (Sorensen, Norton, Callaghan, Hwang, & Van Dillen, 2015; Steinberg, Siev-Ner, Peleg, Dar, Masharawi, Zeev, & Hershkovitz, 2012).

The Modified Thomas Test (MTT) is utilized in clinical practice as a method to assess hip-flexor extensibility (Cejudo, de Baranda, Ayala, & Santonja, 2015; Ferber, Kendall, & McElroy, 2010; Harvey, 1998; Peeler & Anderson, 2007). The illiacus and psoas major muscles, otherwise referred to as the illiopsoas group (ILP); the rectus femoris (RF) and the tensor fascia latte (TFL) are three muscles considered as hip flexors and could be indicated when using the MTT (Hattam & Smeatham 2010; Magee 2008).

When using clinical tests such as the MTT, it is important for them to be clinically valuable. Reliability is the extent to which measurements are consistent, dependable, and free from error (Bolgla & Keskula, 1997; Gisev, Bell, & Chen, 2013). A reliable examiner can make consistent repeated measurements to show how a patient is progressing. Intra rater reliability is the ability of the examiner to make consistent measurements over time, and interrater reliability looks at the consistency of multiple raters (Gisev et al., 2013). This is important in both clinical and sporting practice as patients may see multiple practitioners over time or been seen by the same practitioner but over several weeks.

There has been variation recorded for the interrater and intra rater reliability of the MTT (Cejudo et al., 2015; Clapis, Davis, & Davis, 2008; Dennis, Finch, Elliott, & Farhart, 2008; Gabbe, Bennell, Wajswelner, & Finch, 2004; Fraeulin et al, 2020; Kim & Ha, 2015; Peeler & Anderson 2007; Peeler & Anderson 2008; Peeler & Leiter, 2013; Wakefield, Halls, Difilippo, & Cottrell, 2015). This demonstrates that clinicians cannot be certain of their accuracy when using the MTT to identify dysfunction and monitor improvement whether it be measured using a goniometer or graded as a pass/fail due to the conflict in findings. Suggested explanations for this variation include: inaccuracies identifying bony landmarks, inaccuracies using measuring tools such as a goniometer (Peeler & Anderson, 2008; Wakefield et al., 2015), clinician experience (Calpis et al., 2008; Wakefield et al., 2015), variation in the performance of the MTT (Cejudo et al., 2015; Clapis et al., 2008; Fraeulin et al, 2020; Peeler & Anderson 2007), within subject variation, differences in time between testing sessions, identification of a possible order effect (Clapis et al., 2008; Fraeulin et al, 2020), and variation in the specific ROM assessed. Comparisons of the literature are also difficult due to the varying measurement methods leading to different statistical analysis.

To overcome some of these variables Peeler & Leiter (2013) utilized a digital image of the MTT and found it to have a superior interrater and intrarater reliability for the RF contradicting previous findings (Peeler & Anderson 2008). The Kappa values for the intrarater and interrater reliability demonstrated very good strength of agreement by all examiners irrespective of experience. This was true when measured both with a goniometer (0.98 and 0.97 respectively) and as a pass/fail approach (0.92 and 0.96 respectively). Kim & Ha (2015) also used digital images and identified the MTT to have

very good intrarater reliability in a very small sample size for the RF. However, both studies focus on the RF, whereas others focus on the hip extension ROM relating to the ILP (Calpis et al., 2008; Peeler & Anderson 2007; Wakefield et al., 2015).

Further research is needed to determine the reliability of the MTT for all hip flexor musculature associated with the MTT. This is key due patients seeing multiple clinicians, and in the event of Covid-19 may be communicating with their practitioner digitally. Investigation is needed to greater understand the reliability of a used assessment tool. The purpose of the current study is to investigate the intrarater and interrater reliability of the pass/fail scoring method of the MTT, for the RF, ILP and TFL using digital photographs.

METHODS

Participants

A total of 20 healthy, physically active male rugby players (24.15 ± 3.26 years old) were recruited from a semi-professional rugby club and completed all testing protocols. The exclusion criteria included participants under the age of 18 years, and a history of orthopedic problems such as back and/or lower limb musculoskeletal or neuromuscular injury/pain over the previous six months (Vigotsky et al., 2016; Cejudo et al., 2015). All participants had similar training regimes, consisting of two rugby training sessions per week, (one and a half hours per session), one game at the weekend, and additional individual training sessions during the week. Individuals were advised to continue their activities of daily living, but to avoid vigorous exercise 24 hours before testing sessions

and starting new lower limb stretching (Wakefield et al., 2015). Prior to testing, all participants provided written informed consent, and the study was approved by The University of Gloucestershire ethics committee.

Raters

A total of six raters were used for the current investigation. Three of the raters were university students in their final year of studying BSc Sports Therapy. The other three raters included two Graduate Sports Therapists with experience of one and two years, and one Chartered Physiotherapist with thirteen years of experience. All therapists have routinely used the MTT to evaluate the iliopsoas, rectus femoris and TFL flexibility within a clinical and sporting environment (Peeler & Anderson, 2008). All raters completed an informed consent form prior and were allocated a personal identification number.

Procedure

Participants were instructed to wear loose shorts and remove their shoes and socks for testing and were allocated an identification number to allow for confidentiality and anonymity. This was recorded on the participant information form, in addition to their name, age, position, weight, height and dominant limb. Only the principle investigator had access to these documents throughout the investigation.

Each participant attended one testing session which took place in the medical room prior to a rugby training session. This session was conducted in a standardized testing environment (i.e., room temperature, lighting, and plinth type).

Modified Thomas Test

The procedure for the MTT has been previously used by Wakefield et al, 2015). The dominant leg was used for each test (leg that they used to kick a ball with). Before photographs were taken, a practice test was allowed for participant familiarization prior to data recordings (Gabbe et al., 2004). Participants were instructed to sit on the edge of the plinth. The height of the plinth was adjusted whereby their gluteal fold was on the edge of the plinth, whilst both feet were flat on the ground. Afterwards, participants brought their non-dominant knee towards their chest. Whilst grabbing their knee with both hands, participants slowly rolled back onto the plinth with assistance via the examiner, if needed. The dominant limb then extended unsupported off the end of the plinth via gravity, to allow for full extension. Participants were questioned whether they felt any restriction in hip extension due to the edge of the plinth, or if the plinth was not positioned at the correct height. The test was repeated if the participant felt any restriction.

A passive assistance was then applied to the non-testing limb by the principle investigator. This involved applying a pressure onto the limb the participants were holding, to minimize lumbar extension Participants were asked to notify the examiner when they felt the lower part of their back to be flat against the plinth. A photograph was then taken after ten seconds. To identify TFL flexibility, the tripod and camera was moved two meters anteriorly of the participant to view the frontal plane of the hip joint (Figure 2), and the same procedure was repeated. All photographs were taken by the principle investigator, with the same image background and standardize camera position for each testing session

to produce consistent photographs and to minimize possible sources of measurement error (Peeler & Leiter, 2013).

Digital photographs were taken using a camera mounted on a tripod one meter away from the limb that was being tested (Figure 1). The pan, roll and tilt of the camera was adjusted to line with the central part of the test position to avoid parallax error which occurs when objects move away from the optical axis of the camera (Kirtley, 2006). The patients face and chest were not included in the digital photographs.

(to be inserted here)

Figure 1. Side view of the MTT assessing flexibility of iliopsoas and rectus femoris.

(to be inserted here)

Figure 2. Front view of the MTT assessing the flexibility of TFL.

Intrarater and interrater reliability

The twenty sets photographs were randomly uploaded to a secure online system and were shared with each of the six raters. Raters received an email with their allocated identification number. Raters were asked to record their findings as a pass (1) or fail (0) for the TFL, rectus femoris and iliopsoas. Results were returned to principle investigator after each rating. Each rater was asked to delete their recordings once finalized to avoid reviewing their recordings when repeating the same process later in the study.

Raters were given the following guidelines. Iliopsoas flexibility was determined by the angle of hip flexion. The passive length of rectus femoris was determined by the knee flexion angle, and lastly, the hip abduction angle of the femur relative to the angle of the pelvis represented TFL flexibility (Harvey, 1998). The rater determined their pass/fail score according to the protocol by Magee (2002) and Reid (2017). For a pass score, the participant's test leg remained relaxed, whereby the hip was extended either in line with the plinth or below, knee flexion was at least 90° or below, and if no hip external rotation or hip abduction was present. The presence of hip flexion above the line of the plinth was recorded as a fail. Additionally, knee flexion above 90° and any hip external rotation/hip abduction was categorized as a fail (Magee, 2002; Reid, 2017). All three angles were assessed individually, therefore, each participant may have had a mixture of both pass and fail scores.

Each rater had a maximum of 14 days to score the photographs. Raters were blinded; therefore, they were not aware of the other raters results and any previous results during each scoring period. Two weeks after the first set, the photographs were uploaded in a randomized order. Raters were instructed to repeat the same procedure. A total of three sets of randomized photographs were uploaded, all with an interval of 14 days in-between.

Data Analysis

Data was entered in SPSS (version 24, IBM Corporation, NY, USA), where the intrarater reliability of the pass/fail scoring was calculated using a Cronbach's alpha,

which provided a measure of internal consistency of the test within each rater (Tavakol & Dennick, 2011). The inter-rater reliability of the pass/fail scoring was measured using a Fleiss Kappa statistic (McHugh, 2012). The strength of agreement for kappa values are demonstrated in Table 1 (Bland & Altman, 1999). The alpha level was set at 0.05 for statistical tests.

Table 1. Classification of Fleiss Kappa values (Bland & Altman, 1999).

Value of Kappa	Strength of agreement
< 0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very good

RESULTS

For intrarater reliability Cronbach's alpha values for iliopsoas ranged from 0.84-1.00, demonstrating very good level of reliability between each of the six raters. High levels of intrarater reliability were also found for rectus femoris. Whereas the intrarater scores for TFL ranged from good to very good levels of reliability (Table 2).

Fleiss kappa (Fκ) statistics for testing session (TS) 1 of iliopsoas flexibility revealed a significantly high level of reliability between the six raters ($F\kappa = 0.91$, $p = 0.001$) (Table 3). Likewise, the other two TS for iliopsoas demonstrated consistency within the scores and similar values of very good interrater reliability (TS 2: $F\kappa = 0.71$, $p = 0.001$; TS 3:

Fκ = 0.71, p = 0.001). As for the interrater scores for rectus femoris, a Fleiss kappa could not be calculated due to extremely consistent scores amongst the 6 raters for each testing session. Although, a fair level of reliability was found for TS 3 (Fκ = 0.39, p = 0.001). Lastly, TFL presented with the most inconsistent scores as values ranged from fair to good levels of reliability between raters (TS 1: Fκ = 0.58, p = 0.001; TS 2: Fκ = 0.38, p = 0.001; TS 3: Fκ = 0.71, p = 0.001).

Table 2. Intrarater Cronbach's alpha statistics for pass/fail scoring of each muscle during the modified Thomas test

	Rater A	Rater B	Rater C	Rater D	Rater E	Rater F	Mean (n = 6)
Iliopsoas	0.96	0.92	0.95	1.00	1.00	0.84	0.95
RF**	1.00*	1.00*	1.00*	1.00*	1.00	1.00*	1.00
TFL**	0.95	0.64	0.93	0.88	0.95	0.78	0.86

**1.00- Cronbach's alpha could not be calculated on SPSS due to each rater having 'near to perfect' intrarater reliability scores, therefore, the score for each scale item could not be divided by the total score for each observation.*

***RF=Rectus Femoris, TFL= Tensor Fascia Latae.*

Table 3. Interrater Fleiss kappa statistics of all 6 raters for pass/fail scoring of each muscle during the modified Thomas test (95% Confidence Intervals)

	Testing session 1	Testing session 2	Testing session 3	Mean (n = 3)
Iliopsoas	0.91 (0.80-1.03)	0.71 (0.60-0.83)	0.71 (0.59-0.82)	0.78
RF**	1.00*	1.00*	0.39 (0.27-0.50)	0.80

TFL**	0.58 (0.46-0.69)	0.38 (0.27-0.50)	0.71 (0.58-0.82)	0.56

**1.00- Fleiss kappa statistic could not be correctly calculated within SPSS due to there being extremely minor variations between the raters scores. Therefore, the between examiners reliability is close to 100%.*

***RF=Rectus Femoris, TFL= Tensor Fascia Latae.*

DISCUSSION

This study was conducted to examine the reliability of the MTT to assess iliopsoas, TFL and rectus femoris tightness about the hip and knee joint, using digital photographs. High intrarater reliability was found when scoring both iliopsoas (Chronbacks alpha 0.84-1.0) and rectus femoris flexibility (Chronbacks alpha 1.0). This supports results from Peeler & Leiter, (2013) who found high intrarater reliability values for rectus femoris flexibility. Although this study only assessed rectus femoris flexibility it is the only study in the current literature that uses digital photographs to assess joint range. It is however difficult to compare studies as they used a slightly version of the MTT on their assessment. In argument, Peeler and Anderson (2007) and Peeler and Anderson (2008) both carried out goniometer and pass/fail assessment of the iliopsoas and rectus femoris and found moderate reliability which does not support the findings from this study. However, the study by Peeler and Anderson (2007) used the TT and not the MTT and they used real time assessment rather than digital photographs and so it is difficult to make comparisons.

Although intrarater reliability for rectus femoris and iliopsoas flexibility have been studied previously, there is currently nothing in the literature which looks at reliability of the MTT for TFL flexibility. The results from this study show inconsistency of the values

given within (Cα = 0.64-0.95) raters when scoring the flexibility of TFL with a mean reliability value of (0.86). As this is the first study to look at the TFL, there is a need for further research to be carried out in this area to ensure high reliability for a clinician.

Interrater reliability values varied considerably between raters for iliopsoas (Fκ= 0.71-0.91), rectus femoris (Fκ=0.39-1.0) and TFL flexibility (Fκ = 0.38-0.71) which does differ from previous studies (Peeler and Anderson, 2007, Peeler and Leiter, 2013) which do not show this range of variability. Although it is difficult to make comparisons due to methodological differences more research is required if this test is to be used by multiple clinicians on the same individual. The results show a large variation in inter rater reliability in the third TS for rectus femoris (Fκ = 0.39) in comparison to TS 2 and 3 (Fκ = 1.0). Raters were only briefed on the scoring sessions prior to the first session. As rating of the photographs took place over 6 weeks, a variation could attributed to rater consistency in applying the pass/fail scoring criteria.

The raters of the current study stated that they found it difficult to score the TFL muscle photo as either a pass or fail in relation to the scoring system outlined by Magee (2002) and Reid (2017). This may explain why there was more variance in the reliability of TFL both within (Cα = 0.64-0.95) and between (Fκ = 0.38-0.71) raters. If a kappa result is lower than expected, it may be a result of clinicians needing more training in the scoring technique, or the protocols need to be reworded (Sim & Wright, 2005). Scoring the TFL depended on the alignment of the patella in relation to the ASIS, however, the ASIS could not be seen on the photographs. Therefore, instead of rater error, the study proved to have

limitations of the camera set up and the instructions. The pass/fail method needed to be reworded, or an aerial perspective of the participant would have been more beneficial.

Much of the literature surrounding the MTT does not control for lumbo-pelvic movement, which produces an invalid measure of peak hip extension angle or to identify the presences of hip flexion contracture, which may explain why unreliable results were found (Vigotsky et al., 2016). The weight of the limb itself forces the leg lower to the floor if the contralateral hip is not maximally flexed (Vigotsky et al., 2016). Additionally, if the participants possess a shortened iliopsoas, an increased lumbar lordotic curve occurs (Jorgensson, 1993), further dropping the test thigh below the level of the plinth. This would be categorized as a 'pass' for iliopsoas flexibility, which in fact is a false negative. Therefore, in the current study, the examiner applied an external passive force to maximally flex the non-dominant side hip on each participant. However, the literature suggested that the active lumbo-pelvic stabilization (ALS) method demonstrated the highest reliability when controlling for lumbo-pelvic motion. This is because the ALS method involves co-contraction of the local and global muscles and an increased amount of abdominal muscles activity which provides more stability, minimizing lumbo-pelvic motion (Park et al., 2013; Noh, Kim, Kim, Ha & Oh, 2014). However, after piloting the bio-pressure feedback on a sample of rugby players, this method proved to be difficult and a passive assistance on the contralateral limb was applied to control for lumbo-pelvic motion during the MTT. Therefore, these findings suggest it is important use equipment

or an external force to maximally flex the hip to allow for correction of methodological errors, which can then further improve standardization and the reliability.

Previous studies have used 2-3 testing sessions, with repeated procedures each time (Cejudo et al., 2015; Peeler & Anderson, 2007; Peeler & Anderson, 2008; Wakefield et al., 2015). Therefore, day-to-day variation in the positioning of the participants was not controlled for, which may explain the poor reliability values seen (Wakefield et al., 2015; Atkinson & Nevill, 1998). Additionally, order effects occur during within-participant experiments which may influence the results (McBurney & White, 2009). These order effects include practice, fatigue, boredom, carryover, and sensitization (Mitchell & Jolley, 2012). The current study avoided order effects through only conducting one testing session, whereby the same set of photographs were sent out on 3 different occasions, in a randomized order and could be attributed to the high reliability values seen in this study.

Support for the current findings is derived from a similar study who found the pass/fail method to be highly reliable both within ($\kappa = 0.92$) and between ($\kappa = 0.96$) raters using digital photographs (Peeler & Leiter, 2013). These results suggest that the application of a standardized methodological approach when scoring the MTT from digital photographs is beneficial (Peeler & Leiter, 2013). Digital photographs also allow clinicians to test clients remotely which is extremely beneficial during the covid 19 pandemic.

The different levels of experience the raters possess may explain the variations found within previous studies (Wakefield et al., 2015). However, Peeler and Leiter (2013) indicated that the reliability of digital photographs for the in-experienced examiner group

was equal with that of the experienced group. Photographs minimize the number of variables that an inexperienced therapist needs control in comparison to 'hands on' assessment, therefore, influencing the precision and consistency of their measurements. Therefore, this demonstrates that using digital photographs can help improve the reliability and can be applied to many different special tests and measures used within a clinical setting. Additionally, it allows for more confident results to be given, particularly when one is an inexperienced therapist. Although the current study used 6 raters, there was not enough variation in the level of experience to support the findings from Peeler and Leiter (2013).

Whilst the results of this study have supported the experimental hypothesis and provided information regarding the reliability of this common special test, limitations of the study must be recognized. The age distribution was relatively narrow (21-30 years), along with a very small sample of 20 participants as compared to previous studies who have used over 50 participants (Harvey, 1998; Peeler & Anderson, 2007; Peeler & Anderson, 2008; Peeler & Leiter, 2013; Cejudo et al., 2015). Therefore, the generalizability to the broader population cannot be ascertained, and reliability results may have differed with a larger sample size. The current study only used a healthy male rugby player population for the screening process and therefore it is not known whether the reliability values here could be generalized to the female population or those who participate in varying sporting activities. However, to support the findings, Peeler and Leiter (2013) proved that the pass/fail test was a reliable method amongst those in a wide variety of leisure and sporting opportunities.

Although the current study found reliable results, this study only looked at visual scoring and therefore other measurements were not explored. However, other scoring methods were not used as the purpose of this study was to replicate 'real life' settings whereby time is limited, equipment is not always available, and easier versions need to be performed so that replication can occur on many different occasions and populations.

CONCLUSION

This study provides important information to clinicians and researchers about the reliability of scoring iliopsoas, rectus femoris and TFL flexibility from digital photographs using the MTT and dichotomous data. The low variation both within and between raters for iliopsoas and rectus femoris flexibility supports the use of the pass/fail method and provides evidence that applying the standardized principles to assess muscle flexibility is reliable. The higher variation demonstrated, both within and between raters for TFL flexibility, suggests that the protocol has limitations which may be improved if the correct camera angle is selected for scoring TFL, such as an aerial perspective. Nevertheless, the findings indicated that obtaining digital photographs at the time of the assessment limits measurement error, enhances scoring accuracy, and enables the clinician to establish whether an observed change between baseline and future assessments is in fact real (Peeler & Leiter, 2013). This was the first reliability study to use a valid measure of the MTT, which involved correcting one's pelvic tilt for the iliopsoas, rectus femoris and TFL. Therefore, there is a need for further research to be carried out in this area if we are to continue using this popular flexibility test.

CLINICAL RELEVANCE

- The results of this study show that the use of digital photographs for the pass/fail method of the MTT Is reliable. Therefore, assessment of iliopsoas and rectus femoris can be used via digital imaging which is useful in online consultations.
- There is a need for further study to assess TFL via digital photographs and whether a different camera angle would be beneficial.

Conflict of interest

The authors report not conflict of interest.

Ethical approval

This study has been approved by the appropriate ethical committees related to the institution(s) in which it was performed and that subjects gave informed consent to the work.

Funding

This research did not receive any specific grant from funding agencies in the public, Commercial, or not-for-profit sectors

REFERENCES

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine, 26(4), 217-238.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. Statistical Methods in Medical Research, 8(2), 135-160.

Bolgla, L. A., & Keskula, D. R. (1997). Reliability of lower extremity functional performance tests. Journal of Orthopaedic & Sports Physical Therapy, 26(3), 138-142.

Cejudo, A., de Baranda, P. S., Ayala, F., & Santonja, F. (2015). Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players. Physical Therapy in Sport, 16(2), 107-113.

Clapis, P. A., Davis, S. M., & Davis, R. O. (2008). Reliability of inclinometer and goniometric measurements of hip extension flexibility using the modified Thomas test. Physiotherapy theory and practice, 24(2), 135-141.

Dennis, R. J., Finch, C. F., Elliott, B. C., & Farhart, P. J. (2008). The reliability of musculoskeletal screening tests used in cricket. Physical Therapy in Sport, 9(1), 25-33.

Ferber, R., Kendall, K. D., & McElroy, L. (2010). Normative and critical criteria for iliotibial band and iliopsoas muscle flexibility. Journal of Athletic Training, 45(4), 344-348.

Fraeulin L et al. (2020) Intra- and inter-rater reliability of joint range of motion tests using tape measure, digital inclinometer and inertial motion capturing. PLoS ONE 15(12): e0243646.

Gabbe, B. J., Bennell, K. L., Wajswelner, H., & Finch, C. F. (2004). Reliability of common lower extremity musculoskeletal screening tests. Physical Therapy in Sport, 5(2), 90-97.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. Research in Social and Administrative Pharmacy, 9(3), 330-338.

Harvey, D. (1998). Assessment of the flexibility of elite athletes using the modified Thomas test. British Journal of Sports Medicine, 32(1), 68-70.

Hattam, P., & Smeatham, A. (2010). Special tests in musculoskeletal examination. Churchill Livingstone: Elsevier.

Jorgensson, A. (1993). The iliopsoas muscle and the lumbar spine. Australian Journal of Physiotherapy, 39(2), 125-132.

Kim, G. M., & Ha, S. M. (2015). Reliability of the modified Thomas test using a lumbo-plevic stabilization. Journal of physical therapy science, 27(2), 447-449.

Kirtley, C. (2006). Clinical gait analysis: theory and practice. Washington, USA: Elsevier Health Sciences.

Magee, D. J. (2002). Orthopaedic physical assessment. St. Louis, Mo: Saunders Elsevier.

McBurney, D. H., & White, T. L. (2009). Research methods (8th ed.). USA: Cengage Learning.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica: Biochemia medica, 22(3), 276-282.

Mitchell, M. L., & Jolley, J. M. (2012). Research design explained (8th ed). USA: Cengage Learning.

Noh, K. H., Kim, J. W., Kim, G. M., Ha, S. M., & Oh, J. S. (2014). The influence of dual pressure biofeedback units on pelvic rotation and abdominal muscle activity during the active straight leg raise in women with chronic lower back pain. Journal of Physical Therapy Science, 26(5), 717-719.

Park, K. H., Ha, S. M., Kim, S. J., Park, K. N., Kwon, O. Y., & Oh, J. S. (2013). Effects of the pelvic rotatory control method on abdominal muscle activity and the pelvic rotation during active straight leg raising. Manual Therapy, 18(3), 220-224.

Peeler, J. D., & Anderson, J. E. (2008). Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. Journal of Athletic Training, 43(5), 470-476.

Peeler, J., & Anderson, J. E. (2007). Reliability of the Thomas test for assessing range of motion about the hip. Physical Therapy in Sport, 8(1), 14-21.

Peeler, J. D., & Anderson, J. E. (2008). Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. Journal of athletic training, 43(5), 470-476.

Peeler, J., & Leiter, J. (2013). Using digital photography to document rectus femoris flexibility: a reliability study of the modified Thomas test. Physiotherapy Theory and Practice, 29(4), 319-327.

Reid, A. (2017). Science of Stretching. Wiltshire: The Crowood Press.

Roach, S. M., San Juan, J. G., Suprak, D. N., Lyda, M., Bies, A. J., & Boydston, C. R. (2015). Passive hip range of motion is reduced in active subjects with chronic low back pain compared to controls. International journal of sports physical therapy, 10(1), 13.

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

Sorensen, C. J., Norton, B. J., Callaghan, J. P., Hwang, C. T., & Van Dillen, L. R. (2015). Is lumbar lordosis related to low back pain development during prolonged standing? Manual therapy, 20(4), 553-557.

Steinberg, N., Siev-Ner, I., Peleg, S., Dar, G., Masharawi, Y., Zeev, A., & Hershkovitz, I. (2012). Joint range of motion and patellofemoral pain in dancers. International journal of sports medicine, 33(07), 561-566.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53-55.

Van Dillen, L. R., McDonnell, M. K., Fleming, D. A., & Sahrmann, S. A. (2000). Effect of knee and hip position on hip extension range of motion in individuals with and without low back pain. Journal of Orthopaedic & Sports Physical Therapy, 30(6), 307-316.

Vigotsky, A. D., Lehman, G. J., Beardsley, C., Contreras, B., Chung, B., & Feser, E. H. (2016). The modified Thomas test is not a valid measure of hip extension unless pelvic tilt is controlled. PeerJ, 4, 1-12.

Vigotsky, A. D., Lehman, G. J., Contreras, B., Beardsley, C., Chung, B., & Feser, E. H. (2015). Acute effects of anterior thigh foam rolling on hip angle, knee angle, and rectus femoris length in the modified Thomas test. PeerJ, 3, 1-13.

Wakefield, C. B., Halls, A., Difilippo, N., & Cottrell, G. T. (2015). Reliability of goniometric and trigonometric techniques for measuring hip-extension range of motion using the modified Thomas test. Journal of Athletic Training, 50(5), 460-466.

Conflict of interest

None declared

Ethical Approval

This study has been approved by the appropriate ethical committees related to the institution(s) in which it was performed and that subjects gave informed consent to the work.

Funding

This research did not receive any specific grant from funding agencies in the public, Commercial, or not-for-profit sectors

Acknowledgements

The authors would like to thank the participants for agreeing to be part of the study.