Journal of Athletic Training 2008;43(5):470–476

www.nata.org/jat

Reliability Limits of the Modified Thomas Test for Assessing Rectus Femoris Muscle Flexibility About the Knee Joint

Jason D. Peeler, PhD, CAT(C)*; Judy E. Anderson, PhD†

*University of Winnipeg, Winnipeg, MB, Canada; †University of Manitoba, Winnipeg, MB, Canada

Context: The modified Thomas test is commonly used in the clinical setting to assess flexibility about the thigh region.

Objective: To evaluate the clinical reliability of the modified Thomas test for evaluating the flexibility of the rectus femoris muscle about the knee joint.

Design: Descriptive laboratory study using a test-retest design.

Setting: Institution-based clinical orthopaedic setting.

Patients or Other Participants: Fifty-seven individuals between the ages of 18 and 45 years with no history of trauma participated. Of those, 54 completed the study.

Intervention(s): Three Board-certified athletic therapists with an average of 12.67 years of sport medicine expertise assessed rectus femoris flexibility using pass/fail and goniometer scoring systems. A retest session was completed 7 to 10 days later.

Main Outcome Measure(s): Parametric and nonparametric tests were used to compare participants' test-retest results.

Results: Chance-corrected κ values (intrarater $\bar{x} = 0.40$, 95% confidence interval [CI] = 0.30, 0.54; interrater $\bar{x} = 0.33$, 95% CI = 0.23, 0.41) indicated generally poor levels of reliability for pass/fail scoring. Intraclass correlation coefficient (ICC) values (intrarater $\bar{x} = 0.67$, 95% CI = 0.55, 0.76; interrater $\bar{x} = 0.50$, 95% CI = 0.40, 0.60) indicated fair to moderate levels of reliability for goniometer data. Measurement error values (standard error of measurement = 7°, method error = 6°, and coefficient of variation = 13%) and Bland-Altman plots (with 95% limits of agreement) further demonstrated the degree of intrarater variance for each examiner when conducting the test.

Conclusions: These results call into question the statistical reliability of the modified Thomas test and provide clinicians with important information regarding its reliability limits when used to clinically assess flexibility of the rectus femoris muscle about the knee joint in a physically active population. More research is needed to ascertain the variables that may confound the statistical reliability of this orthopaedic technique.

Key Words: Kendall test, contraction test, quadriceps flexibility

Key Points

- The modified Thomas test demonstrated moderate reliability among examiners during goniometer scoring and poor reliability during pass/fail scoring.
- Measurement error during goniometer evaluation may have resulted in an overemphasis of the difference in test-retest scores for each examiner.
- The modified Thomas test may not demonstrate a high level or reliability even when experienced examiners with similar education and training and with advanced orthopaedic assessment skills use it to assess flexibility of the rectus femoris muscle about the knee joint.
- The clinical reliability of the modified Thomas test may be influenced by many factors that include but are not limited to variations in the application of assessment criteria, the scoring method used, the consistency and accuracy of surface landmarking, the patient population, and examiner experience.

Sports medicine practitioners routinely use the modified Thomas test, which is also called the rectus femoris contraction test or Kendall test, to assess flexibility of the rectus femoris muscle about the knee joint. The face validity of this assessment technique is confirmed by its inclusion in many prominent sports medicine textbooks[1–5] and its use as a measurement tool in applied orthopaedic research examining rectus femoris muscle flexibility about the knee joint.[6–16] Unfortunately, only a few authors have reported rectus femoris flexibility data about the knee that were collected using modified Thomas evaluation criteria with established reliability limits, and much of this research has included a specific population (ie, 1 type of athlete, sport, or disease).[7,11] As a result, the reliability limits of the modified Thomas test need to be determined before it is used to establish normative values for rectus femoris flexibility.

The purpose of our investigation was to test the hypothesis that the modified Thomas test provides reliable assessment of rectus femoris muscle flexibility about the knee joint. Specifically, we had the following aims: (1) to investigate the intrarater reliability of the modified Thomas test, (2) to investigate the interrater reliability of the modified Thomas test, and (3) to compare the reliability of goniometer (continuous data) and pass/fail (dichotomous data) scoring for the modified Thomas test.

METHODS

Participants

After the Research Ethics Board at the University of Manitoba approved this study, we recruited 57 participants between ages 18 and 45 years with no history of surgery or

Table 1. Participants' Anthropometric Data (n = 57)

	Mean ± SD	95% Confidence Interval
Age, y	29 ± 7.3	14.7, 43.3
Height, m	1.68 ± 0.09	1.5, 1.9
Mass, kg	69 ± 11.0	47.4, 93.6
Body mass index	24.3 ± 3.0	18.4, 30.2
Physical activity levels[a]	8.5 ± 1.3	6.0, 11.1

[a] Baecke Questionnaire of Habitual Physical Activity scored out of a maximum of 15.

trauma to the hip, knee, or lower leg region. All participants took part in an initial intake session before beginning the study. During this session, they were assigned an identification number and provided informed consent, and they completed participant information forms. They also completed a physical activity questionnaire to provide information regarding their habitual activity patterns at work, leisure, and play.[17] We recorded anthropometric data, such as height and mass (Table 1). At the conclusion of the intake session, participants were scheduled for 2 separate assessment sessions that were conducted either over the lunch-hour period or during the early evening and were 7 to 10 days apart.[18]

Three experienced examiners were recruited from the community to participate in the study. All were Board-certified athletic therapists who possessed, on average, more than 12 years (range, 6–22 years) of experience clinically assessing and treating musculoskeletal disorders. Additionally, each therapist routinely used the modified Thomas test to evaluate rectus femoris muscle flexibility in his or her patients. Before data collection, each examiner attended 2 instructional workshops to become familiar with the testing protocol and to reinforce the criteria defining pass/fail scoring on the modified Thomas test and the standardized procedures for collecting goniometric data.

Procedures

The study was completed over a 6-month period and was conducted according to detailed published protocols.[19,20] Participants were instructed to wear shorts and T-shirts for all assessments. They were also instructed to refrain from exercise for a minimum of 4 hours before each testing session and to avoid starting new sporting or training activities during their 7 to 10 days of involvement in the study. Each examiner completed the assessment of a participant in approximately 5 minutes, and a maximum of 4 participants were tested each half hour. Participants underwent independent assessment by each of the 3 examiners. Initial testing order was randomized, and the same order was used in the retest session.

Examiners used the modified Thomas test to assess bilateral rectus femoris muscle flexibility in each participant (Figure 1). The participant was placed in a supine position with the knees bent over the edge of the examination table. He or she was instructed to flex one knee to the chest and hold it. At the same time, the angle of the test knee (opposite of the knee held to the chest) was to remain at 90°, and the hip and posterior thigh of the test leg were to remain in a stationary position and flat against the examination table. Examiners determined pass/fail scoring according to the protocol that Magee[1] outlined.

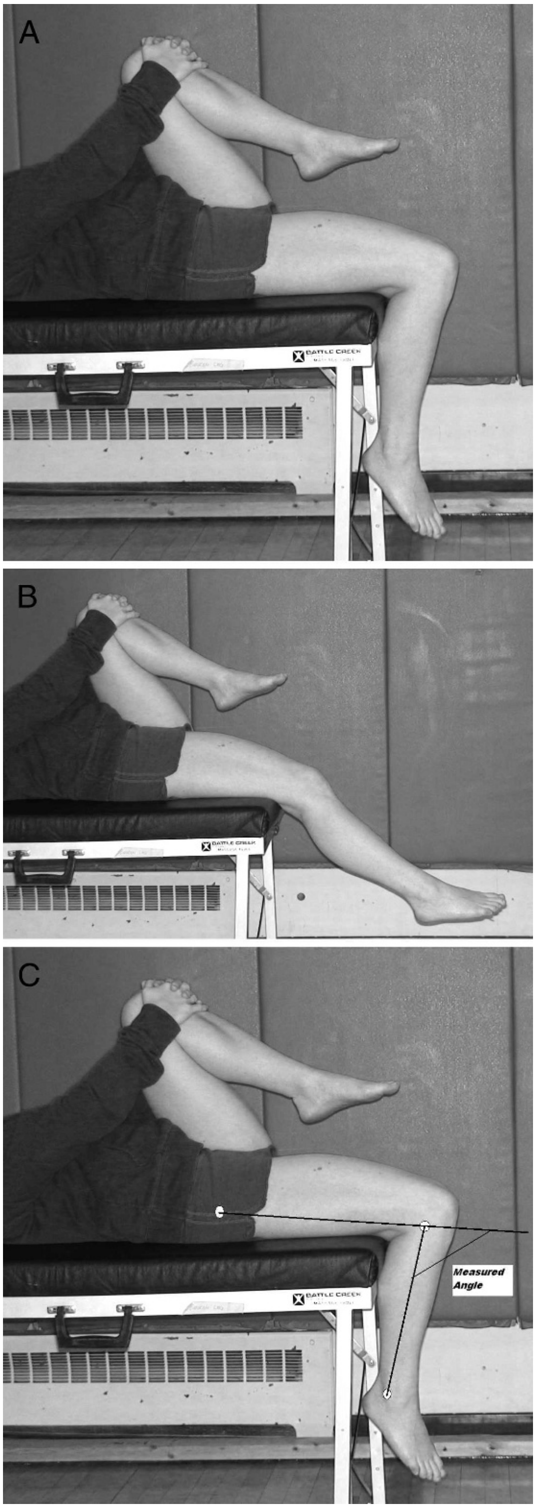

Figure 1. Visual representation of pass/fail grading with the modified Thomas test. A, For a pass score, the participant's test leg remains relaxed with the knee in a position of at least 90° of flexion. B, For a fail score, the participant's test-leg demonstrates a position of less than 90° of knee flexion. C, Visual representation of goniometer scoring. To measure test-leg knee flexion angle, we placed an adhesive marker over the head of the fibula and centered an 18-in (45.72-cm), semirigid plastic goniometer over it.

The assessment was scored as a *pass* if the test knee remained in a stationary 90° position. The assessment was scored as a *fail* if the test knee extended and moved to a position of less than 90°.[1] The same knee was assessed at each session.

Joint range of motion (ROM) was quantified using a flexible 18-in (45.72-cm) adjustable plastic goniometer (Baseline; Fabrication Enterprises Inc, White Plains, NY) that health care practitioners commonly employ in clinical settings.[21] Measurements were taken according to visibly identifiable anatomic landmarks and avoided procedures that would require examiners to estimate the exact center of rotation about which the knee joint moves. Pilot testing with all 3 examiners highlighted several difficulties associated with goniometer measurements about the knee joint during the execution of the modified Thomas test. These problems included but were not limited to (1) inconsistencies among examiners when establishing the exact position of either the medial or lateral joint line of the knee, (2) inconsistencies among examiners when establishing the exact position of either the medial or lateral epicondyle of the femur, (3) inconsistencies among examiners when establishing the center of motion for the knee joint about the sagittal plane, and (4) difficulties in aligning the axis of rotation of the goniometer about the center of motion for the knee joint, especially when trying to maintain the 2 lever arms of the goniometer in line with superior and inferior landmarks. In response to examiner feedback and to minimize the confounding effect that inaccurate establishment of landmarks could have on goniometer measurements,[22,23] we decided to align the axis of rotation of the goniometer over the head of the fibula. Examiners were confident that this anatomic prominence would facilitate the most efficient, accurate, and reliable establishment of surface landmarks and help minimize the confounding effect that inconsistent establishment of surface landmarks could have on goniometer measurements performed by multiple examiners on multiple participants over several testing sessions. Goniometer measurement procedures were standardized; each examiner placed an adhesive marker (1.5 cm in diameter) over the head of the fibula and established the superior (greater trochanter of the femur) and inferior (lateral malleolus of the fibula) landmarks before each measurement. The degree of knee flexion in the test leg was measured about an axis of rotation running through the head of the fibula, and the 2 arms of the goniometer were aligned with the superior and inferior landmarks.

Each examiner recorded modified Thomas test scores (goniometer measurement to the nearest degree and a pass/fail score) for each participant on a standardized data collection sheet for each test session. At both sessions, each examiner was blinded to the scoring by the other examiners, and, at the second session, examiners were blinded to their scoring from the first test session. At the end of each test session, data sheets were collected and collated according to participant identification numbers.

Data Analysis

Data were entered on an Excel spreadsheet (Microsoft Corp, Redmond, WA). Organized by sex, descriptive statistics (mean ± SD) were generated for age, height, and body mass measurements; calculated body mass index (BMI); physical activity levels (scored out of a total of 15 on the Baecke Questionnaire of Habitual Physical Activity); and knee joint angle. Two-way analysis of variance (ANOVA) (sex [man, woman] × rater [1, 2, 3]) was used to compare scoring between the sexes. Intraclass correlation coefficients (ICCs) were calculated to evaluate the intrarater (ICC [3,1]) and interrater (ICC [2,1]) reliability of goniometer scoring.[20,24] Intrarater and interrater reliability of pass/fail scoring was measured using a κ statistic. This statistic uses a simple index of agreement, which is called percent agreement, to measure how often raters agree on scoring for each participant. The advantage of the κ statistic is that it examines the proportion of observed agreement and also considers the proportion of agreement that might be expected by chance. Therefore, the coefficient of agreement (proportion of observations on which there is agreement divided by the number of pairs of scores that were obtained) that the κ test produces is corrected for chance (number of expected agreements divided by number of possible agreements). This calculation provides a reasonable estimate of the reliability of dichotomous pass/fail data.[24,25] As cited by several researchers, ICC and κ values that are more than 0.75 represent high levels of reliability, values that are between 0.4 and 0.75 indicate a fair to moderate level of reliability, and a value that is less than 0.4 represents a poor level of reliability.[24,26–28] We used the standard error of the measurement (SEM), method error (ME), and coefficient of variation (CV) to examine the within-subjects variation between each examiner's testing sessions. The SEM was defined as $SEM = SD1(1\text{-}ICC)^{0.5}$, where SD1 is the SD of all measurements and the ICC value is derived from intrarater analysis. The ME was defined as $ME = SD2/\sqrt{2}$, where SD2 is the SD of the differences between the 2 measurements. The CV was defined as $CV = 100\ ME/X1$, where X1 is the mean for all observations from test sessions 1 and 2.[28] Finally, Bland-Altman graphs with accompanying 95% limits of agreement provided a visual representation of the range of scoring by each of the examiners over the 2 testing sessions and depicted the difference score between test and retest plotted against the mean scores for each participant. The 95% upper and lower limits of agreement represented 2 SDs more than and less than the mean difference score.[29]

RESULTS

The 57 participants represented a young, healthy, and physically active population engaging a wide variety of leisure and sporting opportunities. Of these participants, 54 completed both testing sessions. For analysis, the flexibility measurements from 108 limbs were used to investigate intrarater reliability, and 222 flexibility measurements were used to examine interrater reliability.

Descriptive statistics for knee joint angle are presented in Table 2. The mean knee joint angle for all participants was 50° ± 12° (95% CI = 26, 74), with scores occurring over a large range in men (range, 20°–95°) and women (range, 14°–85°). On average, women demonstrated a larger angle about the knee joint compared with men during modified Thomas testing.

As noted, ICCs and a chance-corrected κ statistic were used to evaluate the relative reliability of intrarater and

Table 2. Goniometer Scoring of Knee Joint Angle (Mean ± SD [95% Confidence Interval])

	Men (34 Limbs)		Women (74 Limbs)		All Participants (108 Limbs)	
	Test (°)	Retest (°)	Test (°)	Retest (°)	Test (°)	Retest (°)
Examiner 1	43 ± 10 (23, 63)	44 ± 14 (17, 71)	47 ± 9 (29, 65)	47 ± 9 (29, 65)	46 ± 9 (28, 64)	46 ± 11 (24, 68)
Examiner 2	54 ± 12 (30, 78)	54 ± 11 (32, 76)	55 ± 11 (33, 77)	59 ± 13 (34, 84)	55 ± 11 (33, 77)	58 ± 13 (33, 83)
Examiner 3	44 ± 9 (26, 62)	45 ± 10 (25, 65)	48 ± 12 (24, 72)	48 ± 10 (28, 69)	47 ± 11 (25, 69)	47 ± 10 (27, 67)
Group average[a]	47 ± 12 (23, 71)		51 ± 12 (27, 75)		50 ± 12 (26, 74)	

[a] The group average represents the mean score of all examiners across both assessments (test and retest).

interrater scoring for the modified Thomas test. Intrarater results are presented in Table 3. Pass/fail chance-corrected κ values ranged from 0.28 (95% CI = 0.06, 0.49) to 0.54 (95% CI = 0.37, 0.74) among the 3 examiners and goniometer ICC values ranged from 0.65 (95% CI = 0.53, 0.76) to 0.72 (95% CI = 0.62, 0.80). Intrarater results demonstrated that, on average, the goniometer method of scoring was more consistent than the pass/fail method of scoring. They also revealed that, independent of scoring method, examiner 1 was generally the most reliable in scoring rectus femoris muscle flexibility and knee joint angle during the test-retest protocol.

Interrater results are presented in Table 4. The goniometer ICC values were, on average, higher than the pass/fail chance-corrected κ values. However, analysis of between-examiners scores using a 2-way ANOVA revealed variation ($P < .01$) in goniometer scoring among the 3 examiners.

The measurement error, which was analyzed using the goniometer data from the 3 examiners, is presented in Table 5. The SEM can be used to determine the minimum difference (MD) that is required to assume that a "real" difference exists between test and retest scores. The MD can be calculated using the formula, SEM $\times 1.96 \times \sqrt{2}$, and enables any change in a participant's score to be considered real if it is either more than or less than the previous score by more than the MD value. The mean MD for the modified Thomas test among examiners was $7 \times 1.96 \times \sqrt{2} = 19°$.

Bland-Altman plots with accompanying 95% limits of agreement (Figure 2) illustrate the wide range in each examiner's scoring over the 2 test sessions and indicate that the difference between test and retest scoring can vary by -1.2 ± 18 (average of 3 examiners) or between $-19.2°$ and $16.8°$. Figure 2 also depicts that examiners were unbiased in scoring over the 2 testing sessions (ie, a higher or lower score was just as likely to occur in test session 1 as in test session 2). In addition, Figure 3 visually represents the large variation ($P < .01$) among the examiners' scoring for participants. It depicts a large number of outlying data points for each examiner and is indicative of systematic examiner-dependent use of the modified Thomas test.

Table 3. Intrarater Chance-Corrected κ Statistics for Pass/Fail Scoring and Intraclass Correlation Coefficient (3,1) Values for Goniometer Scoring (95% Confidence Interval) of the Modified Thomas Test in 108 Limbs

Intrarater Scoring	Pass/Fail	Goniometer
Examiner 1	.54 (0.37, 0.74)	.65 (0.53, 0.76)
Examiner 2	.39 (0.14, 0.62)	.65 (0.53, 0.76)
Examiner 3	.28 (0.06, 0.49)	.72 (0.62, 0.80)
Mean	.40 (0.30, 0.54)	.67 (0.55, 0.76)

DISCUSSION

We conducted this study to examine the reliability of an orthopaedic assessment technique commonly used in a sports medicine setting to assess flexibility of the rectus femoris muscle. To our knowledge, no one has reported the reliability of this special test in a large, normative population sample. The results call into question the clinical reliability of the modified Thomas test during both pass/fail and goniometer scoring. Because reliability is an essential component for the use of clinical assessment techniques in sports medicine, our findings have important implications in both clinical and research settings.

Our results showed that the modified Thomas test demonstrated, on average, only moderate levels of reliability among examiners during goniometer scoring and poor levels of reliability during pass/fail scoring. Although the goniometer (continuous) data were comparable to the values (ICC = 0.69) that Gabbe et al[7] reported in a small sample (n = 15) of health care professionals, they were substantially lower than the values (ICC = 0.91–0.94) that Harvey[11] reported in a large sample (n = 117) of elite athletes. However, our interrater ICC values were substantially lower than those values (ICC = 0.90) that Gabbe et al[7] reported. Differences in the modified Thomas testing procedures (ie, patient positioning, establishing landmarks) and scoring methods (pass/fail criteria, specific placement of the goniometer about the knee joint), as well as sample size and methods of statistical evaluation, may explain the variation in the results between investigations.[24] Despite our examiner workshops that were designed to standardize the assessment protocol and define pass/fail criteria, it appears that each of the examiners used slightly different stringency (ie, specified joint angle) when grading flexibility of the rectus femoris muscle as either pass or fail.

Because we evaluated joint angles to within 1° with the goniometer, small differences in measurements related to inaccurate or inconsistent establishment of surface landmarks during goniometric evaluation also may have resulted in an overemphasis of the variation among examiners' scores. Measurement error values for goniometer data compared favorably with the values (SEM = 5°) that Gabbe et al[7] reported and indicated that considerable

Table 4. Interrater Chance-Corrected κ Statistics for Pass/Fail Scoring and Intraclass Correlation Coefficient (2,1) Values for Goniometer Scoring (95% Confidence Interval) During the Modified Thomas Test of 222 Flexibility Measurements

Interrater Scoring	Examiner 1	Examiner 2	Mean
Pass/Fail			
Examiner 2	.41 (0.26, 0.56)	—	.33 (0.23, 0.41)
Examiner 3	.26 (0.10, 0.40)	.32 (0.14, 0.45)	
Goniometer			
Examiner 2	.44[a] (0.33, 0.54)	—	.50 (0.40, 0.60)
Examiner 3	.59 (0.50, 0.67)	.48[a] (0.38, 0.58)	

[a] $P < .01$.

variation was present in each examiner's scoring over the 2 testing sessions. In fact, 95% limits of agreement illustrated that scores could range by as much as 36° for each examiner's scoring over the 2 testing sessions. The SEM values also depicted that the MD required to assume that a real difference exists between an examiner's test-retest scores for the modified Thomas test was approximately 19°. This information provides valuable insight into the clinical reliability limits of the modified Thomas test and enables practitioners to make knowledgeable decisions regarding whether a real change has occurred between testing sessions or whether the observed change is simply a product of measurement error.

The results of our study have important implications for the application, education, and evaluation of clinical assessment skills within the sports medicine community. The data indicated that even experienced examiners with similar education, training, and advanced orthopaedic assessment skills may have difficulty attaining a high level of reliability when using the modified Thomas test to assess rectus femoris muscle flexibility about the knee joint. The findings suggest that many variables may confound the reliability of the modified Thomas test. These include but are not limited to factors such as variations in the standardized procedures used to execute the modified Thomas test, the precision of the measurement (ie, measuring to within 1°), the ROM that is measured, between-subjects variations in anthropometry and flexibility, and even the clinician's overall comfort and confidence in administering the test with various patient populations and disorders. If this is true, then one could predict that reliability values of the modified Thomas test may be even lower in a "real world" clinical orthopaedic setting in which no standardized procedures for modified Thomas testing exist and in which large variations exist in individual practitioners' educational backgrounds, experiences, clinical environments, patient populations, and musculoskeletal injuries or conditions reviewed in daily practice. These points warrant consideration by clinicians when using the modified Thomas test to evaluate rectus femoris muscle flexibility about the knee in relation to a specific disorder and to record day-to-day progress in rehabilitation programs designed to increase muscle length about the knee joint.

To study the effect of confounding variables on the clinical reliability of the modified Thomas test, alternate methodologic approaches should be investigated. For

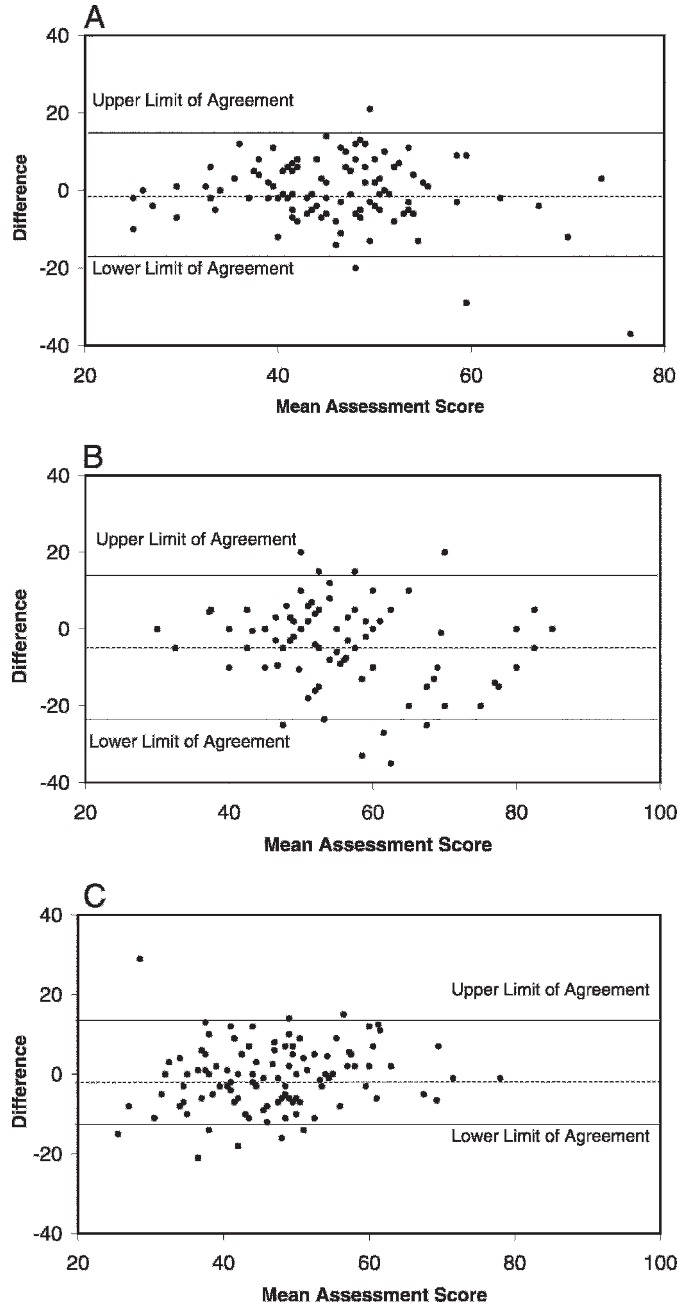

Figure 2. Bland-Altman plots (measured in degrees) depict differences in goniometer scoring between test and retest, plotted against mean scores for each participant. The dashed line shows the mean difference score (2 SDs), and the solid line represents the 95% upper and lower limits. A, Examiner 1: differences in scoring = −0.5 ± 8.4, upper limit = 16.3, lower limit = −17.4. B, Examiner 2: differences in scoring = −3.1 ± 9.9, upper limit = 16.8, lower limit = −23.0. C, Examiner 3: differences in scoring = −0.6 ± 8.0, upper limit = 15.4, lower limit = −16.5.

Table 5. Measurement Error Scores by Examiner for Goniometer Scoring

Intrarater Scoring	Standard Error of Measurement, °	Method Error, °	Coefficient of Variance, %
Examiner 1	6	6	13
Examiner 2	7	7	13
Examiner 3	6	6	13
Mean	7	6	13

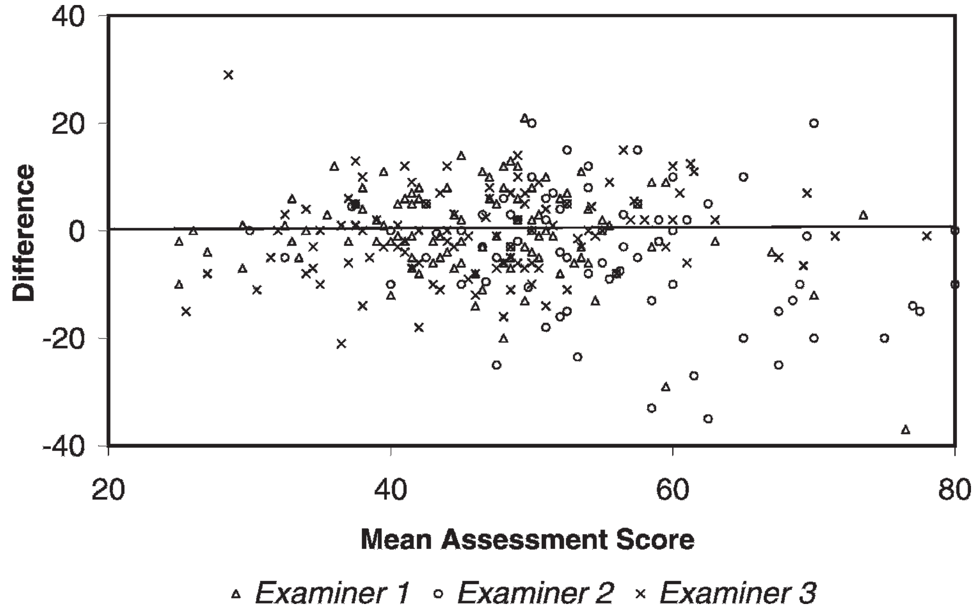

Figure 3. Bland-Altman plots (measured in degrees) illustrate a large amount of interrater variation over the 2 test sessions.

example, clinicians could score rectus femoris flexibility from digital photos or film to minimize participant variations, both within and between assessment sessions. This change would not only serve to further standardize the data collection protocol for the patient but also help limit the number of variables that may confound rater reliability. The results of such a study would help clarify the results of our study and would serve as a valuable comparison of the reliability differences between hands-on and secondary assessment of rectus femoris flexibility.

CONCLUSIONS

Our study provides important information to sports medicine practitioners regarding the reliability of the modified Thomas orthopaedic assessment technique that is commonly used to assess rectus femoris muscle flexibility about the knee joint. The data illustrated that reliable assessment using the modified Thomas test may be influenced by variations in the application of assessment criteria among examiners, the scoring method used, the consistency and accuracy of establishing surface landmarks, and the population from which the sample was selected. The results also provide clinicians with important information regarding the reliability limits of the modified Thomas test and can serve as a guide for establishing whether an observed change between testing sessions is real or simply a product of measurement error. Additionally, the methods employed for this study serve as a template to guide the evaluation or development of clinically reliable musculoskeletal assessment techniques for the lower extremity and to assist with the education of examiners and the application and evaluation of clinical assessment skills within the sports medicine community.

ACKNOWLEDGMENTS

This study was funded through the Canadian Athletic Therapists Association, Calgary, AB; Graduate Studentships and PhD Dissertation Awards from the Manitoba Health Research Council, Winnipeg, MB; health research support to ongoing projects from the Muscular Dystrophy Association, Tucson, AZ; and the University of Winnipeg, Winnipeg, MB.

REFERENCES

1. Magee DJ. Hip. In: *Orthopedic Physical Assessment.* 4th ed. Philadelphia, PA: WB Saunders; 2002:607–660.
2. Kendall FP, McCreary EK, Provance PG, Rodgers MM, Romani WA. Lower extremity. In: *Muscles: Testing and Function, With Posture and Pain.* 5th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005:359–464.
3. Prentice WE. The thigh, hip, groin, and pelvis. In: *Arnheim's Principles of Athletic Training: A Competency-Based Approach.* 11th ed. New York, NY: McGraw Hill; 2003:625–667.
4. Anderson MK, Hall SJ. Thigh, hip, and pelvis injuries. In: Balado D, ed. *Sports Injury Management*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1999:319–358.
5. Richardson JK, Iglarsh ZA. Hip. In: *Clinical Orthopaedic Physical Therapy*. Philadelphia, PA: WB Saunders; 1994:333–398.
6. Tyler T, Zook L, Brittis D, Gleim G. A new pelvic tilt detection device: roentgenographic validation and application to assessment of hip motion in professional ice hockey players. *J Orthop Sports Phys Ther*. 1996;24(5):303–308.
7. Gabbe BJ, Bennell KL, Wajswelner H, Finch CF. Reliability of common lower extremity musculoskeletal screening tests. *Phys Ther Sport*. 2004;5(2):90–97.
8. Young W, Clothier P, Otago L, Bruce L, Liddell D. Acute effects of static stretching on hip flexor and quadriceps flexibility, range of motion and foot speed in kicking a football. *J Sci Med Sport*. 2004;7(1):23–31.
9. Lee LW, Kerrigan DC, Della Croce U. Dynamic implications of hip flexion contractures. *Am J Phys Med Rehabil*. 1997;76(6):502–508.
10. Schache AG, Blanch PD, Murphy AT. Relation of anterior pelvic tilt during running to clinical and kinematic measures of hip extension. *Br J Sports Med*. 2000;34(4):279–283.
11. Harvey D. Assessment of the flexibility of elite athletes using the modified Thomas test. *Br J Sports Med*. 1998;32(1):68–70.
12. Smith AD, Stroud L, McQueen C. Flexibility and anterior knee pain in adolescent elite figure skaters. *J Pediatr Orthop*. 1991;11(1):77–82.
13. Winters MV, Blake CG, Trost JS, et al. Passive versus active stretching of hip flexor muscles in subjects with limited hip extension: a randomized clinical trial. *Phys Ther*. 2004;84(9):800–807.
14. Narvani AA, Tsiridis E, Kendall S, Chaudhuri R, Thomas P. A preliminary report on prevalence of acetabular labrum tears in sports

patients with groin pain. *Knee Surg Sports Traumatol Arthrosc*. 2003;11(6):403–408.
15. Gabbe BJ, Finch CF, Bennell KL, Wajswelner H. Risk factors for hamstring injuries in community level Australian football. *Br J Sports Med*. 2005;39(2):106–110.
16. Tyler TF, Nicholas SJ, Campbell RJ, McHugh MP. The association of hip strength and flexibility with the incidence of adductor muscle strains in professional ice hockey players. *Am J Sports Med*. 2001; 29(2):124–128.
17. Baecke JAH, Burema J, Frijters JER. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutri*. 1982;36(5):936–942.
18. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol*. 2003;56(8):730–735.
19. Peeler J, Anderson JE. Reliability of the Ely's test for assessing rectus femoris muscle flexibility and joint range of motion. *J Orthop Res*. 2008;26(6):793–799.
20. Peeler J, Anderson JE. Reliability of the Thomas test for assessing range of motion about the hip. *Phys Ther Sport*. 2007;8(1):14–21.
21. Rothstein JM, Miller PJ, Roettger RF. Goniometric reliability in a clinical setting: elbow and knee measurements. *Phys Ther*. 1983; 63(10):1611–1615.
22. France L, Nester C. Effect of errors in the identification of anatomical landmarks on the accuracy of Q angle values. *Clin Biomech (Bristol, Avon)*. 2001;16(8):710–713.
23. Szulc P, Lewandowski J, Marecki B. Verification of selected anatomic landmarks used as reference points for universal goniometer positioning during knee joint mobility range measurements. *Med Sci Monit*. 2000;7(2):312–315.
24. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000.
25. Haley SM, Osberg JS. Kappa coefficient calculation using multiple ratings per subject: a special communication. *Phys Ther*. 1989;69(11):970–974.
26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428.
27. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998;26(4):217–238.
28. Holmback AM, Porter MM, Downham D, Lexell J. Reliability of isokinetic ankle dorsiflexor strength measurements in healthy young men and women. *Scand J Rehabil Med*. 1999;31(4):229–239.
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476): 307–310.

Jason D. Peeler, PhD, CAT(C), contributed to conception and design; acquisition and analysis and interpretation of the data; and drafting, critical revision, and final approval of the article. Judy E. Anderson, PhD, contributed to conception and design; analysis and interpretation of the data; and drafting, critical revision; and final approval of the article.

Address correspondence to Jason D. Peeler, PhD, CAT(C), Department of Kinesiology and Applied Health, University of Winnipeg, 515 Portage Avenue, Winnipeg, MB, Canada, R3B 2E9. Address e-mail to j.peeler@uwinnipeg.ca.