

Normative and Critical Criteria for Iliotibial Band and Iliopsoas Muscle Flexibility

Reed Ferber, PhD, CAT(C), ATC*†‡; Karen D. Kendall, MKin, CAT(C)*†;
Lindsay McElroy, BKin, CAT(C)*†

†Running Injury Clinic and Faculties of *Kinesiology and ‡Nursing, University of Calgary, AB, Canada

Context: The Ober and Thomas tests are subjective and involve a “negative” or “positive” assessment, making them difficult to apply within the paradigm of evidence-based medicine. No authors have combined the subjective clinical assessment with an objective measurement for these special tests.

Objective: To compare the subjective assessment of iliotibial band and iliopsoas flexibility with the objective measurement of a digital inclinometer, to establish normative values, and to provide an evidence-based critical criterion for determining tissue tightness.

Design: Cross-sectional study.

Setting: Clinical research laboratory.

Patients or Other Participants: Three hundred recreational athletes (125 men, 175 women; 250 in injured group, 50 in control group).

Main Outcome Measure(s): Iliotibial band and iliopsoas muscle flexibility were determined subjectively using the modified Ober and Thomas tests, respectively. Using a digital inclinometer, we objectively measured limb position. Interrater reliability for the subjective assessment was compared between

2 clinicians for a random sample of 100 injured participants, who were classified subjectively as either negative or positive for iliotibial band and iliopsoas tightness. Percentage of agreement indicated interrater reliability for the subjective assessment.

Results: For iliotibial band flexibility, the average inclinometer angle was $-24.59^\circ \pm 7.27^\circ$. A total of 432 limbs were subjectively assessed as negative ($-27.13^\circ \pm 5.53^\circ$) and 168 as positive ($-16.29^\circ \pm 6.87^\circ$). For iliopsoas flexibility, the average inclinometer angle was $-10.60^\circ \pm 9.61^\circ$. A total of 392 limbs were subjectively assessed as negative ($-15.51^\circ \pm 5.82^\circ$) and 208 as positive ($0.34^\circ \pm 7.00^\circ$). The critical criteria for iliotibial band and iliopsoas flexibility were determined to be -23.16° and -9.69° , respectively. Between-clinicians agreement was very good, ranging from 95.0% to 97.6% for the Thomas and Ober tests, respectively.

Conclusions: Subjective assessments and instrumented measurements were combined to establish normative values and critical criteria for tissue flexibility for the modified Ober and Thomas tests.

Key Words: Ober test, Thomas test, inclinometer, reliability

Key Points

- To provide normative values and critical criteria for tissue flexibility for the modified Ober and Thomas tests, we combined subjective assessments and instrumented measurements.
- Between-examiners percentage-of-agreement measures for the subjective (dichotomous) categorization of tissue flexibility were very high.

Many special tests used by athletic trainers are highly subjective and involve either a negative or positive assessment, making them difficult to apply within evidence-based medicine.¹ Two such special tests are the Ober and Thomas tests for iliotibial (IT) band and hip flexor flexibility, respectively.

The Ober test is a commonly used test for measuring the flexibility of the IT band (Figure 1). The test is administered with the person lying on his or her side, with pelvis and shoulders aligned along the vertical plane and the knee flexed to 90° . The examiner stabilizes the pelvis with 1 hand while the other hand moves the thigh of the tested limb (the top limb) into hip flexion, abduction, and extension and then lowers the limb into adduction until it stops via soft tissue stretch or from posterior rotation of the pelvis or both. A positive test occurs when the leg remains in an abducted position (rests above the horizontal) when the abductor muscles are relaxed.^{2–11}

The modified Thomas test is commonly used to assess a hip flexor contracture or tightness of the iliopsoas muscle

(or both) (Figure 2). The person lies supine with the hip joint positioned over the edge of the examination table and flexes the hip, bringing the knee to the chest and holding it while the low back, sacrum, and pelvis remain horizontal and are stabilized by the examiner. Inability of the opposite thigh to extend to a neutral position or drop below the horizontal constitutes a positive test.^{2–5}

Although some authors have used either goniometers or inclinometers to quantify the Ober or modified Thomas test, very few have established normative values. Corkery et al⁴ reported on values for various muscle lengths, including the iliopsoas, using a goniometer, for 72 college-aged students. However, the modified Thomas test involved having the volunteer lie supine completely on the table, preventing the thigh from dropping below the table and, thus, an average angle of 2.3° above the horizontal was demonstrated. In addition, data for the IT band were not collected. Harvey⁶ assessed the flexibility of the iliopsoas, allowing the thigh to drop below the horizontal and using a goniometer on 117 elite athletes.



Figure 1. Participant position and digital inclinometer positioning for measurement of the modified Ober test.

An average angle of 11.91° below the horizontal was noted. Again, data on IT band flexibility using the Ober test were not reported. These 2 studies provide some data related to iliopsoas flexibility, but different methods of the modified Thomas test were used. Moreover, to our knowledge, very few investigators have provided a normative data range related to IT band flexibility using the Ober test. Finally, no authors have combined the subjective opinion of the clinician with the instrumented measurement from either a goniometer or inclinometer.

Therefore, the purpose of our study was to compare the subjective evaluation of IT band and iliopsoas flexibility and the instrumented measurement of a digital inclinometer with the goal of establishing normative values and providing evidence-based critical criteria for determining tissue tightness. A secondary aim was to determine the interrater reliability for the subjective assessment of IT band and iliopsoas flexibility.

METHODS

Participants

A total of 300 recreational athletes (125 men, 175 women; age = 32.3 ± 9.7 years, height = 167.5 ± 29.2 cm, mass = 73.7 ± 21.4 kg) volunteered, for a total of 600 limbs analyzed. *Recreational* was defined as at least 30 minutes of activity, 3 times per week. Within the 300 participants was a group of 50 uninjured control athletes included for comparison purposes (21 men, 29 women). The Table provides a complete breakdown of injury type and sex comparisons for the injured and control groups. Before data collection began, each volunteer signed a consent form approved by the University of Calgary Conjoint Health Research Ethics Board, which also approved the study.

Procedures

Participants were positioned for the modified Ober and Thomas tests by the clinician (K.D.K.) according to the previous literature^{2,3,6} (Figures 1 and 2). The clinician first determined IT band and iliopsoas muscle flexibility subjectively for all individuals and recorded the outcome as either negative or positive based on the definitions



Figure 2. Participant position and digital inclinometer positioning for measurement of the modified Thomas test.

provided above. Next, a digital inclinometer (Pro 360 digital protractor; SmartTool Technology, Inc, Oklahoma City, OK; accuracy = $\pm 0.1^\circ$, maximum resolution = 0.1°) was placed at the midpoint between the anterior-superior iliac spine and the patella, along the longitudinal axis of either the lateral or anterior aspect of the thigh, to objectively measure IT band or iliopsoas muscle flexibility, respectively (Figures 1 and 2). Measurements from the inclinometer were recorded to the nearest 0.1° . The order of testing and limb tested were randomized, and the procedure was repeated for the opposite limb.

Based on an a priori power analysis (intraclass correlation coefficient ^{2,1}, $r = 0.80$ agreement aspiration, $P = .05$), a subset of injured participants was randomly chosen ($n = 100$; the Table) and another clinician (L.M.) subjectively assessed IT band and iliopsoas flexibility at the same time as the first clinician, with the athlete in the same position; the second clinician privately recorded her subjective opinion. Next, the subjective opinion of the first clinician was recorded and, finally, the inclinometer value was recorded. These steps were consistently performed for

Table. Frequency and Sex Distribution of the Injured and Control Groups

Group	Men (n = 125), n (%)	Women (n = 175), n (%)	Total (n = 300), n
Injured (n = 250)			
Patellofemoral pain syndrome	17 (34)	34 (66)	51
Iliotibial band pain syndrome	10 (33)	21 (67)	31
Metatarsalgia	13 (49)	13 (51)	26
Medial tibial stress syndrome	12 (48)	13 (52)	25
Plantar fasciitis	13 (52)	12 (48)	25
Greater trochanter bursitis	9 (41)	13 (59)	22
Patellar tendinitis	7 (33)	14 (67)	21
Ischial bursitis	7 (33)	14 (67)	21
Achilles tendinopathy	11 (55)	9 (45)	20
Back pain (nonspecific)	5 (63)	3 (38)	8
Control (n = 50)	21 (42)	29 (58)	50

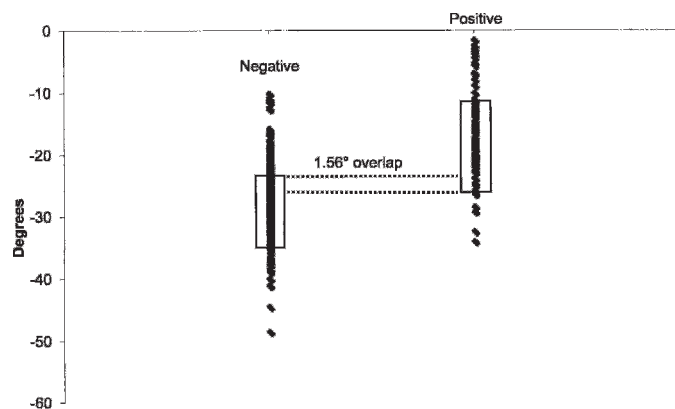


Figure 3. Individual participants' data subjectively categorized as either a negative or positive Ober test along with the ± 1 SD and critical criterion for determining iliotibial band inflexibility.

every instance of data collection and were necessary to minimize bias and examiner influence.

Statistics

Overall group, negative, and positive subjective assessment means, SDs, 95% confidence intervals (CIs), and standard error of measurement (SEM) were calculated for the instrumented (continuous) data. Comparisons of subjective and inclinometer data were analyzed descriptively. By using the +1 SD value for the group assessed as negative and the -1 SD value for the group assessed as positive, a critical criterion for muscle flexibility was established. The ± 1 SD value was used to determine the critical criterion because this value would be the most conservative compared with the 95% CIs.

Percentage of agreement between observers was chosen as the statistical representation of reliability for the subjective (dichotomous) data. This value provides a clear, straightforward, and easily interpreted statistical assessment of the data.¹² Although the κ statistic is used extensively for the purpose of reliability, we did not use it because previous authors¹²⁻¹⁴ have reported that this statistical tool is not appropriate as the primary method of quantifying agreement, particularly with multiple raters and multiple categories. Based on the findings of Sasyniuk et al,¹⁴ we defined a value of agreement between 61% and 80% as good agreement, between 81% and 99% as very good agreement, and 100% as perfect agreement.

RESULTS

For IT band flexibility, the average inclinometer angle was $-24.59^\circ \pm 7.27^\circ$ (95% CI = -23.76° , -25.41° ; SEM = 0.42°). Of the 600 limbs, 432 were subjectively assessed as negative ($-27.13^\circ \pm 5.53^\circ$, 95% CI = -26.60° , -27.66° ; SEM = 0.27°), whereas 168 were assessed as positive ($-16.29^\circ \pm 6.87^\circ$, 95% CI = -15.10° , -17.48° ; SEM = 0.61°). A 1.56° overlap was observed between the +1 SD value of the negative group and the -1 SD value of the positive group. Thus, the critical criterion (Figure 3) for the Ober test was +1 SD of the negative group plus the 1.56° overlap, or -23.16° . This value was chosen to conservatively include those participants whose values fell within the slight overlap between negative and positive.

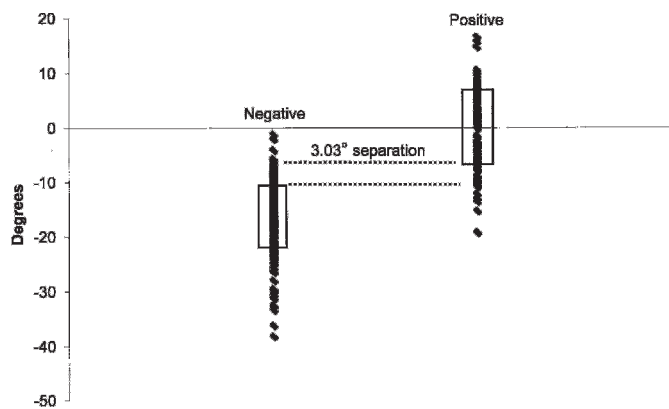


Figure 4. Individual participants' data subjectively categorized as either a negative or positive Thomas test along with the ± 1 SD and critical criterion for determining iliopsoas inflexibility.

For iliopsoas flexibility, the average inclinometer angle was $-10.60^\circ \pm 9.61^\circ$ (95% CI = -9.51° , -11.69° ; SEM = 0.55°). Of the 600 limbs, 392 were subjectively assessed as negative ($-15.51^\circ \pm 5.82^\circ$, 95% CI = -14.92° , -16.10° ; SEM = 0.30°) and 208 were assessed as positive ($0.34^\circ \pm 7.00^\circ$, 95% CI = 1.40° , -0.72° ; SEM = 0.54°). A 3.03° area of separation was observed between the +1 SD value of the negative group and the -1 SD value of the positive group. Thus, the critical criterion (Figure 4) for the Thomas test was -1 SD of the positive group plus the 3.03° gap, or -9.69° . This value was chosen to conservatively include those participants whose values fell in the gap between negative and positive.

The percentage of agreement for subjective assessment of the Ober and Thomas tests between the 2 clinicians was very good (IT band = 97.6% agreement, iliopsoas = 95.0% agreement). When the 2 clinicians did not agree upon the subjective assessment of muscle flexibility, the average measurements for those participants were 22.06° and -8.36° for IT band and iliopsoas flexibility, respectively.

DISCUSSION

The purpose of our study was to compare the subjective evaluation of IT band and iliopsoas flexibility with the instrumented measurement of a digital inclinometer in the hope of establishing normative values and providing critical criteria for muscle tightness to add to the knowledge base of evidence-based practice.

With respect to the Ober test and measures of IT band flexibility, an average angle of -24.59° was measured for all 300 participants, regardless of whether they were subjectively deemed to exhibit a negative or positive position. Using the same positioning and the modified Ober test, Hudson et al⁷ measured IT band tightness in a group of 12 control volunteers and 12 volunteers with patellofemoral pain syndrome (PFPS) using a bubble inclinometer. The authors reported a range of -20.3° to -21.4° for the control group and -14.9° to -17.3° for the PFPS group, values that are similar to our results. In addition, Herrington et al⁸ used similar positioning and measured IT band flexibility for a group of 80 control participants using a fluid goniometer. They noted an average angle of $-16.2^\circ \pm 5.4^\circ$. Furthermore, our findings are also similar to those of other authors⁸⁻¹¹ who have also

measured IT band flexibility using an instrumented device of some type. The fact that the mean IT band measure for our participants is generally higher than that in the few previous studies may simply be the result of the latter's involving a limited number of volunteers ranging from 10 to 80 participants. It is also possible that because we studied both symptomatic and asymptomatic participants and most of the previous authors studied asymptomatic participants, small differences in IT band flexibility were measured.

With respect to the Thomas test and measures of iliopsoas muscle flexibility, we found an average angle of -10.60° , which is similar to the -11.91° reported by Harvey et al,⁶ who used similar positioning and a goniometer for 117 healthy athletes. In contrast, Corkery et al⁴ demonstrated a range from -2.4° to 2.1° using a standard goniometer. However, their 72 healthy volunteers placed their thighs on the examination table, which did not allow for full hip extension. Clapis et al⁵ measured hip extension flexibility of 42 healthy participants with the thigh hanging over the examination table, using both a goniometer and a gravity inclinometer. It is interesting they noted a range of -2.8° to -1.7° , which are values similar to those of Corkery et al.⁴ Only 1 other group¹⁵ used the modified Thomas test and reported average values of -14° to -2° for healthy athletes. Again, the small sample sizes and testing of only asymptomatic volunteers in the previous studies may account for the small differences in iliopsoas flexibility between our study and those of previous investigators.

Our secondary purpose was to establish critical criteria for muscle tightness for both IT band and iliopsoas flexibility based on the subjective assessment of either a negative or positive limb position. We sought to determine interrater reliability for the subjective measures to determine if the criteria were robust and meaningful. No previous authors have reported such criteria, and we thought it was important to do so, especially in light of the growing body of evidence-based literature. By establishing a criterion for tissue flexibility, a clinician's evaluation of muscle or tissue flexibility can be made in an objective manner and clinical scientists can use these values for the purpose of research involving clinical treatment.

Because the subjective (dichotomous) measurements between the clinicians were classified as very high (95.0% to 97.6% agreement), we can conclude that the subjective definitions used by both clinicians were clear and well defined. Most important, when the clinicians did not agree on the subjective assessment of IT band flexibility, the average measurement was -22.06° , which lies within the 1.56° zone of overlap between participants subjectively assessed as either negative or positive (Figure 3). When the clinicians did not agree on the subjective assessment of iliopsoas muscle flexibility, the average measurement was -8.36° , which lies within the 3.03° zone of separation between participants subjectively assessed as either negative or positive (Figure 4).

A limited number of authors have investigated the reliability of the Ober and Thomas tests, providing mixed results. Peeler and Anderson¹⁶ reported poor statistical reliability for both goniometric and negative-positive scoring for the Thomas test. However, they used a different measurement technique for determining hip joint range of

motion than we did; their positioning placed the participant's thigh upon the examination table. In contrast, Clapis et al⁵ reported high correlations for both an inclinometer and a goniometer. These authors used a measurement technique and positioning similar to ours, and they suggested that these instruments demonstrated high reliability for measuring hip flexor muscle flexibility. Moreover, Reese et al⁹ investigated the use of an inclinometer to measure hip adduction with both the Ober test and the modified Ober test and found high reliability. Again, differences in measurement methods and positioning and the fact that the participants in these studies were free of injury may explain these conflicting results.

The critical criterion of -23.16° is the -1 SD value for the group subjectively assessed as positive for IT band inflexibility. This critical criterion thus accounts for the 1.56° overlap between the groups and is a conservative value to account for those participants whose values might fall within the -23.16° to -21.60° range in which the 2 examiners were not in complete agreement. The critical criterion of -9.69° is the $+1$ SD value for the group subjectively assessed as negative for iliopsoas inflexibility. This critical criterion thus accounts for the 3.03° separation between the groups and is a conservative value to account for those participants whose values might fall within the -9.69° to -6.66° range for which the 2 examiners were not in complete agreement.

Among the 250 injured participants, PFPS, iliotibial band syndrome (ITBS), metatarsalgia, and medial tibial stress syndrome were some of the most common injuries. In addition, a sex bias was apparent for PFPS and ITBS, with women (who constituted the majority of patients) exhibiting a 2-fold increase over men. Taunton et al¹⁷ also reported a similar sex bias for PFPS and ITBS and the occurrence of the same injuries for a group of 2002 runners investigated retrospectively using patient charts over a 2-year period. Thus, we are confident that the participant pool was consistent with and representative of the general athletic population.

Significant strengths of our study were the large sample size, the inclusion of both healthy control and injured participants, and the very high interrater percentage of agreement for subjective assessment of tissue flexibility. Still, some limitations are acknowledged. First, the critical criterion values and thus groupings of participants as either negative or positive were based on subjective evaluations of limb position. However, because the grouping of participants was supported by the ± 1 SD and 95% CI values, the separation or near separation of the 2 groups using the digital inclinometer measures, and the very high interrater agreement of the subjective assessments, we feel that our method of classification was robust and sound. Second, no previous authors have used a digital inclinometer; instead, they used bubble inclinometers, fluid inclinometers, or standard goniometers. The use of these different devices could account for the small variations in results between our study and previous investigations. In addition, although the digital inclinometer we used cost approximately \$300, it was chosen based on the accuracy and specificity of measurement. Such an expensive device may not be realistic in most clinical settings, so future studies comparing less expensive devices are warranted. Third, we recognize that subtle differences between examiners can

influence flexibility measures and that the reliability of either subjective or instrumented measures greatly depends on the examiner's positioning of the participant during the test. We attempted to minimize these effects by having the same examiner provide instructions and position the participants and by using positioning criteria based on the literature. As we have already discussed, several discrepancies exist in the literature regarding positioning, measurement method, and positive criteria for the clinical tests we used. Yet very few authors have provided normative data, most have only studied healthy volunteers in relatively small numbers, and none have combined subjective (dichotomous) and instrumented (continuous) data to establish critical criteria for these tests. Thus, the purpose of our study was not to compare different methods of performing or interpreting the Ober or Thomas tests but to provide data that could be used within an evidence-based paradigm. Finally, both examiners' subjective measurements were taken at the same time with the participant in the same position. Although this may not exactly replicate what takes place in "real life" when measurements are taken by multiple examiners in a clinical situation, it was critical and necessary to minimize bias and examiner influence.

CONCLUSIONS

Using the normative data we provide, the clinician or researcher can compare a participant's resting muscle length and make evidence-based decisions. The subjective nature of categorizing the IT band or iliopsoas muscle as inflexible or flexible can be replaced with a simple tool. In our study, subjective assessments and instrumented measurements were combined to establish normative values and critical criteria for tissue flexibility for the modified Ober and Thomas tests. In addition, the percentage-of-agreement measures for the subjective (dichotomous) categorization of tissue flexibility were very high between examiners.

ACKNOWLEDGMENTS

This study was funded by Alberta Innovates: Health Solutions. We thank Treny Sasyniuk for her administrative and logistical assistance.

REFERENCES

1. Hootman JM. Editorial: new section in *JAT*: evidence-based practice. *J Athl Train*. 2004;39(1):9.
2. Kendall FP, McCreary EK, Provance PG. *Muscles: Testing and Function with Posture and Pain*. 5th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005:33–35, 57–59.
3. McGee DJ. *Orthopedic Physical Assessment*. 5th ed. Philadelphia, PA: WB Saunders; 2007:631–632.
4. Corkery M, Briscoe H, Ciccone N, et al. Establishing normal values for lower extremity muscle length in college-age students. *Phys Ther Sport*. 2007;8(2):66–74.
5. Clapis PA, Davis SM, Davis RO. Reliability of inclinometer and goniometric measurements of hip extension flexibility using the modified Thomas test. *Physiother Theory Pract*. 2008;24(2):135–141.
6. Harvey D. Assessment of the flexibility of elite athletes using the modified Thomas test. *Br J Sports Med*. 1998;32(1):68–70.
7. Hudson Z, Darthuy E. Iliotibial band tightness and patellofemoral pain syndrome: a case-control study. *Man Ther*. 2009;14(2):147–151.
8. Herrington L, Rivett N, Munro S. The relationship between patella position and length of the iliotibial band as assessed using Ober's test. *Man Ther*. 2006;11(3):182–186.
9. Reese NB, Bandy WD. Use of an inclinometer to measure flexibility of the iliotibial band using the Ober test and the modified Ober test: differences in magnitude and reliability of measurements. *J Orthop Sports Phys Ther*. 2003;33(6):326–330.
10. Gajdosik RL, Sandler MM, Marr HL. Influence of knee positions and gender on the Ober test for length of the iliotibial band. *Clin Biomech (Bristol, Avon)*. 2003;18(1):77–79.
11. Melchione WE, Sullivan MS. Reliability of measurements obtained by use of an instrument designed to indirectly measure iliotibial band length. *J Orthop Sports Phys Ther*. 1993;18(3):511–515.
12. Feinstein AR, Cicchetti DV. High agreement but low kappa, I: the problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543–549.
13. Cicchetti DV, Feinstein AR. High agreement but low kappa, II: resolving the paradoxes. *J Clin Epidemiol*. 1990;43(6):551–558.
14. Sasyniuk TM, Mohtadi NG, Hollinshead RM, Russell ML, Fick GH. The inter-rater reliability of shoulder arthroscopy. *Arthroscopy*. 2007;23(9):971–977.
15. Wang SS, Whitney SL, Burdett RG, Janosky JE. Lower extremity muscular flexibility in long distance runners. *J Orthop Sports Phys Ther*. 1993;17(2):102–107.
16. 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.
17. Taunton JE, Ryan MB, Clement DB, McKenzie DC, Lloyd-Smith DR, Zumbo BD. A retrospective case-control analysis of 2002 running injuries. *Br J Sports Med*. 2002;36(2):95–101.

Address correspondence to Reed Ferber, PhD, CAT(C), ATC, Faculty of Kinesiology, 2500 University Drive NW, University of Calgary, Calgary, AB, Canada T2N 1N4. Address e-mail to rferber@ucalgary.ca.