

Original Research

Reliability of A Measuring Tape for Assessing Hip Adduction via Ober's Test

Anthony D'Amico, Ph.D.^{1a}, Kevin Silva, Ed.D.¹, Charla Bouranis, Ed.D.¹, Katelyn Nicolay, Ed.D.², Joseph Gallo, Ds.C.²

¹ Sport and Movement Science, Salem State University, ² Sport and Movement Science Department, Salem State University

Keywords: iliotibial band, inclinometer, flexibility, goniometer

<https://doi.org/10.26603/001c.140659>

International Journal of Sports Physical Therapy

Vol. 20, Issue 7, 2025

Background

Hip adduction is typically assessed via Ober's test, using a handheld goniometer or inclinometer. As an alternative, a measuring tape, is sometimes used for this purpose in clinical settings, but there is currently no evidence to support this method's reliability.

Purpose

The purpose of this study was to evaluate the intra-rater and inter-rater reliability, along with the measurement precision of assessing hip adduction via Ober's test with a measuring tape.

Study Design

Descriptive reliability study

Methods

A licensed athletic trainer conducted two Ober's tests on each hip of 31 healthy adults (62 hips), 30 minutes apart, and two separate licensed athletic trainers collected measurements of the test using a measuring tape, a handheld goniometer, and a digital inclinometer. Each athletic trainer was blinded to the other's results. An intraclass correlation coefficient was used to assess the inter-rater and intra-rater reliability of each measure.

Results

Measuring tape (R hip: 0.90, L hip: 0.90), handheld goniometer (R hip: 0.80, L hip: 0.83), and digital inclinometer (R hip: 0.86, L hip: 0.90) all demonstrated 'good' inter-rater reliability. Measuring tape (R hip: 0.76, L hip: 0.83), handheld goniometer (R hip: .65, L hip: 0.77), and digital inclinometer (R hip: 0.73, L hip: 0.81) all demonstrated 'moderate' to 'good' intra-rater reliability.

Conclusion

A measuring tape can be used to reliably evaluate hip adduction via Ober's test. Given its relative simplicity and low cost, it may be considered a viable tool for flexibility assessments, where applicable.

Level of Evidence

Level 3-Reliability study

^a Corresponding Author:

Anthony D'Amico

Salem State University 352 Lafayette Street, Salem, MA, USA 01970 P: 978-542-2904 F: 978-542-6554 adamico@salemstate.edu

INTRODUCTION

Ober's test is a widely used flexibility assessment of the iliotibial band (ITB) and tensor fascia latae (TFL) in fields such as physical therapy and athletic training.¹ Ilio-tibial band syndrome (ITBS) is thought to be an overuse injury typically brought on by repetitive lower-limb exercise such as running and cycling, and typically presents as lateral knee pain.^{2,3} Conservative treatments such as stretching are among the first line of treatment for ITBS, and patients often demonstrate limited hip adduction range of motion (ROM).^{4,5} Ober's test is often used clinically to assess hip adduction ROM for ITBS patients.¹ It is performed with an individual side lying on a table with their hips stacked, where their top leg is placed into a position of hip extension and then passively adducted to the table. A positive test is observed when the thigh remains abducted, whereas a negative test is observed when the thigh falls below the horizontal plane.⁶

Typically, clinicians use goniometers, digital inclinometers, or a visual assessment to determine whether a patient or athlete presents with insufficient hip adduction ROM during Ober's test.⁷ A measuring tape may serve as an alternative means of evaluating Ober's test, and has been used in clinical practice⁸ and research.⁹ The relatively simplistic nature of a measuring tape may lend itself to more precise and reproducible measurement, and it has demonstrated good reliability when evaluating thoracolumbar,¹⁰ cervical,^{10,12} and ankle dorsiflexion ROM,¹³ but its reliability has not been assessed for hip adduction ROM.

If an assessment cannot be evaluated reliably by an individual or group of clinicians, then the collected measurements are of little practical use.¹⁴ Given that a measuring tape is sometimes used to assess hip adduction ROM via Ober's test in clinical⁸ and research⁹ settings, the purpose of this study is to assess the intra-rater and inter-rater reliability, along with the measurement precision of assessing hip adduction ROM via Ober's test with measuring tape. Determining the reliability with which a practitioner can use measuring tape to evaluate Ober's test along with more traditional methods can inform clinical practice. The hypothesis is that intra-rater and inter-rater reliability of measuring tape assessments will be good.

METHODS

An observational, intra-test and inter-test study with repeated measures was conducted to evaluate the reliability of measuring hip adduction ROM via Ober's test using a measuring tape, a handheld goniometer, and a digital inclinometer. The project was approved by the university's institutional review board (IRB # 6028). All participants gave written informed consent prior to participation.

PARTICIPANTS

Thirty-one healthy adults (mean age = 20.45 ± 2.65 years, 17 male, 14 female) (62 hips), aged 18 or older, agreed to participate in the study. Participants were excluded if they

had a current lower limb or torso injury, or any condition that would make Ober's test painful. A minimum sample of 25 participants and thus 50 hips were needed given the expected intraclass correlation coefficient of .9, a power of 80%, and an alpha level of .05.¹⁵ Assessing both of a participant's joints relevant to the measure in question has been previously described in the literature.^{13,16}

RECRUITMENT

Students enrolled in collegiate health and fitness courses were recruited to participate in this investigation. After the objectives were described and the time commitment was clarified, potential participants were invited to schedule an appointment for a data collection session. Participants were screened for potential exclusion criteria in private and gave informed consent at the beginning of the scheduled session.

DATA COLLECTION

Data collection was conducted between October 1, 2023 and April 30, 2024. Each series of Ober's test measurements were conducted individually, with two athletic trainers obtaining ROM measures for comparison, a third athletic trainer performing the Ober's Test but not obtaining ROM measures, and a data scribe in the room at varying times with the participant. The evaluations took place in a private office, atop a treatment table. Each data collection session was comprised of six Ober's test measurements: An Ober's test was performed on each hip with a handheld goniometer, a digital inclinometer, and a measuring tape, with relevant landmarks obtained at the onset of each measurement.

TRAINING FOR RATERS

Prior to data collection, all raters participated in a focused training session that aimed to standardize the identification of relevant anatomical landmarks and specifics for placement of the assessment tools (described below). This preparation was designed to reduce inter-rater variability by reinforcing consistent methods of measurement.

HANDHELD GONIOMETER

A standard 360 degree plastic goniometer was used to measure the quantity of hip adduction ROM during an Ober's testing position. The same goniometer was used for all participants. The licensed athletic trainer identified anatomical landmarks for the axis of rotation, moveable arm, and stationary arm via palpation. The axis of rotation was aligned with the ipsilateral anterior superior iliac spine, the moveable arm was aligned with the mid-point of the patella, and stationary arm was aligned with the contralateral anterior superior iliac spine. The measurement was assessed and reported directly to the scribe by the licensed athletic trainer.



Figure 1. Ober's test assessment with a measuring tape

DIGITAL INCLINOMETER

A Baseline Digital Inclinometer (Fabrication Enterprises, Inc., Elmsford, NY, USA) was used to measure the quantity of hip adduction ROM in an Ober's testing position. Prior to each data collection session, an adequate battery level was confirmed and a calibration check was performed according to the manufacturer's guidelines. Before each measurement, the device was zeroed against the flat surface of the examination table. The same digital inclinometer was used for all participants. After locating and palpating the lateral epicondyle and the greater trochanter of the femur, and assessing the contours of the lateral thigh, the raters positioned the digital inclinometer on the lateral thigh proximal to the distal muscle belly of the vastus lateralis. The specific position varied between participants to ensure that the inclinometer laid flat against the lateral thigh, minimizing potential error due to contour differences. Each rater was responsible for ensuring uniform contact and alignment of the inclinometer in a way that was replicable for every measurement. The measurement was assessed and reported directly to the scribe by the licensed athletic trainer.

MEASURING TAPE

A standard retractable clinical measuring tape was used to measure the distance between the medial joint line of the knee and the table. The licensed athletic trainer identified the medial joint line of the knee as the mid-point of the medial collateral ligament, using the medial epicondyle of the femur and the medial tibial condyle as anatomical points of reference. The zero point of the measuring tape was placed on the medial joint line as described. The measurement was taken at the perpendicular point of the top surface of the table on the measuring tape.

BLINDING AND PRIVACY MEASURES

To prevent any form of bias or influence between the raters, a blinding procedure was employed; (1) Raters were blinded to each other's results. No rater had access to or was informed about the measurements obtained during data collection. (2) Measurements were taken in a private examination room, with only one rater present at any given time. This controlled setting minimized distractions and further ensured that raters could not inadvertently observe or communicate results. (3) The research team also used a scribe to record throughout the data collection to ensure raters could not see the other rater's results.

PROCEDURES

Details of the experimental design are provided in [Figure 2](#). Three licensed athletic trainers were present for each data collection session: Athletic trainer one (AT1) manually guided each participant through each Ober's test. Athletic trainers two (AT2) and three (AT3) took measurements during each test. Ober's test is a clinical examination technique used to assess the extensibility of the iliotibial band from its proximal origin at the tensor fascia latae to its distal insertion in to the anterolateral tibia. Participants were first positioned side lying on the left hip, with the right leg on top in order for that hip to be assessed first, the bottom leg flexed for stability, and the hips stacked perpendicularly to the table. Positioned behind the participant, AT1 stabilized the pelvis proximally and supported the lower leg distally. Next, AT1 passively pulled the leg into abduction and extension, before slowly releasing the lower limb toward the table with the top knee flexed to approximately 10 degrees, which has been reported to limit hip adduction ROM measurements compared to an extended knee.¹⁷ Next, AT1 identified the point of tissue restriction, indicated by restricted adduction of the thigh, and held the limb at that height and angle for the duration of the measurements that followed, approximately four minutes per leg, per measurement session. During that time, AT2 conducted hip adduction ROM assessments for the right hip using a handheld goniometer, then a digital inclinometer, then a measuring tape, for one trial each. Visual aids were not used in this investigation, in order to maintain the independence of each assessment. While AT2 conducted their assessments, AT3 was outside the room beyond listening distance, in order to keep AT2 and AT3 blinded to each other's assessments. While AT1 kept the participant in an Ober's test position, AT2 left the room, and AT3 entered. Next, AT3 conducted hip adduction ROM assessments using each of the three instruments, in the same order. Following this, AT1 instructed the participant to turn over and lay on their right hip, and manually repeated the same Ober's test procedure on the left leg and hip. During this transition, AT3 left the room, and AT2 re-entered. Next, AT2 conducted a hip adduction ROM assessment on the left hip using each of the three instruments, in the same order. While AT1 again kept the participant in an Ober's test position, AT2 left the room, while AT3 entered. Next, AT3 conducted the same assessments on the same hip in the same order as AT2. At

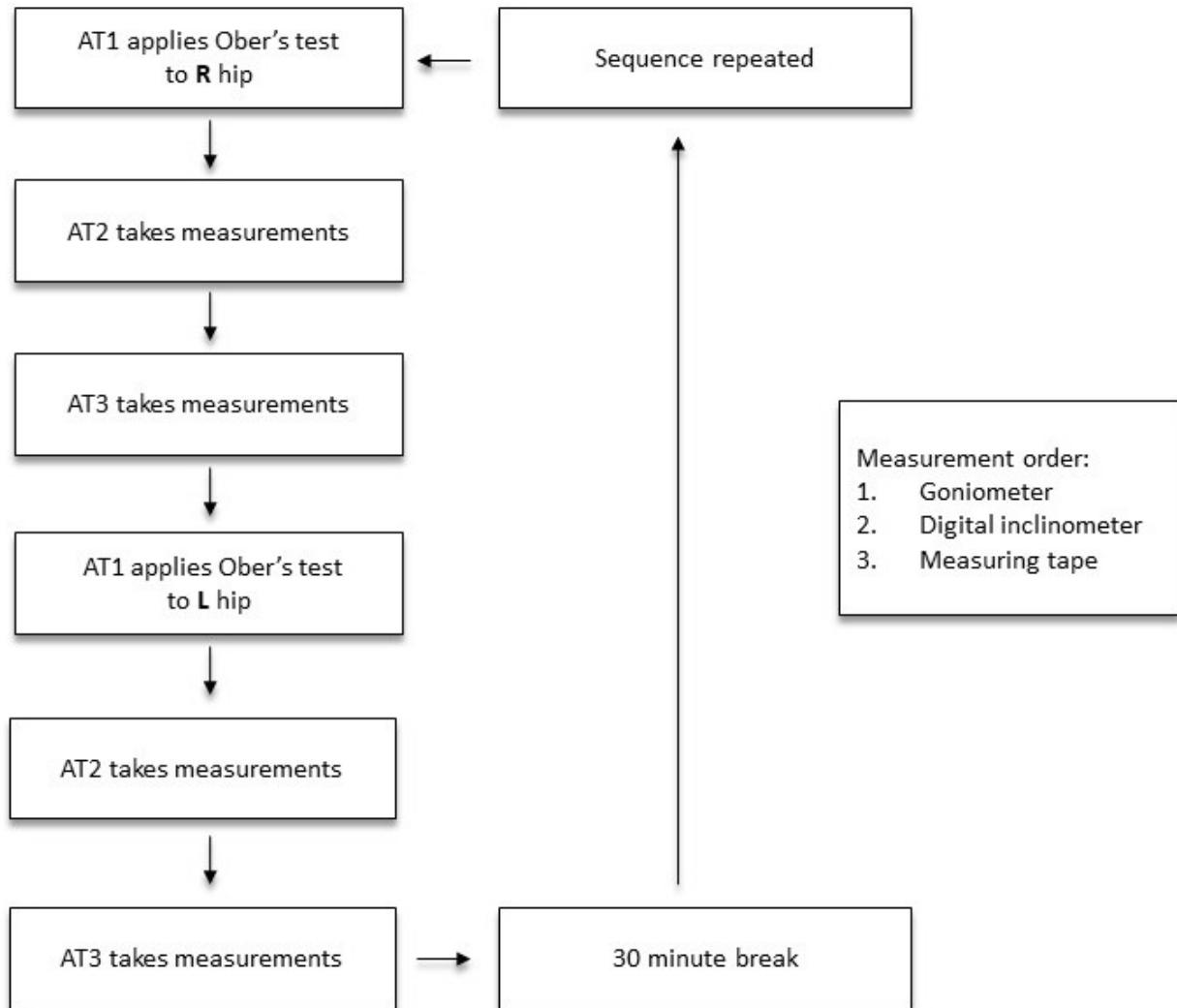


Figure 2. Data collection schematic

the conclusion of the final measurement by AT3, participants were asked to leave the assessment area and return after 30 minutes for an identical second round of assessments, and were asked not to exercise during that time. A 30 minute break period was used to minimize the influence of the stretch participants received during Ober's test in session one on flexibility assessments during session two, which was deemed sufficient based on previous investigations into the time course of acute flexibility increases following a stretching protocol.¹⁸

STATISTICAL ANALYSIS

The significance level was set at $p < 0.05$ with a confidence interval (CI) of 95%. A Kolmogorov-Smirnov test was conducted to examine quantitative data distribution (parametric versus non-parametric). Box plots and histograms were screened for homoscedasticity.¹⁹ Differences between the first-second measurement sessions and between raters were evaluated for each hip using paired samples t -tests

in instances of parametric or normal data distribution, or a Wilcoxon matched-pairs signed rank test in instances of non-parametric or non-normal data distribution. The percentage change was calculated as $((\text{Mean 1} - \text{Mean 2}) / \text{Mean 1}) \times 100$. These analyses were conducted using GraphPad Prism 10.2.3 (GraphPad Software San Diego, CA, USA).

Reliability was calculated using intraclass correlation coefficients (ICCs) via a two-way mixed model to assess intra-rater reliability, using the following equation: $(MS_R - MS_E) / MS_R + (k-1)MS_E + kn(MS_C - MS_E)$.²⁰ A two-way random model was used to establish inter-rater reliability, via the following equation: $MS_R - MS_E(MS_R)$.²⁰ Qualification of correlation strength ascribed to the guide described by Portney & Watkins (2009): (ICC < 0.5 = poor, ICC 0.5 - 0.75 = moderate, ICC > .75 = good).²¹ Where a Pearson/Spearman correlation measures the linear relationship between two distinct variables, ICC measures agreement or reliability between multiple measurements of the same variable, hence its suitability for establishing intra-rater and inter-rater reliability in a study of clinical measurements.²⁰

Table 1. Aggregate hip adduction range of motion across measurement techniques and trials

Test	Mean $\pm$ SD
Measuring Tape	20.57 $\pm$ 6.58 cm
Handheld Goniometer	-0.04 $\pm$ 7.95°
Digital Inclinator	-3.20 $\pm$ 10.01°

Standard error of measurement was assessed to determine precision as $(SD \times \sqrt{1-ICC})$, where smaller SEM values may suggest greater agreement between measures. Standard error of measurement quantifies the amount of error present in a measure due to the imprecision of the measurement tool. A clinical assessment tool must have reliably adequate precision in order to be useful. Minimum detectable change was calculated to provide a threshold at which a measurement is likely due to meaningful change versus measurement variability, calculated as: $SEM \times 1.96 \times \sqrt{2}$. The 95% limits of agreement (LOA) were calculated, which demonstrates the range of measurement error between the two raters.¹⁹ These analyses were conducted using Microsoft Excel 2016 MSO 16.0.4266.1001 (Microsoft Corporation, Redmond, WA, USA)

RESULTS

DESCRIPTIVE STATISTICS

The Kolmogorov-Smirnov test revealed that the measuring tape data were normally distributed ($p = 0.10$), while the handheld goniometer ($p = 0.05$) and digital inclinometer data ($p = 0.00$) were not. Box plots and histograms confirmed homoscedasticity across all measurements. Aggregated Ober's test measurements are provided in [Table 1](#).

COMPARISON BETWEEN MEASUREMENT SESSIONS

Descriptive statistics for trials one and two can be found in [Table 2](#). Given the parametric measuring tape data, paired samples t -test were conducted to assess differences between the first and second measurement sessions for each leg. No significant differences were observed for measuring tape in the right ($p = 0.05$) or left ($p = 0.37$) leg.

Given the non-parametric handheld goniometer and digital inclinometer data, a Wilcoxon matched-pairs signed rank test was conducted to assess differences between the first and second measurement sessions. For handheld goniometer, observed values were greater during measurement one compared to measurement two for the left leg ($1.31 \pm 7.34^\circ$ vs. $-.38 \pm 8.08^\circ$, $p = 0.02$) No significant differences were observed for handheld goniometer for the right leg ($p = 0.12$), or digital inclinometer on the right ($p = 0.83$) or left ($p = 0.74$) leg.

COMPARISON BETWEEN RATERS

Descriptive statistics for AT2 and AT3 can be found in [Table 3](#). Given the parametric measuring tape data, paired sam-

ples t -tests were conducted to assess differences between AT2 and AT3 for each leg. For measurement tape on the right leg, observed values were greater when assessed by AT3 compared to AT2 (20.97 ± 7.42 cm vs. 20.00 ± 6.53 cm, $p = 0.01$). For the left leg, observed values were also greater when assessed by AT3 compared to AT2 (21.65 ± 5.90 cm vs. 19.35 ± 5.90 cm, $p = 0.00$)

Given the non-parametric handheld goniometer and digital inclinometer data, Wilcoxon matched-pairs signed rank tests were conducted to assess differences between the two raters for each assessment. No significant differences were observed for the handheld goniometer on the right ($p = 0.15$) or left leg ($p = 0.05$), or digital inclinometer on the right ($p = 0.40$) or left leg ($p = 0.27$)

INTRA-RATER AND INTER-RATER RELIABILITY

Calculated intraclass correlation (ICC) values for intra-rater reliability from measurement 1 to measurement 2 can be found in [Table 4](#). Intra-rater reliability values ranged from 'moderate' to 'good', with 'moderate' intra-rater reliability observed for the handheld goniometer and digital inclinometer on the right leg, and 'good' intra-rater reliability observed for the measuring tape on both legs, and the handheld goniometer and digital inclinometer on the left leg.

Calculated ICC values for inter-rater reliability between AT1 and AT2 can be found in [Table 5](#). Inter-rater reliability values were 'good' for each measure, on each leg.

PRECISION AND MEASUREMENT ERROR, AND LIMITS OF AGREEMENT

Precision and measurement error statistics can be found in [Tables 4](#) and [5](#). Bland-Altman plots and the associated bias and confidence intervals can be found in [Figure 3](#). For the the measuring tape, the 95% limits of agreement (LOA) were -7.74 cm to 3.97 cm. For handheld goniometer, the 95% limits of agreement (LOA) were -10.23° to 8.34°. For digital inclinometer the 95% limits of agreement (LOA) were -10.45° to 11.49°.

DISCUSSION

The primary aim of this investigation was to evaluate the reliability of a measuring tape as a means of assessing hip adduction ROM via Ober's test. To assess consistency between raters, or inter-rater reliability, an athletic trainer measured hip adduction ROM using a handheld goniometer, a digital inclinometer, and a measuring tape, followed by another athletic trainer immediately repeating those measurements. To assess consistency within raters, or intra-rater reliability, those same measurements were repeated 30 minutes later, and each athletic trainer's first series of measurements was compared to their second. The findings of the present investigation indicate that a measuring tape is a reliable tool for evaluating hip adduction ROM via Ober's test.

Table 2. Hip adduction range of motion by leg and trial.

Test	Leg	Trial 1	Trial 2	P value
Measuring Tape	R	19.87 ± 6.32 cm	21.10 ± 7.48 cm	0.05
	L	20.87 ± 6.35 cm	20.50 ± 6.00 cm	0.37
Handheld Goniometer	R	-0.94 ± 7.3°	0.53 ± 8.54°	0.12
	L	-0.38 ± 8.08°	1.31 ± 7.34°	0.02*
Digital Inclinator	R	-3.02 ± 9.67°	-3.74 ± 10.62°	0.83
	L	-2.20 ± 9.60°	-3.02 ± 8.93°	0.74

*Significant difference between Trial 1 and Trial 2 for handheld goniometer on L leg ($p < 0.05$)

Table 3. Hip adduction range of motion by leg and rater

Test	Leg	AT2	AT3	P value
Measuring Tape	R	20.00 ± 6.53 cm	20.97 ± 7.42 cm	0.01*
	L	19.35 ± 5.90 cm	21.65 ± 5.90 cm	0.00*
Handheld Goniometer	R	-0.66 ± 6.69°	0.26 ± 9.20°	0.15
	L	0.0 ± 7.02°	0.92 ± 8.53°	0.05
Digital Inclinator	R	-3.28 ± 10.17°	-3.47 ± 10.32°	0.40
	L	-2.59 ± 9.53°	-2.15 ± 9.16°	0.27

*Significant difference between AT2 and AT3 for measuring tape on both legs ($p < 0.05$)

Table 4. Intra-rater reliability and precision measures for Ober's test assessments.

Test	R Leg			L Leg		
	ICC (95% CI)	SEM	MDC	ICC (95% CI)	SEM	MDC
Measuring Tape	0.76 (0.62, 0.85)	3.42 cm	9.48 cm	0.83 (0.73, 0.89)	2.56 cm	7.08 cm
Handheld Goniometer	0.65 (0.45, 0.77)	4.75°	13.13°	0.77 (0.62, 0.85)	3.74°	10.36°
Digital Inclinator	0.73 (0.59, 0.83)	5.30°	14.69°	0.81 (0.70, 0.88)	4.06°	11.25°

ICC = intraclass correlation, CI = Confidence interval, SEM = standard error of the measure, MDC = minimum detectable change

Table 5. Inter-rater reliability and precision measures for Ober's test assessments.

Test	R Leg			L Leg		
	ICC (95% CI)	SEM	MDC	ICC (95% CI)	SEM	MDC
Measuring Tape	0.90 (0.84, 0.94)	2.20 cm	6.12 cm	0.88 (0.23, 0.96)	2.08 cm	5.75 cm
Handheld Goniometer	0.80 (0.69, 0.87)	3.59°	9.94°	0.83 (0.74, 0.90)	3.21°	8.90°
Digital Inclinator	0.87 (0.78, 0.92)	3.68°	10.12°	0.90 (0.85, 0.94)	2.94°	8.12°

ICC = intraclass correlation, CI = Confidence interval, SEM = standard error of the measure, MDC = minimum detectable change

A measuring tape demonstrated good inter-rater reliability, with an observed intraclass correlation (ICC) of .90/.88 (R/L, right leg and left leg, respectively). The handheld goniometer and digital inclinometer also demonstrated good inter-rater reliability, with observed ICC values of .80/.83, and .87/.90 (R/L). These findings are consistent with those previously reported in the literature, where reported ICC values for goniometer and inclinometer Ober's test assessments between raters have ranged from .73-.91.^{22,23} It appears that practitioners and researchers can reliably use a measuring tape as an evaluation tool to

assess hip adduction ROM, and consistent ratings can be anticipated between multiple trained evaluators.

In addition to good inter-rater reliability, the measuring tape demonstrated good intra-rater reliability, with observed ICC values of .76/.83 (R/L). Handheld goniometer and digital inclinometer demonstrated moderate to good intra-rater reliability, with observed ICC values of .65/.77, and .73/.81 (R/L), respectively. These values suggest lower intra-rater reliability for assessing Ober's test with these tools in the present investigation compared to previous investigations, where a range of .94-.97 has been reported.^{22, 23} Methodological differences between those studies and

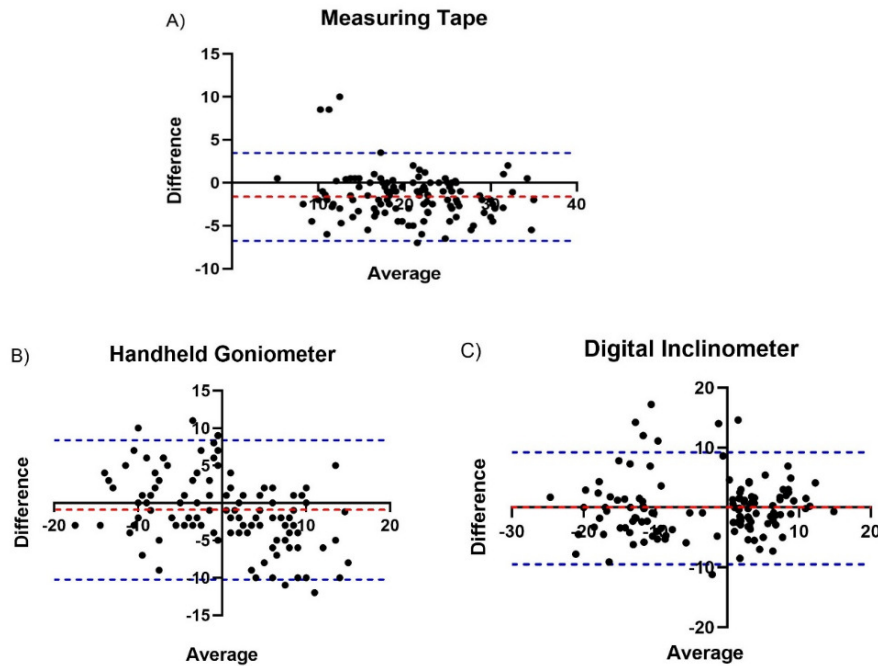


Figure 3. Bland-Altman plots for three hip adduction range of motion assessments between raters, with aggregated data of both hips.

Red dotted lines show mean differences (= bias) between the two raters. Blue dotted lines show the upper and lower 95% limits of agreement (= bias $\pm$ 1.96 x SD)

the present investigation may explain the apparent discrepancy. While also evaluating the intra-rater reliability of Ober's test assessments, Kim et al.²² and Melchione & Sullivan²³ each had a given rater record two measurements in immediate succession. Alternatively, there was a 30 minute gap between each like measurement for a given rater in the present investigation, and participants were placed into a new Ober's test for the second session by AT1, likely resulting in greater hip adduction ROM variability on an intra-rater basis. This discrepancy was not observed on an inter-rater basis, likely owing to AT2 and AT3 taking their comparable assessments a few minutes apart, with each participant held in the same Ober's test position while raters were changing positions. Methodological differences notwithstanding, the current findings appear to support the continued use of handheld goniometers or digital inclinometers for reliable assessment of the Ober's test. Additionally, these findings appear to support the use of a measuring tape as a reliable assessment. Consistent measurements can be anticipated within multiple trials for a trained evaluator, with moderate to good reliability, although the degree of precision may depend in part on the duration between assessments.

A strength of this investigation was the credentials and experience of the raters. Each athletic trainer was licensed and had at least 12 years of clinical experience, thus, the results of the investigation can be generalized to other experienced professionals. Another strength of this investigation was the blinding of the raters; neither athletic trainer was aware of the values the other had reported. Further, given the design, it is unlikely either athletic trainer was influenced by their own previous measurements, given the

amount of time that has passed, and the number of other measurements they had taken on both the participant in question, and others during the interim. A limitation of this study was the use of healthy participants; pertinent injuries or pain, especially in the ITB, may influence measurement reliability, and the generalizability of these findings to a clinical population. Another limitation of this study is the focus on hip adduction ROM and Ober's test; It is unclear whether a measuring tape can be used to reliably assess flexibility in all logistically feasible muscle groups. Another limitation is that the order of measurements was not randomized; it is possible that altering the sequence of measurements, the leg assessed, or the order in which the respective raters conducted their assessments may have influenced the results. Another limitation of the study is the compressed timeline of the assessments; the entirety of each participant's measurements were collected in an approximate 30-40 minute time frame. Measurements taken over multiple days may better reflect the manner in which changes in flexibility are often assessed in a therapeutic setting. Another limitation is that one AT held the participants' leg in the Ober's test position for approximately four minutes per side, which was a function of the study design, but differs from standard clinical practice; A shorter Ober's test may influence the consistency of hip adduction ROM assessments.

CONCLUSIONS

A measuring tape is inexpensive, simple to use, and widely available. The results of this investigation indicate that it is a reliable tool for evaluating hip adduction ROM via Ober's

test. Thus, it can provide an alternative to traditional tools such as the handheld goniometer and digital inclinometer, offering simplicity without compromising precision. Measuring the distance from the knee to the table likely introduces fewer opportunities for inconsistency compared to traditional assessment tools, and could maintain or even improve reliability of Ober's test measurements conducted by practitioners, potentially more so for the less experienced. Additionally, this method could allow for assessments to be conducted in settings with limited resources. Future studies should investigate whether a measuring tape can be used to reliably assess flexibility of other joints, and the reliability of a measuring tape as an Ober's Test assessment tool in clinical populations, such as those with hip mobility restrictions or ITB pain. Additionally, future studies should examine practical considerations for the use of a measuring tape versus other methods, such as cost and time efficiency.

ACKNOWLEDGMENTS

These results were presented in an oral presentation at the New England American College of Sports Medicine Fall Meeting in November 2023

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

Submitted: August 06, 2024 CDT. Accepted: April 30, 2025 CDT. Published: July 01, 2025 CDT.

© The Author(s)



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc/4.0> and legal code at <https://creativecommons.org/licenses/by-nc/4.0/legalcode> for more information.

REFERENCES

1. Lavine R. Iliotibial band friction syndrome. *Curr Rev Musculoskelet Med*. 2010;3:18-22. doi:[10.1007/s12178-010-9061-8](https://doi.org/10.1007/s12178-010-9061-8)
2. Strauss EJ, Kim S, Calcei JG, Park D. Iliotibial band syndrome: Evaluation and management. *J Am Acad Orthop Surg*. 2011;19(12):728-736. doi:[10.5435/00124635-201112000-00003](https://doi.org/10.5435/00124635-201112000-00003)
3. Noehren B, Schmitz A, Hempel R, Westlake C, Black W. Assessment of strength, flexibility, and running mechanics in men with iliotibial band syndrome. *J Orthop Sports Phys Ther*. 2014;44(3):217-222. doi:[10.2519/jospt.2014.4991](https://doi.org/10.2519/jospt.2014.4991)
4. Beals C, Flanagan D. A review of treatments for iliotibial band syndrome in the athletic population. *J Sports Med*. Published online 2013. doi:[10.1155/2013/367169](https://doi.org/10.1155/2013/367169)
5. Unuvar BS, Gercek H, Tufekci O, Torlak MS, Erbas O. The relationship between lower extremity muscle tightness and pain and disability in individuals with non-specific chronic low back pain. *Work*. Published online 2024. doi:[10.3233/WOR-230547](https://doi.org/10.3233/WOR-230547)
6. Gose JC, Schweizer P. Iliotibial band tightness. *J Orthop Sports Phys Ther*. 1989;10(10):399-407. doi:[10.2519/jospt.1989.10.10.399](https://doi.org/10.2519/jospt.1989.10.10.399)
7. 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. doi:[10.2519/jospt.2003.33.6.326](https://doi.org/10.2519/jospt.2003.33.6.326)
8. Doucette SA, Goble EM. The effect of exercise on patellar tracking in lateral patellar compression syndrome. *Am J Sports Med*. 1992;20(4):434-4. doi:[10.1177/036354659202000412](https://doi.org/10.1177/036354659202000412)
9. D'Amico A, Silva K, Gallo J. Time course of acute hip adduction range of motion changes following foam rolling. *Int J Sports Phys Ther*. 2021;16(2):431-437. doi:[10.26603/001c.21417](https://doi.org/10.26603/001c.21417)
10. Alaranta H, Hurri H, Heliövaara M, Soukka A, Harju R. Flexibility of the spine: normative values of goniometric and tape measurements. *J Rehabil Med*. 1994;26(3):147-154. doi:[10.2340/165019771994263147154](https://doi.org/10.2340/165019771994263147154)
11. Johnson M, Mulcahey MJ. Interrater reliability of spine range of motion measurement using a tape measure and goniometer. *J Chiropr Med*. 2021;20(3):138-147. doi:[10.1016/j.jcm.2021.09.003](https://doi.org/10.1016/j.jcm.2021.09.003)
12. Hsieh CY, Yeung BW. Active neck motion measurements with a tape measure. *J Orthop Sports Phys Ther*. 1986;8(2):88-90. doi:[10.2519/jospt.1986.8.2.88](https://doi.org/10.2519/jospt.1986.8.2.88)
13. Correll S, Field J, Hutchinson H, Mickevicius G, Fitzsimmons A, Smoot B. Reliability and validity of the halo digital goniometer for shoulder range of motion in healthy subjects. *Int J Sports Phys Ther*. 2018;13(4):707-714. doi:[10.26603/ijsp20180707](https://doi.org/10.26603/ijsp20180707)
14. Kiatkulanusorn S, Luangpon N, Sriunto W, et al. Analysis of the concurrent validity and reliability of five common clinical goniometric devices. *Sci Rep*. 2023;13(1):20931. doi:[10.1038/s41598-023-48344-6](https://doi.org/10.1038/s41598-023-48344-6)
15. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Arch Orofac Sci*. 2017;12(1):1-11. doi:[10.21315/aos2017.12.1.1](https://doi.org/10.21315/aos2017.12.1.1)
16. Konor MM, Morton S, Eckerson JM, Grindstaff TL. Reliability of three measures of ankle dorsiflexion range of motion. *Int J Sports Phys Ther*. 2012;7(3):279-287. doi:[10.26603/ijsp20120279](https://doi.org/10.26603/ijsp20120279)
17. Gajdosik RL, Sandler MM, Marr HL. Influence of knee positions and gender on the Ober test for length of the iliotibial band. *Clin Biomech*. 2003;18(1):77-79. doi:[10.1016/s0268-0033\(02\)00168-7](https://doi.org/10.1016/s0268-0033(02)00168-7)
18. Spennoga SG, Uhl TL, Arnold BL, Gansneder BM. Duration of maintained hamstring flexibility after a one-time, modified hold-relax stretching protocol. *J Athl Train*. 2001;36(1):44-48.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. doi:[10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)
21. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Pearson Prentice Hall; 2009.
22. Kim MH, Yun SJ, Weon JH. Reliability of three methods for measuring length of iliotibial tract. *Phys Ther Korea*. 2014;21(3):20-29. doi:[10.12674/ptk.2014.21.3.020](https://doi.org/10.12674/ptk.2014.21.3.020)

23. 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.
doi:[10.2519/jospt.1993.18.3.511](https://doi.org/10.2519/jospt.1993.18.3.511)