Introduction

Lateral ankle sprains (LAS) are among the most common musculoskeletal injuries, with a high incidence and recurrence rate among physically active individuals (Doherty et al., 2017). In 2018, the International Ankle Consortium carried out a modified Delphi study to offer guidance to medical professionals regarding the diagnosis and treatment of acute LAS (Delahunt et al., 2019). To individualize the treatment for each patient and prevent chronic ankle instability (CAI), these guidelines, also referred to as the "**Rehabilitation Oriented ASsessmenT**" or "**ROAST**," encourage care professionals to objectively detect mechanical, sensory and functional impairments following an ankle sprain (Delahunt et al., 2019). According to Hertel & al. (2019), CAI is characterized by a recurrent or continuous sensation of instability or giving way in the ankle joint, a tendency for recurrent ankle sprains and persistent symptoms like pain, swelling, decreased range of motion and weakness that maintain for a full year following an initial LAS incident. It is also worth noting that as many as 70% of ankle sprains could turn to CAI, potentially leading to osteoarthrosis and definitive impaired ankle function (Hertel, 2002).

A potential anteriorization of the talus may be one of the causes of CAI. This could lead to arthrokinematics impairments at the talocrural joint (Delahunt et al., 2019; Magerkurth et al., 2010). This may consequently alter the ability to perform functional activities of daily living or also result in a deficit of ankle dorsiflexion range of motion (Denegar et al., 2002; Vicenzino & Branjerdporn, 2006). The International Ankle Consortium recommended the posterior talar glide test (PTGT) to objectively assess this arthrokinematic deficit in the talocrural joint (Delahunt et al., 2019). With the patient in a seated position, the clinician should perform the PTGT by keeping the patient's foot parallel to the ground and exerting a horizontal antero-posterior force on the talus (Vicenzino & Branjerdporn, 2006). The maximal tibial inclination indicates both the degree of joint glide and the restriction of ankle arthrokinematics (Denegar et al., 2002).

Only four studies evaluated the PTGT's reliability over the last 20 years. Three studies evaluated its intra-rater reliability, although as a secondary rather than primary objective. With intraclass coefficient correlation (ICC) values ranging from 0.77 to 0.99, the PTGT demonstrated a good to excellent intra-rater reliability (Cosby & Hertel, 2011; Denegar et al., 2002; Vicenzino & Branjerdporn, 2006). Smith & al. (2015) carried out an investigation with the primary aim of evaluating the PTGT's reliability. With ICC 95% CI values ranging from -0.64 to 0.83 for intra-rater reliability and from 0.29 to 0.92 for interrater reliability, they reported lower intra-rater reliability scores than those mentioned in the previous studies. The authors stated that the PTGT might rely too much on a subjective end feel of the maneuver, making it subject to different factors such as experience of the rater, manual skills as well as the population of interest.

Therefore, the present investigation had three objectives. The primary aim was to assess the PTGT's intra- and inter-rater reliability when delivered by one experienced or one novice clinician. Secondly, we experimented the PTGT in two conditions, i.e., a regular PTGT and a modified PTGT, performed with a referential horizontal plane to standardize the end feel. Lastly, we compared the PTGT measurement into three clinical conditions related to their ankle status, including healthy individuals, Copers, and patients suffering from CAI.

Methods

Study design

The present reproducibility study was carried out in a population of volunteers. They were recruited within the University facilities (UCLouvain, Belgium). The investigation was previously approved by the local ethics committee (CEHF n°: B403201523492). Every participant signed a written consent to participate.

Population

A convenience sample of 28 participants was recruited [15 men and 13 women; Mean (Standard deviation) age 28 (8); mean height: 173,9 (8.1); mean body weight: 69.5 (11) kg].

To be included in the experiment, participants had to be aged from 18 to 65 years and reporting no history of joint dislocation or injury of the lower limbs except for a LAS, no history of surgery to the lower limbs or lower back, no history of hip pain or rheumatic, no neurological or degenerative diseases and not be pregnant. Participants were systematically excluded from the recruitment process if they reported any history of musculoskeletal, traumatic or neurological disorders apart from LAS or underwent any form of orthopedic surgery. They were also systematically excluded if they were pregnant or under any medications.

The participants were divided into three distinct groups: Healthy, Coper and CAI. The Healthy group was composed of individuals who never sustained any ankle injury. The Coper group was composed of participants presenting a history of a LAS that required immobilization and/or non-weight bearing for at least 3 days, but no episodes of ankle giving-way and/or feeling of instability for at least 12 months before enrollment. Coper individual should report a Cumberland Ankle Instability Tool score over 23 (Geerinck et al., 2020). The CAI group was composed of patient presenting a history of at least one ankle sprain associated with pain or swelling and at least one interrupted day of sports/physical activity. CAI patient should have suffered from an initial ankle sprain in the last year and reported multiple episodes (≥2) of ankle giving-way within the last 6 months and/or subjective feeling of instability. They were categorized as CAI if they provide a Cumberland Ankle Instability Tool score of 23 or below (Kwon et al., 2020).

Procedure

In total, there were three raters for the testing session. Raters A and B were blinded to all measurements and were responsible for carrying out the PTGT. With eight years of clinical experience as an Orthopedic Manual Therapist and sport physiotherapist, rater A was regarded as an experienced healthcare professional. With less than 1000 hours of clinical experience and being in the final year of his physiotherapy curriculum, rater B was regarded as a novice clinician. The third rater (C) supervised the collection and reporting of all the data. Throughout the procedure, the raters were not permitted to speak with one another to prevent bias when conducting the PTGT.

During the first session, rater C recorded the age, height, weight, level of physical activity, and occupation of the participant. Then, he measured the ankle dorsiflexion range of motion with the weight bearing lunge test (WBLT) (Powden et al., 2015). During the PTGT procedure and to prevent any movement of the knee, the participant was seated on an examination table with their legs hanging off the edge. Using an elastic band, an iPhone 11 (Apple Inc., Cupertino, United States of America) was placed on each anterior tibial crest (figure 1). The upper end of the telephone was placed two centimeters below the anterior tibial tuberosity. A smartphone app called Clinometer® was used to measure the angle of inclination (App Store, © 2020 Phoenix Solutions, Vietnam). Then, one of the raters, i.e., A or B, performed the PTGT on one of the participant's ankles. The rater started by placing the ankle in a neutral subtalar position, placing the fifth metatarsal of the foot parallel to the ground and the anterior crest of the tibia in the vertical plane, perpendicular to the ground. The rater induced a glide movement on the talus by applying a gentle anteroposterior pressure on the anterior face of the talus, leading to the displacement of the talocrural and knee joints into a flexed position. The rater stopped the test when a firm end feel was felt in the talocrural joint while keeping the foot parallel to the ground. At the end feel, the rater informed the third rater (C) to record the angle of inclination of the tibia.

Figure 1: Posterior Talar Glide Test

The modified PTGT aimed to objectify the end of range assessment. Prior to starting the box test, the rater positioned the heel half a centimeter above the box and checked that the foot's fifth metatarsal was parallel to the box's upper surface. The test was then carried out as previously mentioned, but he made sure the fifth metatarsal remained parallel to the upper surface of the box (figure 2).

The third rater (C) was the only one who could see the measure being streamed on an iPad Pro screen, while the rater conducting the PTGT was unable to see the screen of the phone. Each rater performed a total of four PTGT on each ankle—twice with the box as a guide and twice without. The order of testing was randomized with a sealed envelope. There was a minute of rest between each PTGT.

Figure 2 : Modified Posterior Talar Glide Test With Box

Statistical analysis

All statistical analyses were carried out using MedCalc® Statistical Software version 20.110 (MedCalc Software Ltd, Ostend, Belgium). The normality of the data was analyzed with the Shapiro-wilk test. Descriptive data was compared between group with a One Way ANOVA. A Tukey post hoc analysis was performed if needed.

For the first aim of the present study, an Intraclass correlation coefficient (ICC) was used to estimated intra-rater reliability (ICC 2,1) and inter-rater reliability (ICC 3,k) of PTGT for experienced and novice rater in the two conditions, i.e., PTGT and modified PTGT. ICC values of less than 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicates good reliability and values over 0.90 indicates an excellent reliability (Koo & Li, 2016). A moderate reliability was considered enough for the test to be implemented in a clinical setting. Standard error of measurement

(SEM) was calculated from the equation SEM=SD*√(1-ICC) (SD= standard deviation). The minimum detectable change (MDC) was calculated from the equation MDC=1.96*√2 SEM (Weir, 2005).

For completing the analysis, the raw data of PTGT and modified PTGT measurement was compared for each group between clinical experiences, i.e., experienced and novice clinician and PTGT conditions, i.e. with or without the reference plane. To this end, a two-way (rater x group) ANOVA was used. Bonferroni post hoc analysis was performed if needed. Finally, a One-way ANOVA was performed to compare PTGT measurement between group. If needed, a Tukey post hoc analysis was performed. Significance level was set at alpha= 0.05.

Results

Among the 28 participants, 6 were assigned to the Control group (3 women), 14 were assigned to the Coper group (6 women) and 8 were assigned to the CAI group (4 women). Descriptive statistics for the measurements of PTGT on each clinical group are presented in Table 1.

Table1. Descriptive data of all three clinical conditions

	CAI		**Coper**		**Healthy**		**pvalue**
Age (years)	28.8	(10.1)	28.7	(9.6)	27.1	(4.6)	0.597
Height (cm)	170.68	(5.07)	175.57	(8.82)	174	(9.44)	0.375
Weight (kg)	63	(7.15)	73.64	(11.81)	68.5	(10.15)	0.433
CAIT left (/30)	22.2*	(4.23)	28.5	(1.91)	28.8	(0.98)	<0.001
CAIT right (/30)	19.1*	(4.45)	27.7	(2.09)	28.3	(1.51)	<0.001

Data are presented in mean (standard deviation). Comparison between clinical group was performed using One Way ANOVA (post hoc analysis= Tukey). * Indicated significant difference compared to healthy group. Abbreviations: CAI= Chronic ankle Instability, CAIT= Cumberland Ankle Instability Tool.

Excellent intra-rater reliability for both experienced and novice rater was found with an ICC ranging from 0.930 to 0.980. In contrast, the inter-rater reliability was somewhat lower across both tests, particularly for the PTGT. Inter-rater reliability ranged from 0.662 to 0.852 (moderate). Modified PTGT provides better reproducibility outcome than PTGT with a difference in SEM of 1° for both raters between conditions. The MDC was lower for experienced rater in the modified PTGT (3.3°) and similar for all other conditions (Table 2).

Table 2. Reproducibility of PTGT measurements

			ICC	95%CI	SEM (°)	MDC (°)	Reliability
Posterior Talar Glide test	**Intra-rater (2,1)**	**Novice**	0.968	0.946 to 0.982	2.2	4.1	excellent
		Experienced	0.935	0.886 to 0.963	2.4	4.3	excellent
	Inter-rater (3,k)		0.639	0.469 to 0.754	2.8	4.7	good
Modified Posterior Talar Glide test	**Intra-rater (2,1)**	**Novice**	0.959	0.929 to 0.976	2.3	4.2	excellent
		Experienced	0.964	0.937 to 0.979	1.4	3.3	excellent
	Inter-rater (3,k)		0.777	0.662 to 0.851	2.3	4.2	good

Posterior Talar Glide test: performed without the box, Modified Posterior Talar Glide test: performed with the box. ICC= Intraclass coefficient correlation, CI= confidence interval, SEM= Standard Error of Measurement, MDC= Minimal Detectable Change. ICC classification followed the Shrout & Fleiss convention

Concerning the condition of the PTGT, the two-way ANOVA reported no statistical differences between condition (PTGT or modified PTGT, difference= 0.877°, standard error= 0.469°, pvalue = 0.062), raters (novice or experienced, difference= 0.622°, standard error= 0.469°, pvalue= 0.185) nor interaction (Table 3). One way ANOVA reported lower PTGT measurement in Coper (mean: 13.84°) compared to CAI group (mean: 17.14°) and Healthy individuals (mean: 17.39) (Table 4).

Table 3. Two-way ANOVA of PTGT measurement (in degree) between conditions and raters

	Condition				
	Modified PTGT		**PTGT**		
Rater	**Experienced**	14.38 (3.99)	**Experienced**	15.82 (3.86)	
	Novice	15.68 (5.73)	**Novice**	15.86 (5.41)	
Pairwise comparisons					**Pvalue**
difference between conditions (°)	-0.877				0.062
difference between raters (°)	-0.622				0.185
Interaction					0.213

Data are presented in mean (standard deviation). PTGT= Posterior Talar Glide Test.

Table 4. Posterior Talar Glide test measure between clinical conditions

	CAI		Coper		Healthy		Pvalue
PTGT (°)	17.14	(5.08)	13.84	(4.31) *	17.39	(4.42)	<0.001

Data are presented in mean (standard deviation). Comparison between clinical group was performed using One Way ANOVA (post hoc analysis= Tukey). * Indicated significant difference compared to healthy group. Abbreviations: CAI= chronic ankle instability, CAIT= Cumberland Ankle Instability Tool.

Discussion

To the best of our knowledge, this is the first investigation to address the lack of knowledge concerning the impact of clinical experience as well as the usefulness of a better standardization for the PTGT. Based on our results, clinical experience did not influence the reliability nor the reproducibility of the PTGT. However, modified PTGT demonstrated better reproducibility measures, especially for the experienced clinician and between raters.

Previous investigations demonstrated similar intra-rater reliability levels for PTGT, with ICC 95% CI values ranging from 0.77 to 0.99 (Cosby & Hertel, 2011; Denegar et al., 2002; Smith et al., 2015; Vicenzino & Branjerdporn, 2006). Smith & al. (2015) reported an inter-rater consistency assessment revealing an ICC with a 95% CI ranging from -0.29 to 0.92. The differences in reliability that were observed may be attributed to discrepancies in the statistical models used to assess PTGT reliability. The initial three studies employed ICC (3,1) and (3,3) models, which align with the (3,2) model used in this research, while the fourth study employed the ICC (2,3). Indeed, the controlled conditions and the small samples in which our study aim to analyze reproducibility (i.e. 8 CAI, 14 Coper and 6 Control) aligns well with mixed-effects models (i.e. ICC (3, k)) because those accurately reflects the reliability of PTGT under those subgroups conditions while including the variability existing between experienced and novice raters at the same time (Koo & Li, 2016). Because they use a random-effects model (i.e. ICC (2,k)) in their procedure, Smith & al. (2015) doesn't account for the specific fixed conditions of their participants neither for the specific training of their raters, which might underestimate the PTGT reproducibility values (Koo & Li, 2016). Moreover, using a similar model as previous studies allows us to compare our Modified PTGT reliability values with the previous PTGT ones and encourage clinicians to use both versions of the test with confidence (Cosby & Hertel, 2011; Denegar et al., 2002; Smith et al., 2015; Vicenzino & Branjerdporn, 2006).

One concern might be the relatively small sample size (n=28) used in this study and the absence of a sample size calculation. Our primary aim was to assess reproducibility through ICC, which we expected above 0,90 because of preliminary results (Cosby & Hertel, 2011; Denegar et al., 2002; Smith et al., 2015; Vicenzino & Branjerdporn, 2006). With 2 raters and 28 participants, Free critical F-value is high (≈0.99), justifying the absence of a sample size calculation. Our study found excellent intra-rater reliability (ICC near 0,97 for both experienced and novice raters), validating that a larger simple size is not necessary for our primary aim.

Although the subjective end-feel of the PTGT may negatively impact reproducibility, our results suggest that PTGT could be used differently depending on the clinical expertise of the rater. Similarly, Motantasut & al. (2019) found comparable results when testing novice and expert clinicians in assessing the navicular drop test and foot posture index in healthy individuals. Their findings showed reliable measures, despite the tests could be interpreted subjectively. As the inclinometer is the device used to obtain the measurement in degrees for PTGT, it is the presence or absence of blockage that guides the examiner. Performing PTGT is therefore more like that of a Lachman test (positive or negative) than that of a Passive Accessory Intervertebral Movement (Likert scale from 0 to 3 points), which might explain those high intra- and inter-rater reproducibility values (Mulligan et al., 2011; Stolz et al., 2020). Additionally, novice clinicians appear to be as competent as experts in musculoskeletal testing, such as static stability (Kuo et al., 2021), specific ranges of motion (Whyte et al., 2021), and using classification algorithms (Ramiscal et al., 2022). Therefore, musculoskeletal testing should be highly recommended regardless of expertise level, especially in clinical contexts (i.e.PTGT), as it relates to intra-rater reliability.

An important limitation of the original PTGT is that the rater must keep the foot parallel to the ground as well as correctly identify the sensation of the end of the glide. This limitation led

Smith & al. (2015) to attribute the level of clinical expertise to reduced inter-rater reliability. Similarly, we reported a lower inter-rater reliability for the PTGT compared to its modified version. However, the inclusion of the referential horizontal plane has brought greater benefits to experienced clinicians, as evidenced by improvements in reproducibility measures and overall reliability. It is interesting to note that novice clinicians showed limited effectiveness when using the box, suggesting a learning process for its optimal use. This result could also mean that inexperienced clinicians have not yet been influenced by standard practices (Mulligan et al., 2011; Wilkin et al., 2012). In fact, intra-rater reliability remained high regardless of whether using PTGT or the modified version, providing further evidence of the positive impact of standardization on measurement precision. However, the improvements in SEM and MDC are around 0.5°, which is relatively insignificant in clinical practice. Likewise, Li & al. (2020) reported similar findings with a modified version of the anterior drawer test for detecting talofibular ligament injury. The modified version of the anterior drawer test demonstrated better accuracy and limited false negative result. Therefore, we concluded that while the modified PTGT did not add any clinical relevance, its use or any addition of a referential horizontal plane could be used for research purposes or to compare clinicians' assessments with each other.

The Coper group reported significantly less talus mobility than the other groups. Copers may exhibit altered sensorimotor function and ankle musculoskeletal structure while maintaining normal scores on CAIT (Kwon et al., 2020). No conclusion could be drawn, as the external validity of the clinical measurement is insufficient in the present study. However, this confirmed the ability of clinicians at any experience level to use the PTGT or the modified PTGT in clinical or research practice. To the best of our knowledge, this is one of the first studies to report reproducibility measurements in healthy and CAI patients. Previous studies have emphasized the importance of assessing lower limb function in CAI or LAS patients (Khalaj et al., 2021; Kosik et al., 2020).

The use of the PTGT and the WBLT could make it possible to identify foot and ankle deficiencies more effectively and offer appropriate therapy (Delahunt et al., 2019). Restoring ankle dorsiflexion mobility and arthrokinematic properties is one of the recommendations for the management of LAS and CAI (Delahunt et al., 2019). Studies already stated that no manual technique, i.e., Maitland (Cosby & Hertel, 2011) or Mulligan (Nguyen, Mahaudens, et al., 2021) could provide immediate improvement. However, the application of pre-selected techniques without prior identification of subgroup would be inappropriate when the aim is to improve function. While it is necessary in research, it may not be as relevant in clinical conditions. Nguyen et al. (2021) reported that identifying potential responders could help achieve an average increase of 2.7 cm in ankle dorsiflexion in LAS patients. Similarly, adding the PTGT to the WBLT could help identify responders to hands-on manual therapy.

Several limitations need to be considered. Firstly, while the population included CAI patients, the sample is rather small, and results should be interpreted with caution. No clinical conclusion could be drawn from CAI and Coper group compared to healthy individual. Secondly, this reproducibility study compared rater experience and test conditions but no measures of reliability and reproducibility between groups. Future studies should therefore assess whether clinical history influences the tests. Finally, PTGT measurement for each clinical group should be considered as informative data rather than normative data. Clinicians should not feel restrained from using the PTGT due to concerns that the test is too technical and too dependent on the level of experience. This study has shown the opposite and aims to encourage even novice clinicians to use the PTGT and its modified version, which is easier to perform due to the presence of a referential horizontal plane.

Future studies interested in normative values and cut-off scores of PTGT should consider using the modified version of the PTGT in their experimental protocol. This will provide more reliable inter-rater measurements. In conclusion, skills and knowledge in the use of musculoskeletal test such as PTGT could be important. The modified PTGT could help researchers for more accurate measurement for comparison purposes. This could highlight future guidelines for managing acute LAS or CAI. Based on our results, we confirmed the possibility of using the original PTGT in clinical practice and the intra-rater condition.

The PTGT is reliable enough for clinical use by physiotherapists. The modified version of the test, incorporating a referential horizontal plane, offers greater inter-rater reliability, making it a better option for research purposes. Therefore, integrating this modified version into clinical practices and future studies could enhance the accuracy of assessments and contribute to better management of patients with LAS and CAI.

Acknowledgement

The authors wished to sincerely thank Sasha Legrand and Thomas Hansoulle for their help in the protocol and procedure.

References

Cosby, N. L., & Hertel, J. (2011). Relationships Between Measures of Posterior Talar Glide and Ankle Dorsiflexion Range of Motion. *Athletic Training & Sports Health Care*, *3*(2), 76–85. https://doi.org/10.3928/19425864-20100930-02

Delahunt, E., Bleakley, C. M., Bossard, D. S., Caulfield, B. M., Docherty, C. L., Doherty, C., Fourchet, F., Fong, D. T., Hertel, J., Hiller, C. E., Kaminski, T. W., Mckeon, P. O., Refshauge, K. M., Remus, A., Verhagen, E., Vicenzino, B. T., Wikstrom, E. A., & Gribble, P. A. (2019). *Clinical assessment of acute lateral ankle sprain injuries (ROAST): 2019 consensus statement and recommendations of the International Ankle Consortium*. 1304–1310. https://doi.org/10.1136/bjsports-2017-098885

Denegar, C. R., Hertel, J., & Fonseca, J. (2002). The Effect of Lateral Ankle Sprain on Dorsiflexion Range of Motion, Posterior Talar Glide, and Joint Laxity. In *J Orthop Sports Phys Ther* (Vol. 32). www.jospt.org

Doherty, C., Bleakley, C., Delahunt, E., & Holden, S. (2017). Treatment and prevention of acute and recurrent ankle sprain: An overview of systematic reviews with meta-analysis. In *British Journal of Sports Medicine* (Vol. 51, Issue 2, pp. 113–125). BMJ Publishing Group. https://doi.org/10.1136/bjsports-2016-096178

Geerinck, A., Beaudart, C., Salvan, Q., Van Beveren, J., D'Hooghe, P., Bruyère, O., & Kaux, J. F. (2020). French translation and validation of the Cumberland Ankle Instability Tool, an instrument for measuring functional ankle instability. Foot and ankle surgery : official journal of the European Society of Foot and Ankle Surgeons, 26(4), 391–397. https://doi.org/10.1016/j.fas.2019.05.002.

Hertel, J. (2002). Functional anatomy, pathomechanics, and pathophysiology of lateral ankle instability. *Journal of Athletic Training*, *37*(4), 364–375.

Hertel, J., & Corbett, R. O. (2019). An updated model of chronic ankle instability. *Journal of Athletic Training*, *54*(6), 572–588. https://doi.org/10.4085/1062-6050-344-18

Khalaj, N., Vicenzino, B., & Smith, M. D. (2021). Hip and knee muscle torque and its relationship with dynamic balance in chronic ankle instability, Copers and controls. *Journal of Science and Medicine in Sport*, *24*(7), 647–652. https://doi.org/10.1016/j.jsams.2021.01.009

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kosik, K. B., Johnson, N. F., Terada, M., Thomas-Fenwick, A. C., Mattacola, C. G., & Gribble, P. A. (2020). Health-Related Quality of Life Among Middle-Aged Adults With Chronic Ankle Instability, Copers, and Uninjured Controls. *Journal of Athletic Training*, *55*(7), 733–738. https://doi.org/10.4085/1062-6050-190-19

Kuo, K. T., Hunter, B. C., Obayashi, M., Lider, J., Teramoto, M., Cortez, M., & Hansen, C. (2021). Novice vs expert inter-rater reliability of the balance error scoring system in children between the ages of 5 and 14. *Gait & Posture*, *86*, 13–16. https://doi.org/10.1016/j.gaitpost.2021.02.026

Kwon, Y. U., Harrison, K., Kweon, S. J., & Williams, D. S. B., 3rd (2020). Ankle Coordination in Chronic Ankle Instability, Coper, and Control Groups in Running. Medicine and science in sports and exercise, 52(3), 663–672. https://doi.org/10.1249/MSS.0000000000002170

Li, Q., Tu, Y., Chen, J., Shan, J., Yung, P. S.-H., Ling, S. K.-K., & Hua, Y. (2020). Reverse anterolateral drawer test is more sensitive and accurate for diagnosing chronic anterior talofibular ligament injury. *Knee Surgery, Sports Traumatology, Arthroscopy*, *28*(1), 55–62. https://doi.org/10.1007/s00167-019-05705-x

Magerkurth, O., Frigg, A., Hintermann, B., Dick, W., & Valderrabano, V. (2010). Frontal and lateral characteristics of the osseous configuration in chronic ankle instability. *British Journal of Sports Medicine*, *44*(8), 568–572. https://doi.org/10.1136/bjsm.2008.048462

Motantasut, P., Hunsawong, T., Mato, L., & Donpunha, W. (2019). Reliability of novice and experienced physiotherapists using the normalized navicular height truncated and the foot posture index-6 for classifying static foot posture in adults. *Journal of Physical Therapy Science*, *31*(4), 392–397. https://doi.org/10.1589/jpts.31.392

Mulligan, E. P., Harwell, J. L., & Robertson, W. J. (2011). Reliability and Diagnostic Accuracy of the Lachman Test Performed in a Prone Position. *Journal of Orthopaedic & Sports Physical Therapy*, *41*(10), 749–757. https://doi.org/10.2519/jospt.2011.3761

Nguyen, A. P., Mahaudens, P., Detrembleur, C., Hall, T., & Hidalgo, B. (2021). Inferior tibiofibular joint mobilization with movement and taping does not improve chronic ankle dorsiflexion stiffness: a randomized placebo-controlled trial. *Journal of Manual and Manipulative Therapy*, *29*(2), 73–82. https://doi.org/10.1080/10669817.2020.1805690

Nguyen, A. P., Pitance, L., Mahaudens, P., David, Y., Hall, T., & Hidalgo, B. (2021). Effects of Mulligan Mobilization with Movement in Subacute Lateral Ankle Sprains : A Pragmatic Randomized Trial Effects of Mulligan Mobilization with Movement in Subacute Lateral Ankle. *Journal of Manual & Manipulative Therapy*, *00*(00), 1–12. https://doi.org/10.1080/10669817.2021.1889165

Powden, C. J., Hoch, J. M., & Hoch, M. C. (2015). Reliability and minimal detectable change of the weight-bearing lunge test: A systematic review. In *Manual Therapy* (Vol. 20, Issue 4, pp. 524–532). Churchill Livingstone. https://doi.org/10.1016/j.math.2015.01.004

Ramiscal, L. S., Bolgla, L. A., Cook, C. E., Magel, J. S., Parada, S. A., & Chong, R. (2022). Reliability of the scapular dyskinesis test yes-no classification in asymptomatic individuals between students and expert physical therapists. *Clinics in Shoulder and Elbow*, *25*(4), 321–327. https://doi.org/10.5397/cise.2022.01109

Smith, D. H., Hoch, J. M., Facchini, S. J., & Hoch, M. C. (2015). Intra-rater and Inter-rater Reliability and Responsiveness of the Posterior Talar Glide Test. *Athletic Training & Sports Health Care*, *7*(1), 23–28. https://doi.org/10.3928/19425864-20150121-05

Stolz, M., von Piekartz, H., Hall, T., Schindler, A., & Ballenberger, N. (2020). Evidence and recommendations for the use of segmental motion testing for patients with LBP – A systematic review. *Musculoskeletal Science and Practice*, *45*, 102076. https://doi.org/10.1016/j.msksp.2019.102076

Vicenzino, B., & Branjerdporn, M. (2006). Initial Changes in Posterior Talar Glide and Dorsiflexion of the Ankle After Mobilization With Movement in Individuals With Recurrent Ankle Sprain. *J Orthop Sports Phys Ther*, *36*(6), 464–471. https://doi.org/10.2519/jospt

Whyte, E., Ó Doinn, T., Downey, M., & O'Connor, S. (2021). Reliability of a Smartphone Goniometric Application in the Measurement of Hip Range of Motion Among Experienced and Novice Clinicians. *Journal of Sport Rehabilitation*, *30*(6), 969–972. https://doi.org/10.1123/jsr.2020-0165

Wilkin, E. J., Hunt, A., Nightingale, E. J., Munn, J., Kilbreath, S. L., & Refshauge, K. M. (2012). Manual testing for ankle instability. *Manual Therapy*, *17*(6), 593–596. https://doi.org/10.1016/j.math.2012.03.007