

The Lateral Drawer Test: A New Clinical Test to Assess Mortise Instability in Weber B Fibula Fractures

John Z. Zhao, MD¹ , Eitan M. Ingall, MD¹ , Siddhartha Sharma, MS, FRCS², Soheil Ashkani-Esfahani, MD² , Yuzuru Sakakibara, MD, PhD², Anthony Yi, MD³, Christopher P. Miller, MD, MHS⁴ , and John Y. Kwon, MD⁵

Abstract

Background: Assessment of mortise stability is paramount in determining appropriate management of ankle fractures. Although instability is readily apparent in bimalleolar or trimalleolar ankle fractures, determination of instability in the isolated Weber B fibula fracture often requires further investigation. Prior authors have demonstrated poor predictive value of physical examination findings such as tenderness, ecchymosis, and swelling with instability. The goal of this study is to test the validity of a new clinical examination maneuver, the lateral drawer test, against the gravity stress view (GSV) in a cohort of patients with Weber B fibula fractures. Secondary goals included assessing pain tolerability of the lateral drawer test, as well as testing interobserver reliability.

Methods: Sixty-two patients presenting with isolated fibula fractures were prospectively identified by an orthopaedic nurse practitioner or resident. Three nonweightbearing radiographic views of the ankle as well as a GSV were obtained. Radiographs were not visualized before conducting the lateral drawer test. Two foot and ankle fellowship-trained orthopaedic surgeons performed and graded the lateral drawer test. Radiographs were then examined and medial clear space (MCS) was measured. Visual analog scale (VAS) pain scores were obtained before and after testing. The results of the lateral drawer test were compared with radiographic measurements of MCS on GSV. A cadaveric experiment was devised to assess interobserver reliability of the lateral drawer test.

Results: Thirty (48%) of 62 consecutively enrolled patients demonstrated radiographic instability with widening of the MCS ≥ 5 mm on GSV. When correlated with MCS measurement, the lateral drawer test demonstrated a sensitivity of 83%, specificity of 97%, positive predictive value (PPV) of 96%, and negative predictive value (NPV) of 86%. There was a strong correlation between the lateral drawer test grade and amount of MCS widening (Spearman correlation $\rho = 0.82$, $P < .005$). Patients tolerated the maneuver well with an average increase of 0.7 on the VAS pain scale. Testing of 2 observers utilizing the cadaveric model demonstrated a Cohen's Kappa coefficient of 0.7 indicating moderate interobserver agreement.

Conclusion: The lateral drawer test demonstrates high sensitivity, specificity, PPV, and NPV with moderate interobserver reliability compared with the MCS on GSV in patients presenting with Weber B fibula fractures. Although further external validation is required, the lateral drawer test may offer an adjunct tool via physical examination to help determine mortise stability.

Level of Evidence: Level II, Prospective Cohort Study.

Keywords: ankle fractures, stress radiographs, physical exam

Introduction

Ankle fractures are one of the most common injuries treated by orthopaedic surgeons. Depending on the injury pattern, ankle fractures can be characterized as stable or unstable,

with unstable fractures often requiring operative treatment to stabilize the mortise and achieve anatomic reduction. Although instability can be easily discerned in bimalleolar or trimalleolar ankle fractures, in isolated lateral malleolus



Weber B ankle fractures, this assessment can be more difficult to accurately ascertain.

In the setting of an isolated Weber B ankle fracture, instability is conferred when there is a concomitant injury to the medial structures, whether that be osseous or ligamentous. Lauge-Hansen's original work emphasized the importance of medial structures as the final component to fail before conferring mortise instability.¹⁵ Although a supination external rotation (SER) IV injury readily demonstrates instability on radiography via the fractured medial malleolus, his theory postulated similar instability secondary to radiographically-occult deltoid ligament disruption. Although more recent studies have questioned the validity and reproducibility of the Lauge-Hansen classification,^{9,14,19} the seminal work nevertheless created a framework for understanding the importance of mortise stability.

In the modern-day treatment of Weber B ankle fractures, stability is frequently assessed with radiographic stress views using a medial clear space (MCS) threshold of 4 to 5 mm^{4,7,17,18,20,26} to determine stability. Although initial studies predominantly described use of the manual external rotational stress test,^{7,17} providers frequently found that many patients did not tolerate manual manipulation of the acutely injured ankle. Thus, gravity stress view (GSV) radiographs, using only the weight of the patient's own foot as a deforming force, have been increasingly used.¹⁰ Some studies advocate for the use of weightbearing radiographs to determine mortise stability, with good 1-year clinical outcomes.¹² Magnetic resonance imaging has been found to lack correlation with stress radiographs, suggesting inability to predict mortise instability.²¹

The role of physical examination in determining mortise stability has mostly fallen by the wayside. Previous studies have shown poor predictive value of medial ankle tenderness, swelling, or ecchymosis for the presence of a destabilizing deltoid ligament disruption.^{5,7,17,25} However, no study has documented a physical examination maneuver to test for mortise instability. Multiple investigations have examined the use of the anterior drawer test in assessing ankle instability and have demonstrated reliability, reproducibility, and predictive value.⁶ Similar to determining lateral ankle ligament laxity via the anterior drawer test, we propose a new clinical test, the lateral drawer test, as a tool to help determine mortise instability in isolated Weber B ankle fractures. The goal of this study is to compare the lateral

drawer test to the GSV and determine specificity, sensitivity, positive predictive value (PPV), and negative predictive value (NPV) of this clinical test maneuver.

Materials and Methods

The institutional review board of our hospital system approved this study. Sixty-two patients who presented with an isolated fibula fracture were prospectively enrolled by an orthopaedic nurse practitioner or resident in the senior authors' clinics (JYK, CPM). Inclusion criteria for this study included >18-year-old patients with an acute isolated Weber B ankle fracture presenting within 1 week from time of injury. Exclusion criteria included patients with significant barriers to GSV imaging, injuries >1 week from time of injury, unable to provide independent consent, bimalleolar or trimalleolar ankle fractures, open fractures, and ankle fracture-dislocations. Only the nurse practitioner or resident had visualized the radiographs at initial enrollment; the x-rays remained hidden to the senior authors until after clinical assessment of each patient. The research protocol was completed in all patients prior to any treatment. Patient age, gender, and laterality were recorded.

The foot and ankle fellowship-trained senior orthopaedic surgeons (JYK, CPM) were not allowed to visualize the radiographs prior to the clinical assessment of each patient. Patients were asked to rate their pain at rest via a visual analog scale (VAS) for pain (score 0-10) prior to manipulation. The senior authors then performed the lateral drawer test with the patient in a seated position. One hand stabilized the leg while the other hand performed a direct lateral stress upon the neutrally positioned ankle while gripping the hind-foot (Figure 1). During the stress maneuver, the examiner made sure to minimize inversion or eversion through the tibiotalar and subtalar joints. A thumb was placed over the medial gutter, in order to both minimize inversion/eversion but also to better feel for the translational movement. The lateral drawer test was graded by the senior author as grade 0, I, or II (Figure 2). Grade 0 corresponded to no instability/symmetric to contralateral ankle, grade I corresponded to translation <5 mm by examination, and grade II corresponded to translation \geq 5 mm by examination. Immediately after the lateral drawer test was performed, patients were again asked to rate their pain (as maximal pain experienced during the maneuver) via a VAS for pain (0-10).

¹Harvard Combined Orthopaedic Residency Program, Boston, MA, USA

²Foot & Ankle Research and Innovation Laboratory (FARIL), Massachusetts General Hospital, Weston, MA, USA

³Brigham and Women's Hospital, Boston, MA, USA

⁴Division of Foot and Ankle Surgery, Department of Orthopaedic Surgery, Beth Israel Deaconess Medical Center, Boston, MA, USA

⁵Department of Orthopaedic Surgery, Massachusetts General Hospital, Boston, MA, USA

Corresponding Author:

John Z. Zhao, MD, Harvard Combined Orthopaedic Residency Program, 55 Fruit St, Boston, MA 02114, USA.

Email: zhao.john.z@gmail.com



Figure 1. The lateral drawer test. The examiner uses one hand to stabilize the leg and with the other hand gripping the hindfoot, a laterally directed stress is applied. Care is taken to avoid inversion or eversion of the hindfoot through the talocalcaneal joint. Note the position of the thumb in the medial gutter to aid the examiner in appreciating the amount of lateral translation.

Three views of the nonweightbearing ankle (AP, lateral, mortise) as well as GSV were obtained for each patient. Radiographs were taken with the ankle in neutral dorsiflexion. GSVs were obtained as previously described by Gill et al.¹⁰ Radiographs were visualized for clinical purposes to determine treatment only after the lateral drawer test was performed. For consistency and integrity of data collection, MCS was later measured for all study patients in a deidentified manner by the senior author (JYK) and without knowledge of lateral drawer test results nor final treatment strategies. For consistency, the MCS was measured in millimeters in a standardized fashion previously described: the distance between the medial border of the talus and the lateral border of the medial malleolus on a line parallel and 5 mm below the talar dome.^{4,18}

Given limitations in assessing inter- and intraobserver agreement in the clinical cohort (further explained in the Discussion section), a cadaveric model was created to attempt to perform this analysis. Three normal left-sided transfemoral leg specimens were used after being fully thawed for >24 hours. In specimen 1, an oblique distal fibula fracture was created. In specimen 2, an oblique fibula fracture in addition to a partial deltoid ligament injury was created. In specimen 3, an oblique fibula fracture in

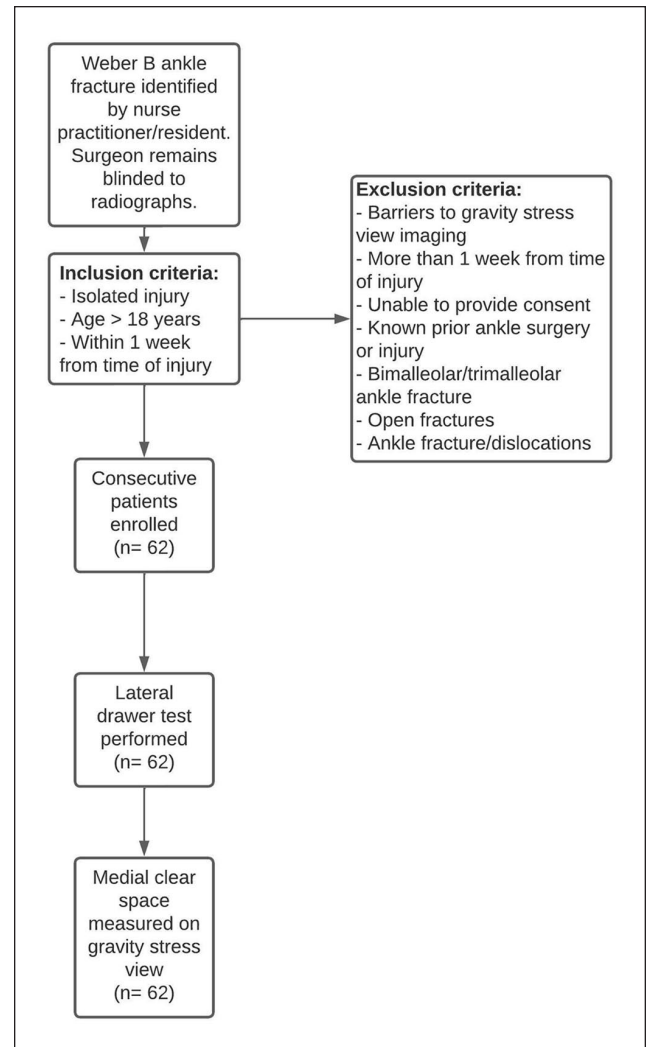


Figure 2. Study workflow. After a patient with a Weber B ankle fracture was identified by a nurse practitioner or a resident, the senior authors (JYK, CPM) performed the lateral drawer test and graded clinical instability. Medial clear space on the patients' gravity stress radiographs was then measured.

combination with a complete superficial and deep deltoid ligament injury was created. GSV was performed using a large c-arm with calibrated image capacity on each cadaver, which demonstrated an MCS of 2.7 mm, 4.4 mm, and 6.9 mm for specimens 1, 2, and 3, respectively. Most of the specimen was gently overwrapped with Coban (3M, Saint Paul, MN) while leaving the ankle area exposed. The cadavers otherwise had no identifiable markings and were similar in appearance.

The same 2 senior orthopaedic surgeons (JYK, CPM) performed the lateral drawer tests while blindfolded to the status of each cadaver. Cadavers were presented in random order after each assessment, and each surgeon performed 10 trials on each of the cadavers for a total of 30 trials per

Table 1. Lateral Drawer Grade (0, I, II) Compared With Measurement of Medial Clear Space on GSV Radiographs.^a

Grade	GSV Stable (<5 mm Widening)	GSV Unstable (>5 mm Widening)	Total
0	16	1	17
I	15	4	19
II	1	25	26
	NPV = 0.86 Sensitivity = 0.83	PPV = 0.96 Specificity = 0.97	

Abbreviations: GSV, gravity stress view; NPV, negative predictive value; PPV, positive predictive value.

^aSensitivity, specificity, PPV, and NPV are listed.

surgeon. Repeat GSV was performed periodically during the experiments to ensure integrity of the specimens and that repeat stressing did not lead to additional destabilization beyond what was initially created.

Power analysis before the study calculated a sample size of at least 40 subjects needed to observe a difference between a sensitivity/specificity null hypothesis of 0.5 and an alternative hypothesis of 0.8.¹ Alpha was set at 0.05, power was set at 0.8, and prevalence was set at 50%. Expected prevalence of unstable ankle fractures among isolated fibula fractures was estimated from prior studies using the GSV as the gold standard.¹⁰

The sensitivity, specificity, PPV, and NPV were calculated for the lateral drawer test, using the GSV as the gold standard and “true” test for instability. Lateral drawer tests with grade 0 or I were classified as stable, and tests with grade II were classified as unstable. Correlation between the lateral drawer test (stable or unstable) and the GSV MCS (MCS <5 mm or MCS ≥5 mm) was calculated using the Spearman correlation coefficient. Comparisons between VAS pain score before and after the lateral drawer test were done via the Wilcoxon signed rank test. Comparisons in VAS pain scores between subgroups grade 0, I, and II were performed via the Kruskal-Wallis test.

Interrater agreement was assessed by the Cohen κ ; values of 0-0.2 were deemed to have no agreement; 0.21-0.39, minimal; 0.4-0.59, weak; 0.6-0.79, moderate; 0.8-0.9, strong; and values >0.9 were deemed to have an almost perfect agreement. Lateral drawer test grades 0 and I were combined and compared to grade II to ensure that the analysis categories were identical to those used for the clinical study arm.

Results

A total of 62 patients met inclusion criteria. There were 21 males and 41 females, with a mean age of 49 years (range, 21-85). Of the 62 patients, there were 31 ankle fractures on the left leg and 31 on the right.

Thirty patients (48%) demonstrated an unstable ankle fracture according to an MCS ≥5 mm on GSV. The stable ankle fracture group had a median MCS of 4.2 mm (IQR = 0.85), whereas the unstable ankle fracture group had a

median MCS of 6.85 mm (IQR = 2.48). In terms of the lateral drawer test, patients were characterized as shown in Table 1. There was a strong correlation between the lateral drawer test grade and extent of MCS widening (Spearman correlation $\rho = 0.82$, $P < .005$). The lateral drawer test demonstrated a sensitivity of 0.83, specificity of 0.97, PPV of 0.96, and an NPV of 0.86 when compared with the GSV.

There was an average increase of 0.7 in the VAS pain score before and after the lateral drawer test (4.2 average pretest, 4.9 average posttest). Although this difference between the pretest and posttest scores was statistically significant ($P < .005$), there was no statistical difference between grade 0, I, or II for pretest VAS scores, posttest VAS scores, or the difference between the 2 scores ($P = .75$, $.35$, $.65$, respectively). No patients demonstrated guarding during the examination that precluded the provider from performing the examination.

Regarding the cadaveric portion of this work, Cohen kappa was 0.7 (agreement = 88.7%, expected agreement = 55.6%, standard error = 0.18, $Z = 3.83$, $P = .0001$), indicating moderate agreement between the 2 observers.

Discussion

Evaluation of mortise instability in the isolated fibula fracture is currently performed radiographically, with measurements performed either with stress maneuvers or after a period of physiologic loading. Although the ability to diagnose mortise instability by physical examination would be useful, clinical examination has essentially been abandoned given prior studies suggesting a lack of diagnostic accuracy. The works by Egol et al,⁷ McConnell et al,¹⁷ DeAngelis et al,⁵ and Stenquist et al²⁵ demonstrated relatively poor sensitivity, specificity, and predictive value of medial-sided tenderness, ecchymosis, and swelling with ankle instability. Instead, manual stress radiography,^{7,17,19} gravity stress radiography,^{10,16,23} weightbearing radiography,¹² and/or a trial of physiologic loading^{13,27} are predominantly used for assessing stability.

The lateral drawer test draws heavy parallels with the anterior drawer test used to test for lateral ankle sprains and instability. The ankle anterior drawer test applies an anterior translational force across the tibiotalar joint and tests

for the amount of laxity and the presence or absence of a firm endpoint. Although some studies have questioned the accuracy of the anterior drawer test,⁸ other studies have shown good sensitivity, specificity, and interrater and intrarater reliability.^{2,6,22} When evaluating patients for ankle instability, the anterior drawer test is nearly ubiquitously performed, is well tolerated by patients, and is frequently a tool for operative decision making. Just as the anterior drawer test is a contributing component in the evaluation of ankle instability, we believe the lateral drawer test may serve a useful role in the evaluation of ankle fractures.

The results of the clinical portion of this study demonstrate that the lateral drawer test has high sensitivity, specificity, PPV, and NPV when compared with the GSV for evaluating mortise instability. The maneuver is well tolerated by patients, resulting in minimal pain increase after the test. The lateral drawer test is quick to perform and would be seamlessly integrated in the physical examination. In terms of how the test could be incorporated in the treatment algorithm of ankle fracture patients, we recognize that a sensitivity of 0.83 with moderate interobserver agreement is still too low and would potentially undertreat some ankle fractures with an unstable mortise. Thus, it cannot be the sole determinant for ankle instability. We advocate that the lateral drawer test is just an additional piece of data that, coupled with additional stress radiographs, may help guide surgical decision making.

Instead, the lateral drawer test may help determine optimal timing of stress radiographs for patients, whether it be manual, gravity, or weightbearing stress views. We propose the following algorithm: patients with a grade 0 or I test may safely forgo additional stress radiographs at the initial clinic visit and may be asked to follow up after 1 week of weightbearing as tolerated. Weightbearing radiographs may be taken at that later time. For patients with a grade II test, however, practitioners will want to obtain a stress radiograph and evaluate for MCS widening at the time of initial visit. In such a fashion, the lateral drawer test would help decrease the number of patients who need additional stress radiographs at the initial clinic visit and shorten the waiting time for patients with grossly unstable ankle fractures who may otherwise be told to trial a period of weightbearing.

This study has a few inherent limitations to consider. First, although the lateral drawer test was performed by 2 independent orthopaedic surgeons, interrater reliability was not feasible to determine on the clinical cohort. Because of the nature that each surgeon had their own clinics on different days and in different locations, the same patient could not be physically evaluated by both providers without requesting the patient to present twice. We felt that this would have been too onerous for the patient. Additionally, intrarater reliability was unable to be tested for. Given the need to initiate treatment for patients, repeat lateral drawer testing at a future time point (with requisite patient

deidentification to reduce bias) was not practical. Although there was only a minimal increase in VAS scores, repeated testing on the same patient may have unacceptably subjected the patient to increased pain, burden, and unknown potential for iatrogenic cartilage injury.

Therefore, in order to further attempt to validate the lateral drawer test, we used an experimental cadaveric model. We felt the use of cadavers would most closely replicate the in vivo clinical situation while overcoming some of the pragmatic issues described above. As interobserver agreement was moderate, we feel our findings give further support for the test's potential clinical applicability but requires external validation.

We did not, however, assess intraobserver agreement as we thought this may be scientifically unsound even using our cadaveric model. An interval period of time is required between 2 observations of the same data set, when assessing intraobserver agreement. Although a standard length of time for this "washout" period is ill-defined, we felt the process of refreezing-thawing the cadavers would have altered the specimens for the second observations because of the known effects of refreezing and repeat thawing on the biomechanical properties of cadaver tissues. Although we recognize the limitations of our model as well as those inherent when extrapolating cadaveric work to in vivo conditions, we feel the results of the cadaveric portion of this investigation further support the clinical utility of the lateral drawer test.

Second, the amount of force applied during the lateral drawer test was not quantified via a force gauge. Although this may result in different providers applying different amounts of force during the test, it should be recognized that similar tests (like the Lachman maneuver of the knee or the anterior drawer test of the ankle) are commonly performed in clinical practice without force quantification. As the VAS pain scores did not markedly increase after the application of the lateral drawer test, we do not believe an excessive amount of force is required such that it generates significant discomfort to the patient. Finally, although precise force measurements using a tensiometer may have improved the scientific rigor of this work, it may also have decreased clinical applicability.

Third, the grading system for the lateral drawer test sets a threshold at 5 mm of translation. Similar to the Lachman grading system, it delineates instability via an assessment in quanta of millimeters. However, the authors acknowledge that small differences, such as the difference between 4 and 6 mm, may be difficult to discern. This weakness in differentiating between grades may impact reliability but is not unlike other physical examination tests in orthopaedics used for instability around joints.

Fourth, the authors recognize that the best way to determine whether a Weber B-type fracture is stable or unstable is still a subject of considerable interest, and specific debate

exists regarding the use of GSV vs weightbearing stress view radiographs.¹² Recent studies suggest that the GSV radiograph may overestimate instability that requires surgery compared with weightbearing stress views,²⁴ potentially leading to avoidable complications.^{3,11} However, it remains unclear if the newer weightbearing stress view should be the new standard. In this study, we used the more traditional GSV as the gold standard to compare the lateral drawer test against; the authors recognize that correlations may be different if the lateral drawer test was compared against weightbearing stress views.

Finally, although a Weber B ankle fracture is predominantly an external rotation injury, the lateral drawer test uses direct lateral translation to gauge mortise instability on physical examination. Although using external rotation for the physical examination maneuver would more directly correlate with the injury pattern, the authors feel that quantifying a physical examination maneuver in number of degrees shifted would be more difficult and less reliable. The amount of lateral translation applied (in millimeters) more directly correlates with MCS widening (also in millimeters), which provides easier conceptual understanding for a physical examination maneuver.

Although we recognize the above limitations, they should be considered in the context of study implications, specifically the development of a physical examination maneuver that aids in assessment of patients sustaining an exceedingly common clinical problem. As the use of physical examination to assess mortise instability has largely been abandoned in modern-day practice, the current report and preliminary validation of the lateral drawer test has moderate implications to clinical practice. The lateral drawer test is well tolerated and demonstrates high sensitivity, specificity, PPV, NPV, and moderate interobserver agreement compared with the GSV for detecting instability in patients with an isolated Weber B ankle fracture. Although this test alone is not sufficient to determine final treatment, it may provide an adjunct tool to help predict mortise instability and to optimize timing for stress radiography in patients.

Ethics Approval

Ethical approval for this study was obtained from the Beth Israel Deaconess Medical Center IRB (approval no. 2020P000449).





Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. ICMJE forms for all authors are available online.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

John Z. Zhao, MD,  <https://orcid.org/0000-0002-8303-0701>
 Eitan M. Ingall, MD,  <https://orcid.org/0000-0003-4842-5801>
 Soheil Ashkani-Esfahani, MD,  <https://orcid.org/0000-0003-2299-6278>
 Christopher P. Miller, MD, MHS,  <https://orcid.org/0000-0001-8259-7629>

References

1. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. *J Clin Diagn Res*. 2016;10(10):1-6.
2. Croy T, Koppenhaver S, Saliba S, Hertel J. Anterior talocrural joint laxity: diagnostic accuracy of the anterior drawer test of the ankle. *J Orthop Sports Phys Ther*. 2013;43(12):911-919. doi:10.2519/jospt.2013.4679
3. Dawe EJC, Shafafy R, Quayle J, Gougoulas N, Wee A, Sakellariou A. The effect of different methods of stability assessment on fixation rate and complications in supination external rotation (SER) 2/4 ankle fractures. *Foot Ankle Surg*. 2015;21(2):86-90. doi:10.1016/j.fas.2014.09.010
4. DeAngelis JP, Anderson R, DeAngelis NA. Understanding the superior clear space in the adult ankle. *Foot Ankle Int*. 2007;28(4):490-493. doi:10.3113/FAI.2007.0490
5. DeAngelis NA, Eskander MS, French BG. Does medial tenderness predict deep deltoid ligament incompetence in supination-external rotation type ankle fractures? *J Orthop Trauma*. 2007;21(4):244-247. doi:10.1097/BOT.0b013e3180413835
6. Dijk CNV, Lim LS, Bossuyt PM, Marti RK. Physical examination is sufficient for the diagnosis of sprained ankles. *J Bone Joint Surg Br*. 1996;78(6):958-962.
7. Egol KA, Amirtharajah M, Tejwani NC, Capla EL, Koval KJ. Ankle stress test for predicting the need for surgical fixation of isolated fibular fractures. *J Bone Joint Surg Am*. 2005;87(8):1885-1886. doi:10.2106/00004623-200508000-00040
8. Fujii T, Luo ZP, Kitaoka HB, An KN. The manual stress test may not be sufficient to differentiate ankle ligament injuries. *Clin Biomech*. 2000;15(8):619-623.
9. Gardner MJ, Demetrakopoulos D, Briggs SM, Helfet DL, Lorch DG. The ability of the Lauge-Hansen classification to predict ligament injury and mechanism in ankle fractures: an MRI study. *J Orthop Trauma*. 2006;20(4):267-272. doi:10.1097/00005131-200604000-00006
10. Gill JB, Risko T, Raducan V, Grimes JS, Schutt RC. Comparison of manual and gravity stress radiographs for the evaluation of supination-external rotation fibular fractures. *J Bone Joint Surg Am*. 2007;89(5):994-999. doi:10.2106/JBJS.F.01002
11. Gougoulas N, Sakellariou A. When is a simple fracture of the lateral malleolus not so simple? How to assess stability, which ones to fix and the role of the deltoid ligament. *Bone Joint J*. 2017;99-B(7):851-855. doi:10.1302/0301-620X.99B7.BJJ-2016-1087.R1
12. Holmes JR, Acker WB, Murphy JM, McKinney A, Kadakia AR, Irwin TA. A novel algorithm for isolated Weber B ankle fractures: a retrospective review of 51 nonsurgically treated

- patients. *J Am Acad Orthop Surg*. 2016;24(9):645-652. doi:10.5435/JAAOS-D-16-00085
13. Hoshino CM, Nomoto EK, Norheim EP, Harris TG. Correlation of weightbearing radiographs and stability of stress positive ankle fractures. *Foot Ankle Int*. 2012;33(2):92-98. doi:10.3113/FAI.2012.0092
 14. Kwon JY, Chacko AT, Kadzielski JJ, Appleton PT, Rodriguez EK. A novel methodology for the study of injury mechanism: ankle fracture analysis using injury videos posted on youtube.com. *J Orthop Trauma*. 2010;24(8):477-482. doi:10.1097/BOT.0b013e3181c99264
 15. Lauge-Hansen N. Fractures of the ankle. II. Combined experimental-surgical and experimental-roentgenologic investigations. *Arch Surg*. 1950;60(5):957-985.
 16. LeBa TB, Gugala Z, Morris RP, Panchbhavi VK. Gravity versus manual external rotation stress view in evaluating ankle stability: a prospective study. *Foot Ankle Spec*. 2015;8(3):175-179. doi:10.1177/1938640014565048
 17. McConnell T, Creevy W, Tornetta P III. Stress examination of supination external rotation-type fibular fractures. *J Bone Joint Surg Am*. 2004;86(10):2171-2178.
 18. Metitiri O, Ghorbanhoseini M, Zurakowski D, Hochman MG, Nazarian A, Kwon JY. Accuracy and measurement error of the medial clear space of the ankle. *Foot Ankle Int*. 2017;38(4):443-451. doi:10.1177/1071100716681140
 19. Michelson J, Solocoff D, Waldman B, Kendell K, Ahn U. Ankle fractures: the Lauge-Hansen classification revisited. *Clin Orthop Relat Res*. 1997;345:198-205.
 20. Murphy JM, Kadakia AR, Irwin TA. Variability in radiographic medial clear space measurement of the normal weight-bearing ankle. *Foot Ankle Int*. 2012;33(11):956-963. doi:10.3113/FAI.2012.0956
 21. Nortunen S, Lepojärvi S, Savola O, et al. Stability assessment of the ankle mortise in supination-external rotation-type ankle fractures: lack of additional diagnostic value of MRI. *J Bone Joint Surg*. 2014;96(22):1855-1862. doi:10.2106/JBJS.M.01533
 22. Parasher RK, Nagy DR, Em AL, Phillips HJ, Mc Donough AL. Clinical measurement of mechanical ankle instability. *Man Ther*. 2012;17(5):470-473. doi:10.1016/j.math.2012.02.015
 23. Schock HJ, Pinzur M, Manion L, Stover M. The use of gravity or manual-stress radiographs in the assessment of supination-external rotation fractures of the ankle. *J Bone Joint Surg Br*. 2007;89(8):1055-1059. doi:10.1302/0301-620X.89B8.19134
 24. Seidel A, Krause F, Weber M. Weightbearing vs gravity stress radiographs for stability evaluation of supination-external rotation fractures of the ankle. *Foot Ankle Int*. 2017;38(7):736-744. doi:10.1177/1071100717702589
 25. Stenquist DS, Miller C, Velasco B, Cronin P, Kwon JY. Medial tenderness revisited: is medial ankle tenderness predictive of instability in isolated lateral malleolus fractures? *Injury*. 2020;51(6):1392-1396. doi:10.1016/j.injury.2020.03.029
 26. Tornetta P. Competence of the deltoid ligament in bimalleolar ankle fractures after medial malleolar fixation. *J Bone Joint Surg Am*. 2000;82(6):843-848. doi:10.2106/00004623-200006000-00011
 27. Weber M, Burmeister H, Flueckiger G, Krause FG. The use of weightbearing radiographs to assess the stability of supination-external rotation fractures of the ankle. *Arch Orthop Trauma Surg*. 2010;130(5):693-698. doi:10.1007/s00402-010-1051-1