Reliability, Surgeon Perioperative Preferences, and Eye-Tracking Assessment of the Stress Examination of the Ankle Syndesmosis

Sara L. Naguib, DPM[a], and Andrew J. Meyr, DPM FACFAS[b]

[a]Resident, Temple University Hospital Podiatric Surgical Residency Program, Philadelphia, Pennsylvania

[e]Clinical Associate Professor and Residency Program Director, Department of Podiatric Surgery, Temple University School of Podiatric Medicine and Temple University Hospital, Philadelphia, Pennsylvania (AJMeyr@gmail.com)*

*Please don't hesitate to contact AJM with any questions/concerns. He's happy to provide you with a .pdf of this poster if you email him.

Statement of Purpose and Literature Review

The diagnosis and stabilization of ankle syndesmotic injuries following acute injury remains an area of controversy within the foot and ankle surgical literature, seemingly without universal consensus [1-6]. Although much of this discussion has focused on hardware constructs and determination of reduction, one specific area within this broad topic that has particularly interested our group is the intra-operative stress examination of syndesmotic stability. Fracture radiographic characteristics and Lauge-Hansen classification might provide surgeons with some degree of pre-test probability of syndesmotic disruption certainly [7,8], but it has been our clinical experience that most surgeons primarily rely on the intra-operative bone hook test, stress dorsiflexion-external rotation test, and/or a combination of these and other clinical tests during ankle fracture open-reduction-internal fixation (ORIF). However, we are unaware of any standardized method for the performance and/or interpretation of these commonly performed surgical techniques, despite the fact that it likely plays a large role in surgical and functional outcomes [1-6]. This introduces the possibility of unwanted subjectivity and variability in the performance of a purportedly objective diagnostic test.

The primary objective of this investigation was to determine the level of agreement and reliability of the stress examination of the ankle syndesmosis. Secondary objectives were to determine surgeon preferences with respect to this testing and to utilize gaze recognition software to perform an eye-tracking assessment during the performance of the test.

Methodology

Following approval by our IRB, 12 board-certified foot and ankle surgeons, 12 senior-level podiatric residents, and 12 4th year podiatric medical students were recruited and consented to participate. Participants were shown 5 intra-operative fluoroscopic images of the stress examination of the ankle syndesmosis. This included 3 still images and 2 videos, and stress examinations performed with the dorsiflexion-external rotation test (maximal dorsiflexion and external rotation of the talus against the lateral ankle mortise) and Cotton hook test (tibial-fibular distraction with bone hooks or clamps). All images contained a distal-lateral fibula hardware construct with a reduced fibular fracture, and 3 (60.0%) of 5 also contained evidence of medial malleolar fixation. No images had clear evidence of posterior malleolar involvement. Participants evaluated each image/video and reported whether they felt the test was "positive" (indicating the syndesmosis was unstable and they would perform operative stabilization) or "negative" (indicating the syndesmosis was stable and they would not perform operative stabilization).

The primary outcome measure was considered the level of agreement between board-certified surgeons with respect to the interpretation of the stress examinations. This was measured with a percent count. However, as there is a 50% likelihood that participants would agree on the interpretation simply as a result of chance within this design, reliability was also measured using the Fleiss' Kappa. This is a measure of agreement between more than two raters when data is categorical, in this case "positive" versus "negative". An established value interpretation of the kappa statistic is as follows: Kappas from 0.01 to 0.20 indicate "slight" agreement, from 0.21 to 0.40 indicate "fair" agreement, from 0.41 to 0.60 indicate "moderate" agreement, from 0.61 to 0.80 indicate "substantial" agreement, and from 0.81 to 1.00 indicate "almost perfect" agreement.

The surgeons and residents additionally completed a survey following completion of their evaluations in an attempt to elucidate perioperative testing protocols dealing with ankle syndesmotic injuries. Surgeons were asked for their current clinical practices while residents were asked what they thought their preferences and protocols would be next year in practice. We chose to not have the students complete the survey as it was our opinion that their level of clinical experience would not be to the point of development of perioperative preferences and protocols.

Further, eye-tracking and gaze recognition software (Gazepoint©, Clemson, South Carolina) was utilized to provide a subjective measure of what specific anatomic structures participants were looking at during testing.

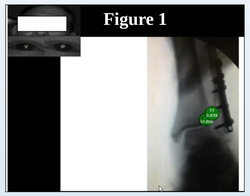
Results

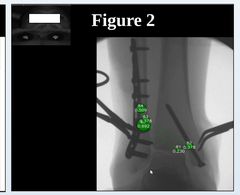
Frequency of agreement with a "positive" result	Still image 1	Still image 2	Still image 3	Video 1	Video 2	Reliability (kappa value)
Surgeons (n=12)	75.0%	75.0%	66.7%	41.7%	25.0%	0.087
Residents (n=12)	91.7%	58.3%	83.3%	58.3%	58.3%	0.019
Students (n=12)	83.3%	50.0%	100.0%	50.0%	25.0%	0.237
Total (n=36)	83.3%	61.1%	83.3%	50.0%	36.1%	0.123

Table 1 (left): We observed an overall level of reliability (as measured with the Kappa coefficient) of 0.123 with respect to the diagnosis of syndesmotic instability. **Surgeons specifically demonstrated a Kappa coefficient of 0.087.** This indicates a **"slight" level of agreement** with respect to the interpretation of this test.

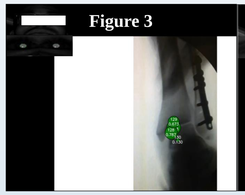
Table 2 (right): We further attempted to objectify eye-tracking results based on the anatomic area of focus. We observed wide variability in the anatomic area of focus with surgeons, residents and students. Examples of this are provided in the following figures.

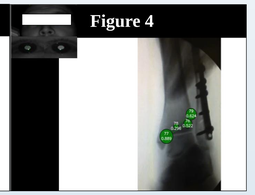
	Exclusively lateral focus	Primarily lateral focus but with glances medially	Equal focus between medial and lateral	Primarily medial focus but with glances laterally	Exclusively medial focus
Surgeons (n=12)	16.7%	33.3%	33.3%	16.7%	0.0%
Residents (n=12)	8.3%	16.7%	25.0%	41.7%	8.3%
Students (n=12)	16.7%	25.0%	58.3%	0.0%	0.0%





Figures 1 and 2: We observed that 2 (16.7%) of the 12 surgeons had "exclusively lateral focus" (an example of which is demonstrated in Figure 1 with near complete attention directed to the tibio-fibular overlap), while 4 (33.3%) of the 12 surgeons had "primarily lateral focus but with glances medially" (an example of which is demonstrated in Figure 2 with primary focus on the fracture line and glances at the medial clear space). Ten (83.3%) of the 12 surgeons reported looking to the tibia-fibula overlap as part of their protocol during testing to assess for syndesmotic stability, with 6 surgeons reporting that they primarily looked to the tibia-fibula overlap while making the determination of syndesmotic stability. The green circles indicate where the subjects were looking, with larger circles indicating consistent focus and gaze.





Figures 3 and 4: We observed that no surgeons had an "exclusively medial focus", while 2 (16.7%) of 12 surgeons had "primarily medial focus but with glances laterally". Figure 3 demonstrates an example of near complete focus on the medial and superior clear spaces. All twelve surgeons (100.0%) reported looking to the medial clear space as part of their protocol during testing to assess for syndesmotic stability, with 3 surgeons reporting that they primarily looked to the medial clear space while making the determination of syndesmotic stability. We observed that 4 (33.3%) of 12 surgeons had "equal focus between lateral and medial" (an example of which is demonstrated in Figure 4).

	Routinely perform stress-dorsiflexion external rotation test to assess for syndesmotic stability (Primary determinant)	Routinely perform Cotton hook test to assess for syndesmotic stability (Primary determinant)	Routinely utilize fracture radiographic characteristics and fracture classification to assess for syndesmotic stability (Primary determinant)	Utilize any other assessment to assess for syndesmotic stability
Surgeons (n=12)	83.3% (50.0%)	58.3% (25.0%)	58.3% (16.7%)	One surgeon (8.3%) reported additionally physically palpating the syndesmosis during testing; One surgeon (8.3%) reported not having a primary determinate of syndesmotic stability and used a combination of tests.
Residents (n=12)	91.7% (66.7%)	50.0% (33.3%)	33.3% (0.0%)	None

	Look to medial clear space in determination (Primary determinant)	Look to tibia-fibula overlap in determination (Primary determinant)	Look to frontal plane rotation of talus in determination (Primary determinant)	Other
Surgeons (n=12)	100.0% (25.0%)	83.3% (50.0%)	8.3% (0.0%)	3 surgeons (25.0%) reported not having a specific anatomic determinant and relied on a combination of findings.
Residents (n=12)	91.7% (66.7%)	66.7% (33.3%)	8.3% (0.0%)	No other findings reported

	Prefer static (screw) fixation	Prefer dynamic (suture) fixation	Prefer combination of static and dynamic fixation	Prefer to always remove syndesmotic fixation	Prefer to never remove syndesmotic fixation	Prefer to sometimes remove syndesmotic fixation based on patient symptoms
Surgeons (n=12)	83.3%	8.3%	8.3%	50.0%	41.6%	8.3%
Residents (n=12)	66.7%	25.0%	8.3%	50.0%	41.6%	8.3%

Tables 3-5: These tables report the perioperative preferences reported by surgeons and residents with respect to syndesmotic testing, syndesmotic stress evaluation, and syndesmotic fixation. The results indicate substantial clinical variation in the testing and fixation of syndesmotic injuries.

Discussion

As with any scientific investigation, critical readers are encouraged to review the study design and results and reach their own conclusions, while the following represents our conclusions based on the specific results. As scientists, we also never consider data to be definitive, but do think that these results are worthy of attention and future investigation:

-First, we conclude the observed levels of agreement were well below what would be expected of a gold standard diagnostic test during stress examination of the ankle syndesmosis. The observed "slight" levels of agreement (as measured by the kappa coefficient) were not much better than would be expected from agreement by chance. **This indicates that the stress examination of the ankle syndesmosis might benefit from the creation of an objective definition and standardized interpretation**. The survey results and eye-tracking/gaze recognition patterns observed in this study might be useful as a starting point in defining such a diagnostic testing protocol.

-Second, the results of the survey indicate variability in clinical practice and teaching with respect to the performance and interpretation of these tests when dealing with the ankle syndesmosis. Although it should certainly not be considered epidemiologic data representing contemporary clinical practice in the US, it does at least show that a relatively wide variety of perioperative preferences and protocols exist.

In conclusion, the results of this investigation provide evidence of reliability well below what would be expected of a gold standard during stress examination of the ankle syndesmosis. These results indicate that future research is required in order to standardize the performance and interpretation of this test. It is our hope that the survey questions and eye-tracking results provided herein might be utilized to do so.

References

[1] McConnell T, Creevy W, Tornetta P 3rd. Stress examination of supination external rotation-type fibular fractures. J Bone Joint Surg Am. 86-A(10): 2171-8, 2004.
[2] 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. 86(11): 2393-8, 2004.
[3] Pakarinen H, Flinkkila T, Ohtonen P, Hyvonen P, Lakovaara M, Leppilahti J, Ristiniemi J. Intraoperative assessment of the stability of the distal tibiofibular joint in supination-external rotation injuries of the ankle: sensitivity, specificity, and reliability of two clinical tests. J Bone Joint Surg Am. 93(22): 2057-61, 2011.
[4] Stoffel K, Wysocki D, Baddour E, Nicholls R, Yates P. Comparison of two intraoperative assessment methods for injuries to the ankle syndesmosis. A cadaveric study. J Bone Joint Surg Am. 91(11): 2646-52, 2009.
[5] Kortekangas TH, Pakarinen HJ, Savola O, Niinimake J, Lepojarvi S, Ohtonen P, Flinkkila T, Ristiniemi J. Syndesmotic fixation in supination-external rotation ankle fractures: a prospective randomized study. Foot Ankle Int. 35(10): 988-95, 2014.
[6] Sman AD, Hiller CE, Rae K, Linklater J, Black DA, Nicholson LL, Burns J, Refshauge KM. Diagnostic accuracy of clinical tests for ankle syndesmosis injury. Br J Sports Med. 49(5): 323-9, 2015.
[7] Beumer A, van Hemert WL, Neising R, Entius CA, Ginai AZ, Mulder PG, Swierstra BA. Radiographic measurement of the distal tibiofibular syndesmosis has limited use. Clin Orthop Relat Res. (423): 227-34, 2004.
[8] Gardner MJ, Demetrakopoulos D, Briggs SM, Helfet DL, Lorich DG. The ability of the Lauge-Hansen classification to predict ligament injury and mechanism in ankle fractures: an MRI study. J Orthop Trauma. 20(4): 267-72, 2006.