# Consistency and Reliability of Ankle Stress Radiography in Patients With Chronic Lateral Ankle Instability

Ji Hye Choi,* MD, Kug Jin Choi,* MD, Chin Youb Chung,* MD, PhD, Moon Seok Park,* MD, PhD, Ki Hyuk Sung,* MD, and Kyoung Min Lee,*† MD, PhD

*Investigation performed at Seoul National University Bundang Hospital, Gyeonggi, Republic of Korea*

**Background:** Ankle stress radiographs are important tools for evaluating chronic lateral ankle instability. The consistency of a patient's ankle condition as it affects the reliability of ankle stress radiographs has never been evaluated.

**Purpose:** To investigate the consistency and reliability of ankle stress radiographs in patients with chronic lateral ankle instability without an ankle injury during the study period.

**Study Design:** Cohort study (diagnosis); Level of evidence, 3.

**Methods:** Included were patients with chronic lateral ankle instability who underwent 2 repeated ankle stress radiographs between January 2014 and July 2019; those with an ankle injury during the study period were excluded. The tibiotalar tilt angle on varus stress radiographs and anterior translation of the talus on anterior drawer stress radiographs were measured at initial presentation and final follow-up examination. Interobserver reliability and consistency of ankle stress radiographs were analyzed using the intraclass correlation coefficient (ICC).

**Results:** A total of 45 patients (mean ± standard deviation age, 36.4 ± 13.4 years; 18 men and 27 women; follow-up duration, 9.1 ± 3.2 months) were included. The mean ± standard deviation tibiotalar tilt angle and anterior talar translation at initial presentation were 10.8° ± 5.2° and 6.9 ± 2.7 mm, respectively. The interobserver reliabilities of the tibiotalar tilt angle and anterior talar translation were excellent (ICC = 0.926 [95% CI, 0.874-0.959] and 0.911 [95% CI, 0.766-0.961], respectively). The consistency between the initial and final radiographs was good for tibiotalar tilt angle (ICC = 0.763 [95% CI, 0.607-0.862]) and poor for anterior talar translation (ICC = 0.456 [95% CI, 0.187-0.660]).

**Conclusion:** Although the interobserver reliability of the radiographic measurements was excellent, the consistency of the ankle stress radiographs was not as acceptable. Surgeons need to be cautious when deciding whether to operate on a patient with chronic lateral ankle instability based on a single ankle stress radiograph.

**Keywords:** consistency; reliability; ankle stress radiograph; chronic lateral ankle instability

Acute ankle sprains are among the most common sports injuries; inappropriate healing of injured lateral ankle ligaments may lead to chronic lateral ankle instability.[4,17] Functional and mechanical factors may be involved in chronic ankle instability. Functional ankle instability is caused by proprioceptive deficits, peroneal muscle weakness, and delayed peroneal reaction time,[12] and it is usually treated nonoperatively.

Mechanical ankle instability is caused by defective ligamentous or bony structures,[3] and the current focus of treatment for severe mechanical instability is surgery such as repair and reconstruction of the defective ankle ligaments.[2,13,15] Defining mechanical ankle instability and suggesting guidelines for surgical indication are matters of ongoing debate; one reason for this is the insufficient evidence of the consistency and reliability of lateral ankle instability.

Ankle stress radiographs, including those in the varus stress and anterior drawer views, are important tools for evaluating mechanical lateral ankle instability.[1,8,9,16] The overall reliability of ankle stress radiographs relies on consistent ankle conditions and reliable radiographic measurements obtained by clinicians. A previous study[11] reported satisfactory interobserver reliability of ankle stress radiographs. However, to the best of our knowledge, the consistency of patients' ankle conditions has never been evaluated.

This retrospective study aimed to investigate the clinical usefulness of repeated ankle stress radiographs performed

The Orthopaedic Journal of Sports Medicine, 9(5), 23259671211004099
DOI: 10.1177/23259671211004099

in patients without further ankle injuries during follow-up and to determine the consistency and reliability of these radiographs. We hypothesized that the consistency of ankle stress radiography would not be as satisfactory as its reliability.

## METHODS

### Patients

This retrospective study was approved by the institutional review board of our hospital, a tertiary referral center for foot and ankle diseases; the need for informed consent from the participants was waived. It was not appropriate or possible to involve patients or the public in the design, conduct, reporting, or dissemination plans of our research.

The study included patients with symptoms of ankle instability after ankle inversion injury who underwent 2 repeated ankle stress radiographs between January 2014 and July 2019. Exclusion criteria were age <20 years; original sprain <6 weeks from the most recent sprain; ankle ligament repair that took place between the 2 ankle stress radiographs; ankle osteoarthritis showing joint space narrowing on weightbearing anteroposterior or lateral radiographs; concomitant osteochondral lesion of the talus; an ankle injury that occurred between the 2 ankle stress radiographs; diagnosis of neuromuscular diseases; history of fracture, infection, tumor, or any other conditions that could change the normal bony structures of the foot and ankle; and inadequate ankle stress radiographs. During the follow-up period, the patients were prescribed nonoperative treatment, which comprised peroneal strengthening exercises and proprioceptive exercises at nearby hospitals. Patient compliance with this protocol was not recorded.

Initially, 90 patients who underwent repeated ankle stress radiographic examination between January 2014 and July 2019 were considered for the study. After the inclusion and exclusion criteria were applied, 45 ankles from 45 patients were included in the final analysis (Figure 1).

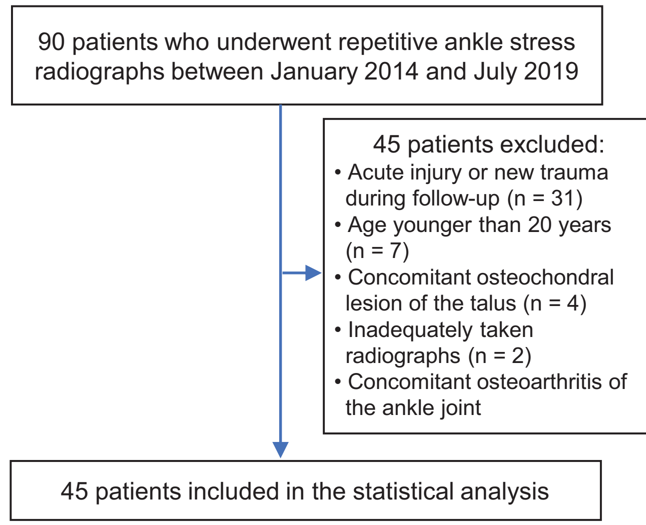


**Figure 1.** Flowchart of patient selection.

### Ankle Stress Radiographs

Ankle radiographs were taken using a standardized technique at a source-to-image distance of 110 cm, a peak potential of 60 kVp, and an amperage of 6.3 mA using a UT-2000 radiographic device (Philips Research). Varus stress radiographs were taken with the patient in the supine position with the knee in 20° of flexion, the tibia in 20° of internal rotation, and the ankle in 10° of plantarflexion. Moreover, 15 daN of varus force was applied to the ankle using a Telometer device (Daiseung Medics), and the x-ray beam was directed in the anterior-to-posterior direction. Anterior drawer stress radiographs were taken with the knee in 20° of flexion and the ankle in 10° of plantarflexion, and a posteriorly directed translating force of 15 daN was applied to the tibia. The x-ray beam was directed in the medial-to-lateral direction, and the lateral view was obtained.

All radiographs were obtained bilaterally, and the images were collected and processed digitally using a picture archiving and communication system (Impax; Agfa). Radiographic measurements were also obtained using the system's software.

### Consensus Building and Reliability of the Radiographic Measurements

A consensus-building session was held among 4 orthopaedic surgeons (K.M.L., K.H.S., K.J.C., and J.H.C.) with 17, 15, 7, and 3 years of experience. On varus ankle stress anteroposterior radiographs, the tibiotalar tilt angle (ie, the angle between the articular surface of the distal tibia [tibial plafond] and talar dome) was measured (Figure 2).[9] The line representing the articular surface was usually the connecting line (tangent line) of the most proximal point of the medial and lateral sides of the tibial plafond because the distal tibial articular surface is distally convex. The same

†Address correspondence to Kyoung Min Lee, MD, PhD, Department of Orthopaedic Surgery, Seoul National University Bundang Hospital, 300 Gumi-dong, Bundang-gu, Seongnam, Gyeonggi, Republic of Korea (email: oasis100@empal.com).

*Department of Orthopaedic Surgery, Seoul National University Bundang Hospital, Gyeonggi, Republic of Korea.

J.H.C. and K.J.C. contributed equally to the writing of this article.

Final revision submitted November 10, 2020; accepted December 29, 2020.

One or more of the authors has declared the following potential conflict of interest or source of funding: This study was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science & ICT (NRF-2017M3A9D8064200) and Seoul National University Bundang Hospital Research Fund (grant No. 13-2018-019). AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from Seoul National University Bundang Hospital.


**Figure 2.** The tibiotalar tilt angle (asterisk) is the angle between a line representing the articular surface of the distal tibia (tibial plafond) and another line representing the articular surface of the talar dome.

**Figure 3.** The anterior talar translation (yellow line) is the distance between the posterior lip of the tibial articular surface and the nearest point of the articular surface of the talar dome.

TABLE 1
Summary of Radiographic Measurements[a]

| | Initial Examination | Final Examination | *P* Value[b] |
|---|---|---|---|
| Index side | | | |
| Tibiotalar tilt angle, deg | 10.8 ± 5.2 | 9.9 ± 4.9 | .096 |
| Anterior talar translation, mm | 6.9 ± 2.7 | 7.0 ± 1.9 | .889 |
| Contralateral side | | | |
| Tibiotalar tilt angle, deg | 6.7 ± 4.4 | 6.0 ± 3.9 | .073 |
| Anterior talar translation, mm | 6.2 ± 2.3 | 5.7 ± 2.0 | .074 |
| Δ index and contralateral sides | | | |
| Tibiotalar tilt angle, deg | 4.1 ± 5.1 | 3.9 ± 3.8 | .679 |
| Anterior talar translation, mm | 0.7 ± 2.1 | 1.3 ± 2.0 | .062 |

[a]Data are expressed as mean ± SD.
[b]Paired *t* test.

reasoning was applied for the articular surface of the talar dome.

On anterior drawer radiographs, anterior talar translation was measured as the distance between the posterior lip of the distal tibial articular surface and the nearest point of the articular surface of the talar dome (Figure 3).[9]

To assess the reliability of the measurements, 3 orthopaedic surgeons (K.M.L., K.J.C., and J.H.C., with 17, 7, and 3 years of experience, respectively) measured the tibiotalar tilt angle and anterior talar translation on 36 images of the index side. All surgeons measured the radiographs independently on the same day without knowing the other surgeons' measurements, and the images were presented in random order by a research assistant who was not involved in this study. After the reliability testing, 1 of the 3 surgeons (K.J.C.) measured the radiographic parameters for all 45 study patients.

The radiographic criteria for chronic lateral ankle instability surgery were >15° or a side-to-side difference >10° of tibiotalar tilt angle on varus stress radiographs and >10 mm or a side-to-side difference of >3 mm of anterior talar translation on anterior drawer stress radiographs.[5]

## Statistical Analysis

Descriptive analysis was performed; data are presented as mean ± standard deviation (SD) and percentages. The Kolmogorov-Smirnov test was conducted to determine data normality, and a paired *t* test was used to compare means between the initial and final radiographic parameters. In addition, the Pearson correlation coefficient was used to analyze the correlation between variables. Pearson correlation coefficients were interpreted as: negligible correlation, 0-0.29 (0 to –0.29); low correlation, 0.3-0.49 (–0.3 to –0.49); moderate correlation, 0.5-0.69 (–0.5 to –0.69); high correlation, 0.7-0.89 (–0.7 to –0.89); very high correlation, 0.9-1.0 (–0.9 to –1.0).

Interobserver reliability was evaluated using the ICC with a 2-way random effect model, assuming a single measurement and absolute agreement. ICC values were interpreted as: poor agreement, <0.5; moderate agreement, 0.5-0.74; good agreement, 0.75-0.9; excellent agreement, >0.9. With an ICC target value of 0.8 and setting 0.2 as the width of the 95% confidence interval, the Bonett approximation gave a minimum sample size of 36 for reliability testing.[14] The consistency of repeated ankle stress radiographs was analyzed using ICCs in patients without additional ankle injuries. Agreement on surgical indication was assessed using the kappa statistic. Kappa values were interpreted as: slight, 0-0.20; fair agreement, 0.21-0.40; moderate agreement, 0.41-0.60; substantial agreement, 0.61-0.80; almost perfect agreement, 0.81-1.

TABLE 2
Interobserver Reliability and Consistency of Ankle Stress Radiographic Measurements

| | ICC (95% CI)[a] | *P* Value |
|---|---|---|
| Interobserver reliability | | |
| Tibiotalar tilt angle, deg | 0.926 (0.874-0.959) | <.001 |
| Anterior talar translation, mm | 0.911 (0.766-0.961) | <.001 |
| Consistency between initial and final examination | | |
| Index side | | |
| Tibiotalar tilt angle, deg | 0.763 (0.607-0.862) | <.001 |
| Anterior talar translation, mm | 0.456 (0.187-0.660) | .001 |
| Δ index and contralateral sides | | |
| Tibiotalar tilt angle, deg | 0.721 (0.544-0.837) | <.001 |
| Anterior talar translation, mm | 0.484 (0.231-0.677) | <.001 |

[a]Intraclass correlation coefficient (ICC) values: <0.5, poor agreement; 0.5-0.74, moderate agreement; 0.75-0.9, good agreement; >0.9, excellent agreement.

All statistical analyses were performed using SPSS Version 20.0 (IBM Corp). Statistical significance was set at $P < .05$.

## RESULTS

### Patient Characteristics

The mean ± SD age of the 45 study patients at initial presentation was 36.4 ± 13.4 years, and there were 18 men and 27 women. The mean ± SD interval between initial presentation and the ankle stress radiographs showing no ankle injury (ie, follow-up duration) was 9.1 ± 3.2 months, and the affected side included 18 right and 27 left ankles.

### Summary of Radiographic Measurements

Table 1 shows a comparison of the radiographic measurements on ankle stress radiographs at initial examination and final follow-up examination. No statistically significant changes were seen in the tibiotalar tilt angle or anterior talar translation between these periods regarding the index side, the contralateral side, or the difference between index and contralateral sides.

### Interobserver Reliability and Consistency of Ankle Stress Radiographs

The ICCs for interobserver reliability of the radiographic parameter measurements among the 3 orthopaedic surgeons indicated excellent agreement (ICCs of 0.926 and 0.911). On comparison of the consistency between the initial and final radiographs of the tibiotalar tilt angle and anterior talar translation, the ICCs for the index side indicated good agreement (ICC = 0.763) and poor agreement (ICC = 0.456), respectively. The consistency of the difference between the index and contralateral sides indicated moderate agreement for tibiotalar tile angle (ICC = 0.721) and poor agreement for anterior talar translation (ICC = 0.484) between the initial and final radiographs (Table 2).

### Correlation Between the Radiographic Measurements

The Pearson correlation coefficients for tibiotalar tilt angle and anterior talar translation were negligible on the index side: $r = 0.201$ ($P = .185$) at initial examination and $r = 0.070$ ($P = .648$) at final examination. The correlation coefficients for the difference in the tibiotalar tilt angle and anterior talar translation were negligible for the index side ($r = –0.203$; $P = .180$) and low for the contralateral side ($r = 0.344$; $P = .021$).

### Agreement on Surgical Indication Based on Repeated Ankle Stress Radiographs

Overall, 22.2% (10/45) and 15.6% (7/45) of patients were indicated for surgical treatment based on the initial and final varus stress radiographs, respectively, whereas 22.2% (10/45) and 20% (9/45) of patients were indicated for surgical treatment based on the initial and final anterior drawer stress radiographs, respectively. The agreement regarding indications for surgery between the initial and final radiographs was fair: 80% ($\kappa = 0.352$; $P = .016$) for varus stress radiographs and 80% ($\kappa = 0.400$; $P = .007$) for anterior drawer stress radiographs.

## DISCUSSION

Ankle stress radiographs are a useful and important tool for evaluating the degree of chronic ankle instability, and some guidelines are available for surgical indication in terms of ankle stress radiographic findings.[2,5,13] However, the clinical judgment for chronic ankle instability based on ankle stress radiographs requires the images to be consistent and reliable. In our study, interobserver reliability of the radiographic measurements was satisfactory, but the consistency of the ankle stress radiographs needs to be interpreted cautiously.

### Clinical Utility of Ankle Stress Radiographs

Although magnetic resonance imaging (MRI) scan can be used to precisely visualize the structural integrity of the ankle joint, including the lateral ankle ligaments and concomitant secondary lesions from chronic ankle instability, MRI scans cannot reflect the functions of the affected structures. The role of the ligaments is to maintain joint stability, which can be evaluated using stress testing. Previous studies have reported the clinical usefulness of ankle stress radiography[2,7-9] and suggested that ankle stress radiographs are more valuable than are MRI scans in examining the mechanical component of chronic lateral ankle instability.[6,10] Therefore, ankle stress radiographs, including those in the varus stress and anterior drawer views, are widely used to evaluate the degree of lateral ankle instability and to define surgical indication, although the criteria for surgical indication or cutoff values are variable.[2,5,13]

### Interobserver Reliability and Consistency of Ankle Stress Radiographs

Lohrer et al[11] reported values for the reliability of ankle stress radiographs that were comparable to our results, and this modality is considered relevant for clinical use. The study suggested that further studies should address the consistency and accuracy of obtaining multiple ankle stress radiographs of the same patients to improve the clinical relevance of ankle stress radiographs. The study found that the mean ± SD difference between tibiotalar tilt angles on repeated varus ankle stress radiographs was 2.6° ± 2.5°, but differences ranged from 0° to 11.7° in patients with a similar ankle condition. The mean ± SD difference between the repeated anterior drawer tests was 1.9 ± 1.5 mm, and the differences ranged from 0.1 to 7.4 mm. Therefore, the individual difference between the repeated ankle stress radiographic examination could be quite variable although the mean difference is small. Although ankle conditions for the same patient can vary greatly between 2 stress radiographs at different time points, the condition itself is difficult to measure using current examination tools. Presumably, consistency of ankle stress radiographs could be affected by muscle guarding due to patient discomfort during the stress test and different soft tissue tension due to changes in ankle effusion or errors in stress testing apparatus and the angle of the x-ray beam, but these factors have not been well documented. Further research on this topic is required.

### Agreement on Surgical Indication Based on Repeated Ankle Stress Radiographs

We found fair agreement regarding surgical indication between the initial and final varus ankle stress radiographs as well as between the initial and final anterior drawer tests. Furthermore, the correlation between the tibiotalar tilt angle and anterior talar translation was not significant. Therefore, surgeons need to be cautious when deciding to operate on a patient with chronic lateral ankle instability based on a single ankle stress radiographic examination. A previous study[16] reported that ankle stress radiography presented high specificity (up to 100%) but low sensitivity (57%), suggesting that the test has a diagnostic value only when it shows positive results. Although ankle stress radiography is a useful dynamic test for detecting the mechanical component of lateral ankle instability, the low sensitivity of this examination could be problematic. We believe that the low sensitivity of ankle stress radiography might be partly attributable to the unsatisfactory consistency of the examination, as was shown in the present study. Therefore, repeated ankle stress radiographic examinations are recommended to determine the mechanical lateral ankle instability requiring surgical treatment.

### Limitations

Some limitations of our study need to be addressed. First, this study had a retrospective design, and selection bias might have affected the study results. More specifically, the assumption that ankles without an injury during the follow-up period represented the same ankle condition was not validated. Second, patients who underwent surgery or sustained additional inversion injuries during the follow-up were excluded from the study. Therefore, patients with more severe ankle instability might have been excluded from the data analysis, so the study results cannot be generalized to all patients with chronic lateral ankle instability. Third, although the ankle stress radiographs were taken using a standardized protocol, they were obtained by multiple technicians, which might have increased variability. Furthermore, errors in stress testing apparatus and the angle of x-ray beam could have contributed to the increased variability of ankle stress radiographic examination. However, this examination setting reflected everyday orthopaedic foot and ankle clinical practice.

## CONCLUSION

Ideally, repeated ankle stress radiographic measurements including the varus tibiotalar tilt angle and anterior talar translation would be equal under identical ankle ligament conditions. This study showed that ankle stress radiographic measurements can vary among patients with presumably identical ankle ligament conditions. Therefore, surgeons must be cautious when deciding whether to operate on a patient with chronic lateral ankle instability based on a single ankle stress radiograph, and repeated examinations might be required along with comprehensive physical examination to determine the mechanical ankle instability indicating the need for surgical treatment for patients with symptoms.

## ACKNOWLEDGMENT

The authors thank Jung Hee Yoon, BS, for her technical support.

## REFERENCES

1. Ahovuo J, Kaartinen E, Slätis P. Diagnostic value of stress radiography in lesions of the lateral ligaments of the ankle. *Acta Radiol*. 1988; 29(6):711-714.
2. Baumhauer JF, O'Brien T. Surgical considerations in the treatment of ankle instability. *J Athl Train*. 2002;37(4):458-462.
3. Bonnel F, Toullec E, Mabit C, Tourné Y; Sofcot. Chronic ankle instability: biomechanics and pathomechanics of ligaments injury and associated lesions. *Orthop Traumatol Surg Res*. 2010;96(4):424-432.
4. Bridgman SA, Clement D, Downing A, Walley G, Phair I, Maffulli N. Population based epidemiology of ankle sprains attending accident and emergency units in the West Midlands of England, and a survey of UK practice for severe ankle sprains. *Emerg Med J*. 2003;20:508-510.
5. Coughlin MJ, Saltzman CL, Anderson RB. *Mann's Surgery of the Foot and Ankle*. 9th ed. Elsevier; 2014.

6. Danna NR, Shakked RJ, Sheskier SC. Ankle stress radiographs predict lateral ankle instability better than MRI. Paper presented at: 2016 Annual Meeting of American Orthopedic Foot & Ankle Society; October 6, 2016; Toronto, ON, Canada. https://doi.org/10.1177/2473011416S00127
7. De Aguiar TO, Oliboni LS, Dezotti VM, Kennedy NI, Ferrari MB, Ellera Gomes JL. Simultaneous radiographic technique to evaluate ankle instability. *Arthrosc Tech*. 2017;6(6):e2187-e2190.
8. Dowling LB, Giakoumis M, Ryan JD. Narrowing the normal range for lateral ankle ligament stability with stress radiography. *J Foot Ankle Surg*. 2014;53(3):269-273.
9. Grace DL. Lateral ankle ligament injuries: inversion and anterior stress radiography. *Clin Orthop Relat Res*. 1984;183:153-159.
10. Kiely PD, Borton DB. The evaluation of chronic ankle instability: MRI or stress X-rays? *Orthop Proc*. 2018;86-B(suppl II).
11. Lohrer H, Nauck T, Arentz S, Schöll J. Observer reliability in ankle and calcaneocuboid stress radiography. *Am J Sports Med*. 2008;36(6):1143-1149.
12. Santos MJ, Liu W. Possible factors related to functional ankle instability. *J Orthop Sports Phys Ther*. 2008;38(3):150-157.
13. Shakked RJ, Karnovsky S, Drakos MC. Operative treatment of lateral ligament instability. *Curr Rev Musculoskelet Med*. 2017;10(1):113-121.
14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420-428.
15. So E, Preston N, Holmes T. Intermediate- to long-term longevity and incidence of revision of the modified Broström-Gould procedure for lateral ankle ligament repair: a systematic review. *J Foot Ankle Surg*. 2017;56(5):1076-1080.
16. Tourné Y, Besse JL, Mabit C; Sofcot. Chronic ankle instability: which tests to assess the lesions? Which therapeutic options? *Orthop Traumatol Surg Res*. 2010;96(4):433-446.
17. van Rijn RM, van Os AG, Bernsen RM, Luijsterburg PA, Koes BW, Bierma-Zeinstra SM. What is the clinical course of acute ankle sprains? A systematic literature review. *Am J Med*. 2008;121:324-331.