



Contents lists available at ScienceDirect

The Journal of Foot & Ankle Surgery

journal homepage: www.jfas.org

Ankle Fracture Classification: An Innovative System for Describing Ankle Fractures



Jan Paul Briet, MD, PhD¹, Falco Hietbrink, MD, PhD², Diederik P. Smeeing, MD³,
Marcel G.W. Dijkgraaf, PhD⁴, Egbert Jan Verleisdonk, MD, PhD⁵, R. Marijn Houwert, MD, PhD^{2,6}

¹ Surgical Resident, Department of Surgery, Diaconessenhuis Utrecht, Utrecht, the Netherlands

² Trauma Surgeon, Department of Surgery, University Medical Center Utrecht, Utrecht, the Netherlands

³ Surgical Resident, Utrecht Traumacenter, Utrecht, the Netherlands

⁴ Statistician, Clinical Research Unit, Academic Medical Center Amsterdam, Amsterdam, the Netherlands

⁵ Trauma Surgeon, Department of Surgery, Diaconessenhuis Utrecht, Utrecht, the Netherlands

⁶ Trauma Surgeon, Utrecht Traumacenter, Utrecht, the Netherlands

ARTICLE INFO

Level of Clinical Evidence: 5

Keywords:

ankle fracture
classification
interobserver variation
intraobserver variation
Lauge-Hansen
malleolar fractures
radiologic classification

ABSTRACT

A good classification system is important for clinical handoffs, research, and clinical treatment guidelines. A reliable classification system shows good interobserver and intraobserver agreement. This study analyzed the interobserver and intraobserver agreement of a descriptive system for ankle fractures and the Lauge-Hansen classification. Three groups of observers (experts, semiexperts, and novices) scored a total of 20 ankle radiographs. All ankle radiographs were classified according to the Lauge-Hansen and Danis-Weber classifications. The ankle fractures were subsequently reviewed in a descriptive manner for the following features: number of affected malleoli, type of fracture of the lateral and medial malleolus, and congruence of the ankle joint. After 2 weeks, the same set of radiographs were reviewed. For interobserver and intraobserver variability, the separate groups were used for analysis, and the Fleiss (multirater) κ values were calculated. The interobserver agreement on the Lauge-Hansen classification was moderate for the experts, fair for semiexperts, and slight for novices ($\kappa = 0.45$, $\kappa = 0.37$, and $\kappa = 0.16$). All factors of the descriptive system had better interobserver agreement than the Lauge-Hansen classification, except for the agreement on the type of fracture of the lateral malleolus. The intraobserver agreement of the Lauge-Hansen classification was substantial for the experts, moderate for the semiexperts, and fair for the novice observers ($\kappa = 0.70$, $\kappa = 0.49$, and $\kappa = 0.26$). The intraobserver agreement was better for all factors of the descriptive system compared with the Lauge-Hansen classification. The descriptive system presented in this study shows less variability between observers than the Lauge-Hansen classification. This system has clinical implications and is easy to use for clinicians with mixed levels of experience. It has the potential to improve clinical and research handoffs and overcome the limitations of current classification systems.

© 2018 by the American College of Foot and Ankle Surgeons. All rights reserved.

An ankle fracture is a frequently encountered injury that can be treated by the most junior physicians in the emergency room or by the most experienced (orthopedic) trauma surgeons. Fracture classification is an important tool for clinical and research purposes. A classification system is of great importance for clinical handoffs and should provide treatment and prognosis guidance for the patient and surgeon (1).

Ideally, a system for describing an injury is simple and has good interobserver and intraobserver agreement (2). A proper understanding of the mechanism and extent of the injury, using a comprehensible classification, may help surgeons choose the proper treatment. In addition

to describing the injured structures, a useful descriptive tool should predict future ankle function. Many classification systems have been developed over the past years. The purely descriptive Danis-Weber and Association for Osteosynthesis/Association for the Study of Internal Fixation (AO/ASIF) systems were developed in response to the more difficult pathologic-anatomic Lauge-Hansen classification (3–5). The Danis-Weber and AO/ASIF classifications are purely radiographic classifications and are easy for daily and clinical use. In the Danis-Weber classification, the fractures are categorized as infratransyndesmotic or suprasyndesmotic. Additionally, the AO/ASIF classification provides information on medial and posterior malleolar involvement. Therefore, the AO/ASIF classification provides valuable information to facilitate communication between physicians; however, it lacks essential information such as the extent of ligament injury. In contrast, the Lauge-Hansen classification emphasizes the different stages of

Financial Disclosure: None reported.

Conflict of Interest: None reported.

Address correspondence to: Jan Paul Briet, MD, PhD, Department of Surgery, Diaconessenhuis Utrecht, Bosboomstraat 1, 3582 KE Utrecht, the Netherlands.

E-mail address: jp.briet@gmail.com (J.P. Briet).

ligamentous injury in addition to the fracture pattern, and it provides options for fracture treatment (3). The Lauge-Hansen classification is based on controlled cadaveric biomechanical studies under physiologic circumstances (3,6). As a result, approximately 5% of the fractures cannot be classified according to the Lauge-Hansen classification in *in vivo* studies models (7–11). Moreover, the extent of the predicted soft tissue and ligament injuries cannot always be reproduced in magnetic resonance evaluation in patients with ankle fractures (9). Furthermore, several prior studies have demonstrated poor interobserver and intraobserver reliability of the Lauge-Hansen classification (1,2,12–15). Despite these shortcomings, the Lauge-Hansen classification remains as the basis of our understanding of the pathomechanics of ankle fractures (16). Experience with the Lauge-Hansen classification is known to reduce individuals' observed variation in reproducibility and reliability (14).

This study aimed to analyze the interobserver variability and intraobserver reproducibility of a novel descriptive system for ankle fractures, which describes both fracture pattern and ligament injury. Additionally, the variability and reproducibility in groups with different levels of experience in the assessment of ankle radiographs were investigated.

Material and Methods

This study was conducted in accordance with the Declaration of Helsinki. All radiographs that were used for this study were anonymized. Informed consent was not deemed necessary for this study, as only clinicians were included.

Three groups of observers were formed. The first group (experts) consisted of 5 (orthopedic) trauma surgeons. The second group (semiexperts) consisted of 10 (orthopedic) surgical residents with some experience in ankle fracture classification. The third group (novices) consisted of 16 medical students/interns without practical experience in ankle fracture classification systems. The last group attended a prestudy lecture on ankle classification systems by the first author (J.B.). All observers were independently asked to classify and describe 20 anonymized radiographs of patients with an ankle fracture.

First, all ankle radiographs were classified and reviewed according to the Lauge-Hansen and Danis-Weber classifications. Second, all ankle fractures were reviewed for the following features in a descriptive manner: number of affected malleoli, type of fracture of the lateral and medial malleolus, and congruence of the ankle joint.

Two weeks later, all participants were asked to classify the same radiographs again.

Lauge-Hansen Classification

The Lauge-Hansen classification is based on the mechanism of injury, and more than 95% of all ankle fractures fit the Lauge-Hansen classification scheme (3,6,10). The Lauge-Hansen classification consists of a 2-criteria descriptor: the position of the foot (supination or pronation) at the time of the traumatic event and the directional force of the ankle (abduction, adduction, or external rotation) (3,6). The 4 main categories of the Lauge-Hansen classification are: supination adduction, supination external rotation, pronation abduction, and pronation external rotation, which can be further subdivided into 13 subgroups (Appendix) (17).

Descriptive System

A simple, descriptive system was used that describes the location of the fracture of the lateral malleolus according to the Weber classification (A, B, C, or no fracture), the number of involved malleoli (1, 2, or 3), and the fracture pattern of the lateral malleolus (comminuted, oblique, or no fracture) and medial malleolus (horizontal, oblique, vertical, or no fracture). Ultimately, it describes the congruity of the ankle joint (yes or no) (Fig. 1).

Statistical Analysis

Univariate statistical methods, based on absolute and relative frequencies and contingency tables, were applied in order to present the results from the classifications on the radiographs. Single random imputation from a uniform distribution of values in the theoretical range was used for each missing value (observation 1: $n = 40$, observation 2: $n = 41$) (18). The Fleiss generalized κ for nominal data for any number of raters and categories was used to evaluate the interobserver agreement for the complete set of features with 95% confidence limits based on bias corrected and accelerated bootstrapping (19). The mean of the Fleiss κ values for nominal data for individual observers within each group of observer was used to evaluate the intraobserver agreements, with 95% confidence intervals calculated following the Fisher z -transformation and back-transformation. An alternative approach with weighted instead of mean κ values for the intraobserver agreement seemingly overestimated the results in these data sets due to the disproportionate impact of the arbitrary capping of κ values with a value of 1 (20).

The κ values were interpreted according to the method of Landis and Koch (Table 1) (21). For interobserver and intraobserver agreement, the separate groups (expert, semiexpert, and novice) were used for analysis, and the κ per group was provided. Whether 2 κ values differed significantly was qualitatively (yes/no) assessed by observing whether the κ point estimate for a group fell outside the 95% confidence intervals of the κ for the comparator group, and vice versa. Data analysis was performed with SPSS version 23 (IBM Corp., Armonk, NY) for Windows.

Results

All observers (5 experts, 10 semiexperts, and 16 novices; Table 2) scored a total of 20 ankle radiographs. The interobserver agreement for Lauge-Hansen classification was moderate for the experts (0.45 [0.33 to 0.63]) and fair for the semiexperts (0.37 [0.27 to 0.51]), which was significantly higher than slight agreement for the novices (0.16 [0.09 to 0.34]). Interobserver agreement for the Weber classification was substantial for the experts (0.71 [0.51 to 0.89]) and semiexperts (0.64 [0.44 to 0.84]), with the former being significantly higher than the moderate agreement (0.46 [0.31 to 0.63]) for novices. Interobserver agreement for the type of fracture of the medial malleolus was substantial for both the experts (0.79 [0.63 to 0.92]) and semiexperts (0.74 [0.63 to 0.87]) but only moderate (0.44 [0.34 to 0.56]) for novices, which was significantly lower (Fig. 2). For the observations in all 3 groups, the descriptive system had a higher interobserver agreement than the Lauge-Hansen classification except for the type of fracture of the lateral malleolus in novice observers (0.12 [0.01 to 0.33]).

Intraobserver agreement for Lauge-Hansen was substantial for the experts (0.70 [0.57 to 0.79]) and higher than the moderate agreement for semiexperts (0.49 [0.36 to 0.59]), which, in turn, was higher than the fair agreement for the novices (0.26 [0.14 to 0.36]). The intraobserver agreement for the Weber classification was substantial for the experts (0.80 [0.72 to 0.87]) and semiexperts (0.74 [0.66 to 0.80]), and both were significantly higher than the moderate agreement for novice observers (0.54 [0.45 to 0.62]). The intraobserver agreement for the congruence of the ankle joint was significantly higher for experts (0.81 [0.72 to 0.87]) and semiexperts (0.76 [0.69 to 0.81]) than for the novices (0.42 [0.31 to 0.51]) (Fig. 3). The intraobserver agreement of all the aspects of the descriptive system was higher than that of the Lauge-Hansen classification.

Discussion

Classification of ankle fractures on plain radiographs is a challenge. A good classification system should be simple, reliable, and reproducible and provide clinical implications for treatment and prognosis. A comprehensive classification is not useful if clinicians cannot reach agreement on classifying these fractures. This study analyzed the interobserver and intraobserver agreement of 2 separate classification systems for ankle fractures: the Lauge-Hansen classification and a simple descriptive system that includes the Weber classification. This simple, descriptive system has better reproducibility and reliability than the Lauge-Hansen classification for clinicians with mixed levels of experience.

The AO/ASIF classification for ankle fractures is easier to understand and has higher interobserver and intraobserver reliability compared with the Lauge-Hansen classification (1). The AO/ASIF classification provides essential information on the type and location of the fracture of the lateral medial and posterior malleolus and the relation. Thus, it describes the relation of the fibular fracture and the syndesmosis, providing an indication of syndesmotic injury. However, the AO/ASIF classification is a purely anatomic-radiographic classification and lacks critical information to make it holistic for diagnosis, treatment, and prognosis. For instance, it does not offer appropriate information for the management of malleolar fractures, because some fractures of one class require surgical fixation whereas others have a good prognosis without surgery (1,7,22). This is mainly because the AO/ASIF classification is a

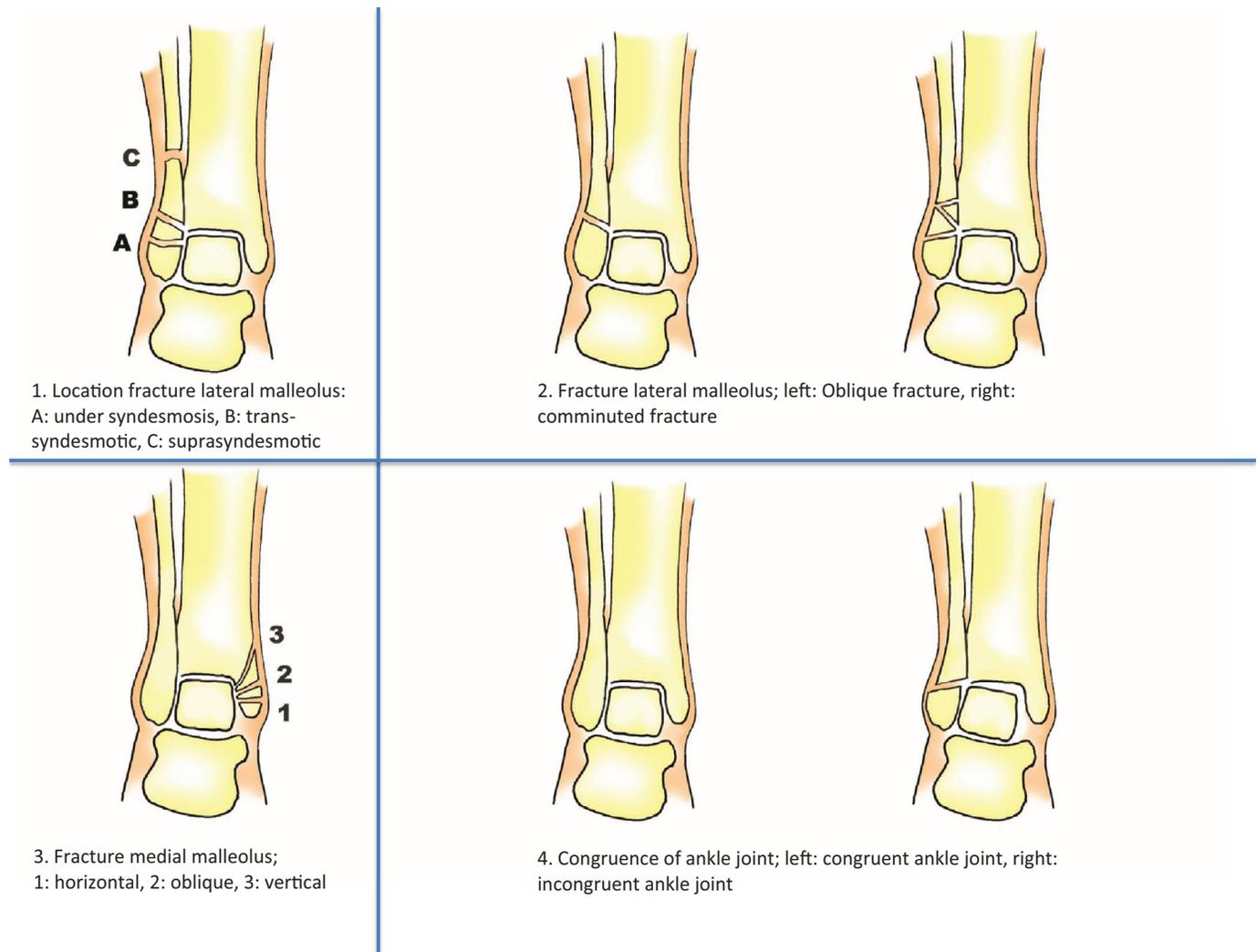


Fig. 1. Descriptive system for classification of ankle fractures.

purely osseous classification. It, therefore, provides less accurate information about the stability of the ankle joint as a whole, compared with the Lauge-Hansen classification (23,24).

Good interobserver and intraobserver agreement is essential for a good classification system. The paradoxical result of a higher intraobserver agreement compared with a lower (slight to moderate) interobserver agreement for the Lauge-Hansen classification in this study seemingly reflects a high level of experience-based self-confidence among observers in the presence of morphologic complexities. Morphologic complexities will make the diagnostic process more challenging, and even consensus meetings with observers, with an exchange of

insights to reduce differences in interpretation, may not be sufficient to overcome such complexity (25).

As shown in this study, some experience is required for the proper staging of an ankle fracture. Novice observers did not perform sufficiently

Table 1
Explanation of κ values

κ Value	Interpretation
0.81 to 1	Almost perfect agreement
0.61 to 0.80	Substantial agreement
0.41 to 0.60	Moderate agreement
0.21 to 0.40	Fair agreement
0.01 to 0.20	Slight agreement
< 0	Poor agreement

Table 2
Level of experience

Level	Years of Experience
Expert 1	5
Expert 2	3
Expert 3	4
Expert 4	20
Expert 5	10
Semiexpert 1	First-year resident
Semiexpert 2	Third-year resident
Semiexpert 3	First-year resident
Semiexpert 4	Fifth-year resident
Semiexpert 5	First-year resident
Semiexpert 6	Third-year resident
Semiexpert 7	Fourth-year resident
Semiexpert 8	Third-year resident
Semiexpert 9	First-year resident
Semiexpert 10	Second-year resident
All novices	Third-year medical school

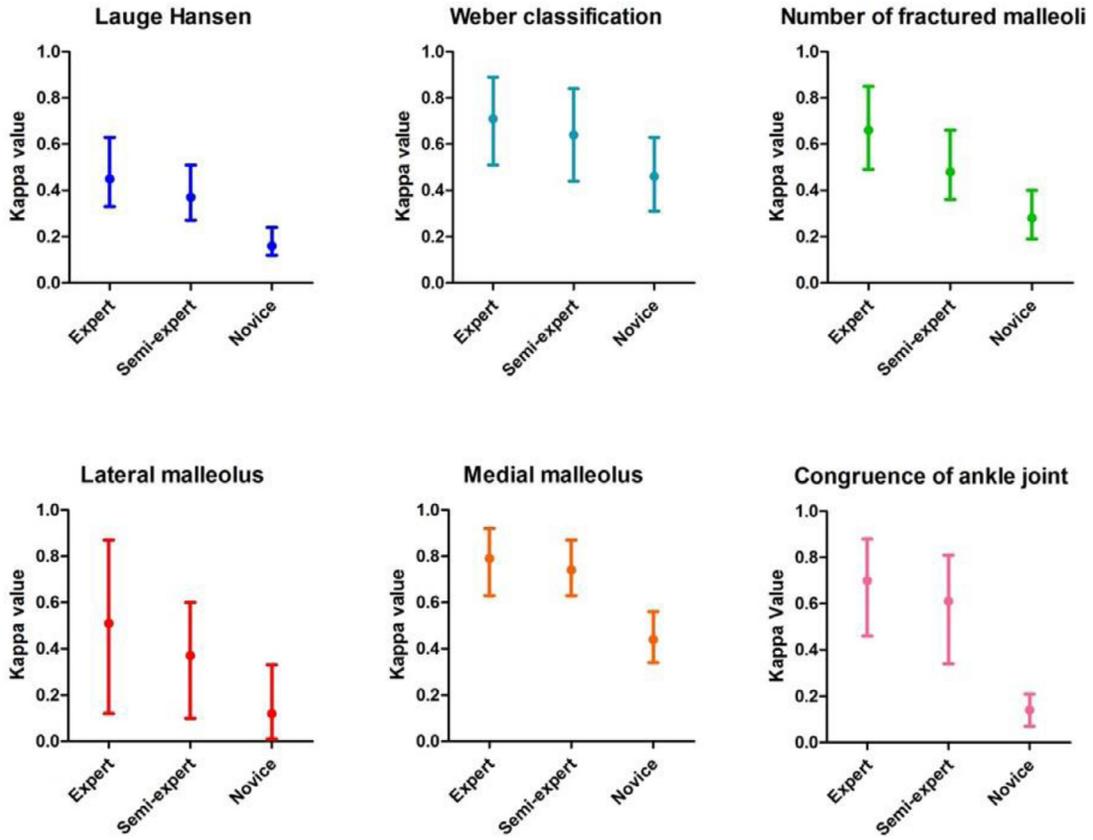


Fig. 2. Interobserver variation.

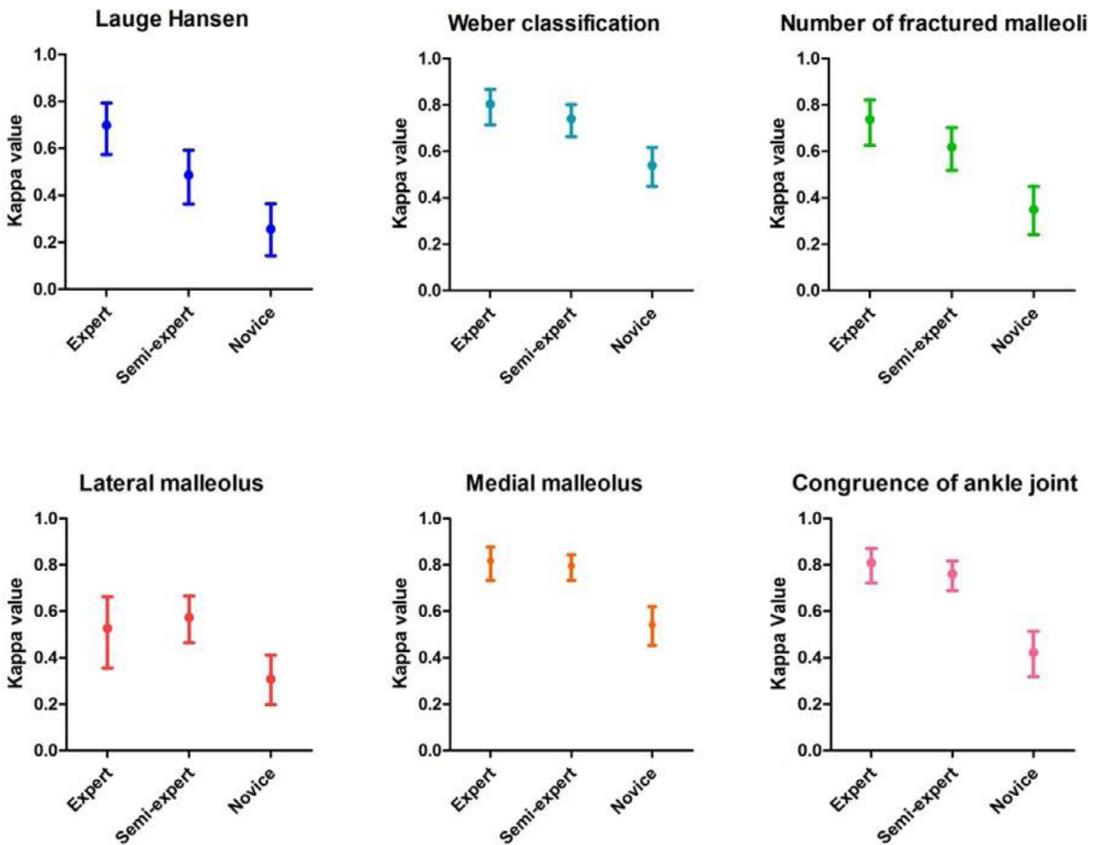


Fig. 3. Intraobserver variation.

on either the Lauge-Hansen or the descriptive system (even the medial malleolar fracture items). This may be due to lack of experience in assessing ankle radiographs. The prestudy lecture that was intended to provide novices with some idea of the different aspects of ankle fractures did not overcome this problem. This suggests that experience is of great importance for the classification of ankle fractures. For the descriptive classification, both experts and semiexperts showed predominantly moderate to substantial interobserver and intraobserver agreement. This consensus is important for the transfer of information between professionals and for patient counseling purposes. Therefore, the descriptive-based classification provides a viable, systematic, and easily comprehensible tool useful for clinical handoffs and research.

This descriptive system also aids in determining the optimal surgical approach. By describing the type of fracture of the medial malleolus, the indication for either zuggurtung osteosynthesis (tension band wiring), compression screws, or (buttress) plate fixation becomes clear. Furthermore, the combination of the congruence of the ankle joint and the location of the fracture of the lateral malleolus assists in determining the need for syndesmotom fixation. Therefore, by describing the ankle fracture, the (orthopedic) trauma surgeon may be as equally well informed about the fracture as if the Lauge-Hansen classification were used.

In recent years, multiple descriptive systems have been developed to classify a broad range of fractures with improved ease-of-use and good observer agreement. For example, the Luo classification for tibial plateau fractures is easier to use than the Schatzker classification (26–28). The simplicity of the Luo “3 column” classification is its greatest strength. This also applies to the Neer or Hertel classification of proximal humeral fractures, 2 descriptive systems that have been shown to be more useful in clinical practice to guide treatment and prognosis than the AO classification (29,30).

This study should be considered in light of its shortcomings. In this study, radiographs of only 20 patients were scored and no clinical information was provided to the observers. The mechanism of injury could have helped the observers classify the fracture.

Furthermore, experience with the Lauge-Hansen classification ranged widely in the semiexpert group, as the group was composed of first- through fifth-year (orthopedic) surgical residents. However, this range in experience reflects real-world differences in healthcare providers with different training backgrounds and as such may not necessarily be considered a limitation.

There were low interobserver and intraobserver agreements for the type of fracture of the lateral malleolus in the descriptive classification. Most fractures of the lateral malleolus were oblique. Overall, approximately 75% of all fractures were classified as an oblique fracture of the lateral malleolus in both time points. This lack of observed variability in the type of fracture pattern limits the agreement above chance, presumably causing the lower interobserver and intraobserver values.

All classification systems have their own advantages, but low reproducibility and high interobserver variability limit the clinical effectiveness of the current classification systems. The descriptive system presented in this study shows less variability between observers, has clinical implications, and is easy to use. To improve clinical and research handoffs and overcome the limitations of current classification systems, the use of this descriptive system is advised.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1053/j.jfas.2018.09.028](https://doi.org/10.1053/j.jfas.2018.09.028).

References

1. Yin MC, Yuan XF, Ma JM, Xia Y, Wang T, Xu XL, Yan YJ, Xu JH, Ye J, Tong ZY, Feng YQ, Wang HB, Wu XQ, Mo W. Evaluating the reliability and reproducibility of the AO and Lauge-Hansen classification systems for ankle injuries. *Orthopedics* 2015;38:e626–e630.
2. Verhage SM, Rhemrev SJ, Keizer SB, Quarles van Ufford HM, Hoogendoorn JM. Inter-observer variation in classification of malleolar fractures. *Skeletal Radiol* 2015;44:1435–1439.
3. Lauge-Hansen N. Fractures of the ankle. II. Combined experimental-surgical and experimental-roentgenologic investigations. *Arch Surg* 1950;60:957–985.
4. Weber B. *Die Verletzungen des oberen Sprunggelenkes*. Huber, Bern, Switzerland; 1972.
5. Müller M, Koch P, Nazarian S, Schatzker J. *The Comprehensive Classification of Fractures of Long Bones*. Springer-Verlag, Berlin/Heidelberg; 1990.
6. Lauge-Hansen N. Fractures of the ankle. IV. Clinical use of genetic roentgen diagnosis and genetic reduction. *AMA Arch Surg* 1952;64:488–500.
7. Alexandropoulos C, Tsourvakas S, Papachristos J, Tselios A, Soukoulis P. Ankle fracture classification: an evaluation of three classification systems: Lauge-Hansen, A.O. and Broos-Bisschop. *Acta Orthop Belg* 2010;76:521–525.
8. Boggs LR. Isolated posterior malleolar fractures. *Am J Emerg Med* 1986;4:334–336.
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;20:267–272.
10. Okanobo H, Khurana B, Sheehan S, Duran-Mendicuti A, Arianjam A, Ledbetter S. Simplified diagnostic algorithm for Lauge-Hansen classification of ankle injuries. *Radiographics* 2012;32:E71–E84.
11. van Wessel KJ, Leenen LP. A rare type of ankle fracture: syndesmotom rupture combined with a high fibular fracture without medial injury. *Injury* 2016;47:766–775.
12. Lindsjo U. Classification of ankle fractures: the Lauge-Hansen or AO system? *Clin Orthop Relat Res* 1985;199:12–16.
13. Nielsen JO, Dons-Jensen H, Sorensen HT. Lauge-Hansen classification of malleolar fractures. An assessment of the reproducibility in 118 cases. *Acta Orthop Scand* 1990;61:385–387.
14. Rasmussen S, Madsen PV, Bennicke K. Observer variation in the Lauge-Hansen classification of ankle fractures. Precision improved by instruction. *Acta Orthop Scand* 1993;64:693–694.
15. Thomsen NO, Overgaard S, Olsen LH, Hansen H, Nielsen ST. Observer variation in the radiographic classification of ankle fractures. *J Bone Jt Surg Br* 1991;73:676–678.
16. Rodriguez EK, Kwon JY, Herder LM, Appleton PT. Correlation of AO and Lauge-Hansen classification systems for ankle fractures to the mechanism of injury. *Foot Ankle Int* 2013;34:1516–1520.
17. Briet JP, Houwert RM, Smeeing DP, Pawiroredjo JS, Kelder JC, Lansink KW, Leenen LP, van der Zwaal P, van Zutphen SW, Hoogendoorn JM, van Heijl M, Verleisdonk EJ, van Lammeren GW, Segers MJ, Hietbrink F. Weight bearing or non-weight bearing after surgically fixed ankle fractures, the WOW! Study: study protocol for a randomized controlled trial. *Trials* 2015;16:175.
18. Engels JM, Diehr P. Imputation of missing longitudinal data: a comparison of methods. *J Clin Epidemiol* 2003;56:968–976.
19. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
20. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–220.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
22. Schatzker J, Tile M. *The Rationale of Operative Fracture Care*. Springer-Verlag, Berlin/Heidelberg; 2013.
23. Fox A, Wykes P, Eccles K, Barrie J. Five years of ankle fractures grouped by stability. *Injury* 2005;36:836–841.
24. Pakarinen H. Stability-based classification for ankle fracture management and the syndesmosis injury in ankle fractures due to a supination external rotation mechanism of injury. *Acta Orthop* 2012;83(suppl):1–26.
25. McCluskey M. Consensus building and verbal desperados. In: *Connected Magazine*, Geneva; 1999.
26. Patange Subba Rao SP, Lewis J, Haddad Z, Paringe V, Mohanty K. Three-column classification and Schatzker classification: a three- and two-dimensional computed tomography characterisation and analysis of tibial plateau fractures. *Eur J Orthop Surg Traumatol* 2014;24:1263–1270.
27. Zhu Y, Yang G, Luo CF, Smith WR, Hu CF, Gao H, Zhong B, Zeng BF. Computed tomography-based three-column classification in tibial plateau fractures: introduction of its utility and assessment of its reproducibility. *J Trauma Acute Care Surg* 2012;73:731–737.
28. Luo CF, Sun H, Zhang B, Zeng BF. Three-column fixation for complex tibial plateau fractures. *J Orthop Trauma* 2010;24:683–692.
29. Papakonstantinou MK, Hart MJ, Farrugia R, Gabbe BJ, Kamali Moaveni A, van Bavel D, Page RS, Richardson MD. Interobserver agreement of Neer and AO classifications for proximal humeral fractures. *ANZ J Surg* 2016;86:280–284.
30. Hertel R, Hempfing A, Stiehler M, Leunig M. Predictors of humeral head ischemia after intracapsular fracture of the proximal humerus. *J Shoulder Elbow Surg* 2004;13:427–433.