DOI: https://doi.org/10.30795/scijfootankle.2019.v13.885

ORIGINAL ARTICLE

Scientific Journal of the Foot & Ankle

Concordance analysis of adult ankle fracture classifications

Análise de concordância das classificações das fraturas de tornozelo do adulto

Antonio Candido de Paula Neto[1], Douglas Hideki Ikeuti[1], Augusto Braga dos Santos[2], Rui dos Santos Barroco[1], Bruno Rodrigues de Miranda[1], Rafael Rocha Macedo[1]

1. Faculdade de Medicina do ABC, Santo André, SP, Brazil.
2. Hospital Ortopédico de Goiânia, Goiânia, GO, Brazil.

ABSTRACT

Objective: To analyze the interobserver agreement among physicians in an orthopedy and traumatology service with knowledge and training in adult ankle fracture classifications through radiographic evaluation.

Methods: A cross-sectional study was carried out in which the records of 20 patients with a diagnosis of ankle fracture in 2016 (uni-, bi- or trimalleolar), aged older than 18 years or with a closed physis, were studied. The radiographs that met the criteria were analyzed and classified according to the Danis-Weber (DW), Lauge-Hansen (LH) and *Arbeitsgemeinschaft für Osteosynthesefragen/Orthopaedic Trauma Association* (AO/OTA) classifications by two first-year residents, two second-year residents, two third-year residents and two preceptors of the service. The conditions were identical for the observers and were performed individually once. The Kappa coefficient (K) was used to assess agreement.

Results: The results showed great interobserver agreement in the DW and LH classifications, with K values equal to 0.69 and 0.64, respectively. The AO/OTA classification showed moderate agreement, with K = 0.47.

Conclusion: The DW classification was the most reliable among the three classifications. The AO/OTA classification showed the lowest agreement and was the least reliable, and the LH classification was in between the others.

Level of Evidence III, Diagnostic Studies; Study of non consecutive patients; without consistently applied reference "gold" standard.

Keywords: Ankle fracture/classification; Ankle fracture/diagnostic imaging; Observer variation; Ankle injuries; Reproducibility of the tests.

RESUMO

Objetivo: Analisar a concordância inter-observador entre médicos de um serviço de ortopedia e traumatologia com conhecimento e treinamento nas classificações de fraturas de tornozelo de adultos, através de avaliação radiográfica.

Métodos: Foi realizado um estudo transversal em que foram estudados prontuários de 20 pacientes com diagnóstico de fratura de tornozelo no ano de 2016 (uni, bi ou trimaleolar), com idade acima de 18 anos ou fises fechadas. As radiografias que preencheram os critérios foram analisadas e classificadas de acordo com as classificações de Danis-Weber, Lauge-Hansen e *Arbeitsgemeinschaft für Osteosynthesefragen - Orthopaedic Trauma Association* (AO-OTA), por dois residentes do primeiro ano, dois residentes do segundo ano, dois residentes do terceiro ano e dois preceptores do serviço. As condições para os observadores foram idênticas e executadas individualmente uma vez. Para avaliação da concordância foi utilizado o coeficiente Kappa (K).

Resultados: Os resultados obtidos foram: nas classificações de Danis-Weber e Lauge-Hansen obteve-se grande concordância inter-observador, com valores de K iguais a 0.69 e 0.64, respectivamente. Quanto à classificação AO-OTA obteve-se concordância moderada com K = 0.47.

Conclusão: A classificação de Danis-Weber foi a mais confiável entre as três. A classificação da AO-OTA foi a que teve menor concordância, sendo a menos confiável, e a classificação de Lauge-Hansen ficou em uma posição intermediária.

Nível de Evidência III; Estudos Diagnósticos; Estudo de pacientes não consecutivo; sem padrão de referência "ouro" aplicado uniformemente.

Descritores: Fratura do tornozelo/classificação; Fratura de tornozelo/diagnóstico por imagem; Variações dependentes do observador; Traumatismos do tornozelo; Reprodutibilidade dos testes.

Work performed at the Faculdade de Medicina do ABC, Santo André, SP, Brazil.

Corresponcence: Antônio Candido de Paula Neto. Rua Xingu, 415, Vila Valparaiso, CEP: 09060-050, Santo André, SP, Brazil

Conflicts of interest: none. **Source of funding:** none.
Date received: December 03, 2018. **Date accepted:** February 05, 2019. **Online:** March 31, 2019

How to cite this article: Paula Neto AC, Ikeuti DH, Santos AB, Barroco RS, Miranda BR, Macedo RR. Concordance analysis of adult ankle fracture classifications. Sci J Foot Ankle. 2019;13(1):10-4.

INTRODUCTION

Adult ankle fractures are among the most commonly found fractures, accounting for approximately 10% of all fractures. Their incidence is higher in young men and elderly women[1]. Smoking and the body mass index (BMI) are important risk factors and predictors of ankle fracture[2,3].

Most ankle fractures are unimalleolar (60-70%), followed by bimalleolar fractures (15-20%) and trimalleolar fractures (7-12%)[4,5]. The diagnosis is made from simple radiography. The main radiographic views used are anteroposterior (AP), AP with 15° of internal rotation (ankle mortise) and lateral[6].

Three classifications are most commonly used today: Danis-Weber (DW), Lauge-Hansen (LH) and *Arbeitsgemeinschaft für Osteosynthesefragen/Orthopaedic Trauma Association* (AO/OTA). Fracture classification together with soft-tissue analysis contributes to the adequate management of this complaint[6]. Studies have evaluated the reproducibility of these classifications, and the DW classification is the most reproducible[7,8].

The objectives of this study were as follows: (1) to assess the overall interobserver agreement for the three most commonly used radiographic classifications for ankle fractures (DW, LH and AO/OTA) and (2) to analyze the interobserver agreement for each subtype in the three classifications.

The motivation for this study was to determine which classification is the most reliable and reproducible, in addition to contributing data for future studies.

METHODS

This study was approved by the Research Ethics Committee with registration in the Brazil Platform under CAAE number: 64525417.2.0000.0035.

This was a cross-sectional study, in which 20 ankle radiographs from the hospital database were evaluated in 2016 that had a radiographic diagnosis of uni-, bi- or trimalleolar ankle fracture. Radiographs that met the inclusion criteria were randomly chosen from the database.

The inclusion criteria for the radiographs were as follows: age older than 18 years or with a closed growth physis; radiographic diagnosis of ankle fracture; and uni-, bi- or trimalleolar fractures in the anteroposterior (AP) and lateral views. The exclusion criteria were as follows: age younger than 18 years or open physis; age older than 80 years; radiographic sign of pathological fracture; and tibial pilon fracture.

The radiographs that met the criteria were analyzed and classified according to the DW, LH and AO/OTA classifications by two first-year residents, two second-year residents, two third-year residents and two service preceptors. All the evaluators had previous knowledge of the classifications and could refer to them. The conditions for the observers were identical and performed individually once.

The DW classification divides ankle fractures into three types, taking into account the height of the lateral malleolus fracture line: A) below syndesmosis, B) at syndesmosis level and C) above syndesmosis[9,10].

The LH classification is based on the foot position and direction of the deforming force, yielding four groups: supination-adduction (SAD), supination-external rotation (SER), pronation-external rotation (PER) and pronation-abduction (PAB)[11,12].

The AO/OTA classification is based on the DW classification and is determined by fracture lines and fibular localization in relation to the syndesmosis level, in addition to medial facet lesions and ligament avulsions of the distal tibia. The classification comprises the malleolar segment (44), which is subdivided into three main groups: A) infrasyndesmotic lesion; B) transsyndesmotic lesion; and C) suprasyndesmotic lesion. It is further subdivided as follows: A1) isolated; A2) bifocal; A3) circumferential; B1) lateral isolated; B2) lateral and medial; B3) medial lateral and posterior; C1) simple diaphyseal; C2) multifragmentary; and C3) proximal[6,13].

The agreement between the observers was analyzed by the Kappa coefficient (K), according to the description of Landis[14]. The value of K may vary from -1 to +1, where K = 1 indicates perfect agreement, K = 0 shows no agreement beyond chance, and K < 0 indicates no agreement.

The K values are presented as absolute values, in addition to their respective confidence intervals (95% CI) and p-value. The p-value was considered statistically significant when $p < 0.05$, indicating that the value obtained is significantly different from 0 ($K \neq 0$).

We used the criteria described in table 1 to interpret the agreement using the K value.

RESULTS

Table 2 shows the K values, 95% confidence interval (95% CI) and p-value for each of the three radiographic classifications. The DW and LH classifications showed high interobserver agreement, while the AO/OTA classification showed moderate agreement. The p-value was <0.001 in all three situations.

In the DW classification, the three fracture types showed a high degree of agreement, all with a statistically significant p-value (Table 3).

In the LH classification, types SER, PER and SAD presented high interobserver agreement, with values of K varying between 0.63 and 0.79. In turn, the PAB type presented a negligible agreement value (K = 0.12) (Table 4).

Table 1. Interpretation of the Kappa Coefficient (K).

Kappa values	Interpretation of agreement
<0	No agreement
0-0.19	Negligible
0.20-0.39	Slight
0.40-0.59	Moderate
0.60-0.79	High
0.80-1.00	Almost perfect

Source: Landis & Koch, 1977.

Table 2. Kappa results of the general analysis of classifications and agreements

Classification	Kappa (95% CI)	p-value	Agreement
Danis-Weber	0.69 (0.63-0.76)	<0.001	High
Lauge-Hansen	0.64 (0.58-0.70)	<0.001	High
AO/OTA	0.47 (0.43-0.51)	<0.001	Moderate

Source: Prepared by the author based on the results of the research.

Table 3. Resultado da análise da concordância de cada um dos três tipos de fratura de acordo com a classificação de Danis-Weber (DW).

Tipo de fratura de acordo com DW	K da categoria	95% CI	p-value
Type A	0.71	0.63 – 0.80	<0.001
Type B	0.73	0.65 – 0.82	<0.001
Type C	0.65	0.57 – 0.73	<0.001

Source: Prepared by the author based on the results of the research.

For the AO/OTA classification, of the seven types present in the sample, two had a high degree of agreement (44B1 and 44B3), two had moderate agreement (44C1 and 44C3), one showed slight agreement (44B2), and two presented negligible agreement (44C2 and 44A1). Only 44A1 presented a nonsignificant p-value (Table 5).

DISCUSSION

The classifications of orthopedic fractures involve certain parameters, among them the ease of understanding and consequent agreement between multiple observers. Because ankle fractures are among the injuries most commonly treated by orthopedic surgeons, their careful identification and the identification of soft-tissue injuries are necessary for proper treatment[15-17].

In the literature, we found studies that evaluated the variation in the intra- and interobserver classification of ankle fractures. All the studies used the K coefficient to analyze the agreement, which was shown to be the most reliable to conduct the study[18-23]. In the Brazilian literature, we found four studies with a similar analysis[7,12,24,25].

Table 4. Results of the agreement analysis of each of the four fracture types according to the Lauge-Hansen (LH) classification.

Type of fracture according to LH	K of the category	95% CI	p-value
SER	0.76	0.67 – 0.84	<0.001
PER	0.63	0.54 – 0.71	<0.001
PAB	0.12	0.04 – 0.20	0.004
SAD	0.79	0.70 – 0.87	<0.001

SER: supination-external rotation; PER: pronation-external rotation; PAB: pronation-abduction; SAD: supination-adduction.

Source: Prepared by the author based on the results of the research.

Table 5. Results of the concordance analysis of each of the seven fracture types according to the AO/OTA classification.

Type of fracture according to AO/OTA classification	K of the category	95% CI	p-value
44A1	0.02	0.06 – 0.10	0.59
44B1	0.61	0.53 – 0.70	<0.001
44B2	0.28	0.20 – 0.36	<0.001
44B3	0.64	0.55 – 0.72	<0.001
44C1	0.50	0.41 – 0.58	<0.001
44C2	0.17	0.10 – 0.26	<0.001
44C3	0.54	0.46 – 0.63	<0.001

* Of the nine possible types in the AO/OTA classification, no cases were identified that fit into types 44A2 and 44A3 in the sample evaluated.

Source: Prepared by the author based on the results of the research.

The overall interobserver agreement and the interobserver agreement for each category of the DW, LH and AO/OTA classifications were analyzed in this study. Based on the data obtained, the DW classification had the highest overall agreement, with a value of 0.69, which represents high agreement according to Landis (95% CI = 0.63-0.76). In the LH classification, we obtained an overall agreement slightly lower than that of the DW classification, with K = 0.64, which is classified as high agreement, with 95% CI = 0.58-0.70. In the AO/OTA classification, we obtained the lowest overall agreement, with a value of 0.47, which is classified as moderate agreement, with 95% CI = 0.43-0.51.

In the analysis of each category, the DW classification was the most reliable, with a small agreement variation in the three categories (A, B and C) and a statistically significant p-value, in addition to having high agreement in all categories. The LH classification showed an intermediate variation in agreement as a function of the categories; the SER, PER and SAD types showed high agreement, whereas the PAB type presented negligible agreement, all with a statistically significant p-value. The AO/OTA classification had the highest agreement variation when the subtypes were analyzed. Two had high agreement (44B1 and 44B3), two had moderate agreement (44C1 and 44C3), one showed slight agreement (44B2), two showed negligible agreement (44C2 and 44A1), and one (44A1) had a nonsignificant p-value.

The problem with the LH classification is that in many cases, one cannot definitively determine the mechanism of the injury that caused the fracture[26]. The difficulties of the DW and AO/OTA classifications include defining the position of tibiofibular syndesmosis and the impossibility of classifying isolated fractures of the medial malleolus[23]. The AO/OTA classification is not widely used in clinical practice because it has many subtypes (nine possibilities), making it difficult to memorize and use in the context of urgency/emergency units. Due mainly to these factors, different interpretations in the mentioned classifications are observed. Routine use of classifications, along with time, considerably reduces errors at the time of classification[24].

Based on these results, the DW classification was the most consistent among observers and is the most reproducible and most reliable for the choice of treatments. The LH classification showed good reproducibility but lower than that of the DW classification. The AO/OTA classification was shown to be the least concordant; therefore, it was less reproducible.

A study found in the literature comparing the three classifications yielded slightly different results, where moderate reproducibility was found for all classifications (DW, K = 0.49; AO/OTA, K = 0.45; LH, K = 0.47). The difference may be due to the greater number of cases analyzed (100), and only four observers had a similar academic background[24].

This study presents some limitations, such as a small sample size, nonrandomization and nonseparation by groups according to academic background.

CONCLUSION

The overall interobserver agreement was higher in the DW and LH radiographic classifications, and they were the most reproducible. The concordance in the subtypes was higher in the DW classification and intermediate in the LH classification and had great variation in the AO/OTA classification subtypes.

Authors' contributions: Each author contributed individually and significantly to the development of this article: ACPN *(https://orcid.org/0000-0001-8459-9601) conceived and planned the activities that led to the study, data collection, wrote the article, participated in the review process, approved the final version; DHI *(https://orcid.org/0000-0003-4337-0880) data collection, wrote the article, interpreted the results of the study, participated in the review process, approved the final version; ABS *(https://orcid.org/0000-0002-0773-9284) participated in the review process, approved the final version; RSB *(https://orcid.org/0000-0002-2870-2261) participated in the review process, approved the final version; BRN *(https://orcid.org/0000-0002-5306-2972) interpreted the results of the study, participated in the review process, approved the final version; RRM *(https://orcid.org/0000-0002-2563-2085) participated in the review process, approved the final version. *ORCID (Open Researcher and Contributor ID).

REFERENCES

1. Jensen SL, Andresen BK, Mencke S, Nielsen PT. Epidemiology of ankle fractures. A prospective population-based stud y of 212 cases in Aalborg, Denmark. Acta Orthop Scand. 1998;69(1):48-50.
2. Valtola A, Honkanen R, Kröger H, Tuppurainen M, Saarikoski S, Alhava E. Lifestyle and other factors predict ankle fractures in perimenopausal women: a population-based prospective cohort study. Bone. 2002;30(1):238-42.
3. Honkanen R, Tuppurainen M, Kröger H, Alhava E, Saarikoski S. Relationships between risk factors and fractures differ by type of fracture: a population-based study of 12,192 perimenopausal women. Osteoporos Int. 1998;8(1):25-31.
4. Daly PJ, Fitzgerald RH Jr, Melton LJ, Ilstrup DM. Epidemiology of ankle fractures in Rochester, Minnesota. Acta Orthop Scand. 1987; 58(5):539-44.
5. Court-Brown CM, McBirnie J, Wilson G. Adult ankle fractures--an increasing problem? Acta Orthop Scand. 1998;69(1):43-7.
6. Court-Brown CM, Heckman JD, McQueen MM, Ricci WM. Rockwood and Green's fractures in adults. 8thed Philadelphia: Wolters Kluwer; 2015. p.2541-2560.
7. Tenório RB, Mattos CA, Araújo LHC, Belangero WD. Análise da reprodutibilidade das classificações de Lauge-Hansen e Danis-Weber para fraturas de tornozelo. Rev Bras Ortop. 2001; 36(11/12):434-7.
8. Michelson JD. Fractures about the ankle. J Bone Joint Surg Am 1995; 77:142.
9. Danis R. Les fractures malleolaires. In: Danis R, editor. Théorie et pratique de l'ostéosynthèse. Masson; Paris: 1949. p. 133-165.
10. Weber BG. Die Verletzungen des oberen Sprung-gelenkes. 2nded. Berne: Verlag Hans Huber; 1972.
11. Lauge-Hansen N. Ligamentous ankle fractures. Diagnosis and treatment. Acta Chir Scand. 1949;97(6):544-50.
12. Pimenta LSM. Estudo experimental e radiográfico das fraturas maleolares do tornozelo baseado na classificação de Lauge-Hansen [Dissertação]. São Paulo: Departamento de Ortopedia e Traumatologia da Faculdade de Medicina da Universidade de São Paulo; 1991.
13. Barbosa P, Bonnaire F, Kojima K. Diagnosis. In: Krikler S, editor. Malleoli. Davos, Switzerland: AO Foundation; 2016. Avaible at: https://www2.aofoundation.org/wps/portal/surgery?showPage=diagnosis&bone=Tibia&segment=Malleoli.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
15. Joy G, Patzakis MJ, Harvey JP Jr. Precise evaluation of the reduction of severe ankle fractures. J Bone Joint Surg Am. 1974;56(5):979-93.
16. Pettrone FA, Gail M, Pee D, Fitzpatrick T, Van Herpe LB. Quantitative criteria for prediction of the results after displaced fracture of the ankle. J Bone Joint Surg Am. 1983;65(5):667-77.
17. Phillips WA, Schwartz HS, Keller CS, Woodward HR, Rudd WS, Spiegel PG, Laros GS. A prospective, randomized study of the management of severe ankle fractures. J Bone Joint Surg Am. 1985;67(1):67-78.
18. Lindsjö U. Classification of ankle fractures: Lauge-Hansen or AO system? Clin Orthop Relat Res. 1985;(199):12-6.
19. Thomsen NO, Overgaard S, Olsen LH, Hansen H, Nielsen ST. Observer variation in the radiographic classification of ankle fractures. J Bone Joint Surg Br. 1991;73(4):676-8.
20. Alexandropoulos C, Tsourvakas S, Papachristos J, Tselios A, Soukouli P. Ankle fracture classification: an evaluation of three classification systems: Lauge-Hansen, A.O. and Broos-Bisschop. Acta Orthop Belg. 2010;76(4):521-5.
21. Juto H, Möller M, Wennergren D, Edin K, Apelqvist I, Morberg P. Substantial accuracy of fracture classification in the Swedish Fracture Register: Evaluation of AO/OTA-classification in 152 ankle fractures. Injury. 2016;47(11):2579-2583.
22. Craig WL 3rd, Dirschl DR. Effects of Binary Decision Making on the Classification of Fractures of the Ankle. J Orthop Trauma 1998;12: 280-283.
23. Verhage SM, Rhemrev SJ, Keizer SB, Quarles van Ufford HM, Hoogendoorn JM. Interobserver variation in classification of malleolar fractures. Skeletal Radiol. 2015;44(10):1435-9.
24. Fonseca LLD, Nunes IG, Nogueira RR, Martins GEV, Mesencio AC, Kobata SI. Reproducibility assessment of the Lauge-Hansen classification for ankle fractures. Sci J Foot Ankle. 2018;12(1):24-8.
25. Fonseca LLD, Nunes IG, Nogueira RR, Martins GEV, Mesencio AC, Kobata SI. Reproducibility of the Lauge-Hansen, Danis-Weber, and AO classifications for ankle fractures. Rev Bras Ortop. 2018;53(1): 101-106.
26. Boszczyk A, Fudalej M, Kwapisz S, Klimek U, Maksymowicz M, Kordasiewicz B, Rammelt S. Ankle fracture - Correlation of Lauge-Hansen classification and patient reported fracture mechanism. Forensic Sci Int. 2018;282:94-100.