



Avaliação da reprodutibilidade das classificações Lauge-Hansen, Danis-Weber e AO para fraturas do tornozelo*

Evaluation of the Reproducibility of Lauge-Hansen, Danis-Weber and AO Classifications for Ankle Fractures

Lucas Sacramento Ramos¹ Henrique Mansur Gonçalves² Anderson Freitas³
 Marcio de Paiva Oliveira³ Diogo Marcelino Santos Lima⁴ Welvis Soares Carmargo¹

¹Departamento de Ortopedia e Traumatologia, Hospital Regional do Gama, Brasília, DF, Brasil

²Departamento de Medicina, Universidade de Brasília, Brasília, DF, Brasil

³Instituto de Pesquisa e Ensino do Hospital Ortopédico e Medicina Especializada, Brasília, DF, Brasil

⁴Departamento de Medicina, Centro Universitário de Brasília (CEUB), Brasília, DF, Brasil

Endereço para correspondência Lucas Sacramento Ramos, SQS 214, bloco C, apartamento 605, Asa Sul, Brasília, DF, 70293-030, Brasil (e-mail: lucas.sacramentoramos@gmail.com).

Rev Bras Ortop 2021;56(3):372–378.

Resumo

Objetivo Avaliar a reprodutibilidade intra- e interobservador das classificações de Lauge-Hansen, Danis-Weber e Arbeitsgemeinschaft für Osteosynthesefragen (AO) para as fraturas de tornozelo, e a influência do estágio de formação dos participantes na avaliação.

Métodos Foram selecionadas radiografias de 30 pacientes com fratura de tornozelo nas incidências anteroposterior (AP), perfil e AP verdadeiro. Todas as imagens foram avaliadas por 11 participantes em diferentes estágios de formação profissional (cinco residentes e seis cirurgiões ortopédicos), em dois momentos distintos. Analisou-se a concordância inter- e intraobservador por meio do coeficiente Kappa ponderado. O teste t de Student para amostras pareadas foi aplicado para verificar se havia diferença significativa no grau de concordância interobservador entre os instrumentos.

Resultado Observou-se que existe concordância significativa em todas as classificações quando da análise intraobservador isolada. Existe concordância interobservador altamente significativa de grau moderado a ótimo na classificação de Danis-Weber ($p \leq 0,0001$). A classificação de Danis-Weber apresentou, em média, grau de concordância significativamente maior que as outras classificações ($p \leq 0,0001$).

Palavras-chave

- ▶ fraturas do tornozelo
- ▶ classificação
- ▶ reprodutibilidade dos testes

* Estudo desenvolvido pelo serviço de ortopedia e traumatologia do Hospital Regional do Gama-DF e pelo Instituto de pesquisa e ensino do Hospital Ortopédico e Medicina Especializada (IPE-HOME-DF) Brasília, DF, Brasil.

recebido
16 de Março de 2020
aceito
06 de Julho de 2020
Publicado on-line
Dezembro 18, 2020

DOI <https://doi.org/10.1055/s-0040-1718508>.
ISSN 0102-3616.

© 2020. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

Conclusão A classificação de Danis-Weber se mostrou a mais reprodutiva entre os instrumentos avaliados, e a pouca experiência do avaliador não influencia negativamente a reprodutibilidade das classificações das fraturas do tornozelo. *Nível de Evidência II, Estudos Diagnósticos - Investigação de um Exame para Diagnóstico.*

Abstract

Objective The present study aims to analyze the intra- and interobserver reproducibility of the Lauge-Hansen, Danis-Weber, and Arbeitsgemeinschaft für Osteosynthesfragen (AO) classifications for ankle fractures, and the influence of evaluators training stage in these assessments.

Methods Anteroposterior (AP), lateral and true AP radiographs from 30 patients with ankle fractures were selected. All images were evaluated by 11 evaluators at different stages of professional training (5 residents and 6 orthopedic surgeons), at 2 different times. Intra- and interobserver agreement was analyzed using the weighted Kappa coefficient. Student t-tests for paired samples were applied to detect significant differences in the degree of interobserver agreement between instruments.

Results Intraobserver analysis alone had a significant agreement in all classifications. Moderate to excellent interobserver agreement was highly significant ($p \leq 0.0001$) for the Danis-Weber classification. The Danis-Weber classification showed, on average, a significantly higher degree of agreement than the remaining classification systems ($p \leq 0.0001$).

Conclusion The Danis-Weber classification presented the highest reproducibility among instruments and the evaluator's little experience had no negative influence on the reproducibility of ankle fracture classifications. *Level of Evidence II, Diagnostic Studies - Investigating a Diagnostic Test.*

Keywords

- ▶ ankle fractures
- ▶ classification
- ▶ reproducibility of results

Introdução

As fraturas de tornozelo compreendem, aproximadamente, 10% de todas as fraturas do corpo humano, são mais comuns em mulheres e estão associadas a obesidade e tabagismo.^{1,2} Com a transição demográfica, houve um aumento na incidência dessas fraturas em pacientes idosos em cerca de 3 vezes nos últimos 30 anos.¹⁻³ A articulação do tornozelo é constituída pelo encaixe do corpo do tálus na pinça maleolar, funcionando como uma tróclea modificada, estabilizada pelos complexos ligamentares lateral, medial e da sinde-mose.⁴ Essa complicada anatomia osteocapsuloligamentar, ao ser submetida a forças deformantes, principalmente rotacionais, sofre uma variedade de lesões que devem ser estudadas. O diagnóstico das fraturas do tornozelo é baseado na história clínica, no exame físico e na avaliação por imagem da região, normalmente obtida com radiografias simples do tornozelo nas incidências anteroposterior (AP), perfil e anteroposterior com 20° de rotação interna (AP verdadeiro).^{1,4}

As classificações são importantes ferramentas que podem definir o prognóstico e guiar o tratamento mais adequado. Uma classificação de boa qualidade deve ter uma linguagem simples e fornecer informações confiáveis para a correta propedêutica.⁵ Além disso, deve ser viável, confiável e reprodutível; esta última característica é conferida pelas concordâncias inter- e intraobservador.^{5,6} Estudos de reprodutibilidade são clássicos na literatura para avaliar a qualidade de um sistema de classificação, principalmente em ortopedia, pois ajudam a definir qual ins-

trumento produz maior concordância e compreensão na comunidade científica.⁷

A classificação para as fraturas do tornozelo de Lauge-Hansen foi, por muitos anos, o sistema mais utilizado, e é baseada no mecanismo de trauma, levando em consideração a posição do pé e a direção da força deformante, podendo ser: pronação com abdução, pronação com rotação externa, supinação com adução e supinação com rotação externa. A classificação de Danis-Weber é majoritariamente anatômica, e se baseia na topografia e tipo de traço no maléolo lateral, dividindo-se em: infrasindesmal – A, transsindesmal – B e suprasindesmal – C. Já a classificação do Grupo Arbeitsgemeinschaft für Osteosynthesfragen (AO) redefine os três tipos da classificação de Danis-Weber levando em consideração as lesões mediais: infrasindesmal (isolada – A1, associada a maléolo medial – A2 ou associada à fratura posteromedial – A3), transsindesmal (isolada – B1, associada à lesão medial – B2 ou associada à lesão medial e lesão posterolateral – B3) e suprasindesmal (simples – C1, cominutiva – C2 ou fratura da fíbula proximal – C3).^{1,8-10}

Apesar de haver alguns estudos na literatura avaliando a reprodutibilidade das diversas classificações das fraturas de tornozelo, eles são controversos e não há um consenso de qual delas é a mais apropriada. Além disso, pouco foi discutido acerca da relação entre a reprodutibilidade dos instrumentos e a experiência do avaliador.^{11,12} Assim, o objetivo do presente estudo é analisar qual das três principais classificações para fratura de tornozelo possui maior

reprodutibilidade intra- e interobservador, e se o estágio de formação dos participantes influencia na avaliação. Acreditamos que quanto mais complexa for a classificação, menor será sua reprodutibilidade e que os avaliadores mais experientes atingirão maior concordância nas respostas.

Material e Métodos

Foram selecionados pacientes com fraturas de tornozelo no ano de 2018, após a aprovação do Comitê de Ética em Pesquisa com parecer N° 2.697.068/18. O estudo atendeu todos os requisitos em relação aos direitos dos seres humanos.

Foram incluídos pacientes esqueleticamente maduros, com diagnóstico de fratura do tornozelo, de forma aleatória, conforme eram atendidos na emergência do hospital, até um total de 30 pacientes com imagens radiográficas dentro dos padrões estabelecidos pelo estudo que eram: radiografias nas incidências AP, perfil e anteroposterior com 20° de rotação interna do tornozelo (AP verdadeiro do tornozelo). Foram excluídos os pacientes que não apresentavam as radiografias nas incidências citadas, os exames com quali-

dade considerada ruim pelos pesquisadores ou aqueles que não concordaram em participar do estudo.

As radiografias foram fotografadas e digitalizadas no arquivo pessoal do pesquisador principal. Todas as imagens foram inseridas no Survey Monkey Canada Inc., Ottawa, Canada, que gerou um questionário virtual no qual o avaliador qualificava cada imagem de acordo com as classificações de Danis-Weber, Lauge-Hansen e do Grupo A.O. O questionário também possuía figuras ilustrativas com as três classificações para que os participantes pudessem consultar a qualquer momento (► Figuras 1, 2, 3). O questionário virtual foi encaminhado para um total de 11 médicos ortopedistas em diferentes estágios de formação: 6 membros da Sociedade Brasileira de Ortopedia e Traumatologia (SBOT), sendo 2 especialistas em cirurgia do pé e tornozelo (ABTPé) e 4 não especialistas; 5 médicos residentes, um no primeiro ano de formação (R1), 2 no segundo ano (R2) e 2 no terceiro ano (R3), para avaliar a concordância interobservadores. A fim de avaliar a concordância intraobservadores, após 1 mês, o mesmo questionário foi reenviado para que os participantes repetissem o processo.

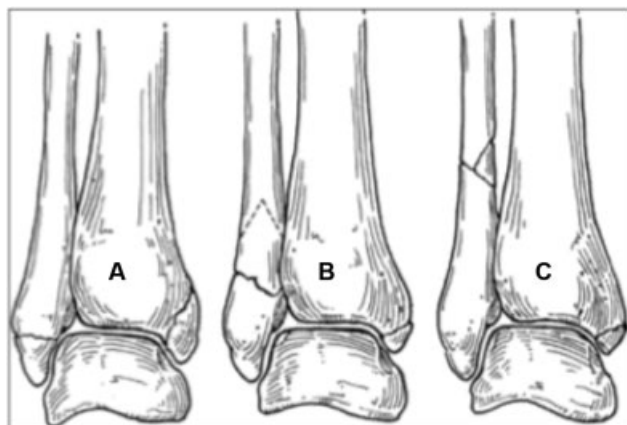


Fig. 1 Classificação de Weber.

Análise Estatística

A análise descritiva apresentou, na forma de tabelas, os dados observados expressos pela frequência (n) e percentual (%). A análise inferencial foi composta pelo coeficiente de Kappa ponderado para análise de concordância intra- e interobservador dos instrumentos de classificação Danis-Weber, Lauge-Hansen e AO em dois momentos de avaliação. O teste t de Student para amostras pareadas foi aplicado para verificar se existia diferença significativa no grau de concordância interobservador entre os instrumentos.

As fidedignidades intra- e interobservador foram avaliadas pelo coeficiente de Kappa ponderado, que verificou se existia concordância significativa na classificação, em escala ordinal, dos instrumentos Danis-Weber (3 níveis), Lauge-Hansen (4 níveis) e AO (9 níveis) entre 2 momentos distintos (M1 e M2, 1 mês após M1) na amostra de 30 estudos

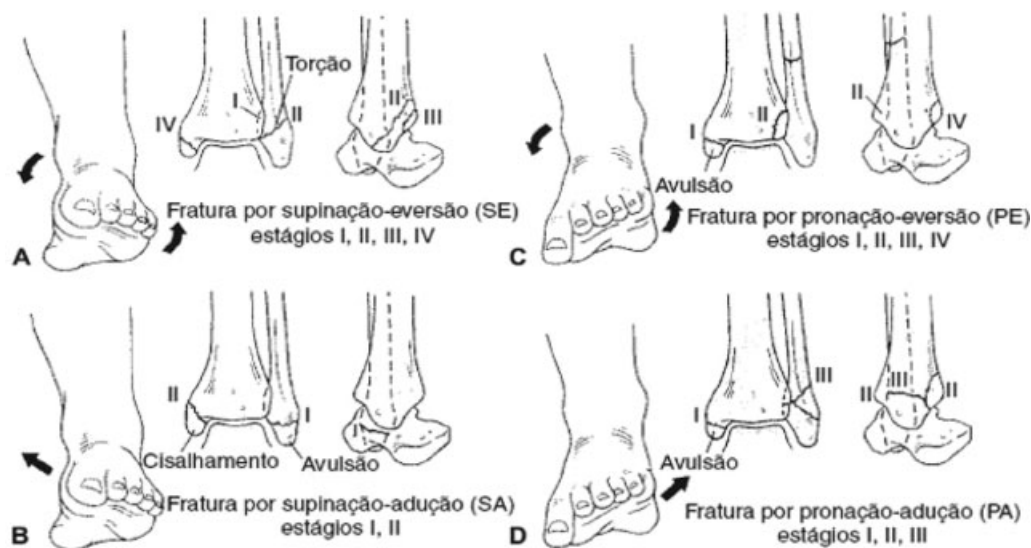


Fig. 2 Classificação de Lauge-Hansen.



Fig. 3 Classificação AO para as fraturas de tornozelo.⁷

radiográficos. Sabe-se que quanto mais próximo o Kappa for de um (1), mais forte (ou perfeita) é a concordância entre os observadores, neste caso, os observadores se assemelham sob o aspecto qualitativo da avaliação. Por outro lado, quanto mais próximo de zero (0), maior é a discordância, ou seja,

significa que não se reproduzem e as diferenças observadas não são ao acaso.

As amostras correspondem aos coeficientes de Kappa das 55 comparações realizadas entre os profissionais, sendo que na amostra total tem 55 comparações, na subamostra de só

Tabela 1 Grau de concordância entre os instrumentos

Amostra*	Instrumento	Momento M1					Momento M2				
		média	±	DP	pares de instrumento	valor-p ^a	média	±	DP	pares de instrumento	valor-p ^a
Total (n = 55)	W (1)	0,81	±	0,11	1 vs 2	< 0,0001	0,79	±	0,11	1 vs 2	< 0,0001
	LG (2)	0,43	±	0,21	1 vs 3	< 0,0001	0,52	±	0,18	1 vs 3	< 0,0001
	AO (3)	0,43	±	0,19	2 vs 3	0,83	0,48	±	0,21	2 vs 3	0,18
entre só especialistas (n = 15)	W (1)	0,76	±	0,10	1 vs 2	< 0,0001	0,79	±	0,13	1 vs 2	< 0,0001
	LG (2)	0,29	±	0,15	1 vs 3	< 0,0001	0,49	±	0,15	1 vs 3	< 0,0001
	AO (3)	0,30	±	0,17	2 vs 3	0,82	0,41	±	0,27	2 vs 3	0,15
entre só residentes (n = 10)	W (1)	0,86	±	0,08	1 vs 2	0,0008	0,76	±	0,08	1 vs 2	0,003
	LG (2)	0,59	±	0,19	1 vs 3	0,0001	0,54	±	0,18	1 vs 3	0,0005
	AO (3)	0,61	±	0,10	2 vs 3	0,85	0,54	±	0,12	2 vs 3	0,84
entre especialistas vs residentes (n = 30)	W (1)	0,82	±	0,11	1 vs 2	< 0,0001	0,80	±	0,11	1 vs 2	< 0,0001
	LG (2)	0,44	±	0,20	1 vs 3	< 0,0001	0,52	±	0,20	1 vs 3	< 0,0001
	AO (3)	0,44	±	0,18	2 vs 3	0,99	0,50	±	0,21	2 vs 3	0,61

Abreviações: AO, Arbeitsgemeinschaft für Osteosynthesefragen; DP, desvio padrão; LG, classificação Lauge-Hansen; W, classificação Danis-Weber.

*amostra de estatísticas Kappa ponderado.

^aTeste t de Student para amostras pareadas.

especialistas tem 15 comparações, na subamostra de só residentes tem 10 comparações e na subamostra de especialistas versus residentes tem 30 comparações.

O critério de determinação de significância adotado foi o nível de 5%. A análise estatística foi processada pelo software estatístico SAS System, versão 6.11 (SAS Institute, Inc., Cary, NC, EUA).

Resultados

Quando avaliamos a reprodutibilidade intraobservador, observou-se que há concordância significativa para todos os 11 profissionais utilizando a classificação de Danis-Weber ($p \leq 0,0001$); para 9 profissionais utilizando a classificação de Lauge-Hansen, com 1 especialista ($p = 0,65$) e 1 residente ($p = 0,30$) sem reprodutibilidade entre os 2 momentos; e para 10 profissionais utilizando a classificação AO, com um especialista ($p = 0,071$) sem reprodutibilidade entre os 2 momentos. De uma forma geral, nota-se que os residentes apresentaram melhores concordâncias intraobservador do que os especialistas.

Quando avaliamos a reprodutibilidade interobservador, observou-se que existe concordância altamente significativa ($p < 0,0001$), de grau moderado a ótimo, na classificação de Danis-Weber para todos profissionais, nos 2 momentos. Existe correlação altamente significativa ($p \leq 0,0001$) na maioria dos pares, na classificação de Lauge-Hansen, porém, 7 pares do momento M1 e 3 pares do momento M2 não mostraram significância estatística. Existe concordância altamente significativa ($p \leq 0,0001$) na maioria dos pares, na classificação AO, porém 7 pares do momento M1 não mostraram relevância estatística.

Avaliou-se também se existe correlação do grau de concordância interobservador entre os instrumentos na amostra total e subamostras de profissionais. A **Tabela 1** fornece a média e desvio padrão (DP) do grau de concordância interobservador (coeficiente Kappa ponderado), segundo os três instrumentos na amostra total e subamostras de profissionais para os momentos M1 e M2. Notamos que o instrumento Danis-Weber apresentou, em média, grau de concordância significativamente maior que os instrumentos Lauge-Hansen e AO, na amostra total e nas subamostras de profissionais, tanto nos momentos M1 como M2. Não existe diferença significativa, ao nível de 5%, no grau de concordância entre os instrumentos Lauge-Hansen e AO na amostra total e nas subamostras de profissionais, tanto nos momentos M1 como M2.

Com relação à influência do estágio de formação dos observadores na reprodutibilidade, observou-se que, de forma geral, o subgrupo de residentes apresentou melhor concordância intraobservador, com valores superiores de κ para as 3 classificações, com diferenças estatisticamente significativas para os dois momentos avaliados ($p < 0,05$).

Discussão

Os principais achados do presente estudo estão parcialmente de acordo com nossas hipóteses iniciais. Observamos que quanto maior a complexidade da classificação das fraturas do tornozelo, menor a sua reprodutibilidade; todavia, diferentemente do que acreditávamos, os avaliadores com mais experiência tiveram menor concordância nas suas respostas em dois momentos distintos.

Em nosso estudo, solicitamos que 11 avaliadores, em diferentes estágios de formação (residentes, ortopedistas e

especialistas em cirurgia de pé e tornozelo), classificassem as fraturas de tornozelo em 30 imagens radiográficas, utilizando o método de Kappa ponderado para analisar estatisticamente os resultados. Audigé et al.¹³ realizaram uma revisão sistemática em estudos de reprodutibilidade de classificações de fraturas e concluíram que o método Kappa foi utilizado por todos, mas com muita variação na interpretação devido ao intervalo de confiança (IC). Para evitar esse viés, utilizamos o IC definido por Landsi et al. em seu trabalho.¹⁴

Fonseca et al.¹¹ avaliaram as mesmas classificações para as fraturas do tornozelo (Danis-Weber, Lauge-Hansen e AO), com 6 avaliadores e 83 imagens; contudo, somente levaram em consideração as incidências em AP e perfil. Nesse estudo, houve maior reprodutibilidade na classificação de Danis-Webber ($\kappa = 0.49$), com taxas inferiores nas classificações de Lauge-Hansen ($\kappa = 0.32$) e AO ($\kappa = 0.38$), todas apresentando concordância considerada baixa. Resultado semelhante foi encontrado por Alexandropoulos et al.,¹² que utilizaram três avaliadores para classificar 294 imagens de fraturas de tornozelos. Os autores reportaram concordância deficiente nas três classificações avaliadas ($\kappa = 0,327-0,408$; $0,174-0,476$; $0,397-0,483$; respectivamente Broos-Bisschop, Lauge-Hansen e AO).¹² No presente estudo, diferentemente, observou-se um grau de concordância interobservador altamente significativo em todas as classificações, com valores de concordância superiores aos estudos anteriores ($\kappa = 0,79$, $0,52$ e $0,48$, respectivamente para Danis-Weber, Lauge-Hansen e AO). Creditamos que o elevado grau de concordância obtido em nosso estudo se relaciona ao fato de termos adicionado uma incidência radiográfica a mais que estudos prévios.^{11,15,16} Certamente, o estudo radiográfico mais completo contribuiu para um diagnóstico com maior precisão, facilitando a classificação das lesões apresentadas aos participantes.

A concordância intraobservador é uma medida não muito avaliada nos estudos similares ao nosso.¹¹ Tenório et al.¹⁵ relataram que a concordância intraobservador foi moderada a alta com a classificação de Lauge-Hansen ($\kappa = 0.58$) e moderada a quase perfeita com a classificação de Danis-Weber ($\kappa = 0.76$). Em nosso estudo, avaliamos que as classificações, de uma maneira geral, apresentam concordância intraobservador significativa ($p < 0.05$), para todos os avaliadores para Danis-Weber, 9 para Lauge-Hansen e 10 para a AO, entre os 11 participantes. Creditamos tal resultado ao fato de o questionário ter sido grande, conseqüentemente, a avaliação no segundo momento foi feita com critério menor. Além deste fato, a maior complexidade das classificações de Lauge-Hansen e AO e a necessidade da compreensão do mecanismo de trauma da fratura para sua correta utilização diminuem sua reprodutibilidade.^{11,15}

Um dos objetivos do presente trabalho era avaliar a influência de estágios de conhecimento diferentes na atividade prática. Esperava-se que à medida que os observadores estudassem e se acostumassem a usar um determinado sistema de classificação, a concordância entre eles e dentro de suas próprias observações aumentasse.⁵ Fonseca et al.¹¹

reportaram que tal variável não influenciou as taxas de reprodutibilidade das classificações estudadas. Entretanto, uma vez que os autores só realizaram a análise da concordância interobservador, a compreensão dos resultados fica parcialmente limitada. Em nosso estudo, o subgrupo de residentes apresentou melhor concordância intraobservador, contrariando o senso comum. Acreditamos que enquanto os residentes recorreram mais vezes às imagens das classificações disponibilizadas como “gabarito”, os avaliadores mais experientes classificaram as fraturas de acordo com sua recordação das classificações. Tal fato nos chama a atenção para a importância do conhecimento dos instrumentos e de seus subtipos ao utilizá-los para classificar uma fratura, uma vez que esta muitas vezes auxilia na tomada de decisão do tratamento cirúrgico.¹⁷

Com base nos resultados do presente estudo, concluímos que um estudo radiográfico completo, incluindo as incidências em AP, perfil e AP verdadeiro do tornozelo, é fundamental para a classificação das fraturas do tornozelo, assim como o conhecimento detalhado do instrumento utilizado, sendo por vezes necessária a utilização de templates. Dentre as classificações avaliadas, apesar da classificação de Danis-Weber ter se mostrado a mais reprodutível, ela fornece informações insuficientes para guiar o tratamento das fraturas, sendo necessária uma avaliação adicional da estabilidade articular do tornozelo para correta indicação cirúrgica. Acreditamos que ainda não há uma classificação radiográfica ideal para as fraturas maleolares, com elevada reprodutibilidade e que, ao mesmo tempo, possibilite o correto planejamento cirúrgico. Assim, nas fraturas mais complexas, a avaliação pré-operatória através da tomografia computadorizada auxilia a compreensão da lesão, principalmente as trimaleolares, devido ao fragmento do maléolo posterior.¹⁸ Black et al. mostraram em seu trabalho que nas fraturas-luxação, fraturas trimaleolares e suprasindesmais, a tomografia computadorizada (CT) tem papel importante em melhorar o estudo pré-operatório e o planejamento cirúrgico.¹⁹

Estamos cientes das limitações do nosso estudo e, como principais, listamos: a quantidade de imagens radiográficas avaliadas inferior ao número observado em trabalhos similares. Existem vários artigos na literatura avaliando a reprodutibilidade das diversas classificações das fraturas do corpo humano na tentativa de definir qual é a melhor.^{10,20} Todavia, ainda não há um consenso sobre a metodologia ideal, uma vez que o número de exames de imagens analisados e a quantidade de avaliadores influenciam na concordância das respostas.^{13,20} Um número muito pequeno ou muito grande diminui a concordância.¹⁴ Tenório et al.¹⁵ utilizaram um total de 50 radiografias e 8 observadores, enquanto em outro estudo¹¹ 6 avaliadores classificaram 83 radiografias. Aumentamos o número de observadores para aumentar o poder estatístico da concordância interobservador. Além disso, o intervalo de 1 mês entre os 2 momentos de aplicação do questionário distinguiu da maioria dos estudos prévios que incluem análise intraobservador. Tal fato pode ter diminuído a concordância, já que a resposta pela memória fica prejudicada,

entretanto, desta forma, conseguimos alcançar o objetivo do trabalho de avaliar de forma isolada cada resposta dos observadores.

Conclusão

A classificação de Danis-Weber se mostrou mais reprodutiva em relação às de Lauge-Hansen e AO, com grau de concordância de moderado a alto tanto intra- quanto interobservador. Já as classificações de Lauge-Hansen e AO demonstraram semelhança entre si com concordância fraca a moderada tanto intra- quanto interobservador.

Além disso, o subgrupo de residentes apresentou concordância intraobservador mais elevada em todas as classificações, demonstrando que a pouca experiência do avaliador não influencia negativamente a reprodutibilidade das classificações das fraturas do tornozelo.

Suporte Financeiro

Não houve suporte financeiro de fontes públicas, comerciais, ou sem fins lucrativos.

Conflito de Interesses

Os autores declaram não haver conflito de interesses.

Referências

- 1 Marsh JL, Saltzman CL. Ankle Fractures. In: Bucholz RW, Heckman JD, Court-Brown CM, Charles A, editors. *Rockwood & Green's Fractures in Adults*. Philadelphia: Lippincott Williams & Wilkins; 2006:2147–2247
- 2 King CM, Hamilton GA, Cobb M, Carpenter D, Ford LA. Association between ankle fractures and obesity. *J Foot Ankle Surg* 2012;51(05):543–547
- 3 Budny AM, Young BA. Analysis of radiographic classifications for rotational ankle fractures. *Clin Podiatr Med Surg* 2008;25(02):139–152, v
- 4 Sinizio H. Xavier. *Ortopedia e traumatologia: princípios e práticas*. Porto Alegre: Artmed; 2009
- 5 Belloti JC, Tamaoki MJ, Franciozi CE, et al. Are distal radius fracture classifications reproducible? Intra and interobserver agreement. *Sao Paulo Med J* 2008;126(03):180–185
- 6 Alla SR, Deal ND, Dempsey IJ. Current concepts: mallet finger. *Hand (N Y)* 2014;9(02):138–144
- 7 Hahn DM, Colton CL. Malleolar fractures. In: Rüedi TP, Murphy WM, editors. *AO Principles of fracture management*. New York: Thieme Stuttgart; 2001:559–582
- 8 Tartaglione JP, Rosenbaum AJ, Abousayed M, DiPrea JA. Classifications in brief: Lauge-Hansen classification of ankle fractures. *Clin Orthop Relat Res* 2015;473(10):3323–3328
- 9 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(02):E71–E84
- 10 Randsborg PH, Sivertsen EA. Classification of distal radius fractures in children: good inter- and intraobserver reliability, which improves with clinical experience. *BMC Musculoskelet Disord* 2012;13:6
- 11 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* 2017;53(01):101–106
- 12 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(04):521–525
- 13 Audigé L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand* 2004;75(02):184–194
- 14 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(01):159–174
- 15 Tenório RB, Mattos CA, Araújo LH, Belangero WD. Análise da reprodutibilidade das classificações de Lauge-Hansen e Danis-Weber para fraturas de tornozelo. *Rev Bras Ortop* 2001;36(11/12):434–437
- 16 Martin JS, Marsh JL. Current classification of fractures. Rationale and utility. *Radiol Clin North Am* 1997;35(03):491–506
- 17 Brage ME, Rockett M, Vraney R, Anderson R, Toledano A. Ankle fracture classification: a comparison of reliability of three X-ray views versus two. *Foot Ankle Int* 1998;19(08):555–562
- 18 Kumar A, Mishra P, Tandon A, Arora R, Chadha M. Effect of CT on Management Plan in Malleolar Ankle Fractures. *Foot Ankle Int* 2018;39(01):59–66
- 19 Black EM, Antoci V, Lee JT, et al. Role of preoperative computed tomography scans in operative planning for malleolar ankle fractures. *Foot Ankle Int* 2013;34(05):697–704
- 20 Berger AJ, Momeni A, Ladd AL. Intra- and interobserver reliability of the Eaton classification for trapeziometacarpal arthritis: a systematic review. *Clin Orthop Relat Res* 2014;472(04):1155–1159