

Avaliação da reprodutibilidade intra- e interobservador da nova classificação AO/OTA para fraturas distais do rádio comparada com a classificação de Fernandez

Evaluation of Intra- and Interobserver Reproducibility of the New AO/OTA Classification for Distal Radius Fractures Compared with the Fernandez Classification

Beatriz Canhoto Carula^{1,2} Giana Silveira Giotri¹ Fabiana Beber¹
Marianna Fergutz dos S. Batista² Flaviana Busignani da Silva¹ Eduardo Murilo Novak¹

¹Hospital Universitário Cajuru, Curitiba, PR, Brasil

²Hospital Universitário Evangélico Mackenzie, Curitiba, PR, Brasil

Endereço para correspondência Beatriz Canhoto Carula, MD, Rua Desembargador Mota, 1648, Curitiba, PR, Brasil (e-mail: carula.beatriz@gmail.com).

Rev Bras Ortop 2022;57(6):917–923.

Resumo

Objetivo Avaliar a confiabilidade e a reprodutibilidade inter- e intraobservadores da nova classificação AO/OTA 2018 para fraturas distais do rádio e compará-la com o sistema classificatório de Fernandez.

Métodos Foi aplicado um questionário no software Qualtrics em 10 especialistas em cirurgia da mão que classificaram 50 radiografias de fraturas distais de rádio de acordo com as classificações de Fernandez e AO/OTA 2018 e, posteriormente, indicaram seu tratamento. Esse questionário foi aplicado em tempo *T0* e repetido após 4 semanas (*t1*). Analisou-se a média de concordância entre as respostas e confiabilidade e reprodutibilidade inter- e intraobservadores utilizando os índices kappa.

Resultados A concordância média interobservador para a classificação de Fernandez foi de 76,4, e de 59,2% para a AO/OTA 2018. A concordância intraobservador foi de 77,3 e 56,6%, respectivamente. O índice de kappa inter- e intraobservador para a classificação de Fernandez foram de 0,57 e de 0,55, respectivamente, e a classificação AO/OTA 2018 obteve 0,34 e 0,31, respectivamente.

Conclusão A classificação AO/OTA 2018 mostrou uma reprodutibilidade intra- e interobservadores baixa quando comparada à classificação de Fernandez. Porém, ambas as classificações apresentam índices intra- e interobservadores baixos. Embora

Palavras-chave

- ▶ fraturas do rádio/classificação
- ▶ traumatismos do punho
- ▶ reprodutibilidade dos testes
- ▶ inquéritos e questionários

Trabalho desenvolvido no Hospital Universitário Cajuru, Curitiba, PR, Brasil

recebido
23 de Abril de 2021
aceito
13 de Agosto de 2021
Publicado on-line
Setembro 26, 2022

DOI <https://doi.org/10.1055/s-0041-1740471>.
ISSN 0102-3616.

© 2022. 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

Abstract

Keywords

- ▶ radius fractures/ classification
- ▶ wrist injuries
- ▶ reproducibility of results
- ▶ surveys and questionnaires

a classificação de Fernandez não tenha obtido resultados excelentes, ela permanece com melhor concordância para o uso rotineiro.

Objective To evaluate the inter- and intraobserver reliability and reproducibility of the new AO/OTA 2018 classification for distal radius fractures and to compare it with the Fernandez classification system.

Method A questionnaire was applied in the Qualtrics software on 10 specialists in hand surgery who classified 50 radiographs of distal radius fractures according to the Fernandez and AO/OTA 2018 classifications and, subsequently, indicated their treatment. The questionnaire was applied in time *T0* and repeated after 4 weeks (*t1*). The mean agreement between the answers, and the reliability and inter- and intraobserver reproducibility were analyzed using kappa indexes.

Results The mean interobserver agreement in the Fernandez classification was 76.4, and it was 59.2% in the AO/OTA 2018 classification. The intraobserver agreements were 77.3 and 56.6%, respectively. The inter- and intraobserver kappa indexes for the Fernandez classification were 0.57 and 0.55, respectively, and, in the AO/OTA 2018 classification, they were 0.34 and 0.31, respectively.

Conclusion The AO/OTA 2018 classification showed a low intra- and interobserver reproducibility when compared with the Fernandez classification. However, both classifications have low intra- and interobserver indexes. Although the Fernandez classification did not obtain excellent results, it remains with better agreement for routine use.

Introdução

A fratura distal do rádio (FDR) é uma das fraturas mais comuns, representando 12% de todas as fraturas na população brasileira, associadas ou não a fraturas da ulna,¹⁻⁵ e é considerada um problema de saúde pública por atingir homens jovens por traumas de alta energia e idosos pela fragilidade óssea.⁵⁻⁸ Na maioria dos casos, para se obter o diagnóstico correto, radiografias do punho em incidências posteroanterior (PA), lateral (L) e oblíquas⁸ são suficientes para estabelecer o tratamento apropriado sem onerar excessivamente o sistema de saúde.^{5,8}

Para que a avaliação da radiografia seja clinicamente útil, um sistema de classificação deve ser abrangente e simples, além de ter confiabilidade intraobservadores e reprodutibilidade interobservadores.^{2,9-12} Em 1967, foi publicada a icônica classificação de Frykman, baseada em características simples da anatomia radiográfica. Posteriormente, seguiu-se uma série de sistemas classificatórios para FDR, entre os quais citamos as classificações propostas por Melone,¹³ Fernández,¹⁴ Universal (Cooney, 1993)¹⁵ e grupo AO (2007),¹⁶ que ordenam basicamente as características radiográficas destas lesões.¹⁰

Atualmente, não existe consenso quanto a qual seria a melhor classificação, devido singularmente às particularidades da FDR.^{6,17-19} A classificação AO/OTA (2007) é muito difundida entre especialistas e, talvez, a mais citada em artigos na literatura. De fácil utilização, ordena as possibilidades mais comuns de FDR sem relacionar com o mecanismo de trauma ou ofertar ideia prognóstica das lesões. Neste sentido, a classificação proposta por Fernandez traz elementos que estabelecem esta ligação com o trauma inicial e o prognóstico, tendendo a identificar mais amplamente a FDR;

contudo, parece apresentar baixa reprodutibilidade na comunicação entre os especialistas.⁴

Em 2018, o grupo AO/OTA efetuou uma atualização da sua classificação com a adição de qualificadores e modificadores em cada subtipo com o propósito de ofertar mais possibilidades na identificação das FDRs.^{9,20} Esta nova classificação AO ficou mais completa, porém aumentou a complexidade na aplicação.²¹

O propósito do presente estudo é avaliar a reprodutibilidade da classificação AO/OTA 2018 entre cirurgiões experientes e compará-la com a classificação de Fernandez, a já utilizada pelo grupo dos autores. O presente estudo também tem como objetivo avaliar a influência do uso destas classificações na tomada de decisão para o tratamento das FDRs.

Materiais e Método

Foi utilizado um questionário, aplicado por meio do aplicativo *online Qualtrics*, contendo perguntas sobre 50 radiografias de FDRs a serem analisadas utilizando 2 diferentes métodos, o de Fernandez e o AO/OTA 2018, para classificar radiografias e escolher o tratamento para cada uma.

As imagens radiográficas foram tomadas retrospectivamente de prontuários de um hospital referência em trauma. Elas eram digitais, tinham boa resolução e estavam padronizadas nas incidências anteroposterior (AP), perfil, oblíqua pronada e oblíqua supinada de fraturas agudas do terço distal do rádio (com ou sem fratura da ulna associada), em pacientes > 18 anos com esqueleto maduro, entre outubro de 2018 e março de 2020. Foram excluídas radiografias de pacientes < 18 anos, fraturas prévias e radiografias de má qualidade.

As imagens foram selecionadas aleatoriamente por ortopedistas cirurgiões da mão, que não foram avaliadores. As

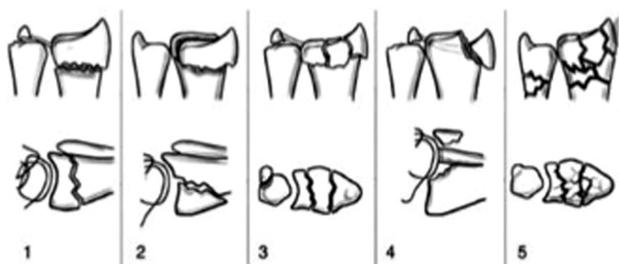


Fig. 1 Classificação Fernandez.

imagens coletadas tiveram sua identificação e data ocultadas em todo o questionário, sendo identificadas apenas com um número, e foram amostradas de maneira randomizada, para reduzir o viés dos avaliadores quanto ao teste de reprodutibilidade intraobservador.

Dez observadores especialistas em cirurgia da mão, de diferentes regiões do Brasil e com > 10 anos de formação na especialidade, foram convidados para voluntariamente avaliar as imagens. Inicialmente, 11 especialistas iniciaram o estudo, porém apenas 10 concluíram todas as etapas. Foram

disponibilizadas cópias das duas classificações para consulta. (► **Apêndice 1**, disponível apenas on-line) Cada avaliador, individualmente, respondeu um bloco de 5 questões para cada uma das 50 radiografias, sem acesso às respostas dos demais. A seguir, é apresentado um exemplo dos guias disponibilizados (► **Figuras 1 e 2**).

O questionário iniciava com o preenchimento obrigatório de um Termo de Consentimento Livre Esclarecido (TCLE) e com a identificação do avaliador. Cada uma das 50 radiografias pertencia a um bloco com 5 perguntas (► **Apêndice 2**, disponível apenas on-line) questionando qual era o tipo de fratura de acordo com a classificação de Fernandez, contendo 5 alternativas e 1 resposta única (tipos 1, 2, 3, 4 e 5). Em seguida, solicitava-se a escolha de preferência para o tratamento, também com resposta única, contendo as seguintes alternativas: (1) redução e tratamento conservador (gesso); (2) redução e pinagem percutânea; (3) redução cirúrgica e fixação com placa em “T” palmar associada ou não a fios de Kirschner; (4) redução e fixação com placa volar bloqueada; (5) redução cirúrgica e fixação com placa dorsal bloqueada; (6) outros.

2R

2R3 Segmento de extremidade distal

| | | |
|--|---|--|
| <p>2R3A1 2R3A2 2R3A3</p>  <p>2U3A Extraarticular 2UA1 Processo estiloide 2UA2 Simples 2UA3 Multifragmentar</p> | <p>2R3B1 2R3B2 2R3B3</p>  <p>2R3B Articular completo 2R3B1 Sagital 2R3B2 Borda dorsal (Barton's) 2R3B3 Aro Volar (reverso do Goyrand-Smith de Barton III)</p> | <p>2R3C1 2R3C2 2R3C3</p>  <p>2R3C Articular completo 2R3C1 Articular simples e metafisário 2R3C2 Metafisário multifragmentário 2R3C3 Articular multifragmentar, metafisário simples ou multifragmentar</p> |
|--|---|--|

2U

2U3 Segmento de extremidade distal

| | | |
|--|---|--|
| <p>2U3A1 2U3A2 2U3A3</p>  <p>2U3A Extraarticular 2UA1 Processo estiloide 2UA2 Simples 2UA3 Multifragmentar</p> | <p>2U3B</p>  <p>2U3B Parcial articular</p> | <p>2U3C</p>  <p>2U3C Articular completa</p> |
|--|---|--|

Fig. 2 Classificação AO/ OTA 2018.

A questão seguinte continha a mesma radiografia, agora para ser classificada de acordo com a classificação AO/OTA 2018, com resposta única entre 2R3A1, 2R3A2, 2R3A3, 2R3B1, 2R3B2, 2R3B3, 2R3C1, 2R3C2 e 2R3C3; o avaliador também era solicitado a responder quanto a fratura da ulna, caso presente, nas opções: 2U3A1, 2U3A2, 2U3A3, 2U3B e 2U3C ou “não se aplica”.

Na sequência, perguntava-se sobre os modificadores da classificação AO/OTA 2018, com as seguintes opções: (0) não se aplica; (1) não desviada; (2) desviada, (3a) impactação articular; (3b) impactação metafisária; (4) não impactada; (5a) desviada anterior (volar); (5b) desviada posterior (dorsal); (5c) desviada ulnar; (5d) desviada radial; (5e) desviada multidirecional; (6a) subluxação – instabilidade ligamentar volar; (6b) subluxação – instabilidade ligamentar dorsal; (6c) subluxação – instabilidade ligamentar ulnar; (6d) subluxação – instabilidade ligamentar radial; (6e) subluxação – instabilidade ligamentar multidirecional; (7) extensão diafisária; e (8) baixa qualidade óssea. Nesta questão, o avaliador poderia escolher mais de uma alternativa. Foram consideradas a quantidade de opções assinaladas para análise posterior: (1) apenas uma opção de modificador; (2) duas opções de modificadores; (3) três opções de modificadores; (4) quatro ou mais opções de modificadores assinaladas.

A última questão de cada bloco de radiografias interrogava se, após classificar esta mesma fratura utilizando o método AO/OTA 2018, o avaliador mudaria ou manteria seu tratamento inicial escolhido após classificar a fratura da radiografia em questão de acordo com a classificação de Fernandez.

A primeira aplicação do questionário ocorreu de maneira concomitante aos 10 avaliadores, considerada como tempo t_0 . Após 4 semanas, os avaliadores responderam novamente ao questionário, com as mesmas 50 radiografias em ordem diferente da anterior, em tempo nominado t_1 .

Foram ponderadas a média de concordância entre os observadores nas respostas de cada questão isoladamente e se a mudança das classificações implicou mudança de conduta e tratamento na análise interobservador, nos tempos t_0 e t_1 . Definiu-se como índice excelente se $> 75\%$ dos participantes concordassem com a mesma resposta, satisfatório se entre 50 e 75%, e insatisfatório se $< 50\%$.²²

Para avaliar a confiabilidade e a reprodutibilidade das classificações e comparar sua aplicabilidade, testou-se a reprodutibilidade interobservador, que analisou a concordância entre os 10 avaliadores acerca de uma mesma fratura em relação às classificações escolhidas, comparando-se as respostas de todas as perguntas de todos os observadores nos dois ciclos (t_0 e t_1).

A reprodutibilidade intraobservador foi testada comparando-se o nível de concordância de um mesmo observador ao responder às mesmas questões em dois tempos diferentes (tempos t_0 e t_1). Considerando que um avaliador se ausentou do questionário no tempo t_1 , tivemos a análise comparativa de apenas nove examinadores nesta segunda etapa.

A avaliação da consistência das respostas inter- e intraobservador foi empregada pelo uso de dois parâmetros: a proporção de concordância e o índice *kappa*. A primeira é a porcentagem média de casos nos quais os avaliadores con-

Tabela 1 Interpretação de Landis et al. para valores de kappa²³

| Interpretação de Landis et al. para valores de <i>kappa</i> | |
|---|------------------------------|
| Valores de <i>kappa</i> | Interpretação |
| < 0 | Sem concordância |
| 0–0,19 | Concordância ruim |
| 0,2–0,39 | Concordância baixa |
| 0,4–0,59 | Concordância moderada |
| 0,6–0,79 | Concordância substancial/boa |
| 0,8–1,0 | Concordância excelente |

cordaram. O segundo avalia a concordância entre os observadores e envolve o ajuste da proporção observada de concordância pela correção da proporção de concordância que surge em cada caso.^{1,7,17,23}

O cálculo dos índices kappa foi realizado em software online (*Online Kappa Calculator*/ Justus Randolph) e a análise dos dados foi realizada utilizando-se *free-marginal kappa*, haja vista que os avaliadores permaneceram livres para escolher as respostas.²⁰ Tradicionalmente, os valores do coeficiente *kappa*, interpretados por Landis et al., variam de 0 a 1, sendo 1 igual a concordância perfeita e 0 quando não há concordância, como especificado na **Tabela 1**.^{1,7,23,24}

O presente trabalho teve a aprovação do Comitê de Ética da instituição sob o número CAEE 22570419.0.0000.0020.

Resultados

Resultados

A análise das respostas do bloco inicial, referente à classificação de Fernandez, mostrou uma concordância média inter- e intraobservador, respectivamente, de 76,40 e 77,33%. No quesito manter o mesmo tratamento nos dois questionários (t_0 e t_1), a concordância foi de 62% na avaliação interobservador e 64,7% intraobservador. Na classificação AO/OTA 2018, as concordâncias inter- e intraobservador para o segmento fratura do rádio foram de 59,2 e 56,66%, respectivamente. No segmento fraturas de ulna, obteve-se 81,2 e 80,44% nas concordâncias intra- e interobservador, respectivamente. E para o segmento quantidade de modificadores utilizados, a concordância foi de 52,6 e 49,55%. Após a classificação pelo sistema AO/OTA nos segmentos rádio, ulna e modificadores, 95,4% dos avaliadores mantiveram a indicação de tratamento feita baseado na classificação de Fernandez. Considerando a concordância intraobservador para este item, 94,17% mantiveram a opção de tratamento (**Tabela 2**).

A média global de concordância interobservador foi considerada moderada para a classificação de Fernandez e baixa para a classificação AO/OTA 2018. A opção de tratamento após os avaliadores classificarem segundo os critérios de Fernandez obteve concordância baixa, e a manutenção do tratamento após a classificação AO/OTA 2018 foi considerada excelente (**Tabela 3**). O resultado da média de concordância intra- e interobservador com todos os participantes se encontra no (**Anexo 3, disponível apenas on-line**).

Tabela 2 Média de concordância intra- e interobservador de acordo com cada quesito avaliado

| Percentual da média de concordância | | | | | | |
|-------------------------------------|----------------------------|---------------------|--|---|--|-----------------------------------|
| | Classificação de Fernandez | Opção de tratamento | Classificação AO/OTA 2018 Segmento Rádio | Classificação AO/OTA 2018 segmento Ulna | Classificação AO/OTA 2018 Segmento quantidade de modificadores | Manutenção da opção de tratamento |
| Interobservador | 76,40% | 62,00% | 59,20% | 81,20% | 52,60% | 95,40% |
| Intraobservador | 77,33% | 61,77% | 56,66% | 80,44% | 49,55% | 94,17% |

Tabela 3 Interpretação dos valores de kappa interobservador

| INTERPRETAÇÃO DOS VALORES KAPPA INTEROBSERVADOR | | |
|---|--------------|--------------|
| Online Kappa Calculator | | |
| | Índice Kappa | Concordância |
| CLASSIFICAÇÃO DE FERNANDEZ | 0,57 | MODERADA |
| ESCOLHA DE TRATAMENTO | 0,33 | BAIXA |
| CLASSIFICAÇÃO AO/OTA 2018- SEGMENTO RÁDIO | 0,34 | BAIXA |
| MANUTENÇÃO DA OPÇÃO DE TRATAMENTO | 0,83 | EXCELENTE |

Na ►Tabela 4, estão demonstrados os resultados da interpretação intraobservador. Os valores de kappa mantiveram-se similares ao tempo t_0 . Observou-se consistência entre os valores de acordo percentual geral em t_0 e t_1 .

Tabela 4 Interpretação dos valores de kappa intraobservador

| Interpretação dos valores kappa intraobservador (acordo percentual geral) | | | | | |
|---|-----------------------|-----------------------|--------------|----------------------------------|----------------------------------|
| Online Kappa Calculator | | | | | |
| | Índice Kappa em t_0 | Índice Kappa em t_1 | Concordância | Acordo Percentual geral em t_0 | Acordo Percentual geral em t_1 |
| Classificação de Fernandez | 0,57 | 0,55 | Moderada | 65,28% | 63,69% |
| Escolha de tratamento | 0,33 | 0,36 | Baixa | 44,22% | 46,27% |
| Classificação AO/OTA 2018- Segmento rádio | 0,34 | 0,31 | Baixa | 41,02% | 38,61% |
| Manutenção da opção de tratamento | 0,83 | 0,81 | Excelente | 91,47% | 92,89% |

Tabela 5 Interpretação dos valores de kappa para fraturas associadas da ulna

| Interpretação dos valores de concordância da classificação AO/OTA 2018–Segmento ulnar e segmento quantidade de modificadores | | | | | |
|--|-----------------------|-----------------------|-----------------|----------------------------------|----------------------------------|
| Online Kappa Calculator | | | | | |
| | Índice kappa em t_0 | Índice kappa em t_1 | Concordância | Acordo percentual geral em t_0 | Acordo percentual geral em t_1 |
| Classificação AO/OTA 2018–Segmento ulnar e | 0,64 | 0,61 | Substancial/Boa | 69,82% | 67,67% |
| Segmento quantidade modificadores | 0,18 | 0,17 | Ruim | 34,53% | 33,56% |

Quanto à classificação AO/OTA 2018 segmento ulna, ou seja, quanto à associação de fratura de ulna, foram obtidos índices kappa de 0,64 e 0,61, nos tempos t_0 e t_1 , respectivamente, como apresentado na ►Tabela 5.

Em relação ao segmento modificadores da classificação AO/OTA 2018, levando em consideração os valores de kappa de 0,18 no tempo t_0 e de 0,17 no tempo t_1 (►Tabela 5), a concordância foi considerada ruim em relação à quantidade de modificadores utilizada para classificar cada fratura. Quando avaliamos individualmente o uso de modificadores, notou-se que o dado mais selecionado foi a opção “2; desviada”, seguida pela opção “5b; desviada posterior”, como detalhado na ►Tabela 6.

Discussão

Os dados obtidos no presente estudo mostraram uma concordância percentual intraobservador média de 77,3% com a classificação de Fernandez e de 56,6% com a AO/OTA 2018. A

Tabela 6 Porcentagem de modificadores selecionados pelos examinadores em t0 e t1

| Modificador | tempo t0 (%)* | tempo t1 (%)* | Modificador | Tempo t0 (%)* | Tempo t1 (%)* |
|----------------------------------|---------------|---------------|--|---------------|---------------|
| (0) Não se aplica | 0,45% | 0,28% | (5d) Desviada radial | 6,00% | 6,32% |
| (1) Não desviada | 1,29% | 1,19% | (5e) Desviada multidirecional | 2,51% | 1,97% |
| (2) Desviada | 25,43% | 23,33% | (6a) Subluxação - Instabilidade Ligamentar Volar | 1,16% | 1,48% |
| (3a) Impactação articular | 12,07% | 11,24% | (6b) Subluxação - Instabilidade Ligamentar Dorsal | 1,74% | 2,53% |
| (3b) Impactação metafisária | 13,23% | 13,00% | (6c) Subluxação - Instabilidade Ligamentar Ulnar | 0,96% | 0,49% |
| (4) Não impactada | 2,32% | 1,76% | (6d) Subluxação - Instabilidade Ligamentar Radial | 0,38% | 0,28% |
| (5a) Desviada anterior (Volar) | 3,80% | 4,08% | (6e) Subluxação - Instabilidade Ligamentar Multidirecional | 0,58% | 0,35% |
| (5b) Desviada posterior (Dorsal) | 19,04% | 21,71% | (7) Extensão Diafisária | 1,87% | 2,25% |
| (5c) Desviada ulnar | 1,03% | 1,41% | (8) Baixa Qualidade Óssea | 6,06% | 6,32% |

concordância percentual interobservador média foi de 76,4 e 59,2% para cada classificação, respectivamente. Esta comparação foi feita com a parte da classificação AO/OTA 2018 que avalia o segmento rádio, pois este sistema separa em segmentos para avaliação isolada de rádio e ulna. Analisando o índice *kappa* para reprodutibilidade inter- e intraobservador, a classificação de Fernandez resultou em concordância moderada (0,57 e 0,59, respectivamente) e a classificação AO/OTA 2018 em concordância baixa (0,34 e 0,31, respectivamente). Os mesmos resultados foram encontrados por Naqvi et al.⁴ em relação à classificação de Fernandez, demonstrando concordância moderada na avaliação de 25 radiografias por especialistas, e no trabalho de Van Leerdam et al.²⁵ a respeito da classificação AO/OTA, com baixa concordância.

Corroborando os resultados do nosso estudo, Yinjie et al.²¹ publicaram uma comparação intra- e interobservador entre as classificações de Fernandez e AO/OTA 2018, na qual 5 cirurgiões experientes avaliaram 160 imagens radiográficas. Os autores obtiveram resultados semelhantes aos nossos: reprodutibilidade intraobservador moderada com Fernandez e baixa com AO. Eles estabeleceram que a reprodutibilidade da classificação AO diminui com o aumento de subgrupos, de modificadores e de qualificadores. Quando comparada com a classificação proposta por Waever et al., eles provaram não haver superioridade nos resultados de reprodutibilidade. Estes autores estudaram 573 radiografias de pacientes com FDR buscando apontar um sistema classificatório unificado e universal.²

É importante ressaltar o aumento da concordância intraobservadores na análise entre os tempos t0 e t1. Depois que a classificação é assimilada pelo avaliador, ela tende a ser usada de forma mais reprodutível. Percebeu-se que quanto menos opções para escolha na classificação, maior será a concordância no seu uso. Na avaliação do segmento ulnar da classificação AO/OTA 2018, obteve-se ~ 80% na concordância com apenas 6 opções. Comparada com o resultado do uso de modificadores, com mais de 15 opções,

observou-se que a concordância caiu para ~ de 50%. Portanto, a concordância tende a diminuir quanto mais opções de escolha houver na classificação, como apontado anteriormente.

A classificação não parece interferir na escolha do tratamento. Após classificar de acordo com os critérios de Fernandez, os avaliadores optaram por uma condução para cada FDR apresentada. Houve manutenção desta conduta na quase totalidade dos casos (94%) após efetuarem a avaliação de acordo com o AO/OTA 2018. Nesta linha, indaga-se que conhecimento e qual experiência pessoal tendem a ser preferíveis ao uso de sistemas classificatórios na predição de prognóstico e na tomada de decisão do tratamento das FDRs.²⁶ Em um estudo multicêntrico, Mulders et al. enfatizaram que, dificilmente, haverá consenso no tratamento destas fraturas caso se seja guiado especificamente por sistemas classificatórios, pois os cirurgiões tenderão sempre a estratégias baseadas em suas experiências.^{9,27,28}

Outro ponto a ser discutido é a familiaridade com a classificação proposta. O sistema de Fernandez vem sendo utilizado desde o início dos anos 2000, podendo explicar os resultados mais reprodutíveis. Talvez, o sistema AO/OTA 2018 proposto, por maior riqueza de detalhes e mais opções para identificação da FDR, aumente sua reprodutibilidade com o passar do tempo e a disseminação do seu uso. Deve-se salientar que nenhum dos vários sistemas classificatórios para FDR apresenta reprodutibilidade alta.²¹

Encontramos várias limitações no presente estudo. Destacamos a aplicação livre do questionário, o que pode evidenciar um viés de resposta. Ressaltamos também o tempo ilimitado concedido para responder o questionário, o qual permitia pausar a pesquisa e retomá-la após o prazo de 1 semana. Outra limitação foi a avaliação apenas quantitativa e não qualitativa dos modificadores da classificação AO/OTA 2018, a qual leva a mais detalhes e dificuldade para classificar. Ademais, avaliamos cirurgiões da mão em apenas um nível de formação, embora experientes. Outra possível limitação foi não avaliar com qual classificação o examinador

estava habituado, pois quanto maior a familiaridade, maior a reprodutibilidade no seu uso. Idealmente, sugere-se pesquisas em maior escala, com avaliadores de diferentes níveis de experiência, com um maior número amostral analisado em um mesmo período, além de com o agrupamento de múltiplos subtipos, de modo a, possivelmente, aumentar a consistência dos resultados obtidos.

Conclusão

No presente estudo, as classificações estudadas não apresentaram concordância alta na reprodutibilidade inter- e intraobservador. Sugere-se que complexidade e o detalhamento da nova classificação AO/OTA 2018 seja a causa da reprodutibilidade baixa quando comparada com a do sistema proposto por Fernandez.

Suporte Financeiro

A presente pesquisa não recebeu nenhum financiamento específico de agências de financiamento dos setores público, comercial ou sem fins lucrativos.

Conflito de Interesses

Os autores declaram não haver conflito de interesses.

Referências

- 1 Tenório PHM, Vieira MM, Alberti A, Abreu MFM, Nakamoto JC, Cliquet A. Evaluation of intra- and interobserver reliability of the AO classification for wrist fractures. *Rev Bras Ortop* 2018;53(06):703–706
- 2 Wæver D, Madsen ML, Rölfing JHD, et al. Distal radius fractures are difficult to classify. *Injury* 2018;49(Suppl 1):S29–S32
- 3 Plant CE, Hickson C, Hedley H, Parsons NR, Costa ML. Is it time to revisit the AO classification of fractures of the distal radius? Inter- and intra-observer reliability of the AO classification. *Bone Joint J* 2015;97-B(06):818–823
- 4 Naqvi SG, Reynolds T, Kitsis C. Interobserver reliability and intraobserver reproducibility of the Fernandez classification for distal radius fractures. *J Hand Surg Eur Vol* 2009;34(04):483–485
- 5 Alffram PA, Bauer GC. Epidemiology of fractures of the forearm. A biomechanical investigation of bone strength. *J Bone Joint Surg Am* 1962;44-A:105–114
- 6 Brogan DM, Richard MJ, Ruch D, Kakar S. Management of Severely Comminuted Distal Radius Fractures. *J Hand Surg Am* 2015;40(09):1905–1914
- 7 Shehovych A, Salar O, Meyer C, Ford DJ. Adult distal radius fractures classification systems: essential clinical knowledge or abstract memory testing? *Ann R Coll Surg Engl* 2016;98(08):525–531
- 8 Arora R, Gabl M, Gschwentner M, Deml C, Krappinger D, Lutz M. A comparative study of clinical and radiologic outcomes of unstable colles type distal radius fractures in patients older than 70 years: nonoperative treatment versus volar locking plating. *J Orthop Trauma* 2009;23(04):237–242
- 9 Meinberg EG, Agel J, Roberts CS, Karam MD, Kellam JF. Fracture and Dislocation Classification Compendium-2018. *J Orthop Trauma* 2018;32(Suppl 1):S1–S170
- 10 Kanakaris NK, Lasanianos NG. Distal Radial Fractures. In: Lasanianos NG, Kanakaris NK, Giannoudis PV, editors. *Trauma and Orthopaedic Classifications*. New York: Springer; 2014:95–105
- 11 Jayakumar P, Teunis T, Giménez BB, Verstreken F, Di Mascio L, Jupiter JB. AO Distal Radius Fracture Classification: Global Perspective on Observer Agreement. *J Wrist Surg* 2017;6(01):46–53
- 12 Lee DY, Park YJ, Park JS. A Meta-analysis of Studies of Volar Locking Plate Fixation of Distal Radius Fractures: Conventional versus Minimally Invasive Plate Osteosynthesis. *Clin Orthop Surg* 2019;11(02):208–219
- 13 Melone CP Jr. Articular fractures of the distal radius. *Orthop Clin North Am* 1984;15(02):217–236
- 14 Fernandez DL. Distal radius fracture: the rationale of a classification. *Chir Main* 2001;20(06):411–425
- 15 Cooney WP. Fractures of the distal radius. A modern treatment-based classification. *Orthop Clin North Am* 1993;24(02):211–216
- 16 Marsh JL, Slongo TF, Agel J, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma* 2007;21(10, Suppl):S1–S133
- 17 Illarramendi A, González Della Valle A, Segal E, De Carli P, Maignon G, Gallucci G. Evaluation of simplified Frykman and AO classifications of fractures of the distal radius. Assessment of interobserver and intraobserver agreement. *Int Orthop* 1998;22(02):111–115
- 18 Koval K, Haidukewych GJ, Service B, Zircibel BJ. Controversies in the management of distal radius fractures. *J Am Acad Orthop Surg* 2014;22(09):566–575
- 19 Thurston AJ. 'Ao' or eponyms: the classification of wrist fractures. *ANZ J Surg* 2005;75(05):347–355
- 20 Porrino JA Jr, Maloney E, Scherer K, Mulcahy H, Ha AS, Allan C. Fracture of the distal radius: epidemiology and premanagement radiographic characterization. *AJR Am J Roentgenol* 2014;203(03):551–559
- 21 Yinjie Y, Gen W, Hongbo W, et al. A retrospective evaluation of reliability and reproducibility of Arbeitsgemeinschaft für Osteosynthesefragen classification and Fernandez classification for distal radius fracture. *Medicine (Baltimore)* 2020;99(02):e18508
- 22 Martin JS, Marsh JL, Bonar SK, DeCoster TA, Found EM, Brandser EA. Assessment of the AO/ASIF fracture classification for the distal tibia. *J Orthop Trauma* 1997;11(07):477–483
- 23 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(01):159–174
- 24 Randolph JJ. Online Kappa Calculator [Computer software]. 2008 Available from: <http://justus.randolph.name/kappa>
- 25 van Leerdam RH, Souer JS, Lindenhovius AL, Ring DC. Agreement between Initial Classification and Subsequent Reclassification of Fractures of the Distal Radius in a Prospective Cohort Study. *Hand (N Y)* 2010;5(01):68–71
- 26 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
- 27 Kleinlugtenbelt YV, Groen SR, Ham SJ, et al. Classification systems for distal radius fractures. *Acta Orthop* 2017;88(06):681–687
- 28 Mulders MA, Rikli D, Goslings JC, Schep NW. Classification and treatment of distal radius fractures: a survey among orthopaedic trauma surgeons and residents. *Eur J Trauma Emerg Surg* 2017;43(02):239–248