

Concordância intra/interobservador dos sistemas de classificação das fraturas de úmero proximal

Douglas Barbosa Lima¹  Alexia Ribas Checchi²  Igor Fiorese Vieira²  José Luiz Lopes Vieira³ 
Giulia Maria Jorge Dinies¹  Renata Mello Calandrini¹  João Vitor Vedam¹ 

¹Hospital Regional de Ponta Grossa. Ponta Grossa/PR, Brasil.

²Centro Universitário Campo Real. Guarapuava/PR, Brasil.

³Universidad Católica del Maule – UCM. Talca, Chile.

E-mail: jllopesvieira@gmail.com

Resumo Gráfico

Highlights

- A fratura do úmero proximal é a quarta fratura mais comum em adultos.
- O método de classificação da fratura define o tratamento e o prognóstico do tratamento.
- Especialista em membros superiores é consistente na avaliação das fraturas do úmero proximal.
- Programas de residência são fundamentais na qualidade da avaliação de fraturas do úmero proximal.



Resumo

Este estudo avaliou a concordância inter e intraobservador dos sistemas de classificações NEER, LEGO/HERTEL, AO/OTA, e MAYO-FJD das fraturas de úmero proximal. Cinco avaliadores sendo, um residente no primeiro ano (R1), segundo ano (R2), terceiro ano (R3), um ortopedista (Orto) e um especialista (Esp) em cirurgia de ombro avaliaram, em dois momentos (T1 e T2), 50 radiografias de pacientes com fratura do úmero proximal. Na análise estatística foi utilizado o coeficiente Kappa de Cohen e Fleiss. Na concordância interobservadores o sistema NEER T1 e T2 ($\kappa=0,444$ e $\kappa=0,496$) apresentou concordância moderada. Os sistemas LEGO/HERTEL T1 e T2 ($\kappa=0,376$ e $\kappa=0,387$), o sistema AO/OTA apresentou T1 e T2 ($\kappa=0,269$ e $\kappa=0,223$) e, o sistema MAYO-FJD T1 e T2 ($\kappa=0,367$ e $\kappa=0,300$) apresentaram concordância baixa. Ocorreu concordância moderada intraobservador, no sistema NEER. O sistema LEGO/HERTEL apresentou concordância substancial com $\kappa=0,528$ (R1) e quase perfeita de $\kappa=0,867$ (Esp). No sistema AO/OTA o Kappa variou de 0,347 (R1) a 0,793 (R3), oscilando entre baixa e substancial. O sistema MAYO-FJD apresentou a menor concordância ($\kappa=0,154$ (R1) a $\kappa=0,665$ (Esp)). Conclui-se que na concordância interobservadores o sistema NEER, obteve resultado moderado. Na concordância intraobservador os avaliadores mais experientes apresentaram maior consistência em comparação aos residentes.

Palavras-chave: Fraturas. Radiografia. Úmero. Interobservador. Intraobservador.

Editor de área: Edison Barbieri
Mundo Saúde. 2026,50:e19382025
O Mundo da Saúde, São Paulo, SP, Brasil.
<https://revistamundodasaude.emnuvens.com.br>

Recebido: 15 dezembro 2025.

Aceito: 24 junho 2026.

Publicado: 14 julho 2026.

INTRODUÇÃO

A fratura do úmero proximal é a quarta fratura mais comum em adultos, ficando atrás apenas das fraturas do rádio distal, do fêmur proximal e das fraturas vertebrais osteoporóticas¹. Representa cerca de 80% de todas as fraturas do úmero e ocorre predominantemente em mulheres idosas com osteoporose, com uma razão aproximada de 3:1 em relação aos homens, sendo geralmente resultante de quedas da própria altura, esses pacientes apresentam ainda aumento do risco de fraturas subsequentes^{2,3}.

Para o manejo adequado da fratura, é essencial que ela seja classificada por meio de radiografias, pois define conduta, tratamento e prognóstico do paciente. No entanto, os sistemas de classificação de fraturas atualmente utilizados para a extremidade proximal do úmero demonstram baixa concordância e reprodutibilidade⁴.

Em 1970, Charles S. Neer II propôs um sistema para classificar as fraturas do úmero proximal com base no número de segmentos ósseos deslocados. O modelo divide o úmero proximal em quatro segmentos anatômicos: cabeça, tubérculo maior, tubérculo menor e diáfise, considerando como “parte fraturada” apenas aquele segmento com deslocamento ≥ 1 cm ou angulação $\geq 45^\circ$, exceto tuberosidade maior a qual considera-se desvio de 0,5cm. Assim, as fraturas podem ser descritas como de duas, três ou quatro partes, incluindo variantes como fraturas com luxação ou do colo anatômico. A classificação de NEER é amplamente utilizada até hoje, servindo de referência para a escolha do tratamento e a avaliação prognóstica⁵.

O sistema LEGO/HERTEL descreve a morfologia das fraturas do úmero proximal em até 12 padrões combinatórios, baseados nas quatro estruturas anatômicas principais (cabeça, tubérculo maior, tubérculo menor e diáfise). Essa abordagem busca padronizar a descrição das fraturas, facilitando a comunicação e a reprodutibilidade entre avaliadores. Hertel *et al.* (2004)⁶ também propuseram critérios prognósticos independentes para estimar o risco de isquemia da cabeça umeral, que incluem a avaliação do calcar medial, da continuidade metafisária e do contato cabeça-diáfise. Assim, enquanto o sistema LEGO é uma classificação morfológica, os chamados “critérios de Hertel” correspondem a preditores clínicos de necrose avascular e não fazem parte formal da classificação.

Por outro lado, o sistema de classificação AO/OTA refere-se a um sistema universal de classificação de fraturas criado pela AO Foundation (Associação para o Estudo da Osteossíntese) em conjun-

to com a OTA (Associação Ortopédica de Trauma) que divide as fraturas em três grupos distintos, sendo cada grupo dividido em três outros subgrupos, totalizando nove tipos de fratura.

As fraturas de duas partes (A) podem acometer as tuberosidades ou o colo cirúrgico. Quando envolvem as tuberosidades, o subtipo A1.1 acomete a grande tuberosidade (GT), enquanto o subtipo A1.2 acomete a pequena tuberosidade (LT). No caso do colo cirúrgico, as fraturas podem ser simples (2.1), em cunha (2.2) ou cisalhamento/verticais (2.3). As fraturas de três partes (B) geralmente envolvem o colo cirúrgico associado a uma tuberosidade. O subtipo B1.1 acomete a grande tuberosidade (GT) e o B1.2 acomete a pequena tuberosidade (LT). Esses subtipos podem ser ainda qualificados quanto à presença de cunha intacta (u) ou cominutiva (v), indicando o grau de fragmentação da fratura.

As fraturas articulares (C) envolvem o colo anatômico, podendo ocorrer de forma isolada ou em associação com o colo cirúrgico. No subtipo C1, o colo anatômico pode sofrer impactação em valgo, sendo qualificado de acordo com a tuberosidade afetada: $n \rightarrow GT$, $o \rightarrow LT$, $p \rightarrow LT + GT$, ou pode ocorrer como colo anatômico isolado (1.3). Nas fraturas de quatro partes (C3), que envolvem colo anatômico e colo cirúrgico, a metáfise e a articulação podem estar intactas (3.1), a metáfise intacta com a articulação cominuída (3.2), ou ambas cominutivas (3.3), representando padrões mais complexos e instáveis⁷.

Em fraturas com grau de complexidade maior podemos utilizar a classificação de MAYO-Fundación Jiménez Díaz (FJD), este sistema utiliza três componentes anatômicos principais, cabeça do úmero, colo do úmero e tubérculos. A classificação contempla padrões de fratura, fratura isolada do tubérculo maior e menor, fratura cirúrgica do pescoço, fratura póstero-medial em varo, fratura impactada em valgo e fraturas onde a cabeça do úmero é deslocada, as quais englobam luxação da cabeça, cabeça dividida e impressão da cabeça⁸.

Sendo a fratura de úmero proximal a quarta mais comum, é necessário padronizar sua classificação dentro da prática médica, visando um plano de tratamento eficaz para cada tipo específico de lesão, hoje dentro da literatura não existem trabalhos relevantes que utilizem todas essas classificações, portanto é necessário realizar uma avaliação inter e intraobservacional das classificações para identificar a mais adequada no manejo da fratura de úmero proximal, principalmente, entre profissionais com diferentes níveis de formação em ortopedia.

MÉTODOS

Este é um estudo observacional retrospectivo descritivo com uma abordagem transversal, que por meio de diferentes sistemas de classificação da fratura de úmero proximal visa avaliar a concordância interobservador. Após a concordância da instituição coparticipante e aprovação do Comitê de Ética em Pesquisa número 7.717.872. O estudo foi realizado com radiografias já existentes nos prontuários de pacientes do Hospital Universitário Regional dos Campos Gerais (HU-UEPG) no período entre julho de 2023 e janeiro de 2024. Como critérios de inclusão foram utilizadas radiografias de pacientes diagnosticados com fratura de úmero proximal, CID S42.2, em duas incidências: anteroposterior e perfil e foram excluídas as radiografias de fraturas de úmero proximal em apenas uma incidência, ou qualquer outro tipo de fratura.

Desta forma, foram analisadas 50 imagens radiográficas de fraturas de úmero proximal que cobriam o espectro completo das classificações NEER, LEGO/HERTEL, AO/OTA e MAYO-FJD, incluindo pelo menos um exemplar de cada tipo de fratura de úmero proximal. Para a avaliação, os 50 casos foram apresentados a um grupo de avaliadores com diferentes estágios de formação em ortopedia: um residente no primeiro ano de formação (R1), um residente no segundo ano de formação (R2), um residente no terceiro ano de formação (R3), um ortopedista (Orto) não especialista em cirurgia de ombro e um especialista (Esp.) em cirurgia de ombro.

As radiografias foram apresentadas de forma randomizada aos avaliadores e com intervalo de tempo entre as duas avaliações foi entre 15 e 30 dias. Os

avaliadores não tiveram acesso às suas respostas da primeira rodada nem ao histórico clínico ou desfecho cirúrgico do paciente. A ordem de exibição das radiografias foi aleatorizada na segunda rodada. Os avaliadores não receberam treinamento prévio, ou seja, cada avaliador baseou sua avaliação no seu conhecimento prévio em aulas, práticas na residência e anos de prática da ortopedia.

Os resultados obtidos com as avaliações dos observadores foram tabulados utilizando o Microsoft Excel 365 (Microsoft Corp., Redmond, WA, EUA) e submetidos aos testes Kappa de Cohen e Fleiss para avaliação intra e interobservadores respectivamente, através do *software Statistical Package for the Social Sciences* (SPSS, IBM Corp., Armonk, NY, USA), versão 30.0, para análises estatísticas. Obtiveram-se as tabelas de avaliação de concordância interobservadores evidenciando a medida do índice de Kappa para cada classe de avaliadores (R1, R2, R3, Orto e Esp.) para cada uma das duas avaliações (T1 e T2).

As tabelas de avaliação de concordância intraobservador, comparando cada uma das avaliações com a outra, bem como a presença dos valores limites superior e inferior para um IC de 95%. Foram considerados significativos os valores de Kappa com $p < 0,05$. A interpretação dos resultados seguindo o entendimento proposto por Landis e Koch (1977)⁹, no qual valores menores ou iguais a zero indicavam concordância pobre, pouca entre 0 e 0,2; razoável entre 0,2 e 0,4; moderada entre 0,4 e 0,6; substancial entre 0,6 e 0,8; e, excelente ou quase perfeita entre 0,8 e 1,0.

RESULTADOS

A Tabela 1 apresenta os valores de concordância interobservadores obtidos por meio do coeficiente

Kappa de Fleiss, considerando os diferentes indicadores avaliados nos dois momentos (T1 e T2).

Tabela 1 - Determinação da concordância interobservadores entre os sistemas de classificação NEER, LEGO/HERTEL, AO/OTA, MAYO-FJD nas fraturas de úmero proximal (HU-UEPG – 07/2023 a 01/2024).

Sistema de Classificação	% Concordância	Fleiss k	IC 95%	p valor	Nível de Concordância
NEER T1	44,40	0,444	0,394-0,493	<0,001	Moderada
NEER T2	49,60	0,496	0,445-0,546	<0,001	Moderada
LEGO/HERTEL T1	37,60	0,376	0,341-0,411	<0,001	Baixa
LEGO/HERTEL T2	38,70	0,387	0,351-0,423	<0,001	Baixa
AO/OTA T1	26,90	0,269	0,242-0,296	<0,001	Baixa
AO/OTA T2	22,30	0,223	0,196-0,251	<0,001	Baixa
MAYO-FJD T1	36,70	0,367	0,342-0,392	<0,001	Baixa
MAYO-FJD T2	30,00	0,300	0,276-0,325	<0,001	Baixa

Nota: % = percentual; IC = intervalo de confiança; T1 = primeira avaliação; T2 = segunda avaliação; p = probabilidade; K = índice Kappa.

Os indicadores NEER T1 e NEER T2 apresentaram concordância moderada, com valores de Kappa de 0,444 (IC95%: 0,394–0,493; $p < 0,001$) e 0,496 (IC95%: 0,445–0,546; $p < 0,001$), respectivamente. Em contrapartida, os indicadores HERTEL-LEGO T1 ($\kappa = 0,376$; IC95%: 0,341–0,411; $p < 0,001$), HERTEL-LEGO T2 ($\kappa = 0,387$; IC95%: 0,351–0,423; $p < 0,001$), AO11 T1 ($\kappa = 0,269$; IC95%: 0,242–0,296; $p < 0,001$) e AO11 T2 ($\kappa = 0,223$; IC95%: 0,196–0,251; $p < 0,001$) demonstraram baixa concordância, evidenciando maior divergência entre os avaliadores.

De forma semelhante, os indicadores MAYO-FJD T1 e MAYO-FJD T2 apresentaram concordância baixa, com valores de Kappa de 0,367 (IC95%: 0,342–0,392; $p < 0,001$) e 0,300 (IC95%: 0,276–0,325; $p < 0,001$), respectivamente. De modo geral, os resultados revelam heterogeneidade na concordância entre os avaliadores, variando de baixa a moderada conforme o indicador analisado e o momento da avaliação. A Tabela 2 apresenta os coeficientes de Kappa de Cohen obtidos na análise da concordância interobservador (comparação entre T1 e T2) para cada observador e indicador.

Tabela 2 - Nível de concordância intraobservador (T1vsT2) nos sistemas de classificação NEER, LEGO/HERTEL, AO/OTA e MAYO-FJD nas fraturas do úmero proximal (HU-UEPG – 07/2023 a 01/2024).

Avaliador	NEER	LEGO/HERTEL	AO/OTA	MAYO-FJD
R1	0,523	0,528	0,347	0,154
R2	0,740	0,696	0,658	0,574
R3	0,876	0,866	0,793	0,349
Ortopedista (Orto)	0,777	0,745	0,683	0,620
Especialista (Esp.)	0,766	0,867	0,763	0,665

Nota: R1 = Residente de primeiro ano; R2 = Residente de segundo ano; R3 = Residente de terceiro ano.

De modo geral, os valores de Kappa indicam níveis variáveis de consistência intravaliadores entre os diferentes indicadores, variando de baixa a perfeita. Para o indicador NEER, observou-se concordância que variou de moderada a quase perfeita entre os avaliadores, com valores de Kappa entre 0,523 (R1) e 0,876 (R3). Esses resultados sugerem boa estabilidade das classificações entre os dois momentos de avaliação. Da mesma forma, o indicador HERTEL-LEGO também apresentou que variou de moderada a quase perfeita, com valores de Kappa entre 0,528 (R1) e 0,867 (Esp.).

Esses resultados sugerem boa estabilidade das classificações entre os dois momentos de avaliação para ambos os indicadores. No caso do indicador AO/

OTA, os valores de Kappa variaram de 0,347 (R1) a 0,793 (R3), indicando concordância que oscilou entre baixa e substancial, conforme o avaliador. Por fim, o indicador MAYO-FJD apresentou os menores índices de estabilidade intravaliador, com valores de Kappa variando de 0,154 (R1) a 0,665 (Esp.), refletindo baixa a moderada concordância entre as duas avaliações.

De forma geral, observou-se que os avaliadores mais experientes (ortopedista e especialista) apresentaram maior consistência intraobservador em relação aos avaliadores residentes, especialmente nos indicadores AO/OTA e MAYO-FJD. Esses achados sugerem que a experiência clínica pode estar associada a maior estabilidade nas avaliações realizadas em momentos distintos.

DISCUSSÃO

Os resultados obtidos neste estudo demonstraram variações significativas na concordância inter e intraobservador entre os diferentes sistemas de classificação de fraturas do úmero proximal, evidenciando que a experiência clínica dos avaliadores exerce papel determinante na consistência das avaliações. Observou-se que o sistema NEER apresentou os melhores índices de concordância interobservador em ambos os momentos (T1 e T2), com concordância moderada, corroborando com resultados similares¹⁰ que relataram confiabilidade razoável a substancial para as classificações

NEER e HERTEL, sendo a primeira superior em termos de estabilidade intraobservador, especialmente em reconstruções tridimensionais mesmo entre os observadores com diferentes níveis de experiência¹¹.

Em contrapartida, os sistemas LEGO/HERTEL e MAYO-FJD apresentaram concordância interobservador baixa, indicando maior divergência entre os avaliadores. Essa limitação é frequentemente relatada na literatura e atribuída à maior complexidade morfológica dessas classificações, que envolvem múltiplos parâmetros anatômicos e prognósticos, dificultando a



reprodutibilidade em análises bidimensionais de radiografias⁸.

Apesar disso, estudos mostram que o sistema MAYO-FJD pode apresentar melhor desempenho quando associado à tomografia computadorizada (TC), uma vez que a visualização multiplanar reduz as ambiguidades de interpretação⁸. Esse achado sugere que o uso exclusivo de radiografias simples, embora prático e amplamente disponível, pode limitar a acurácia diagnóstica de sistemas mais detalhados que observaram melhor concordância com o uso de reconstruções tridimensionais¹¹.

Da mesma forma, o sistema AO/OTA apresentou baixa concordância interobservador, sugerindo divergência considerável entre os avaliadores. Essa limitação é observada também em estudos prévios, que relatam menor confiabilidade interavaliadores nesse sistema, especialmente em suas versões mais antigas¹². A atualização da classificação AO/OTA publicada em 2018 aprimorou a clareza dos critérios e simplificou o processo de codificação, mantendo um poder descritivo adequado e melhorando discretamente a concordância entre os observadores¹³. De forma semelhante, apresenta reprodutibilidade variável conforme o nível de experiência dos avaliadores, reforçando a necessidade de treinamento padronizado para otimizar a consistência diagnóstica¹⁴.

Com relação à concordância intraobservador, os resultados deste estudo estiveram alinhados aos encontrados por Díaz *et al.* (2022)¹⁴, que observaram que o aumento da concordância intraobservador ocorre com o avanço da experiência profissional, destacando que ortopedistas experientes alcançam índices mais altos de confiabilidade em comparação a residentes. Essa tendência também foi observada no presente trabalho, no qual o ortopedista e o especialista em cirurgia de ombro obtiveram os maiores valores de Kappa intraobservador, especialmente nos sistemas AO/OTA e MAYO-FJD. Tais achados ratificam que o tempo de prática e o treinamento específico influenciam diretamente a capacidade de interpretar corretamente as nuances morfológicas das fraturas do úmero proximal.

Embora o sistema MAYO-FJD tenha surgido recentemente como uma proposta mais abrangente e anatômica para a classificação das fraturas proximais do úmero, os resultados deste estudo mostraram que esse sistema de classificação ainda apresenta baixa concordância interobservador. Esse achado pode estar relacionado ao fato de ser um sistema novo, proposto por Foruria *et al.* (2022)⁸, enquanto os outros sistemas NEER (1970), LEGO/HERTEL (2004) e AO/OTA (1990, revisado em 2018) já estão consolidados há mais tempo, apresentando maior familiaridade entre os profissionais e ampla difusão na prática ortopédica.

Apesar de o sistema NEER demonstrar melhor desempenho geral, seu valor clínico ainda é debatido,

uma vez que apresenta limitações em casos de fraturas complexas e em pacientes com alterações anatômicas significativas. Análises baseadas em radiografias simples, classificações como NEER e AO/OTA mantêm concordância apenas moderada entre avaliadores, o que sugere que nenhum dos sistemas atuais atinge confiabilidade ideal para padronização universal¹⁵.

Nesse sentido, revisões recentes defendem que a incorporação de métodos tridimensionais e inteligência artificial pode aumentar substancialmente a reprodutibilidade e a aplicabilidade clínica dessas classificações, reduzindo o viés relacionado à experiência individual^{16,17}.

A formação adequada durante a residência médica em ortopedia é essencial para o desenvolvimento da habilidade diagnóstica e da capacidade de interpretar corretamente as imagens radiográficas, fatores que influenciam diretamente na escolha terapêutica e nos desfechos clínicos^{11,14}. Uma avaliação radiográfica criteriosa é determinante para definir a necessidade de tratamento cirúrgico ou conservador, permitindo decisões mais precisas e personalizadas¹⁵.

Além disso, a correta interpretação das radiografias contribui para a redução de custos hospitalares e para o sucesso funcional e econômico do tratamento, favorecendo a recuperação completa do paciente⁸ e, a impressão tridimensional pode ser aplicada, principalmente, no planejamento de fraturas e no guiamento intraoperatório de sequelas¹⁷.

Embora sendo amplamente utilizadas durante a prática clínica, as classificações clássicas utilizadas para avaliação das fraturas do úmero proximal apresentam limitações significativas, estudos apontaram uma baixa concordância interobservador entre os sistemas de NEER e AO/OTA, o que tem ocasionado as incertezas na padronização direcionada as condutas terapêuticas^{5,18,19}. Em função dessas divergências, a tomada de decisão na atualidade enfatiza uma tendência para a valorização de critérios pato morfológicos específicos, sendo estes, a estabilidade do colo medial e o risco de necrose vascular⁶.

Apesar das divergências entre os sistemas de classificação a longevidade e uso contínuo da classificação NEER ocorre, principalmente, por três fatores: linguagem ortopédica comum por fornecer uma terminologia que é compreendida pelos cirurgiões ao redor do mundo²⁰; a prevenção de necrose pois foi um dos primeiros sistemas a se preocupar criticamente com a vascularização da cabeça umeral e o risco de necrose avascular²¹; e, a orientação do tratamento, aplicando a regra prática do “desvio” (fragmentos deslocados > 1 cm ou angulados > 45°) o que fornece uma diretriz de fácil aplicação para indicar o manejo conservador ou cirúrgico⁵.

Por outro lado, a classificação Mayo-FJD tem como principais vantagens potenciais: a) previsão de falha

conservadora pois agrupa os pacientes em categorias prognósticas claras, identificando com precisão os padrões que evoluirão bem clinicamente daqueles com alto risco de colapso²²; b) menor taxa de falha, auxiliando na tomada de decisão preditiva sobre quando um tratamento não-cirúrgico²² e, c) a elevada confiabilidade baseada em padrões de concordância interobservador substancial (índice Kappa entre 0,69 a 0,81) ao usar tomografia computadorizada⁸.

Finalizando, a interpretação humana de fraturas complexas do úmero proximal é frequentemente heterogênea e, uma janela de oportunidade surge com a inteligência artificial (IA) e os algoritmos de *Deep Learning* estão sendo treinados para superdiagnosticar e padronizar essa avaliação. Este novo contexto permitirá uma precisão elevada, sendo que pesquisas mostram que redes neurais convolucionais (RNCs) alcançam mais de 95% de acurácia em identificar a presença ou ausência da fratura em radiografias simples²³.

Outro avanço será a classificação de padrões complexos, pois a IA atua na identificação automatizada de fraturas de 3 e 4 partes, superando muitas vezes a acurácia de médicos generalistas e equiparando-se ao desempenho de cirurgiões especialistas em ombro²³, gerando maior consistência, pois no futuro a IA poderá remover a subjetividade humana, fornecendo padronização na leitura de imagens que auxiliando rapidamente o cirurgião na formulação do plano cirúrgico²⁴.

Já, modelos tridimensionais demonstraram, a maior concordância interobservador (especialistas *versus* residentes) na escolha do tratamento de fraturas do úmero proximal, em comparação com radiografia, tomografia e holografia. A concordância na escolha do tratamento entre os especialistas que utilizaram tomografia e holografia como métodos diagnósticos foi duas vezes maior em comparação com os residentes¹⁶.

De forma específica, utilizar sistemas de alta complexidade (MAYO-FJD e AO/OTA) sem o suporte de tomografia computadorizada e reconstruções 3D podem induz o cirurgião a um erro de classificação. Sem essas ferramentas, as classificações padecem em sua principal função, que é guiar o planejamento cirúrgico e prever de forma confiável o prognóstico do paciente¹³.

CONCLUSÃO

O presente estudo demonstrou que, entre os sistemas avaliados, a classificação NEER apresentou a melhor concordância interobservador, com resultados moderados, enquanto os sistemas LEGO/HERTEL, AO/OTA e MAYO-FJD exibiram concordância baixa. Na análise intraobservador, observou-se que os avaliadores mais experientes apresentaram maior consis-

Entre as limitações deste estudo, destaca-se a utilização de radiografias simples como método de avaliação, as quais, embora amplamente disponíveis e de baixo custo, podem apresentar variações na qualidade e no posicionamento das incidências, interferindo na interpretação e reduzindo a concordância entre os avaliadores. A ausência de tomografias tridimensionais também limitou a análise de fraturas complexas, que poderiam ser melhor caracterizadas por métodos mais detalhados.

Outra limitação diz respeito ao número reduzido de avaliadores e à inclusão de apenas um observador por grupo, o que pode ter introduzido viés decorrente da variabilidade intrasujeito e do nível de familiaridade com cada sistema de classificação. Além disso, o tempo destinado por cada avaliador à análise das imagens pode ter influenciado a precisão e a consistência dos resultados.

Como aplicações práticas os resultados deste estudo reforçam a necessidade de aprimorar os métodos de ensino e treinamento radiográfico durante a formação médica e a residência em ortopedia. A padronização dos critérios de avaliação e o uso de ferramentas complementares, como modelos 3D e softwares de reconstrução, podem melhorar a confiabilidade das classificações e reduzir a variabilidade entre observadores.

Além disso, a implementação de programas de capacitação contínua em interpretação radiográfica e classificação de fraturas pode auxiliar na tomada de decisão mais precisa entre o tratamento cirúrgico e o conservador, reduzindo taxas de reoperações e otimizando recursos hospitalares.

Por fim, na prática clínica a incorporação de tecnologias digitais e inteligência artificial na rotina diagnóstica pode favorecer a uniformização das condutas, melhorando o prognóstico funcional e a maior eficiência no manejo das fraturas do úmero proximal. Assim, diferentes cirurgiões ao olharem para a mesma radiografia, podem classificar a fratura em categorias distintas e propor abordagens completamente diferentes para o mesmo paciente. Tal fato, tem impacto ao fornecer ao paciente um prognóstico claro sobre complicações ou a necessidade futura de cirurgias de revisão torna-se um desafio no ambiente clínico.

tência nas classificações, indicando que a experiência profissional é um fator determinante para a estabilidade diagnóstica.

Os achados ressaltam a importância de aperfeiçoar os processos de ensino e treinamento em ortopedia, especialmente durante a residência médica, além de sugerirem que o uso de ferramentas de imagem

complementares pode aumentar a acurácia e reduzir a subjetividade das classificações radiográficas. Ainda que o sistema NEER se mantenha o mais reprodutível entre os avaliados, a busca por métodos de classificação mais objetivos, integrados e reprodutíveis continua sendo um desafio essencial para a prática ortopédica moderna.

Para futuros estudos sugere-se análises mais aprofundadas considerando avaliar fraturas simples versus complexas, fraturas em duas, três e quatro partes e o impacto da experiência profissional em cada sistema de classificação de fraturas do úmero proximal.

Declaração do autor CRediT

Conceitualização: Checchi AR; Vieira, IF; Vedam, JV. Metodologia: Vieira, IF; Lima, DB; Checchi AR; Dinies, GMJ. Validação: Vieira, JLL; Lima, DB. Análise estatística: Vieira, JLL; Vieira, IF. Análise formal: Vieira, JLL; Vieira, IF; Lima, DB; Calandrini, RM. Investigação: Vieira, IF; Checchi AR; Lima, DB; Calandrini, RM. Redação – preparação do rascunho original: Checchi AR; Vieira, IF; Vieira, JLL; Dinies, GMJ. Redação – revisão e edição: Vieira, IF; Checchi AR; Vieira, JLL; Calandrini, RM; Vedam, JV. Visualização: Checchi AR; Vieira, JLL; Dinies, GMJ. Supervisão: Vieira, IF; Lima, DB; Vedam, JV. Administração do projeto: Vieira, IF; Checchi AR; Lima, DB.

Todos os autores leram e concordaram com a versão publicada do manuscrito.

Declaração de conflito de interesse

Os autores declaram que não têm interesses financeiros concorrentes ou relações pessoais conhecidas que possam ter influenciado o trabalho relatado neste artigo.

REFERÊNCIAS

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: A review. *Injury*. 2006; 37(8): 691-697. DOI: 10.1016/j.injury.2006.04.130
2. Olsson C, Nordqvist A, Petersson CJ. Increased fragility in patients with fracture of the proximal humerus: a case control study. *Bone*. 2004; 34(6): 1072-1077. DOI: 10.1016/j.bone.2004.1001.009
3. Taskesen A, Göçer A, Uzel K, Yaradılmış YU. Effect of Osteoporosis on Proximal Humerus Fractures. *Geriatr Orthop Surg Rehabil*. 2020; 11:2151459320985399. DOI: 10.1177/2151459320985399
4. Carrerra EF, Wajnsztein A, Lenz M, Archetti Netto N. Reproducibility of three classifications of proximal humeral fractures. *Einstein (São Paulo)*, 2012; 10(4): 473-479. DOI: 10.1590/S1679-45082012000400014
5. Neer CS 2nd. Displaced proximal humeral fractures. I. Classification and evaluation. *J Bone Joint Surg Am*. 1970; 52(6): 1077-1089. DOI: 10.2106/00004623-197052060-00001
6. 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(4): 427-433. DOI: 10.1016/j.jse.2004.01.034
7. Muller ME, Koch P, Nazarian S, Schatzker J. The comprehensive classification of fractures of long bones. Berlin: Springer-Verlag, 1990.
8. Foruria AM, Martinez-Catalan N, Pardos B, Larson D, Barlow J, Sanchez-Sotelo J. Classification of proximal humerus fractures according to pattern recognition is associated with high intraobserver and interobserver agreement. *JSES Int*. 2022; 6(4): 563-568. DOI: 10.1016/j.jseint.2022.03.005
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1): 159-74. DOI: 10.2307/2529310
10. Iordens GI, et al. The reliability and reproducibility of the Hertel classification for comminuted proximal humeral fractures compared with the Neer classification. *J Orthop Sci*. 2016; 21(5): 596-602. DOI: 10.1016/j.jos.2016.05.011
11. Papakonstantinou MK, et al. Interobserver agreement of Neer and AO classifications for proximal humeral fractures. *ANZ J Surg*. 2016; 86(4): 280-284. DOI: 10.1111/ans.13451
12. Marmor MT, et al. Comparison of the Neer classification to the 2018 update of the Orthopedic Trauma Association/AO fracture classification for classifying proximal humerus fractures. *OTA Int*. 2022; 5(3): e200. DOI: 10.1097/OI9.0000000000000200
13. Marongiu G, Leinardi L, Congia S, Frigau L, Mola F, Capone A. Reliability and reproducibility of the new AO/OTA 2018 classification system for proximal humeral fractures: a comparison of three different classification systems. *J Orthop Traumatol*. 2020; 21(1): 4. DOI: 10.1186/s10195-020-0543-1
14. Diaz LEP, Dada Neto F, Lofrano L, Garcia JVDC, Santana MVF, Dobashi ET. Intra- and inter-observer agreement of proximal humeral fractures classifications in adults. *Acta Ortop Bras*. 2022; 30(6): e257229. DOI: 10.1590/1413-785220223006e257229
15. Martínez-Sola R, et al. 'Absolute' inter-observer classifications agreement for proximal humeral fractures with a single shoulder anteroposterior X-ray. *J Orthop Surg (Hong Kong)*. 2021; 29(2): 23094990211010520. DOI: 10.1177/23094990211010520
16. Cocco LF, Aihara AY, Franciozi C, Dos Reis FB, Luzo MVM. Three-dimensional models increase the interobserver agreement for the treatment of proximal humerus fractures. *Patient Saf Surg*. 2020; 14:33. DOI: 10.1186/s13037-020-00258-2
17. de Giovanni R, Coppola M, Rossi V, Mariconda M, Cozzolino A. A Comprehensive Review of 3D Imaging and Printing in Proximal Humerus Fractures and Sequelae. *Journal of Clinical Medicine*. 2025; 14(21): 7711. DOI: 10.3390/jcm14217711
18. Sidor ML, Zuckerman JD, Lyon T, Koval K, Cuomo F, Schoenberg N. The Neer classification system for proximal humeral fractures. An assessment of interobserver reliability and intraobserver reproducibility. *J Bone Joint Surg Am*. 1993. 75(12): 1745-50. DOI: 10.2106/00004623-199312000-00002
19. Siebenrock KA, Gerber C. The reproducibility of classification of fractures of the proximal end of the humerus. *J Bone Joint Surg Am*. 1993; 75(12): 1751-5
20. Vieira AC, Vieira A, Évora PT, Porto PT. Neer Classification – how do we use it and why is it important? European Society Radiology Congress. 2018, Poster section number C-2074. DOI: 10.1594/ecr2018/C-2074
21. Matsumura N, Furuhashi R, Seto T, Takada Y, Shirasawa H, Oki S, Kawano Y, Shiono S. Reproducibility of the modified Neer classification defining displacement with respect to the humeral head fragment for proximal humeral fractures. *J Orthop Surg Res*. 2020; 23;15(1): 438. DOI: 10.1186/s13018-020-01966-2
22. Tafoya-Arreguín GA, Castillo-González JR, Sepúlveda-Cadena CO, de Jesus Martinez-Ruiz J, Briseno-Ramírez J. Prognostic value of the Mayo-FJD classification for nonsurgical proximal humerus fractures: a morphological approach to predicting functional outcomes, *JSES International*.

2026. DOI: 10.1016/j.jseint.2026.101743

23. Chung SW, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthop.* 2018; 89(4): 468-473. DOI: 10.1080/17453674.2018.1453714

24. Sperling Jr. JW, et al. A deep learning algorithm to detect proximal humerus fractures on radiographs. *JSES Reviews, Reports, and Techniques.* 2025; 5:635e640, DOI: 10.1016/j.xrrt.2025.07.025

Como citar este artigo: Lima, D.B., Checchi, A.R., Vieira, I.F., Vieira, J.L.L., Dinies, G.M.J., Calandrini, R.M., Vedam, J.V. (2026). Concordância intra/interobservador dos sistemas de classificação das fraturas de úmero proximal. *O Mundo Da Saúde*, 50. <https://doi.org/10.15343/0104-7809.202650e19382025P>. *Mundo Saúde*. 2026,50:e19382025.