



## CONCORDÂNCIA NA CLASSIFICAÇÃO DE NEER DA FRATURA PROXIMAL DO ÚMERO

*Agreement in the neer classification of proximal humerus fracture*

Francielly Teles Silva<sup>1</sup>, Luiz Carlos Bedin Filho<sup>2</sup>, Sávio Machado Torres<sup>3</sup>, Marcelo Wiltemburg Alves<sup>4</sup>

<sup>1-3</sup>Departamento de Ortopedia e Traumatologia. Centro Médico de Campinas – Campinas, SP.

### Resumo

Introdução: As fraturas de úmero proximal são aproximadamente 4 a 5% de todas as fraturas, acometendo mais idosos do sexo feminino. Como método de tratamento, seja ele cirúrgico ou conservador, é preciso classificá-la para saber o prognóstico da fratura, para isso a utilização da classificação de Neer para úmero proximal, que é uma das mais utilizadas, baseia-se em partes fraturadas do úmero. Apesar de amplamente utilizada há dissociações de avaliação e reprodutibilidade radiográfica, entre examinadores. Objetivo: Avaliar a concordância inter e intraobservacional da classificação de Neer para fraturas da extremidade proximal do úmero em adultos. Método: Foram selecionadas radiografias de 20 casos de fratura de úmero proximal, disponibilizadas para 35 examinadores (Ortopedistas, incluindo especialistas em cirurgia de Ombro e Cotovelo), que as classificaram de acordo com a classificação de Neer, em um período de 60 dias. Resultados: Avaliando inicialmente os 35 avaliadores com diferentes níveis de experiência, em dois momentos diferentes, com comum intervalo de 60 dias entre uma avaliação e outra, a confiabilidade intravaliador, considerando a mediana, foi excelente 0,923 (IC 95% [0,610 - 1,237]), assim como a moda 0,925 (IC 95% [0,626 – 1,224]). Na avaliação dos 20 avaliadores com maior confiabilidade intravaliador ( $k > 0,5$ ), a confiabilidade intravaliador, considerando a mediana entre os 20 avaliadores, foi excelente 0,923 (IC 95% [0,610 - 1,237]), assim como considerando a moda 0,855 (IC 95% [0,569 – 1,141]). Por fim, considerando os 20 avaliadores com maior confiabilidade intravaliador individual, a confiabilidade interavaliador na primeira avaliação (teste) foi leve 0,259 (IC 95% [0,239 – 0,280]), assim como na segunda avaliação (reteste) 0,289 (IC 95% [0,269 – 0,310]). Conclusão: Este estudo evidencia que o sistema de Classificação proposto por Neer, seguindo o índice de Kappa, apresenta como resultado final uma categorização pobre de concordância, tanto intraobservacional quando interobservacional, das fraturas da extremidade proximal do úmero. Observou-se também uma confiabilidade menor (leve a desprezível) na análise interobservador. Desta forma, fica evidente a importância da utilização de outros sistemas de classificação e de métodos diagnósticos complementares, como tomografia computadorizada, para melhor avaliação e condução dos pacientes com fratura da extremidade proximal do úmero, objetivando a otimização do tratamento e reabilitação destes casos.

**Palavras-chave:** Ortopedia, Traumatologia, Úmero proximal, fraturas, Neer.

Doi: <https://doi.org/10.37497/ijhmreview.v8i2.319>

<sup>1</sup> ORCID: <https://orcid.org/0000-0002-2328-7820>

<sup>2</sup> ORCID: <https://orcid.org/0000-0003-2454-5227>

<sup>3</sup> ORCID: <https://orcid.org/0000-0001-6467-5574>

<sup>4</sup> ORCID: <https://orcid.org/0000-0003-1973-885X>



## **Abstract**

**Background:** Proximal humerus fractures account for approximately 4 to 5% of all fractures, affecting more elderly females. As a treatment method, whether surgical or conservative, it is necessary to classify it to know the prognosis of the fracture. Despite being widely used, there are dissociations of evaluation and radiographic reproducibility among examiners. **Aim:** To assess the inter- and intra-observational agreement of Neer's classification for fractures of the proximal humerus in adults. **Method:** Radiographs of 20 cases of proximal humerus fracture were selected and made available to 35 examiners (Orthopedists, including specialists in shoulder and elbow surgery), who classified them according to Neer's classification, over a period of 60 days. **Results:** Initially evaluating the 35 evaluators with different levels of experience, at two different times, with a common interval of 60 days between one evaluation and another, the intra-evaluator reliability, considering the median, was excellent 0.923 (95% CI [0.610 - 1.237]), as well as the mode 0.925 (95% CI [0.626 - 1.224]). In the evaluation of the 20 evaluators with the highest intra-evaluator reliability ( $k > 0.5$ ), the intra-evaluator reliability, considering the median among the 20 evaluators, was excellent 0.923 (95% CI [0.610 - 1.237]), as well as considering the mode 0.855 (95% CI [0.569 - 1.141]). Finally, considering the 20 evaluators with the highest individual intra-evaluator reliability, the inter-evaluator reliability in the first evaluation (test) was light 0.259 (95% CI [0.239 - 0.280]), as well as in the second evaluation (retest) 0.289 (95% CI [0.269 - 0.310]). **Conclusion:** This study evidences that the classification system proposed by Neer, following the Kappa index, presents as a final result a poor categorization of agreement, both intraobservational and interobservational, of fractures of the proximal humerus extremity. A lower reliability (mild to negligible) was also observed in the interobserver analysis. Thus, it is evident the importance of using other classification systems and complementary diagnostic methods, such as computed tomography, for better assessment and management of patients with fractures of the proximal humerus, aiming at optimizing the treatment and rehabilitation of these cases.

**Keywords:** Orthopedics, Traumatology, Proximal humerus, fractures, Neer.

## **Introdução**

A fratura do úmero proximal é a sétima mais comum em frequência, correspondendo a aproximadamente 80% de todas as fraturas no úmero (COURT-BROWN; CAESAR, 2006). A incidência das fraturas nessa região pode variar de 63 a 73 em cada 100 mil habitantes, ocorrendo predominantemente nos pacientes idosos. Essa fratura acomete até 75% dos indivíduos acima de 60 anos (LIND; KRØNER; JENSEN, 1989) e é mais comum no gênero feminino, em uma proporção de três homens para sete mulheres. O número de fraturas do úmero proximal vem crescendo, uma vez que a expectativa de vida está aumentando, ainda mais associado ao aumento de casos de pessoas com osteoporose. Fraturas relacionadas a osteoporose ocorre em média em 2 milhões de pacientes por ano nos Estados Unidos, e fraturas de úmero proximal representam 10% desse número (GRACITELLI et al., 2017).

Ainda há certa controvérsia quanto ao benefício do tratamento cirúrgico e evidencia de parâmetros radiográficos que mudam o prognóstico do tratamento conservador (HANDOLL; BRORSON, 2015). Diversas classificações foram criadas ao longo do tempo como Neer, Patomorfológica (RESCH et al., 2016) AO/ASIF, e outras binárias com o intuito de agrupar casos de fraturas semelhantes em um único tipo e ter uma conduta específica, porém com todas as classificações já publicadas, ainda há divergências entre elas e divergências entre inter e intra-observadores (GRACITELLI et al., 2017; RESCH et al., 2016; SIEBENROCK; GERBER, 1993).

Os sistemas de classificação de fraturas atualmente utilizados para a extremidade proximal do úmero apresentam baixa concordância e reprodutibilidade (SIDOR et al., 1993). Em um estudo feito por Cocco et al. (2019), houve um aumento de concordância nas principais classificações (AO/ASIF



e Neer) entre cirurgiões experientes, quando utilizado realidade aumentada para avaliar a fratura, a concordância não alterou quando o estudo foi realizado ainda com cirurgiões menos experientes.

Mais recentemente, Hertel et al. (2004), propuseram a classificação binária (LEGO), para as fraturas do úmero proximal. Apesar de semelhante à classificação de Neer, onde o úmero proximal é dividido em 4 partes, a classificação ainda correlaciona com a perfusão sanguínea na cabeça do úmero. Foi descrito que os mais importantes preditores para isquemia da cabeça umeral eram a extensão do fragmento metafisário, o desvio da diáfise, o desvio dos tubérculos e a presença de luxação e fratura intra-articular.

O sistema de classificação proposto por Neer (2006) é amplamente utilizado como forma de avaliar e determinar diretrizes para o tratamento das fraturas na região proximal do úmero. A classificação de Neer mesmo sendo amplamente utilizada para extremidade proximal do úmero apresentam baixa concordância e reprodutibilidade. Essa disparidade se deve a múltiplas variáveis, tais como: dificuldade para realizar radiografias e posicionar o paciente de forma correta, radiografias de péssima qualidade, entre outros fatores (SIDOR et al., 1993).

## **Objetivo**

Avaliar a concordância intraobservadores e interobservadores de diferentes parâmetros radiográficos da classificação de Neer para fraturas do úmero proximal em adultos entre cirurgiões ortopédicos com diferentes níveis de experiência, bem como o impacto dessa discordância no meio clínico.

## **Método**

### ***Tipo de estudo***

Estudo observacional retrospectivo.

### ***Seleção de amostras***

Foi realizada análise retrospectiva de todos os pacientes com o diagnóstico de fratura da extremidade proximal do úmero do setor do Pronto Atendimento de Ortopedia e Traumatologia, do período de janeiro de 2019 a março de 2021. Foram incluídos exames realizados em indivíduos esqueleticamente maduros e que apresentem fraturas da extremidade proximal do úmero. Ainda, foram selecionadas radiografias de 20 pacientes, em quatro incidências (anteroposterior, anteroposterior verdadeiro e perfil de escápula e axilar), com boa qualidade radiográfica, conforme padronização do estudo radiográfico da cintura escapular proposto por Doneux, Checchia e Miyazaki (1998). Nenhum novo contato com os pacientes foi realizado. Apenas radiografias já arquivadas no Serviço foram utilizadas.

### ***Processo de classificação das fraturas***

As radiografias foram classificadas por 30 examinadores com diferentes níveis de experiência. Com a finalidade de minimizar o viés devido às dificuldades de interpretação e inexperiência com a classificação, os observadores obtiveram uma explanação prévia do sistema de classificação que se pretende utilizar no presente trabalho.

As radiografias foram organizadas em ordem cronológica em um arquivo digital fechado. As classificações foram realizadas em dois momentos diferentes, com um intervalo de 60 dias entre uma avaliação e a outra. Cada um dos examinadores avaliou independentemente as radiografias e as classificou. Os observadores dispuseram de todo o tempo necessário para avaliar as radiografias, sendo instruídos a não discutir os sistemas de classificação até o fim da fase de classificação. Além disso, não obtiveram acesso à história dos pacientes ou a qualquer dado clínico. Cabe ressaltar que os



avaliadores foram convidados pessoalmente pelos pesquisadores deste projeto, e assinaram Termo de Consentimento Livre e Esclarecido autorizando a sua participação na pesquisa.

### Análise estatística

O *software* utilizado para a análise de dados foi o SPSS 25.0. A confiabilidade foi realizada com o teste Kappa para duas variáveis, e com o teste Kappa de Fleiss (1979) para múltiplas variáveis. Considerou-se um nível de significância de 5% para as análises inferenciais. O coeficiente Kappa ofereceu uma proporção de concordância emparelhada entre observadores. Os valores do coeficiente Kappa variam de -1 a +1. Quando o valor é -1, o significado é de discordância total; quando o valor é +1, indica total concordância; e quando o valor é zero, representa não concordância. A força de concordância do teste de Kappa foi interpretada de acordo com Landis e Koch (1977): < 0 = sem concordância; 0,00 – 0,20 = desprezível; 0,21 – 0,40 = leve; 0,41 – 0,60 = moderada; 0,61 – 0,80 = grande; 0,81 a 1,00 = excelente. Este projeto foi submetido à avaliação por um Comitê de Ética em Pesquisa, sendo aprovado por cumprir os requisitos da Resolução 466/2012 do Conselho Nacional de Saúde, que define os aspectos éticos e legais das pesquisas que envolvem seres humanos.

### Resultados

Foram avaliadas radiografias de 20 pacientes com fratura da extremidade superior do úmero, no período de 2018 a 2020. A análise foi feita por 35 ortopedistas, com diferentes níveis de experiência, em dois momentos diferentes, com comum intervalo de 60 dias entre uma avaliação e outra. A confiabilidade intravaliador variou de excelente 1,000 (IC 95% [0,687 - 1,313]) à desprezível 0,051 (IC 95% [-0,229 - 0,331]), conforme mostra a Tabela 1. A porcentagem de concordância variou de 16,33% e 100,00%, com uma média de 68,71% (DP: 16,33). Observa-se na Tabela 1 que a confiabilidade intravaliador, considerando a mediana entre os 35 avaliadores, foi excelente 0,923 (IC 95% [0,610 - 1,237]), assim como a moda 0,925 (IC 95% [0,626 - 1,224]).

Tabela 1 – Confiabilidade intravaliador (teste e reteste) geral dos 35 avaliadores

Intravaliador	Kappa	Z	p-valor	IC 95% limite inferior	IC 95% limite superior
Mediana	0,923	5,773	0,000	0,610	1,237
Moda	0,925	6,067	0,000	0,626	1,224

Fonte: Dados obtidos pelos pesquisadores.

Foram selecionados os 20 avaliadores com maior confiabilidade intravaliador ( $k > 0,5$ ). Observa-se na Tabela 2 que a confiabilidade intravaliador, considerando a mediana entre os 20 avaliadores, foi excelente 0,923 (IC 95% [0,610 - 1,237]), assim como considerando a moda 0,855 (IC 95% [0,569 - 1,141]).

Tabela 2 – Confiabilidade intravaliador (teste e reteste) geral dos 20 avaliadores com maior confiabilidade intravaliador individual

Intravaliador	Kappa	Z	p-valor	IC 95% limite inferior	IC 95% limite superior
Mediana	0,923	5,773	0,000	0,610	1,237
Moda	0,855	5,855	0,000	0,569	1,141

Fonte: Dados obtidos pelos pesquisadores.

Visualiza-se na Tabela 3 que, considerando os 20 avaliadores com maior confiabilidade intravaliador individual, a confiabilidade interavaliador na primeira avaliação (teste) foi leve 0,259 (IC 95% [0,239 - 0,280]), assim como na segunda avaliação (reteste) 0,289 (IC 95% [0,269 - 0,310]).



Tabela 3 – Confiabilidade interavaliador no teste e no reteste dos 20 avaliadores com maior confiabilidade intravaliador individual

Interavaliador	Kappa	Z	p-valor	IC 95% limite inferior	IC 95% limite superior
Av 1	0,259	25,104	0,000	0,239	0,280
Av 2	0,289	27,778	0,000	0,269	0,310

Fonte: Dados obtidos pelos pesquisadores.

## Discussão

Os sistemas de classificação devem fornecer ferramentas que subsidiem a avaliação clínica. Um bom sistema de classificação precisa ser válido, confiável e reprodutível. No entanto, para ser ideal, é necessário que padronize uma linguagem de comunicação segura, gere diretrizes para o tratamento, e auxilie na obtenção do prognóstico de uma doença. Deve, ainda, ser um instrumento que possibilite avaliar e comparar os resultados obtidos com o tratamento de doenças semelhantes em diversos centros de pesquisa (GRACITELLI et al., 2017).

O conceito de que os principais sistemas de classificação das fraturas da extremidade proximal do úmero variam de uma baixa à média concordância, e são pouco reprodutíveis já é conhecido. No entanto, neste estudo a porcentagem de concordância esteve entre 16,33% e 100,00%, com uma média de 68,71% (DP: 16,33), e a confiabilidade intravaliador, considerando a mediana entre os 20 avaliadores, foi excelente 0,923 (IC 95% [0,610 - 1,237]), bem como a moda 0,855 (IC 95% [0,569 - 1,141]).

A classificação proposta por Neer (2006) é amplamente aceita e comumente usada nos dias de hoje, apesar de ser alvo de inúmeras críticas relacionadas à sua reprodutibilidade. O sistema de classificação de Neer também é limitado em sua capacidade de diferenciar entre os muitos padrões de fraturas com desvio mínimo. Este grupo de fraturas inclui mais de 80% das fraturas do úmero proximal, mas, na verdade, representam uma variedade de padrões e resultados de lesão. Koval et al. (1997), revisaram 104 pacientes com fraturas com desvio mínimo tratados com fisioterapia. Embora a maioria dos pacientes tenha tido um bom resultado, 23% tiveram um resultado regular ou ruim. Pode haver fratura e características do paciente não incluídas no sistema de classificação de Neer que predizem esses resultados menores.

Carofino et al. (2013) apresentaram uma reprodutibilidade inter observador variável de 0.21 e 0.64, enquanto intraobservador entre 0.5 e 0.8. Comparável com a classificação da AO e outros sistemas com de tornozelo, fêmur e escafoide. No entanto foi identificado que nenhum outro Sistema foi tão confiável como a Classificação de Neer.

Bahrs et al. (2010), revisaram uma série de fraturas em uma parte e descobriram que a extensão intra-articular, o deslocamento da tuberosidade maior que 5 mm e a angulação de 15° foram associados a resultados clínicos reduzidos. Esses estudos sugerem que a classificação de Neer pode não prever suficientemente o prognóstico em pacientes com o padrão de lesão mais comum.

Um ensaio clínico randomizado conduzido por Resch et al. (2016), verificou a melhora da concordância da classificação de Neer após treinamento de 14 observadores; o valor do índice kappa variou de 0,34 antes do treinamento para 0,62 após o treinamento desses observadores, o que não se constatou no grupo controle. Já a revisão sistemática de Shrader et al. (2005), não evidenciou que observadores mais experientes apresentem menos discordância do que os menos experientes. No entanto, Gracitelli et al. (2017), apresentaram melhor concordância tanto inter quanto intra-observador entre avaliadores mais experientes.

Devido à baixa concordância e à reprodutibilidade das classificações mais utilizadas para caracterizar as fraturas da região proximal do úmero, propõe-se um novo conceito para classificar



essas fraturas com a finalidade de melhorar essa interpretação e criar novas alternativas para o tratamento.

### **Conclusão**

Este estudo evidenciou que o sistema de Classificação proposto por Neer, seguindo o índice de Kappa, apresenta como resultado final uma categorização pobre de concordância, tanto intraobservacional quanto interobservacional, das fraturas da extremidade proximal do úmero. Observou-se também uma confiabilidade menor (leve a desprezível) na análise interobservador. Dessa forma, fica evidente a importância da utilização de outros sistemas de classificação e de métodos diagnósticos complementares, como tomografia computadorizada, para melhor avaliação e condução dos pacientes com fratura da extremidade proximal do úmero, objetivando a otimização do tratamento e reabilitação destes casos.

### **Referências**

BAHRS, C. et al. Clinical and radiological evaluation of minimally displaced proximal humeral fractures. **Archives of orthopaedic and trauma surgery**, v. 130, n. 5, p. 673–679, 2010.

CAROFINO, B. C.; LEOPOLD, S. S. **Classifications in brief: the Neer classification for proximal humerus fractures**. Springer, 2013.

COCCO, L. F. et al. Inter-observer reliability of alternative diagnostic methods for proximal humerus fractures: a comparison between attending surgeons and orthopedic residents in training. **Patient Safety in Surgery**, v. 13, n. 1, p. 1–13, 2019.

COURT-BROWN, C. M.; CAESAR, B. Epidemiology of adult fractures: a review. **Injury**, v. 37, n. 8, p. 691–697, 2006.

DONEUX, P. S.; CHECCHIA, S. L.; MIYAZAKI, A. N. Padronização do estudo radiográfico da cintura escapular. **Rev Bras Ortop**, v. 33, n. 11, p. 883–8, 1998.

FLEISS, J. L. et al. Inter-examiner reliability in caries trials. **Journal of dental research**, v. 58, n. 2, p. 604–609, 1979.

GRACITELLI, M. E. et al. Intraobserver and interobserver agreement in the classification and treatment of proximal humeral fractures. **Journal of Shoulder and Elbow Surgery**, v. 26, n. 6, p. 1097–1102, 2017.

HANDOLL, H. H.; BRORSON, S. Interventions for treating proximal humeral fractures in adults. **Cochrane Database of Systematic Reviews**, n. 11, 2015.

HERTEL, R. et al. Predictors of humeral head ischemia after intracapsular fracture of the proximal humerus. **Journal of shoulder and elbow surgery**, v. 13, n. 4, p. 427–433, 2004.

KOVAL, K. J. et al. Functional outcome after minimally displaced fractures of the proximal part of the humerus. **JBJS**, v. 79, n. 2, p. 203–7, 1997.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159–174, mar. 1977.



LIND, T.; KRØNER, K.; JENSEN, J. The epidemiology of fractures of the proximal humerus. **Archives of orthopaedic and trauma surgery**, v. 108, n. 5, p. 285–287, 1989.

NEER, C. S. The classic: displaced proximal humeral fractures: part i. classification and evaluation. **Clinical Orthopaedics and Related Research®**, v. 442, p. 77–82, 2006.

RESCH, H. et al. Classification of proximal humeral fractures based on a pathomorphologic analysis. **Journal of shoulder and elbow surgery**, v. 25, n. 3, p. 455–462, 2016.

SHRADER, M. W. et al. Understanding proximal humerus fractures: image analysis, classification, and treatment. **Journal of shoulder and elbow surgery**, v. 14, n. 5, p. 497–505, 2005.

SIDOR, M. L. et al. The Neer classification system for proximal humeral fractures. An assessment of interobserver reliability and intraobserver reproducibility. **The Journal of bone and joint surgery. American volume**, v. 75, n. 12, p. 1745–1750, 1993.

SIEBENROCK, K. A.; GERBER, C. The reproducibility of classification of fractures of the proximal end of the humerus. **The Journal of bone and joint surgery. American volume**, v. 75, n. 12, p. 1751–1755, 1993.