

UC Davis

UC Davis Previously Published Works

Title

Comparison of the Neer classification to the 2018 update of the Orthopedic Trauma Association/AO fracture classification for classifying proximal humerus fractures.

Permalink

<https://escholarship.org/uc/item/8mv318cd>

Journal

OTA International, 5(3)

Authors

Marmor, Meir

Agel, Julie

Dumpe, Jarrod

et al.

Publication Date

2022-09-01

DOI

10.1097/OI9.0000000000000200

Peer reviewed

Comparison of the Neer classification to the 2018 update of the Orthopaedic Trauma Association/AO fracture classification for classifying proximal humerus fractures

Meir T. Marmor, MD^a, Julie Agel, MA, ATC^b, Jarrod Dumpe, MD^c, James F. Kellam, MD^d, Geoffrey S. Marecek, MD^e, Eric Meinberg, MD^a, Mai P. Nguyen, MD^f, Stephen Sims, MD^g, Gillian L. Soles, MD^h, Matthew D. Karam, MDⁱ

Abstract

Background: The classification of fractures is necessary to ensure a reliable means of communication for clinical interaction, education and research. The Neer classification is the most commonly used classification for proximal humerus fractures. In 2018 the Orthopaedic Trauma Association (OTA) and the Arbeitsgemeinschaft für Osteosynthesefragen (AO) Foundation provided an update to the OTA/AO Fracture Classification Scheme addressing many of the concerns about the previous versions of the classification. The objective of the present study was to evaluate the rater reliability of the 2 classifications and if the classifications subjectively better characterized the fracture patterns.

Methods: X-rays and Computerized Tomography scans of 24 proximal humerus fractures were given to 7 independent raters for classification according to the Neer and 2018 OTA/AO classification. Both full-forms and short-forms of the classifications were tested. The Fleiss Kappa statistic was used to assess inter-rater agreement and intra-rater consistency for the 2 classifications. For each case the raters subjectively commented on how well each classification was able to characterize the fracture pattern.

Results: All raters graded the 2018 OTA/AO classification as good as or better than the Neer classification for an adequate description of the fracture patterns. The short-form 2018 OTA/AO classification had the most 4 rater and 5 rater agreement cases and the second most 6 rater agreement cases. The short-form Neer classification had the second most 4 rater and 5 rater agreement cases and the most 6 rater agreement cases. The full 2018 OTA/AO had the least 4, 5, or 6 rater agreement cases of all the classification systems. Inter-rater agreement was fair for the full and short form of both the Neer and 2018 OTA/AO classification. The full and short Neer classifications together with the short 2018 OTA/AO classification had moderate intra-rater consistency, while the full 2018 OTA/AO classification only had slight intra-rater consistency.

Conclusions: The 2018 OTA/AO classification is equivalent in its short-form to the Neer classification in inter-rater reliability and intra-rater consistency; and is superior in its full form for characterizing specific fracture types. The low inter-rater reliability of the full 2018 OTA/AO classification is a concern that may need to be addressed in the future.

Keywords: Neer classification, OTA/AO classification, proximal humerus fractures, reliability study

1. Introduction

The classification of fractures is necessary to ensure a reliable means of communication for clinical interaction, education and research. Often, there are multiple classification schemes, originating in different eras, available for the same fracture.

The Neer classification, described initially in 1970, is the most commonly used classification for proximal humerus fractures.^[1,2] There is ample evidence, however, that observers often disagree on displacement and orientation of fracture lines, which are the necessary factors to accurately use this classifica-

Source of funding: Nil.

The study received institutional review board exempt status in both involved trauma centers.

The authors have no conflicts of interest to disclose.

^a University of California San Francisco, Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco, CA, ^b University of Washington, Harborview Medical Center, Seattle, WA, ^c Orthopaedic Trauma Institute, Atrium Health - Navicent, Macon, GA, ^d Department of Orthopaedic Surgery, McGovern Medical School, University of Texas Health Science Center at Houston, Houston Texas, ^e Cedars - Sinai Medical Center, Los Angeles, CA, ^f University of Minnesota St. Paul, Regions Hospital, Saint Paul, MN, ^g Atrium Health Orthopaedic Surgery, Carolinas Medical Center, Charlotte, NC, ^h University of California Davis, Sacramento, CA, ⁱ University of Iowa, Iowa City, IA.

* Corresponding author. Address: Orthopaedic Trauma Institute, Zuckerberg San Francisco General Hospital & Trauma Center, 2550 23rd Street, Building 9, 2nd floor, San Francisco, CA 94110. Tel: +650 477 6238. E-mail address: meir.marmor@ucsf.edu (M. T. Marmor).

Copyright © 2022 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of the Orthopaedic Trauma Association.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

OTAI (2022) e200

Received: 13 January 2021 / Accepted: 15 February 2022

Published online 16 June 2022

<http://dx.doi.org/10.1097/OI9.000000000000200>

tion.^[3-6] Furthermore, the Neer classification fails to account for clinically significant fracture attributes such as varus/valgus coronal plane alignment.^[7]

Similar to the Neer classification, the original OTA/AO classification^[8,9] was unable to account for fracture features such as fragment displacement.^[7] In 2018 the Orthopedic Trauma Association (OTA) and the AO Foundation provided an update to the OTA/AO Fracture Classification Scheme (2018 OTA/AO classification), addressing many of the concerns about the previous versions of the classification.^[10] The updated compendium included a modification of the classification of proximal humerus fractures that offers a myriad of different fracture descriptions compared to the 16 possible fracture types available in the Neer classification. It is not evident, however, that the new modification will be more useful for research or clinical purposes. The objective of the present study was to evaluate the rater reliability of the 2018 OTA/AO classification compared to the Neer classification. A secondary aim was to survey the raters regarding which of the classifications subjectively better characterized the fracture patterns they evaluated.

2. Methods

The methodology of this study complies with the Guidelines for Reporting Reliability and Agreement Studies for reporting reliability and agreement studies.^[11] The study received institutional review board exempt status in both involved trauma centers.

2.1. Patient and image selection

Two trauma centers submitted de-identified proximal humerus fracture cases for review. An investigator that did not participate in the classification exercise reviewed the cases and chose 24 cases so that they would have equal distribution of OTA/AO types (As, Bs, and Cs) and Neer classification groups (2-parts, 3-parts, and 4-parts). Twenty-four cases were selected to approximate the number of cases in similar, previously published studies^[4,6,12] and in accordance with the recommended sample size of a reliability study.^[13] For each case, injury x-rays, Computerized Tomography (CT) scans with axial, coronal and sagittal reconstructions, and 3D reformations were presented.

2.2. Classification

We used a previously described numbering scheme^[3] to classify fractures according to the Neer classification. According to this scheme, the raters needed to classify any fracture type using a number between 1 and 16 (Fig. 1). A part in the Neer classification was considered a fracture fragment that has 1 cm or 45 degrees of displacement.^[1] Next, the raters classified the fractures, according to the 2018 OTA/AO classification.^[10] For each fracture, the raters recorded the type, group, subgroup (Fig. 2) and added qualifications and universal modifiers as necessary (Fig. 3). For each case, raters also recorded whether they thought each classification adequately reflected the specific fracture pattern.

Seven raters, all members of the OTA Classification and Outcomes Committee, were available to perform the classifications in 2 separate classification rounds. Due to the COVID-19 pandemic, the raters were asked to classify the fractures in 2, 2-hour, zoom video conferences (Zoom version: 5.0.4, Zoom Video Communi-

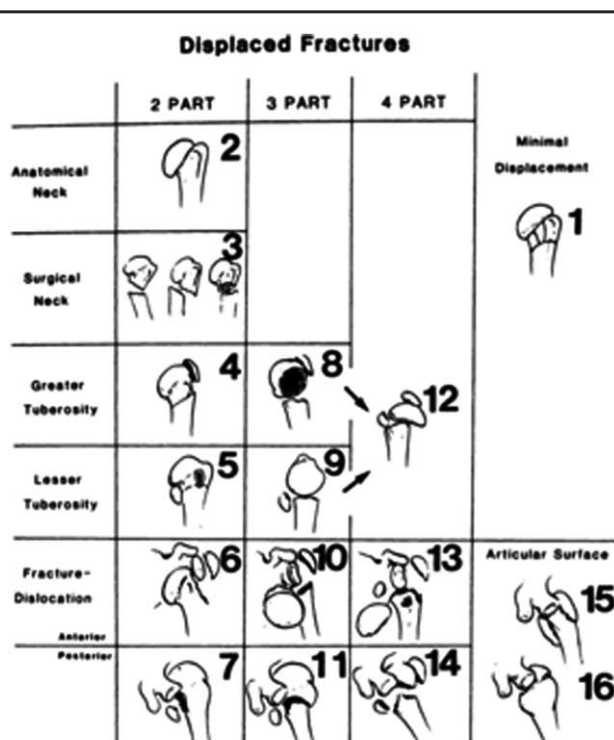


Figure 1. The Neer classification together with a numbering scheme from 1 to 16 to aid in rating evaluation.

ations Inc.). The classification rounds were conducted 6 weeks apart, with the cases presented in a different order in each round, to minimize the risk that raters would remember how they classified a particular fracture pattern in the first round. During the conferences, the investigator not classifying the fractures, acting as the moderator, presented the 24 cases, showing all the available images and going back to review images per the rater's requests. The raters were not limited in time while reviewing the cases but were not allowed to discuss the cases during the call. Once all the raters signaled (verbally or through the Zoom chat/signaling options) that they had finished classifying the fracture, the moderator presented the next fracture. Recordings of the video conferences were available to the raters, and they were allowed to watch the video recording before submitting their final classification scoresheets. Two reviewers in the first round and 6 reviewers in the second round watched only the video recording before submitting their scoresheets. Because of this limitation, the results of this study were analyzed once for all raters and once for the raters that just used video recording in both rounds. Before the classification rounds, the raters received a detailed oral and pdf file description of how to use both classification systems. The raters classified 2 test-cases together as a group exercise to practice the use of the classifications, address any questions and reinforce critical features of each classification before the first session. On both classification rounds, each of the raters classified all 24 proximal humerus cases independently without discussion with the other raters.

2.3. Statistical analysis

All data were collected in Excel 2016 @Microsoft, gathered in a central location and analyzed using Statistical Package for the Social Sciences v 26 @International Business Machines Corporation. We used the Fleiss Kappa statistic to assess inter-rater agreement and intra-rater consistency for the 2 classifica-

11

Location: Humerus, proximal end segment 11



Types:

Humerus, proximal end segment, **extraarticular, unifocal, 2-part fracture** 11A



Humerus, proximal end segment, **extraarticular, bifocal, 3-part fracture** 11B



Humerus, proximal end segment, **articular or 4-part fracture** 11C



Figure 2. The 3 types (A, B, and C) of proximal humerus fractures (bone location 11) according to the 2018 update of the OTA/AO Classification.

tions.^[14] According to this statistical test, a kappa of <0 is considered a poor agreement, 0 to 0.2 a slight agreement, 0.21 to 0.4 a fair agreement, 0.41 to 0.6 a moderate agreement, 0.61 to 0.8 a substantial agreement, and >0.81 an almost perfect

agreement.^[14] Confidence intervals of 95% (95% CI) were calculated for each kappa score.

We also recorded the number of cases in which each reviewer felt that the classifications adequately classified the fractures and

11B

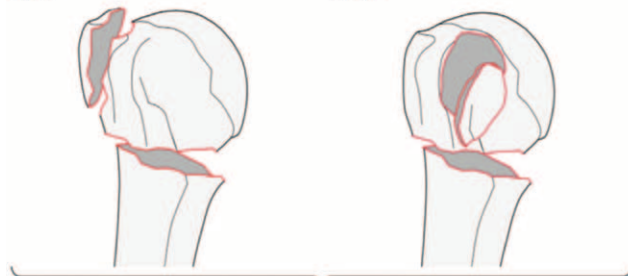
Type: Humerus, proximal end segment, **extraarticular, bifocal, 3-part fracture** 11B

Group: Humerus, proximal end segment, extraarticular, bifocal, 3-part, **surgical neck fracture** 11B1

Subgroups:

With greater tuberosity fracture 11B1.1*

With lesser tuberosity fracture 11B1.2*



*Qualifications:
u Intact wedge
v Fragmentary wedge

List of universal modifiers

| | |
|----|--|
| 1 | Nondisplaced |
| 2 | Displaced |
| 3 | Impaction |
| 3a | Articular |
| 3b | Metaphyseal |
| 4 | No impaction |
| 5 | Dislocation |
| 5a | Anterior (volar, palmar, plantar) |
| 5b | Posterior (dorsal) |
| 5c | Medial (ulnar) |
| 5d | Lateral (radial) |
| 5e | Inferior (with hip is also obturator) |
| 5f | Multidirectional |
| 6 | Subluxation/ligamentous instability |
| 6a | Anterior (volar, palmar, plantar) |
| 6b | Posterior (dorsal) |
| 6c | Medial (ulnar) |
| 6d | Lateral (radial) |
| 6e | Inferior (with hip is also obturator) |
| 6f | Multidirectional |
| 7 | Diaphyseal extension |
| 8 | Articular cartilage injury* |
| 8a | ICRS Grade 0 Normal |
| 8b | ICRS Grade 1 Superficial indentation (A) and/or superficial fissures and cracks (B) |
| 8c | ICRS Grade 2 Abnormal lesions extending down to 50% of cartilage depth |
| 8d | ICRS Grade 3 (A) Severely abnormal with defects extending down >50% of cartilage depth; (B) down to calcified layer; (C) down to subchondral bone but not through; (D) blisters included |
| 8e | ICRS Grade 4 Severely Abnormal Cartilage loss through subchondral bone |
| 9 | Poor bone quality |
| 10 | Replantation |
| 11 | Amputation associated with a fracture |
| 12 | Associated with a nonarthroplasty implant |
| 13 | Spiral type fracture |
| 14 | Bending type fracture |

*This grading system is used with the permission of the International Cartilage Repair Society.**

Figure 3. Examples of the possible subgroups (11B1.1 or 11B1.2), qualifications (u or v) and universal modifiers^[1-14] for the proximal humerus fractures (bone location 11) of the extra-articular, bifocal, 3-part type (B), and the surgical neck group (B1) according to the 2018 OTA/AO Classification.

the maximal number of reviewers that agreed on a specific classification for each case.

In order to account for a more clinically applicable version of the classifications, we repeated the analysis on a truncated “short” version of the 2 classifications. Neer-short included just the number of parts of the fracture: 2, 3, or 4 parts. OTA/AO-short included just the fracture type: A, B, or C.

The full data that was collected is available in the appendix, <http://links.lww.com/OTA/A47>.

3. Results

Seven raters reviewed 24 proximal humerus fracture cases in 2 classification rounds. Of the 168 classifications done in the first round (24 cases by 7 raters) 76 cases were rated as 2-part fractures, 50 as 3-part fractures and 40 as 4-part fractures according to the Neer classification; 51 were rated as type A, 60 as type B, and 57 as type C of the 2018 OTA/AO classification. All raters graded the 2018 OTA/AO classification as good as or better than the Neer classification for an adequate description of the fracture patterns (Fig. 4). Raters uniformly stated that use of advanced imaging in combination with applying the 2018 OTA/AO classification allowed for the most complete description of the injury.

In the Neer classification, complete agreement between all the raters occurred only in 1 case of the full Neer classification (a 4-part fracture dislocation) and only in 2 cases of the short Neer classification (a 4-part and a 2-part fracture). In 5 cases (2 anatomical neck 2-part fractures, 1 surgical neck 2-part fracture, one 4-part fracture and one 4-part fracture-dislocation), only 2 raters out of 7 agreed using the full Neer classification, and in 1 case (a 4-part fracture), only 2 raters agreed using the short Neer classification (Fig. 5).

For the 2018 OTA/AO classification, complete agreement for all reviewers was seen only for the short OTA/AO classification in 4 cases (1 type A and 3 type C). In 7 cases (5 type A and 2 type C), only 2 raters could agree on the full 2018 OTA/AO classification (Fig. 6).

Figure 7 shows the number of cases in which at least 4, 5, or 6 of the raters agreed on the different classifications. OTA/AO short classification had the most 4 rater and 5 rater agreement cases and the second most 6 rater agreement cases. The short Neer classification had the second most 4 rater and 5 rater agreement cases and the most 6 rater agreement cases. The full 2018 OTA/AO had the least 4, 5, or 6 rater agreement cases of all the classification systems.

Overall inter-rater agreement was fair for the full Neer Classification with a kappa of 0.299 (0.266–0.333 95% CI), and fair for the short Neer classification, with a kappa of 0.290 (0.226–0.353 95% CI). For the 2018 OTA/AO classification, overall inter-rater agreement was fair, with a kappa of 0.240 (0.205–0.274 95% CI), and fair for the short OTA/AO (type only) classification, with a kappa of 0.362 (0.300–0.423 95% CI).

Intra-rater consistency was evaluated for all the raters. Both the full and short Neer classifications had moderate intra-rater consistency, with full Neer kappa of 0.443 (0.387–0.50 95% CI), and short Neer kappa of 0.454 (0.348–0.560 95% CI). The short 2018 OTA/AO classification also had moderate intra-rater consistency, kappa of 0.481 (0.374–0.588 95% CI). However, the full 2018 OTA/AO classification only had slight intra-rater consistency, kappa of -0.11 (95% CI, -0.16 , -0.06).

A repeat analysis of the reviewers that only watched a video of the case presentations (did not participate in at least 1 live zoom meeting) did not change any of the inter-rater or intra-rater results for any of the classification systems.

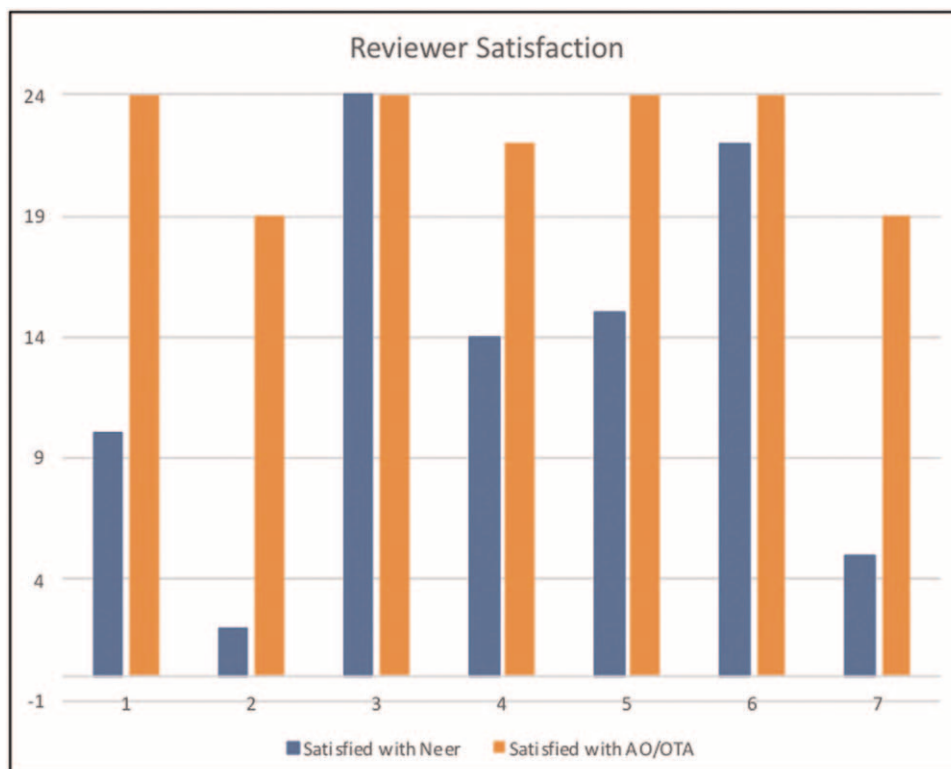


Figure 4. Number of cases (out of 24) each of the 7 reviewers saw as adequately describing each fracture pattern.

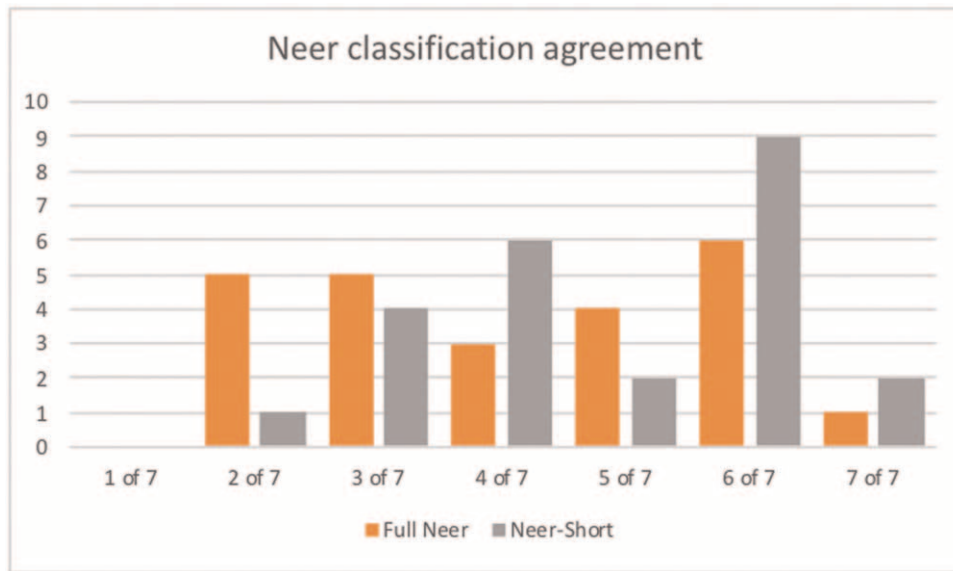


Figure 5. Number of cases that had agreement on the Neer classification between different reviewers (out of 7).

When looking at specific fracture types, only Neer classification type 13, 4-part fracture-dislocation ($\kappa=0.716, 0.626-0.805$ 95%CI), and OTA/AO classification A1.1, isolated greater tuberosity fractures ($\kappa=0.738, 0.624-0.802$ 95% CI) had good agreement. The rest of the specific fracture types in both classifications, including their long and short versions, had fair to poor agreement between raters.

4. Discussion

The purpose of this study was to compare the 2018 update of the OTA/AO classification to the Neer classification of proximal humerus fractures. We were not able to demonstrate the superiority of the newer, 2018 OTA/AO classification over the older, 1970 Neer classification in inter or intra-rater reliability.

However, we did find that all raters felt that the full 2018 OTA/AO classification better characterizes the specific fracture patterns. We also found that the short version of the 2018 OTA/AO classification had equivalent or better reliability than both the full and short Neer classifications.

Management of proximal humerus fractures remains highly controversial, with no single technique, surgical or nonsurgical, consistently demonstrating superior outcomes.^[15] A reliable classification that is also able to capture the intricacies of proximal humerus fracture patterns is necessary for proper standardization of data in outcome studies. Since its first description in 1970, the Neer classification is widely used in outcome studies as well as in clinical practice.^[2] However, previous tests of the Neer classification demonstrated only moderate to fair agreement between raters.^[3,12,16,17] In a study

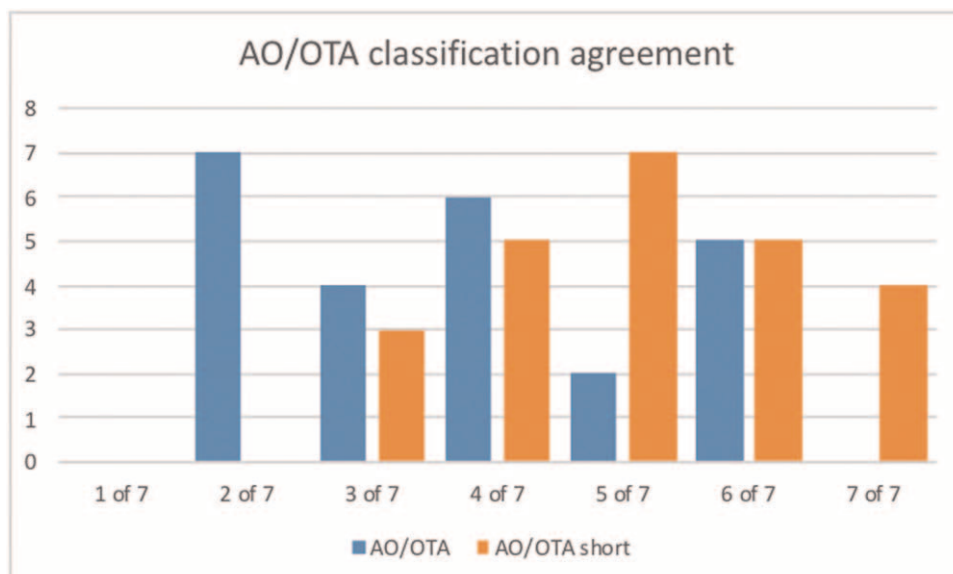


Figure 6. Number of cases that had agreement on the 2018 OTA/AO classification between different reviewers (out of 7).

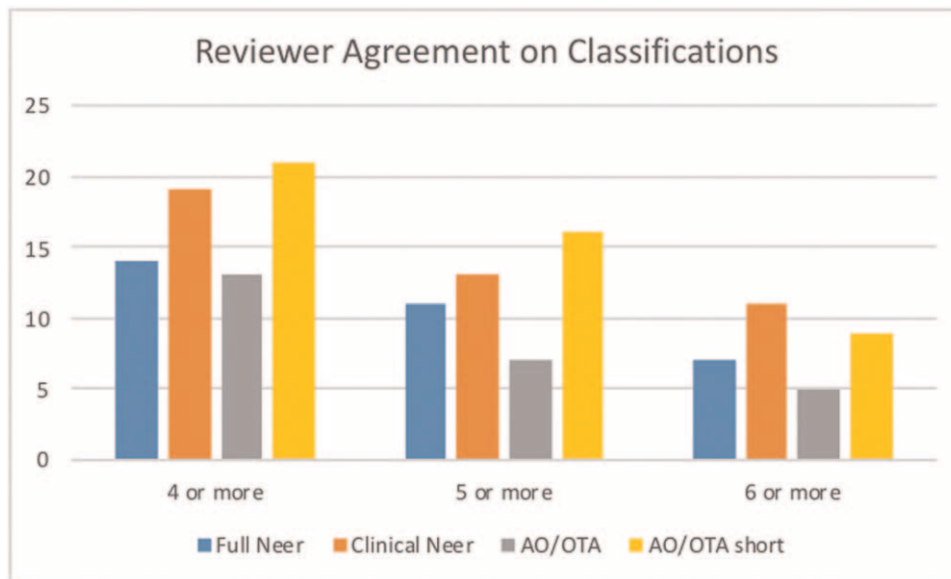


Figure 7. Number of cases in which at least 4, 5, or 6 of the raters agreed on the different classifications.

comparing residents to fellows and specialists, the mean kappa value for inter-rater agreement for the Neer classification was 0.27 (95% CI 0.26–0.28) with no clinically significant difference between orthopedic residents ($n=9$), fellows ($n=6$) and specialists ($n=9$).^[12] Analysis of 250 patient x-rays from the PROximal Fracture of the Humerus: Evaluation by Randomisation trial demonstrated similar findings.^[16] Other studies suggested that with the use of advanced computerized imaging, such as 3D CT reconstructions, this inter-rater agreement could be improved.^[17,18] Our study found only fair agreement for the Neer classification, despite the use of advanced imaging such as CT axial cuts, and sagittal, coronal and 3D reconstructions. A possible explanation of this observation may be the severity of the injuries compared to previous trials since all of the cases were taken from level I trauma centers. It is also worth noting that in more recent studies, such as the PROximal Fracture of the Humerus: Evaluation by Randomisation trial, the inter-rater agreement was similar to our study.^[16]

The OTA/AO classification was first published in 1996^[19] as an expansion of the Comprehensive Classification of Fractures of the Long Bones developed by Müller and collaborators a decade earlier.^[8,9] The classification intended to bring forth a standardized and rational methodology of describing all fractures and dislocations as well as a mechanism to code data for future recall.^[10] The classification was updated in 2008 and 2018, each time addressing concerns about terminology and the relevancy of specific classification schemes.^[10] For proximal humerus fractures, the OTA/AO classification improved on the Neer classification by its account for varus and valgus displacement. However, it was inferior to the Neer classification by its lack of account of displacement.^[7] In the 2018 update to the OTA/AO classification, the Neer classification was integrated into the OTA/AO classification to facilitate clinician comprehension and optimize the best features of both classifications.^[10] The current study is the first study to test the 2018 OTA/AO proximal humerus classification. Our findings suggest that, in its fully detailed form, the 2018 OTA/AO classification is inferior in terms of inter-rater reliability to the Neer classification. However, in its short form, signifying the

fracture type only (A—extraarticular, unifocal, 2-part; B—extraarticular, bifocal, 3-part fracture; and C—articular or 4-part fracture) the inter-rater reliability is better or similar to the Neer classification. This finding is expected since in its short form the 2018 OTA/AO classification closely resembles the short form of the Neer classification. An additional important finding, however, was that 6 out of the 7 reviewers were significantly more satisfied with the ability of the 2018 OTA/AO classification to characterize the various fracture patterns correctly.

The current study demonstrates that both classification systems have advantages as well as significant drawbacks. We tested 2 versions of the Neer classification. The first version was an extended numerical version, where the numbers 1 to 16 signify different fracture patterns. This version has previously been used only for research purposes.^[3] The second version of the Neer classification that we tested was the short version, which is common in clinical practice and where we only counted the number of displaced parts in the fracture.^[2] The familiarity of most surgeons and trainees with this classification is its most significant advantage. However, as demonstrated in this and previous studies, the Neer classification falls short in its ability to describe the more complex fracture patterns. The 2018 OTA/AO classification is exceptionally suited for describing complex fractures. With its types, groups, sub-groups and sub-group qualifications, it offers 6 different unifocal (2-part) fracture types, 6 different bifocal extraarticular (3-part), and 9 multifocal intraarticular (4-part) fracture types. Considering the 14 different “universal modifiers” that account for factors such as displacement, dislocation, extension, bone quality and cartilage injury, 294 classifications are possible if we use 1 modifier and an endless amount if we use more than 1. This is an overwhelming number compared to the 16 possible classifications of the Neer classification, and can explain why the raters uniformly found that the 2018 OTA/AO classification better describes or characterizes the various fractures.

Furthermore, in its most basic (short) form, the 2018 OTA/AO classification retains and maybe slightly improves on the simplicity and inter-rater reliability of the Neer classification. When 1 considers the 4 primary reasons to use a classification—

communication, teaching, research and ease of coding, it appears that the short versions of Neer or OTA/AO are best for communication and ease of coding due to their superior inter-rater reliability. However, for teaching and research purposes and in order to best classify a particular fracture and correlate it to best management and outcome, the full 2018 OTA/AO classification may be the best choice and is supported by the selections of the reviewers in this work. The universality of the OTA/AO classification (i.e., its ability to classify all fractures according to similar principles) is another advantage it has over other classification schemes. Future studies will need to research the utility of the 2018 OTA/AO classification in preoperative planning (e.g., surgical approach, reduction strategy, implant choice) and outcome prediction

The study is the first to compare the 2018 OTA/AO proximal humerus classification to an existing classification. The number of raters and the use of multiple imaging modalities are additional strengths of this study. There are, however, also significant limitations to the study. The first limitation is that a single investigator chose the cases. The investigator, however, did not participate in classifying the fractures. Moreover, the final distribution of the fracture types by the raters was balanced as intended. Another limitation is the nonuniform quality of the images. The low quality of some of the x-ray and CT images may have contributed to the disagreement between reviewers. The authors felt, however, that this also made the cohort resemble “real-life” imaging and the clarity of the images affected both classification systems in a similar way. Another limitation may be the difference in time that different raters took to classify the fractures. Some raters completed their classification during the live zoom call, while others viewed the video recording of the zoom call to do their classifications. The authors do not feel that this skewed the results of the study since there were no time constraints to classify during the live zoom call. A sub-analysis of raters that only used video recordings for both rounds did not change the results. Classification of 2 test cases in group discussion addressed the lack of familiarity with the particulars of the classifications for some raters. The final important limitation is that 2 of the raters were primarily responsible for the 2018 update of the OTA/AO classification and all of the raters were members of the OTA classification committee and, therefore, may have a higher level of interest in this field and may have biased the results of the study in favor of the 2018 OTA/AO classification. This bias, if existent, was not severe since the study did not show a clear superiority of the 2018 OTA/AO classification and since the inter-rater and intra-rater reliabilities of both classifications were similar to those previously reported.

In conclusion, this study showed the equivalence of the short-form versions of the Neer and OTA/AO classifications, with superiority of the OTA/AO classification at describing specific fracture patterns. As a result, the authors believe that overall, this study supports using the 2018 OTA/AO classification over the Neer classification for classifying proximal humerus fractures. For clinical use and data entry into registries, the authors recommend using the short A, B, C (fracture type) version of the 2018 OTA/AO classification as it has the highest inter-rater reliability and it also incorporates the Neer classification. A is typically a 2-part fracture, B a 3-part fracture, and C a 4-part fracture. The authors recommend using the complete 2018 OTA/AO classification for teaching and research purposes, as it offers a much more detailed description of relevant features of the fracture. For research purposes, we recommend using advanced imaging and that 2 or more investigators classify the fractures

regardless of the choice of classification to increase reliability. Accurate classification of fracture patterns is critical to our ability to interpret the results of outcomes studies. However, the low inter-rater reliability of the complete 2018 OTA/AO classification is a concern that may need to be addressed by introducing computer-assisted classification aids in the future.

Acknowledgments

We would like to acknowledge the help of Christine Schreiber from the OTA Classification & Outcomes Committee in the facilitation of this study.

References

1. Neer CS. Displaced proximal humeral fractures. I. Classification and evaluation. *J Bone Joint Surg Am* 1970; 52:1077–1089.
2. Carofino BC, Leopold SS. Classifications in brief: the Neer classification for proximal humerus fractures. *Clin Orthop Relat Res* 2013; 471:39–43.
3. Sidor ML, Zuckerman JD, Lyon T, et al. The Neer classification system for proximal humeral fractures. An assessment of interobserver reliability and intraobserver reproducibility. *J Bone Joint Surg Am* 1993; 75:1745–1750.
4. Bernstein J, Adler LM, Blank JE, et al. Evaluation of the Neer system of classification of proximal humeral fractures with computerized tomographic scans and plain radiographs. *J Bone Joint Surg Am* 1996; 78:1371–1375.
5. Siebenrock KA, Gerber C. The reproducibility of classification of fractures of the proximal end of the humerus. *J Bone Joint Surg Am* 1993; 75:1751–1755.
6. Sjöden GO, Movin T, Güntner P, et al. Poor reproducibility of classification of proximal humeral fractures. Additional CT of minor value. *Acta Orthop Scand* 1997; 68:239–242.
7. Brorson S, Eckardt H, Audigé L, et al. Translation between the Neer and the AO/OTA-classification for proximal humeral fractures: do we need to be bilingual to interpret the scientific literature? *BMC Res Notes* 2013; 6:69.
8. Müller ME, Nazarian S, Koch P. Classification AO des Fractures: Les os Longs. Cham, Switzerland: Springer-Verlag; 1987.
9. Müller ME, Nazarian S, Koch P, et al. The Comprehensive Classification of Fractures of Long Bones. Cham, Switzerland: Springer Science & Business Media; 2012.
10. Meinberg EG, Agel J, Roberts CS, et al. Fracture and dislocation classification compendium-2018. *J Orthop Trauma* 2018; 32 Suppl 1: S1–170.
11. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol* 2011; 64:96–106.
12. Brorson S, Bagger J, Sylvest A, et al. Low agreement among 24 doctors using the Neer-classification; only moderate agreement on displacement, even between specialists. *Int Orthop* 2002; 26:271–273.
13. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; 17:101–110.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174.
15. Lund E, Whiting P, Sethi MK, Obremskey WT, Jahangir AA. Proximal humerus fractures. *Orthopedic Traumatology: An Evidence-Based Approach* Cham, Switzerland: Springer; 2018; 83–108.
16. Handoll HHG, Brealey SD, Jefferson L, et al. Defining the fracture population in a pragmatic multicentre randomised controlled trial: PROFHER and the Neer classification of proximal humeral fractures. *Bone Joint Res* 2016; 5:481–489.
17. Iordens GIT, Mahabier KC, Buisman FE, et al. The reliability and reproducibility of the Hertel classification for comminuted proximal humeral fractures compared with the Neer classification. *J Orthop Sci* 2016; 21:596–602.
18. Mahadeva D, Dias RG, Deshpande SV, et al. The reliability and reproducibility of the Neer classification system—digital radiography (PACS) improves agreement. *Injury* 2011; 42:339–342.
19. Fracture and Dislocation Compendium Orthopaedic Trauma Association committee for coding and classification. *J Orthop Trauma* 1996; 10 Suppl 1:v–ix. 1–154.