RBO REVISTA BRASILEIRA DE ORTOPEDIA
www.rbo.org.br

Original Article

Evaluation of the reproducibility of the AO/ASIF classification for humeral shaft fractures☆

Gustavo Soriano Pignataro, André Elias Junqueira, Fabio Teruo Matsunaga, Marcelo Hide Matsumoto, João Carlos Belloti, Marcel Jun Sugawara Tamaoki*

Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil

ARTICLE INFO

Article history:
Received 8 May 2014
Accepted 25 June 2014
Available online 26 July 2015

Keywords:
Humeral fractures/classification
Humeral fractures/radiograph
Humeral fractures/surgery

ABSTRACT

Objective: To evaluate the reproducibility of the AO/Asif classification for humeral shaft fractures.
Methods: Consecutive radiographs of the arm in both anteroposterior and lateral view from 60 patients with humeral shaft fractures were analyzed. Six observers who were familiar with the AO/Asif classification (three shoulder and elbow surgery specialists and three general orthopedists) were selected to make the analysis, which was done at three different times. The data were subjected to statistical analysis using the kappa coefficient.
Results: The intra and interobserver concordance was statistically significant in all the analyses.
Conclusions: All the evaluators showed concordance between the three evaluations that was considered to be statistically significant. However, the highest values were found among the specialists.

Avaliação da reprodutibilidade da classificação AO/Asif para fraturas diafisárias do úmero

Palavras-chave:
Fraturas do úmero/classificação
Fraturas do úmero/radiografia
Fraturas do úmero/cirurgia

RESUMO

Objetivo: Avaliar a reprodutibilidade da classificação AO/Asif para as fraturas diafisárias do úmero.
Métodos: Foram analisadas radiografias consecutivas em duas incidências (anteroposterior e perfil do braço) de 60 pacientes com fratura do úmero diafisário. Seis observadores familiarizados com a classificação AO/Asif, três especialistas em cirurgia do ombro e cotovelo e três ortopedistas gerais foram selecionados para análise, a qual se deu em três tempos distintos. Os dados foram submetidos à análise estatística com o coeficiente *kappa* (κ).

☆ Work developed in the Discipline of Hand and Upper-limb Surgery, Department of Orthopedics and Traumatology, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.

* *Corresponding author.*
E-mail: gustavosoriano72@yahoo.com.br (G.S. Pignataro).

http://dx.doi.org/10.1016/j.rboe.2015.07.003

Resultados: A concordância intra e interobservadores foi estatisticamente significante em todas as análises.

Conclusões: Todos os avaliadores concordam com as três avaliações consideradas estatisticamente significantes. Porém, os maiores valores são encontrados entre os especialistas.

Introduction

Diaphyseal fractures of the humerus account for 1–3% of all fractures among adults[1,2] and 20% of all fracture of the humerus.[3] The annual incidence is 13 to 14.5 per 100,000 individuals.[4,5] Non-surgical treatment is still the standard for cases of solely diaphyseal fracturing of the humerus.[6,7] On the other hand, surgical treatment is recommended for patients with neurovascular, medullary or brachial plexus lesions, those with exposed fractures, multiple trauma victims and individuals with a floating elbow or unsatisfactory reduction.[8–11] In addition, it is indicated in cases of certain types of fracture that are more unstable: according to AO–OTA, these are type A fractures and oblique fractures of the proximal and distal thirds.[12–14]

The AO classification is an alphanumeric system for all fractures that was created in 1986.[15] In categorizing fractures, this system takes into consideration the bone affected, the region and type of fracture line.

Fractures of the long bones are more common as traumatic injuries than are fractures of the periarticular areas.[16] Several classification standards have been described for fractures. However, diaphyseal fractures are almost exclusively identified in accordance with the AO/ASIF classification system.[17]

An ideal classification system should provide guidance for treatment, indicate possible complications and thus make prognostic predictions for the fracture. In addition, the classification has the functions of standardizing the communication language and providing a mechanism that enables comparisons of the results obtained for a given type of fracture by different centers, in evaluations in the literature that were made at different times. For this reason, it is essential that this system should be valid, reliable and reproducible.

However, the revised version of the AO/ASIF classification system has been criticized as having low reproducibility and inter and intraobserver concordance.[15]

In relation to fractures of the diaphyseal region of the humerus, there have not been any studies testing intra and interobserver concordance regarding these fractures, to the best of our knowledge.[15,18,19]

The objective of the present study was to evaluate intra and interobserver concordance regarding the AO/ASIF classification for diaphyseal fractures of the humerus.

Method

Consecutive radiographs in two views (anteroposterior and lateral views of the arm) of 60 patients with diaphyseal fractures of the humerus were analyzed. These were numbered and the patients' names and ages were concealed. Cases of fractures in patients with an immature skeleton, cases of pathological fractures or cases in which the patient presented previous surgery in the body segment concerned were excluded. The quality of the images was determined by two orthopedists who were not involved in evaluating the concordance. The radiographs were accepted and included in the study only when both of these evaluators considered them to be acceptable.

Six observers who were familiar with the AO/ASIF classification system were selected to perform the analysis. Among these observers, three shoulder and elbow surgery specialists (SES) and three general orthopedists (GO) were chosen. So that the information from all the observers would be standardized, a self-explanatory illustrated diagram of the AO/ASIF classification was handed out to each of the observers. The names and identifications present on the radiographs were covered up and the radiographs were numbered randomly. Each observer classified each fracture in accordance with the AO/ASIF system at three different times. At the first evaluation (T1), the computer-digitized radiographic images were viewed in numerical sequence. Three weeks later, at the second evaluation (T2), the sequence of the radiographs was randomly modified. This was done again for the third evaluation (T3), six weeks after the first evaluation. This randomization sequence was kept secret by an individual who was not involved in evaluating the images.

The data were gathered together on spreadsheets and the kappa coefficient (κ) was used for the analysis, using the method proposed by Fleiss.[20] This made it possible not only to calculate the concordance that would be expected by chance, but also the concordance between multiple observers (i.e. more than two) in evaluating the nominal variables. The κ concordance coefficient provides paired proportions of concordance between the observers, who may have made correct observations. The κ values range from -1 to $+1$; values between -1 and 0 indicate that the observed concordance is less than what would be expected by chance, 0 indicates the level of concordance achievable by chance and $+1$ indicates total concordance. In general terms, κ values lower than 0.5 are considered to be unsatisfactory, between 0.5 and 0.75 fair to satisfactory and higher than 0.75 excellent.[15–18] For this study, we defined a significance level (i.e. the extent to which error in the statistical conclusions would be accepted or the statistical error that would be made in the analyses) of 0.05 (5%) and a confidence interval of 95%.

This project received prior approval from our institution's ethics committee (CEP: 451507), on February 3, 2014.

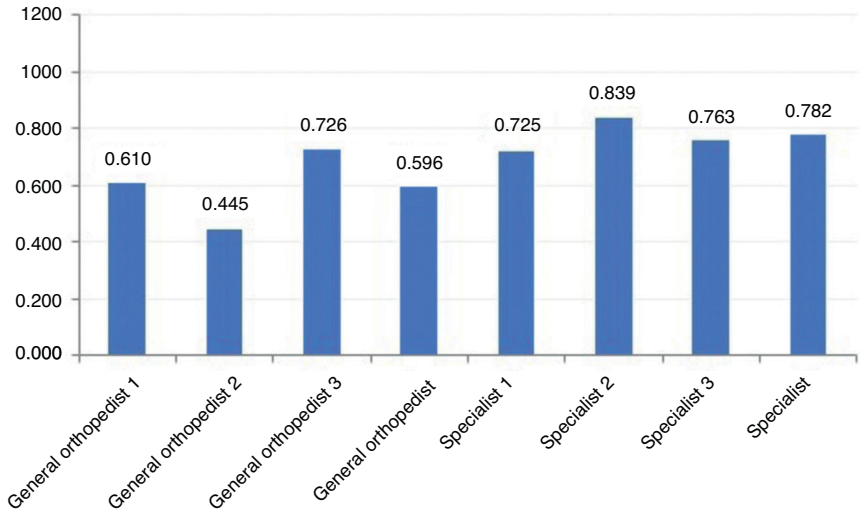

Fig. 1 – Intraobserver concordance for the three evaluations.

Results

The intra and interobserver concordance values were obtained by calculating the κ coefficient, as proposed by Fleiss.[20]

Fig. 1 shows the intraobserver concordance, taking into consideration the three evaluation times, both for the group of general orthopedists and for the specialists. All the evaluators presented concordance between the three evaluations, which were considered to be statistically significant. However, higher values were found among the specialists, and the best among them was for specialist 2, with κ of 0.839, which was considered excellent.

Fig. 2 shows the intraobserver concordance of the three evaluations in pairs.

All of these concordance values were statistically significant (different from zero). It was again seen that the highest concordance occurred in relation to specialist 2, between the second and third evaluations, with a value of 0.980, which was classified as excellent. The intraobserver concordance among the general orthopedists was satisfactory (κ between 0.568 and 0.626), while among the shoulder and elbow specialists, it was excellent (κ between 0.761 and 0.821).

Fig. 3 brings together the general orthopedists and the specialists, to measure the degree of concordance between the groups at each time (as three evaluations) and in general (all the three evaluations together). At all times, there was statistically significant concordance. The highest concordance was in relation to the first evaluation, with a value of 0.819, which was classified as excellent. From grouping the three evaluations, the interobserver concordance was satisfactory ($\kappa = 0.539$).

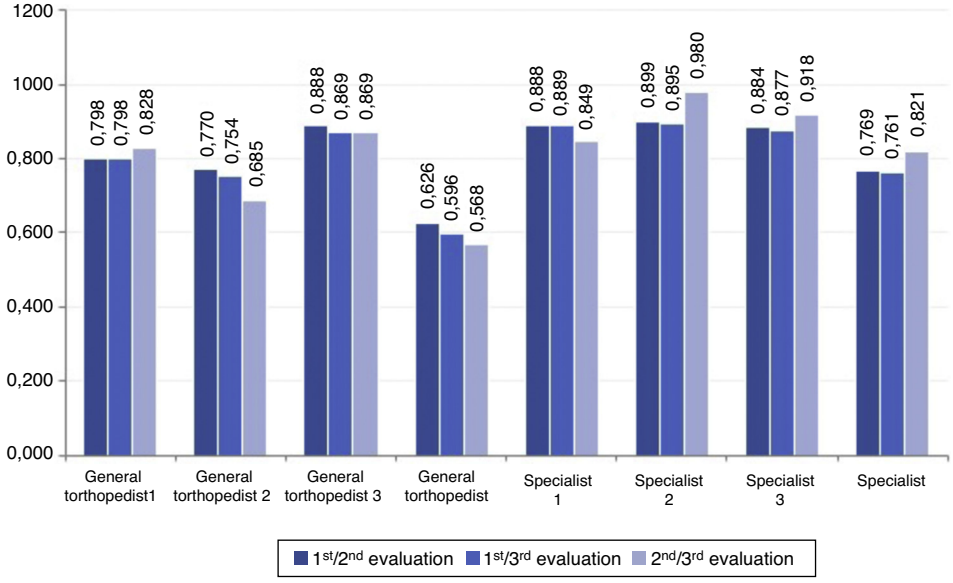

Fig. 2 – Intraobserver concordance for the three evaluations in pairs.

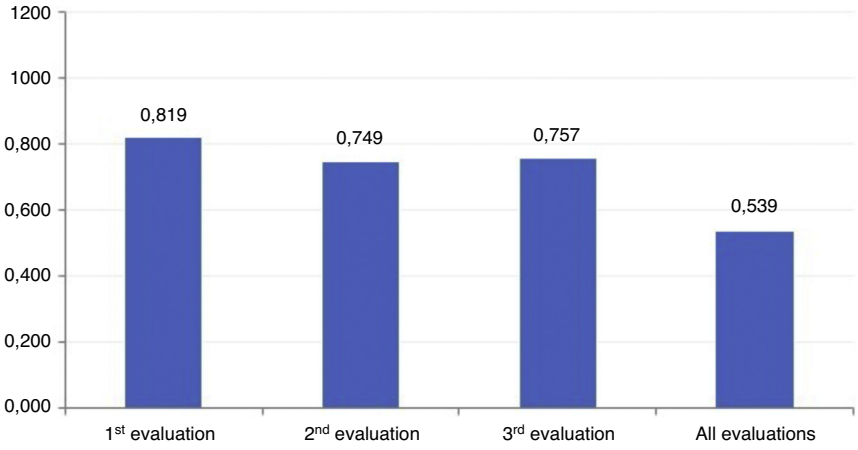

Fig. 3 – Interobserver concordance according to evaluation.

Discussion

Several classification standards for fractures have been described. However, diaphyseal fractures are almost exclusively identified in accordance with the AO/ASIF classification system.[16]

These classification systems are very important within orthopedic practice because they serve to describe the fracture and direct the treatment, as well as standardizing the injuries scientifically and making the treatment reproducible. Therefore, intra and interobserver concordance become essential for any classification system.

In the analysis on the intraobserver concordance, taking the three times into consideration, the mean from the AO classification was satisfactory for the general orthopedists ($\kappa = 0.596$), and particularly so for the specialists ($\kappa = 0.782$). These results were probably due to the simplicity and practicality of this classification. Greater experience among the orthopedists influenced greater reliability.

In analyzing the intraobserver concordance between T1 and T2 and between T2 and T3, it was noted that in the general orthopedist group there was a decrease in the values, while among the specialists there was an increase, probably due to the specialists' experience of analyzing and classifying these fractures, which the other orthopedists did not have. In measuring the degree of concordance between the groups at each time, it was seen that greatest concordance was in relation to the first evaluation. We concluded that such conditioning did not occur generally.

Through this study, it could be seen that the AO/ASIF classification for diaphyseal fractures of the humerus presents good intra and interobserver reproducibility. The greater the familiarity with and applicability of this classification were, the greater the degree of reliability also was.

It is important to emphasize that this study is limited to evaluating the concordance between the opinions of the observers and it was not possible to measure the accuracy of this classification. For this purpose, a diagnostic study comparing each observer with the result from evaluation using a standard diagnostic examination (with high sensitivity and specificity for proving the diagnosis) would be necessary.

Conclusion

The AO/ASIF classification system presented satisfactory intraobserver concordance among the general orthopedists and excellent concordance among the shoulder and elbow specialists, and satisfactory interobserver concordance.

Conflicts of interest

The authors declare no conflicts of interest.

REFERENCES

1. Emmett JE, Breck LW. A review and analysis of 11,000 fractures seen in a private practice of orthopaedic surgery, 1937–1956. J Bone Joint Surg Am. 1958;40(5):1169–75.
2. Schemitsch EH, Bhandari M, Talbot M. Fractures of the humeral shaft. In: Skeletal trauma: basic science, management and reconstruction. 4th ed. Philadelphia: Saunders; 2008. p. 1593–4.
3. Rose SH, Melton LJ 3rd, Morrey BF, Ilstrup DM, Riggs BL. Epidemiologic features of humeral fractures. Clin Orthop Relat Res. 1982;(168):24–30.
4. Brinker MR, O'Connor DP. The incidence of fractures and dislocations referred for orthopaedic services in a capitated population. J Bone Joint Surg Am. 2004;86(2):290–7.
5. Ekholm R, Adami J, Tidemark J, Hanson K, Törnkvist H, Ponser S. Fractures of the shaft of the humerus: an epidemiological study of 401 fractures. J Bone Joint Surg Br. 2006;88(11):1469–73.
6. Sarmiento A, Kinman PB, Galvin EG, Schmitt RH, Phillips JG. Functional bracing of fractures of the shaft of the humerus. J Bone Joint Surg Am. 1977;59(5):596–601.
7. Balfour GW, Marrero CE. Fracture brace for the treatment of humerus shaft fractures caused by gunshot wounds. Orthop Clin North Am. 1995;26(1):55–63.

8. Wallny T, Westermann K, Sagebiel C, Reimer M, Wagner UA. Functional treatment of humeral shaft fractures: indications and results. J Orthop Trauma. 1997;11(4):283–7.
9. Amillo S, Barrios RH, Martinez-Peric R, Losada JI. Surgical treatment of the radial nerve lesions associated with fractures of the humerus. J Orthop Trauma. 1993;7(3):211–5.
10. Foster RJ, Swiontkowski MF, Bach AW, Sack JT. Radial nerve palsy caused by open humeral shaft fractures. J Hand Surg. 1993;18(1):121–4.
11. Pollock FH, Drake D, Bovill EG, Day L, Trafton PG. Treatment of radial neuropathy associated with fractures of the humerus. J Bone Joint Surg Am. 1981;63(2):239–43.
12. Ekholm R, Tidermark J, Törnkvist H, Adami J, Ponzer S. Outcome after closed functional treatment of humeral shaft fractures. J Orthop Trauma. 2006;20(9):591–6.
13. Rutgers M, Ring D. Treatment of diaphyseal fractures of the humerus using a functional brace. J Orthop Trauma. 2006;20(9):597–601.
14. Jawa A, McCarty P, Doornberg J, Harris M, Ring D. Extra-articular distal third diaphyseal fractures of the humerus. A comparison of functional bracing and plate fixation. J Bone Joint Surg Am. 2006;88(11):2343–7.
15. Swiontkowski MF, Agel J, McAndrew MP, Burgess AR, MacKenzie EJ. Outcome validation of the AO/OTA fracture classification system. J Orthop Trauma. 2000;14(8): 534–41.
16. Muller M, Nazarian S, Koch P, Schatzker J. The comprehensive classification of fractures of long bones. Berlin, Germany: Springer-Verlag; 1990.
17. Fracture and dislocation compendium. Orthopaedic Trauma Association Committee for Coding and Classification. J Orthop Trauma. 1996;10 Suppl. 1:1–154.
18. Johnstone DJ, Radford WJ, Parnell EJ. Interobserver variation using the AO/ASIF classification of long bone fractures. Injury. 1993;24(3):163–5.
19. Newey ML, Ricketts D, Roberts L. The AO classification of long bone fractures: an early study of its use in clinical practice. Injury. 1993;24(5):309–12.
20. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378–82.