Injury, Int. J. Care Injured 45S5 (2014) S18–S20

Contents lists available at SciVerse ScienceDirect

Injury

journal homepage: www.elsevier.com/locate/injury

Reproducibility of classifications for olecranon fractures

Marcel Jun Sugawara Tamaoki, Fabio Teruo Matsunaga*, Juliana Doering Silveira, Daniel Balbachevsky, Marcelo Hide Matsumoto, João Carlos Belloti

Department of Orthopaedics and Traumatology, Universidade Federal de São Paulo — Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil

KEYWORDS

Olecranon process
Ulna fractures
Reproducibility of results
Validation studies

ABSTRACT

Introduction: Fractures of the olecranon are relatively common injuries in adults and are of great clinical importance. Classification systems have been developed as tools to assist surgeons in grouping different types of fractures, to facilitate communication and to standardise treatment, but none of the systems used today is universally accepted for olecranon fractures.
Methods: Fifty-nine olecranon fractures were classified according to the Schatzker, Colton, Mayo and AO/ASIF systems by four observers with different levels of expertise. Intra- and inter-observer agreement was assessed. Each observer analysed the images at three different times; the images were randomised and presented in a different sequence at each assessment.
Results: There was higher mean intra-observer agreement in the AO/ASIF (0.60) and Mayo (0.64) classifications compared with the Schatzker (0.49) and Colton (0.38) classifications. Inter-observer agreement was better with AO/ASIF and Mayo (0.35 and 0.32, respectively) than with Schatzker and Colton (0.29 and 0.12, respectively).
Conclusion: The results of this study indicate that the most commonly used classifications for olecranon fractures are associated with low reproducibility.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Olecranon fractures are relatively common in adults and account for about 10% of all elbow fractures [1] and 20% of fractures of the proximal forearm [2]. Fractures of the olecranon play an important role in fractures of the elbow, particularly because potential complications associated with the treatment of olecranon fractures sometimes lead to severe functional deficit due to pain and loss of mobility [3-6].

Different methods of internal fixation have been described for the treatment of these fractures. The characteristics of the fracture are important to indicate the best fixation technique: tension band, intramedullary fixation, excision of the fragment with triceps reattachment, or plate and screws osteosynthesis [1]. The morphology of the fractures is also a major predictor for function and development of osteoarthritis of the elbow after surgery [7].

A reliable classification is essential to guide appropriate treatment. Classification systems have been developed as tools to assist surgeons in grouping different types of fractures, to facilitate communication and to standardise treatments, but none of the systems used today is universally accepted for olecranon fractures [8].

The AO/ASIF (Association Arbeitsgemenschaft Für Osteosynthesefragen) system divides the fractures of the proximal radius and ulna into three broad categories. Type A fractures are extra-articular fractures that involve the metaphysis of the radius or ulna; type B fractures are intra-articular fractures of the radius or the ulna, and type C fractures are intra-articular fractures of both the radial head and olecranon [9].

Morrey [10] described the Mayo classification system of fractures of the olecranon, which is based on the degree of displacement, stability and articular comminution. Type I fractures are undisplaced with minimal or no comminution; type II are displaced fractures, but with stability of the elbow joint, with anterior articular surface integrity to maintain stability, and type III are fractures with unstable elbow involving a major part of the olecranon. Types I, II and III are sub-divided into type A (not comminuted) and type B (comminuted).

The Schatzker classification for olecranon fractures [11] divides the fractures into six types: type A, simple, transverse fracture; type B, transverse fracture with depression of the central part of the olecranon; type C, simple oblique fracture; type D, comminuted fracture; type E, oblique fracture distal to the midpoint of the trochlea; and type F, associated with radial head fracture.

The Colton classification system [12] divides the fractures into two major groups: undisplaced (type I) and displaced (type II). Type II displaced fractures are subdivided into type IIA, avulsion; IIB, oblique or transverse; IIC, comminuted; and IID, fracture-dislocations.

* Corresponding author at: Rua Borges de Lagoa, 786, São Paulo (SP) — Brasil — CEP 04038-002. Tel./fax: (+55 11) 5579-7049.
E-mail address: fteruo@gmail.com (F. Teruo Matsunaga).

0020-1383/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.

A good classification system must be valid, reliable and reproducible. The ideal classification system should standardise a language for reliable communication, provide guidelines for treatment, indicate the possibility of complications, and enable surgeons to predict prognosis for each specific fracture. This ideal system should also provide a mechanism to assess and compare the results with treatment of similar fractures at a variety of centres and at different times [12].

There is currently no consensus on which is the best classification system in terms of reproducibility for fractures of the olecranon and there are no similar studies to assess these fractures in the literature.

The aim of this study, therefore, was to assess intra- and inter-observer reproducibility of the most commonly used classification systems for fractures of the olecranon.

Patients and methods

Fifty-nine fractures of the olecranon in 59 patients aged over 18 years who were treated at the Shoulder and Elbow Surgery Group of Hand and Upper Limb Surgery Clinic - Universidade Federal de Sao Paulo, were evaluated from March 2012 to January 2013. This study was approved by the Institutional Research Ethics Committee.

Initial elbow radiographs were digitised, numbered and ordered in three different series for the analysis.

Four observers participated in the study: one sixth-year student of medicine (S6), one third-year resident of orthopaedics (R3), one orthopaedic trauma surgeon (TS) and one hand and upper limb surgeon (HS).

Observers viewed elbow radiographs in two incidences (anterior-posterior and lateral) and used four systems to evaluate each fracture. Each observer received illustrated diagrams with a description and explanation of each classification system. Each pair of incidences was numbered. The names and ages of the patients were omitted.

Radiographs of patients who presented with open physis or pathological fractures or who had undergone previous elbow surgery were excluded. No communication was allowed between observers.

At first assessment (T1), the radiographs were presented in a particular sequence. Three weeks later a second assessment (T2) was performed, and the radiographs were randomised and presented to observers in a new sequence, this was repeated in the third assessment (T3), six weeks after T1.

All data in this study were used for research purposes only and all patient data were kept confidential.

Statistical analysis

The *kappa* coefficient provides a paired proportion of agreement between observers. Using weighted *kappa*, the level of agreement is obtained by the proportion of agreement that would be expected if the agreement were by chance. The weighted *kappa* takes into account the order of variables, if a gradient between them is found. The relative importance between categories may not be the same between adjacent and distant categories [13-16].

This measure of agreement has the maximum value of 1, where 1 represents total agreement, while values close to and below 0 indicate no agreement or the agreement is exactly what is expected by chance. An eventual negative *kappa* indicates that the correlation found was less than that expected by chance. It therefore suggests disagreement, but its value may not be interpreted as the intensity of disagreement. In general, values of *kappa* coefficient less than 0.5 are considered unsatisfactory; from 0.5 to 0.75 are considered satisfactory and higher than 0.75 are excellent [17].

SPSS V17, Minitab 16 and Excel Office 2010 software were used for statistical analysis.

Table 1
Kappa coefficients for intra-observer comparisons between the three time points (T1, T2 and T3)

	Classification			
Observer	AO/ASIF	Mayo	Schatzker	Colton
S6	0.74	0.43	0.34	0.47
R3	0.52	0.77	0.81	0.47
TS	0.27	0.73	0.10	0.06
HS	0.87	0.64	0.71	0.50
Mean	0.60	0.64	0.49	0.38

Table 2
Kappa coefficients for inter-observer comparisons at each time point (T1, T2 and T3)

	Classification			
Time point	AO/ASIF	Mayo	Schatzker	Colton
T1	0.50	0.50	0.36	0.17
T2	0.30	0.23	0.20	0.08
T3	0.27	0.26	0.30	0.12
Mean	0.36	0.33	0.29	0.12

Results

Table 1 summarises the values found in the intra-observer kappa analysis for each observer in the three analysis times. The best agreement was found in the Mayo and AO/ASIF classifications, and the worst results were obtained in the Colton system.

Table 2 shows the kappa values for inter-observer comparisons obtained in each of the evaluation times for each classification. Again, the best agreement was obtained in the Mayo and AO/ASIF classifications and the worst results were obtained in the Colton classification.

Discussion

Classification systems are very important in orthopaedic practice because they are used to describe fractures, facilitate communication between orthopaedic surgeons, guide treatment and compare treatments through studies in the literature. Reproducibility of these classification systems in terms of intra- and inter-observer agreements is essential to achieve these goals.

The four classification systems used in this study were chosen because they are very popular and commonly used among orthopaedic surgeons.

Unlike other classifications, such as those used in the proximal humerus [18,19] or distal radius [20-22], the systems used in elbow injuries are poorly studied [23,24].

In the analysis of intra-observer agreement, the mean *kappa* value was satisfactory for the ratings of the AO/ASIF (0.60) and Mayo (0.64) and unsatisfactory for Schatzker (0.49) and Colton (0.38). This probably reflects the difficulty in differentiating radiographic fracture patterns and concepts used for grouping fractures, such as oblique fractures and differentiation between comminution and central depression.

The great variability in the results probably relates to the complexity of classification systems; similar results have been presented in studies of classification systems for fractures of other bones [19,20,22].

All ratings for inter-observer agreement were unsatisfactory, with the Colton classification having the worst index. Again, this low agreement is because of the difficulty in definition and conceptualisation of patterns.

A study of elbow fracture-dislocations showed that the identification of a posterior instability with the olecranon fracture shows good inter-observer correlation; however, this study did not assess olecranon fractures alone [24].

A limitation of this study is that there was no measurement of the time spent by each examiner to perform the analyses, and the working conditions of the examination were not recorded; therefore, it is not possible to determine if such conditions favoured the attention of observers during the analysis. Another important aspect would be to highlight the characteristics of each classification system that can lead to greater disagreement, and whether or not this factor influences the decision of treatment.

Recent studies have shown that the use of higher definition scans (PEX 3D CT) may increase the agreement surrounding these classifications [25,26]. A study using these higher definition scans and the classifications used in the current study should be performed to verify if there is an increase in reproducibility.

The aim of this study was to assess only the reproducibility of classifications. It is also necessary to determine the influence of the classifications on the treatment of fractures, and the accuracy of the diagnosis with different imaging tests.

Conclusion

Intra-observer agreement with the Mayo and AO/ASIF classifications were considered satisfactory. Mayo and AO/ASIF classification systems showed better reproducibility compared with other classifications.

Conflict of Interest Statement

This study does not present conflicts of interest among the participants.

References

[1] Newman SD, Mauffrey C, Krikler S. Olecranon fractures. Injury 2009;40:575-81.
[2] Duckworth AD, Clement ND, Aitken SA, Court-Brown CM, McQueen MM. The epidemiology of fractures of the proximal ulna. Injury 2012;43:343-6.
[3] Edwards SG, Martin BD, Fu RH, Gill JM, Nezhad MK, Orr JA, Ferrucci AM, Fraser J, Singer A, Hsieh AH. Quantifying and comparing torsional strains after olecranon plating. Injury 2012;43:712-7
[4] Schmidt-Horlohé K, Wilde P, Bonk A, Becker L, Hoffmann R. One-third tubular-hook-plate osteosynthesis for olecranon osteotomies in distal humerus type-C fractures: a preliminary report of results and complications. Injury 2012;43:295-300.
[5] Sabalic S, Kodvanj J, Pavic A. Comparative study of three models of extra-articular distal humerus fracture osteosynthesis using the finite element method on an osteoporotic computational model. Injury2013;44 Suppl 3:S56-6
[6] Brink PR, Windolf M, de Boer P, Brianza S, Braunstein V, Schwieger K. Tension band wiring of the olecranon: is it really a dynamic principle of osteosynthesis? Injury 2013;44:518-22.
[7] Rommens PM, Küchle R, Schneider RU, Reuter M. Olecranon fractures in adults: factors influencing outcome. Injury 2004;35:1149-57.
[8] Wiegand L, Bernstein J, Ahn J. Fractures in brief: Olecranon fractures. Clin Orthop Relat Res 2012;470:3637-41.
[9] Müller ME, Allgöwer M, Schneider R, Willenegger H (eds): Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group, 3rd ed. Berlin: Springer-Verlag, 1991.
[10] Morrey BF. Current concepts in the treatment of fractures of the radial head, the olecranon, and the coronoid. J Bone Joint Surg Am 1995;77:316–27.
[11] Schatzker J. Fractures of the olecranon, in Schatzker J, Tile M (eds): The Rationale of Operative Fracture Care. Berlin: Springer-Verlag; 1987, p89–95.
[12] Browner BD, Levine AM, Jupiter JB, Trafton PG, Krettek C. Skeletal trauma: Basic science, management, and reconstruction. 4th ed. Saunders-Elsevier; 2009;42, p1525-6.
[13] Fleiss JL, Slakter MJ, Fischman SL, Park MH, Chilton NW. Inter-examiner reliability in caries trials. J Dent Res 1979;58;604-9.
[14] Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
[15] Svanholm H, Starklint H, Gundersen HJ, Fabricius J, Barlebo H, Olsen S. Reproducibility of histomorphological diagnosis with special reference to the kappa statistic. APMIS 1989;97:689-98.
[16] Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York, John Wiley and Sons;1981.
[17] Martin JS, Marsh JL, Bonar SK, DeCoster TA, Found EM, Brandser EA. Assessment of the AO/ASIF fracture classification for the distal tibia. J Orthop Trauma 1997;11:477-83.
[18] Sjödén GO, Movin T, Güntner P, Aspelin P, Ahrengart L, Ersmark H, et al. Poor reproducibility of classification of proximal humeral fractures. Additional CT of minor value. Acta Orthop Scand 1997;68:239-42.
[19] Foroohar A, Tosti R, Richmond JM, Gaughan JP, Ilyas AM. Classification and treatment of proximal humerus fractures: inter-observer reliability and agreement across imaging modalities and experience. J Orthop Surg Res 2011;6:38.
[20] Belloti JC, Tamaoki MJ, Franciozi CE, Santos JB, Balbachevsky D, Chap Chap E, et al. Are distal radius fracture classifications reproducible? Intra and interobserver agreement. Sao Paulo Med J 2008;126:180-5.
[21] Kreder HJ, Hanel DP, McKee M, Jupiter J, McGillivary G, Swiontkowski MF. Consistency of AO fracture classification for the distal radius. J Bone Joint Surg Br 1996;78:726-31.
[22] Flikkila T, Nikkola-Sihto A, Kaarela O, Paakko E, Raatikainen T. Poor interobserver reliability of AO classification of fractures of distal radius. J Bone Joint Surg Br 1998;80:670-2.
[23] Matsunaga FT, Tamaoki MJ, Cordeiro EF, Uehara A, Ikawa MH, Matsumoto MG, et al. Are classifications of proximal radius fractures reproducible? BMC Musculoskelet Disord 2009;10:120.
[24] Doornberg JN, Guitton TG, Ring D; Science of Variation Group. Diagnosis of elbow fracture patterns on radiographs: interobserver reliability and diagnostic accuracy. Clin Orthop Relat Res 2013;471:1373-8.
[25] Guitton TG, Kinaci A, Ring D. Diagnostic accuracy of 2- and 3-dimensional computed tomography and solid modeling of coronoid fractures. J Shoulder Elbow Surg 2013;22:782-6.
[26] Doornberg J, Lindenhovius A, Kloen P, van Dijk CN, Zurakowski D, Ring D. Two- and three-dimensional computed tomography for the classification and management of distal humeral fractures. Evaluation of reliability and diagnostic accuracy. J Bone Joint Surg Am 2006;88:1795-801.