

CLASSIFYING RADIUS FRACTURES WITH X-RAY AND TOMOGRAPHY IMAGING

PAULO ROBERTO MIZIARA YUNES FILHO, MIGUEL VIANA PEREIRA FILHO, FABIANO CORTESI PAULA GOMES, RODRIGO SERIKAWA DE MEDEIROS, EMYGDIO JOSÉ LEOMIL DE PAULA, RAMES MATTAR JUNIOR, ARNALDO VALDIR ZUMIOTTI

ABSTRACT

Introduction: This study evaluated the interobserver reliability of plain radiography versus computed tomography (CT) for the Universal and AO classification systems for distal radius fractures. **Patients and methods:** Five observers classified 21 sets of distal radius fractures using plain radiographs and CT independently. Kappa statistics were used to establish a relative level of agreement between observers for both readings. **Results:** Interobserver agreement was rated as moderate for the Universal classification and poor for the AO classification. Reducing the AO system to 9 categories and to its three main types reliability was raised to a "moderate" level. No difference

was found for interobserver reliability between the Universal classification using plain radiographs and the Universal classification using computed tomography. Interobserver reliability of the AO classification system using plain radiographs was significantly higher than the interobserver reliability of the AO classification system using only computed tomography. **Conclusion:** From these data, we conclude that classification of distal radius fractures using CT scanning without plain radiographs is not beneficial.

Keywords: Wrist injuries. X-Ray computed tomography. Reproducibility of results.

Citation: Yunes Filho PRM, Pereira Filho MV, Gomes FCP, Medeiros RS, De Paula E JL, Mattar Junior R, Zumiotti AV. Classifying radius fractures with x-ray and tomography imaging. *Acta Ortop Bras.* [online]. 2009;17(2):9-13. Available from URL: <http://www.scielo.br/aob>.

INTRODUCTION

Fractures are a public health problem affecting a significant portion of the population. In a study published in Scotland in 2006, an incidence of 11.67/1000 fractures/year was evidenced in men and of 10.65/1000/year in women. The most frequent ones were those located on radius distal third – 1.95/1000/year.¹ Another study, conducted in Sweden in 2007 found an even greater incidence: 2.6/1000/year.²

The way in which these injuries are treated has dramatically changed over the last two decades: from the almost universal use of plastered cast to a large variety of surgical techniques.³ These changes happened after the importance of appropriately restoring radius distal end joints was proven in order to improve prognosis of these fractures.⁴

Joint congruence and other parameters employed on therapeutic decision about these fractures are mainly evaluated by plain X-ray images. The importance of using computed tomography for a more accurate evaluation of these parameters is being discussed by current literature.⁵⁻¹²

Studies on inter- and intraobserver reliability of classifications for radius distal fractures have presented a wide range of reliability rates.^{6,13-18} Studies assessing specific parameters such as step and diastasis have also presented conflicting results.^{5,7,9,19} Most studies assessing reliability on classifications of radius distal fractures in-

volve the use of plain X-ray images only. Few studies use computed tomography images for this kind of assessment.^{6,8,20} In Brazil, we didn't find any interobserver reliability study for Universal and AO classifications using computed tomography images, according to our search on Pubmed, Lilacs and Embase databases with the keywords classification, tomography and radius.

The purpose of this study is to investigate interobserver reliability for AO and Universal classifications using plain X-ray and computed tomography images on patients with radius distal third fractures.

MATERIALS AND METHODS

We captured X-ray and tomography images of 21 adult patients of both genders with radius distal fractures. Only patients with acute fractures not previously treated were included. X-ray images were taken at anteroposterior and lateral planes, while tomography images were captured at sagittal, coronal and axial planes. The images were evaluated by five 3rd year resident Orthopaedic and Traumatology doctors. Fractures were first classified from the X-ray images preventing patient identification and at random order. Then, fractures were classified based on tomography images, also without any patient identification and at random order, so that the observer could not associate X-ray images with tomography images. The AO classification (Chart 1) and Universal classification (Chart 2) were used in this study.^{21,22}

All authors state no potential conflict of interests concerning this article.

Department of Orthopaedics and Traumatology, University of São Paulo Medical School, Hospital das Clínicas, and Musculoskeletal Investigation Laboratory (LIM 41)
Correspondences to: Rua Antenor Guirlanda, 92, apto 122, Casa Verde, São Paulo, SP, Brasil, CEP 02514-010, e-mail: paulorbertomiziara@uol.com.br

Received in 09/26/07; approved in 09/26/07

The AO classification was assessed by stratification. For this, we defined three detail levels: first level corresponding to types A, B or C, that is, it seeks to assess the reproducibility of classification when an observer only needs to determine if the fracture is extra-joint, partially intra-joint or fully intra-joint; the second level corresponds to the 9 subtypes from A1 to C3, while the third level is represented by the full classification, with its 27 sub-items: from A1.1 to C3.3. (Chart 1)

STATISTICAL ANALYSIS

These data were assessed by using the Kappa statistical method. The Kappa coefficient is employed to assess consistency between observers, subtracting the consistency that would be attributed to fate. The interpretation of values was made as recommended by Landis and Koch²³ and as all studies employing Kappa coefficient addressing the topic have used. Kappa values above 0.8 indicate an excellent consistency; between 0.61 and 0.8, a good reproducibility; between 0.41 and 0.60, moderate reproducibility/ between 0.21 and 0.4, low reproducibility; while between zero and 0.2, a poor reproducibility. Negative values represent inconsistency.

Chart 1 - AO classification for radius distal end fractures⁽²²⁾

A1 – extra-joint ulnar fracture with intact radius
A1.1 Styloid process
A1.2 Ulnar metaphysis
A1.3 Fragmented ulnar metaphysis
A2 – extra-joint radius fracture, simple or impacted
A2.1 Without deviation
A2.2 With dorsal deviation (<i>Pouteau-Colles</i>)
A2.3 With volar deviation (<i>Goyran-Smith</i>)
A3 - extra-joint multiple fragmented fracture
A3.1 Impacted with axial shortening
A3.2 Impacted with wedge
A3.3 Complex (metaphyseal comminution)
B1 – partial joint fracture with trace at sagittal plane
B1.1 Simple lateral (<i>Chauffeur</i>)
B1.2 Multiple fragmented
B1.3 Medial
B2 – partial joint fracture with dorsal fragment (<i>Barton</i>)
B2.1 Simple
B2.2 With lateral fracture and sagittal trace
B2.3 With dorsal carpus displacement
B3 – partial joint fracture with volar fragment (<i>Barton reverse</i>)
B3.1 Simple with small fragment
B3.2 Simple with large fragment
B3.3 Multiple fragmented
C1 – full joint fracture (simple articular trace and simple metaphyseal trace)
C1.1 Posteromedial joint fragment
C1.2 Sagittal trace
C1.3 Frontal trace
C2 – full joint fracture (simple joint and multiple fragmented metaphyseal)
C2.1 Sagittal trace
C2.2 Frontal trace
C2.3 extended to shaft
C3 – full joint fracture, multiple fragmented
C3.1 Metaphyseal, simple
C3.2 Metaphyseal, multiple fragmented
C3.3 Extended to shaft

Chart 2 - Universal classification for radius distal end fractures⁽²³⁾

- | |
|--|
| Type 1 – Extra-joint fracture, without deviation |
| Type 2 – Extra-joint fracture, with deviation |
| 2A Reducible and stable |
| 2B Reducible and unstable |
| 2C Irreducible |
| Type 3 – Intra-joint fracture, without deviation |
| Type 4 – Intra-joint fracture, with deviation |
| 4A Reducible and stable |
| 4B Reducible and unstable |
| 4C Irreducible |

RESULTS

The mean interobserver reliability of the Universal classification using plain X-ray images was 0.42. (Table 1) The result for the first level of the AO classification was a mean rate of 0.47, 0.32 for the second level, and 0.21 for the third level. (Table 1)

The mean interobserver reliability of the Universal classification using computed tomography was 0.37. For the AO classification: 0.34 on the first level, 0.21 on the second level, and 0.11 on the third level. (Table 2)

The differences found on interobserver reliability rates for the Universal classification based on plain X-ray images in comparison to those based on computed tomography images were shown to be statistically insignificant after the Wilcoxon's non-parametric paired test.

The differences found between interobserver reliability rates for AO classification based on X-ray images compared to the rates based on tomography images were statistically significant in all detail levels of the AO classification.

DISCUSSION

A good interobserver reliability is of great importance for any classification system. An appropriate classification should consider and evidence the aspects defining the severity of the injury, serving as a basis to decide on the kind of treatment, evaluate its result and predict its prognosis.

The number of patients presented in this study is consistent to the average samples of studies assessing interobserver reliability for radius distal fractures classification, and has shown to be appropriate to provide results with statistical significance.

The selection of observers among 3rd-year resident doctors intended to assess the reliability among observers with a general orthopaedic education with no specific specialization in Hand Surgery.

The studies addressing interobserver reliability for the evaluation of radius distal fractures have several methodologies, making comparisons difficult. Some studies use plain X-ray images only¹³⁻¹⁸, some associate these with the use of computed tomography.⁵⁻¹² Some studies assess fractures by means of anatomical parameters such as ulnar bending, volar bending and shortening.⁸⁻¹⁹ Other studies use measurements such as steps and diastasis^{5,11} while some other still assess these fractures using some well-known classification method: Frykman, AO, Universal, Melone, Mayo or Older's.^{6,8,13-18}

Table 1 – Mean interobserver consistency for Universal and AO classifications using plain X-ray images

			Kappa index by classification			
			Universal	AO – 1 st level	AO – 2 nd level	AO – 3 rd level
Observer 1	X	Observer 2	0.34	0.48	0.55	0.44
Observer 1	X	Observer 3	0.60	0.29	0.23	0.05
Observer 4	X	Observer 1	0.38	0.28	0.24	0.19
Observer 1	X	Observer 5	0.26	0.51	0.33	0.25
Observer 2	X	Observer 3	0.40	0.42	0.29	0.09
Observer 2	X	Observer 4	0.33	0.57	0.40	0.34
Observer 2	X	Observer 5	0.59	0.58	0.33	0.27
Observer 2	X	Observer 4	0.53	0.71	0.40	0.16
Observer 3	X	Observer 5	0.57	0.42	0.14	0.06
Observer 5	X	Observer 4	0.22	0.43	0.32	0.28
		Mean	0.42	0.47	0.32	0.21

Table 2 – Mean interobserver consistency for Universal and AO classifications using computed tomography

			Kappa index by classification			
			Universal	AO -1 st level	AO – 2 nd level	AO – 3 rd level
Observer 1	X	Observer 2	0.27	0.33	0.27	0.20
Observer 1	X	Observer 3	0.19	0.39	0.26	-0.07
Observer 4	X	Observer 1	0.53	0.34	0.02	0.00
Observer 1	X	Observer 5	0.52	0.15	0.12	0.08
Observer 2	X	Observer 3	0.22	0.27	0.39	0.27
Observer 2	X	Observer 4	0.24	0.25	0.20	0.10
Observer 2	X	Observer 5	0.26	0.29	0.20	0.10
Observer 3	X	Observer 4	0.46	0.38	0.27	0.26
Observer 3	X	Observer 5	0.29	0.50	0.26	0.12
Observer 5	X	Observer 4	0.70	0.51	0.10	0.07
		Mean	0.37	0.34	0.21	0.11

Studies using both plain X-ray and computed tomography images usually compare the use of plain X-ray imaging *versus* plain X-ray imaging associated to tomography.^{6,7,9} Johnston et al.⁸ select an observer with access only to tomography images, and another observer with access to both plain X-ray and tomography images. However, they do not perform a comparative analysis between these observers or calculate consistency between them. Cole et al.⁵ make a separate analysis of X-ray and tomography images considering step and joint diastasis, finding a significant difference in terms of reliability of the X-ray images (Kappa: 0.31 – 0.47) and tomography images (0.69 – 0.83). Pruitt et al.¹¹ also presented an evaluation of tomography images independently from X-ray images, but they do not calculate the Kappa index for data analysis. In this study, the evaluation of images by observers is made independently for X-ray and tomography, showing a different approach of tomography images as compared to other studies estimating interobserver reliability for AO and Universal classifications.

The Universal classification for radius distal fractures is divided into four types. Types II and IV can be subdivided into subtypes A, B and C, which refer, respectively, to stable reducible, unstable reducible and irreducible fractures. However, these sub items have not been used in this study because a standardization of images would be necessary, consisting of a baseline examination and an assessment after bloodless reduction, which has not been performed.

X-RAY IMAGES

In this study, the results for the Universal classification (0.47 – moderate reproducibility) had a similar magnitude to those reported by Oliveira-Filho et al.¹⁶ (0.33 – low reproducibility).

For the first level of the AO classification, this study found a mean reliability rate of 0.47 (moderate). Andersen et al.¹³ reported, in their study, a rate of 0.64, Kreder et al.¹⁵ 0.68, while Oskam et al.¹⁸ (who reduced the AO classification to 4 types), 0.65. These values

correspond to a good reproducibility. Flinkkilä et al.⁶, with 5 types, report a rate of 0.23, a low reproducibility level, and, with 2 types, 0,48, which represents a moderate reproducibility.

For the second level (nine subtypes), we found a mean rate of 0.32 (low reliability). Kreder et al.¹⁵ reported, in their study, a Kappa value of 0.48 (moderate).

For the AO classification in its third level (integral, with 27 subtypes), the present study reported a mean Kappa of 0.21 (low reliability). Oliveira Filho et al.¹⁶ reported a Kappa index of 0.21, Illarramendi et al.¹⁴ reported indexes between 0.31 and 0.40, Andersen et al.¹³ 0.25 and Kreder et al.¹⁵ 0.33, all of these in the "low" reliability range. Flinkkilä et al.⁶ present a Kappa index in the "poor" reliability range: 0.18.

As expected, the mean Kappa index has progressively reduced as the detailing of AO classification increased.

TOMOGRAPHY IMAGES

As opposite to our expectations, the reproducibility of Universal and AO classifications using tomography images was systematically lower than the reproducibility calculated from X-ray images.

Concerning tomography images, we found a mean rate of 0.37 (low reliability) for the Universal classification. Despite of the lower reliability rate with this classification using tomography images when compared to X-ray images (0.47), no statistical difference was found after an analysis using the Wilcoxon's non-parametric paired test. We couldn't find another study for comparison.

The first level of the AO classification (three types) showed a mean Kappa index of 0.34 (low reproducibility). Flinkkilä et al.⁶ reduced the AO classification to five types, resulting in a rate of 0.25 (low reproducibility). By reducing the classification to two types, they found a rate of 0.78 (good reproducibility). We found a statistically significant difference ($p < 0.05$ by the Wilcoxon's test) between rates calculated based on tomography images compared to those calculated based on X-ray images (mean index of 0.47). This reflects a better classification reproducibility based on X-ray images compared to the one using tomography images alone as a source.

Concerning the second level of the AO classification (with nine types), the rate for this study was 0.21 (low reproducibility). There is no other study for comparison purposes. Similarly, a significant difference was found in favor of reliability using X-ray images (mean Kappa of 0.32).

The third level of AO classification (integral, with 27 subtypes): mean index of 0.11 (poor reproducibility). No other studies are available for comparison purposes. The difference found when compared to the X-ray images (0.21) is statistically significant.

A potential reason for this difference between reliability rates for the AO classification from X-ray images compared to tomography images would be the difficulty of an observer to precisely determine a three-dimensional morphology of the fracture trace from tomography images without previous aid provided by X-ray images. Flinkkilä et al.⁶ had previously commented this fact in their study, adding that a higher degree of difficulty is seen in determining the degree of metaphyseal comminution from tomography images compared to plain X-ray images. The importance of the parameters "trace morphology" and "comminution degree" seems to be critical in the AO classification for a correct determination of types.

On the other hand, the Universal classification only requires that the fracture is classified as intra-joint or not, and if a deviation is present or not. This could explain the fact that the difference on reliability rates between X-ray and tomography images showed no statistical significance for the Universal classification.

This concept is also consistent with the results reported by Cole et al.⁵ In their study, the authors compare the measurement of fracture steps and diastasis from X-ray images compared to those captured with tomography. They found a better reliability with tomography images. Similarly, in this study, there is no need to determine the morphology of a fracture trace or comminution degree, favoring the use of tomography over X-ray imaging.

When we assessed the images once the calculations were made, we could retrieve some examples of cases in which a significant inconsistency existed between observers. Figure 1 shows an example of a case in which an inconsistency was found concerning the intra- or extra-joint nature of the fracture trace on plain X-ray images. All observers identified an intra-joint fracture when assessing that patient's tomography image.

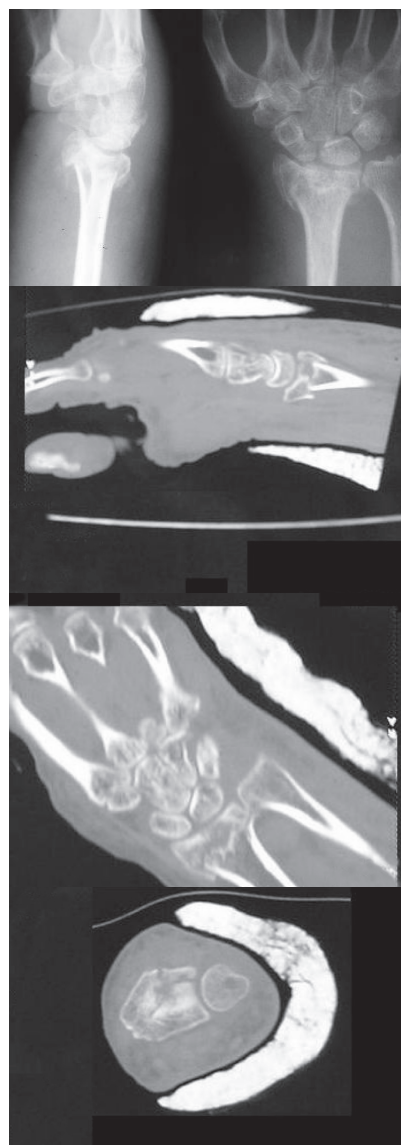


Figure 1 – X-ray and computed tomography images of a wrist at anteroposterior and lateral planes

Figure 2 shows an example of inconsistency between observers when assessing a computed tomography image.

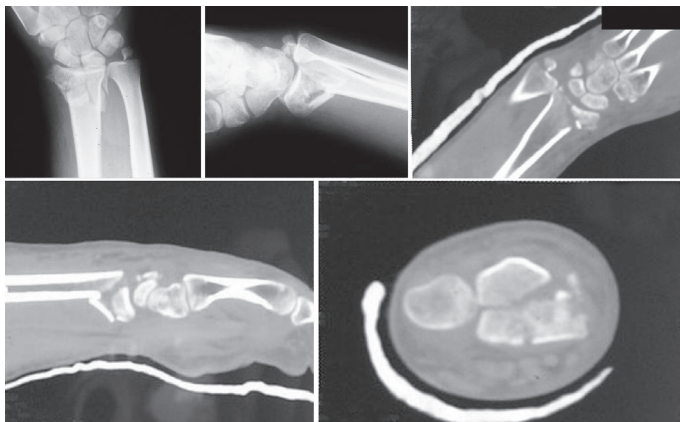


Figure 2 – X-ray and computed tomography images at lateral plane

The observers came to no agreement regarding the determination of fracture as partially articular or fully articular. It is interesting to observe that, in this case, the observers were unanimous to judge the fracture as fully articular when assessing it with X-ray images.

CONCLUSIONS

This study corroborates the low reliability rate reported on other studies addressing AO and Universal classifications for radius distal fractures using plain X-ray images. The use of computed tomography alone is associated to a low reliability rate, which is statistically worse than the reliability of classifications based on plain X-ray images. Therefore, we do not recommend the use of computed tomography alone to classify radius distal fractures. The Universal and AO classifications were historically designed for X-ray images of radius distal fractures, and not for the use of tomography. In the future, a specific classification for tomography images would be able to present improved results.

REFERENCES

- Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. *Injury*. 2006;37:691-7.
- Brogen E, Petranek M, Atroshi I. Incidence and characteristics of distal radius fractures in a southern Swedish region. *BMC Musculoskelet Disord*. 2007;8:48.
- Bucholz RW, Heckman JD, Court-Brown CM. Fractures of the distal radius and ulna. In: Bucholz RW, Heckman JD, Court-Brown C, Koval KJ, Tornetta Paul III, Wirth MA. Editors. *Rockwood & Green's fractures in adults*. 6th ed. Philadelphia: Lippincott Williams & Wilkins. p.910-64.
- Knirk JL, Jupiter JB. Intra-articular fractures of the distal end of the radius in young adults. *J Bone Joint Surg Am*. 1986;68:647-59.
- Cole RJ, Bindra RR, Evanoff BA, Gilula LA, Yamaguchi K, Gelberman RH. Radiographic evaluation of osseous displacement following intra-articular fractures of the distal radius: reliability of plain radiography versus computed tomography. *J Hand Surg Am*. 1997;22:792-800.
- Flikkilä T, Nikkola-Sihto A, Kaarela O, Pääkkö E, Raatikainen T. Poor interobserver reliability of AO classification of fractures of the distal radius. *J Bone Joint Surg Br*. 1998; 80:670-2.
- Harness NG, Ring D, Zurakowski D, Harris GJ, Jupiter JB. The influence of three-dimensional computed tomography reconstructions on the characterization and treatment of distal radius fractures. *J Bone Joint Surg Am*. 2006;88:1315-23.
- Johnston GHF, Friedman L, Krieger JC. Computerized tomographic evaluation of acute distal radial fractures. *J Hand Surg Am*. 1992;17:738-44.
- Katz MA, Beredjikian PK, Bozentka DJ, Steinberg DR. Computed tomography scanning of intra-articular distal radius fractures: does it influence treatment? *J Hand Surg Am*. 2001;26:415-21.
- Mino DE, Palmer AK, Levinsohn EM. Radiography and computerized tomography in the diagnosis of incongruity of the distal radio-ulnar joint. *J Bone Joint Surg Am*. 1985;67:247-52.
- Pruitt DL, Gilula LA, Manske PR, Vannier MW. Computed tomography scanning with image reconstruction in evaluation of distal radius fractures. *J Hand Surg Am*. 1994;19:720-7.
- Rozental TD, Bozentka DJ, Katz MA, Steinberg DR, Beredjikian PK. Evaluation of the sigmoid notch with computed tomography following intra-articular distal radius fracture. *J Hand Surg Am*. 2001;26:244-51.
- Andersen DJ, Blair WF, Steyers CM, Adams BD, El-Khoury GY, Brandser EA. Classification of distal radius fractures: an analysis of interobserver reliability and intraobserver reproducibility. *J Hand Surg Am*. 1996;21:574-82.
- Illarramendi A, González Della Valle A, Segal E, De Carli P, Maignon G, Galluci G. Evaluation of simplified Frykman and AO classifications of fractures of the distal radius. Assessment of interobserver and intraobserver agreement. *Int Orthop*. 1998;22:111-5.
- Kreder HJ, Hanel DP, McKee M, Jupiter J, McGillivray G, Swiontkowski MF. Consistency of AO fracture classification for the distal radius. *J Bone Joint Surg Br*. 1996;78:726-31.
- Oliveira Filho OM, Belangero WD, Teles JBM. Fraturas do rádio distal: avaliação das classificações. *Rev Assoc Med Bras*. 2004;50:55-61.
- Andersen GR, Rasmussen J-B, Dhal B, Solgaard S. Older's classification of Colle's fractures: good intraobserver and interobserver reproducibility in 185 cases. *Acta Orthop Scand*. 1991;62:463-4.
- Oskam J, Kingma J, Klasen HJ. Interrater reliability for the basic categories of the AO/ASIF's system as a frame of reference for classifying distal radial fractures. *Percept Mot Skills*. 2001;92:589-94.
- Bozentka DJ, Beredjikian PK, Westawski D, Steinberg DR. Digital radiographs in the assessment of distal radius fracture parameters. *Clin Orthop Relat Res*. 2002;(397):409-13.
- Velan O, De Carli P, Carreras C. Tomografía computarizada de alta resolución: su valor en las fracturas radiocubitales distales. *Rev Asoc Argent Ortop Traumatol*. 1999;64:186-91.
- Rüedi TP, Murphy WM, Vissoky J. Princípios AO do tratamento de fraturas. Porto Alegre: Artmed; 2002. p.45-7.
- Cooney WP, Agee JM, Hastings H, Melone CP, Rayback JM. Symposium: management of intrarticular fractures of distal radius. *Contemp Orthop*. 1990;21:71-104.
- Landis JR, Koch GK. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-74.