



# Reliability and reproducibility of classifications for distal radius fractures

Levent KÜÇÜK<sup>1</sup>, Mert KUMBARACI<sup>2</sup>, Hüseyin GÜNAY<sup>1</sup>, Levent KARAPINAR<sup>2</sup>, Oğuz ÖZDEMİR<sup>1</sup>

<sup>1</sup>Department of Orthopedics and Traumatology, Faculty of Medicine, Ege University, İzmir, Turkey;

<sup>2</sup>Department of Orthopedics and Traumatology, Tepecik Training and Research Hospital, İzmir, Turkey

**Objective:** The aim of this study was to evaluate the interobserver reliability and intraobserver reproducibility of the Universal, AO, Fernandez and Frykman classifications for distal radius fractures.

**Methods:** Fifty standard sets of posteroanterior and lateral roentgenograms of displaced distal radius fractures were classified two times by two groups of evaluators at 2-month intervals. The first group consisted of 10 orthopedic surgeons with a minimum of 5 years of experience. The second group consisted of 10 orthopedic residents in their first two years of practice. Interobserver reliability and intraobserver reproducibility were assessed using Cohen's kappa test.

**Results:** None of the classifications achieved good – very good reliability levels. The Fernandez classification had a moderate and the others had a fair interobserver agreement kappa coefficient. All classifications had fair kappa intraobserver agreement although the Frykman and Fernandez classifications had better results.

**Conclusion:** None of the classification systems were superior in terms of reliability and reproducibility. The reliability and reproducibility rates of all four classifications were insufficient.

**Key words:** Agreement; classification; Fernandez; Frykman; radius fractures; Universal.

Distal radius fractures are among the most common fracture types. Fractures are classified depending on the trauma mechanism, magnitude of the exposed energy, local anatomical characteristics and bone quality.<sup>[1,2]</sup> More than 20 classification systems have been defined for distal radius fractures.<sup>[3-8]</sup> In practice, fractures are referred to by the name of the writer who first defined that fracture type. Commonly used classification systems are based on the anatomical characteristics of the fracture or trauma mechanism.

Classification systems should act as a guide for treatment planning and assist physician communication and

prognosis prediction.<sup>[9]</sup> “Reliability”, the harmony among different evaluators, and “reproducibility”, harmony within the same evaluator at different periods, are the two most important features of classification systems.<sup>[10]</sup>

The Frykman classification was published in 1967.<sup>[11]</sup> It is based on the existence of a fractured ulnar styloid and extension of the fracture to radiocarpal and radioulnar joints and is commonly used all over the world. The shortcomings of this classification system are its inability to detect the amount of displacement, fragmentation and resulting shortness.<sup>[3]</sup>

**Correspondence:** Levent Küçük, MD. Ege Üniversitesi Tıp Fakültesi, Ortopedi ve Travmatoloji Anabilim Dalı, Bornova, İzmir, Turkey.

Tel: +90 232 390 27 83 e-mail: kucuklevent@yahoo.com

**Submitted:** September 6, 2012 **Accepted:** February 25, 2013

©2013 Turkish Association of Orthopaedics and Traumatology

Available online at  
www.aott.org.tr  
doi:10.3944/AOTT.2013.3038  
QR (Quick Response) Code:



The AO classification aims to provide clues about fracture severity and prognosis. Fracture localization and morphological characteristics must be known for classification. Fractures are divided into three main groups; intra-articular, partial intra-articular, and extra-articular. Detailed classification is formed by dividing each group into smaller main and subgroups.<sup>[12]</sup>

The Fernandez classification is based on trauma mechanism. Understanding the trauma mechanism helps to determine the simultaneous soft tissue damage. Furthermore, it classifies distal radioulnar joint injuries as a separate group, which is effective in prognosis.<sup>[13,14]</sup>

The Universal classification was defined in 1993 and has been a common classification system due to its simplicity. Determiners in this classification system are fracture extension to the radioulnar joint, displacement and stability.<sup>[15]</sup>

The objective of this study was to evaluate the reliability and reproducibility characteristics of the Frykman, AO, Fernandez and Universal classification systems for distal radius fractures.

## Materials and methods

Standard posteroanterior (PA) and lateral radiographs of 50 patients with distal radius fractures taken in the emergency service between 2006 and 2010 were evalu-

ated in our study. A non-homogenous study group including different fracture types was created by considering the extension of the fractures to radiocarpal and distal radioulnar joints, radial shortening, displacement, fragmentation in the dorsal cortex, and direction of angulations. After the X-rays had been digitalized, film sets enabling the radiographs to be seen clearly in each screen were created and numbered from 1 to 50. (Fig. 1)

Two groups of evaluators performed the assessments. The first group consisted of 10 orthopedic surgeons from 2 different training hospitals with a minimum of 5 years of experience. The second group consisted of 10 residents in orthopedic surgery within their first 2 years of residency from the same hospitals.

Radiographs were assessed using the Universal, AO, Fernandez and Frykman classification systems. An assessment report containing the schematic and written explanations of these classification systems were handed out to the evaluators. Assessment results were recorded on pre-prepared forms. Radiographs were evaluated twice by each evaluator with same classification systems, with a minimum of 2 months between the first and the second evaluations.

All data were digitalized and assessed using SPSS 18.0 for Windows (SPSS Inc., Chicago, IL, USA) software. Evaluators' compatibility with their group was analyzed. Compatibility between the first and second evaluation of each evaluator was analyzed for each clas-



**Fig. 1.** Sample X-ray image for evaluation. **(a)** Posteroanterior and **(b)** lateral radiographs are numbered from 1 to 50 after being combined in one screen.

sification system using Cohen's kappa test. Obtained kappa values were assessed in accordance with the Landis and Koch system. Kappa values under 0 was defined as 'poor agreement', 0 to 0.20 as 'slight', 0.21 to 0.40 as 'fair', 0.41 to 0.60 as 'moderate', 0.61 to 0.80 as 'good' and 0.81 to 1 as 'very good'.<sup>[10]</sup>

## Results

Reliability kappa values of each group are shown in Table 1. None of the classification systems achieved good or very good reliability values. The Fernandez classification had moderate and all other groups had fair values. The highest kappa value was 0.46 for the Fernandez system and the lowest was 0.24 for the Universal system. The mean kappa value of the resident group was 0.33 (fair) and 0.36 (fair) in the surgeon group.

Kappa values of the compatibility between the first and the second evaluations of the evaluators are shown in Table 2. In the surgeon group, reproducibility of all the classifications were fair and in the resident group the Universal and AO classification systems received fair and Frykman and Fernandez moderate values. The mean kappa value was 0.40 (fair) in the resident group and 0.55 (moderate) in the surgeon group.

## Discussion

A classification system for fractures should be simple, determine the characteristics of fractures, obtain data

about the accompanying soft tissue injuries, point out the way of treatment and predict prognosis.<sup>[16]</sup> In addition, a classification system should assist the communication between physicians. The most important factors positively affecting communication is high reliability and reproducibility. Recent publications have mostly reported low reliability and reproducibility.<sup>[5,17-20]</sup>

Illarramendi et al.<sup>[21]</sup> assessed the Frykman and AO classification systems and found the Frykman system to be moderate in terms of reliability and good in terms of reproducibility. AO classification results were worse than the Frykman classification. As a result, they reported that both classification systems were not sufficient in terms of reliability and reproducibility for clinical application.

Kreder et al.<sup>[17]</sup> achieved high compatibility rates in a simplified AO classification application. However, significantly lower reproducibility rates were found in a more complicated AO classification with subgroups.

In another study<sup>[18]</sup> evaluating the Frykman, AO, Fernandez and Universal classification systems, reliability values of all the evaluated classification systems were between fair and moderate. The Fernandez classification had the highest reliability value and the Frykman classification the lowest. In terms of reproducibility, all classifications except the AO were found to be satisfying.

Ploegmakers et al.<sup>[22]</sup> evaluated the AO, Frykman, Fernandez and Older classification systems for repro-

**Table 1.** Reliability kappa values for for all classification systems in both resident and surgeon groups after 1st and 2nd evaluations.

|           | Resident      |               | Surgeon       |               |
|-----------|---------------|---------------|---------------|---------------|
|           | 1. Evaluation | 2. Evaluation | 1. Evaluation | 2. Evaluation |
| Universal | 0.24          | 0.33          | 0.24          | 0.32          |
| AO        | 0.29          | 0.31          | 0.30          | 0.34          |
| Fernandez | 0.39          | 0.39          | 0.46          | 0.45          |
| Frykman   | 0.38          | 0.33          | 0.38          | 0.40          |

**Table 2.** The highest, the lowest and the mean reproducibility kappa values for resident and surgeon groups in all classification systems.

|           | Resident |         |      | Surgeon |         |      |
|-----------|----------|---------|------|---------|---------|------|
|           | Lowest   | Highest | Mean | Lowest  | Highest | Mean |
| Universal | 0.12     | 0.63    | 0.30 | 0.31    | 0.91    | 0.49 |
| AO        | 0.15     | 0.66    | 0.37 | 0.37    | 0.74    | 0.50 |
| Frykman   | 0.36     | 0.69    | 0.51 | 0.37    | 0.85    | 0.60 |
| Fernandez | 0.20     | 0.52    | 0.41 | 0.25    | 0.94    | 0.60 |

ducibility. AO classification reproducibility was moderate and the other studies fair. As a result, they reported all the evaluated classification systems to be insufficient for clinical application.

According to our results, the Fernandez classification had the highest reliability rate and the Universal classification the lowest. None of the classifications evaluated could meet the expected reliability level. The Frykman and Fernandez classifications were better in terms of reproducibility. However, none of them had good reproducibility rates.

Reliability rates in both groups in our study were fair. Reproducibility was fair in the resident group and moderate in the surgeon group. Both groups failed to meet the expected values. In this respect, our results are in accordance with previous studies.<sup>[17,21,23,24]</sup>

Advantages of our study were the comparison between 4 common classification systems and that this comparison was performed by 20 different evaluators. However, that the evaluators were picked from only 2 hospitals can be considered a weakness of this study.

In conclusion, the reliability and reproducibility rates of the AO, Frykman, Fernandez and Universal classification systems are insufficient and do not have advantages over each other. New classification systems are required to maintain a common language when defining distal radius fractures.

## Acknowledgement

We would like to thank Ass. Prof. M. N. Orman for his contribution in statistical analysis and evaluation of the data.

**Conflicts of Interest:** No conflicts declared.

## References

- Lill CA, Goldhahn J, Albrecht A, Eckstein F, Gatzka C, Schneider E. Impact of bone density on distal radius fracture patterns and comparison between five different fracture classifications. *J Orthop Trauma* 2003;17:271-8.
- Mackenny PJ, McQueen MM, Elton R. Prediction of instability in distal radius fractures. *J Bone Joint Surg Am* 2006;88:1944-51.
- Fernandez DL. Distal radius fracture: the rationale of a classification. *Chir Main* 2001;20:411-25.
- Oskam J, Kingma J, Klasen HJ. Clinicians' recognition of 10 different types of distal radial fractures. *Percept Mot Skills* 2000;91:917-24.
- Kettler M, Kuhn V, Schieker M, Melone CP. Do we need to include osteoporosis in today's classification of distal radius fractures? *J Orthop Trauma* 2008;22:S79-82.
- Kiliç A, Kabukçuoğlu Y, Ozkaya U, Gül M, Sökücü S, Özdoğan U. Volar locking plate fixation of unstable distal radius fractures. *Acta Orthop Traumatol Turc* 2009;43:303-8.
- Sügün TS, Gürbüz Y, Ozaksar K, Toros T, Kayalar M, Bal E. Results of volar locking plating for unstable distal radius fractures. *Acta Orthop Traumatol Turc* 2012;46:22-5.
- Gereli A, Nalbantoğlu U, Kocaoğlu B, Türkmen M. Comparison of palmar locking plate and K-wire augmented external fixation for intra-articular and comminuted distal radius fractures. *Acta Orthop Traumatol Turc* 2010;44:212-9.
- Audigé L, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma* 2005;19:401-6.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Frykman G. Fracture of the distal radius including sequelae – shoulder-hand-finger syndrome, disturbance in the distal radio-ulnar joint and impairment of nerve function. A clinical and experimental study. *Acta Orthop Scand* 1967;Suppl 108:3+.
- Müller ME, Nazarian S, Koch P, Schatzker J. The comprehensive classification of fractures of long bones. New York: Springer; 1990. p. 54-63.
- Fernandez DL, Jupiter JB. Fractures of the distal radius. Berlin, Heidelberg, New York: Springer; 1995.
- Fernandez DL. Treatment of articular fractures of the distal radius with external fixation and pinning. In: Saffar P, Cooney WP, editors. Fractures of the Distal Radius. London: Martin Dunitz Ltd; 1995. p. 104-17.
- Cooney WP. Fracture of the distal radius. A modern treatment-based classification. *Orthop Clin North Am* 1993;24:211-6.
- Kural C, Sungur I, Kaya I, Ugras A, Ertürk A, Cetinus E. Evaluation of the reliability of classification systems used for distal radius fractures. *Orthopedics* 2010;33:801.
- 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.
- 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.
- Naqvi SG, Reynolds T, Kitsis C. Interobserver reliability and intraobserver reproducibility of the Fernandez classification for distal radius fractures. *J Hand Surg Eur Vol* 2009;34:483-5.
- MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH, Hildebrand KA. Reliability of hand fellows' measurements and classifications from radiographs of distal radius fractures. *Can J Plast Surg* 2001;9:51-8.
- Illarramendi A, González Della Valle A, Segal E, De Carli P, Maignon G, Gallucci 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.
- Ploegmakers JJ, Mader K, Pennig D, Verheyen CC. Four distal radial fracture classification systems tested amongst a large panel of Dutch trauma surgeons. *Injury* 2007;38:1268-72.

23. Andersen DJ, Blair WF, Steyers CM Jr, 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.
24. Flikkilä T, Nikkola-Sihto A, Kaarela O, Pääkkö E, Raatikainen T. Poor interobserver reliability of AO classification of fractures of the distal radius. Additional computed tomography is of minor value. *J Bone Joint Surg Br* 1998;80: 670-2.