SPOTLIGHT ON upper extremity

Evaluation of the Reliability of Classification Systems Used for Distal Radius Fractures

CEMAL KURAL, MD; IBRAHIM SUNGUR, MD; IBRAHIM KAYA, MD; AKIN UGRAS, MD; AHMET ERTÜRK, MD; ERCAN CETINUS, PHD

abstract

Full article available online at OrthoSuperSite.com/view.aspx?rID=76419

The objective of this investigation was to evaluate the reliability of classification systems by determining inter- and intraobserver agreement in displaced distal radius fractures.

Radiographs of 32 patients (21 men and 11 women with a mean age of 41.6 years) who presented with a displaced distal radius fracture were classified by 9 orthopedic surgeons (5-25 years experience) using 5 different classification systems (Fernandez, AO, Frykman, Melone, and Universal Classification systems) twice with 20-day intervals. The results were processed with kappa statistics and used in assessment of inter- and intraobserver agreement of the classification systems. When classification systems were compared, the highest kappa coefficient in intraobserver agreement was determined in Universal classification (0.621). Fernandez (0.474), AO (0.309), Frykman (0.305), and Melone classification systems (0.262) followed the Universal system respectively. Kappa statistical results were evaluated using the Landis Koch score system for the assessment of interobserver agreement. According to the Landis Koch score system, the results were insufficient in all classification systems. Fernandez classification system had the highest interobserver agreement (0.235) and Melone classification system had the lowest interobserver agreement (0.056).

According to the results of our study, the systems used to classify the displaced distal radial fractures are insufficient. A new classification system that ensures the 3-dimensional assessment of the fracture is more user-friendly and a high inter- and intraobserver agreement is necessary.

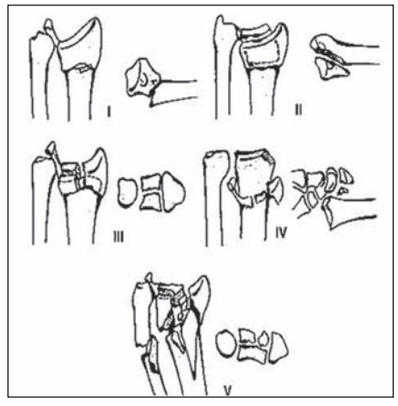


Figure: Fernandez Classification.

Drs Kural, Sungur, Kaya, Uğraş, Ertürk, and Cetinus are from Haseki Training Hospital Orthopedic and Traumatology Clinic.

Drs Kural, Sungur, Kaya, Uğraş, Ertürk, and Cetinus have no relevant financial relationships to disclose.

Correspondence should be addressed to: Cemal Kural, MD, Atakoy 9, kisim A2 Bloc D:92, 34156, Bakirkoy-Istanbul/Turkey.

doi: 10.3928/01477447-20100924-14

Distal radius fractures represent 17% of all fractures treated in emergency departments. These injuries are commonly referred to as fragility fractures and are more prevalent among osteoporotic elderly women.[1] Distal radius fractures are classified with an attempt to help define the trauma mechanism, guide the treatment, and predict prognosis.[2] More than 20 different wrist fracture classifications are in literature but none of these various classifications are reported to function as an excellent guideline for the prediction of prognosis and for the definition of the treatment method.[3,4] In fact, the classification system should be easily understandable for the physicians who use it for the first time. Secondly, it should provide prognostic significance and be a guide in the treatment[4,5] The fracture classification should be known to have interobserver reliability and intraobserver reproducibility regardless of the fracture type.[5,6]

In this study, 5 different popular distal radial fracture classification systems were selected (Fernandez, AO, Frykman, Melone and Universal) to investigate inter- and intraobserver agreement in these classification systems. These classification systems were evaluated in displaced distal radial fractures considering the anteroposterior and lateral radiographic views by 9 investigators. The Fernandez classification system was developed in 1993 by Fernandez and consists of 5 groups based on the mechanism of trauma (Figure 1). This system was designed to determine stability, include associated injuries, and assist treatment recommendations.[7]

AO classification was first used in 1986, and in 1990 it was revised with some additions (Figure 2). Depending on the relationship between the radius joint surface and the fracture, it is divided into 3 main groups; each group is divided into 3 subgroups within itself, and these are divided into 3 subgroups, so that it is composed of 27 subgroups.[8]

Frykman classification, described in 1967, is evaluated under 8 groups according to the involvement of the radiocarpal and distal radioulnar joints (Figure 3), and whether ulna styloid fractures accompany distal radius fractures. It is the most frequently used classification in daily practice.[9]

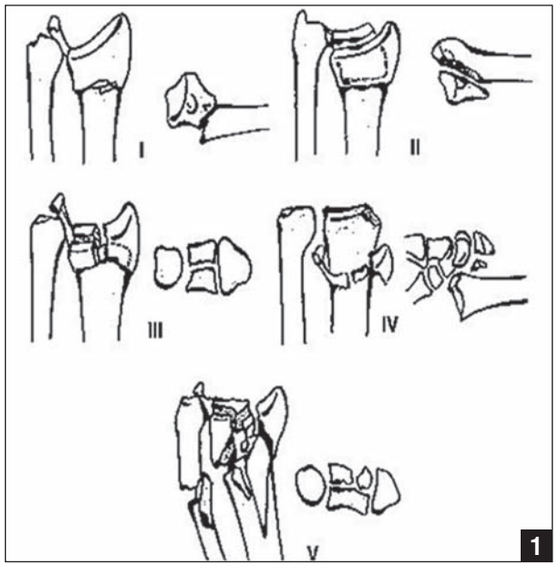


Figure 1: Fernandez Classification.

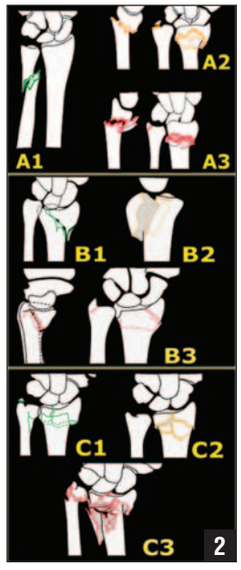


Figure 2: AO Classification.

Melone[10] classification, which was first described in 1984, focuses on 4 main parts in distal radius intraarticular fractures, and is classified under 5 groups depending on whether radius styloid, radius body, dorsal part, and palmar radial part are involved radiologically or not (Figure 4).

Universal classification system, which was described by Cooney[11] in 1993, evaluates distal radial fractures as intra or extraarticular, displacement present or absent, and whether they are stable (Figure 5).

Materials And Methods

Anteroposterior and lateral radiographs of 21 men and 11 women who presented to our emergency department due to displaced distal radius fracture were examined (nondisplaced fractures were excluded). Radiographs were taken from a distance of 105 cm using the same radiography device and the same bathing machine. Mean patient age was 41.6 years (minimum, 18 years, maximum, 74 years). The radiographs were evaluated to determine intraobserver reproducibility (in different time intervals by the same individual physician) and interobserver reliability (among the physicians who were involved in the study). Nine orthopedists and traumatology specialists (minimum, 5 years of experience; maximum, 25 years of experience) participated in this study and the radiographs were evaluated according to 5 different distal radius classification systems by each specialist independently. At the first evaluation, all radiographs were assessed in numerical sequence. Twenty days later, the initial radiograph order was randomly changed to generate a new sequence and all radiographs were reevaluated by the same physicians. During the evaluation of the radiographs, investigators were referred figures of the classification systems used in this study.

Kappa coefficient was used as the statistical method. The inter- and intraobserver compliance was revealed by kappa coefficients. The assessment of the measured kappa coefficients were evaluated according to Landis and Koch classification. According to this classification, kappa values below 0 are considered poor agreement, 0 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 good, and 0.81

to 1.00 near perfect.[11] After all data were recorded to the tables, they were evaluated using SPSS 11.0 Windows statistics package (SPSS Inc, Chicago, Illinois).

RESULTS

The results of inter- and intraobserver agreement are summarized in Tables 1 and 2. For the Universal classification, mean highest intraobserver kappa coefficient was determined (0.621). The mean intraobserver kappa values of the other 4 classifications were determined respectively as Fernandez (0.474), Frykman (0.310), AO (0.309), and Melone (0.285).

When interobserver evaluation of each classification was made after days 1 and 20, mean kappa values for all classification systems were found to be insufficient according to Landis-Koch classification[12] (Table 2). However, Fernandez classification seemed to have a slightly higher mean kappa value (0.2351) with respect to other classifications. First and second mean kappa values of the other 4 classifications were determined to be close and noncompatible.

DISCUSSION

A fracture classification system should be simple, easy to remember, define the trauma mechanism, guide the treatment, and predict prognosis. In addition, it should have acceptable intra- and interobserver agreement. Distal radius fractures like other fractures have been classified by various methods. More than 20 different wrist fracture classifications are described in the literature. Previous studies were conducted to find a classification system that both responded to these goals and had the best inter- and intraobserver agreement.

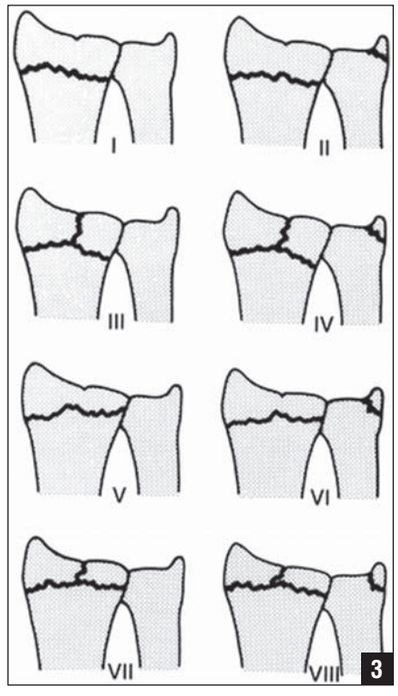
Ploegmakers et al[3] demonstrated moderate intraobserver agreement for the AO/ASIF classification and fair for the Frykman, Fernandez, and Older classifications. They found no correlation and consistency between clinical experience of the investigators and all 4 classifications. Because of these reasons, they did not recommend the use of these classifications for clinical application.[2]

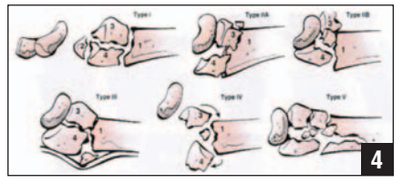
Andersen et al[13] found good intra- and interobserver agreement for the Older classification system in 185 fractures evaluated by 4 hand surgeons and did not observe a difference between the experienced and inexperienced investigators.

Illarramendi et al[4] investigated inter- and intraobserver agreement of the Frykman and AO classification systems in 200 fractures with 6 observers. They found that the Frykman classification showed moderate interobserver reproducibility and good intraobserver reproducibility. The experience of the investigators did not significantly affect either of these. The AO system showed fair interobserver reproducibility and moderate intraobserver reproducibility.[4] Kreder et al[14] evaluated the consistency of the AO classification in 36 physician and nonphysician observers in a series of 30 fractures, and reported diminished interobserver agreement in comminuted fractures and in nonphysicians.

In another study, the Frykman, Melone, Mayo, and AO classification systems in 55 sets of distal radius fractures were evaluated for interobserver reliability and intraobserver reproducibility by 2 orthopedic hand surgeons and 2 radiologists. Interobserver agreement was found as moderate for the Mayo classification and fair for the Frykman, Melone, and AO classifications. Intraobserver agreement was good for only 1 of 4 observers for each of the Frykman, Melone, and Mayo, while the remaining 3 observers achieved only fair to moderate reproducibility. Intraobserver agreement for the AO classification was fair for all 4 observers. No difference was found in inter- or intraobserver agreements between orthopedic hand surgeons and radiologists who were involved in this study.[14]

Belloti et al[15] evaluated the intra- and interobserver agreement of the Universal, AO, Frykman, and Fernandez classification systems and they observed best interobserver reproducibility rate in the Fernandez classification system (0.43) and the worst in the Frykman classification (0.26). They reported that intraobserver





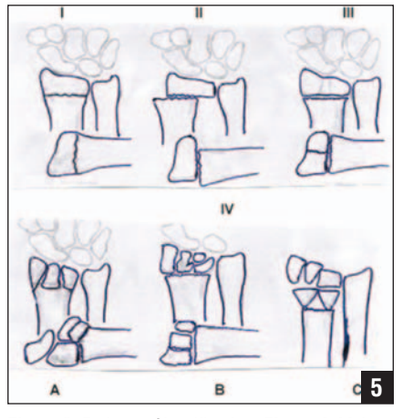


Figure 3: Frykman Classification. **Figure 4**: Melone Classification. **Figure 5**: Universal Classification.

Table 1

Intraobserver Kappa Values after First and Second Evaluations of Observers for Each Classification

	AO	Fernandez	Frykman	Melone	Universal
1. Observer	0.269	0.295	0.196	0.050	0.526
2. Observer	0.477	0.558	0.519	0.556	0.832
3. Observer	0.265	0.775	0.529	0.433	0.441
4. Observer	0.198	0.450	0.355	0.166	0.522
5. Observer	0.302	0.478	0.325	0.080	0.859
6. Observer	0.177	0.547	0.220	0.096	0.679
7. Observer	0.332	0.553	0.459	0.712	0.695
8. Observer	0.468	0.289	0.143	0.180	0.573
9. Observer	0.294	0.324	0.047	0.300	0.463
Mean Values	0.309	0.474	0.310	0.285	0.621

Abbreviation: AO, arbeitgemeinschaft fur osteosynthesfragen.

Table 2

Interobserver Kappa Values of First And Second Readings for Each Classification

	AO	Fernandez	Frykman	Melone	Universal
First Reading	0.0829	0.3022	0.2333	0.0436	0.1848
Second Reading	0.1093	0.1680	0.2058	0.0703	0.2675
Mean Values	0.0961	0.2351	0.2195	0.0569	0.2261

Abbreviation: AO, arbeitgemeinschaft fur osteosynthesfragen.

reproducibility was satisfactory in the Universal (0.61), Fernandez (0.59), and Frykman (0.55) classifications, and unsatisfactory in the AO classification (0.49).[16] Naqvi et al[17] evaluated the Fernandez classification for assessing the interobserver reliability and intraobserver reproducibility and demonstrated poor interobserver reliability and intraobserver reproducibility. They suggested that attention should be spent when using this classification for clinical practice.[15] In the study of McDermid et al,[18] radiographs of 128 distal radius fractures were evaluated by 2 hand surgery fellows, and interrater reliability for Mayo, AO, McMurtry, Universal, and Frykman classification systems were found fair. Only the Older classification system was found good.[17]

In our study, radiographs of 32 patients with displaced distal radius fractures were evaluated by 9 orthopedic surgeons according to 5 different classification systems commonly used at present. Intraobserver evaluation revealed kappa value of 0.621 (good) for Universal classification, 0.474 (moderate) for Fernandez classification, and 0.309, 0.305, and 0.262 (fair) for AO, Frykman, and Melone classifications respectively. When interobserver evaluation of each classification was made, mean kappa values for all classification systems were found fair.

When compared with the literature, our results seem to correlate with the study of Belloti et al,[15] since both studies reveal good intraobserver agreement in Universal classification followed by the Fernandez classification. However, no study in the literature shows excellent results considering intraobserver reproducibility. No study with excellent results among intraobserver literature exists. In our study, interobserver reliability values for all classification systems were found fair and were consistent with the literature.[4,13-17] The only classification system showing good reliability in the literature was the Older system.[12,17] But this classification system was designed for use in extraarticular fractures and presently is not used for intraarticular fractures.

In several studies, no direct proportion has been found between the experience of the observers and the intra- and interobserver agreement ratio.[4,14] In the present study, the relationship between intra- and interobserver agreement and experience of the observers was not investigated.

In the light of these data, none of the classification systems examined in this study approached the excellent or almost perfect level that one would rely on, given their importance in making clinical decisions in daily practice. To date, no classification system evaluating the plain radiographs that can be used by all orthopedic surgeons with high intraobserver agreement and interobserver reproducibility exists. The use of these classifications for direct comparisons of different published series may be beyond the capacity of any of the fracture classifications studied. There exists a need for a new classification system that examines the fracture in 3-dimensional fashion, that is easy to remember and easy to use by all orthopedic surgeons and has a high intraobserver reliability and interobserver reproducibility.

REFERENCES

1. Ruch DS. Fractures of the distal radius and ulna. In: Bucholz RW, Heckman JD, Court-Brown CM, eds. *Rockwood and Green's Fractures in Adults*. 6th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2006:910-964.
2. Colles A. On the fracture of the carpal extremity of the radius. *Clin Orthop Relat Res*. 2006; (445):5-7.
3. 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(11):1268-1272.
4. Illarramendi A, Gonzalez 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(2):111-115.
5. Bernstein J. Fracture classification systems: do they work and are they useful? *J Bone Joint Surg Am*. 1994; 76(5):792-793.
6. Frandsen PA, Andersen E, Madsen F, Skjodt T. Garden's classification of femoral neck fractures: an assessment of inter-observer variation. *J Bone Joint Surg*. 1988; 70(4):588-590.
7. Fleiss JL, Slakter MJ, Fischman SL, Park MH, Chilton NW. Interexaminer reliability in caries trials. *J Dent Res*. 1979; 58(2):604-609.
8. Fernandez DL, Jupiter J. *Fractures of the Distal Radius: A Practical Approach to Management.* New York, NY: Springer Verlag; 1996.
9. Muller ME. The principle of the classification. In: Muller ME, Allgower M, Schneider R, Willenegger H, eds. *Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group*. New York, NY: Springer-Verlag; 1991:118.
10. Melone CP Jr. Articular fractures of the distal radius. *Orthop Clin North Am*. 1984; (15):217-236.
11. Cooney WP. Fracture of the distal radius. A modern treatment-based classification. *Orthop Clin North Am*. 1993; 24(2):211-216.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159-174.
13. Andersen GR, Rasmussen JB, Dahl B, Solgaard S. Older's classification of Colles' fractures. Good intraobserver and interobserver reproducibility in 185 cases. *Acta Orthop Scand*. 1991; 62(5):463-464.
14. 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(5):726-731.
15. Belloti JC, Tamaoki MJ, Franciozi CE, et al. Are distal radius fracture classifications reproducible? Intra and interobserver agreement. *Sao Paulo Med J*. 2008; 126(3):180-185.
16. Andersen DJ, Blair WF, Steyers Jr CM, Adams BD, el-Khouri GY, Brandser EA. Classification of distal radius fractures: an analysis of interobserver reliability and intraobserver reproducibility. *J Hand Surg Am*. 1996; 21(4):574-582.
17. 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(4): 483-485.
18. McDermid 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; 9:51-58.