Injury, Int. J. Care Injured 49S1 (2018) S29–S32

Contents lists available at ScienceDirect

Injury

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

Distal radius fractures are difficult to classify

Daniel Wæver[a,b], Mette Lund Madsen[a], Jan Hendrik Duedal Rölfing[b], Lars Carl Borris[b], Mads Henriksen[c], Lise Loft Nagel[c], Rikke Thorninger[a,*]

[a] Department of Orthopaedics, Regionshospitalet Randers, Randers, Denmark
[b] Department of Orthopaedics, Aarhus University Hospital, Aarhus, Denmark
[c] Department of Radiology, Aarhus University Hospital, Aarhus, Denmark

KEYWORDS

Distal radius fracture
Fracture classification system
AO/OTA
Frykman
Older
Interobserver reliability
Intraobserver reproducibility
Prognosis

ABSTRACT

Background: Traditionally, distal radius fractures (DRFs) have been described using eponyms, e.g. Colles, Smith, Barton, Chauffeur. During the last half of the 20th century several classification systems for DRF have emerged. We evaluated the inter- and intra-observer agreement of the AO/OTA, Frykman and Older classification systems.

Methods: Four observers, an intern, an orthopaedic registrar, an orthopaedic consultant and a radiology consultant, independently evaluated DRF radiograms and classified the fractures according to the AO/OTA, Frykman and Older classification systems. After an interval of 6 months, radiograms of 30 randomly chosen patients were re-evaluated by the same observers.

Results: Radiograms of 573 DRF patients were evaluated in the study. The inter-observer reliability of the AO/OTA fracture types (A, B and C) was 'weak' (kappa = 0.45). The agreement dropped to 'minimal' (kappa = 0.24) regarding the AO/OTA groups (A2, A3, B1, B2, B3, C1, C2 and C3). The reliability of the Frykman classification system was 'weak' (kappa = 0.41), and we observed the lowest inter-observer reliability for the Older classification system (kappa = 0.10). The kappa values for the intra-observer reproducibility of the AO/OTA fracture types (A, B and C) ranged from 0.58 to 0.87. For the AO/OTA groups (A2, A3, B1, B2, B3, C1, C2 and C3) the reproducibility was lower ranging from 'minimal' to 'weak'. The intra-observer reproducibility of the Frykman system was 'weak' to 'moderate' and even worse for the Older classification system.

Conclusion: Based on these findings the AO/OTA classification system seems to be most reliable for routine use, however, with lower kappa values concerning the agreement for the groups. The Frykman and Older classification systems cannot be recommended because of less convincing results.

Introduction

Distal radius fractures (DRFs) are among the most common fractures in the Western world with a reported incidence rate of 190–200 per 100,000 person-years [1,2]. The incidence peaks among adolescent males due to high-energy trauma and in elderly females due to osteoporosis-associated low-energy falls [3,4].

Traditionally, DRFs have been described using eponyms, e.g. Colles, Smith, Barton, Chauffeur. During the last half of the 20th century several classification systems for DRF have emerged [5]. Several studies have evaluated the inter-observer reliability and intra-observer reproducibility of some of these DRF classification systems [6] and have shown both a low reproducibility and a low reliability. Table 1 presents an overview of previous studies and their results. Most of these studies were relatively small in patient volume and used few observers. Thus, the purpose of this large-scale study was to estimate the inter-observer reliability and intra-observer reproducibility of three commonly used classification systems for DRF, namely the AO/OTA, the Frykman and the Older systems.

Materials and methods

Radiographic examinations of patients with DRF were retrospectively obtained from the electronic picture archiving and communication system at Aarhus University Hospital. Details regarding patient selection, inclusion and exclusion criteria and time to follow-up were previously described [7].

The standardized radiographic examination of the wrist consisted of anteroposterior and lateral projections.

Four observers independently evaluated the radiograms and classified the fractures according to the AO/OTA, Frykman and Older classification systems. The four observers were an intern, an orthopaedic registrar, an orthopaedic consultant and a radiology consultant. A visual illustration together with a written explanation were available for each classification system throughout the assessment of the electronic radiograms (Fig. 1). No consensus meeting was held throughout the study.

* Corresponding author at: Department of Orthopaedics, Regionshospitalet Randers, Skovlyvej 15, 8930 Randers, Denmark
E-mail address: rikkthor@rm.dk (R. Thorninger).

0020-1383/

Table 1
Kappa values of reference studies

	AO/OTA groups (A2, A3, B1, B2, B3, C1, C2, C3)		AO/OTA types (A, B, C)		Frykman (I–VIII)		Older (I–IV)	
Study	Inter	Intra	Inter	Intra	Inter	Intra	Inter	Intra
Andersen 1991 [15]							0.69	0.75
Belloti 2008 [2]	<0.31	<0.61			<0.26	<0.59		
Kural 2010 [10]	0.10	0.31			0.22	0.31		
Plant 2015 [11]	0.29	0.53	0.56	0.65				
Ploegmakers 2007 [12]		0.52				0.26		0.27
Siripakarn 2013 [18]	0.34	0.29			0.28	0.31		
Van Buijtenen 2015 [13]	0.48	0.70	0.49	0.70				

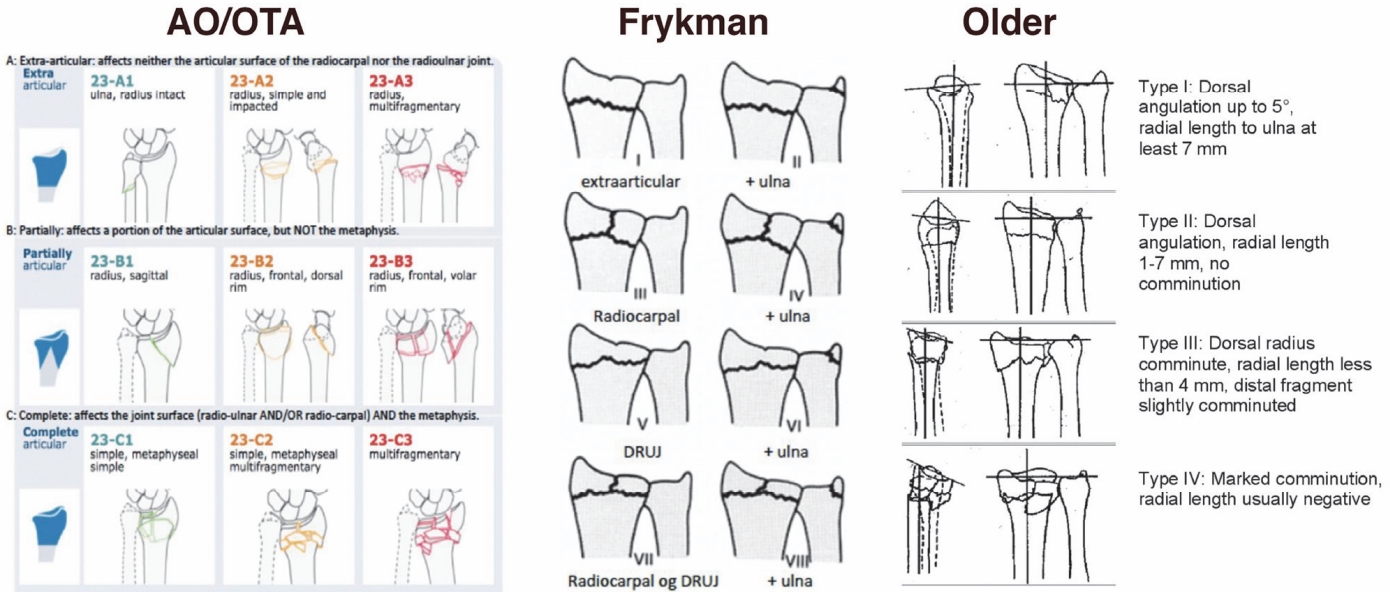

Fig. 1. Visual illustration and written explanation, which were available throughout the study. DRUJ = distal radioulnar joint.

The evaluations were collected on preformatted forms before transfer to our data analysis program.

After six months the electronic radiograms of 30 randomly selected patients were re-evaluated by the same observers in order to estimate the intra-observer reproducibility.

Statistical analysis

Cohen's kappa was calculated using STATA. In the present study the interpretation of the kappa values, e.g. the level of agreement, is presented as suggested by McHugh [8], which is stricter than the originally suggested interpretation by Landis and Koch [9] (Table 2). When referring to previous studies we cite their interpretation of kappa values.

Table 2
Interpretation of kappa values

	Level of agreement		% of data that are reliable
Kappa	Landis and Koch	McHugh	
0–0.20	Slight	None	0–4%
0.20–0.40	Fair	Minimal	4–15%
0.40–0.60	Moderate	Weak	15–35%
0.60–0.80	Substantial	Moderate	35–63%
0.80–0.90	Almost perfect	Strong	64–81%
0.90–1	Almost perfect	Almost perfect	82–100%

Results

Radiographs of 573 patients, who were operatively treated for DRF, were evaluated in the present study.

Inter-observer reliability (Table 3)

The inter-observer reliability of the AO/OTA fracture types (A, B and C) was 'weak' (kappa = 0.45), however, concerning the assessment of the AO/OTA groups (A2, A3, B1, B2, B3, C1, C2 and C3) the agreement dropped to 'minimal' (kappa = 0.24).

The reliability of the Frykman classification system was 'weak' (kappa = 0.41).

The lowest inter-observer reliability was observed for the Older classification system with a kappa value of 0.10.

Table 3
Kappa values of the inter-observer reliability of the AO/OTA, Frykman and Older classification systems, n=573

Classification	kappa
AO/OTA types (A, B and C)	0.45
AO/OTA groups (A2, A3, B1, B2, B3, C1, C2, C3)	0.24
Frykman groups (I–VIII)	0.41
Older types (I–IV)	0.10