

Interobserver and intraobserver variations in radial head fracture classification-assessment of two classification systems

Brijesh Ayyaswamy^{a,*}, Laura Howell^b, Anoop Anand^a, Charalambos P. Charalambous^a

^a Department of orthopaedics, blackpool teaching hospitals nhs trust, UK

^b Associate in medical statistics, university of central Lancashire, UK

ARTICLE INFO

Keywords:

Radial head
Classification
Variations
Reliability
Radiographs
Agreement

ABSTRACT

Objective: The aim of this study was to evaluate the intra- and interobserver variation of two classification systems for radial head fractures (Mason and Charalambous classifications) using plain radiographs.

Material and methods: Five observers classified 30 radial head fractures as per the two classifications using anteroposterior and lateral elbow radiographs. Assessments were done on two occasions, at least 6 weeks apart.

Results: The interobserver and intraobserver variation of the Mason classification showed fair (mean kappa = 0.33) and moderate agreement (mean kappa = 0.43) respectively. The interobserver and intraobserver variation of the Charalambous classification showed moderate agreement for both (mean kappa = 0.42 and 0.49 respectively). A greater proportion of radial head fractures could be classified using the Charalambous classification compared with the Mason classification ($P < 0.001$). With the Charalambous classification, the interobserver variation was better when assessing fracture morphology (4 morphology groups) versus fracture displacement (2 displaced/un-displaced groups) ($p = 0.010$).

Conclusions: The Mason and Charalambous classifications for radial head fractures confer similar reliability when using plain radiographs, but the latter allows a greater proportion of fractures to be classified. Raters may agree more on fracture morphology as compared to fracture displacement. Our findings also demonstrate the limitations of using plain radiographs in classifying radial head fractures for clinical or research purposes, and suggest that evaluation with more sensitive modalities such as Computed Tomography may be preferable.

1. Introduction

Radial head fractures account for approximately one third of all elbow fractures with an estimated annual incidence of 2.5–2.8 per 10000 individuals.⁷ The treatment varies from surgical to non-surgical which can involve fracture fixation, fracture fragment excision, radial head excision or replacement.

Classification systems aim to describe the fracture pattern and help guide communication for clinical or research purposes. Classification systems that indicate severity of injury or guide management are preferable in clinical practice. Several classification systems have been used to describe radial head fractures. The Mason classification and its subsequent modifications are commonly utilized. The original Mason classification assigned radial head fractures to three categories based on comminution and displacement (Table 1). Johnston added as a fourth category radial head fractures associated with elbow dislocation.⁶ Subsequently, Broberg, Morrey² and Hotchkiss⁵ quantified the amount of displacement for type II fractures. The original Mason classification

has been criticized in that it does not allow all fracture patterns to be classified (such as isolated radial neck fractures, marginal fractures involving parts of the radial head other than the lateral sector), and not specifying the extent of comminution or absence of displacement in type III fractures.⁸ A more recently described classification by Charalambous addresses these issues by taking into account fracture morphology, comminution and displacement (Fig. 1). Its grading progresses with fracture severity and can help guide management.

For any classification system to be reliable it must have good inter- and intraobserver reliability, hence the aim of this study was to evaluate these parameters for two currently utilized radial head fracture classification systems. Plain radiographs were examined as these are commonly used in guiding clinical decision making.

2. Material and methods

Thirty acute adult radial head fractures were identified from hospital records. They were chosen to allow a broad inclusion of the

* Corresponding author. Department of orthopaedics, Blackpool teaching hospitals nhs trust, Whinneyheys Road, Blackpool, Lancashire, FY3 8NR, UK.

E-mail address: brijeshayyaswami@gmail.com (B. Ayyaswamy).

<https://doi.org/10.1016/j.jor.2019.05.012>

Received 3 March 2019; Accepted 26 May 2019

Available online 28 May 2019

0972-978X/ © 2019 Professor P K Surendran Memorial Education Foundation. Published by Elsevier B.V. All rights reserved.

Table 1
Mason classification.

Type I: Fissure fracture or marginal sector fracture without displacement (subchondral fracture or fracture involving a sector of the lateral quadrant of the radial head.
Type II: Marginal sector fractures with displacement – a segment of the lateral border of the radial head is separated, impacted, depressed or tilted out of line.
Type III: Comminuted fracture involving the whole of the radial head.

common types of radial head fractures which may be encountered in clinical practice with regards to displacement and configuration. The radiographs had anteroposterior and lateral views. These views were not standardized in terms of exact elbow positioning, but were the ones that had been obtained on the acute presentation, and the ones utilized in guiding fracture management.

The radiographs were presented as a power-point presentation to five Orthopedic Surgeons practicing in the United Kingdom (3 senior staff grade orthopaedic surgeons and 2 Specialist trainees) on two occasions, at least six weeks apart. The order of presentation was altered on the two occasions. The observers were not provided with any feedback after the first viewing and the radiographs were not available to any of them between first and second viewings. Prior to their assessment, all observers were given a verbal and written description of the Mason and Charalambous classifications, and printed figures showing the two classifications. Evaluation was done with observers blinded to other examiners' ratings and patient identity. The study had approval by the Research and Development department of our institution.

2.1. Statistics

The interobserver variation of the first readings was initially assessed using Cohen's kappa.⁴ A kappa was calculated for each pair of raters followed by the mean of the kappa values for each rater and then the overall mean. Intraobserver variation was calculated by comparing the first and second readings for each rater. Kappa is a coefficient of agreement, that varies from +1 (perfect agreement) to 0 (agreement no better than chance), to -1 (representing absolute disagreement). A kappa takes into consideration the number of agreements expected to be seen by chance rather than being a mere reflection of the proportions of cases in which agreement was observed. There are no absolute definitions of interpreting kappa but there are suggested guidelines by Altman 1991¹ (Table 2). This analysis was performed for the Mason classification and Charalambous classification (8 groups of morphology). A sub-analysis into the ability of raters to classify the radial head fractures using the Charalambous classification into the 4 morphology groups and into displaced or un-displaced fractures was also performed.

A second analysis was performed to assess interobserver variation and this involved an analysis-of-variance approach.^{1,2} A kappa was calculated for each category followed by the combined kappa which was the weighted average of the individual kappa for each category. Results from this second analysis were compared with those from the first analysis, to determine whether they yielded similar results. Analysis included those radiographs where one or more raters could not classify them.

The first analysis, using Cohen's kappa, was then repeated but this time we only analyzed agreement between pairs of raters for those radiographs where both raters could classify them. The number of radiographs used for comparison varied for each pair of raters. The results were compared with those produced when we included radiographs which could not be classified for some pairs of raters.

Comparisons of the interobserver and intraobserver variation for the two classifications were made using the Independent Samples *t*-test. All analysis was performed using the statistical package Stata.¹¹

3. Results

The interobserver and intraobserver variation for the Mason classification method showed fair (mean kappa = 0.33) and moderate (mean kappa = 0.43) agreement respectively (Table 3). The interobserver and intraobserver variation for the Charalambous classification (8 groups of morphology) showed moderate agreement for both (mean kappa = 0.42 and mean kappa = 0.49 respectively (Table 4). However, there were no statistically significant differences when comparing the interobserver variation ($p = 0.078$) and intraobserver variation ($p = 0.664$) between the two classifications (Table 5).

A sub-analysis into the ability of raters to classify the radial head fractures using the Charalambous classification solely into the 4 morphology groups or solely into displaced or un-displaced did not show a substantial change in intra-observer variation compared to when utilizing the overall classification (Table 5). However, the inter-observer variation was better when assessing fracture morphology (4 morphology groups) versus fracture displacement (2 displaced/un-displaced groups) ($p = 0.010$). Similarly, the inter-observer variation was better when using just the 4 morphology groups versus the overall classification ($p = 0.001$).

The second analysis approach, involving the analysis of variance method produced a weighted kappa of 0.32 (fair agreement) for the Mason classification (Table 6) and a weighted kappa of 0.41 (moderate agreement) for the Charalambous classification (Table 7). These results replicated those produced by the first method using Cohen's kappa for assessing the interobserver variation of the two methods.

When we repeated the first analysis but excluded radiographs where one or both of the raters could not classify the radiograph, the interobserver and intraobserver variation for the Mason classification showed moderate agreement for both (mean kappa = 0.50 and mean kappa = 0.52 respectively). The interobserver and intraobserver variation for the Charalambous classification (8 groups of morphology) also showed moderate agreement for both (mean kappa = 0.48 and mean kappa = 0.53 respectively). When using this analysis the interobserver variation for the 4 morphology groups of the Charalambous classification improved to good agreement (mean kappa = 0.64) which was significantly higher than that for classifying fractures into displaced or un-displaced (mean kappa = 0.46) ($p = 0.021$) and was also higher than the agreement for the overall classification ($p = 0.005$).

There were only 9 out of 30 radiographs which could be classified by all five raters using the Mason classification, whereas 23 could be classified by all five raters using the Charalambous classification (Table 8) ($p < 0.001$).

Further examination into the assignment ability of the raters to distinguish which fractures were difficult to classify showed that when using the Mason classification the participants could not classify mostly radial neck fractures, followed by displaced marginal fractures. When using the Charalambous classification the most common type of fractures which could not be classified were un-displaced partial articular fractures.

There was less agreement about the rating of radial head fractures into category 3 when using the Mason classification (Table 6). Furthermore, there was less agreement about the rating of radial head fractures into 2u and 3u when using the Charalambous classification (Table 7).

4. Discussion

One of the aims of fracture classification system is to provide a reproducible means by which clinicians and researchers communicate. Classification systems graded by fracture severity and ones which can guide treatment are preferable in clinical practice. Our results have shown similar interobserver and intraobserver variation for two current radial head fracture classification systems, the Mason and Charalambous classifications. Our study also showed that clinicians could not classify many of the radiographs when following the original

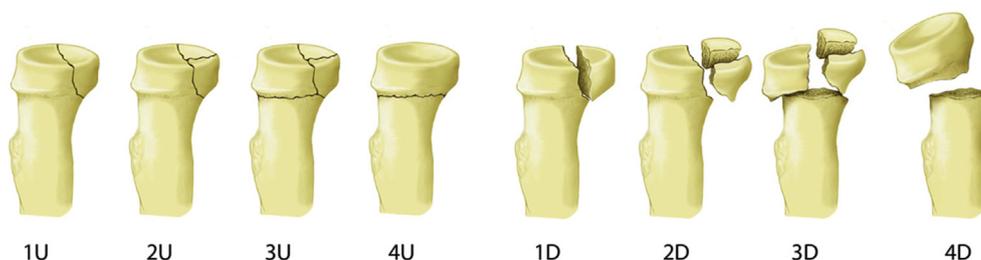


Fig. 1. Charalambous classification: Radial head fractures are classified into 2-part partial articular, comminuted partial articular, comminuted complete articular, and isolated radial neck fractures, all of which can be undisplaced (U) or displaced (D).

Table 2
Interpretation of kappa (Adapted from Altman 1991).

Value of Kappa	Strength of agreement
< 0.20	Poor
0.21–4.0	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81–1.00	Very Good

Table 3
Kappa values for inter- and intraobserver variation for the Mason classification (including ‘could not classify’ as a category and using first analysis approach - Cohen’s Kappa).

	Interobserver					Intraobserver
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	
Rater 1		0.15	0.30	0.40	0.42	0.19
Rater 2			0.17	0.41	0.37	0.41
Rater 3				0.22	0.47	0.18
Rater 4					0.42	0.85
Rater 5						0.51
Mean κ value for each observer	0.32	0.27	0.29	0.36	0.42	
Mean κ value for all observers	0.33					0.43

Table 4
Kappa values for inter- and intraobserver variation for the Charalambous classification (eight categories plus ‘could not classify’ as a category and using first analysis approach Cohen’s Kappa).

	Interobserver					Intraobserver
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	
Rater 1		0.50	0.47	0.31	0.43	0.53
Rater 2			0.38	0.40	0.42	0.44
Rater 3				0.27	0.47	0.32
Rater 4					0.50	0.61
Rater 5						0.54
Mean κ value for each observer	0.43	0.43	0.40	0.37	0.46	
Mean κ value for all observers	0.42					0.49

Mason classification.

The Mason classification was described in 1954 and subdivided radial head fractures into three types.⁸ It was based on fracture comminution and displacement. In the original classification it was not specified whether type II fractures can be comminuted, and there was no description as to whether type III fractures could be displaced or undisplaced. The classification also did not include isolated radial neck fractures. In addition Type I and II fractures referred to the lateral

sector of the radial head, rather than any part of the radial head. The original Mason classification was modified by Johnston in 1962 which included type IV as radial head fractures associated with elbow dislocation.⁶ Later modifications came from Van Riet and Morrey² and Hotchkiss⁵ who subsequently quantified the amount of displacement for type II fractures.

In our study, the Mason classification showed fair agreement in terms of interobserver variation and moderate agreement in terms of intraobserver variation. However, a large proportion of radial head fractures (21/30) could not be classified by all five raters using the Mason classification. The radiographs which could not be classified were Type II marginal fractures (likely because the original classification refers only to those involving the lateral sector of the radial head) and isolated radial neck fractures (as this type is not included in the original classification).

A more recent classification described by Charalambous was based on fracture morphology, comminution and displacement.³ It refers to four types based on fracture morphology, each of which can be undisplaced or displaced. With progression of the classification grading the fracture severity increases and the subdivisions can guide fracture management. Non-surgical treatment is preferable for undisplaced fractures with surgery considered in displaced fractures. Surgical fixation (in the form of screws or pins) may be preferable in two part partial articular fractures where fragment size allows (Type 1) whereas radial head replacement or resection may be preferable in comminuted partial articular (Type 2) and complete articular fractures (Type 3). Isolated radial neck fractures (Type 4) may be treated with surgical fixation, radial head excision or replacement; surgical fixation however is likely to need plate fixation rather than screw or pin fixation. The Charalambous classification does not quantify the amount of displacement to categorize the fracture as displaced or undisplaced, instead it takes any amount of displacement into consideration. This is in line with clinical practice where there is no widely agreed amount of displacement that can guide treatment. A previous study has shown that the Charalambous classification correlates with the severity of associated collateral ligament injuries and possible with the mechanism of elbow injury. Rhyou et al¹⁰ demonstrated that displaced Type 1 and 2 fractures were associated with medial collateral ligament injury suggestive of a valgus load injury mechanism, whereas displaced Type 3 and 4 fracture patterns were associated with lateral ulnar collateral ligament injury, consistent with a posterolateral rotation plus axial loading injury mechanism.

Our study shows the interobserver reliability of the Charalambous classification was slightly better than with Mason classification in spite that the raters were using this method for the first time, although the difference observed was not statistically significant.

Despite the wide use of the Mason classification in describing radial head fractures there is a sparsity of studies evaluating its inter- and intraobserver reliability. We could only identify one study that evaluated these parameters for the original Mason classification. Morgan et al.⁹ evaluated the reliability of the Mason classification in classifying 25 radial head fractures by 20 practicing orthopedic surgeons on two occasions at least three weeks apart. They reported a substantial

Table 5

Comparisons of inter- and intraobserver variation between the different classification methods (including those that could not be classified (cc) as a category).

Classification method (including cc)	Interobserver Kappa Mean (SD)	Intraobserver Kappa Mean (SD)
Mason	0.33 (0.12)	0.43 (0.28)
Charalambous (8 morphology categories)	0.42 (0.08)	0.49 (0.11)
Charalambous (4 morphology categories)	0.54 (0.06)	0.59 (0.10)
Charalambous (un-displaced vs. displaced)	0.35 (0.19)	0.50 (0.14)

Table 6

Agreement between 5 raters on Mason classification for 30 radial head fractures (including could not classify as a category).

Outcome	Kappa
1	0.39
2	0.43
3	0.27
cc ^a	0.22
Combined	0.32

^a cc indicates when a rater could not classify.

Table 7

Agreement between 5 raters on CPC classification (eight categories plus could not classify as a category) for 30 radial head fractures.

Outcome	Kappa
1 d	0.50
1u	0.54
2 d	0.42
2u	−0.03
3 d	0.47
3u	0.03
4 d	0.45
4u	0.51
cc ^a	0.09
Combined	0.41

^a cc indicates when a rater could not classify.

Table 8

Proportion of radial head fractures that could not be classified by raters by each classification (n = 30).

Classification	Mason (n/30(%))	Charalambous (n/30(%))
Raters could classify	9 (30.0)	23 (76.7)
1 rater could not classify	11 (36.7)	4 (13.3)
2 raters could not classify	3 (10.0)	3 (10.0)
3 raters could not classify	4 (13.3)	0 (0.0)
4 raters could not classify	2 (6.7)	0 (0.0)
5 raters could not classify	1 (3.3)	0 (0.0)

interobserver variability with a median kappa value of 0.54 at the first and 0.64 at the second assessment. In addition, they reported an intraobserver reproducibility ranging from 60 to 92% with a mean of 78%, although kappa values were not reported. It is of note however, that in contrast to our study, Morgan et al. made no reference to fractures that could not be classified with the Mason classification, although in three cases no fracture could be identified by some of the surgeons. Furthermore, the level or previous experience of the participating surgeons was not described.

Our findings suggest that plain radiographs may have a limited role in classifying radial head fractures for clinical or research purposes; hence more sensitive radiological imaging such as Computed

Tomography may be necessary. The study by Rhyou¹⁰ et al. reported high interobserver and intraobserver reliability for the Charalambous classification when CT scan with 3D reconstruction images was used for classification. However, in that study only two observers participated.

It is of particular interest that the raters in this study when using the Charalambous classification were able to classify the morphological subgroups more reliably as compared to the displaced/un-displaced groups. This may suggest that when using plain radiographs raters are more likely to reliably identify configuration of fracture lines as compared to detecting displacement, particularly when the amount of displacement is not pre-defined.

There are certain limitations to this study. Although all raters were given verbal and written description of both classifications, they were more familiar with the Mason classification as they have been using it throughout their clinical practice, compared with the more recently described Charalambous classification. Furthermore, the radiographs used in the study were not absolutely standardized but they were the ones utilized in routine clinical practice. When looking at the agreement of only those fracture patterns that could be classified, only 9 out of the 30 radiographs could be classified using the Mason classification by all 5 raters. Although this limits the weight that can be placed on the results of this sensitivity analysis we looked at each pair of raters individually and so sample sizes varied (being higher than 9 in most cases).

5. Conclusion

In conclusion, the Mason and Charalambous classifications for radial head fractures confer similar reliability when using plain radiographs, but the latter allows a greater proportion of fractures to be classified. Raters may agree more on fracture morphology as compared to fracture displacement. Our findings highlight the limitations of using plain radiographs in classifying radial head fractures for clinical or research purposes.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jor.2019.05.012>.

References

- Altman DG. *Practical Statistics For Medical Research*. NOVEMBER 22 London: Chapman & Hall/CRC; 1990 ISBN 9780412276309-CAT C7630.
- Broberg MA, Morrey MF. Results of treatment of fracture –dislocations of the elbow. *Clin Orthop Relat Res*. 1987 Mar;216:109–119.
- Charalambous CP, Stanely JK, Mills SP, Hayton JM, Hearnden A, Trail I. Comminuted radial head fractures: aspect of current management. *Shoulder Elbow Surg*. 2011;1–12 <https://doi.org/10.1016/j.jse.2011.02.013>.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:27–46.
- Hotchkiss RN. Displaced fractures of the radial head: internal fixation or excision? *J Am Acad Orthop Surg*. 1997;5(1):1–10.
- Johnston GW. A follow –up of one hundred cases of fracture of the head of the radius with a review of the literature. *Ulster Med J*. 1962;31:51–56.
- Kodde IF, Kass L, Flipson M, Bekerom M PJ Van den, Eygendaal D. Current concepts in the management of radial head fractures. *World J Orthoped*. 2015 Dec 18;6(11):954–960 <https://doi:10.5312/wjo.v6.i11.954>.
- Mason ML. Some observations on fractures of the head of the radius with a review of one hundred cases. *Br J Surg*. 1954;42:123–132.

9. Morgan SJ, Groshen SL, Itamura JM, Kuschner SH. Reliability evaluation of classifying radial head fractures by the system of Mason. *Bull Hosp Jt Dis.* 1997;56:95–98.
10. Rhyou IH, Kim KC, Kim KW, Ji-Ho Lee, Kim SY. Collateral ligament injury in the displaced radial head and neck fracture: correlation with fracture morphology and management strategy to the torn ulnar collateral ligament. *J Shoulder Elb Surg.* 2013 Feb;22(2):261–267 <https://doi.org/10.1016/j.jse.2012.10.020>.
11. StataCorp. *Stata Statistical Software: Release 13.* College Station TSL. StataCorp. 2013. College Station, TX: StataCorp L; 2013 Stata Statistical Software: Release 13.