Evaluation of Reliability of the Mason Classification for Radial Head Fractures

Mark T. Dillon, MD, Charles L. Getz, MD, Pedro K. Beredjiklian, MD, Brent B. Wiesel, MD, Gregory F. Carolan, MD, and Matthew L. Ramsey, MD

Abstract

In the study reported here, we sought to determine the interobserver reliability and the intraobserver reproducibility of the Mason classification. We also evaluated the effect of having an external rotation oblique view on agreement in radiographic readings. Four readers reviewed 50 radial head fracture radiographs approximately 2 months apart. Half the radiographs had an anteroposterior view and a lateral view; the other half had an additional external rotation oblique view. There was a trend toward improved interobserver agreement in the 3-view radiographs. Three of the 4 readers demonstrated substantial intraobserver reproducibility, which was noted to be higher when 3 views were available.

Radial head and neck fractures are relatively common injuries involving the proximal aspect of the radius at its articulation with the capitellum of the distal humerus. Mason[1] first categorized proximal radius fractures more than 50 years ago. Type I fractures are nondisplaced; this category was modified by Broberg and Morrey[2] to include all radial head and neck fractures displaced less than 2 mm. In the Broberg–Morrey modification, type II fractures have more than 2 mm of displacement involving at least 30% of the radial head. Type III fractures have significant comminution.[2] Johnston[3] added to the classification a type IV fracture, a radial head or neck fracture associated with an elbow dislocation.

Dr. Dillon is Attending Surgeon, Department of Orthopaedic Surgery, The Permanente Medical Group, Sacramento, California.
Dr. Getz is Assistant Professor of Orthopaedic Surgery, Rothman Institute and Thomas Jefferson University, Philadelphia, Pennsylvania.
Dr. Beredjiklian is Associate Professor of Orthopaedic Surgery, Thomas Jefferson University, and Chief of Hand Surgery Division, Rothman Institute, Philadelphia, Pennsylvania.
Dr. Wiesel is Chief of Shoulder Service, Georgetown University Hospital, Washington, DC.
Dr. Carolan is Assistant Clinical Professor of Orthopaedic Surgery, Temple University School of Medicine, Philadelphia, Pennsylvania, and St. Luke's Orthopaedic Specialists, Bethlehem, Pennsylvania.
Dr. Ramsey is Associate Professor of Orthopaedic Surgery, Rothman Institute and Thomas Jefferson University, Philadelphia, Pennsylvania.

Address correspondence to: Matthew L. Ramsey, MD, Rothman Institute, Thomas Jefferson University, 925 Chestnut St, Philadelphia, PA 19107 (fax, 215-503-0580; e-mail, matthew.ramsey@rothmaninstitute.com).

Am J Orthop. 2010;39(9):430-432.

The modified Mason classification is used to evaluate radiographs and help determine treatment regimens. In most cases, type I fractures are treated conservatively, with brief immobilization followed by early motion, while Mason type II and III fractures often lend themselves to surgical intervention.

Given the widespread use of the modified Mason classification in determining treatment regimens for radial head fractures, it is surprising that there are few studies evaluating its effectiveness.[4,5] In the present study, we sought to evaluate the interobserver reliability and the intraobserver reproducibility of the modified Mason classification for radial head and radial neck fractures. We also wanted to determine whether having an oblique external rotation view would increase agreement among observers.

"We also wanted to determine whether having an oblique external rotation view would increase agreement among observers."

Patients and Methods

All patients older than 18 years and presenting with a diagnosis of a radial head or radial neck fracture to either the Shoulder and Elbow service or the Hand service between the years 2000 and 2006 were identified. Radiology reports were subsequently reviewed, and patients with available digitized injury radiographs were used for this study. Patients with Mason I, II, and III fractures were included, and patients with Mason IV fracture-dislocations or other associated injuries about the elbow were excluded.

Radiographs were chosen by 2 of the authors not involved in evaluations of the films. Radiographs of sufficient quality were chosen to ensure an appropriate variety of fracture patterns. All radiographs were deidentified to protect patient privacy. Fifty elbows were used in this study. Twenty-five of them had only an anteroposterior (AP) view and a lateral view; the other 25 had an additional external rotation oblique view. Oblique views were obtained with the forearm externally rotated from a position used for a standard AP view.[6] In addition, of the 50 studies, 6 (3 with 2 views, 3 with 3 views) did not demonstrate a fracture.

Table I. Interobserver Reliability

	First Reading	Second
Reading		
All studies	.44	.37
2 views only	.43	.31
3 views only	.46	.43

All images were saved as high-resolution JPEG files and viewed on a ViewSonic VP201b monitor (ViewSonic, Walnut, California) in portrait mode using Microsoft Windows Picture and Fax Viewer (Microsoft, Redmond, Washington). Contrast levels were set as needed to best visualize osseous details and were generally adjusted sparingly. Therefore, reviewers could not alter the contrast but could rotate the images and zoom as needed.

Four orthopedic surgeons were recruited to participate in this study: a fellowship-trained shoulder and elbow surgeon in an academic practice, a fellowship-trained hand surgeon in an academic practice, and 2 residents in their fifth year of training. All were familiar with the Mason classification. They were given a chart of the Broberg–Morrey modification and were able to refer to it as often as needed during their evaluation of radiographs. They also were to view each radiograph as long and often as needed, but, after recording an answer, they were not allowed to return to a series already reviewed.

The radiographs were reviewed on 2 occasions at least 2 months apart. The order in which radiographs were presented was changed between sittings. All participants were asked to avoid discussing the study with one another between viewings.

With use of Kappa values (κs), interobserver reliability was calculated between all 4 surgeons, and intraobserver reproducibility was calculated for each participant. Kappa values allow for determining agreement when the contribution caused by chance is taken into account. Values of 1.00 indicate perfect agreement, and values of .00 indicate agreement caused by chance alone. Unweighted κs were used for this study.

Kappa values were determined by examining all 50 radiographs as a group. The group was then divided into 2 subgroups according to availability of a third, oblique view to see if having this additional view led to agreement better than that obtained with only an AP view and a lateral view. The method of Fleiss[7] was used to determine agreement among several raters. According to the system of Landis and Koch,[8] κs of .00 to .20 indicate slight agreement; .21 to .40, fair agreement; .41 to .60, moderate agreement; .61 to .80, substantial agreement; and .81 to 1.00, almost perfect agreement. In this study, no attempt was made to determine the accuracy of readings against a known interpretation, including the presence or absence of a fracture.

Table II. Intraobserver Reproducibility

	Reader A	Reader B	Reader C	Reader D
All studies	.67	.16	.63	.67
2 views only	.50	.22	.59	.66
3 views only	.84	.09	.65	.68

Results

There was agreement among all 4 observers on 34% of the studies on the first reading and 32% of the studies on the second reading. Computed κs were .44 (first reading, moderate agreement) and .37 (second reading, fair agreement) (Table I). The effect of the additional view was then assessed. For the first reading, κs were .43 (2 views) and .46 (3 views); for the second reading, κs were .31 (2 views) and .43 (3 views).

Intraobserver reproducibility for 3 of the 4 readers was in the range of substantial agreement, while 1 demonstrated slight agreement. For the 2 attendings, readers A and B, κs were .67 and .16, respectively; for the 2 residents, readers C and D, κs were .63 and .67, respectively (Table II). For the attending surgeons, κs for studies with only 2 views were .50 for reader A and .22 for reader B, and κs for studies with 3 views were .84 for reader A and .09 for reader B. Reader C had a κ of .59 for only 2 views and .65 for 3 views. Reader D had κs of .66 and .68, respectively. The only κ representing almost perfect agreement was intraobserver reliability for reader A for 3 views.

Discussion

Interobserver reliability for categorizing radial head fractures was found to be moderate to fair (κ = .44 for first reading, κ = .37 for second reading). We also noted a trend toward improved interobserver agreement for series for which an additional external oblique view was available. This was especially true for the second reading, in which fractures with 3 views were read with moderate agreement (κ = .43) and fractures with only an AP view and a lateral view were read with fair agreement (κ = .31).

Our data compare well with the limited information available for the reliability of evaluating radial head fractures with the Mason system. Morgan and colleagues[4] reported on the reliability of the Mason classification with 20 surgeons classifying 25 sets of AP and lateral radiographs. Pairwise κs were computed between all surgeons, with values for the first reading ranging from .142 to .821 and for the second reading from .186 to .940. However, overall κs were not computed. Sheps and colleagues[5] recently examined interobserver reliability for the Hotchkiss[9] modification of the Mason system (κ = .585) and the Arbeitsgemeinschaft für Osteosynthesefragen classification system[10] (κ = .261). They combined Hotchkiss system groups II and III and created 1 group for nonoperative treatment and the other for operative treatment. With this modification, κ increased into the range of substantial agreement (κ = .760).

Morgan and colleagues[4] also examined intraobserver agreement. Their results demonstrated that reproducibility was fair to poor 60% of the time; the authors did not comment on the influence of experience level on their results. When we examined intraobserver reproducibility in our study, we found that 3 of the 4 readers had substantial agreement (κs = .67, .63, .67). One reader, a fellowship-trained orthopedic surgeon, had a κ of .16, indicating only slight agreement. In addition, 1 resident and the other attending surgeon had equally high κs, which suggests that intraobserver reproducibility is not experience dependent. This is in contrast to studies involving other anatomical areas, such as in the evaluation of proximal humerus fractures.[11,12]

There also appeared to be a trend toward improving intraobserver agreement with the addition of an external rotation oblique view. Again, this result did not prove to be experience dependent, as one attending had slightly worse reproducibility when reading studies with 3 views. However, the results for the other 3 surgeons showed improved agreement for 3 views compared with only 2 views. This was most apparent for the attending surgeon (reader A) who had moderate agreement for only 2 views (κ = .50) but almost perfect agreement for 3 views (κ = .84).

It is important to point out that this study does not demonstrate that adding a third image improves the reproducibility of reading any single fracture. Rather, this study points out that series with 3 views are read more consistently between observers and among observers than series with only 2 views.

The effect of additional radial head views on diagnosing fractures has been examined before. On the basis of a cadaver study, Greenspan and colleagues[13] recommended using an additional radial head-capitellum view in diagnosing radial head fractures. Grundy and colleagues,[14] who had 4 physicians review radiographs of radial head fractures, found that adding a radial head–capitellum view provided additional information in 21% of cases. However, most studies have found that an additional radial-capitellum view does little to help in identifying presence of a fracture.[15-17] Traditionally, the need for agreement is highest not in diagnosing a fracture but in using radiographs to differentiate between indications for nonoperative treatment for Mason I fractures and indications for surgery for Mason II and III fractures. However, even the need for surgery on all Mason II fractures is being questioned.[18]

We have shown that the modified Mason classification has only moderate interobserver and intraobserver agreement, as shown in previous studies. However, we have demonstrated improved reproducibility when 3 views are available. In the evaluation of a patient with a radial head fracture, adding a third external rotation view may improve the consistency of treatment.

Authors' Disclosure Statement

The authors report no actual or potential conflict of interest in relation to this article.

References

1. Mason ML. Some observations on fractures of the head of the radius with a review of one hundred cases. *Br J Surg.* 1954;42(172):123-132.
2. Broberg MA, Morrey BF. Results of treatment of fracture-dislocations of the elbow. *Clin Orthop Relat Res.* 1987;(216):109-119.
3. 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.
4. Morgan SJ, Groshen SL, Itamura JM, Shankwiler J, Brien WW, Kuschner SH. Reliability evaluation of classifying radial head fractures by the system of Mason. *Bull Hosp Jt Dis*. 1997;56(2):95-98.
5. Sheps DM, Keifer KR, Boorman RS, et al. The interobserver reliability of classification systems for radial head fractures: the HotchKiss modification of the Mason classification and the AO classification systems. *Can J Surg.* 2009;52(4):277-282.
6. Ballinger PW. *Merril's Atlas of Radiographic Positions and Radiologic Procedures*. 8th ed. St. Louis, MO: Mosby; 1995.
7. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bull*. 1971;76(5):378-382.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
9. Hotchkiss RN. Displaced fractures of the radial head: internal fixation or excision? *J Am Acad Orthop Surg.* 1997;5(1):1-10.
10. Muller ME, Nazarian S, Koch P, et al. *The Comprehensive Classification of Fractures of Long Bones.* Berlin, Germany: Springer-Verlag; 1990.
11. Bernstein J, Adler LM, Blank JE, et al. Evaluation of the Neer system of classification of proximal humerus fractures with computerized tomographic scans and plain radiographs. *J Bone Joint Surg Am.* 1996;78(9):1371-1375.
12. Sidor ML, Zuckerman JD, Lyon T, et al. The Neer classification system for proximal humerus fractures: an assessment of interobserver reliability and intraobserver reproducibility. *J Bone Joint Surg Am.* 1993;75(12):1745-1750.
13. Greenspan A, Norman A, Rosen H. Radial head-capitellum view in elbow trauma: clinical application and radiographic-anatomic correlation. *AJR Am J Roentgenol.* 1984;143(2):355-359.
14. Grundy A, Murphy G, Barker A, et al. The value of the radial head-capitellum view in radial head trauma. *Br J Radiol.* 1985;58(694):965-967.
15. Hall-Craggs MA, Shorvon PJ, Chapman M. Assessment of the radial head-capitellum view and the dorsal fat-pad sign in acute elbow trauma. *AJR Am J Roentgenol*. 1985;145(3):607-609.
16. Manns RA, Lee JR. Critical evaluation of the radial head-capitellum view in acute elbow trauma with effusion. *Clin Radiol.* 1990;42(6):433-436.
17. Page AC. Critical evaluation of the radial head-capitellum view in elbow trauma. *AJR Am J Roentgenol.* 1986;146(1):81-82.
18. Akesson T, Herbertsson P, Josefsson P, Hasserius R, Besjakov J, Karlsson MK. Primary nonoperative treatment of moderately displaced two-part fractures of the radial head. *J Bone Joint Surg Am.* 2006;88(9):1909-1914.