Observer variation in MRI for suspected scaphoid fractures

[1]F J P BEERES, PhD, MD, [2]M HOGERVORST, MD, [3]L M KINGMA, PhD, MD, [4]S LE CESSIE, PhD, [3]E G COERKAMP, MD and [1]S J RHEMREV, MD

[1]*Department of Surgery, Medical Centre Haaglanden, The Hague,* [2]*Department of Surgery, Gelre Ziekenhuizen, Apeldoorn,* [3]*Department of Radiology, Medical Centre Haaglanden, The Hague and* [4]*Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands*

ABSTRACT. The aim of this study was to evaluate prospectively the observer variation in the detection of pathology on MRI for suspected acute scaphoid fracture. 79 consecutive MR scans were included to calculate the inter-observer variation. All patients were suspected of having a scaphoid fracture but had no fracture on radiographs. 38 randomly chosen MR scans were used to calculate the intra-observer variation. Four observers, with varying levels of expertise, blinded scored three items: (i) scaphoid fracture, (ii) localization of a scaphoid fracture, and (iii) another fracture. The observer variation was analysed using the kappa statistic. The inter-observer variation for a scaphoid fracture showed substantial agreement. For the localization of a scaphoid fracture and another fracture, there was a moderate and substantial agreement, respectively. The intra-observer variation for a scaphoid fracture had an almost perfect agreement. For the localization of a scaphoid fracture and another fracture, there was an almost perfect and substantial agreement, respectively. In conclusion, the observer variation in MRI of suspected scaphoid fractures was low. The influence of expertise with MRI in daily practice should be taken into consideration. Observers with little experience of MRI will identify all scaphoid fractures but are likely to over-diagnose injuries. Based on these results, it is recommended that all scans are reviewed by an experienced radiologist.

Received 17 June 2007
Revised 3 November 2007
Accepted 9 December 2007

DOI: 10.1259/bjr/72567977

The scaphoid bone plays a pivotal role in the function of the wrist. Lesions of this bone can therefore lead to major functional impairment. In addition, the scaphoid bone is the most frequently fractured carpal bone, accounting for approximately 70% of all carpal fractures [1–3].

Delayed treatment of a scaphoid fracture increases the risk of non-union [2–6]. A fast and accurate imaging method can therefore enhance treatment outcome. Conventional imaging methods are not sufficient, as initial radiographs detect only 79% of all scaphoid fractures [7]. Repeated radiographs do not lead to an improvement in diagnostic accuracy, as the added sensitivity is low [8–11]. Because of its high sensitivity (95%), the use of bone scintigraphy is often recommended in the case of a suspected scaphoid fracture that is not detected on initial radiographs. However, the outcome of bone scintigraphy has to be interpreted with care because the specificity is between 60% and 95% [1, 8, 10, 12–14]. Moreover, nuclear imaging requires intravenous radioactive isotopes and, in general, a delay of 72 h after injury. Data regarding CT are limited and suggest that CT is not an appropriate diagnostic tool [15, 16].

Recently, MRI has gained widespread interest for the assessment of suspected scaphoid fractures. In a well-controlled study environment, the results of MRI look particularly promising. [13, 16–20]. To determine the role of MRI in daily practice, however, the observer variation and the influence of expertise have to be evaluated.

Address correspondence to: F J P Beeres, Medical Centre Haaglanden, T.a.v. Secretariaat Heelkunde, Postbus 432, 2501 CK The Hague, The Netherlands. E-mail: f.j.p.beeres@lumc.nl

Methods and materials

Patients

This prospective study was conducted in accordance with the standards of the regional ethical committee. Between April 2004 and November 2005, 79 consecutive patients (43 men, 36 women; age range, 18–84 years; mean age, 41 years) were enrolled. Both oral and written informed consent was obtained. Patients older than 18 years with a tender anatomical snuffbox, pain on application of axial pressure and a recent history of acute trauma but no radiological evidence of a scaphoid fracture were eligible for inclusion. Patients who refused or had a contraindication for MRI and polytrauma patients were excluded.

Study protocol

Patients with a suspected scaphoid fracture but no radiological evidence of a scaphoid fracture underwent MRI of the wrist, without a plaster and within 24 h. At two different points in time, 10 months apart, four observers filled in a standardized scoring sheet for 38 randomly chosen patients from the group of 79 patients.

Scaphoid radiographs

Initial scaphoid radiographs were carried out in three planes: (i) a posteroanterior view with the hand in the neutral position; (ii) an oblique view with the hand at 10° supination and maximal ulnar deviation; and (iii) a lateral view with the wrist resting in the ulnar position. Initially, all radiographs were judged by the attending resident surgeon in the Emergency Department and a resident radiologist. Subsequently, the consultant surgeon and consultant radiologist judged the radiographs.

MRI

A 1.5 Tesla MR scanner (Siemens, Erlangen, Germany) was used. The patient lay down in a supine position on the scanner couch — head first with the arm containing the suspected scaphoid fracture stretched forward. A flexible surface coil was wrapped around the wrist. The MRI protocol included coronal T_1 weighted turbo spin-echo images with a repetition time (TR) of 450 ms, an echo time (TE) of 13 ms, a field view of 180 mm, a matrix of 210 × 512, two averages, a slice thickness of 3 mm with a distance factor of 10%, and a scan time of 2.17 min. The parameters for the coronal fat-suppressed T_2 weighted fast spin-echo images were 5220/73 ms (TR/TE), a field of view of 220 mm, a matrix of 216 × 448, three averages, a slice thickness of 3 mm with a distance factor of 10%, and a scan time of 4.33 min.

Observers

Observer 1 was a consultant radiologist with specific skeletal MR scanning expertise. Observer 2 was an experienced emergency consultant radiologist with no specific experience of skeletal MR scanning. Observers 3 and 4 were residents with 5 years and 3 years of experience, respectively, and a completed MR scanning training period.

Measurements

All MR scans were blind coded, *i.e.* without any patient-identifying factors. The four observers scored all 79 MR scans using a soft-copy reading on a Leonardo workstation (Siemens, Erlangen, Germany). They filled out a standard scoring sheet blind to each other and blind to any other data (*i.e.* scaphoid radiographs, other diagnostic modalities, clinical outcome, etc). For every MR scan, each observer scored the following three items:

1. Scaphoid fracture (yes/no).
2. Localization of a scaphoid fracture (proximal third/middle third/distal third).
3. Another fracture (other carpal, metacarpal or ulnar fracture/distal radius fracture/bone bruise of the scaphoid/bone bruise of another carpal bone/bone bruise of the distal radius/no traumatic injury).

10 months apart, 38 scans randomly chosen from the group of 79 patients were evaluated by all four observers. The same blinded and standardized scoring sheet was used.

Statistical analysis

Both the inter- and intra-observer variation was calculated for the four observers. For a scaphoid fracture, the simple kappa (κ) coefficient was calculated because the data were nominal (scaphoid fracture — yes or no). For the localization of a scaphoid fracture or another fracture, the weighted κ coefficient was calculated because the data were ordinal.

The inter-observer variation was calculated for the 79 scans. For the inter-observer variations, an overall κ statistic, along with the pairwise κ statistics, is provided [21]. As overall weighted κ, an intraclass correlation coefficient was calculated [22]. For the intra-observer variation, the randomly chosen 38 scans were used.

The κ statistic is a chance corrected measure of agreement for data. The κ value has a range from +1 (for perfect agreement) to –1 (which corresponds to absolute disagreement). A value of 0 indicates no more agreement than expected by change alone. Interpretation of the κ value was based on Landis and Koch [23] guidelines which suggest that values between 0 and 0.2 represent slight agreement, 0.21 and 0.40 fair agreement, 0.41 and 0.60 moderate agreement, and 0.61 and 0.80 substantial agreement. A value above 0.80 is considered an almost perfect agreement.

Results

An overview of the injuries scored (in 79 scans) by the four observers is shown in Table 1. An example of an MR scan of a patient with an occult scaphoid fracture is

Table 1. Overview of the injuries scored in the 79 MR scans

	Observer 1	Observer 2	Observer 3	Observer 4
Scaphoid fracture	8	14	15	11
Other carpal fracture	15	7	11	9
Distal radius fracture	10	12	11	11
Bone bruise scaphoid	1	4	1	4
Bone bruise other carpal bone	8	6	11	12
Bone bruise distal radius	1	6	0	0
No injury	36	30	30	32
Total	79	79	79	79

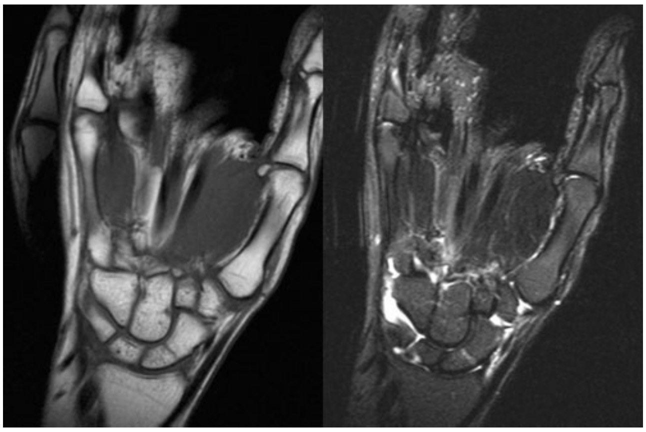

Figure 1. All four observers identified an occult scaphoid fracture and all observers localized it to the scaphoid waist on the MRI scan.

shown in Figure 1. Figure 2 shows a number of MR images with discrepancies between the observers.

1. Inter-observer variation

The inter-observer variation was calculated for four observers for 79 scans.

A scaphoid fracture

There was substantial agreement as shown in Table 2. The pairwise and overall κ statistic was 0.67 (95% confidence interval of 0.44–0.90).

The localization of a scaphoid fracture

There was moderate agreement as shown in Table 3. The pairwise and overall κ statistic was 0.57 (95% confidence interval of 0.47–0.68).

Another fracture

There was substantial agreement. The paiwise and overall κ statistic was 0.61 (95% confidence interval of 0.51–0.71).

2. Intra-observer variation

The intra-observer variation was calculated for 38 patients.

A scaphoid fracture

There was almost perfect agreement with a median pairwise κ of 0.96 (minimum 0.69, maximum 1.0) as shown in Table 4.

The localization of a scaphoid fracture

There was almost perfect agreement with a median pairwise κ of 0.90 (minimum 0.60, maximum 0.97) as shown in Table 5.

Another fracture

There was substantial agreement with a median pairwise κ of 0.76 (minimum 0.53, maximum 0.79).

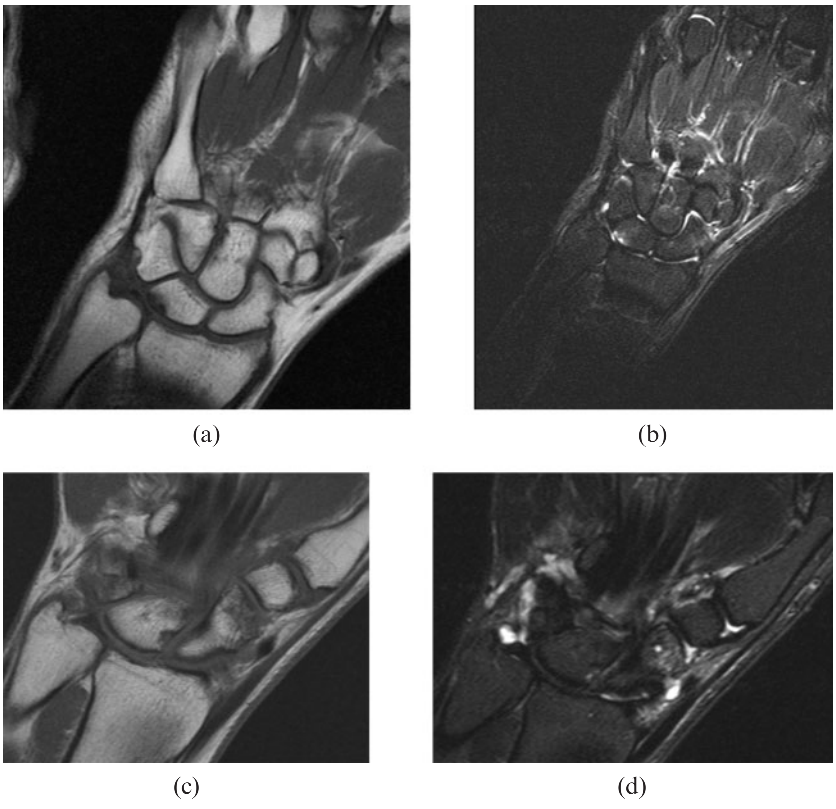

Figure 2. (a) T_1 and (b) T_2 weighted images from a patient for whom there were discrepancies between the observers regarding the presence of a capitate fracture. (c) T_1 and (d) T_2 weighted images from a patient for whom there were discrepancies between the observers for a scaphoid fracture.

Table 2. Cross-tabulation showing the inter-observer variation (κ coefficient) for 79 MR scans relating to a scaphoid fracture (95% confidence intervals in parentheses)

Observer	1	2	3	4
1	X	0.69 (0.46–0.92)	0.55 (0.30–0.80)	0.70 (0.46–0.95)
2		X	0.70 (0.50–0.91)	0.76 (0.57–0.96)
3			X	0.63 (0.40–0.86)
4				X

Table 3. Cross-tabulation showing the inter-observer variation (κ coefficient) for 79 MR scans relating to the localization of a scaphoid fracture (95% confidence intervals in parentheses)

Observer	1	2	3	4
1	X	0.58 (0.33–0.84)	0.46 (0.23–0.69)	0.54 (0.27–0.81)
2		X	0.60 (0.40–0.81)	0.68 (0.46–0.91)
3			X	0.56 (0.33–0.78)
4				X

Table 4. Results at the first and second evaluation, and the intra-observer variation (κ coefficient) for a scaphoid fracture (38 patients) (95% confidence intervals in parentheses)

	Observer 1		Observer 2		Observer 3		Observer 4	
Evaluation	1	2	1	2	1	2	1	2
Scaphoid fracture	4	4	10	6	6	7	7	7
No scaphoid fracture	34	34	28	32	32	31	31	31
Simple κ	1.00 (1.00– 1.00)		0.69 (0.41–0.96)		0.91 (0.73–1.00)		1.00 (1.00–1.00)	

Table 5. Results at the first and second evaluation, and the intra-observer variation (κ coefficient) for the localization of a scaphoid fracture (38 patients) (95% confidence intervals in parentheses)

	Observer 1		Observer 2		Observer 3		Observer 4	
Evaluation	1	2	1	2	1	2	1	2
Proximal third	0	0	1	1	0	0	0	1
Middle third	4	3	3	2	2	1	3	2
Distal third	0	1	6	3	4	6	4	4
Weighted κ	0.93 (0.83–1.00)		0.60 (0.30–0.90)		0.87 (0.67–1.00)		0.97 (0.90–1.00)	

Discussion

This study shows that the inter-observer variation in MRI shows (i) substantial agreement for a scaphoid fracture (average κ, 0.67), (ii) moderate agreement for the localization of a scaphoid fracture (average κ, 0.57), and (iii) substantial agreement for another fracture (average κ, 0.61).

For the intra-observer variation in MRI, we found (i) almost perfect agreement for a scaphoid fracture (median κ, 0.90), (ii) almost perfect agreement for the localization of a scaphoid fracture (median κ, 0.85), and (iii) substantial agreement for another fracture (median κ, 0.71).

To our knowledge, there are no other reports evaluating the observer variation in MRI of (occult) scaphoid fractures in daily practice. There is an obvious need for a fast and reliable diagnostic procedure in cases of a suspected scaphoid fracture. It is important to diagnose and treat every patient with a scaphoid fracture because even occult fractures can lead to impairment. Moreover, differentiation between a scaphoid fracture, a bone bruise and another fracture has clinical implications. MRI has been proposed as the modality of choice for detecting suspected scaphoid fractures [13, 16–20]. However, the observer variation in daily practice should be assessed in order to define the place of MRI in the management of these fractures. This is important because MR scans of suspected scaphoid fractures are often (initially) evaluated by the attending radiologist, who often has no specific MR expertise.

One could consider Observer 1, being the consultant with the most experience of MRI, as the gold standard. Even though this was not the aim of this study, it appeared that the other, less experienced observers were over-diagnosing scaphoid fractures, with approximately one-third of their diagnoses being false-positive. It is important not to miss fractures and, therefore, over-diagnosing scaphoid fractures is preferred to under-diagnosing them. A characteristic of the κ statistic is that its value depends upon the number of categories in the outcome scale. In general, the results regarding a scaphoid fracture (yes/no) are likely to be higher than the other evaluations (localization of a scaphoid fracture and another fracture), as these other evaluations have more possible answers. This could explain the lower weighted κ values.

In order to assess a fast protocol that could be easily implemented in daily practice, no sagittal plane was used. Abnormal signal intensity could result from a partial volume artefact, which might cause false-positive findings. A cyst or bone bruise could have been interpreted as a fracture, as shown in Figure 2.

In conclusion, the observer variation in MRI of suspected scaphoid fractures was low. The influence of expertise with MRI in daily practice should be taken into consideration. Observers with little experience of MRI will identify all scaphoid fractures but are likely to over-diagnose injuries. Based on these results, it is recommended that all scans are reviewed by an experienced radiologist.

References

1. Bayer LR, Widding A, Diemer H. Fifteen minutes bone scintigraphy in patients with clinically suspected scaphoid fracture and normal x-rays. Injury 2000;31:243–8.
2. Cooney WP III. Scaphoid fractures: current treatments and techniques. Instr Course Lect 2003;52:197–208.
3. Tiel-van Buul MM, Roolker W, Verbeeten BW, Broekhuizen AH. Magnetic resonance imaging versus bone scintigraphy in suspected scaphoid fracture. Eur J Nucl Med 1996;23:971–5.
4. Krasin E, Goldwirth M, Gold A, Goodwin DR. Review of the current methods in the diagnosis and treatment of scaphoid fractures. Postgrad Med J 2001;77:235–7.
5. Prosser GH, Isbister ES. The presentation of scaphoid non-union. Injury 2003;34:65–7.
6. Roolker W, Maas M, Broekhuizen AH. Diagnosis and treatment of scaphoid fractures, can non-union be prevented? Arch Orthop Trauma Surg 1999;119:428–31.
7. Beeres FJ, Hogervorst M, Hollander den P, Rhemrev S. Outcome of routine bone scintigraphy in suspected scaphoid fractures. Injury 2005;36:1233–6.
8. Gaebler C, Kukla C, Breitenseher M, Trattnig S, Mittlboeck M, Vecsei V. Magnetic resonance imaging of occult scaphoid fractures. J Trauma 1996;41:73–6.
9. Tiel-van Buul MM, van Beek EJ, Borm JJ, Gubler FM, Broekhuizen AH, van Royen EA. The value of radiographs and bone scintigraphy in suspected scaphoid fracture. A statistical analysis. J Hand Surg [Br] 1993;18:403–6.
10. Tiel-van Buul MM, Roolker W, Broekhuizen AH, Van Beek EJ. The diagnostic management of suspected scaphoid fracture. Injury 1997;28:1–8.
11. Low G, Raby N. Can follow-up radiography for acute scaphoid fracture still be considered a valid investigation? Clin Radiol 2005;60:1106–10.
12. Fowler C, Sullivan B, Williams LA, McCarthy G, Savage R, Palmer A. A comparison of bone scintigraphy and MRI in the early diagnosis of the occult scaphoid waist fracture. Skeletal Radiol 1998;27:683–7.
13. Beeres FJP, Hogervorst M, Rhemrev SJ, den Hollander P, Jukema GN. A prospectice comparison for suspected scaphoid fractures: Bone scintigraphy versus clinical outcome. Injury 2007;38:769–74.
14. Hambidge JE, Desai VV, Schranz PJ, Compson JP, Davis TR, Barton NJ. Acute fractures of the scaphoid. Treatment by cast immobilisation with the wrist in flexion or extension? J Bone Joint Surg [Br] 1999;81:91–2.
15. Adey L, Souer SJ, Lozano-Calderon S, Palmer W, Lee SG, Ring D. Computed tomography of suspected scaphoid fractures. J Hand Surg 2007;32:61–6.
16. Memarsadeghi M, Breitenseher MJ, Schaefer-Propok C, Weber M, Aldrian S, Gäbler C, et al. Occult scaphoid fractures: comparison of multidetector CT and MR imaging – initial experience. Radiology 2006;240:169–76.
17. Brydie A, Raby N. Early MRI in the management of clinical scaphoid fracture. Br J Radiol 2003;76:296–300.
18. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, et al. Radiographically occult scaphoid fractures: value of MR imaging in detection. Radiology 1997;203:245–50.
19. Hunter JC, Escobedo EM, Wilson AJ, Hanel DP, Zink-Brody GC, Mann FA. MR imaging of clinically suspected scaphoid fractures. AJR Am J Roentgenol 1997;168:1287–93.
20. Nikken JJ, Oei EH, Ginai AZ, Krestin GP, Verhaar JA, van Vugt AB, et al. Acute wrist trauma: value of a short dedicated extremity MR imaging examination in prediction of need for treatment. Radiology 2005;234:116–24.
21. Fleiss JL. Measuring nominal scale agreement among many raters. Psychological Bulletin 1971;76:378–82.
22. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions, 3rd edition, New York: John Wiley & Sons, 2003: 598–626.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.