Comparison of CT and MRI for Diagnosis of Suspected Scaphoid Fractures

By Wouter Mallee, MSc, Job N. Doornberg, MD, PhD, David Ring, MD, PhD, C. Niek van Dijk, MD, PhD, Mario Maas, MD, PhD, and J. Carel Goslings, MD, PhD

Investigation performed at Academic Medical Center of Amsterdam, Amsterdam, The Netherlands

Background: There is no consensus on the optimum imaging method to use to confirm the diagnosis of true scaphoid fractures among patients with suspected scaphoid fractures. This study tested the null hypothesis that computed tomography (CT) and magnetic resonance imaging (MRI) have the same diagnostic performance characteristics for the diagnosis of scaphoid fractures.

Methods: Thirty-four consecutive patients with a suspected scaphoid fracture (tenderness of the scaphoid and normal radiographic findings after a fall on the outstretched hand) underwent CT and MRI within ten days after a wrist injury. The reference standard for a true fracture of the scaphoid was six-week follow-up radiographs in four views. A panel including surgeons and radiologists came to a consensus diagnosis for each type of imaging. The images were considered in a randomly ordered, blinded fashion, independent of the other types of imaging. We calculated sensitivity, specificity, and accuracy as well as positive and negative predictive values.

Results: The reference standard revealed six true fractures of the scaphoid (prevalence, 18%). CT demonstrated a fracture in five patients (15%), with one false-positive, two false-negative, and four true-positive results. MRI demonstrated a fracture in seven patients (21%), with three false-positive, two false-negative, and four true-positive results. The sensitivity, specificity, and accuracy were 67%, 96%, and 91%, respectively, for CT and 67%, 89%, and 85%, respectively, for MRI. According to the McNemar test for paired binary data, these differences were not significant. The positive predictive value with use of the Bayes formula was 0.76 for CT and 0.54 for MRI. The negative predictive value was 0.94 for CT and 0.93 for MRI.

Conclusions: CT and MRI had comparable diagnostic characteristics. Both were better at excluding scaphoid fractures than they were at confirming them, and both were subject to false-positive and false-negative interpretations. The best reference standard is debatable, but it is now unclear whether or not bone edema on MRI and small unicortical lines on CT represent a true fracture.

Level of Evidence: Diagnostic Level I. See Instructions to Authors for a complete description of levels of evidence.
tenderness and normal findings on radiographs actually has a scaphoid fracture and approximately 84% of patients may have unnecessary cast immobilization, resulting in a substantial loss of productivity.

In 2006, Groves et al. performed a worldwide survey and found substantial variation in imaging and treatment protocols for acute scaphoid injuries. Current treatment protocols most commonly include repeat physical examination and radiographs two weeks after the initial presentation, or earlier bone scan, computed tomography (CT), or magnetic resonance imaging (MRI). Current evidence regarding the optimum protocol for the diagnosis of suspected scaphoid fractures lacks methodological quality, which may contribute to a lack of consensus. It has been recommended that MRI is the best radiographic test for the diagnosis of suspected scaphoid fractures, but bone scans, CT, and ultrasound may also be useful, particularly when MRI is not readily available.

CT and MRI have both been studied for their usefulness in establishing the diagnosis of scaphoid fractures. MRI has been reported in case series as having a high sensitivity (98% to 100%) and specificity (100%). The disadvantages of MRI include lack of availability and scheduling issues and cost. CT is more readily available and less costly. CT is more reliable than radiographs are, and it has a greater sensitivity (89% to 100%) and specificity (85% to 100%) than radiographs do. All diagnostic radiographic studies are better at excluding the diagnosis of a nondisplaced scaphoid fracture than they are at confirming a fracture because the prevalence of true fractures among patients with radial-sided wrist pain is low, which serves to magnify the impact of false-positive results.

We are aware of only one prospective study comparing CT and MRI for establishing the correct diagnosis in patients with suspected scaphoid fractures. MRI was both 100% specific and 100% sensitive, whereas CT was 100% specific but only 73% sensitive; however, the CT images were made in planes relative to the wrist and forearm as opposed to the recommended reconstructions in planes defined by the long axis of the scaphoid.

The present prospective study evaluates the sensitivity, specificity, and accuracy as well as the prevalence-adjusted positive and negative predictive values of CT (with reconstructions in the long axis of the scaphoid) and MRI for the diagnosis of suspected scaphoid fractures. We tested the null hypothesis that CT and MRI have the same performance characteristics for the diagnosis of suspected scaphoid fractures.

**Materials and Methods**

The present study was designed and reported according to the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) guidelines (see Appendix). The study was approved by our institutional review board, and all patients gave written informed consent.

### TABLE I CT Versus MRI

<table>
<thead>
<tr>
<th>Scaphoid Fracture</th>
<th>No Scaphoid Fracture</th>
<th>CT</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaphoid fracture</td>
<td></td>
<td>4/4 + 2 = 67%*</td>
<td>4/4 + 2 = 67%*</td>
</tr>
<tr>
<td>No scaphoid fracture</td>
<td></td>
<td>27/1 + 27 = 96%†</td>
<td>25/3 + 25 = 89%†</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>4 + 27/4 + 2 + 1 + 27 = 91%‡</td>
<td>4 + 25 + 4 + 2 + 3 + 25 = 85%‡</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>6/12 = 50%*</td>
<td>6/12 = 50%*</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td></td>
<td>27/53 = 51%†</td>
<td>25/53 = 47%†</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>4 + 27/4 + 2 + 1 + 27 = 91%‡</td>
<td>4 + 25 + 4 + 2 + 3 + 25 = 85%‡</td>
</tr>
<tr>
<td>Positive predictive value (PPV)</td>
<td>4/4 + 1 = 80%§</td>
<td>4/4 = 100%§</td>
<td></td>
</tr>
<tr>
<td>PPV accounting for prevalence and incidence</td>
<td>76%</td>
<td>54%</td>
<td></td>
</tr>
<tr>
<td>Negative predictive value (NPV)</td>
<td>27/27 = 100%*#</td>
<td>25/25 = 100%*#</td>
<td></td>
</tr>
<tr>
<td>NPV accounting for prevalence and incidence</td>
<td>49%</td>
<td>93%</td>
<td></td>
</tr>
</tbody>
</table>

*The proportion of patients who had a scaphoid fracture according to the reference standard and who were classified as having a positive MRI (true-positive). †The proportion of patients who had no scaphoid fracture according to the reference standard and who were classified as having a false-positive MRI result. ‡The proportion of patients who were correctly classified by CT/MRI. §The probability that a patient with a positive CT/MRI has a scaphoid fracture. #The probability that a patient with a negative CT/MRI does not have a fracture.
Patients
Between April 2008 and October 2008, all patients with clinical symptoms of tenderness in the anatomic snuffbox and normal scaphoid-specific radiographs after a fall on an outstretched hand were invited to enroll in a prospective comparison of CT and MRI with regard to their diagnostic utility. The scaphoid-specific radiographs consisted of four views: (1) a posteroanterior view with the wrist in ulnar deviation, (2) a lateral view with the wrist in 15° extension, (3) a lateral view with the wrist in 30° of pronation, and (4) a posteroanterior view with the x-ray beam directed from distal to proximal and with the wrist positioned in 40° of angulation. To be included in this study, the patient had to present within twenty-four hours after injury, have tenderness in the anatomic snuffbox, and have normal scaphoid-specific radiographs. Exclusion criteria were an age of less than eighteen years; any concurrent distal ulnar, radial, or carpal fracture; previous scaphoid fracture; rheumatoid arthritis; and cognitive dysfunction that would limit the physical examination. A radiologist and trauma surgeon evaluated the radiographs.

Forty patients (twenty-five men and fifteen women) were enrolled. Twenty-four patients injured the right hand (twenty-two of these patients were right-hand dominant), and sixteen patients injured the left hand (none of these patients were left-hand dominant). Both CT and MRI were performed on the same day and within an average of 3.6 days (range, zero to ten days) after injury. All wrists with a suspected scaphoid fracture were immobilized in a thumb spica splint or cast until a definitive diagnosis was established. Five wrists diagnosed with a fracture of the waist of the scaphoid were immobilized for ten weeks in a below-the-elbow thumb-spica cast. One patient with a diagnosis of fracture of the distal pole of the scaphoid was immobilized in a below-the-elbow thumb-spica cast for six weeks. Thirty-four patients returned for follow-up radiographs at approximately seven weeks after the injury (average, forty-eight days; range, thirty-five to seventy-four days) and five did not. One patient was a tourist at the time of injury and was no longer in the area, three patients were lost to follow-up, and one withdrew from the study. One patient was excluded because of inadequate image quality due to a motion artifact.

Fig. 1-A

Normal scaphoid-specific radiographs. Fig. 1-B Evidence of a nondisplaced cortical fracture (arrow) on CT (sagittal plane).
Imaging

MRI Protocol

MRI was performed in all patients with use of a 1.0-Tesla open MRI system (Panorama 1.0T; Philips Medical Systems, Eindhoven, The Netherlands). The standard scaphoid protocol (SENSE wrist coil) had a slice thickness of 3 mm and a 0.6-mm gap and included the following series: a localizer image, a coronal slice of a short tau inversion recovery (STIR) sequence, and a coronal slice of a spin-echo T1-weighted sequence, in coronal views. The patient was positioned supine, with the forearm and wrist alongside the body. The open MRI allowed for central placement of the hand relative to the magnetic field, resulting in improved image quality as compared with the image quality obtained with off-centered scanning in the conventional-tube MRI.

CT Protocol

Multidetector high-resolution CT scanning was performed in all patients with use of a Brilliance CT scanner (64 slice; Philips Medical Systems) in a sequence with a high-resolution 0.5-mm-slice section thickness. The scan covered the wrist from the distal radioulnar joint to the carpometacarpal joints. Patients were positioned prone, with the affected arm above the body and with the palm down. Reconstructions in planes defined by the long
axis of the scaphoid were made. Sagittal-plane images of the scaphoid were defined as reconstructions that provided a lateral view of the scaphoid bone, as defined by the central longitudinal axis of the scaphoid. Coronal-plane images were defined as images that provided a posteroanterior view of the scaphoid in the anatomic plane and in line with the axis of the scaphoid.

Reference Standard
Six weeks after the initial injury, scaphoid-specific radiographs were made. This is the most commonly used reference standard in studies of tests for diagnosis of suspected scaphoid fractures and the one used by Memarsadeghi et al. in their comparison of multidetector CT and MR imaging. An abnormal lucent line within the scaphoid was considered to be evidence of a fracture.

Study Design
CT, MRI, and six week follow-up radiographs were separated into three groups and presented to a panel of three observers: an attending musculoskeletal radiologist, an attending trauma surgeon who treats fractures, and an attending orthopaedic surgeon. The panel evaluated the images for the presence of a scaphoid fracture until a consensus opinion was reached. In the absence of consensus, the panel openly discussed the case. The images were blinded, randomly ordered according to a computer random-number generator, and reviewed in two rounds. In the first round, the panel evaluated the initial radiographs and the CT scan; in the second evaluation, they evaluated the initial radiographs and the MRI. The panel was thereby blinded to the CT results during the MRI evaluation and to the MRI results during the CT evaluation. An interval of two weeks between each round of interpretations was used to limit recognition of the radiographs and recall of the CT scan or MRI.

Criteria for a scaphoid fracture on CT images were in accordance with the study protocol of Memarsadeghi et al.: the presence of a sharp lucent line within the trabecular bone pattern, a break in the continuity of the cortex, a sharp step in the cortex, or a dislocation of bone fragments.

Criteria for a fracture on MRI included the presence of a cortical fracture line, a trabecular fracture line, or a combination of both. These criteria are the same as those used by
Memarsadeghi et al. In addition to these criteria, any extensive focal zone of edema without a clear cortical fracture line, comparable with that seen with a stress fracture, was discussed to decide if the findings represented a fracture or not.

**Statistical Analysis**

The sensitivity, specificity, and accuracy for the detection of a scaphoid fracture with CT and with MRI were calculated according to standard formulas (Table I) and with 95% confidence intervals constructed with use of Pratt’s normal approximation method for binomial proportions. The significance of differences was evaluated with use of the McNemar test for paired binary data for each imaging modality. The positive predictive value and negative predictive value were determined with use of the Bayes theorem, which requires an a priori estimate of the prevalence (pretest probability) of the presence of scaphoid fractures. The positive predictive value is the patient’s probability of having a scaphoid fracture when the test is positive, and the negative predictive value is the probability of a patient not having a scaphoid fracture when the test is negative. The predictive values of any imaging modality depend critically on the prevalence of the characteristic in the patients being tested; hence the use of the appropriate Bayesian analysis is important. For the determination of positive and negative predictive values, we estimated an average prevalence of scaphoid fractures of 16% on the basis of the best available data. The positive predictive value was
calculated as sensitivity × prevalence/(sensitivity × prevalence) + [(1 − specificity) × (1 − prevalence)], and the negative predictive value was calculated as specificity × (1 − prevalence)/[(1 − sensitivity) × prevalence] + [specificity × (1 − prevalence)]\(^a\).

Statistical analysis and power analysis were performed to establish the number of patients required for the comparison of diagnostic performance characteristics (sensitivity, specificity, accuracy, positive predictive value, and negative predictive value) between CT and MRI. With use of the McNemar test of equality of paired proportions\(^b\), a sample size of thirty-two patients provided 80% power (\(\alpha = 0.05, \beta = 0.20\)) to detect significant differences in proportions of 20% in each performance characteristic between the two imaging protocols with use of a two-sided significance level of 0.05.

Source of Funding
No external funding source played a role in this study.

Results
Reference Standard
Ten patients (29%) had a fracture of the wrist identified on radiographs made six weeks after injury, six (18%) of whom had a fracture of the scaphoid (one fracture of the distal pole and five waist fractures). One evident fracture that was diagnosed on CT and MRI was not seen on the six-week scaphoid series (Figs. 1-A through 1-D). The remaining four fractures were located in the triquetral (two patients), the capitate (one patient), and the distal part of the radius (one patient). No patients were diagnosed with multiple fractures on the basis of the conventional radiographs.

CT Imaging
CT imaging resulted in a diagnosis of twenty fractures in seventeen patients. Fractures were located in the scaphoid (five fractures: four waist fractures and one fracture of the distal pole), the lunate (two fractures), the triquetral (four fractures), the trapezium (one fracture), the capitate (one fracture), the hamate (one fracture), the distal part of the radius (four fractures), and the metacarpal of the little finger (two fractures). Three patients were diagnosed with multiple fractures. One patient had a fracture of the distal part of the radius and the scaphoid bone, one fractured both the metacarpal of the little finger and the triquetral, and one fractured the trapezium and the capitate. CT identified all of the nonscaphoid fractures seen in the capitate (one fracture), the trapezium (two fractures), the hamate (one fracture), the distal part of the radius (three fractures), and the metacarpal bone of the little finger (one fracture). Three patients were diagnosed with multiple fractures: two patients had a fracture of the distal part of the radius and the scaphoid bone, and one patient fractured both the trapezium and the capitate. Four of the six scaphoid fractures that were diagnosed on the six-week follow-up radiographs were found on MRI. Three additional scaphoid fractures were diagnosed on MRI. The panel diagnosed a fracture of the scaphoid on the basis of the MRI in two patients in whom CT and six-week follow-up radiographs were negative for fractures (Figs. 2-A through 2-D). According to the reference standard, the sensitivity of MRI for correct diagnosis of an occult scaphoid fracture was 67% (95% confidence interval, 35% to 88%), the specificity was 89% (95% confidence interval, 76% to 96%), and the accuracy was 85%.

The difference between the performance characteristics of CT and those of MRI were not significant with the numbers available, according to the results of the McNemar test for paired binary data\(^b\) (\(p > 0.05\)). For MRI, the prevalence-adjusted positive predictive value was 0.54 (95% confidence interval, 0.29 to 0.81) and the prevalence-adjusted negative predictive value was 0.93 (95% confidence interval, 0.80 to 0.98) (Table I).

Discussion
A systematic review\(^c\) of studies that evaluated imaging techniques for the diagnosis of suspected scaphoid fractures found that MRI had an average sensitivity of 98%, a specificity of 99%, an accuracy of 96%, a prevalence-adjusted negative predictive value of 1.00 (meaning that an MRI showing no fracture always corresponded with a true absence of fracture), and a prevalence-adjusted positive predictive value of 0.88 (meaning that a positive MRI corresponded with a true fracture in 88%). That same review showed that CT had an average sensitivity of 94%, a specificity of 96%, an accuracy of 98%, a prevalence-adjusted positive predictive value of 0.75, and a prevalence-adjusted negative predictive value of 0.99. Both MRI and CT were reported to be better at excluding than they were at confirming scaphoid fractures, and MRI has performed slightly better than CT in these noncomparative cohort studies. We could not reproduce these excellent diagnostic performance characteristics in our study. We found a sensitivity of 67% for both MRI and CT, and positive predictive values of 0.54 and 0.76, respectively. These results might be explained by our strict inclusion criteria, as we did not include any fractures that were diagnosed on scaphoid-specific radiographs that were made at the time of injury.

We found that the interpretation of bone-marrow edema on MRI is questionable. In this study, our panel decided that a focal zone of bone edema was considered a fracture. In the study of Memarsadeghi et al., evidence of a zone of diffusely increased...
signal intensity on STIR images was interpreted as bone-marrow edema and not a fracture. In their study, edema had to be accompanied by a cortical fracture line, a trabecular fracture line, or a combination of both to be compatible with a fracture. If we had used these criteria when evaluating MRI in our study, it would have resulted in a diagnosed fracture in four patients (rather than the seven that were identified according to our criteria), with one false-positive result (rather than three in our study), three false-negative results (rather than two in our study), and three true-positive results (rather than four in our study). Sensitivity, specificity, and accuracy for MRI would have been 50%, 96%, and 88% as compared with the actual results of 67%, 89%, and 85%, respectively, in our study. Positive predictive value and negative predictive value, according to the Bayes theorem, would have been 0.70 and 0.91 instead of the actual values of 0.54 and 0.93, respectively, that we found with use of our criteria.

One difficulty that is encountered in a study of suspected scaphoid fractures is the absence of a consensus reference standard for a true fracture. While it is accepted that both MRI and CT scanning may have findings that can be misinterpreted as a fracture (i.e., possible bone bruise on MRI and possible vascular channels on CT), it is not clear that a six-week post-injury radiograph can be used to diagnose all fractures; however, there is currently no viable alternative. In our study, one evident fracture that was diagnosed with perfect agreement by our observers on CT and MRI was not seen on the six-week scaphoid series (Figs. 1-A through 1-D). Therefore, we suspect that the reference standard of six-week post-injury radiographs is inadequate. Subsequent to the design and execution of this study, we became aware of latent class analysis as a technique for analyzing diagnostic performance characteristics in the absence of a consensus reference standard.

Instead of relying on a reference standard to determine diagnostic categories, this statistical technique looks for separate diagnostic groups (classes) in the data. Given that there may never be an adequate reference standard for the diagnosis of true fractures among suspected fractures, latent class analysis may be useful here and will be incorporated in future studies.

The potential weaknesses of our study include the use of a 1.0-Tesla MRI unit, the use of an open MRI unit, the involvement of trauma surgeons, and the use of a consensus panel to diagnose fracture (which might introduce selection bias). We used a 1.0-Tesla open MRI system mainly for practical purposes but also with the understanding that, to achieve the goal of obtaining a high-quality image, the use of a dedicated coil and the central placement of the wrist in a comfortable position within the magnet so as to limit motion artifact would be more important than a stronger magnet or the choice of a closed rather than an open MRI machine. Interpretation of the involvement of a trauma surgeon as a weakness may be a cultural bias or misunderstanding because, in the Netherlands, trauma surgeons treat fractures as often as orthopaedic surgeons do.

Considered in the light of these shortcomings, the null hypothesis that CT and MRI have the same diagnostic performance characteristics for diagnosis of suspected scaphoid fractures was confirmed. CT had accuracy comparable with MRI (91% versus 85%, respectively). Both imaging modalities were better at excluding a scaphoid fracture (negative predictive value, accounting for prevalence and incidence, was 0.94 for CT vs. 0.93 for MRI) than they were at confirming a scaphoid fracture (positive predictive value, accounting for prevalence and incidence, was 0.76 for CT vs. 0.54 for MRI). While additional study is needed, on the basis of our study and the results available in the literature, CT with reconstructions made in planes defined by the long axis of the scaphoid is comparable with MRI for diagnosis of suspected scaphoid fractures.

Appendix

A table showing the quality assessment of the study according to the QUADAS guidelines is available with the electronic version of this article on our web site at jbjs.org (go to the article citation and click on “Supporting Data”).

References