Chapter 7

F.J.P. Beeres¹
M. Hogervorst²
L.M. Kingma¹
S. Le Cessie⁴
E.G. Coerkamp³
S.J. Rhemrev¹

¹Department of Surgery, Medical Centre Haaglanden, The Hague, the Netherlands
²Department of Surgery, Gelre Hospitals, Apeldoorn, the Netherlands
³Department of Radiology, Medical Centre Haaglanden, The Hague, the Netherlands
⁴Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, the Netherlands
Observer variation of MR imaging for suspected scaphoid fractures

Based on:
Observer variation of MR imaging for suspected scaphoid fractures
British Journal of Radiology 2008; in press
Abstract

The objective of the present study was to prospectively evaluate the observer variation in the detection of pathology on MR imaging in suspected scaphoid fractures.

Methods: Seventy-nine consecutive MR scans were included to calculate the inter-observer variation. All patients were suspected of a scaphoid fracture but had no fracture on radiographs. Thirty-eight randomly chosen MR scans were used to calculate the intra-observer variation. Four observers, with varying levels of expertise, blinded scored 3 items; 1) scaphoid fracture, 2) localisation of a scaphoid fracture, and 3) another fracture. The observer variation was analysed using the kappa statistic.

Results: The inter-observer variation for a scaphoid fracture had a substantial agreement. Concerning the localisation 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. Concerning the localisation of a scaphoid fracture and another fracture, there was an almost perfect and substantial agreement respectively.

Conclusion: The observer variation of MR imaging of suspected scaphoid fractures was low. The influence of expertise with MR imaging in daily practice should be taken into consideration. Observers with little experience with MR imaging 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.
Materials and methods

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, mean age 41 years) were enrolled. Both oral and written informed consent was obtained. Patients older than 18 years with a tender anatomic snuffbox, pain when applying axial pressure, a recent history of acute trauma but no radiological evidence of a scaphoid fracture were eligible for inclusion. Patients who refused or had a contra-indication for MR imaging and poly-trauma patients were excluded.

Study protocol
Patients with a suspected scaphoid fracture but no radiological evidence of a scaphoid fracture underwent MR imaging of the wrist, without a plaster and within 24 hours. At 2 different points in time, 10 months apart, 4 observers filled in a standardised scoring sheet for 38 randomly chosen patients from the group of 79 patients.

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

Measurements
All MR scans were blind coded, i.e. without any patient identifying factors. The 4 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 3 items:
1. Scaphoid fracture (yes / no);
2. Localisation 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).

Ten months apart 38 scans, randomly chosen from the group of 79 patients, were evaluated by all 4 observers. The same blinded and standardised scoring sheet was used.

Statistical analysis
Both the inter- and intra-observer variation was calculated for the 4 observers.
The inter-observer variation was calculated for the 79 scans. For the intra-observer variation, the randomly chosen 38 scans were used.

For a scaphoid fracture the simple kappa coefficient was calculated because the data were nominal (scaphoid fracture; yes or no). For the localisation of a scaphoid fracture or another fracture, the weighted kappa coefficient was calculated because the data were ordinal.

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

**Results**

An overview of the injuries scored by the 4 observers of 79 scans is shown in Table 7.1. An example of a MR scan of a patient with an occult scaphoid fracture is shown in Figure 7.1. Figure 7.2 shows a number of images of MR scans with discrepancies between the observers.

![Figure 7.1](image-url)

MR scan of a suspected scaphoid fracture. All 4 observers identified an occult scaphoid fracture and all observers localised it to the scaphoid waist.

**Inter-observer variation**

The inter-observer variation was calculated for 4 observers for 79 scans. Concerning:

1. A scaphoid fracture:
   There was substantial agreement with an average kappa of 0.67 (minimum 0.55, maximum 0.76) as shown in Table 7.2;

2. The localisation of a scaphoid fracture:
Chapter 7
Observer variation of MR imaging

Table 7.1
Overview of the injuries scored in the 79 MR scans.

<table>
<thead>
<tr>
<th></th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
<th>Observer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaphoid fracture</td>
<td>8</td>
<td>14</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>Other carpal fracture</td>
<td>15</td>
<td>7</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Distal radius fracture</td>
<td>10</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Bone bruise scaphoid</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Bone bruise other carpal bone</td>
<td>8</td>
<td>6</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Bone bruise distal radius</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43</strong></td>
<td><strong>49</strong></td>
<td><strong>49</strong></td>
<td><strong>47</strong></td>
</tr>
</tbody>
</table>

There was moderate agreement with an average kappa of 0.57 (minimum 0.46, maximum 0.68) as shown in Table 7.3;

3. Another fracture:
There was substantial agreement with an average kappa of 0.61 (minimum 0.49, maximum 0.66).
Table 7.2
Cross tabulation showing the inter-observer variation (kappa coefficient) for 79 MR scans concerning a scaphoid fracture.

<table>
<thead>
<tr>
<th>Observer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>0.69</td>
<td>0.55</td>
<td>0.70</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>X</td>
<td>0.70</td>
<td>0.76</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>X</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 7.3
Cross tabulation showing the inter-observer variation (kappa coefficient) for 79 MR scans concerning the localisation of a scaphoid fracture.

<table>
<thead>
<tr>
<th>Observer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>0.58</td>
<td>0.46</td>
<td>0.54</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>X</td>
<td>0.60</td>
<td>0.68</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>X</td>
<td>0.56</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 7.4
Results at the first and second evaluation moment and the intra-observer variation (kappa coefficient) for a scaphoid fracture (38 patients).

<table>
<thead>
<tr>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
<th>Observer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation moment</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Scaphoid fracture</td>
<td>4</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>No scaphoid fracture</td>
<td>34</td>
<td>34</td>
<td>28</td>
</tr>
<tr>
<td>Simple kappa</td>
<td>1</td>
<td>0.69</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 7.5
Results at the first and second evaluation moment and the intra-observer variation (kappa coefficient) for the localisation of a scaphoid fracture (38 patients).

<table>
<thead>
<tr>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
<th>Observer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation moment</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Proximal 1/3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Middle 1/3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Distal 1/3</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Weighted kappa</td>
<td>0.94</td>
<td>0.60</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Intra-observer variation

The intra-observer variation was calculated for 38 patients. Concerning:

1. A scaphoid fracture:
   There was almost perfect agreement with an average kappa of 0.90 (minimum 0.69, maximum 1) as shown in Table 7.4;

2. The localisation of a scaphoid fracture:
   There was almost perfect agreement with an average kappa of 0.85 (minimum 0.60, maximum 0.97) as shown in Table 7.5;

3. Another fracture:
   There was substantial agreement with an average kappa of 0.71 (minimum 0.53, maximum 0.79).

Discussion

This study shows that the inter-observer variation of MR imaging has 1) substantial agreement for a scaphoid fracture, 2) moderate agreement for the localisation of a scaphoid fracture, and 3) substantial agreement for another fracture.

For the intra-observer variation of MR imaging, we found 1) almost perfect agreement for a scaphoid fracture, 2) almost perfect agreement for the localisation of a scaphoid fracture, and 3) substantial agreement for another fracture.

To our knowledge, there are no other reports evaluating the observer variation of MR imaging of (occult) scaphoid fractures in daily practice. There is an obvious need for a fast and reliable diagnostic procedure in case 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. MR imaging has been suggested as the modality of choice for detecting suspected scaphoid fractures.\textsuperscript{1,13,18-19,22-23}

But, the observer variation in daily practice should be assessed in order to define the place of MR imaging 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 most experience with MR imaging, 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 around a 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.

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 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 7.2.

In conclusion, the observer variation of MR imaging of suspected scaphoid fractures was low. The influence of expertise with MR imaging in daily practice should be taken into consideration. Observers with little experience with MR imaging 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.
Chapter 7
Observer variation of MR imaging