CHAPTER 4

RECOGNITION OF 10 DISTAL RADIAL FRACTURE TYPES BY RESIDENTS.


Departments of Surgery, University Hospital Groningen and Deventer Hospital
the Netherlands
Patients with a distal radial fracture are commonly treated by physicians with little experience. Since, many specific types of distal radial fractures have been described, and different therapeutic regimens can be chosen, there is a need for a classification method which is easy to handle. A search in Medline showed that at least thirteen classification systems have been reported since 1960. It seems, that no classification system has been proven to be superior, and a generally accepted frame of reference for inexperienced physicians to classify distal radial fractures is still lacking. Although, universal classification systems has been introduced, it appears not to be uncommon that fractures are usually provided with labels referring to either the first author describing the particular fracture type (synonyms), or referring to an injury mechanism (eponyms) (2).

Given the fact that a universal classification system is often not being used, a question to be answered is how the base-line of the clinician’s recognition of distal radial fractures can be established. Knowledge about the performance of recognizing the different fracture types may reflect the actual classification ability, and may serve as a starting point to develop useful programs to teach the essence of distal radial fractures (7). In addition, a strategy may be designed how modern, treatment-based classification systems might be introduced to clinicians (5). The purpose of the present study is to investigate the verbal and visual recognition of 10 commonly cited distal radial fractures (1,10).

MATERIAL AND METHODS

Participants were 30 surgical residents from five teaching hospitals, who had on the average 2 years clinical experience (range 1–4 years). The residents participated in a test in which they were asked to assess a series of 10 different distal radial fracture types. The series of fracture types was developed with an increasing level of complexity. Several simple fractures and some specific intra-articular fractures, fracture-dislocation, and combination fractures were included. The series of 10 fracture types consisted of respectively: Colles’, Smith’s, distal
forearm, a combination of radius & scaphoid, radial styloid process, dorsal Barton’s, volar Barton’s, pilon, chauffeur’s, and lunate load fracture (1, 4, 8, 10). The test consisted of two parts; a verbal and a visual part with 10 items each. Each verbal item consisted of a description of a distal radial fracture. At least two relevant distinctive features of the specific fracture type were used for the description of each verbal item (see Appendix). In each verbal item the subject was asked whether he recognized the particular fracture in the description. The 10 questions were printed on one sheat and each verbal statement was followed by a question about whether the clinician recognized the particular fracture in a “yes” or “no” format.

The 10 corresponding visual items contained an X-ray (AP and lateral projection) of each fracture type. Only those X-rays were included in which there was complete agreement between the authors about the type of fracture and the clearness of the X-ray. Each fracture type on an X-ray corresponded with the concomitant verbal item. The administration of the tasks was in a random order. The subjects were asked to write down the name (diagnosis) of the fracture type on a sheat. In case they didn’t know an exact label of the fracture type, they were asked to write down the relevant distinctive features of the particular fracture. Before test administration, two introductionary examples were given for each kind of test to accustom the subjects to the questions that were posed. The verbal test was administered at first, followed by the visual test. To avoid bias due to repeated measurement the verbal and visual recognition items were administered in a random counterbalanced order for each test (9). The administration of the 10 verbal and visual items together took about 45 minutes. For each verbal item affirmative responses (“yes”) were scored as 1, and “no” responses were scored as 0. For each visual item, a correct response (score 1) was defined as either a correct diagnostic label, or a description in which at least two distinct features from the corresponding fracture description were correctly used. A zero score was given in all other instances. The maximal total score per participant ranged from 0 to 10 for both the verbal and visual counterparts.

The percentage of positive (“yes”) answered verbal questions, and correctly diagnosed X-rays was computed per fracture type for the whole group of subjects.
By definition, the criterion of a score more than 80 percent correct responses per item for both verbal and visual recognition was chosen as a standard of adequately diagnosing the particular fracture (9). The binomial (Z) test was used to test the percentage of correct responses against the criterion of 80 percent for each fracture type. The Spearman rank correlation coefficient (Rs) was employed to show the extent of agreement between the rank order of scores in the verbal task and the rank order of scores in the visual task. For each individual fracture type the number of correct responses on the verbal items was tested against the number of correct responses on the visual counterpart with the McNemar test. In all tests, a p-value < 0.05 was considered as the level of significance.

RESULTS

VISUAL TASK

VERBAL TASK

*Figure 1.* Frequency distributions of correct answers in the visual task (left diagram), and affirmative responses ("yes") in the verbal recognition task (right diagram). The median score per participant was 3 (range 0-10) in the visual task (left diagram), and 6 (range 0-10) in the verbal task (right diagram). Legend: On the X-axis the total score of a subject is displayed. On the Y-axis the number of subjects with a particular total score is displayed.

The frequency distributions of the total scores per participant are shown in Figure 1. It can be seen that the median was at the score of 3 on the visual recognition task, whereas the median for the total score on the verbal recognition task was 6.
Although, the overall performance differed between the two types of tasks, the rankorder of the total scores per participant on the verbal task showed a very strong association with the rankorder of the total scores on the visual recognition task (Spearman rank correlation coefficient \(Rs = 0.91\)).

Figure 2 shows, that the mean score of verbal recognition (68% “yes”) was statistically significantly greater than the mean score of visual recognition (33% correct). Turning to the performance per fracture type, the highest scores on both verbal and visual recognition were observed in respectively: Colles’, Smith’s, and distal forearm fracture. The percentage of correct responses per fracture type was also tested against our criterion of more than 80% correct responses. Although, 8 verbal tasks did not differ statistically significantly on the 80% criterion (Z-test), only 3 corresponding items (Colles’, Smith’s, and distal forearm fracture) satisfied the criterion in the visual task.

According to the McNemar test, visual and verbal recognition were in accordance in 6 fracture types. However, a statistically significant difference was found in dorsal Barton’s, volar Barton’s, pilon, and radius & scaphoid fracture. Taken together, it may be concluded that in the group of 30 residents only Colles’, Smith’s, and distal forearm fracture met the 80% performance criterion. Almost no correct radiographic assessment was found in 5 particular fracture types (pilon fracture, radial styloid process, volar and dorsal Barton’s fracture, and radius & scaphoid fractures). In addition, the lowest performance on the corresponding verbal recognition task was found in chauffeur’s and lunate load fracture. Apparently, specific intra-articular and fracture-dislocations were the difficult fracture types to recognize. Finally, an inventory was made of the incorrect descriptions on the 10 visual items for all 30 residents. It was found that the total of 68% \((n=202)\) of responses was incorrect. A division could be made in either an incorrect diagnostic label, or an incomplete description. Remarkably, the labels Colles’ or Smith’s were used in 98 out of 202 incorrect diagnoses.

---

Figure 2. (see next page) Survey of the results of self-assessment of verbal knowledge, and the assessment of wrist radiographs (visual recognition). The data clearly show an empirical division in a group with more than 80% verbal and visual recognition, and a group of 7 fracture types with less adequate diagnostic performance. In the gray areas the observed measurement did not differ statistically significantly (Z-test) from the 80% performance criterion. Legend: The
maximum score per fracture item is 30.

<table>
<thead>
<tr>
<th>Fracture type</th>
<th>Visual recognition</th>
<th>Verbal recognition</th>
<th>McNemar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n correct (%)</td>
<td>n “yes” (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>per item</td>
<td>per item</td>
<td></td>
</tr>
<tr>
<td>Colles’</td>
<td>25 (83)</td>
<td>30 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>Distal forearm</td>
<td>25 (83)</td>
<td>25 (83)</td>
<td>NS</td>
</tr>
<tr>
<td>Smith’s</td>
<td>23 (77)</td>
<td>30 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>Radial styloid process</td>
<td>16 (53)</td>
<td>20 (66)</td>
<td>NS</td>
</tr>
<tr>
<td>Barton’s dorsal</td>
<td>4 (13)</td>
<td>20 (66)</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Barton’s volar</td>
<td>2 (7)</td>
<td>20 (66)</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Pilon</td>
<td>2 (7)</td>
<td>19 (63)</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Radius &amp; Scaphoid</td>
<td>1 (3)</td>
<td>22 (73)</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Chauffeur’s</td>
<td>0</td>
<td>9 (30)</td>
<td>NS</td>
</tr>
<tr>
<td>Lunate load</td>
<td>0</td>
<td>8 (27)</td>
<td>NS</td>
</tr>
</tbody>
</table>
DISCUSSION

This study showed that many residents have difficulties in recognizing the specific types of distal radial fractures. It appeared, that only Colles’, Smith’s, and distal forearm fracture were adequately recognized. It also occurred that the Colles’ and Smith's labels were most often used in incorrect visual assessments. In other words, the participants tend to label complex visual pictures in fracture types they already know. Most participants seemed to reduce the actual number of fracture types to 3 items in order to cope with 10 different distal radial fractures (6). Most likely, many subjects rely on a base-line knowledge which represents 3 fracture types. To our knowledge this observation has not been reported before, although it may be obvious to those involved in the management of wrist fractures and the training of residents. However, the results confirm the idea that a simple eponymous or synonymous classification system may not be the best strategy to deal with all distal radial fractures (2).

A discrepancy between verbal and visual recognition was found in all fracture types. The observed difference may be explained by two reasons. Firstly, it appeared that visual and verbal recognition were statistically significant different in respectively, dorsal Barton’s, volar Barton’s, pilon, and radius & scaphoid fracture. The high scores on verbal recognition in these four items may reflect a tendency to overestimate the skill to visually recognize a fracture type on the X-ray. Overestimation in inexperienced physicians has been reported before and is also known as “overconfidence bias” (3). Verbal performance was significantly greater than visual performance in the mentioned four fracture types, and it seemed that overconfidence bias influenced diagnostic performance. Secondly, an explanation for the observed visual recognition rate of 12% in 7 fracture types may be a lack of verbal knowledge about the relevant distinctive features of distal radial fractures (3,6). The relevant distinctive features may not be known by inexperienced
physicians, because definitions and descriptions are often not described explicitly in publications. It seems plausible, that recognition might be improved if relevant distinctive features of distal radial fractures are described more systematically in handbooks or at educational courses.

The drawback of simple classification systems with synonyms and eponyms, is that it is historical, not universal, and not treatment-based (7). The AO/ASIF's system for classification of fractures has been designed to overcome the above mentioned shortcomings, and may be a worthwhile tool for inexperienced physicians (5). The precise clinical value and the rate of agreement of the AO system for distal radial fractures has yet to be assessed. But, it has also been advocated by Mueller that the reliability of the AO system can only be improved if the relevant distinctive features of each fracture are emphasized even more better (5).

REFERENCES