Improved acetabular fracture diagnosis after training in a CT-based method

P. Jouffroy, A. Sebaaly, T. Aubert, G. Riouallon*
Service de chirurgie orthopédique et traumatologique, Centre hospitalier Paris Saint Joseph, 185 rue Raymond Losserand, 75014 Paris, France

**ARTICLE INFO**

**Abstract**

**Background:** Acetabular fractures remain challenging to diagnose, particularly when they are complex. An accurate diagnosis is nevertheless crucial to select the best surgical strategy. None of the training methods described to date relies on the Letournel classification with a detailed analysis of each abnormality seen by computed tomography (CT). We therefore prospectively assessed a CT-based diagnostic method by (1) determining the rate of correct diagnoses by orthopaedic surgeons before and after training in the method, (2) comparing the times needed to read the CT images before and after training, (3) and assessing the repeatability of the method.

**Hypothesis:** Training in the CT-based diagnostic method significantly increases the rate of correct diagnoses.

**Method:** The CT-based diagnostic method involves analysing eight anatomical landmarks in the anterior, posterior, and no man’s land zones. From our institutional database (450 cases between 2007 and 2016), we selected 35 acetabular fractures that replicated the overall distribution of fracture types. The images were reviewed by 10 inexperienced and 3 experienced readers before and after they received training in the CT-based diagnostic method. The rates of correct diagnoses and times needed to read the images were compared. Finally, an additional reading was performed to allow an assessment of reproducibility.

**Results:** After training, the rate of correct diagnoses by the inexperienced readers improved by 16.64% for all fractures combined (from 212/350, 60.5% [37–83%] to 270/350, 77.14% [63–86%]; P = 0.001) and by 25.0% for associated fractures (from 90/180, 50% [11–89%] to 114/140, 75.6% [61–90%]; P = 0.003). Mean time required by the inexperienced readers to interpret the 35 sets of images decreased after training, from 66.1 to 47.6 min (i.e., a 1.22-minute decrease per patient, P = 0.001). None of the study variables changed significantly after training of the experienced readers (P > 0.05). Reproducibility among the inexperienced readers was 0.78.

**Conclusion:** Analysing the eight anatomical landmarks located in the anterior, posterior, and no man’s land zones is a simple and reproducible method for diagnosing all fracture patterns defined by the Letournel classification.

**Level of evidence:** Level III, non-randomised prospective case–control diagnostic study.

© 2016 Elsevier Masson SAS. All rights reserved.

1. Introduction

The diagnosis of acetabular fractures is challenging, yet of the utmost importance to select the best surgical strategy [1,2]. The most widely used classification system was developed by Judet and Letournel, who differentiated ten types, five so-called elementary fractures and five associated fractures [3]. Letournel et al. reported that this classification system ensured the correct diagnosis in nearly 95% of cases [4,5].

Nevertheless, diagnostic errors remain common, particularly among inexperienced surgeons and radiologists [6,7], not only because acetabular fractures are rare, but also because their frequencies vary widely across fracture types and some fracture types may be confused, particularly by orthopaedic surgeons and radiologists with limited experience [8,9]. Several diagnostic methods have been developed to avoid confusing fracture types and to improve the diagnostic rate [9–13]. However, the complexity of the decision trees obtained with these methods is ill-suited to everyday practice. Furthermore, training of residents in a radiographic...
The diagnostic method failed to improve the diagnosis of acetabular fractures [14].

In our department, the diagnosis of acetabular fractures rests on computed tomography (CT) with 3D reconstructions. In previous studies, 3D CT was superior over conventional radiography [15–17]. We therefore conducted a prospective study to evaluate a 3D CT-based diagnostic method, to measure (1) the rate of correct diagnoses by physicians before and after training in this method, (2) the times needed to read the images before and after training, and (3) and repeatability of the method. Our hypothesis was that the 3D CT diagnostic method described here significantly improved the rate of correct diagnoses of acetabular fractures.

2. Material and methods

2.1. Patients

Two orthopaedic surgeons (PJ and GR) selected 35 cases from our institution’s database, which included 450 acetabular fractures managed between 2007 and 2016. The 35 cases were classified according to Letournel based on the radiographs, CTs (multiplanar reconstruction [MPR] and 3D reconstructions), and intra-operative findings. Cases for which the two surgeons disagreed about the fracture type were excluded.

The 35 fractures were distributed as follows: 14 elementary fractures (anterior wall, n = 2, 5.8%; anterior column, n = 3, 8.6%; posterior wall, n = 5, 14.2%; posterior column, n = 1, 2.8%; and transverse, n = 3, 8.6%), 18 associated fractures (posterior column with posterior wall, n = 2, 5.8%; transverse with posterior wall, n = 5, 14.2%; T-shaped, n = 2, 5.8%; anterior column with posterior hemi-transverse, n = 4, 11.4%; and both columns, n = 5, 14.2%). Two (5.8%) atypical fractures that did not fall in any of the Letournel fracture types, and 1 (2.8%) normal pelvis with no acetabular fracture. The distribution of the fracture types within the 35 fractures replicated that in the entire set of 450 fractures and was consistent with other reports [18].

2.2. Description of the method

The diagnostic method involves analysing eight anatomical landmarks on MPR and 3D CT images (Fig. 1). The fracture lines are examined carefully to determine which landmarks they involve. The evaluation starts with three anterior landmarks: the iliac wing, linea arcuata, and anterior wall of the acetabulum. Three landmarks located in the no man’s land zone are then examined, namely, the roof of the acetabulum, quadrilateral surface, and obturator ring. Finally, two posterior landmarks are assessed, the posterior border of the iliac bone and the posterior wall of the acetabulum. Table 1 shows which landmarks are involved in each of the ten fracture types described by Judet and Letournel.

2.3. Assessment methods

Among inexperienced and untrained medical students and residents who performed rotations in our department, 10 were selected at random to participate in this prospective study. The controls were 3 orthopaedic surgeons who were well beyond the learning curve for acetabular surgery (over 50 procedures for each) [6].

For each acetabular fracture, the CT image set included native axial slices, MPR reconstructions, and the following standardised 3D reconstructions: anterior and posterior views of the pelvis, an endopelvic view of the fractured hemi-pelvis, and an exopelvic view of the fractured hemi-pelvis after subtraction of the femoral head (Fig. 2).

Table 1

<table>
<thead>
<tr>
<th>Anatomical landmarks</th>
<th>Landmarks in the no man’s land zone</th>
<th>Posterior landmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iliac wing</td>
<td>Linea arcuata</td>
<td>Anterior wall</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

* Posterior wall of the acetabulum.
* Posterior border of the iliac bone.
When the readers interpreted the images for the first time, before training, they were given only the list of fracture types described by Letournel, with no instructions regarding image interpretation. The readers received training in the diagnostic method 2 weeks later. To minimise the risk of recognition bias, the second reading session occurred 2 weeks after training, i.e., 4 weeks after the first reading session, and the 35 cases were given in random order. The readers were also instructed to refrain from obtaining additional training between the two reading sessions and to base their diagnosis solely on the diagram of the Letournel classification. The total time spent on each of the two reading sessions was recorded. Finally, a third session was held 6 weeks after the second to allow an evaluation of reproducibility.

2.4. Statistical analysis

SPSS 18.0 software (International Business Machines Corporation, Armonk, NY) was used. Fisher’s test was applied to compare the proportion, percentage, standard deviation, and range of correct diagnoses between groups. The times needed for the reading sessions were compared using Student’s t test. We also computed the mean time needed to establish one correct diagnosis, as the total time/35 diagnostic rate, to look for an association between training and the time needed to make one correct diagnosis. The intra-class correlation coefficient (ICC) was computed to assess reproducibility then converted to the level of agreement according to Landis and Koch: 0.00–0.20, poor agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.0 near-perfect agreement [19]. Values of P lower than 0.05 were taken as indicating significant differences.

The smallest required sample size was computed as follows:

\[ n = \frac{p(1-p) \left( \frac{Z_{1-\alpha}^2 + Z_{1-\beta}^2}{\rho - p \bar{q}} \right)}{\bar{p} \bar{q}} \]

where \( p \) was the predicted diagnostic rate, \( \alpha \) the risk of Type I error, and \( \beta \) the risk of Type II error. The power of the test is then \( 1 - \beta \). For our model, power was 0.9. Thus, with \( \alpha = 0.05 \) and \( \beta = 0.1 \), at least 62 read cases were required. The 350 cases in our study therefore ensured high statistical power in order to detect significant differences.

3. Results

3.1. Rate of correct diagnoses

In the inexperienced readers, the overall diagnostic rate increased significantly after training in the diagnostic method, from 212/350 (60.5%; SD, 14; range, 37–83%) to 270/350 (77.14%; SD, 8; range, 63–86%), a difference of 16.86% \( (P = 0.001) \). We then conducted separate analyses of elementary fractures (anteroinferior, anteroposterior, posterior, anteroinferior and posteroinferior, and transverse) and associated fractures (posteroinferior and posteroinferior, transverse and posteroinferior, T-shaped, and posteroinferior and posteroinferior, and both columns). Training was not followed by a significant improvement in the diagnosis of elementary fractures (104/140, 74.3%, before training and 112/140, 81.4%, after training, a 7.1% increase; \( P > 0.05 \)). In contrast, for associated fractures, the diagnostic rate increased significantly, by 25.6% (90/180, 50%, before training and 132/180, 75.6%, after training; \( P = 0.003 \)) (Table 3).

None of the study variables differed significantly before and after training in the group of 3 experienced readers. After training, the diagnostic rate increased by 2% overall, 2.3% for elementary fractures, and 5.6% for associated fractures; none of these increases was statistically significant \( (P \) values, 0.6, 0.7, and 0.46, respectively). The diagnostic rate after training was not significantly different between the inexperienced and experienced groups for elementary fractures \( (P > 0.05) \). However, for associated fractures and all fractures pooled, the diagnostic rate remained significantly higher in the experienced group \( (P = 0.015 \) and \( P = 0.017 \), respectively) (Table 2). The most common error among inexperienced readers consisted in classifying both-column fractures as transverse or transverse plus posterior wall before training, and as anterior column plus posterior hemi-transverse after training. This last type of error was the most common error among experienced readers both before and after training.
Table 2
Diagnostic rate in each group.

<table>
<thead>
<tr>
<th></th>
<th>Before training</th>
<th>After training</th>
<th>Difference</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Percentage</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td><strong>Unexperienced readers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct diagnosis overall</td>
<td>206/350</td>
<td>60.5%</td>
<td>14%</td>
<td>37–83%</td>
</tr>
<tr>
<td>Correct diagnosis of elementary fractures</td>
<td>104/140</td>
<td>74.3%</td>
<td>11%</td>
<td>43–93%</td>
</tr>
<tr>
<td>Correct diagnosis of associated fractures</td>
<td>90/180</td>
<td>50%</td>
<td>23%</td>
<td>11–89%</td>
</tr>
<tr>
<td><strong>Experienced readers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct diagnosis overall</td>
<td>95/105</td>
<td>90.3%</td>
<td>6%</td>
<td>86–97%</td>
</tr>
<tr>
<td>Correct diagnosis of elementary fractures</td>
<td>37/42</td>
<td>88.1%</td>
<td>14%</td>
<td>71–100%</td>
</tr>
<tr>
<td>Correct diagnosis of associated fractures</td>
<td>49/54</td>
<td>90.7%</td>
<td>11%</td>
<td>78–100%</td>
</tr>
</tbody>
</table>

SD, standard deviation, bold characters indicate a significant difference.

Table 3
Comparison of the time needed to establish the diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>Before training</th>
<th>After training</th>
<th>Difference</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Percentage</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td><strong>Unexperienced readers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total reading time</td>
<td>66.6 ± 26</td>
<td>3' 10 ± 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time needed to make one correct diagnosis</td>
<td>3' 10 ± 1</td>
<td>1’48 ± 0.8’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experienced readers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total reading time</td>
<td>40.36 ± 11\’</td>
<td>39' ± 9’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time needed to make one correct diagnosis</td>
<td>1’16’ ± 0.43’</td>
<td>1’12’ ± 0.43’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bold characters indicate a significant difference.

3.2. Time required to read the CT images

Mean total time needed for the inexperienced readers to interpret the three CT image sets was 66.6 min before training and 47.36 min after training; the difference was of borderline significance (P = 0.09). The time needed by the inexperienced readers to achieve one correct diagnosis diminished significantly, from 3’10” before training to 1’48” after training (P = 0.001) (Table 2). In the experienced group, training was not followed by any significant decrease in either the total reading time (40.36” before training and 39” after training) or the time needed to achieve one correct diagnosis (1’16” before training and 1’12” after training, P = 0.84).

3.3. Reproducibility

Reproducibility of the diagnosis after training was substantial (ICC, 0.78) in the inexperienced group and near-perfect (ICC, 0.91) in the experienced group.

4. Discussion

Acetabular fractures are challenging to diagnose, although they are better understood since studies reported by Judet and Letournel [20–22], Potok et al. [23], however, described the Letournel classification as complicated and poorly understood, but also as being the only classification that provides guidance for selecting the surgical approach. However, this classification remains the most widely used. Our study demonstrates persisting diagnostic insufficiencies despite the use of CT with 3D reconstructions. Adding a simple diagnostic method (based on identifying eight anatomical landmarks) improved the rate of correct diagnoses both for all fractures pooled and for the more challenging associated fractures. In addition, the diagnostic method decreased the time needed to achieve a correct diagnosis. Finally, the diagnoses made after training were reproducible, with a substantial level of agreement.

A few limitations of our study should be acknowledged. First, the number of participants (n = 10) may seem small. Selection bias may have occurred, since the participants were all residents or medical students performing rotations in a tertiary referral department for acetabular fractures. However, the large number of cases analysed (n = 35) minimises the risk of bias related to the small sample size, and pooling residents and medical students mitigated any selection bias. Second, we selected the 35 cases from a database instead of enrolling consecutive patients. However, the selected cases included examples of all ten fracture types in the Letournel classification, in proportions that replicated our overall case-series and other large studies in the international literature [24].

The Letournel classification is the most widely used today [1]. However, in several studies done since its publication, the rate of correct diagnoses by orthopaedic surgeons remains low, even with 2D and 3D CT. At the 2009 SOFCOT round table, the mean diagnostic rate among participating centres was only 30% despite the use of 2D CT [25]. In early studies done at a time when CT image quality was poor, CT failed to contribute to the correct diagnosis of acetabular fractures [6,26]. The ability to establish the correct diagnosis depends heavily on image quality. O’Toole et al. [27] used 3D anterior, posterior, and three-quarters oblique views; whereas Schäffler et al. [28] used one sagittal, two coronal, and five axial CT views and found that 93% of readers felt a larger number of views was necessary and that 3D reconstructions probably improved the diagnostic rate. Hüfner et al. [7] reported that using standard radiographs alone resulted in very low diagnostic rates, which were not better than chance alone among residents; with 3D CT and femoral head subtraction, the results were similar to those of other studies. A consistent finding is that the diagnosis is improved by using CT, particularly with 3D reconstructions. These results have prompted attempts to develop diagnostic algorithms for analysing acetabular fractures. Thus, Petrisor et al. [29] reported that inter-observer reproducibility of the diagnosis by experts, although improved when a simplified system of six radiographic lines on standard antero-posterior radiographs was used, nevertheless remained limited, with an ICC of only 0.56. Prasartitha et al. [30] relied on 3D CT views with femoral head subtraction to simplify fracture analysis but focused solely on the direction of fracture lines involving the quadrilateral surface. A complex algorithm with more than 20 branches developed by Ly et al. [9] based on plain radiographs had only 50% to 59% diagnostic accuracy. Lawrence et al. [12] also reported a diagnostic algorithm, based on six landmarks identified on standard radiographs (ilio-ischial line, posterior wall, linea arcuata, anterior wall, iliac wing, and obturator ring).
However, this algorithm fails to consider several landmarks that are crucial to fracture classification, such as the roof of the acetabulum and the posterior border of the iliac bone. Furthermore, its reliance on the appearance of the fracture line on axial views makes it difficult to use [22]. In a study of orthopaedic residents, Polesello et al. [14] found that training in a method for interpreting radiographs of acetabular fractures failed to improve the diagnosis. Prevezes et al. [31] reported that using an algorithm to interpret standard radiographs improved the diagnostic rate by 22%. Similarly, in a study by Patel et al. [32], identifying specific radiological features significantly improved the diagnostic rate; however, their method is difficult to learn and involves a number of modifiers, which results in a low level of agreement (ICC, 0.24). Garrett et al. [33] demonstrated that adding 3D CT to standard radiographs increased the diagnostic rate to about 58% in fellowship-trained orthopaedic traumatologists, i.e., to the value noted in inexperienced readers in our study. These methods for improving the correct diagnosis appear too complex to be easily applied and fail to ensure the diagnosis of the full range of fracture types. Furthermore, they are not suited to all types of imaging studies. Our diagnostic method based on eight anatomical landmarks, in contrast, is simple, reproducible regardless of the type of reconstruction used (MPR or 3D), and effective in identifying all the fracture types in the Letournel classification. Its results provide guidance regarding surgical indications and selection of the approach and internal fixation method. Moreover, this method is a reproducible diagnostic tool. However, the final surgical decision requires examination of all the images, including the multiplanar and other 3D CT slices, as these are far more sensitive for diagnosing concomitant abnormalities such as intra-articular foreign bodies, impactions, the degree of comminution, and intermediate fragments.

5. Conclusion

This study demonstrates that analysing eight anatomical landmarks in three categories—anterior, posterior, and in the no man’s land zone—is a simple diagnostic method that improves the diagnosis of associated fractures by readers with limited experience, while diminishing the time needed to interpret the images. In addition, this method is reproducible for all fracture types in the Letournel classification. Its contribution is limited, however, among experienced readers. A multicentre study would be useful to assess the external validity of this diagnostic method.

Disclosure of interest

The authors declare that they have no competing interest.

References