Evaluation of Letournel and Judet classification of acetabular fracture with plain radiographs and three-dimensional computerized tomographic scan

Boonyarak Visutipol, Pornchai Chobtangsin, Bunyat Ketmalasiri, Narongchai Pattarabanjird and Namchai Varodompun
Institute of Orthopaedics, Lerdsin General Hospital, Bangkok, Thailand.

ABSTRACT

Letournel and Judet classification of acetabular fracture is widely used. The classification is based on the identification of fracture lines on plain radiographs. Three-dimensional CT scan was claimed to give a better view of the fracture line.

Our study showed that intraobserver reproducibility and interobserver reliability were almost the same when classification was done by using plain radiographs and 3D-CT scan. And 3D-CT scan did not increase either the interobserver reliability or the intraobserver reproducibility in classifying the fracture.

Key words: Letournel and Judet classification, acetabular fracture, three dimensional CT scan, intraobserver reproducibility, interobserver reliability

INTRODUCTION

Letournel and Judet classification (L-J classification) of the acetabular fracture is widely used (Fig. 1). Plain radiographs in AP, iliac and obturator oblique are commonly used to evaluate the fracture lines in each area of the pelvis and acetabulum. However, due to the complexity of the fracture, the accuracy of the diagnosis is in doubt.

It is claimed that two-dimensional computerized tomography (2D-CT scan) (Fig. 2) gives more detail of the fracture lines and might give more accuracy in classification. However, there are some difficulties because a lot of images have to be viewed. Modern three-dimensional computerized tomography (3D-CT scan) using a software program to create three-dimensional images from 2D-CT scan is said to give better details of the fracture lines and gives more accuracy in classifying the fracture (Fig. 3).

One way to measure the accuracy and efficiency of a classification is to measure the intraobserver reproducibility and interobserver reliability, which is the measurement of the agreement of one observer at different periods of time and among different observers, respectively.

To our knowledge, there is no study which compares the use of plain radiographs and 3D-CT scan to classify acetabular fracture in terms of intra and interobserver variation.
Figure 1  L-J classification which was divided into 10 subtypes.

Figure 2  2D-CT scan
The objectives of this study were:
1. To find the intraobserver reproducibility and interobserver reliability of L-J classification with the use of plain radiographs (AP, iliac and obturator oblique) and 3D-CT scan.
2. To evaluate if 3D-CT scan could increase the intraobserver reproducibility and interobserver reliability from the use of plain radiographs.

MATERIALS AND METHODS

20 sets of plain radiographs (AP, iliac and obturator oblique) from 20 patients who had fractures of the acetabulum were studied. The names, hospital number and date of X-ray were covered. The radiographs were marked with serial numbers from 1 to 20.

These radiographs were reviewed by 5 orthopaedic surgeons. Before the review, all the observers were supplied with information and diagrams of the L-J classification of the acetabulum. The sets of plain radiograph were reviewed twice with a 2-month interval. The serial numbers of the radiographs were changed before the second review.

The 3D-CT scans, which were created from 3 mm cut interval 2D-CT scans, from another set of 20 patients, were reviewed in the same manner as the plain radiographs.

Then intraobserver reproducibility and interobserver reliability were evaluated by using the Kappa test introduced by Cohen. The Kappa coefficient (K), which is the level of agreement corrected by proportion of agreement by chance, was calculated using an SPSS 7.5 software program and interpreted by using Landis and Koch guidelines: values of less than 0.00 indicate poor agreement; 0.00 to 0.20, slight agreement; 0.21 to 0.40, fair agreement; 0.41 to 0.60, moderate agreement; 0.61 to 0.80, substantial agreement; and more than 0.80, excellent agreement.

RESULTS

From our study, intraobserver reproducibility of the plain radiographs and 3D-CT scan were almost the same (0.42 for plain radiographs and 0.44 for 3D-CT scan), and both were in a moderate level of agreement (Table 1). Interobserver reliability in both groups were 0.24, which was in the fair level of agreement. Therefore 3D-CT scan did not help increasing the reliability (Table 2).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Intraobserver reproducibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa coefficients for intraobserver reproducibility</td>
<td></td>
</tr>
<tr>
<td>Observer</td>
<td>K plain radiograph</td>
</tr>
<tr>
<td>1</td>
<td>0.33</td>
</tr>
<tr>
<td>2</td>
<td>0.59</td>
</tr>
<tr>
<td>3</td>
<td>0.24</td>
</tr>
<tr>
<td>4</td>
<td>0.56</td>
</tr>
<tr>
<td>5</td>
<td>0.39</td>
</tr>
</tbody>
</table>

mean Kappa for plain radiographs = 0.42 (0.24 to 0.59)
mean Kappa for 3D-CT scan = 0.44 (0.30 to 0.59)

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Interobserver reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa coefficients for interobserver reliability</td>
<td></td>
</tr>
<tr>
<td>Plain radiography</td>
<td>mean K 0.24 (0.15 to 0.37)</td>
</tr>
<tr>
<td>3D-CT scan</td>
<td>mean K 0.24 (0.30 to 0.46)</td>
</tr>
</tbody>
</table>
DISCUSSION

Plain radiography in AP and both oblique views of the pelvis, as described by Judet et al.\(^8\), has long served as the standard for diagnosis of acetabular fractures. Modern computerized scan in 2 dimensions has been suggested to be superior in the detection of fracture involving the sacrum, quadrilateral surface, acetabular roof, posterior acetabular lip and the abnormality of the joint space.\(^6,12\) The disadvantage is that many images have to be viewed. 3D-CT scan, which can be rotated in different views and some parts can be erased for a better vision, was claimed to be more useful in characterizing complex displacement of the acetabulum.\(^1,2\)

Our study was not designed to evaluate the efficacy of the 3D-CT scan in identifying the fracture lines. The objective of our study was to evaluate the level of agreement of one observer and among different observers in classifying acetabular fracture using L-J classification.

From our study, moderate intraobserver reproducibility (K = 0.42) and fair interobserver reliability (K = 0.24) were observed when the classification was done using plain radiographs. When the classification was done using 3D-CT scan, the intraobserver reproducibility increased slightly but remained in the moderate level of agreement (K = 0.42) while the interobserver reliability remained the same (K = 0.24). We think the reasons why both the intraobserver reproducibility and interobserver reliability were low in L-J classification is because the classification is too complicated (there are 10 types) and too difficult to differentiate one type from the other especially in complex fracture groups. Though 3D-CT scan might have given a better detail of the fracture line, it did not help to improve either the intraobserver reproducibility or the interobserver reliability. This suggested that even when the fracture lines are more clearly viewed, the classification is so complicated that an accurate classification cannot be made.

Among the observers, observer number 4 who is specialized in trauma seems to have the highest Kappa both in plain radiographs and 3D-CT scan. This suggests that the level of experience might be an important factor in determining the L-J classification. The effect of experience on determining L-J classification needs further study. However, even with the most experienced surgeon, the Kappa coefficient was still in the moderate level of agreement which, again, reflects the difficulty of the classification.

The question is whether it is really necessary to differentiate one subtype from another in complex fracture groups when all of them require an extensive approach and the prognosis depends on the comminution and the degree of reduction. Thus, simplification of the classification may be helpful in improving the intraobserver reproducibility and interobserver reliability.

CONCLUSION

Intraobserver reproducibility and interobserver reliability were found to be low in L-J classification and 3D-CT scan did not seem to increase them. The results of this study suggest that Letournel and Judet classification might be too complicated and too difficult for the average orthopaedic surgeon.

REFERENCES

