Original Article

Do computed tomography and its 3D reconstruction increase the reproducibility of classifications of fractures of the proximal extremity of the humerus?∗,**

Thaís Matsushigue, Valmir Pagliaro Franco, Rafael Pierami∗, Marcel Jun Sugawara Tamaoki, Nicola Archetti Netto, Marcelo Hide Matsumoto

Universidade Federal de São Paulo, São Paulo, SP, Brazil

ARTICLE INFO

Article history:
Received 4 September 2012
Accepted 1 August 2013
Available online 27 March 2014

Keywords:
Fractures of the humerus/radiography
Fractures of the humerus/classification
Tomography

ABSTRACT

Objective: to determine whether 3D reconstruction images from computed tomography (CT) increase the inter and intraobserver agreement of the Neer and Arbeitsgemeinschaft für Osteosynthesefragen (AO) classification systems.

Methods: radiographic images and tomographic images with 3D reconstruction were obtained in three shoulder positions and were analyzed on two occasions by four independent observers.

Results: the radiographic evaluation demonstrated that using CT improved the inter and intraobserver agreement of the Neer classification. This was not seen with the AO classification, in which CT was only shown to increase the interobserver agreement.

Conclusion: use of 3D CT allows better evaluation of fractures with regard to their component parts and their displacements, but nevertheless the intraobserver agreement presented is less than ideal.

© 2014 Sociedade Brasileira de Ortopedia e Traumatologia. Published by Elsevier Editora Ltda. All rights reserved.

A tomografia computadorizada e sua reconstrução 3D aumentam a reprodutibilidade das classificações das fraturas da extremidade proximal do úmero?

RESUMO

Objetivo: determinar se as imagens da reconstrução 3D da tomografia computadorizada (TC) aumentam a concordância inter e intraobservador dos sistemas de classificação de Neer e Arbeitsgemeinschaft für Osteosynthesefragen (AO).

Palavras-chave:
Fraturas do úmero/radiografia
Fraturas do úmero/classificação
Tomografia


** Work performed in the Department of Orthopedics, Escola Paulista de Medicina, Universidade Federal de São Paulo.

* Corresponding author.

E-mail: rpierrami@gmail.com (R. Pierami).

2255-4971/$ – see front matter © 2014 Sociedade Brasileira de Ortopedia e Traumatologia. Published by Elsevier Editora Ltda. All rights reserved.

http://dx.doi.org/10.1016/j.rboe.2014.03.016
Introduction

Fractures of the proximal extremity of the humerus have an incidence of approximately 63–105 per 100,000 per year and account for 5% of all injuries to the appendicular skeleton. Their incidence is low among individuals under the age of 40 years and increases exponentially after this age. There is greater prevalence of these fractures among women, and female cases account for around 70–80%. The characteristics of the fracture (line, location, joint involvement, comminution and degree of soft-tissue involvement) are directly related to the force of the trauma, position of the limb at the time of the trauma and bone quality. Several classification systems have been developed in an attempt to guide treatments and compare results. For a classification system to be considered good, it needs to be validated, reliable and reproducible, as well as guiding the treatment, predicting possible complications and providing a prognosis. Furthermore, it should function as a mechanism for comparing the results obtained from the different types of treatment.

The classification method of the Arbeitsgemeinschaft für Osteosynthese/Association for the Study of Internal Fixation (AO/ASIF), which was created in 1986 and revised in 1990, uses an alphanumeric system for dividing fractures of the proximal extremity of the humerus into 27 subgroups. Three basic types of injury are taken into consideration in this classification method: extra-articular single-focus fractures, extra-articular bifocal fractures and joint fractures. The three groups are organized in increasing order of complexity and treatment difficulty and according to the prognosis. This is one of the most complete classification systems, but its intra- and interobserver reproducibility have been shown to be problematic with regard to the divisions between groups and subgroups.

Neer used the parts defined by Codman to propose the classification system that today is most frequently used. The four parts of the proximal extremity of the humerus defined in this classification method are the greater tuberosity, lesser tuberosity, diaphysis of the humerus and humeral head. For these parts to be considered to be fractured, the fragment should have a displacement greater than 1 cm or 45°, except for the greater tuberosity, which is considered to be a fractured part if there is a displacement greater than 0.5 cm or an angle of 45°. Thus, the fractures can be classified as affecting one, two, three or four parts. One criticism of Neer’s classification is that it does not consider the possibility of glenohumeral dislocation associated with the fracture, whereas this is encompassed in the AO classification.

Materials and methods

This study was submitted for appraisal by the Research Ethics Committee of the Federal University of São Paulo (UNIFESP) under the number 0212/11, on February 24, 2011, and was approved.

A retrospective analysis was conducted on all the patients with a diagnosis of fracturing of the proximal extremity of the humerus attended in the Shoulder and Elbow Surgery Sector of Hospital São Paulo, UNIFESP, between August 2009 and April 2012.

Seventy-two consecutive patients were selected, of whom 37 could be included in the study. The images selected were analyzed according to the views obtained (anteroposterior, lateral scapular and axillary), the total number of views, the quality of the radiographs and the use of CT with 3D reconstruction at the time of the trauma. Radiographs were excluded if their quality was poor, if the views needed for the study were not available or if no CT images with 3D reconstruction were available. Furthermore, all patients who had undergone previous surgical procedures on the limb under examination, those who had previously had fractures in the...
region studied and those presenting fractures that were considered to be pathological were also excluded.

The images were analyzed by four independent observers: one third-year resident in orthopedics and traumatology (C); one trainee orthopedist in the shoulder and elbow sector (B); and two orthopedics and traumatology specialists in the field of shoulder and elbow surgery (A and D). These observers independently classified the fractures in accordance with the AO/ASIF and Neer classification methods, by means of analysis on images that had previously been digitized. These analyses were done twice, at two separate times with a one-week interval between them. At the two evaluations, the images were randomized into different sequences so as to avoid bias.

The data were gathered and subjected to statistical analysis. The kappa coefficient (κ) was obtained in order to determine the inter- and intraobserver concordance of the classifications. The kappa values varied from −1 to +1; values between −1 and 0 indicated that the concordance was less than expected and attributable purely to chance; values of 0 indicated that the concordance with similar to chance; and values of +1 indicated total concordance. Generally, values of 0.5 are considered to be unsatisfactory, values between 0.5 and 0.75 are satisfactory and appropriate and values greater than 0.75 are excellent.17

**Results**

The method with greatest interobserver concordance was the Neer classification using CT (κ = 0.57). CT provided greater interobserver concordance in both classifications (Table 1).

In relation to intraobserver concordance, no increase was observed through using CT applied to the AO classification (κ = 0.39 for radiography and κ = 0.33 for CT). However, for the Neer classification, there was an increase in this concordance (κ = 0.45 for radiography and κ = 0.56 for CT) (Table 2).

<table>
<thead>
<tr>
<th>Examination</th>
<th>Classification</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiography</td>
<td>Neer</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>AO</td>
<td>0.25</td>
</tr>
<tr>
<td>Tomography</td>
<td>Neer</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>AO</td>
<td>0.36</td>
</tr>
</tbody>
</table>

**Table 2 – Mean intraobserver concordance for the Neer and AO classification systems when CT and X-ray examinations were used.**

<table>
<thead>
<tr>
<th>Examination</th>
<th>Classification</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiography</td>
<td>Neer</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>AO</td>
<td>0.39</td>
</tr>
<tr>
<td>Tomography</td>
<td>Neer</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>AO</td>
<td>0.33</td>
</tr>
</tbody>
</table>

**Discussion**

Over recent decades, with the introduction of new technologies for diagnosing fractures of the proximal extremity of the humerus, it has been asked whether using CT with three-dimensional reconstruction (3D CT) might provide benefits with regard to identifying the fracture pattern and guiding the type of treatment to use.11–16 However, CT is not a harmless method, because it exposes the patient to a high dose of radiation. Moreover, the cost of this method is still much higher than that of simple radiography. For these reasons, new studies should be conducted to define the usefulness and possible indications for using CT.

In our study, the interobserver concordance using the Neer classification by means of radiographs was unsatisfactory (κ = 0.37). This finding is in agreement with other published studies that had this aim.12,13,18 When the evaluation was done using CT, the interobserver concordance became satisfactory (κ = 0.57), as also seen in other studies already published,12,13 which justifies the use of CT. In relation to the AO classification, the values found using radiography and CT were considered to be unsatisfactory, even though there was an increase in the kappa value (κ = 0.25 for radiography and κ = 0.36 for CT). These unsatisfactory values can perhaps be explained by the complexity of the classification system. The same is observed when the system is used to classify fractures in other segments, such as the distal extremity of the radius, the ankle or the femoral neck.19–22

In relation to interobserver concordance, CT was shown to be useful when the Neer classification was used, and produced a satisfactory value (κ = 0.56). This has also been shown by other studies.11,12,18 For the AO classification, CT was not shown to be useful and led to a decrease in the kappa value (from 0.39 to 0.33), which perhaps can be explained by the complexity of the classification system, as discussed earlier.

Thus, our study demonstrated that the Neer classification was more reproducible when CT with 3D reconstruction was used, which therefore justifies its use in classifying fractures of the proximal extremity of the humerus, which is what it is most used for in our setting. Nonetheless, in absolute values, the reproducibility still remains low. This was not observed when the AO classification was used.

One limitation of this study was the low number of cases evaluated, which may have led to bias. Furthermore, because this study was conducted at a reference service for trauma, it is likely that there were greater numbers of complex cases among the sample, which would lead to lower intra- and interobserver concordance, considering that in cases of fractures of the proximal extremity of the humerus of greater complexity, it becomes more difficult to measure the angular deviation, displacement and impaction. These characteristics of the fracture, which form part of the criteria for the classification, have been indicated to be the factors that cause low reproducibility of the classifications. It is important to emphasize that the present study only aimed to investigate the reproducibility of the classification systems using conventional radiographs and CT with 3D reconstruction. This was not an accuracy study, nor did it assess the merit of these examinations regarding surgical indications.
Conclusion

CT with 3D reconstruction improved the intra- and interobserver concordance for the Neer classification method. This was not observed for the AO classification system, in which only interobserver concordance was seen to improve with the use of CT with 3D reconstruction.

Conflicts of interest

The authors declare no conflicts of interest.

REFERENCES