CrossMark

ORIGINAL ARTICLE • SHOULDER - TRAUMA

Analysis of four-fragment fractures of the proximal humerus: the interest of 2D and 3D imagery and inter- and intra-observer reproducibility

X. Ohl[1] · P. Mangin[1] · C. Barbe[2] · V. Brun[1] · C. Nerot[1] · F. Sirveaux[3]

Abstract The treatment of the four-part fractures of the proximal humerus remains a therapeutic challenge. The decision-making is based on preoperative criteria concerning the displacement of the fracture, the stability of the fracture and the risk of avascular necrosis of the humeral head. The aim of this study was to analyse the inter- and intra-observer reproducibility of those criteria previously described by Hertel. Three observers analysed three times 20 radiologic files comprising 2D X-rays, 2D CT scan and 3D reconstructions for the intra-observer study, and an expert committee was used to assess the inter-observer reproducibility. The Kappa coefficient was used to measure agreement. The Kappa coefficient founded poor to moderate agreement for the majority of the criteria after the 2D X-ray analysis. This coefficient was improved with the use of 2D CT scan and 3D reconstructions, in particular for the medial hinge assessment, the humeral head fracture and the metaphyseal extension. The reproducibility of the criteria described by Hertel on 2D X-rays is at least moderate. Reproducibility could be considerably improved by associating 2D scans and 3D reconstruction, in particular for the criteria related to prognosis for the vascularisation of the humeral head.

✉ X. Ohl
xohl@chu-reims.fr

1   Department of Orthopaedic Surgery, Reims Teaching
     Hospital, CHU Maison-Blanche, 45 rue Cognacq-Jay,
     51092 Reims, France

2   Research Methodology Unit, Reims Teaching Hospital, CHU
     Robert Debré, Avenue du Général Koenig, 51092 Reims,
     France

3   Department of Orthopaedic Surgery, Centre Chirurgical
     Emile Gallé, CHU de Nancy, 49 rue Hermitte, 54000 Nancy,
     France

Introduction

Fractures of the proximal humerus are increasingly frequent, with numbers tripling between the 1970s and the 2000s [1, 2]. Among these fractures, those involving the tuberosities and also the anatomic neck are a therapeutic challenge. For this type of fracture (Neer's four-part fracture) the degree of displacement of the fracture needs to be understood in order to provide suitable treatment and apprehend the risks in its evolution. There is indeed, in this type of fracture, a risk of humeral head ischaemia, which will carry considerable weight in the therapy adopted.

The usual classifications, such as the AO or the Neer classification, have shown their limitations in terms of reproducibility [3, 4] and are not suitable for the prognostic assessment of these four-fragment fractures of the proximal humerus. The radiographic parameters described by Hertel in 2004, on the other hand, seem to us to be far more relevant to routine clinical practice [5]. They enable an analysis of the fracture and an assessment of the risk of post-traumatic necrosis of the humeral head. They also enable therapeutic care to be apprehended more efficiently, in particular via the assessment of the medial cortex (calcar) which is an important stability criterion for osteosynthesis [6–8]. These parameters were initially described on standard radiographs, but they are fully operative for reading 2D scans or 3D reconstructions.

The use of the scanner to improve reproducibility of the classification of these proximal humerus fractures is still controversial [9–11]. In case of complex fractures, however, the scanner is most often the rule to guide therapeutic

🖄 Springer

strategy, although the reproducibility of the different assessment criteria has never been studied.

The aims of the present study were firstly to analyse inter-observer and intra-observer reproducibility for the different criteria proposed by Hertel, using three types of imagery (standard radiographies, 2D and 3D scans) and secondly to assess the relevance of the use of the scanner to improve reproducibility.

Material and method

Twenty radiographic files derived from the 2014 SOFCOT symposium on four-part fractures of the proximal humerus were chosen randomly among the 384 complete files. This multi-centre study involved 11 centres specialised in surgery of the shoulder, and it obtained approval from CPP-Est (2013-A00050-36). The radiological files all contained standard radiographs performed in emergency settings when the patient arrived in the care facility, a 2D scan with axial, sagittal and coronal sections, and high-resolution 3D reconstructions.

Analysis protocol

Three independent observers (PM, VB and XO) with differing experience assessed each radiographic file three times for the implementation of a study of intra-rater reproducibility. An expert committee (CN and FS) convened for the analysis of the radiological files. Inter-rater reproducibility was thus assessed by comparing the first assessment of each observer with the expert committee assessment. The assessment of the different files was always conducted in the same manner: analysis of the standard radiographies, then analysis of the 2D scan sections and finally analysis of the 3D reconstructions. An interval of at least two weeks was allowed between each round of assessments. A computerised document presenting the different parameters and explaining them was issued to each participant before the start of the study.

The eight criteria assessed were:

- Displacement of the humeral head on the frontal plane: not displaced/varus/valgus.
- Displacement of the humeral head on the sagittal plane: not displaced/angle $<20°$/angle $>20°$.
- Humeral head split: yes/no.
- Calcar comminution: yes/no.
- Medial hinge: yes/no.
- Length of metaphyseal extension: <8 mm/>8 mm.
- Displacement of the greater tuberosity: not displaced/ >5 mm/<5 mm.
- Displacement of the lesser tuberosity: not displaced/ >5 mm/<5 mm.

Statistical analysis

Intra-rater reproducibility was assessed using Cohen's Kappa coefficient across three series of measures. Inter-rater reproducibility was also assessed using the Kappa coefficient comparing rater series 1 and the expert committee assessment. We thus assessed the extent of the influence of the types of imagery on concordance with expert opinion.

The Kappa coefficient is a tool designed to measure agreement between two qualitative variables based on the same terms. The Kappa coefficient values are always between 0 (no agreement) and 1 (perfect agreement). A Kappa coefficient comprised between 0.00 and 0.20 significate a very poor agreement, between 0.21 and 0.40 significate a poor agreement, between 0.41 and 0.60 significate a moderate agreement, between 0.61 and 0.80 significate a satisfactory agreement and between 0.81 and 1.00 significate an excellent agreement.

The analyses were performed on SAS software (version 9.3, SAS Institute Inc., Cary, NC, USA) and SPAD (version 8.2, Société Coheris, Suresnes, France).

Results

Tables 1, 2 and 3 describe the results for intra-observer reproducibility for the analysis of standard radiographs, 2D scan sections and 3D scan reconstructions, respectively.

The overall analysis of intra-observer reproducibility for the criteria studied reached poor to moderate agreement in 13 out of 24 cases (8 criteria and 3 observers)—Kappa values ranging from 0.21 to 0.60—when the analyses were performed on standard radiographs. The number was 6 out of 24 (25%) for the 2D scans and 5 out of 24 (21%) for the 3D reconstructions.

In the other instances, the Kappa coefficient was satisfactory (8 for the standard radiographs, 11 for the 2D scans and 13 for the 3D reconstructions) or excellent (3 for the standard radiographs, 7 for the 2D scans and 6 for the 3D reconstructions.

Agreement between the observers and the expert committee was studied via inter-rater reproducibility (Table 4). The Kappa coefficient increased with imagery type in formal manner for the criteria median hinge and length of metaphyseal extension. For the other criteria, the influence of the type of imagery on rater/expert agreement was less marked. The criterion calcar comminution was very poorly reproducible between raters 2 and 3 and the expert committee, while rater 1 reached Kappa coefficients between 0.29 and 0.60.

Table 1 Intra-rater reproducibility after analysis using standard radiographies

	Observer 1	Observer 2	Observer 3
Displacement of the humeral head, frontal plane	0.85 [0.63; 1.07]	0.61 [0.41; 0.81]	0.79 [0.59; 0.99]
Displacement of the humeral head, sagittal plane	0.79 [0.61; 0.97]	0.43 [0.25; 0.61]	0.74 [0.56; 0.92]
Fracture of the humeral head	0.61 [0.35; 0.87]	0.61 [0.35; 0.87]	0.85 [0.59; 1.11]
Calcar comminution	0.53 [0.27; 0.79]	0.64 [0.38; 0.90]	0.41 [0.15; 0.65]
Existence of a medial hinge	0.39 [0.13; 0.65]	0.60 [0.34; 0.86]	0.85 [0.59; 1.11]
Metaphyseal extension	0.58 [0.32; 0.84]	0.39 [0.13; 0.65]	0.56 [0.30; 0.82]
Displacement of the greater tuberosity	0.62 [0.42; 0.82]	0.54 [0.30; 0.78]	0.51 [0.31; 0.71]
Displacement of the lesser tuberosity	0.47 [0.27; 0.67]	0.48 [0.30; 0.66]	0.32 [0.14; 0.50]

Kappa coefficient and 95% CI

Table 2 Intra-rater reproducibility after analysis of 2D scans

	Observer 1	Observer 2	Observer 3
Displacement of the humeral head, frontal plane	0.66 [0.44; 0.88]	0.64 [0.44; 0.84]	0.67 [0.49; 0.85]
Displacement of the humeral head, sagittal plane	0.85 [0.67; 1.03]	0.66 [0.48; 0.84]	0.90 [0.72; 1.08]
Fracture of the humeral head	0.70 [0.44; 0.96]	0.82 [0.56; 1.08]	0.78 [0.52; 1.04]
Calcar comminution	0.59 [0.33; 0.85]	0.51 [0.25; 0.77]	0.33 [0.07; 0.53]
Existence of a medial hinge	0.80 [0.54; 1.06]	0.78 [0.52; 1.04]	0.77 [0.51; 1.03]
Metaphyseal extension	0.91 [0.65; 1.17]	0,93 [0.67; 1.19]	0.77 [0.51; 1.03]
Displacement of the greater tuberosity	0.59 [0.39; 0.79]	0.84 [0.58; 1.11]	0.52 [0.34; 0.70]
Displacement of the lesser tuberosity	0.88 [0.66; 1.11]	0.48 [0.30; 0.66]	0.61 [0.41; 0.81]

Kappa coefficient and 95% CI

Table 3 Intra-rater reproducibility after analysis of 3D reconstructions

	Observer 1	Observer 2	Observer 3
Displacement of the humeral head, frontal plane	0.81 [0.57; 1.05]	0.61 [0.41; 0.81]	0.64 [0.44; 0.84]
Displacement of the humeral head, sagittal plane	0.89 [0.69; 1.09]	0.66 [0.48; 0.84]	0.80 [0.62; 0.98]
Fracture of the humeral head	0.72 [0.46; 0.98]	0.72 [0.46; 0.98]	0.82 [0.58; 1.08]
Calcar comminution	0.80 [0.54; 1.06]	0.63 [0.37; 0.89]	0.52 [0.26; 0.78]
Existence of a medial hinge	0.73 [0.47; 0.99]	0.72 [0.46; 0.98]	0.85 [0.59; 1.11]
Metaphyseal extension	0.91 [0.65; 1.17]	0.86 [0.60; 1.12]	0.68 [0.42; 0.94]
Displacement of the greater tuberosity	0.64 [0.42; 0.86]	0.66 [0.40; 0.92]	0.26 [0.04; 0.48]
Displacement of the lesser tuberosity	0.51 [0.29; 0.73]	0.47 [0.29; 0.65]	0.29 [0.11; 0.47]

Kappa coefficient and 95% CI

Discussion

Intra-rater reproducibility on the Hertel criteria was moderate for assessments based on standard radiographies, and it improved with the use of 2D scans and 3D reconstructions. Inter-rater reproducibility was better for several criteria in case of use of 2D and 3D scans compared to standard radiographies.

Numerous authors have shown that the most widely used international classifications (Neer, AO, Codman, Duparc, AST, etc.) were not very reproducible, with at best only moderate inter-rater reproducibility (Kappa values

between 0.41 and 0.60) [4, 10–13]. The main value of these classifications resides in the scope for comparing clinical series one to the other. Thus, all the fractures in the present series are type 12 fractures for Codman-Hertel, 11-C1 or 11-C2 for AO and four-part fracture for Neer.

The interest of the Hertel criteria is their prognostic and therapeutic value. Indeed, certain criteria, such as displacement of the greater tuberosity or humeral head angle greater than 20° in the sagittal plane will orient towards surgical treatment. Other criteria, such as medial hinge or absence of calcar comminution, provide information on the stability of the fracture [5, 6, 8]. Finally, parameters such as

Table 4 Inter-rater reproducibility between first assessment of each rater and expert opinion

	Observer 1	Observer 2	Observer 3
Displacement of the humeral head frontal plane			
Rx std	0.65 [0.36; 0.94]	0.60 [0.31; 0.88]	0.55 [0.25; 0.86]
2D scans	0.41 [0.10; 0.73]	0.63 [0.33; 0.92]	0.45 [0.12; 0.78]
3D reconstructions	0.57 [0.28; 0.85]	0.57 [0.28; 0.85]	0.46 [0.17; 0.75]
Displacement of the humeral head sagittal plane			
Rx std	0.27 [−0.08; 0.61]	0.24 [0.07; 0.55]	0.16 [−0.14; 0.46]
2D scans	0.16 [−0.33; 0.36]	0.16 [−0.14; 0.46]	0.10 [−0.20; 0.40]
3D reconstructions	0.60 [0.24; 0.96]	0.57 [0.20; 0.93]	0.33 [0.04; 0.63]
Fracture humeral head			
Rx std	0.32 [−0.08; 0.71]	0.62 [0.15; 1.00]	0.47 [0.04; 0.90]
2D scans	0.47 [0.04; 0.90]	1.00 [1.00; 1.00]	0.63 [0.24; 1.00]
3D reconstructions	0.57 [0.20; 0.93]	0.88 [0.64; 1.00]	0.52 [0.12; 0.93]
Calcar comminution			
Rx std	0.29 [−0.14; 0.71]	0.14 [−0.20; 0.48]	−0.15 [−0.55; 0.26]
2D scans	0.39 [−0.03; 0.81]	0.18 [−0.14; 0.49]	0.06 [−0.37; 0.49]
3D reconstructions	0.60 [0.24; 0.95]	0.12 [−0.10; 0.35]	0.19 [−0.24; 0.62]
Medial hinge			
Rx std	0.25 [0.01; 0.49]	0.29 [−0.14; 0.73]	0.30 [0.04; 0.57]
2D scans	0.69 [0.38; 1.00]	0.69 [0.38; 1.00]	0.43 [0.11; 0.74]
3D reconstructions	0.71 [0.41; 0.99]	0.69 [0.38; 1.00]	0.53 [0.21; 0.84]
Metaphyseal extension			
Rx std	0.17 [−0.29; 0.62]	0.17 [−0.29; 0.62]	0.41 [0.01; 0.83]
2D scans	0.88 [0.64; 1.00]	0.78 [0.50; 1.00]	0.66 [0.31; 1.00]
3D reconstructions	0.86 [0.59; 1.00]	0.69 [0.38; 1.00]	0.63 [0.28; 0.99]
Displacement of the greater tuberosity			
Rx std	0.30 [−0.03; 0.64]	0.16 [−0.09; 0.42]	0.17 [−0.08; 0.41]
2D scans	0.36 [0.04; 0.69]	0.43 [0.11; 0.74]	0.41 [0.15; 0.68]
3D reconstructions	0.30 [−0.03; 0.62]	0.56 [0.22; 0.91]	0.39 [0.03; 0.78]
Displacement of the lesser tuberosity			
Rx std	0.41 [0.12; 0.70]	0.36 [0.01; 0.71]	0.22 [−0.02; 0.46]
2D scans	0.2 [−0.03; 0.67]	0.32 [−0.01; 0.65]	0.33 [0.01; 0.66]
3D reconstructions	0.31 [0.01; 0.61]	0.40 [0.14; 0.67]	0.29 [−0.14; 0.73]

Kappa coefficient and 95% CI

fracture of the humeral head or metaphyseal extension under 8 mm are risk factors for osteonecrosis of the humeral head [5].

In assessments using standard radiographs, certain criteria exhibit moderate to good intra-observer reproducibility, for instance the greater tuberosity criterion, displacement of the humeral head in the frontal plane or fracture of the humeral head. These criteria are indeed fairly easy to interpret on standard radiographs such as those performed in emergency setting. In contrast, intra-rater reproducibility was not as good for the assessment of the medial hinge (observer 1), metaphyseal extension (observer 2) or calcar comminution (observer 3). The assessment of the lesser tuberosity on radiographs was in all cases difficult and poorly reproducible. For inter-rater reproducibility, it was poor to moderate in all cases for radiographic analyses, with the exception of displacement of the humeral head in the frontal plane (Kappa range 0.55–0.65).

The use of 2D and 3D scans improved intra-observer reproducibility for most of the criteria considered for all three raters, with satisfactory to excellent Kappa coefficients for five of the eight criteria. The three criteria that were more problematic were the assessments relating to the greater and lesser tuberosities and calcar comminution.

This last criterion was not initially described by Hertel, but consideration of this parameter developed with the use of locking plates. It is in fact an important criterion for the

assessment of the stability of the fracture, in particular when plate osteosynthesis is envisaged. In 2012 Osterhoff [8] defined calcar comminution as being present if there was an intermediate fragment on the medial curve of the humeral metaphysis below the anatomic neck. The rather uncertain definition of this criterion could well explain the very poor intra-rater reproducibility whatever the type of imagery used. The intra-rater reproducibility for this criterion, however, showed improvement in the assessment of 3D reconstructions, since this type of imagery enables the medial cortical to be viewed as a whole and thus the detection of any intermediate fragment. This criterion is thus worth studying, but requires a better definition to gain in reproducibility.

The analysis of lesser and greater tuberosity displacement is moderately reproducible on 2D and 2D imagery both intra- and inter-rater. This can certainly be explained by the lack of precision in the distinction "not displaced" and "displacement under 5 mm". Indeed, in case of a (very minor) displacement of 1 or 2 mm, the raters may have assessed differently without actually misinterpreting. To improve reproducibility for these parameters, the following response options could be suggested: not fractured/non-displaced fracture (<5 mm)/displaced fracture (>5 mm).

The value of 2D scans and 3D reconstructions for the assessment of proximal humerus fractures is still a subject of debate [9–11]. However, for complex fractures where treatment can be guided by precise criteria such as metaphyseal extension or the presence of a medial hinge, we have seen that these scans are very useful. They improve inter-rater assessment reproducibility for most of the criteria that are relevant to therapeutic decisions.

Conclusion

The criteria we have retained for four-fragment complex fractures of the proximal humerus exhibit intra- and inter-rater reproducibility that is at least moderate. Reproducibility could be considerably improved by associating 2D scans and 3D reconstruction, in particular for the criteria related to prognosis for the vascularisation of the humeral head.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest with this study and nothing to disclose.

References

1. Palvanen M, Kannus P, Niemi S, Parkkari J (2006) Update in the epidemiology of proximal humeral fractures. Clin Orthop Relat Res 442:87–92
2. Resch H (2011) Proximal humeral fractures: current controversies. J Shoulder Elbow Surg 20(5):827–832. doi:10.1016/j.jse.2011.01.009
3. Maier D, Jaeger M, Izadpanah K, Strohm PC, Suedkamp NP (2014) Proximal humeral fracture treatment in adults. J Bone Joint Surg Am 96(3):251–261. doi:10.2106/jbjs.l.01293
4. Majed A, Macleod I, Bull AMJ, Zyto K, Resch H, Hertel R, Reilly P, Emery RJH (2011) Proximal humeral fracture classification systems revisited. J Shoulder Elbow Surg 20(7):1125–1132. doi:10.1016/j.jse.2011.01.020
5. Hertel R, Hempfing A, Stiehler M, Leunig M (2004) Predictors of humeral head ischemia after intracapsular fracture of the proximal humerus. J Shoulder Elbow Surg 13(4):427–433. doi:10.1016/j.jse.2004.01.034
6. Gardner MJ, Weil Y, Barker JU, Kelly BT, Helfet DL, Lorich DG (2007) The importance of medial support in locked plating of proximal humerus fractures. J Orthop Trauma 21(3):185–191. doi:10.1097/BOT.0b013e3180333094
7. Lescheid J, Zdero R, Shah S, Kuzyk PR, Schemitsch EH (2010) The biomechanics of locked plating for repairing proximal humerus fractures with or without medial cortical support. J Trauma 69(5):1235–1242. doi:10.1097/TA.0b013e3181beed96
8. Osterhoff G, Hoch A, Wanner GA, Simmen HP, Werner CM (2012) Calcar comminution as prognostic factor of clinical outcome after locking plate fixation of proximal humeral fractures. Injury 43(10):1651–1656. doi:10.1016/j.injury.2012.04.015
9. Bahrs C, Rolauffs B, Südkamp NP, Schmal H, Eingartner C, Dietz K, Pereira PL, Weise K, Lingenfelter E, Helwig P (2009) Indications for computed tomography (CT-) diagnostics in proximal humeral fractures: a comparative study of plain radiography and computed tomography. BMC Musculoskeletal Disord 10(1):1–7. doi:10.1186/1471-2474-10-33
10. Matsushigue T, Pagliaro Franco V, Pierami R, Jun Sugawara Tamaoki M, Archetti Netto N, Hide Matsumoto M (2014) Do computed tomography and its 3D reconstruction increase the reproducibility of classifications of fractures of the proximal extremity of the humerus? Rev Bras Ortop 49(2):174–177. doi:10.1016/j.rboe.2014.03.016
11. Sjoden GO, Movin T, Aspelin P, Guntner P, Shalabi A (1999) 3D-radiographic analysis does not improve the Neer and AO classifications of proximal humeral fractures. Acta Orthop Scand 70(4):325–328
12. Brorson S, Hróbjartsson A (2008) Training improves agreement among doctors using the Neer system for proximal humeral fractures in a systematic review. J Clin Epidemiol 61(1):7–16. doi:10.1016/j.jclinepi.2007.04.014
13. Cuny C, Baumann C, Mayer J, Guignand D, Irrazi M, Berrichi A, Ionescu N, Guillemin F (2013) AST classification of proximal humeral fractures: introduction and interobserver reliability assessment. Eur J Orthop Surg Traumatol 23(1):35–40. doi:10.1007/s00590-011-0916-5