Diagnosing displaced four-part fractures of the proximal humerus: a review of observer studies

Stig Brorson · Jens Bagger · Annette Sylvest ·
Asbjørn Hrobjartsson

Received: 3 April 2008 / Revised: 15 April 2008 / Accepted: 15 April 2008 / Published online: 7 June 2008
© Springer-Verlag 2008

Abstract Displaced four-part fractures comprise 2–10 % of all proximal humeral fractures. The optimal treatment is unclear and randomised trials are needed. The conduct and interpretation of such trials is facilitated by a reproducible fracture classification. We aimed at quantifying observer agreement on the classification of displaced four-part fractures according to the Neer system. Published and unpublished data from five observer studies were reviewed. Observers agreed less on displaced four-part fractures than on the overall Neer classification. Mean kappa values for interobserver agreement ranged from 0.16 to 0.48. Specialists agreed slightly more than fellows and residents. Advanced imaging modalities (CT and 3D CT) seemed to contribute more to classification of displaced four-part patterns than in less complex fracture patterns. Low observer agreement may challenge the clinical approach to displaced four-part fractures and poses a problem for the interpretation and generalisation of results from future randomised trials.

Résumé Les fractures à quatre parts représentent 2 à 10% du total des fractures de l'extrémité supérieure de l'humérus. Le traitement optimal de ces fractures n'est pas évident. Une randomisation est nécessaire. De tels essais randomisés sont possibles si l'on dispose d'une classification fiable et

reproductible de ces fractures. Les données publiées ou non publiées concernant 5 études ont été examinées. Notre but était également d'évaluer la classification en fonction de l'évolution. Pour les observateurs, sur l'ensemble de cette classification, les fractures à quatre parts étaient les plus difficiles à évaluer et la consolidation moins fiable. Les observations inter observateurs sont étalées de 0,16 à 0,48 sur le plan statistique (tests Kappa), l'avis des médecins spécialistes étant légèrement différent des « fellows » et des « résidents ». L'utilisation d'un scanner et notamment d'un scanner à trois dimensions a permis semble-t-il de mieux contribuer à la classification de ces fractures. Tout ceci intervient dans l'évaluation clinique de ces patients et pose le problème de l'interprétation, des résultats pour de futures études randomisées.

Introduction

Fractures of the proximal humerus are common injuries. The incidence is about 73 per 100,000 and is likely to increase [1, 2]. Displaced four-part fractures are among the most severe injuries and comprise 2–10% of all proximal humeral fractures [3, 4]. The appropriate treatment for displaced four-part fractures is uncertain. Several techniques for percutaneous pinning, external fixation or open reduction and internal fixation including plate fixation, tension-band wiring, Kirschner wiring, transosseous suturing, intramedullary rod fixation, screw fixation and, recently, locking plates, have been proposed in observational studies. Most techniques have shown unsatisfactory outcomes except in 'valgus impacted' four-part fractures (Fig. 1) or in selected younger populations. Thus, most textbooks recommend primary hemiarthroplasty in 'classical' four-part fractures (Fig. 2) and four-part 'fracture-dislocations' (Fig. 3) in elderly patients with poor bone quality, and open

S. Brorson (✉) · J. Bagger
Department of Orthopaedic Surgery, Herlev University Hospital,
Herlev, Denmark
e-mail: sbrorson@hotmail.com

A. Sylvest
Department of Orthopaedic Surgery,
Bispebjerg University Hospital,
Copenhagen, Denmark

A. Hrobjartsson
The Nordic Cochrane Centre,
Copenhagen, Denmark

Fig. 1 'Valgus impacted' four-part fracture characterised by impaction of the humeral head into the humeral shaft with variable displacement of the tuberosities. **a** Humeral shaft. **b** Greater tuberosity. **c** Lesser tuberosity. **d** Head of humerus. Modified with permission from Murray and Zuckerman [4]

Fig. 2 'Classical' four-part fracture characterised by lateral displacement and rotation of the humeral head. **a** Humeral shaft. **b** Greater tuberosity. **c** Lesser tuberosity. **d** Head of humerus. Modified with permission from Murray and Zuckerman [4]

reduction and internal fixation in younger patients and in 'valgus impacted' fractures [4–6].

Only few and disparate randomised clinical trials have tested various interventions for displaced four-part fractures [8–11], and systematic reviews have requested further randomised trials [12–15].

A reproducible identification of displaced four-part fractures is a challenge and poses a problem for the interpretation and generalisation of results of randomised trials. Observer studies have reported low agreement among doctors classifying proximal humeral fractures according to the Neer system [16]. Some authors have reported a trend of lower agreement on more complex fractures, but no observer study has reported agreement among doctors classifying displaced four-part fractures. The aim of this study was to quantify observer agreement on classification of displaced four-part fractures according to the Neer system.

Materials and methods

We conducted a review of observer studies including data on displaced four-part fractures. Studies were identified from a systematic review of observer studies of proximal humeral fractures classified according to the Neer system

Fig. 3 'Fracture-dislocation' characterised by the humeral head being detached and dislocated. Reprinted with permission from Neer CS [7]

[16]. Authors of all studies were contacted by phone or e-mail and asked for additional data. We extracted published and unpublished data on agreement among doctors classifying displaced four-part fractures. Mean kappa values for inter- and intraobserver agreement were calculated. Major differences in the prevalence of four-part fractures within the populations studied precluded a direct comparison of kappa values between studies [17]. We therefore compared mean kappa values for agreement on displaced four-part fractures to the overall agreement on the Neer system within each study. We also studied the agreement on classification of four-part fractures on different levels of clinical experience. In studies involving training of observers we analysed the pattern of changes after training.

Mean kappa values for inter- and intraobserver agreement and 95% confidence intervals were calculated according to Svanholm et al. [18]. Kappa values were interpreted qualitatively according to Landis and Koch [19].

Results

Five observer studies provided data allowing for calculation of interobserver agreement on displaced four-part fractures [20–24]. In three studies kappa values for interobserver agreement could be calculated from published data [20, 23, 24]. In two studies [21, 22] the original data sheets were still available. Three studies provided data allowing calculation of intraobserver agreement [22–24].

The prevalence of displaced four-part fractures ranged from 2 to 23% in the populations studied (Table 1). In the study involving training of observers the prevalence of

displaced four-part fractures decreased from 10% (28 out of 294) to 2% (6 out of 294) after training. The decrease was found on all levels of clinical experience. Most commonly the initial assignment of displaced four-part fractures was changed to displaced two-part fractures (Fig. 4).

In six out of eight classification sessions performed in the five studies, observers agreed less on displaced four-part fractures than on the overall Neer classification (Table 1). Mean kappa values for interobserver agreement on displaced four-part fractures ranged from 0.16 to 0.48 (Table 1) indicating slight to moderate agreement. Specialists agreed more (mean kappa = 0.26) than fellows (mean kappa = 0.12) and residents (mean kappa = 0.10) [21]. A higher level of agreement was found in selected populations of displaced fractures evaluated by computed tomography (CT) or 3D CT [23, 24] than in unselected populations of consecutive patients [21, 22]. In the study involving training of observers [22] mean kappa value for interobserver agreement on displaced four-part fractures increased from 0.22 to 0.32 compared to an increase in agreement on the entire Neer system from 0.27 to 0.62.

In the three studies allowing for calculation of intraobserver agreement on displaced four-part fractures [22–24] mean kappa values ranged from 0.26 to 0.55 as compared to 0.51 to 0.56 for overall agreement on the Neer system.

Discussion

We found a slight to moderate agreement among doctors classifying displaced four-part fractures according to the Neer system. The agreement on four-part fractures was

Table 1 Mean kappa values (95% confidence intervals) for interobserver agreement on displaced four-part fractures compared to overall interobserver agreement on all other categories within the Neer system

Study	Cases (n)	Observers (n)	Imaging modality	Prevalence of four-part fractures ^b	Agreement on four-part fractures (mean κ)	Overall agreement (mean κ)
Brien et al. [20]	28 (21) ^a	4	Radiographs only	0.16 (13/84)	0.48 (0.31–0.66)	0.43 (0.32–0.53)
Sjöden et al. [23]	26	10	Radiographs and CT	0.23 (59/260) ^c 0.20 (51/260) ^d	0.373 (0.31–0.42) 0.47 ^d (0.41–0.52)	0.42 ^c (0.39–0.45) 0.43 ^d (0.40–0.46)
Sjöden et al. [24]	24	7	Radiographs and 3D-CT	0.22 (37/168) ^c 0.20 (33/168) ^d	0.33 ^c (0.24–0.42) 0.46 ^d (0.37–0.55)	0.44 ^c (0.41–0.50) 0.49 ^d (0.44–0.53)
Brorson et al. [21]	42	24	Radiographs only	0.07 (68/1008)	0.16 (0.14–0.18)	0.27 (0.26–0.28)
Brorson et al. [22]	42	7	Radiographs only	0.10 (28/294) ^e 0.02 (6/294) ^f	0.22 ^e (0.16–0.29) 0.32 ^f (0.25–0.39)	0.27 ^e (0.23–0.31) 0.62 ^f (0.57–0.67)

CT computed tomography, 3D three-dimensional

^a 28 cases were included in the study but only 21 reported in published data

^b Proportion of cases classified as displaced four-part fractures by any observer out of all observations

^c First classification session

^d Second classification session

^e Before training

^f After training

Fig. 4 Pattern of changes in assignment of displaced four-part fractures after training.

*‘Two-part fractures’ include isolated, displaced anatomical neck, surgical neck, greater tuberosity, and lesser tuberosity fractures

28 four-part assignments before training were reduced to 6 four-part fractures after training

- 27 four-part fractures were changed to:
 - 5 non-displaced fractures
 - 16 two-part* fractures
 - 6 three-part fractures
- 1 four-part fracture remained unchanged
- 2 fracture-dislocations were changed into 2 four-part fractures
- 1 two-part* fracture was changed into a four-part fracture
- 1 three-part fracture was changed into a four-part fracture
- 1 articular surface fracture was changed into a four-part fracture

lower than the overall agreement on the entire Neer system in six out of eight classification sessions.

Strengths and weaknesses of the present study

Selection of five studies due to lack of data in another seven studies from the systematic review [16] may potentially bias the results due to an association between reporting of the data and the nature of the results. However, the systematic review found no association between high methodological quality and level of agreement regardless of explorative sensitivity analyses [16]. A re-analysis of quality assessment data (on a scale from 0 to 9) from the systematic review reveals no difference in methodological quality between the five studies reporting agreement on four-part fractures and the other seven studies (mean quality score 6.0 [range, 3–9] as compared to 4.9 [range, 2–7]).

Kappa and the prevalence of the diagnosis

Kappa decreases with very high or very low prevalence of the category under study [17] and a lower kappa value is expected when an uncommon category is studied. Thus, the reported kappa values should be interpreted in the light of the characteristics of the population under study. In a specialised shoulder clinic a higher prevalence of displaced four-part fractures can be expected than in a general trauma unit. This may explain the lower kappa values reported in the studies [21, 22], including unselected cases from a general orthopaedic population with a low prevalence of four-part fractures.

Advanced imaging modalities and observer agreement

In studies including supplemental CT and 3D CT scans a higher agreement on displaced four-part fractures was found [23, 24]. Advanced imaging modalities have not been found to improve overall observer agreement on the Neer system in previous studies [16]. However, advanced imaging modalities may contribute more to classification of

more complex fracture patterns, for example, fractures involving the lesser tuberosity. This may explain the somewhat higher observer agreement found in studies including CT and 3D CT scans.

Training of observers

The lower proportion of displaced four-part fractures reported after training in the Neer system could indicate that these fracture patterns are less common than previously reported. The observed changes in classification practise after training (Fig. 4) may be attributed to the emphasis during training sessions to include only displaced segments in the classification. If, for example, a displaced two-part surgical neck fracture involves undisplaced fractures of the tuberosities, it may have been classified as a four-part fracture prior to training and, correctly, as a two-part fracture after training. It is not clear, however, why specialists did not agree more after training (mean kappa = 0.33) as compared to their overall agreement on the Neer system (mean kappa = 0.79).

Patterns of displaced four-part fractures

Displaced four-part fractures can be divided into three patho-anatomical patterns of increasing prognostic severity: ‘valgus impacted’ fractures, ‘classical’ fractures, and ‘fracture dislocations’ (Figs. 1, 2, and 3, respectively). There is, however, no consensus about the clinical importance of identifying these different patterns of displaced four-part fractures.

The valgus impacted fracture pattern was not included in the original Neer system. In a later review paper [25] Neer proposed consideration of four-part fractures as a continuum of lateral displacement progressing from minimal displacement (Neer group 1) to displaced four-part fractures (Neer group 12) and four-part fracture dislocations (Neer group 13 and 14). The four-part valgus impacted fracture pattern can be regarded a ‘precursor’ to the ‘classical’ four-part fracture with lateral displacement (Neer group 12). According to Neer, it may be difficult for surgeons to agree on the incidence and treatment of a borderline lesion.

Previous observer studies have not distinguished four-part fracture dislocations from two- and three-part fracture dislocations. In our review, only displaced, but not dislocated, four-part fractures were included. Some authors have found no statistically significant difference in clinical outcome between four-part fracture dislocations and displaced four-part fractures without dislocation [26]. Thus, to distinguish ‘classical’ four-part fractures from four-part fracture dislocations may be less important from a clinical perspective. Furthermore, it has been suggested that fracture dislocations should not be classified until the dislocation is reduced [27].

Unanswered questions and future research

The conduct and reporting of randomised trials on interventions for displaced four-part fractures is facilitated by a reproducible fracture classification. No observer study has attempted to quantify the agreement on different patterns of four-part fractures and treatment decisions among shoulder surgeons. We think that such a study would be valuable. Until then, reported effects of interventions for displaced four-part fractures should be interpreted cautiously.

References

- Bengner U, Johnell O, Redlund-Johnell I (1988) Changes in the incidence of fracture of the upper end of the humerus during a 30-year period. A study of 2125 fractures. *Clin Orthop Relat Res* 231:179–182
- Palvanen M, Kannus P, Niemi S, Parkkari J (2006) Update in the epidemiology of proximal humeral fractures. *Clin Orthop Relat Res* 442:87–92
- Court-Brown CM, Garg A, McQueen MM (2001) The epidemiology of proximal humeral fractures. *Acta Orthop Scand* 72:365–371
- Murray D, Zuckerman JD (2005) Four-part fractures and fracture-dislocations. In: Zuckerman JD, Koval KJ (eds) *Shoulder fractures. The practical guide to management*. Thieme, New York, pp 295–298
- Lee EW, Flatow EL (2005) Arthroplasty for proximal humerus fractures, nonunions, and malunions. In: Bigliani LU, Flatow EL (eds) *Shoulder arthroplasty*. Springer, New York, pp 86–116
- Williams GR, Yamaguchi K, Ramsey ML, Galatz LM (2005) *Shoulder and elbow arthroplasty*. Lippincott Williams & Wilkins, Philadelphia
- Neer CS (1955) Articular replacement for the humeral head. *J Bone Joint Surg Am* 37A:215–228
- Holbein O, Bauer G, Hoellen IP, Keppler P, Hehl G, Kinzl L (1999) Is primary endoprosthetic replacement of the humeral head an alternative treatment for comminuted fractures of the proximal humerus in elderly patients? *Osteosynthese International* 7:207–210
- Kristiansen B, Kofoed H (1988) Transcutaneous reduction and external fixation of displaced fractures of the proximal humerus. A controlled clinical trial. *J Bone Joint Surg Br* 70:821–824
- Stableforth PG (1984) Four-part fractures of the neck of the humerus. *J Bone Joint Surg Br* 66:104–108
- Zyto K, Ahrengart L, Sperber A, Tornkvist H (1997) Treatment of displaced proximal humeral fractures in elderly patients. *J Bone Joint Surg Br* 79:412–417
- Bhandari M, Matthys G, McKee MD (2004) Four part fractures of the proximal humerus. *J Orthop Trauma* 18:126–127
- Handoll HHG, Madhok R (2001) Interventions for treating proximal humeral fractures in adults. *Cochrane Database of Systematic Reviews*, 1:CD000434
- Misra A, Kapur R, Maffulli N (2001) Complex proximal humeral fractures in adults—a systematic review of management. *Injury* 32:363–372
- Tingart M, Bathis H, Bouillon B, Tiling T (2001) The displaced proximal humeral fracture: is there evidence for therapeutic concepts? *Chirurg* 72:1284–1291
- Brorson S, Hrobjartsson A (2008) Training improves agreement among doctors using the Neer system for proximal humeral fractures in a systematic review. *J Clin Epidemiol* 61:7–16
- Gjorup T (1988) The kappa coefficient and the prevalence of a diagnosis. *Methods Inf Med* 27:184–186
- Svanholm H, Starklint H, Gundersen HJ, Fabricius J, Barlebo H, Olsen S (1989) Reproducibility of histomorphologic diagnoses with special reference to the kappa statistic. *APMIS* 97:689–698
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- Brien H, Nofall F, MacMaster S, Cummings T, Landells C, Rockwood P (1995) Neer's classification system: a critical appraisal. *J Trauma* 38:257–260
- Brorson S, Bagger J, Sylvest A, Hrobjartsson A (2002) Low agreement among 24 doctors using the Neer-classification; only moderate agreement on displacement, even between specialists. *Int Orthop* 26:271–273
- Brorson S, Bagger J, Sylvest A, Hrobjartsson A (2002) Improved interobserver variation after training of doctors in the Neer system. A randomised trial. *J Bone Joint Surg Br* 84:950–954
- Sjöden GO, Movin T, Guntner P, Aspelin P, Ahrengart L, Ersmark H et al (1997) Poor reproducibility of classification of proximal humeral fractures. Additional CT of minor value. *Acta Orthop Scand* 68:239–242
- Sjöden 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* 70:325–328
- Neer CS (2002) Four-segment classification of proximal humeral fractures: purpose and reliable use. *J Shoulder Elbow Surg* 11:389–400
- Trupka A, Wiedemann E, Ruchholtz S, Brunner U, Habermeyer P, Schweiberer L (1997) Dislocated multiple fragment fractures of the head of the humerus. Does dislocation of the humeral head fragment signify a worse prognosis? *Unfallchirurg* 100:105–110
- Habermeyer P, Schweiberer L (1989) Fractures of the proximal humerus. *Orthopäde* 18:200–207