

Reliability and Reproducibility of the OTA/AO Classification for Humeral Shaft Fractures

Kiran C. Mahabier, MD¹, Esther M.M. Van Lieshout, MSc PhD¹, Boyd C. Van Der Schaaf¹, Gert R. Roukema, MD², Bas J. Punt, MD³, Michael H.J. Verhofstad, MD PhD¹, Dennis Den Hartog, MD PhD^{1*}, On behalf of the HUMMER trial investigators

¹ Trauma Research Unit Department of Surgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

² Department of Surgery, Maastad Hospital, Rotterdam, The Netherlands

³ Department of Surgery, Albert Schweitzer Hospital, Dordrecht, The Netherlands

* Corresponding author:

D. Den Hartog, MD PhD

Erasmus MC, University Medical Center Rotterdam

Trauma Research Unit, Department of Surgery

P.O. Box 2040

3000 CA Rotterdam

The Netherlands

Phone: +31.10.703.1050

Fax: +31.10.7032396

Mail: d.denhartog@erasmusmc.nl

Kiran C. Mahabier, MD

Trauma Research Unit Department of Surgery

Erasmus MC, University Medical Center Rotterdam

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

k.mahabier@erasmusmc.nl

Esther M.M. Van Lieshout, MSc PhD

Trauma Research Unit Department of Surgery

Erasmus MC, University Medical Center Rotterdam

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

e.vanlieshout@erasmusmc.nl

Boyd Van der Schaaf

Trauma Research Unit Department of Surgery

Erasmus MC, University Medical Center Rotterdam

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

b.schaaf@erasmusmc.nl

Gert R. Roukema, MD

Department of Surgery

Maasstad Hospital Rotterdam

P.O. Box 9100, 3007 AC Rotterdam, The Netherlands

roukemag@maasstadziekenhuis.nl

Bas J. Punt, MD

Department of Surgery

Albert Schweitzer Hospital,

P.O. Box 444, 3300 AK Dordrecht, The Netherlands

b.j.punt@asz.nl

Michael H.J. Verhofstad, MD PhD

Trauma Research Unit Department of Surgery

Erasmus MC, University Medical Center Rotterdam

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

m.verhofstad@erasmusmc.nl

Dennis Den Hartog, MD PhD

Trauma Research Unit Department of Surgery

Erasmus MC, University Medical Center Rotterdam

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

d.denhartog@erasmusmc.nl

HUMMER TRIAL INVESTIGATORS

Hugo W. Bolhuis, MD⁴
Email: h.bolhuis@gelre.nl

P. Koen Bos, MD PhD⁵
Email: p.k.bos@erasmusmc.nl

Maarten W.G.A. Bronkhorst, MD⁶
Email: mbronkhorst@bronovo.nl

Milko M.M. Bruijninx, MD⁷
Email: mbruijninx@ysl.nl

P. Ted Den Hoed, MD PhD⁸
Email: pt.hoed@ikazia.nl

Boudewijn J. Dwars, MD PhD⁹
Email: bj.dwars@slz.nl

J. Carel Goslings, MD PhD¹⁰
Email: j.c.goslings@amc.nl

Robert Haverlag, MD¹¹
Email: r.haverlag@olvg.nl

Martin J. Heetveld, MD PhD¹²
Email: heetveld@kg.nl

Albert J.H. Kerver, MD PhD¹³
Email: b.kerver@sfg.nl

Karel A. Kolkman, MD¹⁴
Email: kkolkman@rijnstate.nl

Peter A. Leenhouts, MD¹⁵
Email: p.a.leenhouts@amc.nl

Ron Onstenk, MD¹⁶
Email: ron.onstenk@ghz.nl

Martijn Poeze, MD PhD¹⁷
Email: m.poeze@maastrichtuniversity.nl

Rudolf W. Poolman, MD PhD¹⁸
Email: r.w.poolman@olvg.nl

W. Herbert Roerdink, MD PhD¹⁹
Email: roerdink@dz.nl

Jan Bernard Sintenie, MD²⁰
Email: jbs@elkerliek.nl

Nicolaj M.R. Soesman, MD²¹
Email: nsoesman@vlietlandziekenhuis.nl

Frank H.W.M. Van der Heijden, MD PhD²²
Email: f.vd.heijden@elisabeth.nl

Peer Van der Zwaal, MD PhD²³
Email: p.van.der.zwaal@mchaaglanden.nl

Jan P. Van Dijk, MD²⁴
Email: dijkj@zgv.nl

Hans-Peter W. Van Jonbergen, MD PhD²⁵
Email: vanjonbergen@dz.nl

Egbert J.M.M. Verleisdonk, MD PhD²⁶
Email: ejverlei@diakhuis.nl

Jos P.A.M. Vroemen, MD PhD²⁷
Email: jvroemen@amphia.nl

Marco Waleboer, MD²⁸
Email: m.waleboer@adrz.nl

Wietse P. Zuidema, MD²⁹
Email: w.zuidema@vumc.nl

AFFILIATIONS

⁴ Department of Surgery, Gelre Hospital, P.O. Box 9014, 7300 DS Apeldoorn, The Netherlands

⁵ Department of Orthopaedic Surgery, Erasmus MC, University Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

⁶ Department of Surgery, MC Haaglanden Bronovo, P.O. Box 96900, 2509 JH The Hague, The Netherlands

⁷ Department of Surgery, IJsselland Hospital, P.O. Box 690, 2900 AR Capelle a/d IJssel, The Netherlands

- ⁸ Department of Surgery, Ikazia Hospital, P.O. Box 5009, 3008 AA Rotterdam, The Netherlands
- ⁹ Department of Surgery, Slotervaart Hospital, P.O. Box 90440, 1006 BK Amsterdam, The Netherlands
- ¹⁰ Trauma Unit Department of Surgery, Academic Medical Center, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands
- ¹¹ Department of Surgery, Onze Lieve Vrouwe Gasthuis, P.O. Box 95500, 1090 HM Amsterdam, The Netherlands
- ¹² Department of Surgery, Spaarne Gasthuis, P.O. Box 417, 2000 AK Haarlem, The Netherlands
- ¹³ Department of Surgery, Sint Franciscus Gasthuis, P.O. Box 10900, 3004 BA Rotterdam, The Netherlands
- ¹⁴ Department of Surgery, Rijnstate, P.O. Box 9555, 6800 TA Arnhem, The Netherlands
- ¹⁵ Department of Surgery, Zaans Medical Center, P.O. Box 210, 1500 EE Zaandam, The Netherlands. Currently: Trauma Unit Department of Surgery, Academic Medical Center, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands
- ¹⁶ Department of Orthopaedic Surgery, Groene Hart Hospital, P.O. Box 1098, 2800 BB Gouda, The Netherlands
- ¹⁷ Department of Trauma Surgery, Maastricht University Medical Center, PO Box 5800, 6202 AZ Maastricht, The Netherlands
- ¹⁸ Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis, P.O. Box 95500, 1090 HM Amsterdam, The Netherlands
- ¹⁹ Department of Surgery, Deventer Hospital, P.O. Box 5001, 7400 GC Deventer, The Netherlands
- ²⁰ Department of Surgery, Elkerliek Hospital, P.O. Box 98, 5700 AB Helmond, The Netherlands
- ²¹ Department of Surgery, Vlietland Hospital, P.O. Box 215, 3100 AE Schiedam, The Netherlands
- ²² Department of Surgery, St Elisabeth Hospital, P.O. Box 90151, 5000 LC Tilburg, The Netherlands
- ²³ Department of Orthopaedic Surgery, MC Haaglanden Bronovo, P.O. Box 432, 2501 CK The Hague, The Netherlands
- ²⁴ Department of Surgery, Hospital Gelderse Vallei, P.O. Box 9025, 6710 HN Ede, The Netherlands

²⁵ Department of Orthopaedic Surgery, Deventer Hospital, P.O. Box 5001, 7400 GC Deventer, The Netherlands

²⁶ Department of Surgery, Diaconessenhuis, P.O. Box 80250, 3508 TG Utrecht, The Netherlands

²⁷ Department of Surgery, Amphia Hospital, P.O. Box 90158, 4800 RK Breda, The Netherlands

²⁸ Department of Surgery, Admiraal De Ruyter Hospital, P.O. Box 3200, 4380 DD Vlissingen, The Netherlands

²⁹ Department of Trauma Surgery, VU University Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands

Conflict of interest statement and Source of Funding

The authors declare that they have no conflict of interest. No external funding was received for this study.

Ethical approval

The study was exempted by the Medical Research Ethics Committee Erasmus MC (No. MEC-2014-107).

ABSTRACT

Objectives: This study aimed to determine inter-observer reliability and intra-observer reproducibility of the OTA/AO classification for humeral shaft fractures, and to evaluate differences between fracture types, fracture groups, and surgical specializations.

Methods: 30 observers (25 orthopedic trauma surgeons and five general orthopedic surgeons) independently classified 90 humeral shaft fractures according to the OTA/AO classification. Patients of 16 years and older were included. Periprosthetic, recurrent, and pathological fractures were excluded. Radiographs were provided in random order, and observers were blinded to clinical information. To determine intra-observer agreement, radiographs were reviewed again after two months in a different random order. Agreement was assessed using kappa statistics.

Results: Inter-observer agreement for the three fracture types was moderate ($\kappa=0.60$; 0.59-0.61). It was substantial for type A ($\kappa=0.77$; 0.70-0.84), and moderate for type B ($\kappa=0.52$; 0.46-0.58) and type C fractures ($\kappa=0.46$; 0.42-0.50). Inter-observer agreement for the nine fracture groups was moderate ($\kappa=0.48$; 95% CI 0.48-0.48). Orthopedic trauma surgeons had better overall agreement for fracture types, and general orthopedic surgeons had better overall agreement for fracture groups. Observers classified 64% of fractures identically in both rounds. Intra-observer agreement was substantial for the three types ($\kappa=0.80$; 0.77-0.81) and nine groups ($\kappa=0.80$; 0.77-0.82). Intra-observer agreement showed no differences between surgical disciplines.

Conclusions: The OTA/AO classification for humeral shaft fractures has a moderate inter-observer and substantial intra-observer agreement for fracture types and groups.

INTRODUCTION

A fracture classification system should not only provide a reliable and reproducible means of communication between physicians, but also provide for repeated viewings of the same material.¹ Ideally it should have a prognostic value for the outcome of patients, assist in managing fractures, and assist documentation and research.² Humeral shaft fractures are most widely classified using the OTA/AO classification system.^{1, 3}

Despite its widespread use, the OTA/AO classification has not been validated for humeral shaft fractures, including the complete range of fracture types and groups. The primary aim of this study was therefore to determine the inter-observer reliability and intra-observer reproducibility of the OTA/AO classification for humeral shaft fractures. The secondary aims were to evaluate if reliability and reproducibility differed between the three different fracture types or the nine fracture groups, and to assess if agreement was dependent on surgical specialization or time spent on classifying the radiographs.

METHODS

Classification

In the OTA/AO classification, number 1 stands for the humerus and number 2 for the diaphyseal segment. As shown in Figure 1, three types of fractures are defined and coded with letters: type A consists of simple fractures, type B of wedge-type fractures, and type C of complex fractures. Each of these three types can be further subdivided into groups 1, 2, or 3. Overall, the OTA/AO classification system for humeral shaft fractures has nine groups (12-A1/2/3, 12-B1/2/3, 12-C1/2/3).²

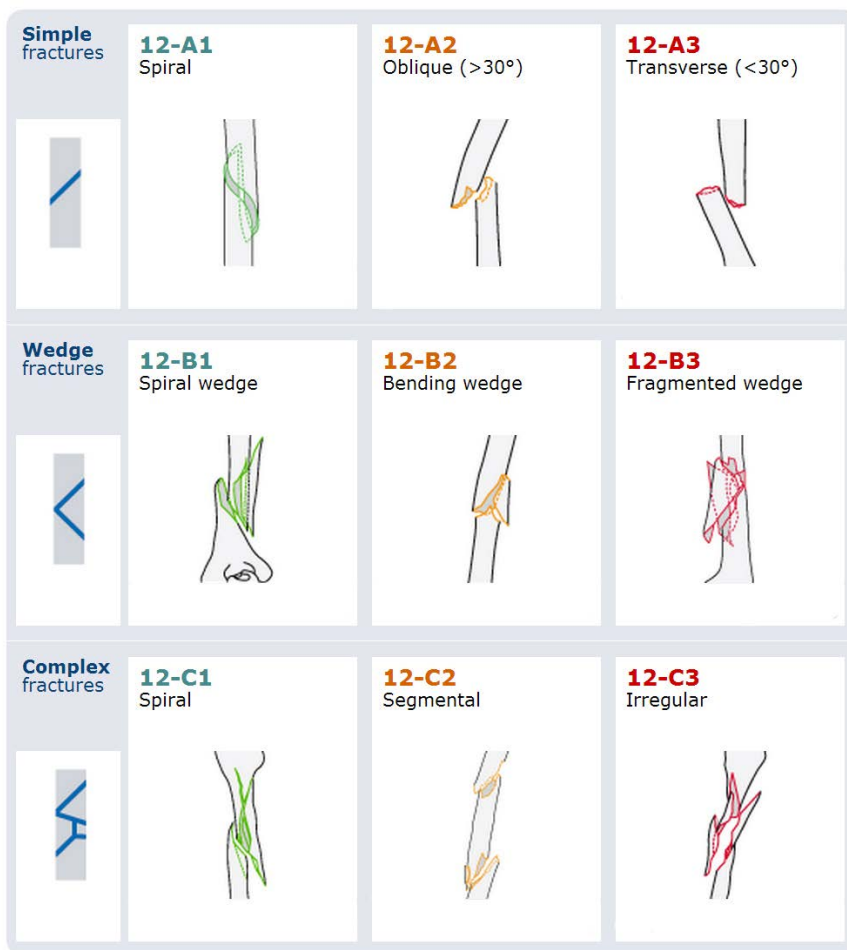


Figure 1. OTA/AO classification for humeral shaft fractures^{1,3}

Study subjects

Patients were selected from the hospital records and from the radiology system PACS (Picture Archiving and Communication System) of three hospitals. Eligible patients had already been identified from hospital databases as part of another study.⁴ All patients aged 16 years or older treated for a humeral shaft fracture in one of three hospitals were included in this study. The humeral shaft was defined as the area between the surgical neck and the area immediately above the supracondylar ridge. Radiographs had to include initial (i.e., before treatment) anterior-posterior or lateral images. Patients with periprosthetic, recurrent, or pathological fractures were excluded. Patients with fractures extending outside the predefined shaft area were excluded as well. A total of 90 patients representing the full spectrum of humeral shaft fractures were selected by the clinical investigator (KCM). The investigator was adequately trained, had sufficient experience to select the radiographs of humeral shaft fractures and was not involved as observer. The sample size of 90 patients allowed for all groups to be represented with ten subjects. The first ten subjects per group were included. In order to reflect routine day-to-day practice the quality of the images was not used as an exclusion criterion. The radiographs that were accepted for clinical decision making were also considered adequate for this classification study. All radiographs available (two or three for 85 patients and one for the other five) were used. Radiographs had no identifying information. Following randomization using a web-based list randomizer (www.random.org), they were imported into an open-source Digital Imaging and Communications in Medicine (DICOM) compliant viewer (RadiAnt DICOM Viewer 1.9.14, Medixant, Poznan, Poland). This viewer provided all necessary tools for adequate viewing (e.g., fluid zooming and panning, brightness and contrast adjustments, and angle measurements). The same workstation and DICOM viewer were used for all observations in order to guarantee identical viewing conditions.

Observers

Thirty-seven consultant upper extremity (orthopedic) trauma surgeons experienced in the treatment of humeral fractures were invited to act as observer. All surgeons act as site principal investigator in a multicenter clinical study comparing the operative and non-operative treatment of humeral shaft fractures (HUMMER study).⁵ Years of independent practice and whether the OTA/AO classification is used in daily practice were noted for each observer.

Study procedure

Each observer independently classified 90 humeral shaft fractures according to the classification system. All radiographs were provided in random order, and the observers were given as much time as needed for accurate assessment. Observers were blinded to clinical information and were not allowed to discuss their observations with other investigators. All observers were familiar with the OTA/AO classification system used in this study. In order to ensure unambiguous application of the fracture classification system, an overview of the classification system was available to the surgeons during the classification (Figure 1). The amount of time needed to classify all radiographs was recorded.

In order to determine the intra-observer agreement, all radiographs were reviewed a second time at least two months after the first review. On the second occasion, images were provided in a different random order. Inter-observer reliability is the degree of agreement when two or more independent observers classify the same fracture. Intra-observer reproducibility is agreement when one observer classifies the same fracture more than once.

Statistical analysis

Normality of continuous data was judged from frequency histograms and Q-Q plots, homogeneity of variances was tested using the Levene's test. Data were analyzed using Kappa statistics, as described by Cohen.⁶ The kappa coefficient represents the agreement between two sets of observations compared with the likelihood of agreement based on chance alone. The kappa coefficient ranges from 1 (perfect agreement) to <0 (systematic disagreement, or no more agreement than would be expected by chance alone). The kappa value for inter-observer agreement was calculated for each possible pair of observers before calculating the mean kappa value.⁷ Interpretation of the values were carried out according to the guidelines of Landis and Koch, which suggest that values <0 represent poor reliability; 0.00-0.20 slight agreement; 0.21-0.40 fair agreement; 0.41-0.60 moderate agreement; 0.61-0.80 substantial agreement; and 0.81-1.00 almost perfect agreement.⁸ The kappa value for intra-observer agreement were calculated for each of the individual observers before calculation the mean kappa value. The kappa values were classified according to Landis and Koch as described in the previous section.

Kappa values for both inter- and intra-observer agreement were assessed for the nine groups (12-A1/2/3, 12-B1/2/3, 12-C1/2/3), as well as for the three types (A, B and C) in order to judge if kappa values differ between fractures. Statistical significance of differences in the kappa values across these groups and types were tested with a one-way Analysis of Variance (ANOVA). Statistical significance of differences between orthopedic trauma and general orthopedic surgeons and time spent on the classification of all radiographs were tested with the a Student's t-test. Data were analyzed using the Statistical Package for the Social Sciences (SPSS) version 21 or higher (SPSS, Chicago, Ill., USA).

RESULTS

Twenty-five of the observers were orthopedic trauma surgeons and five were general orthopedic surgeons. Of the observers 11 worked ten years or less in an independent practice and 19 had more than ten years of experience in an independent practice.

As shown in Table 1, the inter-observer reliability was moderate for the three fracture types ($\kappa=0.60$; 95% CI 0.59-0.61). It was substantial for type A fractures ($\kappa=0.77$; 95% CI 0.70-0.84), and moderate for type B ($\kappa=0.52$; 95% CI 0.46-0.58) and type C fractures ($\kappa=0.46$; 95% CI 0.42-0.50).

The inter-observer agreement for the nine groups was moderate ($\kappa=0.48$; 95% CI 0.48-0.48). It was highest for 12-A3 fractures ($\kappa=0.68$; 95% CI 0.64-0.71) and 12-C3 fractures ($\kappa=0.63$; 95% CI 0.60-0.66) and lowest for 12-C1 fractures ($\kappa=0.24$; 95% CI 0.21-0.27). The overall inter-observer agreement for three fracture types and nine groups both showed statistical significance differences between orthopedic trauma and general orthopedic surgeons. Orthopedic trauma surgeons had better overall agreement for fracture types ($\kappa=0.60$; 95% CI 0.61-0.59 and $\kappa=0.58$; 95% CI 0.53-0.62, respectively). For overall inter-observer agreement for the nine groups it was the other way around and general orthopedic surgeons had better overall agreement ($\kappa=0.47$; 95% CI 0.47-0.48 and $\kappa=0.51$; 95% CI 0.50-0.53, respectively).

No differences were found between surgical specialization and agreement of specific fracture types or groups, except for the 12-B1 fracture group. Orthopedic trauma surgeons had a lower inter-observer agreement for that specific group than general orthopedic surgeons ($\kappa=0.33$; 95% CI 0.28-0.37 and $\kappa=0.47$; 95% CI 0.31-0.62, respectively).

The intra-observer reproducibility was substantial as shown in Table 2 ($\kappa=0.80$; 95% CI 0.77-0.81) for the three types, as well as for the nine groups ($\kappa=0.80$; 95% CI 0.77-0.82). Observers classified 64% (95% CI 62-67%) of the fractures identically in both rounds. Intra-

observer agreement for types, groups or the percentage of identically classified fractures in both rounds did not differ between surgical specializations.

Both the inter- and intra-observer agreement were not significantly associated with the time spent on the classification of all radiographs.

Table 3 shows an overview of the number of dominant classifications of the 90 fractures classified. In addition, it shows the mostly chosen alternative per fracture type and group.

When the 12-A fracture type was the dominant classification, 12-B was the mostly chosen alternative (61% of classifications). For type 12-B, the mostly chosen alternative was type 12-C (66%). The type 12-B classification was mostly chosen as alternative (91%) when the type 12-C classification was dominant. For the type 12-A1 (simple spiral fracture), the 12-B1 (spiral wedge fracture) was the mostly chosen alternative. The 12-A2 and 12-A3 groups (oblique and transverse fractures, respectively), were both chosen mostly as alternative when these were the dominant classification. For 12-B1 and 12-C1 groups (spiral wedge and complex spiral fractures, respectively) and for 12-B2 and 12-B3 groups (bending wedge and fragmented wedge fractures, respectively) this was also the case. When the 12-C2 segmental group was dominant, the 12-A3 transverse group was chosen mostly as alternative.

Figure 2 shows an example of a fracture with perfect agreement. This fracture was classified in the 12-A2 group by all observers in both rounds. This was also the only fracture with perfect agreement in the entire study. An example of poor agreement is shown in Figure 3.

This fracture was classified in six different groups.



Figure 2. Example of a fracture with perfect agreement

Classified as 12-A3 by all observers in both rounds.

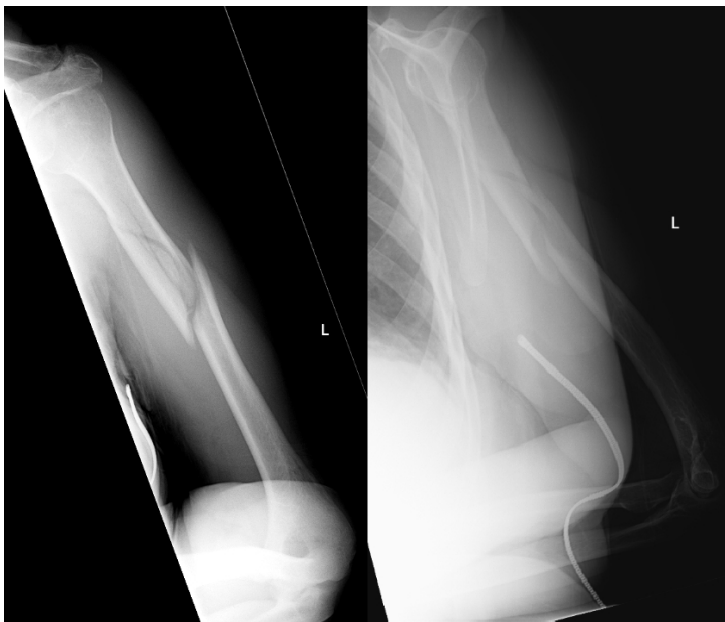


Figure 3. Example of a fracture with poor agreement

Classified into six different groups.

DISCUSSION

The main purpose of the current study was to determine whether the OTA/AO classification is a reliable and reproducible system for the classification of humeral shaft fractures. The inter-observer reliability was moderate and the intra-observer reproducibility was substantial. Although the usability of the OTA/AO classification has previously been questioned, it remains the most widely used classification system in the research of humeral shaft fractures.⁹

The validity of the classification has also been studied in various bone segments, but specific results of the classifications used for humeral shaft fractures are scarce.¹⁰⁻¹⁶ Johnstone et al. concluded in 1993 that the classification system for long bone fractures demonstrated a significant inter-observer variation, but no humeral shaft fractures were included in that study.¹⁷ In the same year, Newey et al. concluded that the classification system was only useful for audit options, but again no humeral shaft fractures were included.¹⁸ Meling et al. reported a moderate inter-observer agreement for the OTA/AO classification for long bone fractures ($\kappa=0.67$ (95% CI: 0.62-0.72)).¹⁹ A study comparing the OTA/AO classification for long bone fractures with a newly proposed classification, including 40 humeral shaft fractures classified by six observers, reported a fair inter-observer agreement ($\kappa=0.30$) and also a fair intra-observer agreement ($\kappa=0.38$) for the OTA/AO classification.²⁰ That classification system, describing fractures by location (proximal, middle, distal, or in combinations when the fracture is located in multiple zones) and morphology (simple [transverse, oblique or spiral], intermediate and complex), had a good inter-observer ($\kappa=0.66$) as well as a moderate intra-observer ($\kappa=0.56$) agreement.

As shown in Table 3, observers did not agree on specific fracture patterns. When most observers classified a fracture as a simple spiral fracture (12-A1), the remaining observers classified it as a spiral wedge (12-B1). When most observers classified a fracture as a spiral

wedge the remaining observers chose the complex spiral fracture (12-C1). Apparently, the fracture lines discriminating these fracture types were easily missed (or thought to be seen). Also, the angle of the fracture seemed difficult to determine. The angle separating the oblique (12-A2) and transverse (12-A3) fracture groups of 30 degrees seemed to cause observers to disagree. For future classifications, specific attention should be paid to these items.

CONCLUSION

The OTA/AO classification system for humeral shaft fractures has a moderate inter-observer agreement for the fracture types and fracture groups. Apart from a substantial agreement for type A fractures the agreement for the other fracture types was moderate. Agreement for specific fracture groups ranged from fair to substantial. The intra-observer agreement was substantial for the fracture types and groups, with 64% fractures classified identically in both rounds. Specific attention should be paid to discriminating A1 from B1, B1 from C1, and A2 from A3 as fracture lines or angles discriminating these two were often misinterpreted.

Acknowledgments

Hugo W. Bolhuis, MD, Department of Surgery, Gelre Hospital, The Netherlands; P. Koen Bos, MD PhD, Department of Orthopaedic Surgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; Maarten W.G.A. Bronkhorst, MD, Department of Surgery, MC Haaglanden Bronovo, The Hague, The Netherlands; Milko M.M. Bruijninx, MD, Department of Surgery, IJsselland Hospital, Capelle aan den IJssel, The Netherlands; P. Ted Den Hoed, MD PhD, Department of Surgery, Ikazia Hospital, Rotterdam, The Netherlands; Boudewijn J. Dwars, MD PhD, Department of Surgery, Slotervaart Hospital, Amsterdam, The Netherlands; J. Carel Goslings, MD PhD, Trauma Unit Department of Surgery, Academic Medical Center, Amsterdam, The Netherlands; Robert Haverlag, MD, Department of Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands; Martin J. Heetveld, MD PhD, Department of Surgery, Spaarne Gasthuis, Haarlem, The Netherlands; Albert J.H. Kerver, MD PhD, Department of Surgery, Sint Franciscus Gasthuis, Rotterdam, The Netherlands; Karel A. Kolkman, MD, Department of Surgery, Rijnstate, Arnhem, The Netherlands; Peter A. Leenhouts, MD, Department of Surgery, Zaan Medical Center, Zaandam, The Netherlands; Ron Onstenk, MD, Department of Orthopaedic Surgery, Groene Hart Hospital, Gouda, The Netherlands; Martijn Poeze, MD PhD, Department of Trauma Surgery, Maastricht University Medical Center, Maastricht, The Netherlands; Rudolf W. Poolman, MD PhD, Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands; W. Herbert Roerdink, MD PhD, Department of Surgery, Deventer Hospital, Deventer, The Netherlands; Jan Bernard Sintenie, MD, Department of Surgery, Elkerliek Hospital, Helmond, The Netherlands; Nicolaj M.R. Soesman, MD, Department of Surgery, Vlietland Hospital, Schiedam, The Netherlands; Frank H.W.M. Van der Heijden, MD PhD, Department of Surgery, St Elisabeth Hospital, Tilburg, The Netherlands; Peer Van der Zwaal, MD PhD, Department of Orthopaedic Surgery, MC

Haaglanden Bronovo, The Hague, The Netherlands; Jan P. Van Dijk, MD, Department of Surgery, Hospital Gelderse Vallei, Ede, The Netherlands; Hans-Peter W. Van Jonbergen, MD PhD, Department of Orthopaedic Surgery, Deventer Hospital, Deventer, The Netherlands; Egbert J.M.M. Verleisdonk, MD PhD, Department of Surgery, Diakonessenhuis, Utrecht, The Netherlands; Jos P.A.M. Vroemen, MD PhD, Department of Surgery, Amphia Hospital, Breda, The Netherlands; Marco Waleboer, MD, Department of Surgery, Admiraal De Ruyter Hospital, Vlissingen, The Netherlands; Wietse P. Zuidema, MD, Department of Trauma Surgery, VU University Medical Center, Amsterdam, The Netherlands.

REFERENCES

1. Marsh JL, Slong TF, Agel J, Broderick JS, Creevey W, DeCoster TA, et al. Fracture and dislocation classification compendium-2007 - Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma*. 2007 Nov-Dec;21(10):S1-S133. Epub 2008/03/07.
2. Kellam JF, Augdigé L. Fracture classification. AO Foundation Publishing; [cited 2015]; Available from: <https://www2.aofoundation.org/wps/portal/surgery?showPage=diagnosis&bone=Humerus&segment=Shaft>.
3. Müller ME, Koch P, Nazarian S, Schatzker J. *The Comprehensive Classification of Fractures of Long Bones*. Berlin: Springer-Verlag; 1990.
4. Mahabier KC, Vogels LMM, Punt BJ, Roukema GR, Patka P, Van Lieshout EMM. Humeral shaft fractures: Retrospective results of non-operative and operative treatment of 186 patients. *Injury*. 2013 Apr;44(4):427-30. Epub 2012/09/04.
5. Mahabier KC, van Lieshout EMM, Bolhuis HW, Bos PK, Bronkhorst MWGA, Bruijninx MMM, et al. HUMeral Shaft Fractures: MEasuring Recovery after Operative versus Non-operative Treatment (HUMMER): a multicenter comparative observational study. *Bmc Musculoskel Dis*. 2014 Feb 11;15:39. Epub 2014/02/13.
6. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968 Oct;70(4):213-20. Epub 1968/10/01.
7. Svanholm H, Starklint H, Gundersen HJG, Fabricius J, Barlebo H, Olsen S. Reproducibility of Histomorphologic Diagnoses with Special Reference to the Kappa-Statistic. *APMIS*. 1989 Aug;97(8):689-98. Epub 1989/08/01.

8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74. Epub 1977/03/01.
9. Swiontkowski MF, Agel J, McAndrew MP, Burgess AR, MacKenzie EJ. Outcome validation of the AO/OTA fracture classification system. *J Orthop Trauma*. 2000 Nov;14(8):534-41. Epub 2001/01/10.
10. Majed A, Macleod I, Bull AMJ, Zyto K, Resch H, Hertel R, et al. Proximal humeral fracture classification systems revisited. *J Shoulder Elb Surg*. 2011 Oct;20(7):1125-32. Epub 2011/04/13.
11. Maripuri SN, Rao P, Manoj-Thomas A, Mohanty K. The classification systems for tibial plateau fractures: how reliable are they? *Injury*. 2008 Oct;39(10):1216-21. Epub 2008/04/29.
12. Matsunaga FT, Tamaoki MJ, Cordeiro EF, Uehara A, Ikawa MH, Matsumoto MH, et al. Are classifications of proximal radius fractures reproducible? *BMC Musculoskelet Disord*. 2009;10:120. Epub 2009/10/02.
13. Neuhaus V, Bot AGJ, Guitton TG, Ring DC. Scapula Fractures: Interobserver Reliability of Classification and Treatment. *J Orthop Trauma*. 2014 Mar;28(3):124-9. Epub 2013/05/01.
14. Pervez H, Parker MJ, Pryor GA, Lutchman L, Chirodian N. Classification of trochanteric fracture of the proximal femur: a study of the reliability of current systems. *Injury*. 2002 Oct;33(8):713-5. Epub 2002/09/06.
15. Szwebel JD, Ehlinger V, Pinsolle V, Bruneteau P, Pelissier P, Salmi LR. Reliability of a Classification of Fractures of the Hand Based on the Ao Comprehensive Classification System. *J Hand Surg-Eur Vol*. 2010 Jun;35E(5):392-5. Epub 2010/01/27.

16. van Embden D, Rhemrev SJ, Meylaerts SA, Roukema GR. The comparison of two classifications for trochanteric femur fractures: the AO/ASIF classification and the Jensen classification. *Injury*. 2010 Apr;41(4):377-81. Epub 2009/11/13.
17. Johnstone DJ, Radford WJP, Parnell EJ. Interobserver Variation Using the Ao Asif Classification of Long-Bone Fractures. *Injury*. 1993 Mar;24(3):163-5. Epub 1993/03/01.
18. Newey ML, Ricketts D, Roberts L. The Ao Classification of Long-Bone Fractures - an Early Study of Its Use in Clinical-Practice. *Injury*. 1993 May;24(5):309-12. Epub 1993/05/01.
19. Meling T, Harboe K, Enoksen CH, Aarflot M, Arthursson AJ, Soreide K. How reliable and accurate is the AO/OTA comprehensive classification for adult long-bone fractures? *J Trauma Acute Care Surg*. 2012 Jul;73(1):224-31. Epub 2012/06/20.
20. Garnavos C, Kanakaris NK, Lasanianos NG, Tzortzi P, West RM. New Classification System for Long-bone Fractures Supplementing the AO/OTA Classification. *Orthopedics*. 2012 May;35(5):E709-E19. Epub 2012/05/17.