Diagnostic Accuracy of Two-Dimensional and Three-Dimensional Imaging and Modeling of Radial Head Fractures

Thierry G. Guitton · Kim Brouwer · Anneluuk L. C. Lindenhovius ·
George Dyer · David Zurakowski · Chaitanya S. Mudgal · David C. Ring

Received: 14 April 2013 / Accepted: 28 October 2013 / Published online: 12 November 2013
© Society of the Hand & Microsurgeons of India 2013

Abstract To tests the hypothesis that classification and characterization of fractures of the radial head is more accurate with 3D than 2D computed tomography images and radiographs, using a prospective study design with intraoperative inspection as the reference standard. Treating surgeons and first assistants completed a questionnaire assigning a fracture type according to the Broberg and Morrey modification of Mason's classification, evaluating selected fracture characteristics, and electing preferred management based upon radiographs and 2D images alone; then adding 3D-CT; then 3D printed physical models; and finally intra-operative visualization. The addition of the 3D CT and physical models improved the sensitivity for fracture line separating the entire head from the neck, comminution of the radial neck, fracture involving the articular surface, articular fracture gap greater than 2 mm, impacted fracture fragments, greater than 3 articular

fragments, and articular fragments judged too small to repair. There were no significant differences in diagnostic performance with the addition of 3D models. The addition of 3D CT and models improved the reliability of Broberg and Morrey classification. We conclude that 3DCT and 3D physical modeling provide more accurate fracture classification and characterization of fracture of the radial head with less proposed variability in treatment. We did not demonstrate a clear advantage for modeling over 3DCT reconstructions.

Level of Evidence: Diagnostic, Level I

Keywords Three-dimensional computed tomography · Modeling · Radial head fracture · Reliability · Accuracy

Electronic supplementary material The online version of this article (doi:10.1007/s12593-013-0107-1) contains supplementary material, which is available to authorized users.

T. G. Guitton · A. L. C. Lindenhovius
Department of Orthopaedic Surgery, Academic Medical Center
Amsterdam, University of Amsterdam, Amsterdam, The Netherlands

K. Brouwer · C. S. Mudgal · D. C. Ring (✉)
Orthopaedic Hand and Upper Extremity Service, Massachusetts
General Hospital, Harvard Medical School, Yawkey Center,
Suite 2100 55 Fruit Street, Boston, MA 02114, USA
e-mail: dring@partners.org

G. Dyer
Orthopedic Surgery-Brigham & Women's Hospital,
Brigham and Women's Hospital, Harvard Medical School,
75 Francis St, Boston, MA, USA

D. Zurakowski
Departments of Anesthesiology and Surgery, Children's Hospital
Boston, Harvard Medical School, 300 Longwood Avenue, Boston,
MA 02115, USA

Introduction

Accurate preoperative radiological characterization of the fracture may facilitate management. Prior studies have demonstrated improved agreement in characterization and classification of various fractures with three-dimensional (3DCT) compared to two-dimensional computed tomography (2DCT) images and radiographs [1–9]. These studies were based upon retrospective data and the reference standard was based upon surgeon recollection and the medical record (e.g. operative notes).

Three-dimensional (3D) models that are constructed based on CT images and can be held in the hand and may facilitate fracture characterization and surgical planning. Computer-generated bone models have been used in the planning of osteotomy of multidirectional distal radius malunions [10, 11].

This investigation tests the hypothesis that 3D computed tomography images identify and predict fracture characteristics more accurately than 2D computed tomography images and radiographs, using a prospective study design with intraoperative inspection as the reference standard. A secondary hypothesis was that 3D printed, hand-held, physical models

predict fracture characteristics more accurately than 2D and 3D CT images and radiographs. We also tested interrater reliability.

Material and Methods

Under an IRB approved protocol, we prospectively included patients between 2007 and 2010 with a radial head fracture seen at two Level 1 trauma centers. Inclusion criteria were 1) fracture of the radial head; 2) election of operative treatment; 3) availability of Computed Tomography scan (CT); 4) age of 18 years or older. Exclusion criteria were pregnant women and patients unable to give informed consent. Forty-one patients satisfied the inclusion and exclusion criteria. Two patients were excluded for incomplete questionnaires, resulting in a final cohort of 39 patients.

Among the 39 patients, the mean age was 52 years (range, 23 to 92 years). There were 18 men (46 %) and 21 (54 %) women. The radial head fracture was an isolated injury in 4 patients (10 %), and was associated with an elbow dislocation in 7 (18 %) patients, an elbow dislocation and coronoid fracture (the so-called terrible triad injury) in 15 patients (38 %), a posterior-olecranon-fracture dislocation (POFD) in 8 patients (21 %), metaphyseal fracture of the proximal ulna (posterior Monteggia fracture) in 1 patient (3 %), a complex fracture of the distal humerus in 1 patient (3 %), an Essex-Lopresti lesion in 1 patient (3 %), an anterior-transolecranon-fracture dislocation in 1 patient (3 %), and a capitellum/trochlea fracture in 1 patient (3 %). The left side was injured in 22 patients (56 %) and the right side in 17 patients (44 %). Twenty-three patients (59 %) fractured their elbow in a fall from a standing height, 13 (33 %) from a greater height, 2 (5 %) patients in a motor vehicle collision (MVC) and one (3 %) in a crush injury.

Several different CT scanners were used with up to 140Kv and 500–700 mAs and slices from 8 to 64/Dual Source. Two-dimensional images were created in axial, coronal and sagittal planes. Three-dimensional images were created with rotations in both the sagittal and coronal planes with and without the ulna and the distal humerus extracted from the image.

3DCT reconstructions were ordered for all patients. Computed Tomography scans were sent to Medical Modeling LLC (Golden, CO, USA) for printing of the 3D physical hand-held models. These were plastic models of the bones made using a three-dimensional printer. The resolution of the 3D physical models was 1:1.

The treating surgeons (7 surgeons and 27 resident or fellow first assistants) completed a questionnaire assigning a fracture type according to the Broberg and Morrey modification of Mason's classification and important fracture characteristics and management. Broberg and Morrey modified Mason's

classification as follows: Type 1 fractures involve less than 30 % of the articular surface or are displaced fewer than 2 mm; Type 2 fractures are partial head fractures involving at least 30 % of the articular surface and displaced at least 2 mm; Type 3 fractures are displaced articular fractures involving the entire head of the radius; and Type 4 fractures have an associated elbow dislocation [12]. A standardized graphic instruction about the Broberg and Morrey modification of Mason's classification system was used. There was no increase in radiation exposure to achieve the detail needed for the 3D modeling.

The questionnaire was completed four times: initially based upon radiographs and 2D images alone; a second time based on radiographs, 2D and 3D-CT images; a third time on radiographs, 2D, 3D-CT and 3D physical models; and a fourth time based on intra-operative visualization of the fracture characteristics. For practical reasons this was usually done at a single time, after surgery. The fourth questionnaire completed by the surgeon represented the reference standard. Both the surgeon and the first assistant rated the fractures, allowing us to calculate interobserver agreement. Sensitivity and specificity were calculated for 2DCT and radiographs, 3DCT, and 3D physical models as compared to the intraoperative direct observation of the surgeon. Observers only evaluated 2D and 3D-CT images and 3D physical models from cases in which they participated.

Interobserver agreement regarding fracture characteristics and treatment proposal was measured for each method by the chance-corrected kappa (κ) coefficient with strength of agreement assessed using the benchmarks of Landis and Koch [13]. Logistic regression was applied using a generalized estimating equations (GEE) strategy in order to account for the same 39 cases evaluated by multiple surgeons using each of 4 different methods (2D, 2D/3D, 2D/3D with physical model, direct operative view) with a binomial distribution used for binary yes/no fracture characteristics and a multinomial logit distribution for Broberg-Morrey classification (Types I–IV) and treatment plan (5 options: nonoperative management; ORIF with wires, screws or pins; ORIF with plate and screws; radial head excision; radial head replacement/arthroplasty). Differences between the methods were determined using the maximum likelihood Wald chi-square test with a two-tailed $p < 0.05$ as the criterion for statistical significance [14]. Power analysis revealed that a minimum sample size of 30 fractures would provide 80 % power ($\alpha = 0.05$, $\beta = 0.20$) to detect significant intra- and inter-observer agreement using the kappa coefficient [13] as well as in comparing diagnostic characteristics between the two imaging modalities.

Sensitivity, specificity, and accuracy for detection of each of the fracture characteristics and type of treatment with two-dimensional images, 3D reconstructions and 3D Model was calculated with the intra operative findings of the attending surgeon as the gold standard. The statistical significance of

these differences was evaluated using McNemar's test for paired binary data [15]. Statistical analysis was performed using SPSS version 18.0 (SPSS Inc./IBM, Chicago, IL, USA).

No funding was received in direct support of this study. Medical Modeling LLC (Golden, CO) provided free 3D physical models by an agreement approved by our Human Research Committee and Research Contracting Department.

Results

The addition of 3DCT and the 3D models to 2DCT and radiographs led to significant improvements in sensitivity for diagnosis of fracture line separation of the entire articular surface from the radial neck, comminution of the radial neck, involvement of the articular surface, articular gap or step of 2 mm or greater, central impaction of the articular surface, presence of more than 3 articular fragments and to the presence of articular fragments too small to repair (all $p < 0.01$, Table 1). There were no significant differences with the addition of 3D models. There were no significant changes in specificity with more sophisticated imaging. (Figs. 1, 2 and 3).

Compared to standard radiographs and 2D CT scans, the addition of 3DCT reconstructions and 3D models improved the interobserver reliability of fracture classification according to the Broberg and Morrey modification of the Mason classification, diagnosis of comminution of the radial neck, involvement of the articular surface, articular gap or step of 2 mm or greater, central impaction of the articular surface, presence of more than 3 articular fragments, presence of articular fragments too small to repair, and proposed treatment (Table 2). Compared to 3DCT, the addition of models led to a significant

Table 1 Interobserver agreement for classification and treatment of radial head and neck fractures for each CT method and direct operative view

Characteristic	2DCT	2D3D CT	2D3D CT With Model	Operative View (Gold Standard)
Broberg-Morrey	0.23*	0.26*	0.37*	0.38*
Fracture line	0.69†	0.54†	0.59†	0.54†
Comminution	0.31*	0.48†	0.57†	0.40*
Articular surface	0.12	0.28	0.22	0.28
Gap >2 mm	0.37*	0.59†	0.57†	0.37*
Impaction	0.17	0.24	0.12	0.23
>3 fragments	0.29	0.50†	0.64†	0.57†
Small fragments	0.26	0.34*	0.33*	0.42*
Proposed Treatment	0.47†	0.53†	0.67†	0.85†

Data are kappa (κ) values based on 39 cases evaluated by two independent surgeons. Guidelines for strength of observer agreement: $\kappa = 0-0.20$ (slight), $\kappa = 0.21-0.40$ (fair), $\kappa = 0.41-0.60$ (moderate), $\kappa = 0.61-0.80$ substantial, $\kappa = 0.81-1.00$ almost perfect. Significant interobserver agreement beyond chance level (* $p < 0.05$; † $p < 0.01$)

Fig. 1 Agreement on fracture characteristics by 2D-CT and 2D/3D-CT

difference in agreement on treatment ($p = 0.046$), but did not affect the other factors.

Discussion

We have studied the use of 3DCT for fracture classification before [9, 16–18], but never prospectively for fractures of the radial head. We also had the opportunity to look at the influence of physical models. We found that more sophisticated images and models improve accuracy.

The strengths of this investigation include the prospective design, the relatively large number of patients, and an intra-operative reference standard. The limitations of this investigation include the fact that images were usually rated after surgery (in part due to the inherent delay in receiving the physical 3D model), so that ratings of the radiological images were—in essence—retrospective; the injuries were relatively complex resulting in a spectrum bias in terms of all fractures of the radial head, although our work is representative of the

Fig. 2 Agreement on treatment proposal stratified by 2D-CT and 2D/3D-CT

Fig. 3 Sensitivity and specificity according to Broberg and Morrey modification of the Mason classification stratified by 2D-CT, 2D/3D-CT and 2D/3D-CT with model

types of fractures that would be studied with CT and operated on; two patients (one with addition of a capitulum/trochlea fracture and one anterior-transolecranon fracture dislocation) had non-nondisplaced fractures of the radial neck, which are relatively unusual – both fractures were seen only on intraoperative exposure; and multiple physicians were involved in the ratings at two sites, which makes the results more generalizable, but less consistent. These data should also be interpreted in light of the fact that the first assistant was usually a resident or fellow, so that the observer variability may largely reflect differences in training and experience. The low reliability for even the intra-operative view may be a result of the kappa paradox, where substantial agreement results in a low kappa due to the effect of low numbers of certain observations [16]. In addition, we did not measure intra-observer reliability because it adds a great deal more effort and it's always greater than interobserver reliability. Finally, ideally we should not have used Broberg and Morrey Type 4—we should have

forced surgeons to classify the radial head fracture independent of associated injuries such as dislocation.

It is not always feasible to have models available prior to operative treatment at this point, but three-dimensional reconstructions of computed tomography images can be easily produced by orthopaedic surgeons using the DICOM (Digital Imaging and Communications in Medicine) files from the patient's CT scan. Three-dimensional reconstructions are made from CT-scans and therefore do not require additional scanning and do not expose the patient to additional radiation. It has been calculated at the investigators institution, that the cost for additional 3D reconstructions are an additional 20 % of the cost of a CT-scan. Free software such as OsiriX [19] is available which makes it possible for every orthopedic surgeon to quickly and easily create 3D reconstructions themselves with minimal training. Contrast this with the cost of the models, which was estimated at approximately US\$2,000 at the start of the study in 2007, but that is now the price for some 3D printers capable of making the models, and the models can be outsourced to commercial printing companies for under \$100 and likely will become even less expensive. The models offered little advantage over 3DCT images in this study of radial head fractures, but might prove useful for specific complex fractures or reconstructions. With respect to the radial head, use of 3D imaging techniques to plan the surgery might be particularly helpful for surgeons that do not have implants for fixing or replacing radial head fractures readily available.

This study found that increasing levels of sophistication in imaging/modeling: 1) improved the sensitivity for diagnosis of numerous fracture characteristics using the surgeon's interpretation of the intraoperative findings as the reference standard; and 2) decreased observer variation between surgeon and first assistant. The greatest improvements were observed in the jump between two and three-dimensional imaging. This

Table 2 Sensitivity and specificity characteristics for 2D and 3D CT methods

Variable	2D CT Alone		2D3D CT		2D3D CT with Model	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Fracture Line	85 (91–99)	100 (90–100)	95* (78–100)	94 (78–100)	95* (81–100)	94 (78–100)
Comminution	58 (31–83)	93 (79–99)	83* (57–97)	96 (85–100)	92* (68–100)	93 (79–99)
Articular Surface	94 (84–99)	67 (16–98)	97 (88–100)	67 (16–98)	97 (88–100)	67 (16–98)
Gap >2 mm	93 (81–99)	78 (46–96)	97 (86–100)	78 (46–96)	97 (86–100)	89 (59–99)
Impaction	57 (32–80)	88 (72–97)	86* (62–97)	84 (67–95)	93* (72–100)	96 (84–100)
>3 Fragments	55 (34–74)	82 (60–95)	100* (90–100)	85 (66–96)	91* (74–98)	100 (89–100)
Small Fragments	74 (56–88)	83 (57–97)	100* (93–100)	67 (39–88)	100* (93–100)	75 (47–93)
Proposed Treatment	74 (56–88)	83 (57–97)	89* (74–97)	75 (47–93)	100* (93–100)	75 (47–93)

Values are percentages with results are based on attending surgeon for each method compared to intraoperative direct view gold standard ($N=39$ paired cases)

* Statistically significant compared to sensitivity for 2D CT alone (all $p < 0.01$). There were no significant differences were detected in specificity between the three methods

is concordant with prior studies that have demonstrated improved agreement in characterization and classification of fractures with 3DCT compared to 2DCT and radiographs alone [1–9]. Prior studies that addressed the classification of radial head fractures specifically found substantial observer variation [20–23] when fractures were evaluated by radiographs only. However, these studies differed in that they are based upon retrospective data in small groups of observers/patients and the reference standard was based upon surgeon recollection and the medical record (e.g. operative notes). Three-dimensional images are more intuitive than two-dimensional images—they look more like the actual bone on operative exposure. The models add the ability to feel and manipulate the bones.

We interpret this combination of findings to indicate that fracture classification and characterization based on three-dimensional imaging and models is more accurate and reliable, essentially helping to narrow the experience and training gap. While some surgeons suggest that there is marginal utility in more sophisticated imaging, our data show that there is a benefit to these advanced imaging techniques. However, recommendations regarding the use of a new technology should be based on both diagnostic performance characteristics and clinical impact. The next steps are to investigate whether more sophisticated imaging leads to more effective treatment as measured by fewer complications with less functional impairment.

Disclaimer Research Specific Grants: Stryker; Biomet
Honoraria/Travel: AO North America, AO International
Consultant: Wright Medical, Skeletal Dynamics, Biomet
Royalties: Wright Medical; Skeletal Dynamics; Biomet
Stock Options: Illuminos, Inc

IRB: The Massachusetts General Hospital IRB has approved the human protocol for this investigation under number 2006-P-000869/16.

References

1. Borrelli J Jr, Goldfarb C, Catalano L, Evanoff BA (2002) Assessment of articular fragment displacement in acetabular fractures: a comparison of computerized tomography and plain radiographs. *J Orthop Trauma* 16(7):449–456, discussion 56–7
2. Castagno AA, Shuman WP, Kilcoyne RF, Haynor DR, Morris ME, Matsen FA (1987) Complex fractures of the proximal humerus: role of CT in treatment. *Radiology* 165(3):759–762
3. Edelson G, Kelly I, Vigder F, Reis ND (2004) A three-dimensional classification for fractures of the proximal humerus. *J Bone Joint Surg Br* 86(3):413–425
4. Katz MA, Beredjikian PK, Bozentka DJ, Steinberg DR (2001) Computed tomography scanning of intra-articular distal radius fractures: does it influence treatment? *J Hand Surg [Am]* 26(3):415–421
5. Kilcoyne RF, Shuman WP, Matsen FA 3rd, Morris M, Rockwood CA (1990) The Neer classification of displaced proximal humeral fractures: spectrum of findings on plain radiographs and CT scans. *AJR Am J Roentgenol* 154(5):1029–1033
6. Kuhlman JE, Fishman EK, Ney DR, Magid D (1989) Two- and three-dimensional imaging of the painful shoulder. *Orthop Rev* 18(11):1201–1208
7. Siebenrock KA, Gerber C (1993) The reproducibility of classification of fractures of the proximal end of the humerus. *J Bone Joint Surg Am* 75(12):1751–1755
8. Smith RM (2000) The classification of fractures. *J Bone Joint Surg Br* 82(5):625–626
9. Lindenhovius A, Karanicolas PJ, Bhandari M, van Dijk N, Ring D (2009) Interobserver reliability of coronoid fracture classification: two-dimensional versus three-dimensional computed tomography. *J Hand Surg [Am]* 34(9):1640–1646
10. Jupiter JB, Ruder J, Roth DA (1992) Computer-generated bone models in the planning of osteotomy of multidirectional distal radius malunions. *J Hand Surg [Am]* 17(3):406–415
11. Murase T, Oka K, Moritomo H, Goto A, Yoshikawa H, Sugamoto K (2008) Three-dimensional corrective osteotomy of malunited fractures of the upper extremity with use of a computer simulation system. *J Bone Joint Surg Am* 90(11):2375–2389
12. Broberg MA, Morrey BF (1987) Results of treatment of fracture-dislocations of the elbow. *Clin Orthop Relat Res* 216:109–119
13. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159–174
14. Harrell F (2001) Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. Springer, New York, pp 331–342
15. Kocher MS, Zurakowski D (2004) Clinical epidemiology and biostatistics: a primer for orthopaedic surgeons. *J Bone Joint Surg Am* 86-A(3):607–620
16. Doornberg J, Lindenhovius A, Kloen P, van Dijk CN, Zurakowski D, Ring D (2006) Two and three-dimensional computed tomography for the classification and management of distal humeral fractures. Evaluation of reliability and diagnostic accuracy. *J Bone Joint Surg Am* 88(8):1795–1801
17. Harness NG, Ring D, Zurakowski D, Harris GJ, Jupiter JB (2006) The influence of three-dimensional computed tomography reconstructions on the characterization and treatment of distal radial fractures. *J Bone Joint Surg Am* 88(6):1315–1323
18. Guitton TG, Ring D, Science of Variation Group (2011) Interobserver reliability of radial head fracture classification: two-dimensional compared with three-dimensional CT. *J Bone Joint Surg Am* 93(21):2015–2021
19. Rosset A, Spadola L, Ratib O (2004) OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images. *Journal of Digital Imaging* 17(3):205–16
20. Doornberg J, Elsner A, Kloen P, Marti RK, van Dijk CN, Ring D (2007) Apparently isolated partial articular fractures of the radial head: prevalence and reliability of radiographically diagnosed displacement. *J Shoulder Elbow Surg* 16(5):603–608
21. Matsunaga FT, Tamaoki MJ, Cordeiro EF, Uehara A, Ikawa MH, Matsumoto MH et al (2009) Are classifications of proximal radius fractures reproducible? *BMC Musculoskelet Disord* 10:120
22. Morgan SJ, Groshen SL, Itamura JM, Shankwiler J, Brien WW, Kuschner SH (1997) Reliability evaluation of classifying radial head fractures by the system of Mason. *Bull Hosp Jt Dis* 56(2):95–98
23. Sheps DM, Kiefer KR, Boorman RS, Donaghy J, Lalani A, Walker R et al (2009) The interobserver reliability of classification systems for radial head fractures: the Hotchkiss modification of the Mason classification and the AO classification systems. *Can J Surg* 52(4):277–282