

Inter-observer variation in the diagnosis of coronal articular fracture lines in the lunate facet of the distal radius

Mathieu M. E. Wijffels · Thierry G. Guitton ·
David Ring · Science of Variation Group

Published online: 26 May 2012
© American Association for Hand Surgery 2012

Abstract

Background Several studies support the use of CT for diagnosing coronal fractures of the distal radius but the inter-observer reliability of these observations is less well studied. We tested the null hypothesis that radiographs alone and the combination of radiographs and two-dimensional computed tomography scans (2DCT) have the same inter-observer variation for the diagnosis of coronal articular fracture lines in the distal radius.

Methods Using a web-based survey, 63 surgeons were randomized to evaluate 16 fractures of the distal radius on radiographs alone or radiographs and 2DCT for the presence or absence of a coronal fracture line of the lunate facet and, if present, the stability of the fracture. The kappa multirater measure was calculated to estimate agreement between observers.

Results The inter-observer variation in diagnosis of a coronal fracture line was fair with both radiographs and 2DCT, as was the diagnosis of instability of the volar lunate facet fracture when present.

The Science of Variation group is composed of Ponsen KJ., Rocha S., Athwal G., Elmans L., Fanuele J., Sodha S., Swigart C., Barquet A., Cassidy C., Papandrea R., Shyam A., de Bedout R., Nolla J., Richard MJ., Hanel D., Poolman R., Wright T., Allan C., Rhemrev S., Gosens T., Baskies M., Grosso E., Swiontkowski M., Harris I., Boretto J., Brink P., Biert J., Osterman A. L., Schep N., Coles CP., Feibel RJ., Abrams R., Hughes T., Nelissen R., Ladd A., McAuliffe J., Duncan S., Beingessner D., Conflitti JM., Goslings JC., Crist B., Frihagen F., Dyer G., Page RS., Zalavras C., Jeray K., Jebson P., Thomas G., Kalainov D., Hammerberg EM., Chen N., Pesantez R., Taras J., Goldfarb C., Prayson M., Zura R., Richardson M., Boyer M., Ilyas A., Farsr., Kloen P., Evans PJ., and Segalman K.

M. M. E. Wijffels · T. G. Guitton · D. Ring (✉)
Orthopaedic Hand and Upper Extremity Service,
Harvard Medical School, Massachusetts General Hospital,
Yawkey Center, Suite 2100, 55 Fruit Street,
Boston, MA 02114, USA
e-mail: dring@partners.org

Conclusion Two-dimensional computed tomography does not improve observer agreement on the diagnosis of coronal plane articular fracture lines in the lunate facet of the distal radius.

Keywords Distal radius fracture · Radiographs · Computed tomography · Coronal plane fracture line

Introduction

Melone emphasized the importance of a coronal split in the lunate facet of the distal radius, suggesting that an unstable volar fragment would require open reduction and internal fixation and could not be managed percutaneously. [6] Such fractures also involve the distal radioulnar joint. [1, 3]

It has been suggested that computed tomography aids identification of coronal fracture lines. Rozental and colleagues identified an articular fracture involving the distal radioulnar joint (DRUJ) on 13 of 20 CT scans, but only on seven of 20 radiographs. [4, 9, 10] These data support the use of CT for diagnosing coronal fractures of the distal radius, but the inter-observer reliability of these observations has not been studied.

A new collaboration motivated to better understand inter-observer variation consists of observers that have completed all training and are independently treating patients (www.scienceofvariationgroup.nl). This provides an opportunity to further investigate inter-observer variability and how to reduce it.

This study tested the null hypothesis that the combination of radiographs and two-dimensional computed tomography scans (2DCT) has the same inter-observer variation as radiographs alone in the diagnosis of coronal articular fracture lines.



Fig. 1 Example of radiograph with CT-scan uploaded to the research group's website. These radiographs and CT scan were evaluated as a coronal fracture being present by all raters, but only 29 of 31 raters, only using radiographs, evaluated a coronal fracture being present

Materials and methods

Study design

Independent observers (all orthopedic surgeons) from several countries were invited to evaluate 16 distal radius fractures in an online survey. Participants were randomly assigned on a 1:1 basis to review either radiographs alone or radiographs and 2DCT to identify a coronal fracture of the lunate facet and to determine its stability. The study was performed under a protocol approved by the Institutional Research Board at the principal investigators hospital.

The Science of Variation group uses internet survey software to facilitate large international inter-observer studies. With multiple fully trained surgeons from diverse countries and institutions participating in studies, this

approach should provide a powerful and externally valid forum for studying, understanding, and ultimately reducing inter-observer variation in aspects of patient care.

Observers

A total of 206 surgeons were invited via e-mail to join the Science of Variation Group. Other than an acknowledgement as part of the author collaborative in the paper, no incentives were provided. Sixty-three surgeons were interested in participation and logged on to the website. Thirty-three surgeons were randomized to conventional radiographs and 30 to radiographs with additional 2DCT scans using a computer random number generator. Group members that agreed to participate were sent the appropriate link after randomization. Three surgeons were excluded because

of incomplete answers; two in the CT group and one in the radiograph group. Four weekly reminders to complete the online survey were e-mailed. This study presents an analysis of the 60 observers that completed the study; 31 in the radiograph group and 29 in the radiograph with 2DCT group.

Fractures

Radiographs and computed tomography scans of distal radius fractures were identified from a consecutive list of 355 cases treated by the senior investigator between 2001 and 2006 at one level-1 trauma center. Inclusion criteria were (1) articular fracture of the distal radius fractures with adequate quality radiographs (posteroanterior and lateral views with no plaster or other obscuring features) and 2DCT (slice thickness less than 1 mm); and (2) Age 18 or older. Radiographs and CT scans of 16 articular fractures of the distal radius, eight with a coronal articular fracture line in the lunate facet, were blinded by an independent research fellow for use in this study. The fractures were all C-type fractures classified according to the Comprehensive Classification of Fractures. [7]

For each case, videos with 2DCT images along the sagittal and axial axes were created. Observers could scroll through the videos or play them automatically. Radiographs and 2DCT scans were uploaded to the research group’s website (Fig. 1).

Evaluation

Observers logged in independently on the website. Upon login to the website, they were asked the number of distal radius fractures treated per year. Subsequently, observers were asked for (1) the presence of an intra-articular fracture line in the coronal plane and (2) if present, is this fracture unstable?

Observers were blinded to clinical information. Observers had the option to comment on each case and all questions had to be completed in order to continue with the next case. The observers completed the study at their own time and pace.

Statistical analysis

The multirater kappa measure (κ) was used to estimate agreement among surgeons with respect to (1) presence of a fracture in the coronal plane and (2) the stability of this fracture. It is a commonly used statistic to describe chance-corrected agreement in a variety of intra-observer and inter-observer studies. [2, 5, 8] Agreement among observers was calculated with use of the multirater kappa measure described by Siegel and Castellan. [11] Kappa values were interpreted using the guidelines proposed by Landis and Koch

[5]; values of 0.01–0.20 indicate slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and more than 0.81, almost perfect agreement. Zero indicates no agreement beyond that expected due to chance alone, –1.00 means total disagreement, and +1.00 represents perfect agreement. [5, 8] The agreement based on radiographs (κ_R) was compared to the agreement based on radiographs with 2DCT (κ_{CT}) using *Z* tests, which assume that the two samples are independent. Since the samples compared in this study were not independent (the same set of patients were rated by the radiographs and radiographs with 2DCT group), this method produced conservative estimates of the *p* values. The study had 80 % power to detect a difference in kappa scores of 0.1 between results based on radiographs vs. radiographs with 2DCT.

Table 1 Observer demographics

	Radiographs (N=31)		Radiographs+ 2DCT (N=29)		Total (N=60)	
	N	%	N	%	N	%
Gender						
Male	29	94	26	90	55	92
Female	2	6	3	10	5	8
Practice						
Asia	1	3	1	3	2	4
Australia	2	6	1	3	3	5
Canada	0	0	3	10	3	5
Europe	4	13	8	28	12	20
UK	0	0	0	0	0	0
USA	22	71	13	45	35	58
other	2	6	3	10	5	8
Years in practice						
0–5	7	23	12	41	19	32
6–10	9	29	3	10	12	20
11–20	13	42	8	28	21	35
21–30	2	6	6	21	8	13
Supervise						
Yes	29	94	26	90	55	92
No	2	6	3	10	5	8
Fractures per year						
0–5	3	10	0	0	3	5
6–10	2	6	1	3	3	5
11–20	2	6	6	21	8	13
>20	24	78	22	76	46	77
Specialization						
General orthopedics	0	0	1	3	1	2
Orthopedic traumatology	11	35	11	38	22	37
Shoulder and elbow	7	23	3	10	10	17
Hand and wrist	10	32	12	41	22	37
Other	3	10	2	7	5	8

Table 2 Fracture line in the coronal plate: differences in inter-observer variability

	Radiographs				2-Dimensional CT				P value
	N	Categorical	κ	SE	N	Categorical	κ	SE	
Specialty									
Orthopedic traumatology	11	Fair	0.36	0.07	11	Slight	0.15	0.10	0.12
Shoulder and elbow	7	Slight	0.15	0.13	3	Substantial	0.67	0.22	0.04*
Hand	10	Fair	0.32	0.11	12	Slight	0.20	0.13	0.52
Other	3	Fair	0.38	0.16	2	Poor	0.19	0.42	0.21

N Number of Observers, SE standard error

* $P < 0.05$ (Significant)

Results

Surgeons

Thirty-five (58 %) observers practiced in USA, 12 (20 %) in Continental Europe, three (5 %) in Canada, three (5 %) in Australia, two (3 %) in Asia, and five (8 %) in Latin America. Nineteen (32 %) observers had been in independent practice for fewer than 5 years, 12 (20 %) for 5–10 years, 21 (35 %) for 11–20 years, and eight (13 %) for more than 21 years. Fifty-five (92 %) observers supervised surgical trainees in the operating room. Most observers (77 %) treated more than 20 distal radius fractures per year. One (2 %) observer was a general orthopedic surgeon, 22 (37 %) were orthopedic trauma surgeons, 10 (17 %) were shoulder and elbow surgeons, 22 (37 %) were hand and wrist surgeons, and five (8 %) were other surgeons (Table 1).

Inter-observer reliability

Presence of a fracture in the coronal plane

In the X ray group, a total of 364 coronal fractures were found, while 372 coronal fractures of the distal radius were found in the CT-group. The inter-observer variation for the presence of an articular fracture of the

lunate facet in the coronal plane was fair both with use of radiographs alone and with use of radiographs and 2DCT ($\kappa_R = 0.29$, SE 0.05 and $\kappa_{CT} = 0.24$, SE 0.07; $p = 0.55$).

Instability of the coronal plane fracture

In the X-ray group, 228 of the 364 coronal fractures were found unstable. In the CT group, 232 of the 372 coronal fractures of the distal radius were found unstable. Inter-observer agreement on instability of the coronal split fracture was fair for both radiographs and radiographs with 2DCT ($\kappa_R = 0.29$, SE 0.02, and $\kappa_{CT} = 0.32$, SE 0.02, $p = 0.28$).

Observer demographics

With respect to the presence of a coronal fracture line in the lunate facet, the only significant finding with respect to observer demographics was that shoulder and elbow specialists had better agreement with 2DCT than radiographs (Table 2). With respect of instability of the lunate facet fracture, the only significant findings were that hand and wrist surgeons had a significant greater agreement than other specialists, and experienced surgeons had greater agreement than less experienced surgeons (Table 3).

Table 3 Fracture instability: differences in inter-observer variability

	N	Categorical	κ^1	SE ¹	N	Categorical	κ^2	SE ²	P value
Fractures per year									
10 or less vs. 11 or more	5	Slight	0.16	0.06	26	Fair	0.31	0.02	0.01
10 or less vs. 20 or more	5	Slight	0.16	0.06	24	Fair	0.32	0.02	0.01
Specialty									
Trauma vs. Hand	11	Fair	0.25	0.03	10	Fair	0.36	0.04	0.04
Hand vs. non Hand	10	Fair	0.36	0.04	21	Fair	0.27	0.02	0.04

N Number of observers, SE standard error

Discussion

Inter-observer agreement regarding the diagnosis of a coronal articular fracture of the lunate facet of the distal radius and instability of that fracture when present are both fair. Additional 2DCT did not improve agreement over radiographs alone on average. These findings were quite consistent across training, experience, and region. The only positive findings were that observers that treat more than 10 distal radius fractures a year and specialized hand and wrist surgeons seem more consistent in defining a fracture as unstable, but diagnosis of stability was a secondary study question and the significant findings must be interpreted in the context of multiple testing.

We are not aware of any prior studies of the reliability of diagnosing coronal plane fractures of the lunate facet, but three prior studies address diagnosis of articular fracture lines with CT compared to radiographs. In one study of suspected distal radius fractures, CT scans demonstrated a fracture in 22 patients, only 19 of which were visible on radiographs (16 intra-articular) [4]. In another study of 18 intra-articular distal radius fractures, CT scans demonstrated seven fractures involving the DRUJ, five involving the radial styloid that were not seen on radiographs [9]. The third study identified an articular fracture involving the DRUJ on 13 of 20 CT scans, but only seven of 20 radiographs [10]. None of these studies had a reference standard for true fracture such as intraoperative visualization of the articular surface.

This study should be interpreted in light of its potential shortcomings: (1) We did not provide coronal plane images; (2) the observers may have interpreted the question differently when faced with a multifragmented fracture, with some only applying that diagnosis to a single coronal split fracture; (3) the observers were not provided any clinical data (e.g., age, sex, mechanism); (4) the quality of the radiographs was limited to what had been obtained at the time of injury, which reflects usual practice, but not what might be achieved with specific protocols; (5) there was a spectrum bias by selecting cases to represent the known variety of injuries, with the result that less common complex fractures were over represented compared to the more common minimally or slightly displaced fractures; (6) we asked numerous secondary study questions, all of which should be regarded as hypothesis generating and not as robust as the primary study question; (7) we don't know how carefully each observer studied the

radiographs and scans; and (8) having only 16 patients may limit the external validity of the findings. However, according to statisticians, the power of a reliability study is determined by the total number of observations, which given the number of observers in this study was quite high.

This study does not support the use of CT scans for the diagnosis of coronal articular fracture lines in the lunate facet of the distal radius, at least in terms of reliability. This study did not evaluate the diagnostic performance characteristics of CT or radiographs for coronal plane fractures. We believe that future studies should attempt to clarify the definition of these fracture lines and further address the role of experience and training in observer variation.

Conflict of interest M.M.E. Wijffels and T.G. Guitton have no conflicts to declare. D. Ring is a consultant to and Royalties contracted with Wright, Skeletal Dynamics, and Biomet; research contract with Biomet; stock options in Illuminoss.

References

1. Bowers WF. The distal radioulnar joint. In: Hucherson DC, editor. *Operative hand surgery*. 3rd ed. New York: Churchill Livingstone; 1993. p. 973–1019.
2. Cohen J. A coefficient of agreement for nominal scales. *Edoc Psychol Meas*. 1960;20:37–46.
3. Hucherson DC. The Darrach operation for lower radioulnar derangement. *Am J Surg*. 1941;53:237–41.
4. Johnston GHF, Friedman L, Kriegler JC. Computerized tomographic evaluation of acute distal radial fractures. *J Hand Surg*. 1992;17A:737–44.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biom*. 1977;33:159–74.
6. Melone Jr CP. Articular fractures of the distal radius. *Orthop Clin N Am*. 1984;15:217–36.
7. Müller ME, Allgöwer M, Schneider R, et al. *Manual of internal fixation*. 3rd ed. New York: Springer; 1991.
8. Posner KL, Sampson PD, Caplan RA, et al. Measuring interrater reliability among multiple raters: an example of methods for nominal data. *Stat Med*. 1990;9:1103–15.
9. Pruitt DL, Gilula LA, Manske PR, et al. Computed tomography scanning with image reconstruction in evaluation of distal radius fractures. *J Hand Surg*. 1994;19A:720–7.
10. Rozental TD, Bozentka DJ, Katz MA, et al. Evaluation of the sigmoid notch with computed tomography following intra-articular distal radius fracture. *J Hand Surg [Am]*. 2001;26A:244–51.
11. Siegel S, Castellan NJ. *Nonparametric statistics for the behavioral sciences*. 2nd ed. New York: McGraw-Hill; 1988.