Clin Orthop Relat Res (2013) 471:1373–1378
DOI 10.1007/s11999-012-2742-4

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CLINICAL RESEARCH

Diagnosis of Elbow Fracture Patterns on Radiographs: Interobserver Reliability and Diagnostic Accuracy

Job N. Doornberg MD, PhD, Thierry G. Guitton PhD, David Ring MD, PhD, Science of Variation Group

Received: 17 May 2012 / Accepted: 3 December 2012 / Published online: 18 December 2012

Abstract

Background Studies of traumatic elbow instability suggest that recognition of a pattern in the combination and character of the fractures and joint displacements helps predict soft tissue injury and guide the treatment of traumatic elbow instability, but there is no evidence that patterns can be identified reliably.

Questions/Purposes We therefore determined (1) the interobserver reliability of identifying specific patterns of traumatic elbow instability on radiographs for subgroups of orthopaedic surgeons; and (2) the diagnostic accuracy of radiographic diagnosis.

Methods Seventy-three orthopaedic surgeons evaluated 53 sets of radiographs and diagnosed one of five common patterns of traumatic elbow instability by using a web-based survey. The interobserver reliability was analyzed using Cohen's multirater kappa. Intraoperative diagnosis was the reference for fracture pattern in calculations of the sensitivity, specificity, accuracy, and positive predictive and negative predictive values of radiographic diagnosis.

Results The overall interobserver reliability for patterns of traumatic elbow instability on radiographs was $\kappa = 0.41$. Treatment of greater than five such injuries a year was associated with greater interobserver agreement, but years in practice were not. Diagnostic accuracy ranged from 76% to 93% and was lowest for the terrible triad pattern of injury.

Conclusions Specific patterns of traumatic elbow instability can be diagnosed with moderate interobserver reliability and reasonable accuracy on radiographs.

Level of Evidence Level III, diagnostic study. See Guidelines for Authors for a complete description of levels of evidence.

Science of Variation Group: Andrew Schmidt, Antonio Barquet, Ashok K. Shyam, Brett D. Crist, Carrie Swigart, Chad Coles, Charalampos Zalavras, Charles A. Goldfarb, Charles Cassidy, Daphne Beingessner, David M. Kalainov, David Ruchelsman, Denise Eygendaal, Elena Grosso, Eric Mark Hammerberg, Frede Frihagen, George Dyer, George S. Athwal, George Thomas, Gregory J. Della Rocca, Ian Harris, J.C. Goslings, Jan Biert, Jason Fanuele, Jeff Lawton, Jeffrey A. Greenberg, John Jiuliano, John McAuliffe, John T. Capo, John Taras, Jorge Orbay, Joseph M. Conflitti, Keith Segalman, Kenneth Egol, Kyle Jeray, Lee Osterman, Leon Elmans, Lisa Lattanza, Lisa Taitsman, Louis Catalano III, M Baumgaertner, Marc Swiontkowski, Martin Boyer, Martin Richardson, Matthew Camuso, Michael Baskies, Michael Prayson, Neal Chen, P.V. van Eerten, Parag Sancheti, Paul Levin, Paul Tornetta, Peter Brink, Peter J. Evans, Peter Jebson, Peter Kloen, Phani Dantuluri, Philip Blazar, Richard S. Page, Rick Papandrea, Rob Nelissen, Robert D. Zura, Robert J. Feibel MD, Robert Probe, Rodrigo Pesantez, S. Rhemrev, Samir Sodha, Scott Duncan, Scott Wolfe, Taco Gosens, Thomas Higgins, Thomas Hughes, Thomas Wright.
One of the authors certifies that he (DR) has or may receive payments or benefits, during the study period, an amount of less than $10,000 from Wright Medical (DR) (Arlington, TN, USA) and Biomet (DR) (Warsaw, IN, USA). The institution of one or more of the authors (DR) has received funding from Skeletal Dynamics (DR)(Miami, FL, USA) and Biomet (DR)(Warsaw, IN, USA).

Introduction

Numerous studies indicate traumatic elbow instability occurs in specific patterns [3, 5, 9, 12, 17]. Some authors [3, 5, 9, 12, 17] suggest that identifying the pattern of injury (ie, the combination and character of the fractures and joint displacements) helps to predict soft tissue injury and morphologic features of fractures, both of which may guide treatment. For examples, specific coronoid fracture types are associated with injury pattern [3]; radial head fractures occur in posterior, but not anterior olecranon fracture dislocations [2]; the anterior band of the medial collateral ligament remains intact when apposition of the articular surfaces is maintained (disruption or subluxation, rather than dislocation of the elbow) [2]; lateral collateral ligament injury occurs in most varus posteromedial rotational instability injuries [4, 8, 10]—approximately ½ of posterior olecranon fracture dislocations and rarely among anterior olecranon fracture dislocations [2, 13]. These studies are limited to cases series from dedicated elbow surgeons [2, 3, 7, 9, 14, 15, 18]. To date, there are no studies investigating the reliability and diagnostic accuracy of these injury patterns and it is unclear how orthopaedic surgeons recognize these patterns of traumatic elbow instability.

We therefore determined (1) the interobserver reliability of identifying specific patterns of traumatic elbow instability on radiographs for subgroups of orthopaedic surgeons; and (2) diagnostic accuracy of radiographs for recognizing these patterns of injury.

Patients and Methods

In 2010, 148 independent orthopaedic surgeons from several countries were invited via email to evaluate radiographs of 53 patients with one of five common patterns of traumatic elbow instability, treated by the senior author (DR) between 2000 and 2006. Seventy-six surgeons who treat elbow trauma and were interested in participating logged onto the website (https://www.surveymk.com); 73 of these 76 surgeons (96%), completed the study. The study was performed under a protocol approved by the institutional research board at the principal investigator's hospital. Seventy-three observers agreed to participate.

Patients with traumatic elbow fractures and/or dislocations were identified from a list of all patients who were treated by the principle investigator between 2000 and 2006 at one level 1 trauma center. Inclusion criteria were: (1) traumatic elbow instability; (2) available injury radiographs of adequate quality (that included the entire injury and with adequate penetration; rotation and angulation issues typical of initial postinjury radiographs were not a reason for exclusion); and (3) age 18 years or older. Fifty-three cases were randomly selected from 80 patients that satisfied the inclusion criteria by two of the authors (one subspecialty-trained upper extremity surgeon, and one research fellow in upper extremity trauma), subject to the constraint that the sample contain at least four fractures of each of five common injury patterns [2, 3, 7, 9, 14, 15, 18]: (1) posterior radial head fracture dislocation (seven cases); (2) terrible triad injury (19 cases); (3) varus posteromedial rotational injury (four cases); (4) anterior olecranon fracture dislocation (five cases); and (5) posterior olecranon fracture dislocation (18 cases). Injury patterns were verified (eg, presence or absence, and the type of coronoid fracture and soft tissue injury) on operative exposure.

The senior author (DR) operated on all 53 fractures; the author's intraoperative diagnosis of injury pattern was the reference for fracture pattern when calculating the sensitivity, specificity, and accuracy and the positive predictive value (PPV) and negative predictive value (NPV) of radiographic diagnosis.

All radiographs were blinded by an independent research fellow for use in this study. The blinded radiographs were uploaded to the website of the Science of Variation Group (https://www.surveymk.com). On login to the website, observers received a short descriptive summary of the definitions of the five patterns of traumatic elbow instability, including references (Fig. 1). Injury radiographs (of attempted AP and lateral views) of all 53 injured elbows were presented to the observers without clinical information (eg, age, sex, mechanism). Observers were asked only to classify each injury into one of the five groups. This question had to be completed to continue with the next case. Observers could comment on each case. The observers completed the study at their own pace.

The agreement among the observers was calculated by using a multirater kappa, described by Siegel and Castellan [16]. It is a commonly used statistic to describe chance-corrected agreement in various intraobserver and interobserver studies [1, 6, 11]. Zero indicated that there was no agreement beyond what was expected attributable to chance alone. The value of −1.00 meant total disagreement and +1.00 represented perfect agreement [6, 11]. The sensitivity, specificity, accuracy, and PPV and NPV of radiographic diagnosis with respect to the intraoperative reference standard were calculated using standard formulas. The PPV is the patient's probability of having an injury fracture pattern when the test is positive, and NPV is the probability of a patient not having an injury pattern when the test is negative.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request.
This work was performed at University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands and Harvard Medical School, Orthopaedic Hand and Upper Extremity Service, Massachusetts General Hospital, Boston, MA, USA.

J. N. Doornberg, T. G. Guitton
University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands

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

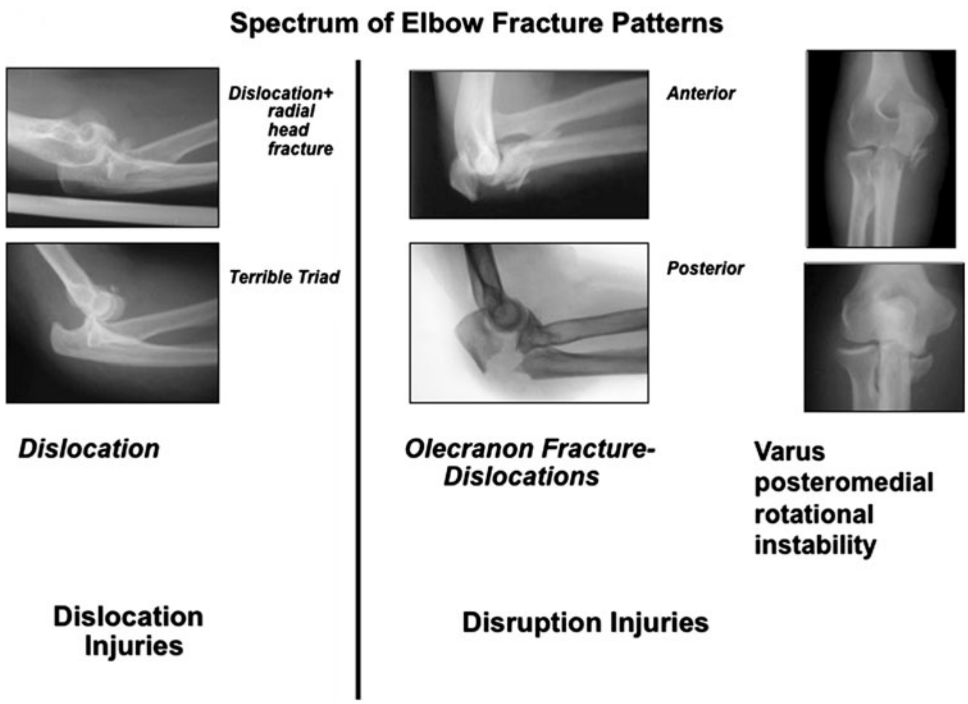

Fig. 1 The spectrum of elbow fracture patterns is shown.

Results

The overall interobserver reliability of identifying specific patterns of traumatic elbow instability on radiographs for subgroups of orthopaedic surgeons was κ = 0.41 (SE, 0.003; range, 0.22 to 0.56). The surgeon observers were mostly men practicing in the United States (Table 1). Additional analyses identified fair agreement in the following subgroups: European, Australian, and Asian observers; those who were in practice 5 years or fewer, or from 11 to 20 years; those who did not supervise trainees in the operating room; those who treated five or fewer elbow fractures per year; and shoulder and elbow specialists.

There were substantial differences in agreement between US and European observers (z = 22, p < 0.001), those who were in practice for 5 or fewer years and 21 to 30 years (z = 3, p = 0.007), observers who supervised or did not supervise in the operating room (z = 2, p = 0.045), observers who treated five or fewer and greater than 20 elbow fractures per year (z = 13, p = < 0.001), and orthopaedic trauma specialists and shoulder and elbow specialists (z = 11, p = < 0.001). The variability was greatest for posterior fracture dislocation with or without fracture of the coronoid process (the terrible triad lesion) (Table 1).

The observers' diagnoses were greater than 80% accurate for all patterns except for the terrible triad lesion (Table 2; Appendix 1).

The diagnostic performance characteristics were best for the olecranon fracture dislocations (Table 3; Appendix 2).

Discussion

Numerous studies suggest that recognition of an injury pattern that helps predict soft tissue injury and guide treatment of traumatic elbow instability are limited to cases series from dedicated elbow surgeons [2, 3, 7, 9, 14, 15, 18]. Using a large online collaborative of experienced surgeons we studied the reliability and accuracy of the diagnosis of patterns of traumatic elbow instability on radiographs. If these patterns can be identified reliably and they are accurate, they would be useful to surgeons treating these complex injuries. We therefore determined (1) the interobserver reliability of identifying specific patterns of traumatic elbow instability on radiographs for subgroups of orthopaedic surgeons; and (2) the diagnostic accuracy of radiographic diagnosis.

We acknowledge the inherent weaknesses of this study. First, we did not determine intraobserver reliability. Single round studies are more practical and, we believed interobserver reliability was more important. Second, 90% of the participating observers were in academic practice (in terms of supervising trainees), which could create bias by overestimating agreement as surgeons in academic referral clinics might be more familiar with these relatively uncommon injuries. Third, by design the ratings were determined in part by the quality of the initial injury radiographs that included the entire injury with adequate penetration, but also rotation and angulation issues typical of initial postinjury radiographs. The latter was not a reason for exclusion. We believe the quality of available radiographs reflected daily practice, which facilitates the application of our findings to current practice. Fourth, we did not study a variety of potential sources of variation, including cultural differences, standardized training of observers, electronically transmitted radiographs evaluated in regular JPEG format, computer and screen quality, and a designated viewer with which to view the radiographs. However, studies using a designated viewer might produce different results.

We found the overall agreement was acceptable with the exception, perhaps, of identifying a small coronoid fracture on radiographs. The subgroup analysis suggests that greater ongoing experience with these injuries improves awareness and recognition of injury patterns, which might translate to improved treatment of these complex injuries. However, few of the observers treated more than 10 patients a year, reflecting the fact that these injuries are uncommon. We can only speculate about some of the differences in culture, training, and specialty. These data do confirm that one sees to some extent what one knows. Future studies will investigate the influence of training and more simplified ratings on interobserver reliability.

We found the lowest accuracy was in distinguishing between posterior elbow dislocation with fracture of the radial head and associated coronoid fracture. Radiographs alone may not be sufficient to identify a small coronoid facture. The ability of CT to improve reliability and accuracy of the diagnosis of injury pattern merits study.

This argues that, where possible, these injuries might be best triaged to select surgeons who are interested in developing their expertise. Future studies might address the influence of CT on reliability and accuracy and the influence of injury pattern recognition on the treatment.

Table 1. Observer demographics and kappa values*

Variable	Number	%	Kappa	SE
Observers' gender (total = 73)				
Male	67	92	0.41	0.00
Female	6	8	0.44	0.02
Location of practice				
Asia	3	4	0.27	0.05
Australia	3	4	0.22	0.04
Canada	3	4	0.45	0.04
Europe	12	16	0.25	0.01
United States	50	68	0.47	0.00
Other	2	3	0.56	0.08
Years in practice				
0–5	15	21	0.40	0.01
6–10	21	29	0.46	0.01
11–20	28	38	0.36	0.01
21–30	9	12	0.46	0.02
Supervise				
Yes	66	90	0.41	0.00
No	7	10	0.37	0.02
Fractures per year				
0–5	11	15	0.28	0.01
6–10	18	25	0.41	0.01
11–20	23	32	0.46	0.01
> 20	21	29	0.42	0.01
Specialization				
General orthopaedics	1	1		
Orthopaedic traumatology	32	44	0.43	0.01
Shoulder and elbow	11	15	0.28	0.01
Hand and wrist	24	33	0.43	0.01
Other	5	7	0.47	0.02

* Overall kappa = 0.41; SE, 0.003; < 0.001.

Table 2. Percentage of agreement by injury pattern

Gold standard	Number of fractures	Mean agreement per fracture (%)	Median (%)	Minimum (%)	Maximum (%)	SD
Posterior radial head fracture dislocation	7	60	55	47	79	13.7
Terrible triad	19	53	53	19	84	18.7
Posteromedial varus rotational injury	4	76	86	40	92	24.6
Anterior fracture dislocation	5	81	82	73	82	3.6
Posterior olecranon fracture dislocation	18	72	75	42	86	12.6
Total	53	68	70	19	92	14.6

Table 3. Diagnostic accuracy of radiographs for identification of elbow fracture patterns

Elbow fracture pattern	PRH	No PRH		TT	No TT		PMVRI	No PMVRI		AOFD	No AOFD		POFD	No POFD	
Reference standard															
Correct	307	204	511	731	656	1387	222	70	292	294	71	365	939	375	1314
Other pattern	517	2841	3358	263	2219	2482	245	3332	3577	213	3291	3504	138	2417	2555
Totals	824	3045	3869	994	2875	3869	467	3402	3869	507	3362	3869	1077	2792	3869
Diagnostic performance characteristics															
Sensitivity**	60%			53%			76%			81%			71%		
Specificity†	85%			89%			93%			94%			95%		
Accuracy‡	81%			76%			92%			93%			87%		
Positive predictive value§	37%			74%			48%			58%			87%		
Negative predictive value#	93%			77%			98%			98%			87%		

AOFD = anterior fracture dislocation; PMRVI = posteromedial varus rotational injury; POFD = posterior olecranon fracture dislocation; PRH = posterior radial head fracture dislocation; TT = terrible triad; ** proportion of patients with an elbow fracture pattern X classified as pattern X (true positives); †proportion of patients with no elbow fracture pattern X fracture dislocation classified as not having pattern X (true negatives); ‡proportion of patients who are correctly classified by the test. Proportion of true positive and true negatives; §probability that a patient with a positive plain radiograph for elbow fracture pattern X has pattern X; #probability that a patient with negative plain radiographs does not have pattern X.

Appendices

Appendix 1. Agreement per case

Case	Standard	Agreement (number)	%
1	Posteromedial varus rotational injury	67	92
2	Terrible triad	51	70
3	Posterior radial head fracture dislocation	58	79
4	Terrible triad	38	52
5	Terrible triad	26	36
6	Terrible triad	58	79
7	Terrible triad	39	53
8	Posterior olecranon fracture dislocation	63	86
9	Posterior olecranon fracture dislocation	55	75
10	Posterior olecranon fracture dislocation	33	45
11	Posterior olecranon fracture dislocation	53	73
12	Terrible triad	17	23
13	Posterior olecranon fracture dislocation	31	42
14	Posterior olecranon fracture dislocation	58	79
15	Anterior fracture dislocation	61	84
16	Posterior olecranon fracture dislocation	51	70
17	Posteromedial varus rotational injury	60	82
18	Posterior olecranon fracture dislocation	61	84
19	Terrible triad	14	19
20	Posterior radial head fracture dislocation	50	68
21	Terrible triad	42	58
22	Posterior olecranon fracture dislocation	54	74
23	Posterior olecranon fracture dislocation	55	75
24	Posteromedial varus rotational injury	66	90
25	Terrible triad	22	30
26	Terrible triad	61	84
27	Terrible triad	37	51
28	Terrible triad	31	42
29	Anterior fracture dislocation	60	82
30	Terrible triad	47	64
31	Terrible triad	38	52
32	Posterior olecranon fracture dislocation	44	60
33	Posterior radial head fracture dislocation	54	74
34	Terrible triad	50	68
35	Terrible triad	41	56
36	Anterior fracture dislocation	53	73
37	Posterior olecranon fracture dislocation	43	59
38	Anterior fracture dislocation	62	85
39	Posterior radial head fracture dislocation	37	51
40	Terrible triad	51	70
41	Posterior radial head fracture dislocation	40	55
42	Posterior olecranon fracture dislocation	49	67
43	Posterior olecranon fracture dislocation	55	75
44	Anterior fracture dislocation	58	79
45	Terrible triad	21	29
46	Posterior olecranon fracture dislocation	62	85

Appendix 1. continued

Case	Standard	Agreement (number)	%
47	Posterior radial head fracture dislocation	34	47
48	Posterior olecranon fracture dislocation	61	84
49	Terrible triad	47	64
50	Posteromedial varus rotational injury	29	40
51	Posterior radial head fracture dislocation	34	47
52	Posterior olecranon fracture dislocation	54	74
53	Posterior olecranon fracture dislocation	57	78

Appendix 2. Diagnostic accuracy of using plain radiographs to identify elbow fracture patterns - Statistical Analysis

Elbow fracture pattern	Pattern X	Not Pattern X	
Reference standard			
Correct	True positive (TP)	False negative (FN)	TP + FN
Other pattern	False positive (FP)	True negative (TN)	FP + TN
Totals	TP + FP	FN + TN	3869
Diagnostic performance characteristics			
Sensitivity**			TP/ TP + FN
Specificity†			TN/FP + TN
Accuracy‡			TP + TN/ TOTAL
Positive predictive value §			TP/ TP + FP
Negative predictive value #			TN/FN + TN

**Proportion of patients with an elbow fracture pattern X classified as pattern X (true positives); †proportion of patients with no elbow fracture pattern X fracture dislocation classified as not having pattern X (true negatives); ‡proportion of patients who are correctly classified by the test. Proportion of true positive and true negatives; §probability that a patient with a positive plain radiograph for elbow fracture pattern X has pattern X; #probability that a patient with negative plain radiographs does not have pattern X.

References

1. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37–46.
2. Doornberg J, Ring D, Jupiter JB. Effective treatment of fracture-dislocations of the olecranon requires a stable trochlear notch. *Clin Orthop Relat Res.* 2004;429:292–300.
3. Doornberg JN, Ring D. Coronoid fracture patterns. *J Hand Surg Am.* 2006;31:45–52.
4. Doornberg JN, Ring DC. Fracture of the anteromedial facet of the coronoid process. *J Bone Joint Surg Am.* 2006;88:2216–2224.
5. Giannicola G, Sacchetti FM, Greco A, Cinotti G, Postacchini F. Management of complex elbow instability. *Musculoskelet Surg.* 2010;94(suppl 1):S25–36.
6. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
7. McKee MD, Pugh DM, Wild LM, Schemitsch EH, King GJ. Standard surgical protocol to treat elbow dislocations with radial head and coronoid fractures: surgical technique. *J Bone Joint Surg Am.* 2005;87(suppl 1):22–32.
8. McKee MD, Schemitsch EH, Sala MJ, O'driscoll SW. The pathoanatomy of lateral ligamentous disruption in complex elbow instability. *J Shoulder Elbow Surg.* 2003;12:391–396.
9. O'Driscoll SW, Jupiter JB, Cohen MS, Ring D, McKee MD. Difficult elbow fractures: pearls and pitfalls. *Instr Course Lect.* 2003;52:113–134.
10. Pollock JW, Brownhill J, Ferreira L, McDonald CP, Johnson J, King G. The effect of anteromedial facet fractures of the coronoid and lateral collateral ligament injury on elbow stability and kinematics. *J Bone Joint Surg Am.* 2009;91:1448–1458.
11. Posner KL, Sampson PD, Caplan RA, Ward RJ, Cheney FW. Measuring interrater reliability among multiple raters: an example of methods for nominal data. *Stat Med.* 1990;9:1103–1115.
12. Ring D, Jupiter JB. Fracture dislocation of the elbow. *J Bone Joint Surg Am.* 1998;80:566–580.
13. Ring D, Jupiter JB. Reconstruction of posttraumatic elbow instability. *Clin Orthop Relat Res.* 2000;370:44–56.
14. Ring D, Jupiter JB, Sanders RW, Mast J, Simpson NS. Transolecranon fracture-dislocation of the elbow. *J Orthop Trauma.* 1997;11:545–550.
15. Ring D, Jupiter JB, Zilberfarb J. Posterior dislocation of the elbow with fractures of the radial head and coronoid. *J Bone Joint Surg Am.* 2002;84:547–551.
16. Siegel S, Castellan NJ. *Nonparametric Statistics for the Behavioral Sciences.* New York, NY: McGraw-Hill; 1988.
17. Sotereanos DG, Darlis NA, Wright TW, Goitz RJ, King GJ. Unstable fracture dislocations of the elbow. *Instr Course Lect.* 2007;56:369–376.
18. Tashjian RZ, Katarincic JA. Complex elbow instability. *J Am Acad Orthop Surg.* 2006;14:278–286.