J Shoulder Elbow Surg (2015) 24, e21-e26

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

The factors influencing the decision making of operative treatment for proximal humeral fractures

Michiel G.J.S. Hageman, MD[a], Prakash Jayakumar, MD[a], John D. King, BSc[a], Thierry G. Guitton, MD, PhD[b], Job N. Doornberg, MD, PhD[b], David Ring, MD, PhD[c,*] for the Science of Variation Group[1]

[a]*Orthopaedic Hand and Upper Extremity Service, Massachusetts General Hospital, Boston, MA, USA*
[b]*Academic Medical Center Amsterdam, Amsterdam, The Netherlands*
[c]*Yawkey Center, Massachusetts General Hospital, Boston, MA, USA*

Background: The factors influencing the decision making of operative treatment for fractures of the proximal humerus are debated. We hypothesized that there is no difference in treatment recommendations between surgeons shown radiographs alone and those shown radiographs and patient information. Secondarily, we addressed (1) factors associated with a recommendation for operative treatment, (2) factors associated with recommendation for arthroplasty, (3) concordance with the recommendations of the treating surgeons, and (4) factors affecting the inter-rater reliability of treatment recommendations.

Methods: A total of 238 surgeons of the Science of Variation Group rated 40 radiographs of patients with proximal humerus fractures. Participants were randomized to receive information about the patient and mechanism of injury. The response variables included the choice of treatment (operative vs nonoperative) and the percentage of matches with the actual treatment.

Results: Participants who received patient information recommended operative treatment less than those who received no information. The patient information that had the greatest influence on treatment recommendations included age (55%) and fracture mechanism (32%). The only other factor associated with a recommendation for operative treatment was region of practice. There was no significant difference between participants who were and were not provided with information regarding agreement with the actual treatment (operative vs nonoperative) provided by the treating surgeon.

Conclusion: Patient information—older age in particular—is associated with a higher likelihood of recommending nonoperative treatment than radiographs alone. Clinical information did not improve agreement of the Science of Variation Group with the actual treatment or the generally poor interobserver agreement on treatment recommendations.

The Massachusetts General Hospital Institutional Research Board approved this study (IRB No.: 2009P001019).

*Reprint requests: David Ring, MD, PhD, Chief Orthopaedic Hand Service, Yawkey Center, Ste 2100, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114, USA.

E-mail address: dring@partners.org (D. Ring).

[1] **The Science of Variation Group** member list is available at http://dx.doi.org/10.1016/j.jse.2014.05.013.

http://dx.doi.org/10.1016/j.jse.2014.05.013

Level of evidence: Basic Science, Survey of Experts.

Keywords: Proximal humeral fractures; factors; decision-making; operative versus nonoperative treatment

The role of operative treatment for fractures of the proximal humerus is debated. Surgery is considered for approximately 1 in 5 patients, but there is no consensus on which fractures benefit from surgery or which procedure to perform.[7] The data to date are limited and inconclusive.[7,9] A recent Cochrane review found no statistically significant difference between operative and nonoperative treatment regarding patient-reported functional scores and EuroQoL results at 1 year from 3 randomized control trials with a total of 153 participants.[7] However, compared to nonoperative treatment, operative treatment had superior EuroQoL scores at 2 years of follow-up in 2 randomized control trials with a total of 101 participants.[7]

Among a small group of surgeons at 2 level 1 trauma centers, Okike et al[11] identified younger age, operative treatment of other musculoskeletal injuries, Arbeitsgemeinschaft für Ostesynthesefragen (AO) classification, translation-type displacement, associated glenohumeral dislocation, and surgeon subspecialty (upper extremity specialists were more likely to operate than traumatologists) as factors associated with operative intervention. The use of arthroplasty rather than internal fixation was associated with a higher Charlson score and more severe Neer and AO classifications.[4,10,11] Many of these factors relate to the radiographic appearance of the fracture, whereas some relate to patient or surgeon factors.

We were curious about the factors that influence agreement between surgeons on treatment recommendations and the factors that lead a surgeon to recommend operative treatment and type of surgery (ie, fixation vs arthroplasty). We used the Science of Variation Group (SVOG), an international Web-based collaborative of practicing surgeons, to test the primary null hypothesis that there is no difference in treatment recommendations regarding operative vs nonoperative treatment between surgeons shown radiographs alone and those shown radiographs and patient information such as age, sex, hand dominance, and fracture mechanism. Secondarily, we addressed (1) factors associated with a recommendation for operative treatment, (2) factors associated with recommendation for arthroplasty, (3) concordance with the recommendations of the treating surgeons, and (4) factors affecting inter-rater reliability of treatment recommendations.

Materials and methods

We asked the surgeons of the SOVG to complete a survey regarding the recommendation of operative or nonoperative treatment for a series of proximal humeral fractures. The SOVG is an international collaboration of fully trained surgeon observers that studies variation in the definition, interpretation, classification, and treatment of human illness. Collaborative authorship, scientific curiosity, and camaraderie are the only incentives for participation.

Participating members viewed the radiographs of 20 fractures of the proximal humerus treated operatively and 20 treated nonoperatively. These were the radiographs used by the surgeon caring for the patient and were not standardized. Participants were randomized to receive information about the patient, including sex, age, American Society of Anesthesiologists (ASA)

Table I Demographics of the participating surgeons

Parameters	Participants with radiographs (N = 228)	
	With additional information	Without additional information
	No. (%)	No. (%)
Sex		
Men	122 (52)	101 (42)
Women	8 (3.2)	7 (2.8)
Location of practice		
Asia	9 (3.6)	2 (0.8)
Australia	4 (1.6)	2 (0.8)
Canada	7 (2.8)	4 (1.6)
Europe	40 (17)	27 (12)
United Kingdom	5 (2.0)	0
United States of America	59 (24)	62 (25)
Other	12 (4.8)	11 (4.4)
Years in practice		
0-5	41 (18)	40 (17)
6-10	29 (12)	18 (7.3)
11-20	34 (15)	33 (13)
21-30	26 (11)	17 (6.9)
Supervise trainees		
Yes	119 (50)	95 (39)
No	11 (5.6)	13 (5.6)
Fractures per year		
0-5	11 (4.8)	14 (5.6)
6-10	32 (13)	23 (9.7)
11-20	41 (17)	34 (14)
>20	50 (21)	37 (15)
Specialization		
General orthopedics	7 (2.8)	6 (2.4)
Orthopedic traumatology	58 (24)	35 (14)
Shoulder and elbow	27 (11)	27 (11)
Hand and wrist	32 (15)	39 (16)
Other	6 (2.4)	1 (0.4)

Table II Bivariable analysis about the participating surgeons' recommendation

Parameters	Operative treatment, %		Min-Max	Nonoperative treatment, %		Min-Max
	Operative treatment	Nonoperative		Operative treatment	Nonoperative	
Overall						
Observers	64	38	0.23-1	46	54	0-1
			P			*P*
Information group						
Radiographs with additional information	61	39	<.01	43	57	<.01
Radiographs without additional information	66	34		50	50	
Sex						
Men	63	37	.8	46	54	.93
Women	65	35		47	53	
Location of practice						
Asia	72	28		58	42	
Australia	69	31		53	47	
Canada	51	49		30	70	
Europe	62	38	<.01	45	55	<.01
United Kingdom	54	46		37	63	
United States of America	61	39		45	55	
Other	73	27		60	40	
Years in practice						
0-5	62	38		44	56	
6-10	62	38	.24	44	56	.25
11-20	66	34		49	51	
21-30	64	36		49	51	
Supervise trainees						
Yes	63	37	.48	46	54	.54
No	65	35		47	53	
Fractures per year						
0-5	63	37		47	53	
6-10	61	39	.57	41	59	.18
11-20	65	35		49	51	
>20	65	35		47	53	
Specialization						
General orthopedics	62	38		42	58	
Orthopedic traumatology	63	37	.41	46	54	.62
Shoulder and elbow	62	38		44	56	
Hand and wrist	66	34		49	51	
Other	60	40		50	50	

classification, and hand dominance, and mechanism of injury or not, in a 1-to-1 allocation.

Evaluation

The 40 proximal humeral fractures were selected from a separate case-control study in which 66 patients, 33 treated operatively and 33 treated nonoperatively, were matched for fracture type, age, sex, and ASA classification. The treating surgeons classified those fractures as 2-part surgical neck fractures in 7 pairs of patients, 3-part fractures in 9 pairs, and 4-part fractures in 4 pairs. Participants viewed anterior-posterior and lateral radiographs.

Each participant provided demographic and professional information: sex, world region of practice, years in independent practice, supervision of trainees, and surgical subspecialty. Each observer was asked 2 questions for each set of images: (1) Would you recommend surgery? And, if so, (2) What is your preference: open reduction and internal fixation, percutaneous pinning, or arthroplasty? In addition, the observers who received patient information were asked a third question: What information was most influential? They could choose from the following answers: (1) patient characteristics (sex, age, ASA, hand dominance), (2) fracture mechanism, (3) other (Supplementary Table I).

Of the 238 surgeons who completed the survey, 130 were randomized to receive information and 108 were randomized to receive radiographs alone. The cohorts were comparable (Table I).

Statistical analysis

The response variables included the choice of treatment (operative vs nonoperative) and the percentage of matches with the actual treatment. Associations between response variables and categoric

explanatory variables were assessed using χ^2 tests. Factors with $P < .10$ in bivariate analysis were entered into a multiple logistic regression analysis.

The percentage of agreement (1) with other observers and (2) with the original treating surgeon was calculated for surgeons who received information or not, and the κ multirater measure was also measured. The κ values were interpreted with use of the guidelines proposed by Landis and Koch.[5,8]

An a priori power analysis indicated that a cohort of 200 surgeons randomized equally to review radiographs with or without patient information would provide 80% power to detect a mean difference of 0.10 in interobserver reliability based on multirater κ, assuming a pooled standard deviation of 0.25 (moderate effect size: 0.10/0.25 = 0.40) using a parametric Z test with a two-tailed α level of 0.05 and assuming an underlying normal distribution in the patient population.

Results

Participants who received patient information recommended operative treatment less than those who received no information (60% vs 66%; $P < .01$; Table II). The only other factor associated with a recommendation for operative treatment was region of practice: participants from Asia were more likely than participants from Canada to recommend surgery (72% vs 51%; $P < .01$; Table II). In multivariable analysis, the only factor associated with a recommendation of nonoperative treatment was receiving clinical information in addition to radiographs, which explained 10% of the variation in recommendation for nonoperative treatment (odds ratio, 4.3; $r^2 = 0.10$; $P < .01$).

Participants provided with patient information were more likely to recommend arthroplasty (24% vs 17%; $P < .01$) and less likely to recommend open reduction and internal fixation (69% vs 76%; $P < .01$; Table III). The patient information that had the greatest influence on treatment recommendations included age (55%) and fracture mechanism (32%; Table IV).

There was no significant difference between participants provided with information and participants not provided information regarding agreement with the actual treatment (operative vs nonoperative) provided by the treating surgeon (65% vs 67% on average; $P = .11$). Agreement with the actual treatment provided by the treating surgeon varied significantly by location of practice ($P < .01$; Table V).

Interobserver agreement regarding recommendations for operative treatment was poor (average, −0.008; range, −0.007 to –0.008; Supplementary Table II).

Table III Comparison of percentage of arthroplasty as preferred osteosynthesis for the information and noninformation group

Parameters	Mean, %	*P*	95% CI	
			Lower	Upper
Arthroplasty				
Information group				
Radiographs with additional information	24	<.01	0.04	0.1
Radiographs without additional information	17			
Open reduction, internal fixation				
Information group				
Radiographs with additional information	69	<.01	0.71	0.75
Radiographs without additional information	76			
Pin				
Information group				
Radiographs with additional information	5.9	.54	−0.04	0.02
Radiographs without additional information	6.8			

CI, confidence interval.

Table IV Additional information used by the information group for decision making regarding the preferred treatment

Parameters	No.	%
Sex	9	7.2
Age	72	55
ASA	20	15
Hand dominance	12	8.9
Fracture mechanism	42	32
Other	68	52

ASA, American Society of Anesthesiologists.

Discussion

The recommendations for managing proximal humeral fractures vary substantially. Recent studies demonstrate the poor levels of reliability in the treatment of these injuries.[3,6] We were interested in the relative influence of patient information and surgeon characteristics on the decision-making process in treating proximal humeral fractures.

This study should be considered in light of its shortcomings. Most regions were represented by small numbers of observers, and the findings may not be representative of the average surgeon in those regions. The low κ values may be due to the κ paradox: when the prevalence of an outcome is low, it could cause an imbalance that generates a lower κ than one might expect based on the agreement. Also, our study did not include fracture classification as an explanatory variable, although interobserver reliability of

Table V Bivariable analysis comparing the recommendation of the participating surgeons and the actual treatment

Parameters	Operative treatment, %		Min-Max	Nonoperative treatment,%		Min-Max
	Operative treatment	Nonoperative		Operative treatment	Nonoperative	
Overall						
Observers	80	20	0.3-1.0	46	54	0.0-1.0
			P			*P*
Information group						
Radiographs with additional information	77	23	<.01	43	57	<.01
Radiographs without additional information	83	17		50	50	
Sex						
Men	80	20	.45	46	54	.93
Women	82	20		47	53	
Location of practice						
Asia	85	15		58	42	
Australia	86	14		53	47	
Canada	71	29		30	70	
Europe	79	21	<.01	45	55	<.01
United Kingdom	71	29		37	63	
United States of America	79	21		45	55	
Other	86	14		60	40	
Years in practice						
0-5	79	21		44	66	
6-10	77	23	.39	44	66	.26
11-20	81	19		48	51	
21-30	79	21		49	51	
Supervise trainees						
Yes	80	20	.52	46	54	.93
No	78	22		47	53	
Fractures per year						
0-5	78	22		47	52	
6-10	78	22		41	59	
11-20	80	20	.62	49	51	.14
>20	81	19		47	53	
Specialization						
General orthopedics	82	18		42	58	
Orthopedic traumatology	79	21		46	54	
Shoulder and elbow	80	20	.23	44	57	.62
Hand and wrist	80	20		49	51	
Other	69	31		50	50	

fracture classification is limited amongst orthopedic surgeons.[2,3] In addition, a consecutive selected case series would have limited spectrum bias. However, a relatively even distribution between operatively and nonoperatively treated patients is needed to avoid the κ paradox. Readers should also keep in mind that observers were shown the unstandardized anterior-posterior and lateral radiographs used by the treating surgeons to direct management, thereby reflecting daily practice. Some surgeons use other radiographic views and computed tomography scans, including 3-dimensional reconstructions, more routinely in the decision process.

We found that patient information—older age in particular—is associated with a higher likelihood of recommending nonoperative treatment than radiographs alone. This is consistent with the observations of Okike et al.[11]

We also identified regional variations in treatment recommendations. This was in concordance with earlier published reports that demonstrated wide regional variations for surgical treatment adjusted for age, sex, and race in populations of elderly patients with proximal humeral fractures ranging from 0% in many regions to almost 70% in Duluth, Minnesota, USA.[1,13]

Of interest, the sex, level of experience in years in practice, number of proximal humeral fractures treated per year, and specialization did not have a significant influence on the decision to operate. This is in contrast to

other studies, which report shoulder and upper extremity specialists are more likely to choose operative intervention than general orthopedic trauma specialists.[11] The effect of specialty may vary in specific centers with small samples of each type of surgeon, but our larger, broader cohort suggests that specialty has relatively little influence.

The provision of clinical information also influenced the recommended type of operative treatment, with arthroplasty (as with nonoperative treatment) favored by surgeons who received information about the patients. This suggests that—with current biases—older, more infirm, and inactive patients are less likely to be treated operatively and are more likely to be treated with arthroplasty if they do have operative treatment.

Clinical information did not improve agreement of the SOVG participants with the actual treatment or the generally poor interobserver agreement on treatment recommendations. The poor agreement may be unreliable due to the κ paradox. A study by Petit et al[12] documented moderate interobserver agreement on the surgical management (nonoperative, closed manipulation and reduction, open reduction and internal fixation, and hemiarthrosplasty) of 38 proximal humeral fractures among 8 fellowship-trained orthopedic surgeons.

Treatment recommendations for proximal humeral fractures are influenced by patient information—older age in particular—but most of the variation in recommendations remains unaccounted for. The highly variable and inconsistent influence of patient factors on surgeon recommendations belies variations in surgeon preferences and values that are likely at the root of the substantial treatment variations documented in this and other studies. We speculate that greater involvement of the patient in decision making is likely to decrease variation across caregivers.

Conclusion

Patient information—older age in particular—is associated with a higher likelihood of recommending nonoperative treatment than radiographs alone. Clinical information did not improve agreement of the SOVG with the actual treatment or the generally poor interobserver agreement on treatment recommendations.

Disclaimer

Michiel G.J.S. Hageman is supported by Dutch research grants from Marti-Keunig Eckhart Stichting, Anna Foundation, and Spinoza Foundation.

The authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jse.2014.05.013.

References

1. Bell JE, Leung BC, Spratt KF, Koval KJ, Weinstein JD, Goodman DC, et al. Trends and variation in incidence, surgical treatment, and repeat surgery of proximal humeral fractures in the elderly. J Bone Joint Surg Am 2011;93(2):121-31. http://dx.doi.org/10.2106/JBJS.I.01505
2. Brorson S, Rasmussen JV, Frich LH, Olsen BS, Hrobjartsson A. Benefits and harms of locking plate osteosynthesis in intraarticular (OTA Type C) fractures of the proximal humerus: a systematic review. Injury 2012;43(7):999-1005. http://dx.doi.org/10.1016/j.injury.2011.08.025
3. Bruinsma WE, Guitton TG, Warner JJ, Ring D. Science of Variation G. Interobserver reliability of classification and characterization of proximal humeral fractures: a comparison of two and three-dimensional CT. J Bone Joint Surg Am 2013;95(17):1600-4. http://dx.doi.org/10.2106/JBJS.L.00586
4. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-83.
5. Cohen JA. Coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37-46.
6. Foroohar A, Tosti R, Richmond JM, Gaughan JP, Ilyas AM. Classification and treatment of proximal humerus fractures: inter-observer reliability and agreement across imaging modalities and experience. J Orthop Surg Res 2011;6:38. http://dx.doi.org/10.1186/1749-799X-6-38
7. Handoll HH, Ollivere BJ, Rollins KE. Interventions for treating proximal humeral fractures in adults. Cochrane Database Syst Rev 2012;(12):CD000434. http://dx.doi.org/10.1002/14651858.CD000434.pub2
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
9. Misra A, Kapur R, Maffulli N. Complex proximal humeral fractures in adults–a systematic review of management. Injury 2001;32(5):363-72.
10. Neer CS 2nd. Displaced proximal humeral fractures. I. Classification and evaluation. J Bone Joint Surg Am 1970;52(6):1077-89.
11. Okike K, Lee OC, Makanji H, Harris MB, Vrahas MS. Factors associated with the decision for operative versus non-operative treatment of displaced proximal humerus fractures in the elderly. Injury 2013;44(4):448-55. http://dx.doi.org/10.1016/j.injury.2012.09.002
12. Petit CJ, Millett PJ, Endres NK, Diller D, Harris MB, Warner JJ. Management of proximal humeral fractures: surgeons don't agree. J Shoulder Elbow Surg 2010;19(3):446-51. http://dx.doi.org/10.1016/j.jse.2009.06.012. S1058-2746(09)00294-8 [pii].
13. Sporer SM, Weinstein JN, Koval KJ. The geographic incidence and treatment variation of common fractures of elderly patients. J Am Acad Orthop Surg 2006;14(4):246-55.