

# Treatment of Proximal Humerus Fractures: Comparison of Shoulder and Trauma Surgeons

Andrew Jawa, MD, Paul H. Yi, MD, Robert E. Boykin, MD, Michael J. Gardner, MD, Christian Gerber, MD, Dean G. Lorich, MD, Gilles Walch, MD, and Jon J. P. Warner, MD

## Abstract

Surgeons' disagreement about ideal treatment for proximal humerus fractures (PHFs) may reflect a difference in training.

We conducted a study to compare treatment decision-making by experienced shoulder and trauma fellowship-trained surgeons. Two expert shoulder surgeons and 2 expert trauma surgeons reviewed 100 consecutive PHFs surgically treated at another institution. Using available imaging, the examiners assigned scores for agreement with treatment decisions and for ratings of reduction/arthroplasty placement, fixation method, and radiographic outcomes. The scores were evaluated for interobserver

reliability using intraclass correlation coefficients.

Overall, these experienced surgeons agreed poorly with treatment decisions and fixation methods but agreed moderately on acceptable reductions/arthroplasty placement and final radiographic outcomes. Agreement on the final radiographic outcomes was more uniform and acceptable for both shoulder and trauma surgeons. Trauma surgeons agreed more with each other about treatment decisions than shoulder surgeons agreed with each other.

In this study, surgeon disagreement and an aging population highlight the need for better evidence regarding optimal treatment for PHFs in order to improve consensus.

Proximal humerus fractures (PHFs), AO/OTA (Arbeitsgemeinschaft für Osteosynthesefragen/Orthopaedic Trauma Association) type 11,<sup>1</sup> are common, representing 4% to 5% of all fractures in adults.<sup>2</sup> However, there is no consensus as to optimal management of these injuries, with some reports supporting and others rejecting the various fixation methods,<sup>3</sup> and there are no evidence-based practice guidelines informing treatment decisions.<sup>4</sup> Not surprisingly, orthopedic surgeons do not agree on ideal treatment for PHFs<sup>5,6</sup> and differ by region in their rates of surgical management.<sup>2</sup> In addition, analyses of national databases have found variation in choice of surgical treatment for PHFs between surgeons and between hospitals of different patient volumes.<sup>4</sup> Few studies have assessed surgeon agreement on treatment decisions. Findings from these limited investigations indicate there is little agreement on treatment choices, but training may have some impact.<sup>5-7</sup> In 3 studies,<sup>5-7</sup> shoulder and trauma fellowship-trained surgeons differed in their management of PHFs both in terms of rates of operative treatment<sup>5,7</sup> and specific operative management choices.<sup>5,6</sup> No study has assessed surgeon agreement on radiographic outcomes.

We conducted a study to compare expert shoulder and

trauma surgeons' treatment decision-making and agreement on final radiographic outcomes of surgically treated PHFs. We hypothesized there would be poor agreement on treatment decisions and better agreement on radiographic outcomes, with a difference between shoulder and trauma fellowship-trained surgeons.

## Materials and Methods

After receiving institutional review board approval for this study, we collected data on 100 consecutive PHFs (AO/OTA type 11<sup>1</sup>) surgically treated at 2 affiliated level I trauma centers between January 2004 and July 2008. None of the cases in the series was managed by any of the surgeons participating in this study.

We created a PowerPoint (Microsoft, Redmond, Washington) survey that included radiographs (preoperative, immediate postoperative, final postoperative) and, if available, a computed tomography image. This survey was sent to 4 orthopedic surgeons: Drs. Gardner, Gerber, Lorich, and Walch. Two of these authors are fellowship-trained in shoulder surgery, the other 2 in orthopedic traumatology with specialization in treating PHFs. All are internationally renowned in PHF

---

**Authors' Disclosure Statement:** The authors report no actual or potential conflict of interest in relation to this article.

management. Using the survey images and a 4-point Likert scale ranging from *disagree strongly* to *agree strongly*, the examiners rated their agreement with treatment decisions (arthroplasty vs fixation). They also rated (*very poor* to *very good*) immediate postoperative reduction or arthroplasty placement, immediate postoperative fixation methods for fractures treated with open reduction and internal fixation (ORIF), and final radiographic outcomes.

Interobserver agreement was calculated using the intraclass correlation coefficient (ICC),<sup>8,9</sup> with scores of <0.2 (*poor*), 0.21 to 0.4 (*fair*), 0.41 to 0.6 (*moderate*), 0.61 to 0.8 (*good*), and >0.8 (*excellent*) used to indicate agreement among observers. ICC scores were determined by treating the 4 examiners as independent entities. Subgroup analyses were also performed to determine ICC scores comparing the 2 shoulder surgeons, comparing the 2 trauma surgeons, and comparing the shoulder surgeons and trauma surgeons as 2 separate groups. ICC scores were used instead of  $\kappa$  coefficients to assess agreement because ICC scores treat ratings as continuous variables, allow for comparison of 2 or more raters, and allow for assessment of correlation among raters, whereas  $\kappa$  coefficients treat data as categorical variables and assume the ratings have no natural ordering. ICC scores were generated by SAS 9.1.3 software (SAS Institute, Cary, North Carolina).

## Results

The 4 surgeons' overall ICC scores for agreement with the rating of immediate reduction or arthroplasty placement and the rating of final radiographic outcome indicated moderate levels of agreement (Table 1). Regarding treatment decision-making and ratings of fixation, the surgeons demonstrated poor and fair levels of agreement, respectively.

The ICC scores comparing the shoulder and trauma surgeons revealed similar levels of agreement (Table 2): moderate levels of agreement for ratings of both immediate postoperative reduction or arthroplasty placement and final radiographic outcomes, but poor and fair levels of agreement regarding treatment decision-making and the rating of immediate postoperative fixation methods for fractures treated with ORIF, respectively.

Subgroup analysis revealed that the 2 shoulder surgeons had poor and fair levels of agreement for treatment decisions and rating of immediate postoperative fixation, respectively, though they moderately agreed on rating of immediate postoperative reduction or arthroplasty placement and rating of final radiographic outcome (Table 3). When the 2 trauma surgeons were compared with each other, ICC scores revealed higher levels of agreement overall (Table 4). In other words, the 2 trauma surgeons agreed with each other more than the 2 shoulder surgeons agreed with each other.

## Discussion

This study had 3 major findings: (1) Surgeons do not agree on treatment decisions, including fixation methods, regarding PHFs; (2) regardless of their opinions on ideal treatment, they moderately agree on reductions and final radiographic

outcomes; (3) expert trauma surgeons may agree more on treatment decisions than expert shoulder surgeons do. In other words, surgeons do not agree on the best treatment, but they radiographically recognize when a procedure has been performed technically well or poorly. These results support our hypothesis and the limited current literature.

An analysis of Medicare databases showed marked regional variation in rates of operative treatment of PHFs.<sup>2</sup> Similarly, a Nationwide Inpatient Sample analysis revealed nationwide variation in operative management of PHFs.<sup>4</sup> Both findings are consistent with our results of poor agreement about treatment decisions and ratings of postoperative fixation of PHFs. In 2010, Petit and colleagues<sup>6</sup> reported that surgeons do not agree on PHF management. In 2011, Foroohar and colleagues<sup>10</sup> similarly reported low interobserver agreement for treatment recommendations made by 4 upper extremity orthopedic specialists, 4 general orthopedic surgeons, 4 senior residents, and 4 junior residents, for a series of 16 PHFs—also consistent with our findings.

The lack of agreement about PHF treatment may reflect a difference in training, particularly in light of the recent expansion of shoulder and elbow fellowships.<sup>2</sup> Three separate studies performed at 2 affiliated level I trauma centers demonstrated significant differences in treatment decision-making between shoulder and trauma fellowship-trained surgeons.<sup>5-7</sup> Our results are consistent with the hypothesis that training affects treatment decision-making, as we found poor agreement between shoulder and trauma fellowship-trained surgeons regarding treatment decision for PHFs. Subanalyses revealed that expert trauma surgeons agreed with each other on treatment decisions more than expert shoulder surgeons agreed with each other, further suggesting that training may affect how surgeons manage PHFs. Differences in fellowship training even within the same specialty may account for the observed lesser levels of agreement between the shoulder surgeons, even among experts in the field.

The evidence for optimal treatment historically has been poor,<sup>4,6</sup> with few high-quality prospective, randomized controlled studies on the topic up until the past few years. The most recent *Cochrane Review* on optimal PHF treatment concluded that there is insufficient evidence to make an evidence-based recommendation and that the long-term benefit of surgery is unclear.<sup>11</sup> However, at least 5 controlled trials on the topic have been published within the past 5 years.<sup>12-16</sup> The evidence is striking and generally supports nonoperative treatment for most PHFs, including some displaced fractures—contrary to general orthopedic practice in many parts of the United States,<sup>2</sup> which hitherto had been based mainly on individual surgeon experience and the limited literature. Without strong evidence to support one treatment option over another, surgeons are left with no objective, scientific way of coming to agreement.

Related to the poor status quo of evidence for PHF treatments is new technology (eg, locking plates, reverse total shoulder arthroplasty) that has expanded surgical indications.<sup>2,17</sup> Although such developments have the potential to improve surgical treatments, they may also exacerbate the disagreement

between surgeons regarding optimal operative treatment of PHFs. This potential consequence of new technology may be reflected in our finding of disagreement among surgeons on immediate postoperative fixation methods. Precisely because they are new, such technological innovations have limited evidence supporting their use. This leaves surgeons with little to nothing to inform their decisions to use these devices, other than familiarity with and impressions of the new technology.

Our study had several limitations. First is the small sample size, of surgeons who are leaders in the field. Our sample therefore may not be generalizable to the general population of shoulder and trauma surgeons. Second, we did not cal-

culate intraobserver variability. Third, inherent to studies of interobserver agreement is the uncertainty of their clinical relevance. In the clinical setting, a surgeon has much more information at hand (eg, patient history, physical examination findings, colleague consultations), thus raising the possibility of underestimations of interobserver agreements.<sup>18</sup> Fourth, our comparison of surgeons' ratings of outcomes was purely radiographic, which may or may not represent or be indicative of clinical outcomes (eg, pain relief, function, range of motion, patient satisfaction). The conclusions we may draw are accordingly limited, as we did not directly evaluate clinical outcome parameters.

**Table 1. Interobserver Agreement: Overall**

	ICC	95% CI	Agreement Strength
Agreement with treatment decision	0.14	0.03–0.26	Poor
Rating of:			
Immediate postoperative reduction or arthroplasty placement	0.50	0.40–0.58	Moderate
Immediate postoperative fixation methods	0.24	0.13–0.35	Fair
Final radiographic outcome	0.59	0.51–0.66	Moderate

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

**Table 2. Interobserver Agreement Between Shoulder and Trauma Surgeons**

	ICC	95% CI	Agreement Strength
Agreement with treatment decision	0.06	0.08–0.21	Poor
Rating of:			
Immediate postoperative reduction or arthroplasty placement	0.45	0.33–0.55	Moderate
Immediate postoperative fixation methods	0.28	0.14–0.40	Fair
Final radiographic outcome	0.56	0.46–0.65	Moderate

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

**Table 3. Interobserver Agreement Among Shoulder Surgeons**

	ICC	95% CI	Agreement Strength
Agreement with treatment decision	0.18	0.03–0.38	Poor
Rating of:			
Immediate postoperative reduction or arthroplasty placement	0.55	0.40–0.68	Moderate
Immediate postoperative fixation methods	0.28	0.09–0.45	Fair
Final radiographic outcome	0.45	0.28–0.59	Moderate

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

**Table 4. Interobserver Agreement Among Trauma Surgeons**

	ICC	95% CI	Agreement Strength
Agreement with treatment decision	0.50	0.34–0.63	Moderate
Rating of:			
Immediate postoperative reduction or arthroplasty placement	0.52	0.37–0.65	Moderate
Immediate postoperative fixation methods	0.08	0.11–0.28	Poor
Final radiographic outcome	0.70	0.59–0.79	Good

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

Our study had several strengths as well. First, to our knowledge this is the first study to assess interobserver variability in surgeons' ratings of radiographic outcomes. Its findings may provide further insight into the reasons for poor agreement among orthopedic surgeons on both classification and treatment of PHFs. Second, our surveying of internationally renowned expert surgeons from 4 different institutions may have helped reduce single-institution bias, and it presents the highest level of expertise in the treatment of PHFs.

The overall disagreement among the surgeons in our study... highlights the need for better evidence for the optimal treatment of proximal humerus fractures in order to improve consensus.

Although the surgeons in our study moderately agreed on final radiographic outcomes of PHFs, such levels of agreement may still be clinically unacceptable.<sup>19</sup> The overall disagreement on treatment decisions highlights the need for better evidence for optimal treatment of PHFs in order to improve consensus, particularly with anticipated increases in age and comorbidities in the population in coming years.<sup>4</sup> Subgroup analysis suggested trauma fellowships may contribute to better treatment agreement, though this idea requires further study, perhaps by surveying shoulder and trauma fellowship directors and their curricula for variability in teaching treatment decision-making. The surgeons in our study agreed more on what they consider acceptable final radiographic outcomes, which is encouraging. However, treatment consensus is the primary goal. The recent publication of prospective, randomized studies is helping with this issue, but more studies are needed. It is encouraging that several are planned or under way.<sup>20-22</sup>

### Conclusion

The surgeons surveyed in this study did not agree on ideal treatment for PHFs but moderately agreed on quality of radiographic outcomes. These differences may reflect a difference in training. We conducted this study to compare experienced shoulder and trauma fellowship-trained surgeons' treatment decision-making and ratings of radiographic outcomes of PHFs when presented with the same group of patients managed at 2 level I trauma centers. We hypothesized there would be little agreement on treatment decisions, better agreement on final radiographic outcome, and a difference between decision-making and ratings of radiographic outcomes between expert shoulder and trauma surgeons. Our results showed that surgeons do not agree on the best treatment for PHFs but radiographically recognize when an operative treatment has

been performed well or poorly. Regarding treatment decisions, our results also showed that expert trauma surgeons may agree more with each other than shoulder surgeons agree with each other. These results support our hypothesis and the limited current literature. The overall disagreement among the surgeons in our study and an aging population that grows sicker each year highlight the need for better evidence for the optimal treatment of PHFs in order to improve consensus.

Dr. Jawa is Assistant Professor of Orthopaedic Surgery, Tufts University School of Medicine, Boston, Massachusetts. Dr. Yi is Resident, University of California San Francisco, San Francisco, California. Dr. Boykin is Attending Surgeon, Blue Ridge Bone and Joint Clinic, Asheville, North Carolina. Dr. Gardner is Associate Professor, Washington University, St. Louis, Missouri. Dr. Gerber is Professor, Balgrist Hospital, University of Zurich, Zurich, Switzerland. Dr. Lorich is Associate Professor, Hospital for Special Surgery, New York, New York. Dr. Walch is Professor, Unit e Epaul e (Shoulder Unit), Centre Orthop edique Santy, Lyon, France. Dr. Warner is Professor, Massachusetts General Hospital, Boston, Massachusetts.

Acknowledgments: The authors thank Hongshu Guan for assisting with statistical analysis and Paul Yannopoulos for assisting with logistics.

Address correspondence to: Paul H. Yi, MD, University of California, San Francisco, 500 Parnassus Avenue, MU 320 West, San Francisco, CA 94143 (fax, 888-435-5351; e-mail, paulyi88@gmail.com).

*Am J Orthop.* 2015;44(2):77-81. Copyright Frontline Medical Communications Inc. 2015. All rights reserved.

### References

1. Marsh JL, Slongo TF, Agel J, et al. Fracture and dislocation classification compendium – 2007: Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma.* 2007;21(10 suppl):S1-S133.
2. Bell JE, Leung BC, Spratt KF, 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):121-131.
3. McLaurin TM. Proximal humerus fractures in the elderly are we operating on too many? *Bull Hosp Jt Dis.* 2004;62(1-2):24-32.
4. Jain NB, Kuye I, Higgins LD, Warner JJP. Surgeon volume is associated with cost and variation in surgical treatment of proximal humeral fractures. *Clin Orthop.* 2012;471(2):655-664.
5. Boykin RE, Jawa A, O'Brien T, Higgins LD, Warner JJP. Variability in operative management of proximal humerus fractures. *Shoulder Elbow.* 2011;3(4):197-201.
6. Petit CJ, Millett PJ, Endres NK, Diller D, Harris MB, Warner JJP. Management of proximal humeral fractures: surgeons don't agree. *J Shoulder Elbow Surg.* 2010;19(3):446-451.
7. 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):448-455.
8. Kodali P, Jones MH, Polster J, Miniaci A, Fening SD. Accuracy of measurement of Hill-Sachs lesions with computed tomography. *J Shoulder Elbow Surg.* 2011;20(8):1328-1334.
9. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428.
10. 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.
11. Handoll HH, Ollivere BJ. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev.* 2010;(12):CD000434.
12. Boons HW, Goosen JH, van Grinsven S, van Susante JL, van Loon CJ. Hemiarthroplasty for humeral four-part fractures for patients 65 years and older: a randomized controlled trial. *Clin Orthop.* 2012;470(12):3483-3491.
13. Fjalestad T, Hole MØ, Hovden IAH, Bl ucher J, Str omsv e K. Surgical treatment with an angular stable plate for complex displaced proximal humeral fractures in elderly patients: a randomized controlled trial. *J Orthop Trauma.* 2012;26(2):98-106.

14. Fjalestad T, Hole MØ, Jørgensen JJ, Strømsøe K, Kristiansen IS. Health and cost consequences of surgical versus conservative treatment for a comminuted proximal humeral fracture in elderly patients. *Injury*. 2010;41(6):599-605.
15. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Internal fixation versus nonoperative treatment of displaced 3-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg*. 2011;20(5):747-755.
16. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Hemiarthroplasty versus nonoperative treatment of displaced 4-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg*. 2011;20(7):1025-1033.
17. Agudelo J, Schürmann M, Stahel P, et al. Analysis of efficacy and failure in proximal humerus fractures treated with locking plates. *J Orthop Trauma*. 2007;21(10):676-681.
18. Brorson S, Hróbjartsson A. Training improves agreement among doctors using the Neer system for proximal humeral fractures in a systematic review. *J Clin Epidemiol*. 2008;61(1):7-16.
19. Brorson S, Olsen BS, Frich LH, et al. Surgeons agree more on treatment recommendations than on classification of proximal humeral fractures. *BMC Musculoskelet Disord*. 2012;13:114.
20. Handoll H, Brealey S, Rangan A, et al. Protocol for the ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial: a pragmatic multi-centre randomised controlled trial of surgical versus non-surgical treatment for proximal fracture of the humerus in adults. *BMC Musculoskelet Disord*. 2009;10:140.
21. Den Hartog D, Van Lieshout EMM, Tuinebreijer WE, et al. Primary hemiarthroplasty versus conservative treatment for comminuted fractures of the proximal humerus in the elderly (ProCon): a multicenter randomized controlled trial. *BMC Musculoskelet Disord*. 2010;11:97.
22. Verbeek PA, van den Akker-Scheek I, Wendt KW, Diercks RL. Hemiarthroplasty versus angle-stable locking compression plate osteosynthesis in the treatment of three- and four-part fractures of the proximal humerus in the elderly: design of a randomized controlled trial. *BMC Musculoskelet Disord*. 2012;13:16.