JB&JS OPEN ACCESS

Let Us Agree to Disagree on Operative Versus Nonoperative Treatment for Proximal Humerus Fractures

A Multicenter International Prospective Cohort Study of Gray-Zone, Clinical Equipoise Fractures

Ruben J. Hoepelman, MD, Frank. J.P. Beeres, MD, PhD, Bryan J.M. van de Wall, MD, PhD, Christian Michelitsch, MD, Isabelle R. Bünter, MD, Reto Babst, Prof, MD, PhD, Christoph Sommer, MD, Egbert-Jan M.M. Verleisdonk, MD, PhD, Detlef van der Velde, MD, PhD, Roderick M. Houwert, MD, PhD, Rolf H.H. Groenwold, MD, PhD, and Mark van Heijl, MD, PhD, on behalf of the NEXT study group

Investigation performed at Diakonessenhuis, Sint Antonius Hospital, and University Medical Center, Utrecht, Utrecht, The Netherlands; Luzerner Kantonsspital and Kantonsspital Graubünden, Lucerne, Switzerland.

Background: Internationally, the optimal treatment strategy of proximal humerus fractures remains much debated.

Methods: To investigate whether operative treatment of displaced proximal humerus fractures is superior to nonoperative treatment, this international multicenter prospective natural experiment based on clinical equipoise was performed. Two hundred twenty-six patients with acute proximal humerus fractures presenting from July 2020 to March 2022 were included after expert panel evaluation, consisting of Dutch and Swiss surgeons with diverse ideas on optimal treatment. Patients were included when no consensus on optimal treatment was reached i.e., clinical equipoise. Follow-up was completed after 1 year (n = 191 [84%]). The primary outcome was Quick Disability of the Arm, Shoulder, and Hand (QuickDASH) after 1 year. Secondary outcomes included QuickDASH at 6 weeks and EQ5D, Subjective Shoulder Value (SSV), Numeric Rating Scale (NRS) for pain at 6 weeks and 1 year.

Results: No difference in QuickDASH score after 1 year (12.8 vs. 16.2, p = 0.73) was found. At 6 weeks, operative treatment resulted in lower NRS (4.3 vs. 3.0, p < 0.001), higher EQ5D (0.59 vs. 0.67, p = 0.011), and higher SSV (40.9

continued

The authors have no conflicts of interest to declare.

This work was part of the activities of the Natural Experiments Study Group (www.next-studygroup.org).

All available anonymized data can be shared on reasonable request after careful consideration.

The Medical Ethics Committee Utrecht (METC-U) and the Medical Research Ethics Committees United (MEC-U) confirmed that the Medical Research Involving Human Subject Act (WMO) does not apply to the study (METC-protocol number 20-169/C). The Ethikkommission Nordwest-und Zentralschweiz (EKNZ) also approved the study (proposal number 2020-00961). Patients were fully informed of the purpose and procedures of the study, and signed informed consent was obtained in agreement with the General Data Protection Regulation.

Registered in the Netherlands Trial Register NTR9357 and Swiss trial register CH 2020-00961.

This work was financially supported by the DePuy Synthes (grant number DPS-TCMF-2019-033). The organization has reviewed the manuscript, however, was not involved in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. All researchers are independent from the funder. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The study was Registered in Netherlands trial register NL9357 and Swiss trial register CH 2020-00961.

Disclosure: The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSOA/A760).

JBJS Open Access • 2025:e24.00170. openaccess.jbjs.org 2

vs. 51.4, p = 0.005). At 1 year, operative treatment resulted in higher SSV (72.1 vs. 83.7, p = 0.002), while EQ5D was comparable (0.87 vs. 0.85, p = 0.95).

Conclusion: No difference between treatments was observed in the primary outcome. Patient-tailored care may still include counseling operative treatment to patients to reduce short-term pain and/or facilitate early return to sport/work in young active patients.

Level of Evidence: Level II. See Instructions for Authors for a complete description of levels of evidence.

Introduction

Proximal humeral are the fourth most common fracture in adults and third in the elderly population[1,2]. Nonetheless, the optimal treatment strategy, i.e., operative or nonoperative treatment, remains debated, with considerable differences between countries[3,4]. Several meta-analyses have been inconclusive regarding which treatment is superior[3,4].

The largest study to date, the Proximal Fracture of the Humerus Evaluation by Randomization (PROFHER), found no differences in functional or quality-of-life (QOL) outcomes between operative and nonoperative treatments[5]. However, there has been criticism regarding the generalizability of these findings[6]. The inconsecutive inclusion of patients, low participation rate (54% declined randomization), and low number of cases per surgeon could have affected the results.

These problems are commonly encountered in surgical randomized controlled trials (RCT)[7]. The patient populations treated in daily clinical practice often differ from the highly selected populations enrolled in RCTs. By contrast, observational studies provide a better reflection of clinical practice, suffer less from low inclusion rates, and thus may provide evidence that is more applicable to clinical practice. The results of observational studies could therefore complement those of RCTs and provided potential biases are adequately controlled for[7,8]. A natural experiment, e.g., based on practice variation, could be a suitable design, when recommendations for operative or nonoperative management are largely influenced by training of the treating surgeons instead of patient characteristics[9].

The aim of this multicenter international prospective cohort study was to compare operative and nonoperative treatment of proximal humeral fractures in 2 countries with different preferences for treatment of proximal humerus fractures, i.e., Switzerland and the Netherlands[10,11]. Patients were only included when there was no consensus among experts on optimal treatment (clinical equipoise).

Methods

A detailed description of the methods used in this study is available in the published protocol and further elaboration in the supplementary material[12]. An abbreviated description is provided below. This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[13].

Study Design

This international multicenter prospective cohort study included patients from 3 Dutch trauma centers with a predominant preference for nonoperative treatment: Diakonessenhuis, level-2 with 4 trauma specialists; Sint Antonius Hospital, level-2 with 6 trauma specialists; and the University Medical Center, Utrecht, level-1 with 8 trauma specialists and 2 trauma centers from Switzerland with a predominant preference for operative treatment: Luzerner Kantonsspital, level-1 with 5 trauma (shoulder) specialists and Kantonsspital Graubünden, level-1 with 3 trauma (shoulder) specialists. These differences in treatment preference between Dutch and Swiss trauma centers are the result of local education, culture, and conviction. These hospitals were expected to treat similar groups of patients, since trauma patients typically receive care from the nearest appropriate hospital. Treatment allocation in this study is therefore largely dependent on the location of the incident (i.e., geographical randomization) and only marginally influenced by patients' characteristics, thus providing the circumstances for a natural experiment[9]. To further reduce potential confounding, inclusion was restricted to patients in whom there was clinical equipoise (according expert panel evaluation), and therefore, this study investigates a gray zone of patients in whom optimal treatment is unknown[9].

Patients

All consecutive patients, aged 18 years and older, with an acute displaced proximal humerus fracture were eligible for inclusion. Exclusion criteria were delayed presentation (>3 weeks), open fractures, pathological fractures, associated injuries requiring immediate surgery (e.g., soft tissue or neurovascular injury), dislocation of the shoulder joint, concomitant ipsilateral upper extremity fractures, and inability to complete follow-up questionnaires (e.g., cognitive impairment or patients not living in hospital area and unable to complete follow-up).

Patient characteristics included age, gender, body mass index (BMI), smoking status, osteoporosis, diabetes, American Society of Anesthesia (ASA) score[14], trauma mechanism (high or low according to the Advances Trauma Life Support [ATLS] guidelines)[15], and fracture classification (according to AO foundation/Orthopaedic Trauma Association [AO/OTA] 2018)[16,17].

JBJS Open Access • 2025:e24.00170. openaccess.jbjs.org 3

Interventions

Nonoperative treatment consisted of a sling immobilization for 3 to 6 weeks, adequate pain management, and guided physiotherapy according to local hospital protocol and needs of the patients. Operative treatment consisted of minimal invasive plate osteosynthesis (MIPO), open reduction and internal fixation, intramedullary nailing, or reversed shoulder arthroplasty (RSA). Implants used were the PHILOS plate, MultiLoc humeral nail by DePuy Synthes, and the Medacta reverse shoulder prothesis.

Expert Panel

A visual representation of the expert panel process is available in Figure S1. The expert panel consisted of 3 Dutch trauma surgeons and 3 Swiss trauma surgeons. All relevant data to reach a clinical decision were made available to panel members, including demographic and clinical information radiographs and Computed Tomography (CT) scans. Panel members decided on the preferred management, operative or nonoperative treatment, for each individual case. The decision was made independently from other members and blinded for the treatment that had already been started. Patients were included in the study if the majority of experts in one country, i.e., minimally 2 of 3, disagreed with the received treatment in the other country. For example, patients were eligible if they received operative treatment based on the expert opinions in e.g., Switzerland, while in fact they received nonoperative treatment in the Netherlands, and vice versa. This process led to a group of patients in whom there is disagreement on the optimal treatment management. Hereby, we identified patients with a similar profile and by zooming in on a patient cohort with a similar profile, confounding by patient characteristics will, to a large extent, be mitigated. The disadvantage of this method is that expert opinion is biased by their clinical experience. This methodology is explained in more detail in the published study protocol and other articles from the NEXT study group[9,12].

Outcomes

The primary outcome was the Quick Disability of the Arm, Shoulder, and Hand (QuickDASH) questionnaire at 12 months after treatment. The QuickDASH is a patient-reported outcome instrument developed to measure upper extremity disability and symptoms, resulting in a score ranging from no disability (0) to most severe disability (100)[18]. A difference of 8 points on the QuickDASH is considered a clinically relevant difference[19].

Secondary outcomes were the QuickDASH at 6 weeks, the Subjective Shoulder Value (SSV)[20], EuroQol five-dimensional questionnaire (EQ-5D)[21], Numerical Rating Scale (NRS) Pain score, and return to sporting activity and return to work activity at 6 weeks and 12 months.

Other secondary outcomes included revision surgery, implant removal, and complications, including nonunion[22], malunion[22], infection[23], implant failure, and revision surgery, defined as the need for secondary surgical treatment other than implant removal.

Sample Size and Statistical Analyses

A previous study found that patients who underwent MIPO surgery had a QuickDASH standard deviation (SD) of 14[24]. To detect a difference of 8 points with 80% power and a type-I error probability (alpha) of 0.05 would require 2 × 50 participants. Accounting for 20% loss to follow-up and additional 40% loss due to propensity score (PS) matching (see below), the total sample size was calculated to be 220 participants (i.e., 110 per treatment arm).

All analyses were performed using R statistical software v4.1.2[25]. To control for potential confounding, PS matching was performed. The PS was estimated using binary logistic regression analysis, with operative treatment as the dependent variable and age, sex, BMI, ASA score, trauma mechanism, and fracture classification as prespecified covariates in the model. A 1:1 nearest neighbor matching was performed, with a maximum caliper of 0.2 of the SD of the natural logarithm of the PS using the MatchIt algorithm in R.

Results

Expert Panel Evaluation

From July 2020 to March 2022, 830 patients were sent for expert panel evaluation. There was 66% agreement of treatment among experts in our cohort. In 33% (129 operative and 150 nonoperative) of the cases, experts from the Netherlands and Switzerland disagreed with each other on optimal treatment and, thus, clinical equipoise was found (Fig. 1).

Patients

After expert panel evaluation, 226 patients were included (Fig. 1). All 115 nonoperative patients were from Dutch hospitals, and all 111 operative patients were from Swiss hospitals. Follow-up was completed within March 2023 with a completion rate of 84%. Baseline characteristics before and after PS matching are presented in Table I, showing adequate balance on potential confounders after matching 2 × 98 patients. Baseline characteristics and outcomes stratified per center are available in Tables S1 and S2. The mean age of included patients was 65.3 ± 13.6 years, 72% (n = 162) was female, ASA score was predominantly I and II (n = 179 [79%]), and distribution of AO classification was as follows: type AO 11A n = 66 (29%), type AO 11B n = 67 (30%), and type AO 11C n = 93 (41%). The mean time from fracture to clinical and radiological reassessment at the outpatient clinic visit was 8.4 ± 2.8 days for nonoperatively treated patients. The mean time from fracture to surgery was 3.7 ± 4.8 days for operatively treated patients.

Primary Outcome

No differences were observed between nonoperative and operative treatment in QuickDASH scores after 12 months: 12.8 (±14.9) vs. 16.2 (±16.7) for nonoperative and operative treatment, respectively (regression coefficient [b] 1.4, 95% confidence interval [CI] −6.2 to 8.9; p = 0.73), see Table II. Multiple regression analysis of the multiple imputed data sets, while adjusting for confounding factors, yielded similar results for all outcomes (see Table S3).

Secondary Outcomes

Secondary outcomes are reported in Table II. Six weeks after treatment, operative treatment was associated with a higher

JBJS Open Access • 2025:e24.00170. openaccess.jbjs.org

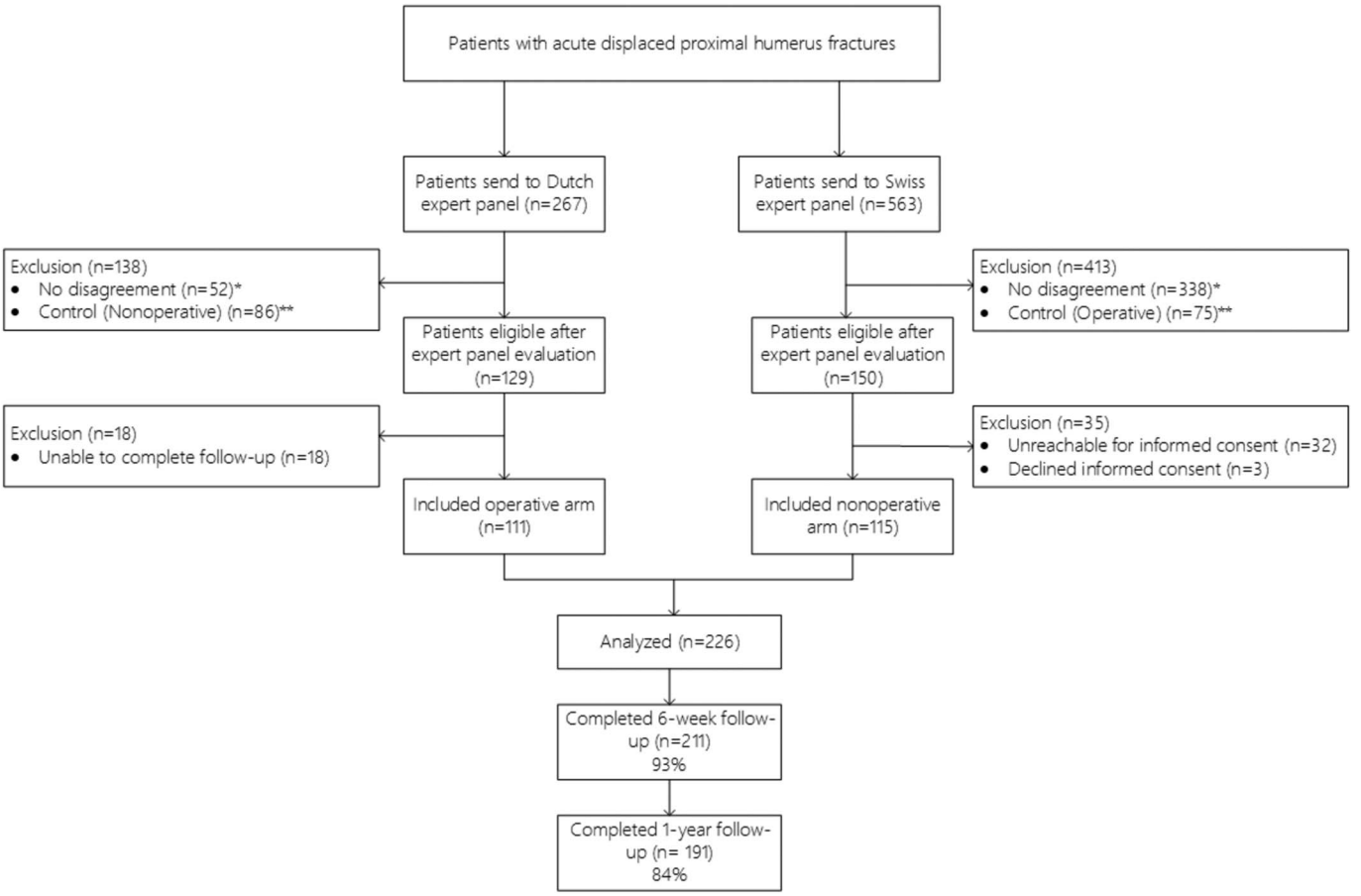

Fig. 1

Flow chart of inclusions and follow-up per treatment type. *No disagreement = These were operatively treated patients sent to the Dutch expert panel and nonoperatively treated patients sent to the Swiss expert panel where there was consensus between experts from the different countries on optimal treatment. Therefore, these patients were not included in the study as there was no clinical equipoise. **Control = These were nonoperatively treated patients in Switzerland sent to the Dutch expert panel, and operatively treated patients in the Netherlands sent to the Swiss expert panel, to check the hypothesis that patients treated operatively in the Netherlands would also be operated on in Switzerland and vice versa. Therefore, these patients were not included in the study as there was no clinical equipoise.

EQ5D -index score 0.59 vs. 0.67 ($p = 0.015$), a higher SSV 40.0 vs. 51.4 ($p = 0.005$), and lower NRS Pain score 4.3 vs. 3.0 ($p < 0.001$) compared with nonoperative treatment. No differences were observed in QuickDASH scores (50.4 vs 52.9) and EQ5D health scores (68.1 vs. 69.0) between operative and nonoperative treatment, respectively.

One year after trauma, operative treatment was associated with a higher SSV 72.1 vs. 83.7 ($p = 0.002$) and higher EQ5D health score 74.1 vs. 81.5 ($p = 0.002$) compared with nonoperative treatment.

In both the nonoperative and operative treatment groups, 10 patients underwent (secondary) surgery. Six patients in the nonoperative group underwent RSA: in 3 cases due to symptomatic malunion, one symptomatic nonunion, one avascular head necrosis, and one because of unmanageable pain. Four patients underwent plate osteosynthesis due to unmanageable pain within 3 weeks after trauma. Ten patients underwent secondary surgery in the operative group: 5 to implant failure, 3 to loss of reduction, and 2 to infections. Twenty patients underwent implant removal within the first year.

No differences were observed in patients aged 60 years or younger, except for the SSV at 6 weeks and 1 year (Table III).

Discussion

This natural experiment shows an agreement rate of 66% between 2 countries with diverging preferences on treatment strategy for proximal humerus fractures. The remaining 33%, consisting mostly of 3-part and 4-part fractures, is investigated here. No difference in QuickDASH score was found (12.8 vs. 16.2, 95% CI −6.2 to 8.9). Operative treatment resulted in favorable secondary outcomes compared with nonoperative treatment (SSV after 6 weeks and 1 year, EQ5D index score after 6 weeks, and EQ5D visual health score after 1 year).

Some of these results are consistent with the PROFHER, which found no significant differences in Oxford Shoulder Score,

JBJS Open Access • 2025:e24.00170. openaccess.jbjs.org

TABLE I Baseline Characteristics Before and After Propensity Score Matching of Patients with Proximal Humerus Fractures

	Before PS Matching			After PS Matching		
Baseline variables	Nonoperative (n = 115)	Operative (n = 111)	SMD	Nonoperative (n = 98)	Operative (n = 98)	SMD*
Age (mean ± SD)	66.5 ± 12.3	64 ± 14.7	0.176	65.9 ± 12.7	65.0 ± 14.8	0.064
Male (n, %)	26 (23)	38 (34)	0.260	24 (25)	27 (28)	0.070
ASA-score (n, %)			0.084			0.025
1-2	93 (81)	86 (77)		77 (79)	76 (78)	
>2	22 (19)	25 (23)		21 (21)	22 (22)	
BMI (mean ± SD)	27.6 ± 5.0	26.8 ± 5.4	0.379	27.4 ± 5.1	27.1 ± 5.5	0.045
Current smoker (n, %)	12 (11.4)	16 (14.4)	0.089	13 (13)	11 (11)	0.062
Osteoporosis (n, %)	6 (5)	6 (5)	0.008	5 (5)	6 (6)	0.044
Trauma mechanism (n, %)			0.012			0.043
Low energy fall	75 (65)	73 (66)		64 (66)	64 (66)	
Other	40 (35)	38 (34)		32 (33)	34 (35)	
AO classification/Neer classification			0.042			0.096
11A/2-part	34 (30)	32 (29)		26 (27)	30 (30)	
11B/3-part	33 (29)	34 (31)		32 (33)	29 (30)	
11C/4-part	48 (42)	45 (41)		40 (40)	39 (40)	
Treatment type						
Nonoperative	115 (100)	0 (0)		97 (100)	0 (0)	
Minimal invasive plate osteosynthesis	0 (0)	58 (52)		0 (0)	54 (55)	
Intramedullary nailing	0 (0)	9 (8)		0 (0)	8 (8)	
Open reduction and internal fixation (plate)	0 (0)	27 (25)		0 (0)	21 (22)	
Reversed shoulder arthroplasty	0 (0)	17 (15)		0 (0)	15 (15)	
Days to clinical control or surgery	8.4 ± 2.8	3.7 ± 4.8		8.3 ± 2.9	3.6 ± 4.4	

*ASA = American Society of Anesthesiologists score, BMI = body mass index, and SMD = standardized mean difference; SMD <0.1 indicates adequate matching.

at any time point between either treatment groups[5]. Contrary to our study, no differences were observed in QOL measurements. It should be mentioned that most of the patient included in the PROFHER had a 2-part (n = 128, 51%) or 3-part fracture (n = 90, 36%), and only 11 (4%) patients had a 4-part fracture, while in this study, 67 (30%) patients had a 3-part and 93 (42%) patients had a 4-part fracture. Another multicenter RCT did not find any differences between treatment modalities either[26]. They compared Philos plating with nonoperatively treated patients with displaced 2-part proximal humerus fractures. Their primary outcome (DASH-score) and all secondary outcomes did not show any differences. Neither did a final study comparing plate fixation and hemiarthroplasty to nonoperative treatment in 3-part and 4-part fractures[27]. Older RCTs and meta-analysis did not show convincing benefit of surgery either[4,28–30].

Although this study is not an RCT, which is still considered to provide the highest level of evidence, its results and design are complementary to current evidence and an example of a more feasible, cheaper, and faster alternative study design, providing valuable evidence. Despite the COVID pandemic which hindered patient inclusion, this study took 2.5 years from the start of inclusion to completion of follow-up with 5 participating centers, recruiting 226 participants, whereas the PROFHER took 5 years with 33 participating centers. Furthermore, one of the critiques of the PROFHER is the low surgeon to patient ratio (1.5 case per surgeon). In this study, all 111 patients were operated by shoulder specialists in 2 centers. Finally, its methodological design, where patients were included in case of clinical equipoise, which was due to cultural/geographical differences and not on patient characteristics, shows the potential for natural experiments. This is confirmed by the comparability of the groups before matching and low dropout rate after matching.

There were some limitations. Although we believe that this design limits potential bias by unmeasured confounding factors, it is not a RCT and therefore does not rule out unmeasured confounding. Regarding the observed potential confounders, the standardized mean difference was below 0.1 for all baseline characteristics, which indicates adequate matching, and therefore, bias due to confounders was mostly

JBJS Open Access • 2025:e24.00170.

TABLE II Outcomes in a Propensity Score Matched Cohort of Nonoperatively and Operatively Treated Patients with Proximal Humerus Fractures*

	Outcome Variables					
6-wk Outcomes Mean ± SD or n (%)	Nonoperative (n = 98)	Operative (n = 98)	Regression Coefficient (b)	95% CI	SE	p
QuickDASH	50.4 ± 15.4	52.9 ± 19.2	1.3	−4.9 to 7.7	3.2	0.68
EQ5D index score	0.59 ± 0.2	0.67 ± 0.2	0.08	0.0 to 0.1	0.0	0.011
EQ5D health score	68.1 ± 17.9	69.0 ± 18.6	2.9	−2.9 to 8.9	3.0	0.32
Subjective Shoulder Value (SSV)	40.9 ± 21.9	51.4 ± 22.2	10.1	2.9 to 17.1	3.6	0.005
Numeric Rating Scale (NRS) Pain score	4.3 ± 2.0	3.0 ± 2.0	−1.3	−1.9 to −0.6	0.4	<0.001
1-yr Outcomes Mean ±SD or n (%)			Regression Coefficient (b)	95% CI	SE	p
QuickDASH	12.8 ± 14.9	16.2 ± 16.7	1.4	−6.2 to 8.9	3.9	0.73
EQ5D index score	0.87 ± 0.1	0.85 ± 0.2	0.0	−0.1 to 0.1	0.0	0.95
EQ5D health score	74.1 ± 14.2	81.5 ± 15.3	8.8	3.2 to 14.3	2.8	0.002
Subjective Shoulder Value (SSV)	72.1 ± 19.0	83.7 ± 15.1	11.7	4.3 to 19.1	3.8	0.002
Numeric Rating Scale (NRS) Pain score	1.2 ± 1.9	1.6 ± 1.8	0.3	−0.5 to 1.0	0.4	0.52
Time to work in wk	4.6 ± 5.5	4.7 ± 4.3	−0.9	−7.1 to 5.3	3.2	0.77
Time to sport in wk	14.8 ± 13.1	9.5 ± 7.1	−6.2	−16.2 to 3.8	5.1	0.22
Secondary Outcomes/Complications			OR	95% CI	SE	p
Revision/secondary surgery	10 (10)	10 (10)	0.9	0 to inf	NA	0.99
Implant removal	0 (0)	20 (20)	—	—	—	—
Secondary RSA	6 (6)	—	—	—	—	—
Implant failure	—	5 (5)	—	—	—	—
Loss of reduction	—	3 (3)	—	—	—	—
Nerve palsy	—	2 (2)	—	—	—	—
Infection	—	2 (2)	—	—	—	—
Persistent pain	6 (7)	1 (1)	NA	NA	NA	0.99
AVN	1 (1)	0 (0)	NA	NA	NA	0.99
Symptomatic malunion	3 (3)	0 (0)	NA	NA	NA	0.99
Symptomatic nonunion	1 (1)	1 (1)	NA	NA	NA	0.99

*AVN = avascular head necrosis, CI = confidence interval, DASH = Disability of the Arm, Shoulder, and Hand Questionnaire, OR = odds ratio, SD = standard deviation, SE = standard error, and RSA = reverse shoulder arthrodesis.

eliminated[31]. Moreover, multiple regression and the complete case analysis showed similar results, and treatment allocation was more dependent on the geographical location of the accident than on patient characteristics; therefore, we believe the potential impact of unmeasured confounding is limited.

Second, we investigated several operative techniques, which make it difficult to generalize a particular procedure; however, this does not steer away from the primary comparison, operative vs. nonoperative treatment.

Third, although all questionnaires are validated for the languages and we presume the Dutch and Swiss society to be comparable, we do not have evidence to support this. Therefore, it is possible that patient-reported outcomes are susceptible to bias introduced by cultural differences.

Last, we opted for short-term (6-week) outcomes instead of 3-month outcomes. Three-month outcomes would have been informative and could be considered in future research. Furthermore, a longer follow-up (such as 2 or 5 years) is needed to accurately capture outcomes and complications such as implant removal, avascular (head) necrosis, and arthrosis.

Considering these limitations and unique methodological aspects, it should be mentioned that even for experts from 2 countries with diverse preference in treatment modality, there was agreement in 52% of the conservative cases and 15% of the operative cases. Agreement for operative treatment between the

JBJS Open Access • 2025:e24.00170.

TABLE III Outcomes of Nonoperatively and Operatively Treated Patients of 60 Years and Younger with Proximal Humerus Fractures, Analyzed With Multivariable Regression Analysis

Outcome Variable			Complete Case Analysis				Multiple Imputed Data			
6 wk Outcomes Mean ± SD or n (%)	Nonoperative (n = 30)	Operative (n = 39)	Regression Coefficient (b)	95% CI	SE	p	Regression Coefficient (b)	95% CI	SE	p
QuickDASH*	51.9 ± 15.2	52.7 ± 19.2	0.9	−11.3 to 13.1	6.0	0.89	2.3	−8.4 to 13.1	5.5	0.67
EQ5D index score*	0.65 ± 0.1	0.70 ± 0.2	0.07	−0.1 to 0.2	0.1	0.23	0.05	0.04 to 0.15	0.1	0.28
EQ5D health score*	64.8 ± 15.9	69.2 ± 15.7	8.4	−3.6 to 20.5	5.9	0.16	6.5	−5.1 to 18.1	5.9	0.27
Subjective Shoulder Value (SSV)*	38.6 ± 15.7	51.7 ± 23.4	27.9	12 to 43.4	3.6	<0.001	14.5	2.1 to 26.9	6.3	0.022
Numeric Rating Scale (NRS) Pain score*	4.1 ± 2	3.1 ± 2	−0.7	−2.0 to −0.5	0.6	0.26	−0.9	−2.1 to −0.3	0.3	0.15
1 yr Outcomes Mean ± SD or n (%)										
QuickDASH*	12.6 ± 13.5	12.3 ± 16.8	−3.1	−12.0 to 5.9	4.4	0.49	1.2	−8.8 to 11.4	5.1	0.80
EQ5D index score*	0.92 ± 0.1	0.88 ± 0.2	0.0	−0.1 to 0.1	0.0	0.88	0.0	−0.2 to 0.1	0.1	0.51
EQ5D health score*	82.6 ± 10.4	88.8 ± 11.1	7.3	4.5 to 14.0	3.6	0.052	5.5	−1.9 to 12.9	3.7	0.14
Subjective Shoulder Value (SSV)*	75. 0 ± 19.3	88.3 ± 15.2	17.3	5.3 to 29.3	5.8	0.006	11.5	0.7 to 22.4	5.5	0.038
Numeric Rating Scale (NRS) Pain score*	1.3 ± 2.1	1.3 ± 1.9	−0.7	−2.2 to 0.7	0.7	0.31	−0.1	−1.3 to 1.5	0.6	0.93
Time to work in wk*	9.9 ± 6.5	8.1 ± 5.6	−2.5	−7.7 to 2.8	2.5	0.34	−0.9	−6.1 to 4.4	2.7	0.74
Time to sport in wk*	21.3 ± 12.7	11.7 ± 5.9	−8.6	−18.2 to 0.9	4.6	0.074	−7.4	−17.1 to 2.9	4.9	0.13
Secondary Outcomes/Complications			OR	95% CI	SE	p	OR	95% CI	SE	p
Revision/secondary surgery†	3 (10)	2 (5)	1.3	0.02 to 60.0	15	0.87	1.2	0.1 to 18.3	1.4	0.91
Implant removal	0 (0)	14 (36)	—	—	—	—	—	—	—	—
Secondary RSA	1 (3)	0 (0)	—	—	—	—	—	—	—	—
Implant failure	0 (0)	0 (0)	—	—	—	—	—	—	—	—
Loss of reduction	0 (0)	0 (0)	—	—	—	—	—	—	—	—
Nerve palsy	0 (0)	2 (5)	—	—	—	—	—	—	—	—
Infection	0 (0)	0 (0)	—	—	—	—	—	—	—	—
Persistent pain†	6 (20)	1 (3)	3.8	0.0 to 5.6	1.4	0.49	0.05	0.0 to 0.8	1.4	0.030
AVN†	1 (1)	0 (0)	—	—	—	—	—	—	—	—
Symptomatic malunion†	1 (1)	0 (0)	—	—	—	—	—	—	—	—
Symptomatic nonunion†	1 (1)	1 (1)	—	—	—	—	—	—	—	—

*Multivariate linear regression analysis. †Multivariate binary logistic regression analysis; AVN = avascular head necrosis, DASH = Disability of the Arm, Shoulder, and Hand Questionnaire, OR = odds ratio, and RSA = reverse shoulder arthrodesis. Covariates included in the regression model were; age, sex, BMI, ASA score, trauma mechanism, and fracture classification.

Swiss and the Dutch Team was particularly high in for severely displaced fractures, which means that in these cases there is no discussion for the need of operative treatment. As mentioned previously, results apply to the remaining 33% in whom there is no consensus.

The QuickDASH was the primary outcome as it is the most widely used questionnaire for shoulder function and has been validated for proximal humerus fractures. A score <15 is considered as excellent function, which was reached by most patients (median of 9 nonoperative vs. 14 operative) in both treatment groups[32]. More importantly, no differences were found at either follow-up points. These results are in line with previous studies, and therefore, we argue that the patients in whom there is no clear consensus, there is no benefit 1 year after injury between nonoperative or operative treatment. This would imply that conservative management should be used.

Contrary, it could be argued that no long-term benefit was detected because of strong coping mechanisms of the patients in both groups. Treatment effects tend to dissolve over time as coping mechanisms dealing with limitations in

shoulder function, increase. This would suggest that the benefit of surgery is mainly measurable in the first months after surgery. Interestingly, this appears to be the case.

At 6 weeks, the differences in SSV (40.9 vs. 51.4), NRS scores (4.3 vs. 3.0), and the EQ5D index score (0.59 vs. 0.67) favored operative treatment. Therefore, in our opinion, patients could be counseled that surgery might reduce pain and that surgery is considerable for patients with inadequately manageable during nonoperative treatment, the latter being the reason for surgery in 4 cases in the nonoperative group.

Last, surgery might be considered in some young/active patients for early return to sports/work. This study was designed to investigate the average effect of operative and nonoperative treatment on a group level. Although the study was underpowered to identify subgroups in whom operative treatment is superior to nonoperative treatment, the results regarding time to return to work and time to return to sports are suggestive of a beneficial effect of operative treatment. This may indicate that in clinical practice, there is a place for operative treatment in proximal humerus fracture patients who are in severe pain or young and active. This study does not confirm operative treatment to be superior in that patient subgroup, and future studies could shed more light on that particular question. In the meantime, operative treatment should be part of the shared decision-making process.

In conclusion, no difference between treatments of equipoise fractures was observed in the primary outcome. Patient-tailored care should include counseling on conservative vs. operative treatment, whereas operative treatment may reduce short-term pain, facilitate early return to work and sport, while the longer term benefits of either treatment have not yet been formally investigated.

Appendix

eA Supporting material provided by the author is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSOA/A761). This content has not been copyedited or verified. ■

Note: This work was part of the activities of the Natural Experiments Study Group (www.next-studygroup.org).NEXT-study group members: Collaborators (not meeting inclusion criteria): Dr. Diederik P.J. Smeeing, Rijnstate Hospital, Department of Trauma Surgery, Arnhem, the Netherlands. Prof. Dr. Job N. Doornberg, University Medical Center Groningen, Department of Orthopedic Surgery, Groningen, the Netherlands. Dr. Frank F.A. IJpma, University Medical Center Groningen, Department of Trauma Surgery, Groningen, the Netherlands. Dr. Bas van Wageningen, Radboud University Medical Center, Department Trauma Surgery, Nijmegen, the Netherlands. Advisory board (not meeting inclusion criteria): Dr. Marilyn Heng, MD, MPH, FRCSC, Orthopaedic Trauma Surgeon Massachusetts General Hospital|Assistant Professor of Orthopaedic Surgery, Harvard Medical School. Drs. Mike Hogervorst, University Medical Center Groningen, Department of Trauma Surgery, Groningen, the Netherlands. Prof. Dr. Moyo C. Kruyt, University Medical Center Utrecht, Department of Orthopedic Surgery, Utrecht, the Netherlands. Simon M. Lambert, BSc MBBS FRCS FRCSE Orth, University College London Hospital NHS Foundation Trust. Dr. Luke P.H. Leenen, University Medical Center Utrecht, Department of Trauma Surgery, Utrecht, the Netherlands. Prof. Dr. Cumhur F, Öner, University Medical Center Utrecht, Department of Orthopedic Surgery, Utrecht, the Netherlands.Prof. Dr. Rudolf W. Poolman, Leiden University Medical Centre, Leiden, the Netherlands || OLVG, Amsterdam, the Netherlands.

Ruben J. Hoepelman, MD[1]
Frank. J.P. Beeres, MD, PhD[2,3]
Bryan J.M. van de Wall, MD, PhD[2,3]
Christian Michelitsch, MD[4]
Isabelle R. Bünter, MD[2]
Reto Babst, Prof, MD, PhD[2,3]
Christoph Sommer, MD[4]
Egbert-Jan M.M. Verleisdonk, MD, PhD[5]
Detlef van der Velde, MD, PhD[6]
Roderick M. Houwert, MD, PhD[1]
Rolf H.H. Groenwold, MD, PhD[7,8]
Mark van Heijl, MD, PhD[5]

[1]Department of Trauma Surgery, University Medical Center Utrecht, Utrecht, the Netherlands

[2]Department of Orthopedic and Trauma Surgery, Luzerner Kantonsspital, Lucerne, Switzerland

[3]University of Lucerne, Faculty of Health sciences and Medicine, Lucerne, Switzerland

[4]Department of Trauma Surgery, Kantonsspital Graubünden, Chur, Switzerland

[5]Department of Trauma Surgery, Diakonessenhuis, Utrecht, the Netherlands

[6]Department of Trauma Surgery, St. Antonius Hospital Nieuwegein, the Netherlands

[7]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

[8]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

E-mail address for R.J. Hoepelman: r.j.hoepelman-2@umcutrecht.nl

References

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. Injury. 2006;37(8):691-7.
2. Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures. Clin Orthop Relat Res. 2006;442:87-92.
3. Handoll HHG, Brorson S. Interventions for treating proximal humeral fractures in adults. Cochrane Database Syst Rev. 2015;2015(11):CD000434.
4. Beks RB, Ochen Y, Frima H, Smeeing DPJ, van der Meijden O, Timmers TK, van der Velde D, van Heijl M, Leenen LPH, Groenwold RHH, Houwert RM. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. J Shoulder Elbow Surg. 2018;27(8):1526-34.
5. Rangan A, Handoll H, Brealey S, Jefferson L, Keding A, Martin BC, Goodchild L, Chuang LH, Hewitt C, Torgerson D, PROFHER Trial Collaborators. Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus the PROFHER randomized clinical trial. JAMA. 2015;313(10):1037-47.
6. von Keudell A, Vrahas MS, Weaver MJ. Surgical versus nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. J Orthop Trauma. 2016;30(4):e143.
7. Houwert RM, Beks RB, Dijkgraaf MGW, Roes KCB, Öner FC, Hietbrink F, Leenen LPH, Groenwold RHH. Study methodology in trauma care: towards question-based study designs. Eur J Trauma Emerg Surg. 2021;47(2):479-84.
8. Beks RB, Bhashyam AR, Houwert RM, van der Velde D, van Heijl M, Smeeing DPJ, Hietbrink F, Leenen LPH, Groenwold RHH. When observational studies are as helpful as randomized trials: examples from orthopedic trauma. J Trauma Acute Care Surg. 2019;87(3):730-2.

9. van de Wall BJM, Stadhouder A, Houwert RM, Oner FC, Beeres FJP, Groenwold RHH, NEXT Study Group. Natural experiments for orthopaedic trauma research: an introduction. Injury. 2023;54(2):429-34.

10. Beerekamp MSH, de Muinck Keizer RJO, Schep NWL, Ubbink DT, Panneman MJM, Goslings JC. Epidemiology of extremity fractures in The Netherlands. Injury. 2017;48(7):1355-62.

11. Tepass A, Blumenstock G, Weise K, Rolauffs B, Bahrs C. Current strategies for the treatment of proximal humeral fractures: an analysis of a survey carried out at 348 hospitals in Germany, Austria, and Switzerland. J Shoulder Elbow Surg. 2013; 22(1):e8-e14.

12. Hoepelman RJ, Ochen Y, Beeres FJP, Frima H, Sommer C, Michelitsch C, Babst R, Buenter IR, van der Velde D, Verleisdonk EJMM, Groenwold RHH, Houwert RM, van Heijl M. Let's Agree to Disagree on Operative versus Nonoperative (LADON) treatment for proximal humerus fractures: study protocol for an international multicenter prospective cohort study. PLoS One. 2022;17(2):e0264477.

13. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP., STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806-8.

14. Abouleish AE, Leib ML, Cohen NH. ASA provides examples to each ASA physical status class. ASA Newsl. 2015;79(6):38-49.

15. ATLS Subcommittee, American College of Surgeons' Committee on Trauma, International ATLS Working Group. Advanced trauma life support (ATLS): the ninth edition. J Trauma Acute Care Surg. 2013;74(5):1363-6.

16. Meinberg E, Agel J, Roberts C, Karam M, Kellam J. Fracture and dislocation classification compendium—2018. J Orthop Trauma. 2018;32(suppl 1):S1-S170.

17. Marongiu G, Leinardi L, Congia S, Frigau L, Mola F, Capone A. Reliability and reproducibility of the new AO/OTA 2018 classification system for proximal humeral fractures: a comparison of three different classification systems. J Orthop Traumatol. 2020;21(1):4.

18. Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. BMC Musculoskelet Disord. 2006;7:44.

19. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the arm, shoulder, and hand questionnaire (QuickDASH) and numeric pain rating scale in patients with shoulder pain. J Shoulder Elbow Surg. 2009;18(6):920-6.

20. Gilbart MK, Gerber C. Comparison of the subjective shoulder value and the Constant score. J Shoulder Elbow Surg. 2007;16(6):717-21.

21. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.

22. Vigni GE, Sabbioni G, Tomasi A, Gervasi E. Nonunion and malunion in proximal humerus fractures. Lo Scalpello Otodi Educ. 2020;34(1):3-8.

23. Metsemakers WJ, Morgenstern M, McNally MA, Moriarty TF, McFadyen I, Scarborough M, Athanasou NA, Ochsner PE, Kuehl R, Raschke M, Borens O, Xie Z, Velkes S, Hungerer S, Kates SL, Zalavras C, Giannoudis PV, Richards RG, Verhofstad MHJ. Fracture-related infection: a consensus on definition from an international expert group. Injury. 2018;49(3):505-10.

24. Frima H, Michelitsch C, Beks RB, Houwert RM, Acklin YP, Sommer C. Long-term follow-up after MIPO Philos plating for proximal humerus fractures. Arch Orthop Trauma Surg. 2019;139(2):203-9.

25. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available at: http://www.r-project.org/index.html. Accessed March 3, 2022.

26. Launonen AP, Sumrein BO, Reito A, Lepola V, Paloneva J, Jonsson KB, Wolf O, Ström P, Berg HE, Felländer-Tsai L, Jansson KÅ, Fell D, Mechlenburg I, Døssing K, Østergaard H, Märtson A, Laitinen MK, Mattila VM, as the NITEP group. Operative versus non-operative treatment for 2-part proximal humerus fracture: a multicenter randomized controlled trial. Plos Med. 2019;16(7):e1002855.

27. Launonen AP, Sumrein BO, Reito A, Lepola V, Paloneva J, Berg HE, Felländer-Tsai L, Kask K, Rahnel T, Tootsi K, Märtson A, Jonsson KB, Wolf O, Ström P, Døssing K, Østergaard HK, Mechlenburg I, Mattila VM, Laitinen MK. Surgery with locking plate or hemiarthroplasty versus nonoperative treatment of 3-4-part proximal humerus fractures in older patients (NITEP): an open-label randomized trial. PLoS Med. 2023; 20(11):e1004308.

28. Sanders RJ, Thissen LG, Teepen JC, van Kampen A, Jaarsma RL. Locking plate versus nonsurgical treatment for proximal humeral fractures: better midterm outcome with nonsurgical treatment. J Shoulder Elbow Surg. 2011;20(7):1118-24.

29. 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):747-55.

30. Fjalestad T, Hole MØ, Hovden IAH, Blücher J, Strømsøe 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.

31. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011;46(3):399-424.

32. Angst F, Schwyzer H, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: disabilities of the arm, shoulder, and hand questionnaire (DASH) and its short version (*Quick* DASH), shoulder pain and disability index (SPADI), American shoulder and elbow surgeons (ASES) society standardized shoulder assessment form, constant (Murley) score (CS), simple shoulder test (SST), Oxford shoulder score (OSS), shoulder disability questionnaire (SDQ), and western ontario shoulder instability index (WOSI). Arthritis Care Res (Hoboken). 2011;63(suppl 11):S174-S188.