Challenges and opportunities in trauma research: study designs and patient-reported outcome measures

Ochen, Y.

Author: Ochen, Y.
Title: Challenges and opportunities in trauma research: study designs and patient-reported outcome measures
Issue Date: 2021-03-31

CHAPTER 3

Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials

Reinier B. Beks, Yassine Ochen, Herman Frima, Diederik P.J. Smeeing, Olivier van der Meijden, Tim K. Timmers, Detlef van der Velde, Mark van Heijl, Luke P.H. Leenen, Rolf H.H. Groenwold, R. Marijn Houwert

(Journal of Shoulder and Elbow Surgery)

Abstract

Background

There is no consensus on the choice of treatment for displaced proximal humeral fractures in older (>65 years) patients. The aim of this systematic review and meta-analysis was (1) to compare operative with nonoperative management of displaced proximal humeral fractures and (2) to compare effect estimates obtained from randomized controlled trials (RCTs) and observational studies.

Methods

The databases of MEDLINE, Embase, CENTRAL, and CINAHL were searched on September 5, 2017 for studies comparing operative versus nonoperative treatment of proximal humeral fractures; both RCTs and observational studies were included. The MINORS criteria, a validated instrument for methodological quality assessment, were used to assess study quality. The primary outcome measure was physical function as measured by the absolute Constant-Murley score after operative or nonoperative treatment. Secondary outcome measures were major reinterventions, nonunion, and avascular necrosis.

Results

We included 22 studies comprising 7 RCTs and 15 observational studies, resulting in 1743 patients total: 910 treated operatively and 833 nonoperatively. The average age was 68.3 years, and 75% were female. There was no difference in functional outcome between operative and nonoperative treatment with a mean difference of -0.87 (CI, -5.13 – 3.38; P=0.69; I^2=69%). Major reinterventions occurred more often in the operative group. Pooled effects of RCTs were similar to pooled effects of observational studies for all outcome measures.

Conclusion

We recommend nonoperative treatment for the average elderly (aged >65 years) patient with a displaced proximal humeral fracture. Pooled effects of observational studies were similar to those of RCTs, and including observational studies led to more generalizable conclusions.

Introduction

The proximal humeral fracture is the third most common fracture seen in elderly persons, with an incidence of 82 per 100,000 person-years, with an annual increase in the rate by 13.7% over the past 33 years[1–3] The typical patient is a female aged 65 or older.[4] Nearly 75% of patients are treated nonoperatively, and one out of five will undergo surgery depending on fracture type and displacement.[5]

Depending on related factors such as patient age, activity, and fracture pattern, operative treatment options include minimally invasive reduction and intramedullary fixation, open reduction and internal plate fixation, or arthroplasty of the glenohumeral joint. Nonoperative treatment usually starts with immobilization followed by passive and active rehabilitation.[5] Despite the fact that the available literature is inconclusive regarding the superiority of either treatment option, it is common practice to attempt joint-saving operative procedures in younger patients.[5,6] In addition, there is no consensus on whether surgery is beneficial for the older patient with a displaced proximal humeral fracture.

Increasing scientific evidence has demonstrated that meta-analyses of both high-quality observational studies and randomized controlled trials (RCTs) can be similar in value to meta-analyses of RCTs alone in the field of orthopedic trauma surgery.[7–10] Observational studies may give better insight into infrequent outcome measures, rare complications, and small effects of operative treatment while also increasing the generalizability of the results owing to an increase in patient numbers available for analysis or meta-analysis.

The aims of this systematic review and meta-analysis were (1) to compare operative versus nonoperative treatment of displaced proximal humeral fractures and (2) to compare effect estimates obtained from RCTs and observational studies. We hypothesized that (1) operative treatment of proximal humeral fractures does not improve functional outcomes as compared with nonoperative treatment and (2) including observational studies in this meta-analysis will lead to more robust conclusions without decreasing the quality of the results.

Methods

This systematic review and meta-analysis followed guidelines published by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and MOOSE (Meta-Analysis of

Observational Studies in Epidemiology).[11,12] These checklists aim to improve the reporting of systematic reviews and meta-analyses for RCTs and observational studies, respectively.

Search strategy and eligibility criteria

Two reviewers (R.B.B. and Y.O.) independently searched the MEDLINE, Embase, CENTRAL, and CINAHL databases on September 5, 2017, for studies comparing operative and nonoperative treatment of proximal humeral fractures. The search syntax is provided in supplementary Table S1. Both RCTs and observational studies were included. After screening of the titles and abstracts of identified records, studies were independently assessed based on full text. The eligibility criteria were proximal humeral fracture; operative versus nonoperative treatment; and reporting of functional outcomes, as well as complications. The exclusion criteria were language other than English, Dutch, or German; no availability of full text; inclusion of patients younger than 18 years; letters, meeting proceedings, and case reports; and external osteosynthesis as operative treatment. Disagreement over eligibility was resolved by discussion with a third reviewer (R.M.H.). The references of the included studies were screened for eligibility, and citation tracking was performed by using Web of Science to identify articles not found in the original search. Authors were approached via ResearchGate when no full-text version of the article was available.

Data extraction

Data extraction was done independently by two reviewers (R.B.B. and Y.O.) with a data extraction file. The following data were extracted: first author, journal, year of publication, study period, study design, country or countries in which the study was performed, fracture displacement, fracture classification system (Neer classification), follow-up, treatment groups, operative treatment, nonoperative treatment, number of patients, loss to follow-up, implant removal, and outcome measures. Definitions of fracture characteristics, such as displacement, were applied according to the description in the original study. A major reintervention was defined as an additional, initially unplanned surgical procedure for implant failure, deep infection, symptomatic nonunion, subacromial impingement, or avascular necrosis. Planned implant removal was not considered a major reintervention. Fjalestad et al.[13,14] reported additional follow up of previously published data that were merged with the original article for this meta-analysis.

Quality assessment

Two reviewers (R.B.B. and H.F.) independently assessed the methodological quality of all included studies with the Methodological Index for Non-Randomized Studies (MINORS).[15] The MINORS is a validated instrument for methodological quality assessment and clear reporting of observational studies of surgical interventions.[15] Other quality assessment tools focus on a specific study design, while the MINORS is externally validated on RCTs by comparison with the CONSORT statement, making it a suitable instrument for meta-analyses of different study designs. The MINORS score ranges from 0 – 24; a higher score represents better methodological quality. Further details on the MINORS criteria and scoring system are provided in supplementary Table S2. Disagreements were resolved by involving a third reviewer (R.M.H.).

Outcome measures

The primary outcome measure was physical function as measured by the absolute Constant-Murley score[16] at least one year after initialization of either treatment. Normalized (sex- and age-adjusted) Constant-Murley scores were converted to absolute Constant-Murley scores using normal population-based values.[17] Secondary outcome measures were major reinterventions, nonunion, and avascular necrosis. If available, other functional outcome measures, such as the American Shoulder and Elbow Surgeons Shoulder Score[18] or the Neer score[19], were extracted as well.

Statistical analysis

Statistical analyses were performed using Review Manager (RevMan, Version 5.3.5. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). All continuous variables were converted to means and standard deviations (SD) when sufficient information was available using methods described in the Cochrane Handbook for Systematic Reviews of Interventions.[20]
All analyses were performed stratified by study design (i.e. RCTs and observational studies separately) as well as including both designs. Outcomes reported by two or more studies were pooled in a meta-analysis. Pooled effects of operative versus nonoperative treatment of dichotomous outcome measures were presented as risk ratios with confidence intervals (CI) using the Mantel-Haenszel method.[20] Pooled effects of continuous outcome measures were presented as mean differences with CI using the inverse variance weighting method.[20] Heterogeneity between studies was assessed by visual inspection of the forest plots and by estimating statistical measures for heterogeneity, i.e. the I^2 statistic and the Chi-square test. The main quantitative

assessment of heterogeneity was the I^2 statistic where the following interpretation was used: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100% considerable heterogeneity.[20] When heterogeneity was present a random-effects models was used instead of a fixed-effects model. Inspection of a funnel plot of the primary outcome measure against its standard error was done to detect potential publication bias.

Sensitivity analyses

Several sensitivity analyses were performed for study quality, year of publication, osteosynthesis by (locking) plate fixation and arthroplasty, and Neer classification. For the analysis of study quality only studies with an arbitrarily chosen MINORS score of 16 or higher were included, similar to previously published meta-analyses in orthopedic trauma surgery studying both study designs.[8,21] To assess the influence of the period in time in which the study was performed (and, consequently, development of different operative techniques), only studies published after 2005 were included in a separate analysis. Since the locking plate is the most commonly used type of osteosynthesis, another sensitivity analysis was conducted with studies where at least 80% of patients were treated with a locking plate. Furthermore, a sensitivity analysis was performed for all studies in which arthroplasty was the operative intervention. Finally, to explore the impact of fracture type on the functional outcome, a sensitivity analysis was performed including only Neer 3-part and 4-part fractures.

Different methods of meta-analysis may be differentially sensitive to studies with zero events on one or both study arms. Therefore, a sensitivity analysis to the choice of method of analysis was performed by means of the DerSimonian Laird method with correction and the inverse variance with and without correction for zero event data.[22]

Results

Figure 1 shows a flowchart of the literature search. In the end, 22 studies were included.[4,13,14,23–42] There were seven RCTs and 15 observational studies, of which nine were retrospective, four prospective, and two a combination of retrospective and prospective design.

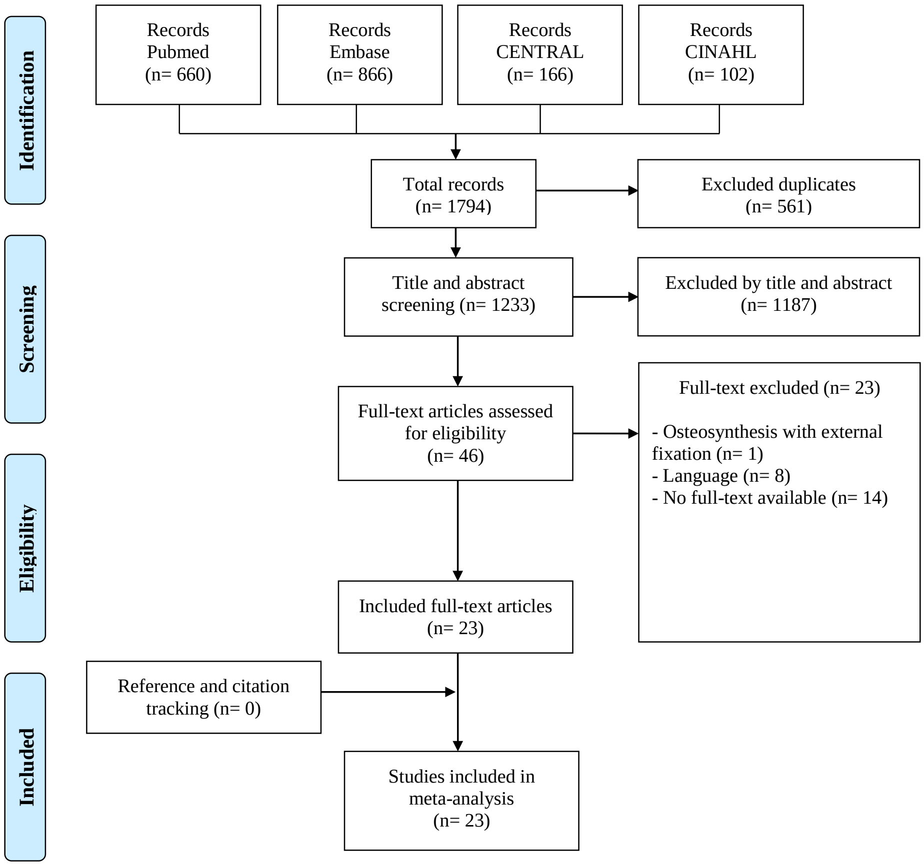

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram representing search and screen process of studies comparing operative versus nonoperative treatment of proximal humeral fractures. Central, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

Quality assessment

The MINORS score for all included studies ranged from 12 to 22 with a median of 17.5 (IQR 14-21). The MINORS score ranged from 16 to 22 with a median of 21 (IQR 21-22) for RCTs and from 12 to 21 with a median of 16 (IQR 14-18) for observational studies. Study-specific MINORS scores are provided in supplementary Table S3. The MINORS criteria for unbiased assessment of study endpoints and prospective calculation of study size were rarely met.

Baseline characteristics of study participants

Details of the included studies and patients are provided in Table 1. The 22 studies included a total of 1743 patients for meta-analysis: 910 treated operatively and 833 nonoperatively. The weighted average age was 68.3 years, and 75% were female. Follow-up ranged from 12 to 86 months.

Table 1. Baseline characteristics of studies included in a meta-analysis of proximal humerus fractures comparing operative with nonoperative treatment

Study	Study design	Country	Fracture classification	Treatment groups	Number patients	Follow-up (months)	Age (years, range/SD)	Female/ Male
Boons 2012	RCT	Netherlands	Neer 4	Operative: arthroplasty	25	12	76,4 (5,6)	24/1
				Nonoperative: Sling	25	12	79,9 (7,7)	23/2
Fjalestad 2012-14*	RCT	Norway	AO type B2-C2	Operative: LP	25	24	72,2 (60-86)	20/5
				Nonoperative: Sling + closed reduction	25	24	73,1 (60-88)	24/1
Olerud 2011a	RCT	Sweden	Neer 4	Operative: hemiarthroplasty	27	24	75,8 (58-90)	23/4
				Nonoperative: Sling	28	24	77,5 (60-92)	24/4
Olerud 2011b	RCT	Sweden	Neer 3	Operative: LP	30	24	72,9 (56-92)	24/6
				Nonoperative: Sling	29	24	74,9 (58-88)	24/5
Rangan 2015	RCT	England	Neer 2,3,4	Operative: PHN, LP, TB, arthroplasty, screw	125	24	66,6 (11,8)	97/28
				Nonoperative: Sling or hanging cast	125	24	65,43 (12,09)	95/30
Stableforth 1984	RCT	England	Neer 4	Operative: arthroplasty	16	All: 18-144	65,6 (52-88)	12/4
				Nonoperative: Sling	16		70,1 (60-85)	13/3
Zyto 1997	RCT	Sweden	Neer 3,4	Operative: Tension band	20	50	73 (7,5)	18/2
				Nonoperative: Sling	20	50	75 (6,7)	17/3
Court-Brown 2001	PC	Scotland	Neer 2	Operative: PHN + tension band fixation	18	12	73	NR
				Nonoperative: Sling	31	12	78	
Hauschild 2013	PC	Germany	AO type A2, A3	Operative: PHN, LP	133	12	62,9 (17,2)	97/36
				Nonoperative: Sling	31	12	65,6 (13,3)	22/9
Innocenti 2013	PC	Italy	Neer 2,3,4	Operative: K-wire	23	All: 86	73,92 (6,01)	All: 38/13
				Nonoperative: Sling	19		77,47 (6,95)	
Noureai 2014	PC	Iran	Neer 2,3,4	Operative: LP, Tension band, K-wire	57	12	All: 52,9 (15,0)	All: 70/44
				Nonoperative: Sling	57	12		

*Fjalestad 2012 and 2014 were analyzed as one study as both described the same patient cohort. RC retrospective cohort study; RCT randomized controlled trial; PC prospective cohort study; NR not reported; TB tension band; PHN proximal humerus nail; LP locking plate

Table 1. Continued

Study	Study design	Country	Fracture classification	Treatment groups	Number patients	Follow-up (months)	Age (years, range, SD)	Female/ Male
Fjalestad 2005	RC+PC	Norway	AO type A,B,C	Operative: K-wire(n=4), LP(n=5), Screws(n=4), Screws + cerclage(n=2)	15	12	All: 70 (25-95)	All: 50/20
				Nonoperative: Sling	55	12		
Ilchman 1998	RC+PC	Sweden / Swiss	Neer 3,4	Operative: Tension band	18	63	61 (23-80)	13/5
				Nonoperative: Sling (n=10), Closed reduction(n=4), Open reduction(n=2)	16	70	70 (23-91)	13/3
Blonna 2009	RC	Italy	AO type A2,2	Operative: K-wire	42	32	73 (7,83)	20/12
				Nonoperative: Sling	37	35	75,1 (8,0)	26/9
vd Broek 2007	RC	Netherlands	Neer 4,5,6	Operative: PHN	27	16	64,6 (27-87)	NR
				Nonoperative: Sling	17	69	69,4 (35-84)	
Hageman 2016	RC	Netherlands	Neer 2,3,4	Operative: PHN (n=3); LP (n=23), K-wire (n=2), Screws (n=5)	33	37	59,0 (12,5)	22/11
				Nonoperative: Sling	33	70	60,1 (15,3)	24/9
Kollig 2003	RC	Germany	Neer 4,5,6	Operative: LP(n=2), Screw + cerclage(n=7), K-wire(n=4)	13	82	52,5 (14,7)	NR
				Nonoperative: Sling	9	76	52,7 (11,5)	
Lange 2016	RC	Germany	Neer 2,3,4	Operative: PHN	41	All: 55	69,1 (37-88)	35/6
				Nonoperative: hanging cast or dessault dressing	41		68,9 (42-93)	35/6
Okike 2015	RC	US	Neer 2,3,4	Operative: LP	61	All:40	All: 76,9	109/37
				Nonoperative: Sling	146			46/15
Roberson 2017	RC	US	Neer 3,4	Operative: reversed arthroplasty	20	53	All: 71	19/1
				Nonoperative: Sling	19	29	(52-88)	15/4
Sanders 2011	RC	Australia	Neer 2,3,4	Operative: LP	17	37	58 (14)	9/8
				Nonoperative: Sling	18	42	64 (15)	12/6
Tamimi 2015	RC	Canada	Neer 2,3,4	Operative: PHN(n=19), LP(n=44), K-wire(n=25)	88	26	All: 65,3 (15,2)	All: 57/31
				Nonoperative: Sling	25	28		

*Fjalestad 2012 and 2014 were analyzed as one study as both described the same patient cohort. RC retrospective cohort study; RCT randomized controlled trial; PC prospective cohort study; NR not reported; TB tension band; PHN proximal humerus nail; LP locking plate

All studies but one included displaced proximal humeral fractures in their study. The majority of the included studies excluded patients with pathological fractures, open fractures, fractures of the skeletally immature, and other sustained injury to the affected side. Most studies (n=18, 82%) used the Neer classification and included patients with a Neer 2,3 or 4-part proximal humeral fracture. In seven studies at least 80% of patients were treated with a locking plate.[13,14,25,27,33,35,41,42] Four studies investigated arthroplasty; three hemiarthroplasty and one reverse shoulder arthroplasty[29,30,32,34]; three studies assessed proximal humeral nails[4,28,39], and eight studied fixation by means of Kirschner wires, screws, tension band, or a combination of techniques.

Functional outcome

Fourteen studies (64%, n=817) reported the Constant-Murley score after at least one year of follow-up (supplementary Table S4).[13,14,37,39,40,42,23,25–28,32–34] In patients with a proximal humeral fracture, the functional outcome as measured by the Constant-Murley score showed no difference in operative versus nonoperative treatment with a mean difference of -0.87 (CI, -5.13 – 3.38; P=0.69; I^2=69%) (Figure 2). Pooled effects of RCTs were similar to those of observational studies for all outcome measures (Table 2). Figure 3 shows a funnel plot of the mean difference and standard error of the included studies using the Constant-Murley score; there was no important asymmetry observed.

For studies that did not use the Constant-Murley score, we performed additional analysis with the standardized mean difference of different functional outcome measures which yielded the same result as the primary analysis (SMD -0.06; CI, -0.25 – 0.12; P=0.52; I^2=53%) (supplementary Figure S1). Seven studies (n=327) reported functional outcome of patients treated with a Neer 3-part or 4-part fracture.[14,28,29,31–34] Forty-three percent of patients with Neer 4-part fractures were initially treated with arthroplasty (Table 1). A subgroup analysis of these studies showed no difference in standardized mean difference of functional outcome measures between operative and nonoperative treatment with a mean difference of 0.02 (CI, -0.20 – 0.24; P=0.86; I^2=0%) (supplementary Figure S2).

Major reinterventions

Fifteen studies (68%, n=938) reported on major reinterventions (supplementary Table S4).[13,14,34–36,39–41,23,25,27–30,32,33] Two studies had no major reintervention in either treatment arm

Study or Subgroup	Operative Mean	Operative SD	Operative Total	Nonoperative Mean	Nonoperative SD	Nonoperative Total	Weight	Mean Difference IV, Random, 95% CI
1.1.1 Observational studies								
Blonna 2009	72	25.4	32	58.2	25.4	25	5.5%	13.80 [0.51, 27.09]
Broek vd 2007	67.1	11.8	23	81.4	9.75	16	9.1%	-14.30 [-21.09, -7.51]
Hageman 2016	63.3	35.2	33	87.5	35.2	33	4.1%	-24.20 [-41.18, -7.22]
Hauschild 2013	74.2	13	103	74.3	9.9	26	10.5%	-0.10 [-4.66, 4.46]
Innocenti 2013	56.5	5.24	23	52	9.03	19	10.5%	4.50 [-0.09, 9.09]
Kollig 2003	72.1	21.1	13	82	15.6	9	4.7%	-9.90 [-25.24, 5.44]
Lange 2016	72.3	18	41	72	19	41	8.4%	0.30 [-7.71, 8.31]
Okike 2015	58	16.9	25	59.7	17.5	22	7.3%	-1.70 [-11.57, 8.17]
Tamimi 2015	65.1	19	88	57.2	12.7	25	9.4%	7.90 [1.53, 14.27]
Subtotal (95% CI)			**381**			**216**	**69.4%**	**-1.50 [-7.33, 4.33]**
Heterogeneity: Tau² = 56.90; Chi² = 40.29, df = 8 (P < 0.00001); I² = 80%								
Test for overall effect: Z = 0.50 (P = 0.61)								
1.1.2 RCT								
Boons 2012	64	15.8	23	60	17.6	24	7.4%	4.00 [-5.55, 13.55]
Fjalestad 2012-14	51.8	30.6	23	53.2	26.8	19	4.0%	-1.40 [-18.77, 15.97]
Olerud HA 2011	48.3	16.4	24	49.6	20.5	25	7.0%	-1.30 [-11.67, 9.07]
Olerud LP 2011	61	19.2	27	58.4	23.1	26	6.4%	2.60 [-8.86, 14.06]
Zyto 1997	60	19	14	65	15	15	5.9%	-5.00 [-17.52, 7.52]
Subtotal (95% CI)			**111**			**109**	**30.6%**	**0.40 [-4.76, 5.56]**
Heterogeneity: Tau² = 0.00; Chi² = 1.55, df = 4 (P = 0.82); I² = 0%								
Test for overall effect: Z = 0.15 (P = 0.88)								
Total (95% CI)			**492**			**325**	**100.0%**	**-0.87 [-5.13, 3.38]**
Heterogeneity: Tau² = 39.60; Chi² = 41.84, df = 13 (P < 0.0001); I² = 69%								
Test for overall effect: Z = 0.40 (P = 0.69)								
Test for subgroup differences: Chi² = 0.23, df = 1 (P = 0.63), I² = 0%								

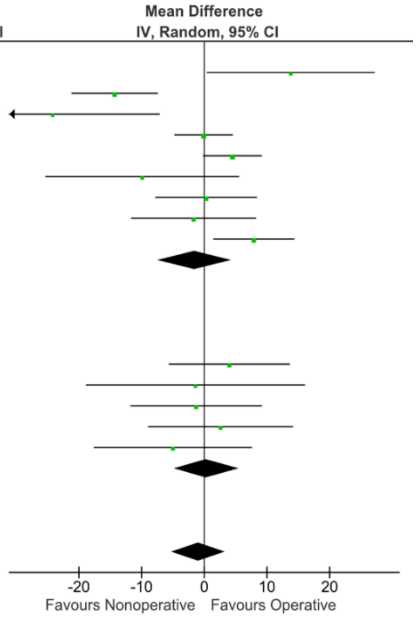

Figure 2. Functional outcome as measured with Constant-Murley score in systematic review of proximal humeral fractures comparing operative versus nonoperative treatment. *SD*, standard deviation; *IV*, inverse variance; *CI*, confidence interval; *RCT*, randomized controlled trial.

at follow-up. Major reinterventions occurred more often in the operative group than the nonoperative group with a risk ratio (RR) of 2.72 (CI, 1.71 – 4.34; P< 0.0001; I^2=0%) (supplementary Figure S3). Using different methods of incorporating studies in the meta-analysis with zero event data in one or both arms yielded similar results (supplementary Table S5). Implant removal was reported in 10 studies (45%). The mean percentage of implant removal across studies was 21% (range 0–100%). When stratified by study design, observational studies showed a greater risk for major reinterventions in the operative treatment group compared with the nonoperative group (RR 5.43; CI 2.51–11.74; P< 0.0001; I^2=0%) (Table 2). Five studies specified their reinterventions for nonoperatively treated patients: four patients received arthroplasty for displacement and malunion, two patients received ORIF for displacement, and two patients received acromioplasty for impingement complaints.

Nonunion

A total of thirteen studies (59%) reported on nonunion (supplementary Table S4). Operative treatment of proximal humeral fractures resulted in fewer nonunion than nonoperative treatment with a RR of 0.45 (CI, 0.23–0.89; P=0.02; I^2=0%) (supplementary Figure S4). When stratified by study design, both subgroups showed a similar, non-significant, pooled effect (Table 2).

Table 2. Subgroup & sensitivity analyses of studies included in a meta-analysis of proximal humerus fractures comparing operative to nonoperative treatment

Analysis description	CS				MR				Nonunion				AVN			
	n	MD	(95% CI)	P-value	n	RR	(95% CI)	P-value	n	RR	(95% CI)	P value	n	RR	(95% CI)	P-value
All studies	14	-0.87	(-5.13; 3.38)	0.69	15	2.72	(1.71; 4.34)	< 0.0001	13	0.45	(0.23; 0.89)	0.02	13	1.24	(0.87; 1.77)	0.24
Subgroup analysis																
RCTs	5	0.40	(-4.76; 5.56)	0.88	6	1.45	(0.78; 2.70)	0.25	6	0.48	(0.19; 1.20)	0.12	6	0.88	(0.55; 1.41)	0.59
Observational studies	9	-1.50	(-7.33; 4.33)	0.61	7	5.43	(2.51; 11.74)	< 0.0001	7	0.41	(0.15; 1.16)	0.09	7	1.93	(1.11; 3.37)	0.02
Sensitivity analysis																
High-quality studies	11	0.55	(-2.93; 4.03)	0.76	11	2.52	(1.55; 4.11)	0.0002	11	0.44	(0.21; 0.93)	0.03	10	1.14	(0.74; 1.74)	0.55
Studies after 2005	12	-0.14	(-4.65; 4.38)	0.95	14	2.58	(1.59; 4.20)	0.0001	12	0.41	(0.18; 0.89)	0.03	10	1.10	(0.72; 1.69)	0.65
Locking plate	5	-0.15	(-0.43; 0.13)	0.30	7	1.81	(1.04; 3.16)	0.04	6	0.37	(0.12; 1.17)	0.09	6	1.35	(0.86; 2.11)	0.19
Arthroplasty	2	1.50	(-5.24; 8.23)	0.66	4	2.66	(0.72; 9.77)	0.14	3	0.52	(0.13; 1.99)	0.34	2	0.17	(0.02; 1.37)	0.10

N Number; CS Constant score MR Major reintervention; AVN avascular necrosis; RCT randomized controlled trial; RR risk ratio; MD mean difference; CI confidence interval; Sensitivity analysis of locking plate includes studies comparing locking plate to nonoperative treatment; Sensitivity analysis of arthroplasty includes studies comparing hemiarthroplasty and reversed arthroplasty to nonoperative treatment

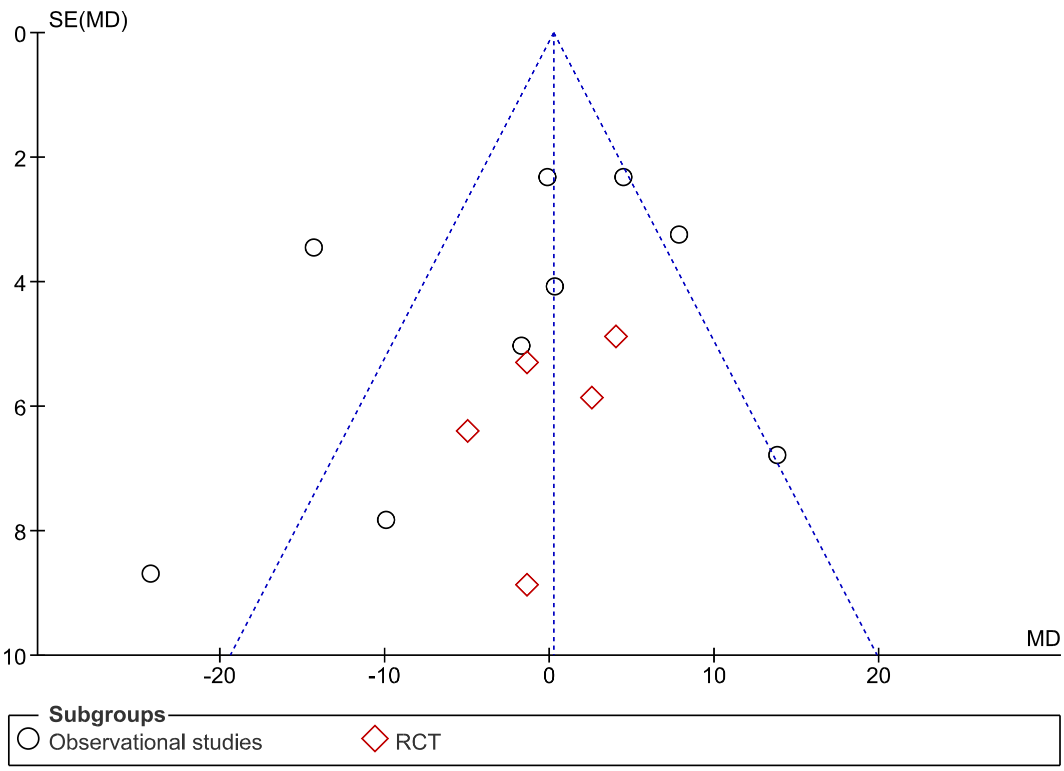

Figure 3. Funnel plot of studies included in meta-analysis reporting Constant-Murley scores after operative or nonoperative treatment of proximal humeral fractures. *SE*, standard error; *MD*, mean difference; *RCT*, randomized controlled trial.

Avascular necrosis

A total of thirteen studies (59%) reported on avascular necrosis (supplementary Table S4). There was no difference in the rate of avascular necrosis between operative and nonoperative treatment for proximal humeral fractures with a RR of 1.24 (CI, 0.87–1.77; P=0.24; I^2=24%) (supplementary Figure S5). When stratified by study design, observational studies showed a higher risk of avascular necrosis for the operative group compared with the nonoperative group (RR 1.93; CI 1.11–3.37; P=0.02; I^2=9%) (Table 2).

Sensitivity analysis

Sensitivity analysis did not significantly alter the primary and secondary outcome measures (Table 2).

Discussion

In this systematic review and meta-analysis of patients with displaced proximal humeral fractures, there was no difference in physical function as measured with the Constant-Murley score after operative or nonoperative treatment. Subgroup analysis for Neer 3-part or 4-part fractures neither showed differences in functional outcome. Results of the primary and secondary outcome measures were similar from the pooled effects of RCTs and observational studies. There was a higher risk for major reinterventions and a lower risk of nonunion after operative treatment compared with nonoperative treatment. This the largest meta-analysis in the current literature by including both RCTs and observational studies.

Compared with nonoperative treatment, there is no improved in functional outcome after operative treatment for displaced proximal humeral fractures, which confirms findings of previous meta-analyses.[6,43] A recent systematic review of displaced proximal humeral fractures is based on only 7 RCTs with just over 500 patients.[6] With a total of 250 patients, the PROFHER trial represents the most substantial evidence currently available.[35] The patient demographic characteristics of the PROFHER trial are comparable with those of the included studies in this meta-analysis (Table 1). However, only 4.4% of patients in the PROFHER trial suffered a Neer 4-part fracture compared with 21% of patients in this meta-analysis. Therefore, compared with previous, smaller magnitude meta-analyses, this review contributes substantially to the current evidence and enables recommendations for a broader patient population. Furthermore, this is the first meta-analysis in which subgroup analysis for Neer 3-part and 4-part fractures was possible, and the results showed no differences in operative versus nonoperative treatment.

This review showed similar pooled effects of observational studies and RCTs for the primary and secondary outcome measures. This finding is similar to previous meta-analyses in orthopedic trauma surgery including both study designs.[7–10,44] As such, this review speaks to the growing potential of observational studies in orthopedic trauma surgery and contributes to the expanding discussion about the value of different study designs.[45]

In this review, the major reintervention rate included every additional surgery except for implant removal because of patient preference, implant-related irritation, or a stiff shoulder. Therefore, the major reintervention rate in this review is a surrogate marker for severe complications (e.g. implant failure, deep infection, nonunion, impingement, or avascular necrosis) after operative and nonoperative treatment of displaced proximal humeral fractures. This is the first review to show significantly more severe complications requiring surgical re-intervention after operative treatment of displaced proximal humeral fractures. These procedures add up to the additional surgery for implant removal for 21% of the patients for a less serious indication.

Another new finding is the higher risk of nonunion for nonoperatively treated patients. RCTs and observational studies alone were not able to detect a significant difference in this outcome. This demonstrates the added value of increasing study power by including observational studies in order to detect rare outcomes. It is important to note that this difference is supported by the sensitivity analysis including only high-quality RCTs and observational studies (Table 2).

This review found no difference in the rate of avascular necrosis between the nonoperative and operative management. However, it should be noted that three of the 15 studies reporting on avascular necrosis had a follow-up of 12 months while avascular necrosis can be detected up to two years of follow-up. For this outcome measure, the pooled effect of observational studies was significantly different than the pooled effect of RCTs. However, in the sensitivity analysis with high quality studies, this contrasting result did not yield, and pooled effects of both study designs were similar again. This demonstrates the importance of evaluating the quality of the included studies (Table 2). Therefore, including a study in a meta-analysis should be based on the quality of the study regardless of the study design.[44] Generally, RCTs will be of higher quality and thus included for analysis, however, a high-quality observational study should be chosen over a low-quality RCT.

The results of this systematic review and meta-analysis should be interpreted in the light of several limitations. First, the results of the meta-analysis may be influenced by missed studies in the database search or by publication bias. However, an extensive search was performed using multiple databases, and the citations and references of included studies were also screened. Furthermore, a funnel plot of the primary outcome measure did not suggest possible bias due to selective publication. Second, results of observational studies are more heterogeneous than those of RCTs in the meta-analysis of the Constant-Murley score. Still, it should be noted despite heterogeneity in mean differences of the observational studies, the observed effects all are within a range of the Constant-Murley score which is clinically nonimportant.[46] Third, in the analysis of functional outcome, we did not distinguish between 12 or more than 12 months of follow-up since prior studies have shown the greatest increase in functional outcome takes place in the first six months and no significant improvement is to be expected after 12 months.[4,14,32,33,35] This is further supported by an additional sensitivity analysis that showed no differences in functional outcome at 12 months and at 24 or more months. Fourth, the Neer classification for proximal humeral fractures is the most frequently used classification system in the literature even though it has been considered to have important limitations. However, no other system for evaluating these fractures is consistently more reliable than the Neer classification.[47] Fifth, The majority of the included studies were European, and only three studies described patients from Northern America, let alone other continents. However, subgroup analyses revealed no differences for the primary and secondary outcome measures between these continents (data not shown). Finally, it should be noted that the majority of studies in this review excluded patients with pathological fractures, patients with open fractures, fractures of skeletal immature patients, and patients with other sustained injuries to the affected arm. As a result, recommendations from this review are not applicable to these patients.

Although we acknowledge the vast amount of existing systematic reviews on this topic[6,43,48,49], we believe that the several unique qualities of this meta-analysis contribute to the existing knowledge. Strengths of this study include the consistent results of the different sensitivity analyses for time of publication, type of osteosynthesis, and arthroplasty. Furthermore, by including observational studies in addition to the highly selective patient population of RCTs, the analyzed patients may be more representative of patients encountered in daily clinical practice and improve the generalizability of our results. We also demonstrated that the findings were consistent across study designs with respect to different outcome measures. Although no subgroup analysis of elderly patients (aged > 65 years) could be performed, the mean age of all patients in this review was 68 years, with a relatively small standard deviation for the majority of the included studies; therefore, we feel confident that recommendations from this review apply to the average elderly patient. Finally, this is the largest meta-analysis in the literature with the highest number of patients available for analysis of proximal humeral fractures.

Conclusion

We recommend nonoperative treatment for the average elderly patient (aged > 65 years) with a displaced proximal humeral fracture. Pooled effects of observational studies were similar to those of RCTs, and the inclusion of observational studies improved the generalizability of findings.

References:

1. Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures. *Clin Orthop Relat Res*. 2006;442:87-92.
2. Roux A, Decroocq L, El Batti S, et al. Epidemiology of proximal humerus fractures managed in a trauma center. *Orthop Traumatol Surg Res*. 2012;98(6):715-719.
3. Launonen AP, Lepola V, Saranko A, Flinkkilä T, Laitinen M, Mattila VM. Epidemiology of proximal humerus fractures. *Archives of Osteoporosis*. 2015;10(1):2.
4. Court-Brown CM, Garg A, McQueen MM. The translated two-part fracture of the proximal humerus. Epidemiology and outcome in the older patient. *Journal of Bone and Joint Surgery - Series B*. 2001;83(6):799-804.
5. Jawa A, Burnikel D. Treatment of Proximal Humeral Fractures. *JBJS Reviews*. 2016;4(June):1-9.
6. Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. 2015;(11):CD000434.
7. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *Am J Sports Med*. 2016.
8. Houwert RM, Smeeing DP, Ahmed Ali U, Hietbrink F, Kruyt MC, van der Meijden OA. Plate fixation or intramedullary fixation for midshaft clavicle fractures: a systematic review and meta-analysis of randomized controlled trials and observational studies. *J Shoulder Elbow Surg*. 2016;25(7):1195-1203.
9. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *J Clin Epidemiol*. 2010;63(3):238-245.
10. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev*. 2014;(4):MR000034.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100.
12. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of Observational Studies in Epidemiology<SUBTITLE>A Proposal for Reporting</SUBTITLE> *JAMA*. 2000;283(15):2008.
13. Fjalestad T, Hole MO, Hovden IA, Blucher J, Stromsoe 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 MO. Displaced proximal humeral fractures: operative versus non-operative treatment--a 2-year extension of a randomized controlled trial. *Eur J Orthop Surg Traumatol*. 2014;24(7):1067-1073.
15. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg*. 2003;73(9):712-716.
16. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res*. 1987;(214):160-164.
17. Constant CR, Gerber C, Emery RJ, Sojbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg*. 2008;17(2):355-361.
18. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: Reliability, validity, and responsiveness. *Journal of Shoulder and Elbow Surgery*. 2002;11(6):587-594.
19. Neer CS. Displaced Proximal Humeral Fractures. *J Bone Joint Surg Am*. 1970;52(6):1077-1089. 20. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley; 2011.
21. Smeeing DP, van der Ven DJ, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older. *Am J Sports Med*. 2016:363546516673615.
22. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*. 2007;26(1):53-77.
23. Innocenti M, Carulli C, Civinini R, Matassi F, Tani M, Muncibi F. Displaced fragility fractures of proximal humerus in elderly patients affected by severe comorbidities: percutaneous fixation and conservative treatment. *Aging Clin Exp Res*. 2013;25(4):447-452.
24. Nouraei MH, Majd DA, Zamani F. Comparing the treatment results of proximal humerus fracture based on surgical or nonsurgical methods. *Adv Biomed Res*. 2014;3:253.
25. Okike K, Lee OC, Makanji H, Morgan JH, Harris MB, Vrahas MS. Comparison of locked plate fixation and nonoperative management for displaced proximal humerus fractures in elderly patients. *Am J Orthop (Belle Mead NJ)*. 2015;44(4):E106-12.
26. Tamimi I, Montesa G, Collado F, et al. Displaced proximal humeral fractures: when is surgery necessary? *Injury*. 2015;46(10):1921-1929.
27. Hageman MGJS, Meijer D, Stufkens SA, Ring D, Doornberg JN, Steller EP. Proximal humeral fractures: Nonoperative versus operative treatment. *Arch Trauma Res*. 2017;6(1).

28. Lange M, Brandt D, Mittlmeier T, Gradl G. Proximal humeral fractures: non-operative treatment versus intramedullary nailing in 2-, 3- and 4-part fractures. *Injury*. 2016;47 Suppl 7:S14-s19.
29. Roberson TA, Granade CM, Hunt Q, et al. Nonoperative management versus reverse shoulder arthroplasty for treatment of 3- and 4-part proximal humeral fractures in older adults. *Journal of Shoulder and Elbow Surgery*. 2017;26(6):1017-1022.
30. Stableforth PG. Four-part fractures of the neck of the humerus. *J Bone Joint Surg Br*. 1984;66(1):104-108.
31. Zyto K, Törnkvist CH, Ahrengart L, et al. Treatment of displaced proximal humeral fractures in elderly patients. *J Bone Joint Surg Br*. 1997;79(3):412-417.
32. 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.
33. 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.
34. 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 Relat Res*. 2012;470(12):3483-3491.
35. Rangan A, Handoll H, Brealey S, et al. Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA*. 2015;313(10):1037.
36. Ilchmann T, Ochsner PE, Wingstrand H, Jonsson K. Non-operative treatment versus tension-band osteosynthesis in three- and four-part proximal humeral fractures. A retrospective study of 34 fractures from two different trauma centers. *Int Orthop*. 1998;22(5):316-320.
37. Kollig E, Kutscha-Lissberg F, Roetman B, Dielenschneider D, Muhr G. [Complex fractures of the humeral head: which long-term results can be expected?]. *Zentralbl Chir*. 2003;128(2):111-118.
38. Fjalestad T, Strømsøe K, Blücher J, et al. Fractures in the proximal humerus: functional outcome and evaluation of 70 patients treated in hospital. *Archives of orthopaedic and trauma surgery*. 2005;125(5):310-316.
39. Van-den-Broek CM, van den Besselaar M, Coenen JMF, Vegt PA. Displaced proximal humeral fractures: Intramedullary nailing versus conservative treatment. *Arch Orthop Trauma Surg*. 2007;127(6):459-463.
40. Blonna D, Rossi R, Fantino G, Maiello A, Assom M, Castoldi F. The impacted varus (A2.2) proximal humeral fracture in elderly patients: is minimal fixation justified? A case control study. *J Shoulder Elbow Surg*. 2009;18(4):545-552.
41. 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):1118-1124.
42. Hauschild O, Konrad G, Audige L, et al. Operative versus non-operative treatment for two-part surgical neck fractures of the proximal humerus. *Arch Orthop Trauma Surg*. 2013;133(10):1385-1393.
43. Xie L, Ding F, Zhao Z, Chen Y, Xing D. Operative versus non-operative treatment in complex proximal humeral fractures: a meta-analysis of randomized controlled trials. *Springerplus*. 2015;4:728.
44. Smeeing DPJ, Houwert RM, Kruyt MC, van der Meijden OAJ, Hietbrink F. Clinical research on postoperative trauma care: has the position of observational studies changed? *European journal of trauma and emergency surgery : official publication of the European Trauma Society*. 2017;43(1):43-51.
45. Frieden TR. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. Drazen JM, Harrington DP, McMurray JJV, Ware JH, Woodcock J, eds. *New England Journal of Medicine*. 2017;377(5):465-475.
46. van de Water AT, Shields N, Davidson M, Evans M, Taylor NF. Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. *Disabil Rehabil*. 2014;36(13):1072-1079.
47. Carofino BC, Leopold SS. Classifications in brief: The neer classification for proximal humerus fractures. *Clinical Orthopaedics and Related Research*. 2013;471(1):39-43.
48. Fu T, Xia C, Li Z, Wu H. Surgical versus conservative treatment for displaced proximal humeral fractures in elderly patients: a meta-analysis. *International journal of clinical and experimental medicine*. 2014;7(12):4607-4615.
49. Rabi S, Evaniew N, Sprague SA, Bhandari M, Slobogean GP. Operative *vs* non-operative management of displaced proximal humeral fractures in the elderly: A systematic review and meta-analysis of randomized controlled trials. *World Journal of Orthopedics*. 2015;6(10):838.

Supplementary materials to Chapter 3

Table S1. Search syntax performed last on March 30, 2017

Database	Syntax
PubMed/MEDLINE (n= 660)	(Humeral Fractures[MeSH Terms] OR Shoulder Fractures[MeSH Terms] OR ((humeral[Title/Abstract] OR humerus[Title/Abstract] OR humeri[Title/Abstract] OR humor[Title/Abstract] OR (upper[Title/Abstract] AND arm[Title/Abstract] AND bone[Title/Abstract]) OR (upperarm[Title/Abstract] AND bone[Title/Abstract])) AND fractur*[Title/Abstract])) AND (proximal[Title/Abstract] OR sub-capital[Title/Abstract] OR subcapital[Title/Abstract] OR neck[Title/Abstract]) AND (surgery[subheading] OR Fracture Healing[MeSH Terms] OR Fracture Fixation[MeSH Terms] OR Surgical Procedures, Operative[MeSH Terms] OR orthopedics[MeSH Terms] OR orthopedics[Title/Abstract] OR orthopaedics[Title/Abstract] OR orthopedic[Title/Abstract] OR orthopaedic[Title/Abstract] OR surgery[Title/Abstract] OR surgical[Title/Abstract] OR operative[Title/Abstract] OR operate[Title/Abstract] OR operating[Title/Abstract] OR operated[Title/Abstract] OR operation[Title/Abstract]) AND (conservative[Title/Abstract] OR conventional[Title/Abstract] OR non-operative[Title/Abstract] OR non-surgical[Title/Abstract] OR non surgical[Title/Abstract] OR nonoperative[Title/Abstract] OR Physical Therapy Modalities[MeSH Terms] OR sling[Title/Abstract] OR collar[Title/Abstract] OR cuff[Title/Abstract] OR bandages[Title/Abstract] OR bandage[Title/Abstract])
Embase (n= 866)	('humerus'/exp OR humerus:ti,ab OR humeri:ti,ab OR humer:ti,ab OR humor:ti,ab OR 'corpus humeri':ti,ab OR 'upper arm bone':ti,ab OR 'upperarm bone':ti,ab OR humeral:ti,ab) AND ('fracture'/exp OR fracture:ti,ab OR fractured:ti,ab OR fractures:ti,ab) AND (proximal:ti,ab OR 'sub capital':ti,ab OR 'subcapital':ti,ab OR neck:ti,ab) AND ('surgery'/exp OR surgery:ti,ab OR surgical:ti,ab OR operative:ti,ab OR operation:ti,ab OR 'Fracture Healing':ti,ab OR 'Fracture fixation':ti,ab OR 'Surgical Procedures':ti,ab OR orthopedics:ti,ab OR orthopedic:ti,ab OR orthopaedics:ti,ab OR orthopaedic:ti,ab OR operate:ti,ab OR operating:ti,ab OR operated:ti,ab) AND ('conservative treatment'/exp OR 'conservative treatment':ti,ab OR conservative:ti,ab OR conventional:ti,ab OR 'non-operative':ti,ab OR nonoperative:ti,ab OR non-surgical:ti,ab OR 'non surgical':ti,ab OR sling:ti,ab OR collar:ti,ab OR cuff:ti,ab OR bandages:ti,ab OR bandage:ti,ab)
CENTRAL (n= 166)	humerus AND fracture AND (proximal OR neck OR sub capital OR subcapital)
CINAHL (n= 102)	(humerus OR humeri OR humer OR humor OR corpus humeri OR upper arm bone OR upperarm bone OR humeral) AND (fracture OR fractured OR fractures) AND (proximal OR sub capital OR neck OR subcapital) AND (surgery OR surgical OR operative OR operation OR Fracture Healing OR Fracture fixation OR Surgical Procedures OR orthopedics OR orthopedic OR orthopaedics OR orthopaedic OR operate OR operating OR operated) AND (conservative treatment OR conservative OR conventional OR non-operative OR nonoperative OR non-surgical OR non surgical OR sling OR collar OR cuff OR bandages OR bandage)

Table S2. Quality assessment according to the MINORS criteria in a meta-analysis of proximal humeral fractures

Criteria	Reported and adequate (2)	Reported but inadequate (1)	Not reported (0)
Clearly stated aim	Aim including outcomes reported	Aim reported without outcomes	Not reported
Inclusion consecutive patients	Inclusion/exclusion criteria reported	Unclear description inclusion/exclusion criteria	Not reported
Prospective collection data	Prospective	Retrospective	Not reported
Appropriate endpoints	Appropriate endpoints to aim study	Endpoints not appropriate to aim study	Not reported
Unbiased assessment	Blinded evaluation of outcomes	Reason not blinding stated	Not reported
Appropriate follow-up	≥ 1 year	< 1 year	Not reported
Loss to follow-up < 5%	≤ 5%	> 5% and ≤ 20%	Not reported/>20%
Prospective calculation study size	Power-analysis performed	Explanation number without power-analysis	Not reported
Adequate control group	Operative versus nonoperative treatment	Not applicable	Not reported
Contemporary groups	Study/control group managed during same period	Study/control not managed during same period	Not reported
Baseline equivalence groups	Baseline characteristics described and comparable	Baseline characteristics not comparable	Not reported
Adequate statistical analyses	Statistical analysis described including type of analyses	Inadequate description statistical analysis	Not reported

Items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S3. Quality assessment of all included studies in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment

Criteria	Blonna 2009	Boons 2012	vd Broek 2007	Court-Brown 2001	Fjalestad 2005	Fjalestad 2012-14	Hageman 2016	Hauschild 2013	Ilchman 1998	Innocenti 2013	Kollig 2003	Lange 2016	Noureai 2014	Okike 2015	Olerud 2011a	Olerud 2011b	Rangan 2015	Roberson 2017	Sanders 2011	Stableforth 1984	Tamimi 2015	Zyto 1997
Clearly stated aim	2	2	2	1	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
Inclusion of consecutive patients	2	2	2	2	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
Prospective collection of data	0	2	0	2	1	2	0	2	0	2	0	0	2	0	2	2	2	0	0	2	0	2
Appropriate endpoints	2	2	2	2	1	2	2	2	1	2	2	2	1	2	2	2	2	2	2	1	2	2
Unbiased assessment endpoints	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	0	0	1
Appropriate follow-up	2	2	1	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2
Loss to follow-up < 5%	1	1	1	0	0	2	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	2
Prospective calculation study size	0	2	0	0	0	1	0	1	0	0	0	0	0	0	1	1	2	0	0	0	0	0
Adequate control group	2	2	2	2	2	2	2	2	1	2	1	2	2	2	2	2	2	2	2	2	1	2
Contemporary groups	2	2	1	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2
Baseline equivalence of groups	2	2	1	2	1	2	2	2	2	2	0	2	0	1	2	2	2	2	2	2	1	2
Adequate statistical analysis	2	2	0	2	2	2	2	2	1	2	0	2	1	2	2	2	2	2	2	0	2	2
Total MINORS score	**18**	**22**	**13**	**18**	**14**	**22**	**17**	**21**	**14**	**19**	**12**	**16**	**13**	**16**	**21**	**21**	**22**	**16**	**18**	**16**	**14**	**21**

Items are scored 0 (not reported/ not applicable), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S4. Outcome measures in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment

Study		Constant score (± SD)	Revision surgery	Non-union	AVN	DASH score (± SD)	Implant removal
Blonna 2009	Operative	72 (25.4)	0	0	NR	15.0 (3.0)	32
	Nonoperative	58.2 (25.4)	0	0		30.5 (5.1)	
Boons 2012	Operative	64 (15.8)	1	2	0	NR	0
	Nonoperative	60 (17.6)	1	3	2		
vd Broek 2007	Operative	67.1 (11.8)	3	0	0	NR	5
	Nonoperative	81.4 (9.8)	0	1	0		
Court-Brown 2001	Operative	NR	NR	1	NR	NR	NR
	Nonoperative			4			
Fjalestad 2005	Operative	NR	NR	1	3	NR	NR
	Nonoperative			5	2		
Fjalestad 2012-14*	Operative	51.8 (30.6)	1	1	12	NR	7
	Nonoperative	53.2 (26.8)	1	2	15		
Hageman 2016	Operative	63.3 (35.2)	5	NR	1	22 (13.9)	2
	Nonoperative	87.5 (35.2)	2		0	10.3 (13.9)	
Hauschild 2013	Operative	74.2 (13)	NR	1	1	NR	NR
	Nonoperative	74.3 (9.9)		0	0		
Ilchman 1998	Operative	NR	4	NR	9	NR	NR
	Nonoperative		1		7		
Innocenti 2013	Operative	56.5 (5.2)	0	NR	0	NR	23
	Nonoperative	52 (9.0)	0		0		
Kollig 2003	Operative	72.1 (21.1)	NR	NR	NR	NR	NR
	Nonoperative	82 (15.6)					
Lange 2016	Operative	72.3 (18)	13	NR	NR	NR	NR
	Nonoperative	72 (19)	0				

*In this analysis Fjalestad 2012 and 2014 were seen as one study as both studies describe the same patient cohort AVN avascular necrosis; NR not reported; SD standard deviation

Table S4. Continued

Study		Constant score (± SD)	Revision surgery	Non-union	AVN	DASH score (± SD)	Implant removal
Noureai 2014	Operative	NR	NR	NR	NR	NR	NR
	Nonoperative						
Okike 2015	Operative	58 (16.9)	8	0	10	26.5 (17.8)	NR
	Nonoperative	59.7 (17.5)	2	2	3	25.1 (18.2)	
Olerud 2011a	Operative	48.3 (16.4)	2	0	0	30.2 (18.3)	1
	Nonoperative	49.6 (20.5)	1	1	3	36.9 (21.3)	
Olerud 2011b	Operative	61 (19.2)	4	1	3	26.4 (25.2)	5
	Nonoperative	58.4 (23.1)	1	1	2	35 (26.8)	
Rangan 2015	Operative	NR	11	0	4	NR	NR
	Nonoperative		11	5	1		
Roberson 2017	Operative	NR	3	NR	NR	NR	0
	Nonoperative		0				
Sanders 2011	Operative	NR	3	0	8	NR	7
	Nonoperative		0	1	5		
Stableforth 1984	Operative	NR	1	NR	NR	NR	1
	Nonoperative		0				
Tamimi 2015	Operative	65.1 (19)	NR	NR	NR	33 (21.8)	NR
	Nonoperative	57.2 (12.7)				38.4 (19.2)	
Zyto 1997	Operative	60 (19)	NR	1	1	NR	1
	Nonoperative	65 (15)		0	0		

*In this analysis Fjalestad 2012 and 2014 were seen as one study as both studies describe the same patient cohort AVN avascular necrosis; NR not reported; SD standard deviation

Study or Subgroup	Operative Mean	Operative SD	Operative Total	Nonoperative Mean	Nonoperative SD	Nonoperative Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.5.1 Observational studies								
Blonna 2009	72	25.4	32	58.2	25.4	25	5.8%	0.54 [0.00, 1.07]
Broek vd 2007	67.1	11.8	23	81.4	9.75	16	4.3%	-1.27 [-1.98, -0.57]
Court-Brown 2001	71.1	11.5	18	73.6	11.5	31	5.3%	-0.21 [-0.80, 0.37]
Hageman 2016	63.3	35.2	33	87.5	35.2	33	6.2%	-0.68 [-1.18, -0.18]
Hauschild 2013	74.2	13	103	74.3	9.9	26	7.0%	-0.01 [-0.44, 0.42]
Innocenti 2013	56.5	5.24	23	52	9.03	19	4.9%	0.61 [-0.01, 1.24]
Kollig 2003	72.1	21.1	13	82	15.6	9	3.2%	-0.50 [-1.36, 0.37]
Lange 2016	72.3	18	41	72	19	41	6.9%	0.02 [-0.42, 0.45]
Okike 2015	58	16.9	25	59.7	17.5	22	5.4%	-0.10 [-0.67, 0.48]
Roberson 2017	72	40	20	72	40	19	4.9%	0.00 [-0.63, 0.63]
Sanders 2011	71.6	21.1	18	82.5	21.1	18	4.6%	-0.51 [-1.17, 0.16]
Tamimi 2015	65.1	19	88	57.2	12.7	25	6.7%	0.44 [-0.01, 0.89]
Subtotal (95% CI)			**437**			**284**	**65.1%**	**-0.11 [-0.39, 0.18]**
Heterogeneity: Tau² = 0.17; Chi² = 34.62, df = 11 (P = 0.0003); I² = 68%								
Test for overall effect: Z = 0.73 (P = 0.46)								
1.5.2 RCT								
Boons 2012	64	15.8	23	60	17.6	24	5.4%	0.23 [-0.34, 0.81]
Fjalestad 2012-14	51.8	30.6	23	53.2	26.8	19	5.1%	-0.05 [-0.66, 0.56]
Olerud HA 2011	48.3	16.4	24	49.6	20.5	25	5.5%	-0.07 [-0.63, 0.49]
Olerud LP 2011	61	19.2	27	58.4	23.1	26	5.7%	0.12 [-0.42, 0.66]
Rangan 2015	40.1	14.1	114	40.4	13.7	117	9.2%	-0.02 [-0.28, 0.24]
Zyto 1997	60	19	14	65	15	15	4.1%	-0.29 [-1.02, 0.45]
Subtotal (95% CI)			**225**			**226**	**34.9%**	**-0.00 [-0.19, 0.18]**
Heterogeneity: Tau² = 0.00; Chi² = 1.53, df = 5 (P = 0.91); I² = 0%								
Test for overall effect: Z = 0.03 (P = 0.98)								
Total (95% CI)			**662**			**510**	**100.0%**	**-0.06 [-0.25, 0.12]**
Heterogeneity: Tau² = 0.08; Chi² = 36.35, df = 17 (P = 0.004); I² = 53%								
Test for overall effect: Z = 0.64 (P = 0.52)								
Test for subgroup differences: Chi² = 0.36, df = 1 (P = 0.55), I² = 0%								

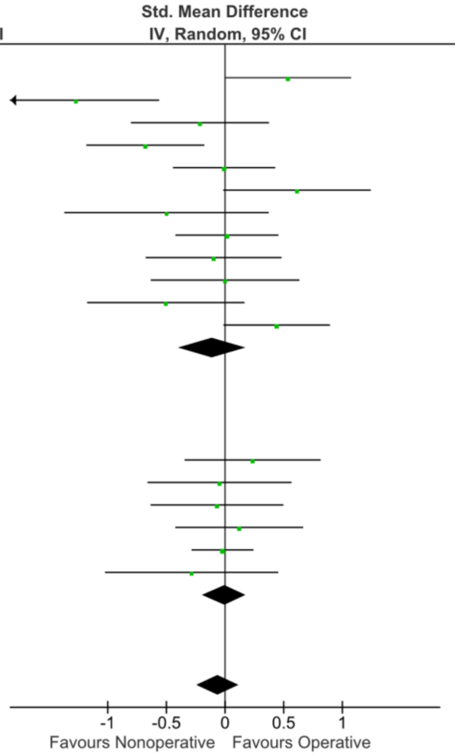

Figure S1. Standardized mean difference of functional outcome scores in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

Study or Subgroup	Operative Mean	Operative SD	Operative Total	Nonoperative Mean	Nonoperative SD	Nonoperative Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
1.7.1 Observational studies								
Lange 2016	72.5	19.5	34	71.8	20.7	34	20.9%	0.03 [-0.44, 0.51]
Roberson 2017	72	40	20	72	40	19	12.0%	0.00 [-0.63, 0.63]
Subtotal (95% CI)			**54**			**53**	**32.8%**	**0.02 [-0.36, 0.40]**
Heterogeneity: Chi² = 0.01, df = 1 (P = 0.93); I² = 0%								
Test for overall effect: Z = 0.11 (P = 0.91)								
1.7.2 RCT								
Boons 2012	64	15.8	23	60	17.6	24	14.3%	0.23 [-0.34, 0.81]
Fjalestad 2012-14	51.8	30.6	23	53.2	26.8	19	12.8%	-0.05 [-0.66, 0.56]
Olerud HA 2011	48.3	16.4	24	49.6	20.5	25	15.0%	-0.07 [-0.63, 0.49]
Olerud LP 2011	61	19.2	27	58.4	23.1	26	16.2%	0.12 [-0.42, 0.66]
Zyto 1997	60	19	14	65	15	15	8.8%	-0.29 [-1.02, 0.45]
Subtotal (95% CI)			**111**			**109**	**67.2%**	**0.02 [-0.25, 0.28]**
Heterogeneity: Chi² = 1.48, df = 4 (P = 0.83); I² = 0%								
Test for overall effect: Z = 0.13 (P = 0.90)								
Total (95% CI)			**165**			**162**	**100.0%**	**0.02 [-0.20, 0.24]**
Heterogeneity: Chi² = 1.49, df = 6 (P = 0.96); I² = 0%								
Test for overall effect: Z = 0.17 (P = 0.86)								
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.99), I² = 0%								

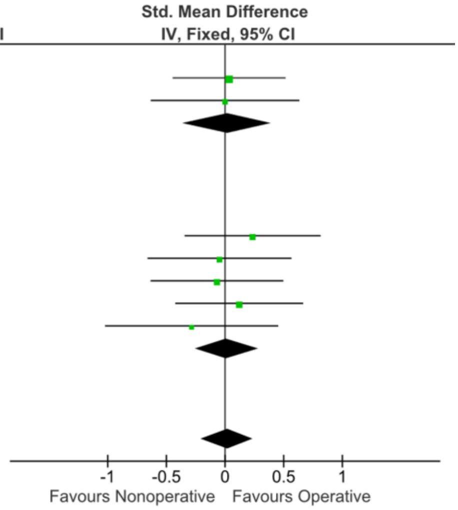

Figure S2. Subgroup analyses looking at standardized mean difference for functional outcome measures including only studies reporting on Neer 3-part or 4-part fractures in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

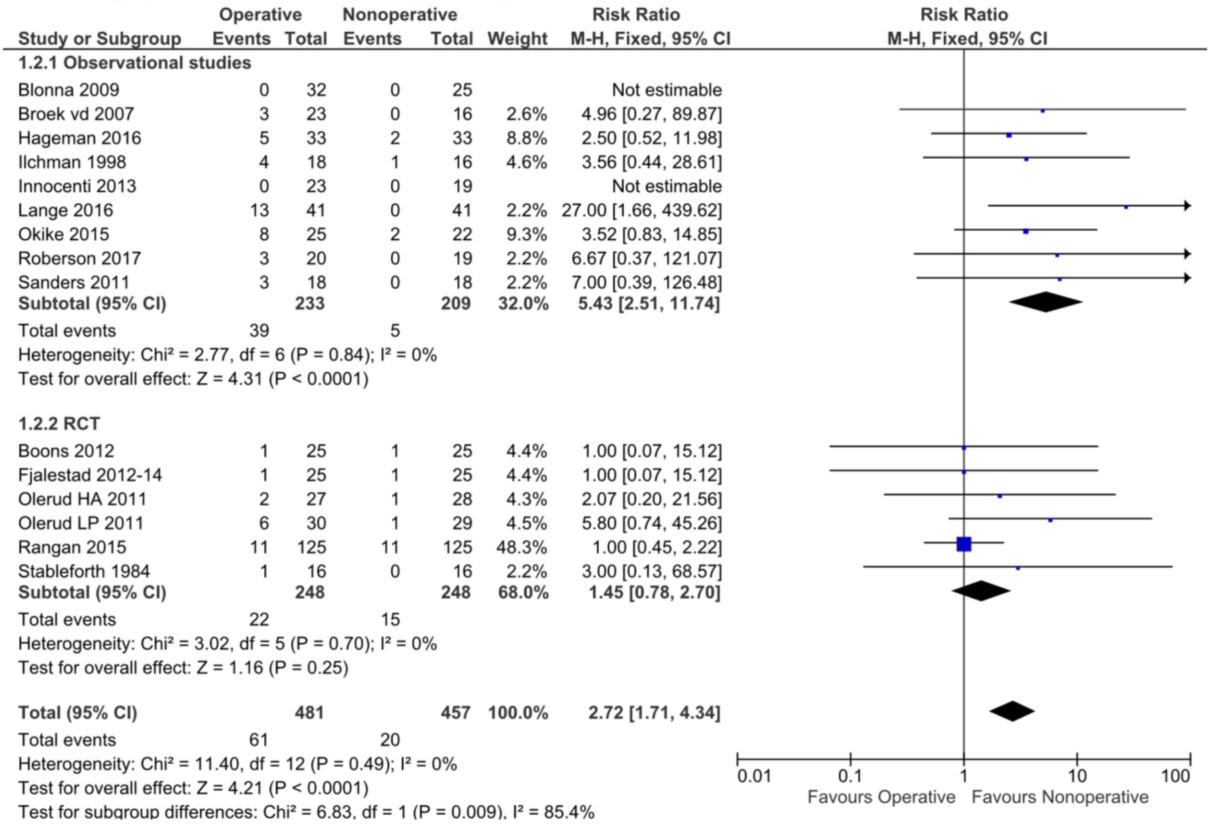

Figure S3. Revision surgery in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

Table S5. Impact of different methods to handle zero-event data

Method	Observational studies	RCT	Total
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Mantel-Haenzel*	5.46 (2.29, 13.01)	1.37 (0.85, 2.77)	2.32 (1.34, 4.02)
Inverse variance - no correction	3.76 (1.30, 10.91)	1.32 (0.64, 2.71)	1.83 (1.01, 3.33)
Inverse variance - with correction	4.64 (2.03, 10.62)	1.37 (0.68, 2.77)	2.29 (1.30, 7.28)
DerSimonian Laird with correction	4.75 (1.43, 15.73)	1.71 (0.57, 5.13)	2.96 (1.26, 7.00)

* Method used in meta-analysis; OR odds-ratio; CI confidence interval. In a model with correction 0.5 is added to every table of the 2x2 table

Study or Subgroup	Operative Events	Operative Total	Nonoperative Events	Nonoperative Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.3.1 Observational studies						
Blonna 2009	0	32	0	25		Not estimable
Broek vd 2007	0	23	1	16	6.9%	0.24 [0.01, 5.45]
Court-Brown 2001	1	18	4	31	11.6%	0.43 [0.05, 3.56]
Fjalestad 2005	1	15	5	55	8.5%	0.73 [0.09, 5.81]
Hauschild 2013	1	103	0	26	3.1%	0.78 [0.03, 18.59]
Okike 2015	0	25	2	22	10.5%	0.18 [0.01, 3.50]
Sanders 2011	0	18	1	18	5.9%	0.33 [0.01, 7.68]
Subtotal (95% CI)		**234**		**193**	**46.6%**	**0.41 [0.15, 1.16]**
Total events	3		13			
Heterogeneity: Chi² = 0.90, df = 5 (P = 0.97); I² = 0%						
Test for overall effect: Z = 1.68 (P = 0.09)						
1.3.2 RCT						
Boons 2012	2	25	3	25	11.9%	0.67 [0.12, 3.65]
Fjalestad 2012-14	1	25	2	25	7.9%	0.50 [0.05, 5.17]
Olerud HA 2011	0	27	1	28	5.8%	0.35 [0.01, 8.12]
Olerud LP 2011	1	30	1	29	4.0%	0.97 [0.06, 14.74]
Rangan 2015	0	125	5	125	21.8%	0.09 [0.01, 1.63]
Zyto 1997	1	20	0	20	2.0%	3.00 [0.13, 69.52]
Subtotal (95% CI)		**252**		**252**	**53.4%**	**0.48 [0.19, 1.20]**
Total events	5		12			
Heterogeneity: Chi² = 3.02, df = 5 (P = 0.70); I² = 0%						
Test for overall effect: Z = 1.57 (P = 0.12)						
Total (95% CI)		**486**		**445**	**100.0%**	**0.45 [0.23, 0.89]**
Total events	8		25			
Heterogeneity: Chi² = 4.03, df = 11 (P = 0.97); I² = 0%						
Test for overall effect: Z = 2.29 (P = 0.02)						
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.82), I² = 0%						

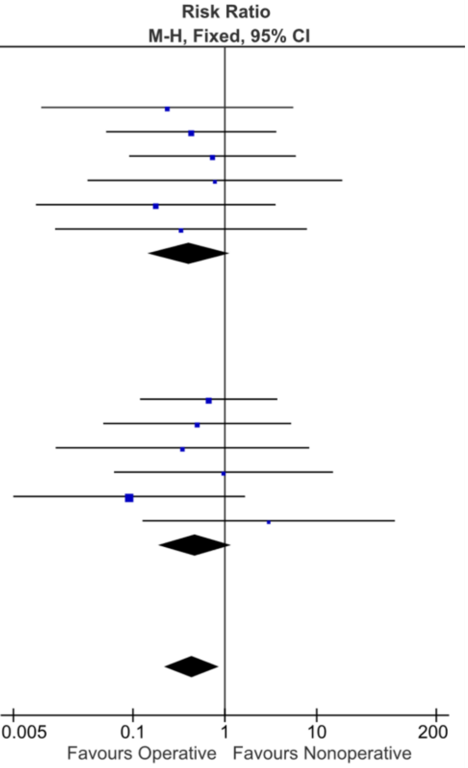

Figure S4. Nonunion in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

Study or Subgroup	Operative Events	Operative Total	Nonoperative Events	Nonoperative Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.6.1 Observational studies						
Broek vd 2007	0	23	0	16		Not estimable
Fjalestad 2005	3	15	2	55	2.3%	5.50 [1.01, 29.98]
Hageman 2016	1	33	0	33	1.3%	3.00 [0.13, 71.07]
Hauschild 2013	1	103	0	26	2.1%	0.78 [0.03, 18.59]
Ilchman 1998	9	18	7	16	19.9%	1.14 [0.55, 2.35]
Innocenti 2013	0	23	0	19		Not estimable
Okike 2015	10	25	3	22	8.6%	2.93 [0.92, 9.32]
Subtotal (95% CI)		**240**		**187**	**34.3%**	**1.93 [1.11, 3.37]**
Total events	24		12			
Heterogeneity: Chi² = 4.38, df = 4 (P = 0.36); I² = 9%						
Test for overall effect: Z = 2.33 (P = 0.02)						
1.6.2 RCT						
Boons 2012	0	25	2	25	6.7%	0.20 [0.01, 3.97]
Fjalestad 2012-14	12	25	15	25	40.3%	0.80 [0.48, 1.34]
Olerud HA 2011	0	27	3	28	9.2%	0.15 [0.01, 2.74]
Olerud LP 2011	3	30	2	29	5.5%	1.45 [0.26, 8.06]
Rangan 2015	4	125	1	125	2.7%	4.00 [0.45, 35.29]
Zyto 1997	1	20	0	20	1.3%	3.00 [0.13, 69.52]
Subtotal (95% CI)		**252**		**252**	**65.7%**	**0.88 [0.55, 1.41]**
Total events	20		23			
Heterogeneity: Chi² = 5.28, df = 5 (P = 0.38); I² = 5%						
Test for overall effect: Z = 0.54 (P = 0.59)						
Total (95% CI)		**492**		**439**	**100.0%**	**1.24 [0.87, 1.77]**
Total events	44		35			
Heterogeneity: Chi² = 13.19, df = 10 (P = 0.21); I² = 24%						
Test for overall effect: Z = 1.19 (P = 0.24)						
Test for subgroup differences: Chi² = 4.51, df = 1 (P = 0.03), I² = 77.8%						

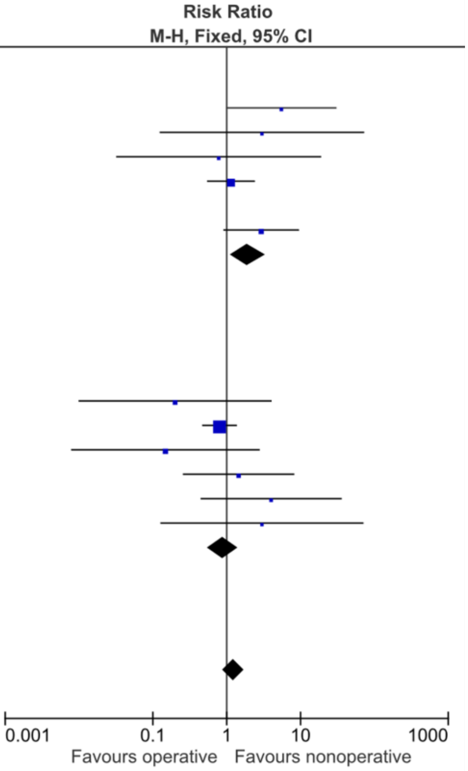

Figure S5. Avascular necrosis in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.