# Surgical Treatment of Proximal Humerus Fractures. A systematic review and meta-analysis

**Running Title: Meta-Analysis Surgical Treatment Proximal Humerus Fractures**


**Erik Hohmann**

MBBS, FRCS, FRCS (Tr & Orth), MD, PhD

Medical School, Faculty of Health Sciences, University of Pretoria

Pretoria, South Africa

Department of Orthopaedic Surgery and Sports Medicine

Valiant Clinic/Houston Methodist Group

**Natalie Keough**

PhD

Department of Anatomy and Cellular Biology, College of Medicine and Health Sciences, Khalifa University, Abu Dhabi, United Arab Emirates

Department of Anatomy, School of Medicine, Faculty of Health Sciences, University of Pretoria, South Africa

**Vaida Glatt**

PhD

Department of Orthopaedics, University of Texas Health Science Center, San Antonio, Texas, USA

Orthopaedic Research Centre of Australia, Brisbane, Australia

**Kevin Tetsworth**

MD, FRACS

Department of Orthopaedic Surgery, Royal Brisbane Hospital, Herston, Australia

Department of Surgery, School of Medicine, University of Queensland, Australia

Limb Reconstruction Center, Macquarie University Hospital, Macquarie Park, Australia

Orthopaedic Research Centre of Australia, Brisbane, Australia

**Corresponding Author:**
Erik Hohmann
Burjeel Hospital for Advanced Surgery
Dubai, United Arab Emirates
drerik@burjeelspecialty.com

**Conflict of Interest Statement:**
The authors declare that they have no conflict of interest

**Funding Sources:**
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Author contribution**

EH, KT and VG substantially contributed to acquisition, analysis and interpretation of data and drafted the article. Statistical analyses were made by EH. All authors contributed substantially to conception and design and revised it critically for important intellectual content and provided final approval of the version to be published. They agreed to act as guarantor of the work ensuring that questions related to any part of the work are appropriately investigated and resolved.

**Availability of data and materials**

Not applicable

## Compliance with ethical standards

## Conflict of interest

The authors declare no conflict of interest with the presented body of research.

## Ethical Approval

Not applicable

## Informed Consent

For this type of study informed consent is not required.


## Abstract

**Introduction:**

The purpose of this study was to perform a systematic review and meta-analysis of both randomized controlled and observational studies comparing surgical interventions for proximal humerus fractures.

**Methods:**

Systematic review of Medline, Embase, Scopus, and Google Scholar, including all level 1-3 studies from 2000 to 2022 comparing surgical treatment with ORIF, IM nailing, hemiarthroplasty, total and reverse shoulder arthroplasty (RTS) was conducted. Clinical outcome scores, range of motion (ROM), and complications were included. Risk of bias was assessed using the Cochrane Collaboration's ROB2 tool and ROBINs-I tool. The GRADE system was used to assess the overall quality of the body of evidence. Heterogeneity was assessed using $\chi^2$ and $I^2$ statistics.

**Results:**

Thirty-five studies were included in the analysis. Twenty-five studies had a high risk of bias and were of low and very low quality. Comparisons between ORIF and hemiarthroplasty favored ORIF for clinical outcomes (p=0.0001), abduction (p=0.002), flexion (p=0.001), and external rotation (p=0.007). Comparisons between ORIF and IM nailing were not significant for clinical outcomes (p=0.0001) or ROM. Comparisons between ORIF and RTS were not significant for clinical outcomes (p=0.0001) but favored RTS for flexion (p=0.02) and external rotation (p=0.02). Comparisons between hemiarthroplasty and RTS favored RTS for clinical outcomes (p=0.0001), abduction (p=0.0001) and flexion (p=0.0001). Complication rates between groups were not significant for all comparisons.

**Conclusions:**

This meta-analysis for surgical treatment of proximal humerus fractures demonstrated that ORIF is superior to hemiarthroplasty, ORIF is comparable to IM nailing, reverse shoulder arthroplasty is superior to hemiarthroplasty but comparable to ORIF with similar clinical outcomes, ROM and complication rates. However, the study validity is compromised by high risk of bias and low level of certainty. The results should therefore be interpreted with caution. Ultimately, shared decision making should reflect the fracture characteristics, bone quality, individual surgeon's experience, the patient's functional demands, and patient expectations.

**Keywords:**

Proximal humerus fractures; surgical treatment; displaced fractures; Neer proximal humerus; clinical outcomes; meta-analysis

**Level of evidence**

Level III; systematic review and meta-analysis

## Introduction

Proximal humerus fractures comprise approximately five percent of all fractures [1]. These fractures are more commonly observed in elderly patients and are often associated with osteoporosis [2]. As demographics in the developed world are changing and the aging population increases, one can expect an increase in the prevalence of these fractures [1,2]. Between 1970 and 1998 an annual increase of 15% of proximal humerus fractures has been observed and further increases can be expected [3]. Most proximal humerus fractures can be treated conservatively with good functional outcomes, and generally only 25% of fractures require operative intervention [4]. Furthermore, surgical intervention often reflects multiple factors including patient age, physical activity, bone quality, fracture patterns, and displacement [5].

The treatment of displaced fractures is still controversial and often depends on the surgeons' experience and personal preferences [6]. For example, shoulder surgeons were more likely to consider arthroplasty and trauma surgeons were more likely to use ORIF or suggest conservative treatment [6]. The current evidence does not seem very helpful and is often limited by study heterogeneity, risk of bias, and low samples sizes [7].

A Cochrane review from 2015 concluded that there is moderate evidence that non-surgical treatment with displaced two-part fractures of the surgical neck is comparable to surgical treatment [8]. However, displaced 3-part and 4-part fractures more likely benefit from surgical treatment, which seems to result in superior outcomes but is also associated with higher rates of reoperations [9]. A broad spectrum of operative interventions have been described and include open reduction and internal fixation

(ORIF) with plates or locking plates, intramedullary nailing, hemiarthroplasty, total shoulder arthroplasty, to reverse shoulder replacement [9,10]. A recent network meta-analysis suggested that reverse shoulder arthroplasty has the best clinical outcomes whilst ORIF had the highest reoperation rates with poor clinical outcomes [10]. However, risk of bias, study heterogeneity, low methodological quality, and omission of randomized but also observational studies substantially reduces the validity of these conclusions [7,11].

The inclusion of observational studies into systematic reviews and meta-analysis is a feasible concept and potentially has several advantages [5]. Inclusion of these studies increases sample size for meta-analysis, and can also increase the generalizability of the pooled results [5]. The inclusion of observational studies into meta-analysis does not result in differences in the risk estimate of treatment effects of an intervention derived from meta-analysis of randomized controlled trials or observational studies [12,13]. The purpose of this study was therefore to perform a systematic review and meta-analysis of both randomized controlled and observational studies comparing surgical interventions for proximal humerus fractures.

## Methods

The study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [14], and the updated guidelines described in the Cochrane Handbook [15].

**Eligibility criteria**

All randomized level of evidence (LOE) grade 1 and 2 and observational LOE 3 studies that compared surgical treatment of proximal humerus fractures were included, if they fulfilled the following inclusion and exclusion criteria. Studies were included if they were published between January 2000 and May 2022. This timeframe was selected to include studies which describe more contemporary techniques. Studies were included if they compared any of the following surgical interventions: ORIF with proximal humerus locking plates, proximal humeral intramedullary nails, hemiarthroplasty, total and reverse shoulder arthroplasty. Further inclusion criteria included: minimum follow-up of 6 months; inclusion of at least one functional validated outcome score such as Constant, ASES, DASH, VAS, OSS, UCLA, or SANE, and preferably the inclusion of range of motion and complications; and minimum mean age of 50 years. Clinical case LOE IV studies and case series were excluded as well as abstracts and conference proceedings. It is accepted that the omission of "grey" data sources could potentially result in publication bias. However, it was considered unlikely that these publications would have fulfilled the eligibility criteria.

**Literature search**

A systematic review of the literature was performed to identify all publications in English and German, screening the databases Medline, Embase, Scopus, and Google Scholar. These databases were screened using the following terms and Boolean operators: "proximal humerus" AND/OR "fracture" AND/OR "Neer proximal humerus fracture" AND/OR "3-part"; AND/OR "4-part" AND/OR "2-part" AND/OR surgical treatment" AND/OR "operative treatment". For the Medline search the following MeSH terms were used in addition to the above search strategy: "fracture, proximal humerus", "fractures, proximal humerus", "proximal humerus fracture", "humerus surgical neck fracture", "humeral surgical neck fracture" and "proximal humerus fractures". Two reviewers conducted independent title and abstract screening. Disagreements between reviewers were resolved by consensus, and if no consensus was reached, they were carried forward to the full text review. All eligible articles were manually cross-referenced to ensure that other potential studies were included.

**Data extraction and quality assessment**

An electronic data extraction form was used to obtain the following data from each article: level of evidence, country, age, gender, length of follow-up, sample size, clinical outcome scores, range of motion and complications. Risk of bias for LOE I and II studies was assessed using the Cochrane Collaboration's Risk of Bias Tool [15]. For LOE III studies the ROBINS-I tool was used. The Methodological Index for Non-Randomized Studies (MINORS) was used to assess study quality for non-randomized studies, and the Modified Coleman Methodology Score (CMS) was used as a validated instrument to assess the quality of both randomized and non-randomized studies. The MINORS score was categorized into the following categories: <6 very low quality, <10 poor quality, <14 fair quality, and >16 good quality. The CMS was categorized as follows: 85-100 excellent quality; 70-84 good quality; 55-69 fair quality; <55 poor quality. Any disagreement between reviewers was resolved by consensus and/or by arbitration between the two senior authors.

The GRADE system was used by two reviewers to assess the certainty of evidence for each outcome measure [15]. The recommendations from the Cochrane Handbook were followed, and an initial level of certainty assigned. Outcomes were downgraded if there was a high risk of bias, inconsistency and imprecision of the results, and indirectness of evidence. Studies were upgraded if there were large treatment effects, a dose-response, or reasons to oppose plausible residual bias and confounding effects. Any disagreement between reviewers was resolved by consensus and/or by arbitration between the two senior authors.

**Statistical analysis**

Inter-observer differences for study eligibility and risk of bias were measured using Cohen's kappa coefficient. Heterogeneity of the data was assessed using $\chi^2$ and $I^2$ statistics. Outcomes were pooled using a random effects model if the $I^2$ statistic was >25%, and a fixed model was used if the statistic was <25%. Pooling of data for clinical outcomes, stability measures, and functional testing was only performed if a minimum of three studies were available. The prevalence of osteoarthritis between groups was pooled as a binary yes/no variable, and analyzed by calculating the odds ratios. If standard deviations were not reported, the standard deviation was calculated using the following formula: SD = max-min/4 [15,16]. All tests of significance were two-tailed, and an $\alpha$ of less than 0.05 was considered significant. Publication bias was assessed using funnel plots and Egger's test. Funnel and forest plots, and all statistical analyses, were performed using STATA SE (Version 13.0; StataCorp, College Station, Texas, USA) for Windows, and the comprehensive meta-analysis software package (CMA), version 3 (Biostat Inc, Englewood, NJ, USA).

## Results

### Study selection and characteristics

The initial literature search identified 3,694 studies for consideration. Of those, 2,225 studies were excluded for duplication, and the titles of the remaining 1,469 publications were checked for eligibility. Another 842 studies were excluded, and following abstract review of 627 studies, the full text manuscripts of 82 studies were examined. Thirty-five studies met all of the eligibility criteria and were included in the analysis [16-50] (Figure 1). Five studies were considered level I evidence [20,28,32,38,50], four were level II evidence [17,24,41,48] and twenty-six studies were level III evidence [16,18,19,21-23,25-27,29-31,33-37,39,40,42-47,49]. All 35 studies were published in English between 2008 and 2022 with a cumulative total of 2672 cases. Of these patients 1061 underwent surgical reconstruction with locked plates, 787 had hemiarthroplasty surgery, 446 were treated with locked intramedullary nailing and 377 underwent reverse total shoulder arthroplasty. The study characteristics are summarized in Tables 1 to 4. Overall agreement between the two reviewers for final eligibility was excellent (kappa value 0.89, 95% CI 0.85-0.92).

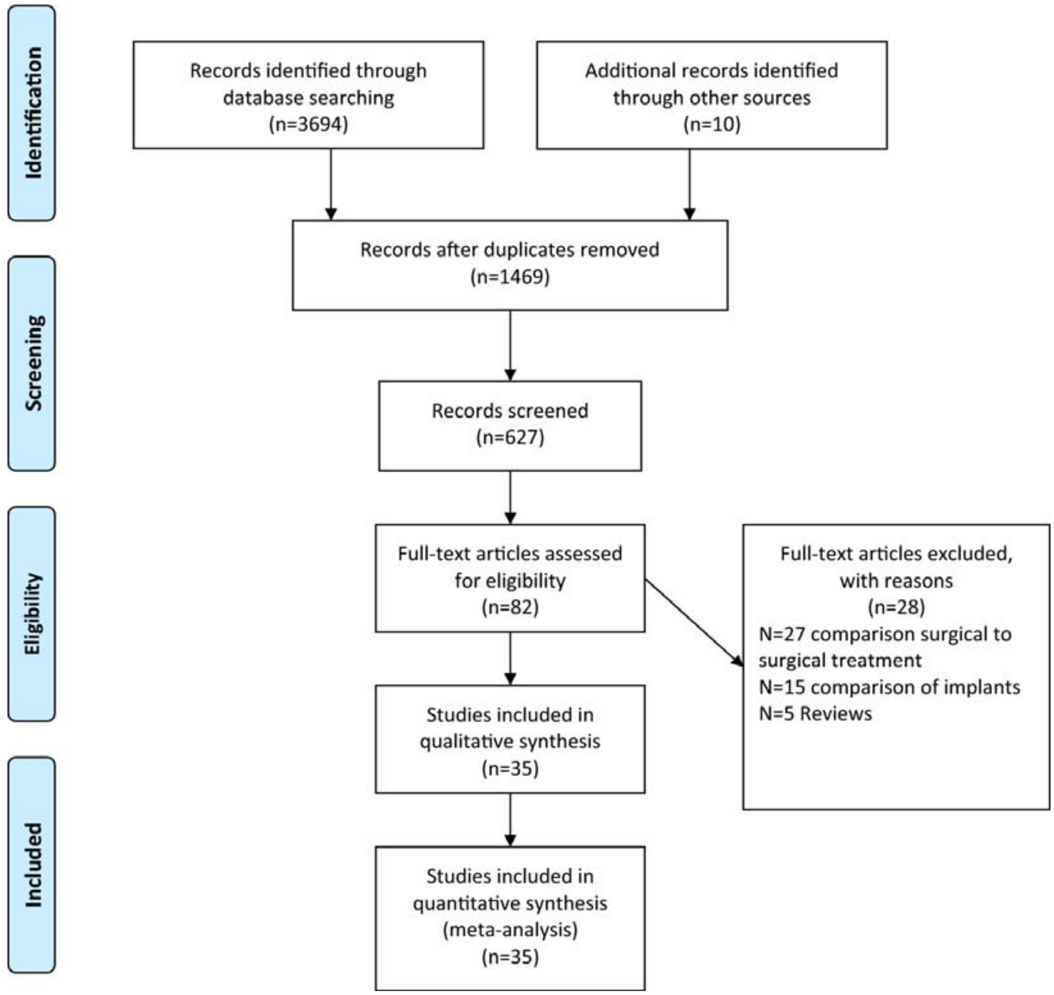


**Figure 1.** PRISMA flow diagram. From the initial 3704 records, 35 studies were included in the quantitative synthesis

**Table 1: ORIF versus Hemiarthroplasty**

| Authors | LOE | Country | Patients (n) Hemi-ORIF | Age (years) | Gender | Neer Classification | Follow-Up (months) | Outcomes |
|---|---|---|---|---|---|---|---|---|
| Dietrich 2008 | III | Germany | 59<br>52 | 80 (70-92)<br>82 (71-92) | M=7 F=52<br>M=10 F=42 | 3 part 19-11<br>4 part 40-40 | Not mentioned | Constant 41 (17-77) – 71 (31-92)<br>OSS: 29 (14-41) – 20 (12-51)<br>Complications: not mentioned |
| Bastian 2009 | II | Switzerland | 51<br>49 | 66 (38-87)<br>54 (21-88) | M=45 F=55 | 2 part 0-6<br>3 part 19-29<br>4 part 30-6 | 60 (39-85) | Constant 70 (39-84) 77 (37-98)<br>SSV 90 (40-100) – 92 (40-100)<br>Complications 16% - 37% |
| Solberg 2009 | III | USA | 48<br>38 | 67.4+6.3<br>66.5+8.6 | M=14 F=34<br>M=12 F=26 | 3 part: 25-23<br>4 part: 15-23<br>Dislocation: 8:10 | 35.3+8.9<br>36.1+10 | Constant 68.6+9.5-60.6+5.9<br>Complications: 21% - 39% |
| Kim 2011 | III | Korea | 26<br>38 | 67+14.8<br>63.5+9.2 | M=10 F=16<br>M=19 F=19 | 3 part: 11 – 24<br>4 part: 15-14 | 24 | Constant 70+7.4 – 75+6.5<br>Complications: not mentioned |
| Cai 2012 | I | China | 19<br>15 | 71.1 (67-85)<br>72.4 (71-86) | M=3 F=16<br>M=4 F=11 | 4 part 19-15 | 24 | Constant: 72.9-60.1<br>DASH 9.2-15.3<br>Complications 16%-23% |
| Spross 2012 | III | Switzerland | 22<br>22 | 76 (55-92)<br>75 (42-93) | M=3 F=19<br>M=4 F=18 | 3 part: 0-3<br>4 part: 6-10<br>Dislocation: 2-1<br>Head impression 7-7<br>Head split 7-1 | 36 (12-83)<br>21 (12-60) | Constant 54.5 (38-86) – 65.2 (41-100)<br>Complications: 4.5%-64% |
| Chalmers 2014 | III | USA | 9<br>9 | 72+7<br>71+7 | M=2 F=7<br>M=2 F=7 | 3+4 part | 35+14<br>36+18 | VAS 3+3 – 1+1<br>ASES 66+31 – 75+15<br>ROM Flexion 106+29 -108+40<br>ROM Ext Rotation 28+19 – 46+21<br>Complications: 11% -11% |
| Repetto 2017 | III | Italy | 24<br>19 | 67.5+10.2<br>65.3+12.4 | | | 42.2+18.7<br>36.5+17.9 | Constant 48.4+27.3 – 61.8+14.7<br>DASH 33.8+28.7 – 33.8+28.7<br>SST 5.6+3.9 – 11.7+3.1<br>ROM Flexion 103+58 – 131+50<br>ROM Ext Rotation 16.5+16 – 23.2+8.8<br>ROM Abduction 90.5+24.7 – 104.4+21.3 |
| Thorsness 2017 | III | USA | 15<br>15 | 73<br>73 | Not mentioned | Dislocation: 3-9<br>Head Spilt: 8-2<br>Both: 4-4 | 60 | Constant: 53.6+19.2 – 69.1+15.7<br>DASH: 29.2+17.6 – 13.8+18.2<br>ASES 66.2+22.1 – 86.5+13.5 |
| Spross 2019 | II | Switzerland | 4<br>36 | 66 | M=58 F=134 | 3+4 part | 12 | Constant 44-63<br>Complications 50% - 44% |
| Erpala 2021 | III | Turkey | 15<br>18 | 68.5+11.3<br>69.5+11.5 | M=5 F=10<br>M=6 F=12 | 3 part 10-12<br>4 part 5-6 | 42.3 (21-54)<br>24.2 (20-38) | DASH 23.3 (14.6-36.2) – 12.1 (5.2-24.2)<br>Constant 49.7+11.8 – 71.6+16.2<br>ASES 54.9 (42-78) – 77.5 (51-97)\ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | ROM Flexion 61+23 – 106+34<br>ROM Abduction 55+18-102-38<br>ROM External Rotation 30 (10-60) – 52 (30-60) |
| Gurnani 2022 | III | India | 10<br>10 | 72.8<br>67.2 | M=5 F=5<br>M=5 F=5 | 3 and 4 part | | Constant 62.6 – 71.3<br>ROM Flexion 89-109<br>ROM Abduction 85-98<br>ROM External Rotation 54-64 |
| Peker 2022 | III | Turkey | 18<br>30 | 67.3+10.1<br>60+9.4 | M=7 F=11<br>M=16 F=30 | 3 part 5-18<br>4 part 6-5<br>Dislocation 7-7 | 18.7+16.4 –<br>18.3+10.2 | Constant 59.1+17.8 – 66.8+17.6<br>UCLA 24.9+4.1 – 26.1+4.4<br>VAS 1.9+1.9 – 2.5+1.8<br>ROM Flexion 63.9+45.4 – 88.3+41.1<br>ROM Abduction 65.6+39 – 80.2+33.1<br>Complications 44% - 30% |

**Table 2: ORIF versus Intramedullary Nailing**

| Authors | LOE | Country | Patients (n) ORIF-Nail | Age (years) | Gender | Neer Classification | Follow-Up (months) | Outcome |
|---|---|---|---|---|---|---|---|---|
| Gradl 2009 | III | Germany | 76<br>76 | 63+16 | M=24F=52<br>M=24 F=52 | 2 part: 26-26<br>3 part: 30 – 30<br>4 part: 16- 16<br>Dislocation: 8 - 8 | 12.8+0.4 | Constant 77+19 – 80+19<br>Complications: 28% - 22% |
| Matziolis 2010 | III | Germany | 11<br>11 | 54.8+16.9<br>55.6+16.5 | M=4 F=7<br>M=4 F=7 | Displaced 2 part | 36 | Constant 83+29-78+17<br>Complications 36% - 36% |
| Zhu 2011 | I | China | 25-25 | 50.5+19.9<br>54.8+17.1 | M=8 F=18<br>M=6 F=16 | 2 part: 25-25 | 36 | VAS: 0.5+1.8 – 1+1<br>ASES: 94+6.3 – 90+8.1<br>Constant: 94.5+5.8 – 93.3+6.7<br>ROM Flexion : 157.3+15.1 – 160.8+11.9<br>ROM External Rotation: 40.4+17.4 – 47.8+17.3<br>Complications: 31% - 4%<br>Infection 0 – 0<br>Screw Penetration 5 – 0 |
| Konrad 2012 | III | Germany | 153<br>58 | 65.5+15.6<br>64.8+13 | M=40 F=113<br>M=11 F=47 | 3 part: 153-58 | 12 | Pain: n= 54 – 9<br>Constant: 87+14 – 89+11<br>Complications: 31% - 21%<br>Intra-operative: 27-5<br>Post-operative: 28-12 |
| Lekic 2012 | III | USA | 12<br>12 | 58.8+17.3<br>60.2+17.7 | M=3 F=9<br>M=2 F=10 | 2 part: 12-12 | 14.7+5.8<br>12.9+9.5 | ROM Flexion : 157.3+15.1 – 160.8+11.9<br>ROM External Rotation: 141.2+30.2 – 120.8+34.5<br>Complications: 33% - 42%<br>Painful hardware 0-2<br>Heterotopic ossification 0-3<br>Screw Penetration 3-0<br>Osteonecrosis 1-0 |
| Trepat 2012 | III | Spain | 11<br>13 | 68.3+17.3<br>64.5+20.7 | M=3 F=8<br>M=6 F=7 | AO Type A31, A 32, A33 | 6 | Constant 64-63.8<br>OSS 19-19.8<br>ROM Flexion: 130-131<br>ROM Abduction 126-121<br>ROM Ext Rotation: 35-39<br>Complications 45%-77% |
| Urda 2012 | III | Spain | 14<br>24 | 71+13.5<br>70.9+11.4 | M=3 F=11<br>M=6 F=18 | 2 part | 40.7+18 | Constant 82.45+17.7 – 72.7+15.9<br>ROM Abduction 134.7+35 – 115.8+33.5<br>ROM Flexion 153+28 – 128+49 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | ROM Ext Rotation 40+17 – 31+15.4<br>Complications 33% - 31% |
| Tamimi 2015 | III | Canada | 44<br>19 | 65.3+15.2 | M=15 F=29<br>M=9 F=10 | 2 part 8-13<br>3+4 part 36-6 | 25.9+15<br>22.5+9 | Constant 62.9+16.8 – 63.9+23.6<br>DASH 38.4+19.2 – 34.9+26.5<br>Complications: 18% - 7% |
| Gracitelli 2016 | I | Brazil | 36<br>36 | 66.4+8.1<br>64.5+9.6 | M=9 F=27<br>M=10 F=25 | 2 part: 16-16<br>3 part: 17-16 | 12 | Constant 71.5+12.8 – 70.3+15.8<br>DASH: 14.3+13 – 18.1+18.8<br>VAS: 1.3+2.1 – 1.7+2.3<br>Complications: 30% - 87% |
| Lee 2017 | I | Korea | 31<br>38 | 58.6+12.3<br>59.7+12 | M=11 F=20<br>M=12 F=26 | 2 part: 31-38 | 24 | VAS 0.9+1.1 – 1.3+1.3<br>ASES 91.9+6.8 – 90.2+7.1<br>UCLA 31.8+2.7 – 30.7+3.1 |
| Ge 2017 | III | China | 69<br>72 | 75.1+8.5<br>76.9+8.1 | M=24 F=45<br>M=22 F=50 | 2 part: 38 – 36<br>3 part: 31 - 36 | 24 | Constant 82.3+9 – 65.4+10.6<br>ASES 80.1+9.0 – 81.5+8.9<br>VAS: 0.81+0.6 – 0.83+066<br>ROM Flexion: 164+9 – 161+12.7<br>ROM External Rotation: 48.3+6.9 – 45.9+8.5 |
| Plath 2019 | II | Germany | 27<br>28 | 77.1 (60-92)<br>71.1 (60-87) | M=7 F=25<br>M=10 F=26 | 2 part: 4-5<br>3 part: 24-25<br>4 part: 4-6 | 12.8 | Constant 64+20-67+20.2<br>DASH 42+1934+17.8<br>VAS 1+1.6-0+1.8<br>ROM Flexion 130+43.5-124.1+45.4<br>ROM Abduction: 123.3+52-120.9+49.1<br>Complications: 34.3% - 33.3% |
| Bu 2021 | III | China | 48<br>34 | 66.3+7<br>67.2+6.3 | M=17 F=31<br>M=11 F=2326 | 2 part: 28-19<br>3 part: 13-10<br>4 part: 7-5 | 15.6 (12-24) | Constant 78.7+.1 – 81.9+9.7<br>ASES 82.8+7.1 – 85.4+9.2<br>VAS 0.52+0.51 – 0.38+0.55<br>Complications 54% - 9% |

**Table 3: ORIF versus Reverse Total Shoulder Arthroplasty**

| Authors | LOE | Country | Patients (n) Reverse-ORIF | Age (years) | Gender | Neer Classification | Follow-Up (months) | Outcome |
|---|---|---|---|---|---|---|---|---|
| Chalmers 2014 | III | USA | 9<br>9 | 77+6<br>71+7 | M=2 F=7<br>M=2 F=7 | 3+4 part | 14+6<br>36+18 | VAS 1+1 – 1+1<br>ASES 80+11 – 75+15<br>ROM Flexion 133+20 -108+40<br>ROM Ext Rotation 41+19 – 46+21<br>Complications: 11% - 11% |
| Repetto 2017 | III | Italy | 27<br>19 | 71.2+7.5<br>65.3+12.4 | | | 41.7+17.1<br>36.5+17.9 | Constant 58.5+8.5 – 61.8+14.7<br>DASH 28.6+12.3 – 33.8+28.7<br>SST 6.7+2.1 – 11.7+3.1<br>ROM Flexion 125+45 – 131+50<br>ROM Ext Rotation 20.3+10.6 – 23.2+8.8<br>ROM Abduction 110+32 – 104.4+21.3 |
| Spross 2019 | II | Switzerland | 20<br>36 | | | | | Constant 69 – 63<br>Complications 0% - 44% |
| Fraser 2020 | I | | 64<br>60 | 75.7+6.1<br>74.7+6.5 | M=5 F=59<br>M=8 F=52 | AO<br>B2 26-29<br>C2: 38-31 | 24 | Constant 68 (64-72) – 54.6 (48.5-60.7)<br>OSS 40.8 (38.8-42.7) – 36.5 (34-39)<br>Complications 11%-20% |
| Klug 2020 | III | Germany | 30<br>30 | 73.9+6.7<br>72.5+6.3 | M=5 F=25<br>M=5 F=25 | 3-part 1:10<br>4-part: 17-18<br>head splitting 12/2 | 49 (12-83)<br>38 (12-50) | ASES 74.6+21.6 – 83.4+17.2<br>OSS 37.7+10.3-42.8+5.7<br>Constant 69.9+26-81.4+17.2<br>DASH 25.3+20—14.3+14.7<br>ROM External Rotation 39+25-52+23<br>Complications 10% - 30% |

**Table 4: Hemiarthroplasty versus Reverse Total Shoulder Arthroplasty**

| Authors | LOE | Country | Patients (n) Reverse-Hemi | Age (years) | Gender | Neer Classification | Follow-Up (months) | Outcome |
|---|---|---|---|---|---|---|---|---|
| Gallinet 2009 | III | France | 16<br>17 | 74 (58-84)<br>74 (49-80.5 | M=3 F=13<br>M=2 F=15 | 3 and 4 part | 12.4 (4-18)<br>16.5 (6-55) | Constant 53 (34-76) – 39 (19-61)<br>DASH 37.4 (11.7-65) – 41.2 (18.3-60.7)<br>ROM Flexion 97 (10-150) – 53 (30-100)<br>ROM Abduction 91 (10-150) – 60 (30-90)<br>ROM Ext Rot 9 (0-80) – 13.5 (0-30)<br>Complications: 19% - 18% |
| Garrigues 2009 | III | USA | 10<br>9 | 80.5 (67-97)<br>69.3 (57-87) | Not mentioned | 3 and 4 part | 42 | ASES 81.1 (75-88) – 47.4 (30-81)<br>SANE 85 (70-95) – 38.5 (0-90)<br>ROM Flexion 121 (90-145) – 91 (30-140)<br>ROM Ext Rotation: 34 (10-45) – 31 (5-60)<br>Complications: 10% - 44% |
| Young 2009 | III | New Zealand | 10<br>10 | 77.2<br>75.5 | M=0 F=10<br>M=2 F=8 | 3 part 2-2<br>4 part: 8-8 | 77.2<br>75.5 | ASES 65 (40-88) – 67 (26-100)<br>OSS 28.7 (15-56) – 22.4 (12-34)<br>ROM Flexion 115 (45-140) – 108 (50-180)<br>ROM Ext Rot 49 (5-105) – 48 (10-90)<br>Complications: 0% - 20% |
| Boyle 2012 | III | New Zealand | 55<br>313 | 79.6 (57-90)<br>71.9 (27-96) | M=4 F=51<br>M=69 F=244 | 3 and 4 part | 60 | OSS 41.5+2.3-32.3+1.2<br>Complications: 3.6-.3.5 |
| Cuff 2013 | II | USA | 24<br>23 | 74.4 | Not mentioned | 4 part | 33 (24-48) | ASES 77 (67-82) – 62 (28-84)<br>SST 7.4 (6-9) – 5.8 (1-9)<br>ROM Flexion 139 (102-172) – 100 (30-170)<br>ROM Ext Rotation: 24 (8-42) – 25 (0-48)<br>Complications: 8% - 9% |
| Baudi 2014 | III | Italy | 25<br>28 | 77<br>70 | | | 26 | Constant 56.2+14.9 – 42.3+16.6<br>DASH 40.4+25 – 46.1+27.9<br>ASES: 59.8+22.9 – 43.5+26.5<br>ROM Flexion 131+36 – 89+44<br>ROM Abduction 128+36-82+40<br>ROM Ext rotation: 15+11-23+15<br>Complications 4% - 24% |
| Chalmers 2014 | III | USA | 9<br>9 | 77+6<br>72+7 | M=2 F=7<br>M=2 F=7 | 3+4 part | 14+6<br>59+14 | VAS 1+1 – 3+3<br>ASES 80+11 – 66+31<br>ROM Flexion 133+20 -106+29<br>ROM Ext Rotation 41+19 – 28+19<br>Complications: 11% - 11% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Sebastia-Forcada 2014 | I | Spain | 31<br>30 | 74.7 (70-85)<br>73.3 (70-83) | M=4 F=27<br>M=5 F=25 | 3 part: 5-4<br>4 part: 21-20<br>4 part dislocation: 5-6 | 24 | Constant 56.1 (24-80) – 40 (8-74)<br>UCLA: 29.1 (16-34) – 21.1 (6-34)<br>DASH: 17.5 (12-30) – 24.4 (13-41)<br>ROM Flexion: 120 (40-180) – 80 (20-180)<br>ROM Abduction: 113 (50-170) – 79 (30-150)<br>ROM External Rotation: 5 (0-10)- 3 (0-10)<br>Complications:6.5% - 26% |
| Repetto 2017 | III | Italy | 27<br>24 | 71.2+7.5<br>67.5+10.2 | | | 41.7+17.1<br>42.2+18.7 | Constant 58.5+8.5 – 48.4+27.3<br>DASH 28.6+12.3 – 33.8+28.7<br>SST 6.7+2.1 – 5.6+3.9<br>ROM Flexion 125+45 – 91+25<br>ROM Ext Rotation 20.3+10.6 – 16.5+16<br>ROM Abduction 110+32 – 90.5+24.7 |
| Spross 2019 | II | Switzerland | 20<br>4 | | | | | Constant 69 – 44<br>Complications 0% - 50% |

**Risk of bias and quality assessment**

The findings of the risk of bias assessment are summarized in Tables 5 and 6.

*Risk of bias Cochrane Assessment Tool Version 2*

Three of the five level I evidence studies [20,28,38] were assessed as having a high risk of bias. Two studies had used an intention to treat protocol [28,38] and one study [20] did not include any information on the randomization process and how bias due to missing outcome data was handled. All four level II evidence studies [17,24,41,48] were assessed as having a high risk of bias. This assessment was based on the high risk of bias introduced by bias from the randomization process. In addition, two studies [17,47] had a loss to follow-up rate exceeding 20%.

*Risk of bias ROBINS-I Assessment Tool*

The ROBINS-I tool was used to assess the risk of bias for observational studies. Sixteen studies [18,19,25,26,29,30,31,34,39,42,43,44,45,46,47,49] had a moderate risk of bias. The main reason for this assessment was based on bias in the selection of participants for all 16 studies. Seven studies [22,23,2735,36,37,40] overall risk of bias was serious. The main reason for this assessment was based on bias due to confounding for all seven. Three studies [16,21,33] had a critical risk of bias based on bias due to deviations from the intended interventions. Publication bias was not detected. The funnel plot was symmetric and Egger's regression intercept (Intercept -1.66, t-value 0.42, p-level 0.65) was not significant (Figure 2).

**Table 5 Risk of Bias Cochrane Risk of Bias Assessment Tool Version 2 for Randomized Controlled Trials**

| Authors | LOE | Bias from Randomization | Bias from Deviations from Intended Interventions | Bias due to Missing Outcome Data | Bias in Measurement of the Outcome | Bias in Selection of the Reported Results | Overall Risk of Bias |
|---|---|---|---|---|---|---|---|
| Zhu 2011 | I | Low | Low | Low | Low | Low | Low |
| Cai 2012 | I | High | Low | High | Low | Low | High |
| Sebastia-Forcada 2014 | I | Low | Low | Low | Low | Low | Low |
| Gracitelli 2016 | I | Low | High | Low | Low | Low | High |
| Fraser 2020 | I | Low | High | Low | Low | Low | High |
| | | | | | | | |
| Bastian 2009 | II | High | Low | High | Low | Low | High |
| Cuff 2013 | II | High | Low | Low | Low | Some | High |
| Plath 2019 | II | High | Low | Low | Low | Low | High |
| Spross 2019 | II | High | High | High | Low | High | High |
| | | | | | | | |

**Table 6 Risk of Bias Assessment for Non-Randomized Studies with the ROBINS-I tool**

| Authors | LOE | Bias due to Confounding | Bias in Selection of Participants | Bias in Classification of Interventions | Bias due to Deviations from Intended Interventions | Bias due to Missing Data | Bias in Measurement of Outcomes | Bias in Selection of the Reported Results | Overall Bias |
|---|---|---|---|---|---|---|---|---|---|
| Dietrich 2008 | III | Critical | Moderate | Moderate | Critical | Moderate | Low | Low | Critical |
| Gallinet 2009 | III | Low | Moderate | Low | Low | Low | Low | Moderate | Moderate |
| Gradl 2009 | III | Moderate | Moderate | Low | Low | Moderate | Low | Low | Moderate |
| Solberg 2009 | III | Low | Moderate | Low | Low | Low | Low | Moderate | Moderate |
| Matziolis 2010 | III | Moderate | Moderate | Moderate | Moderate | Low | Low | Moderate | Moderate |
| Young 2010 | III | Moderate | Moderate | Moderate | Moderate | Low | Low | Low | Moderate |
| Kim 2011 | III | Serious | Moderate | Moderate | Moderate | Low | Low | Low | Moderate |
| Boyle 2012 | III | Low | Moderate | Low | Low | Moderate | Low | Moderate | Moderate |
| Garrigues 2012 | III | Low | Moderate | Low | Low | Moderate | Low | Moderate | Moderate |
| Konrad 2012 | III | Critical | Moderate | Moderate | Critical | Moderate | Low | Moderate | Critical |
| Lekic 2012 | III | Moderate | Moderate | Moderate | Moderate | Low | Low | Moderate | Moderate |
| Spross 2012 | III | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Low | Moderate |
| Trepat 2012 | III | Serious | Moderate | Moderate | Moderate | Moderate | Low | Moderate | Serious |
| Urda 2012 | III | Serious | Moderate | Moderate | Serious | Low | Low | Low | Serious |
| Baudi 2014 | III | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Moderate | Moderate |
| Chalmers 2014 | III | Serious | Moderate | Serious | Critical | High | Low | Low | Critical |
| Ge 2017 | III | Serious | Moderate | Serious | Critical | Low | Low | Low | Serious |
| Lee 2017 | III | Low | Moderate | Low | Low | Moderate | Low | Low | Moderate |
| Tamimi 2015 | III | Serious | Moderate | Serious | Critical | Serious | Low | Low | Serious |
| Repetto 2017 | III | Serious | Moderate | Moderate | Serious | Low | Low | Low | Serious |
| Thorsness 2017 | III | Serious | Moderate | Moderate | Moderate | Serious | Low | Moderate | Serious |
| Klug 2020 | III | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Moderate | Moderate |
| Bu 2021 | III | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Low | Moderate |
| Erpala 2021 | III | Critical | Moderate | Moderate | Moderate | Moderate | Low | Low | Moderate |
| Gurnani 2022 | III | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Moderate | Moderate |
| Peker 2022 | III | Serious | Moderate | Moderate | Serious | Serious | Low | Low | Serious |

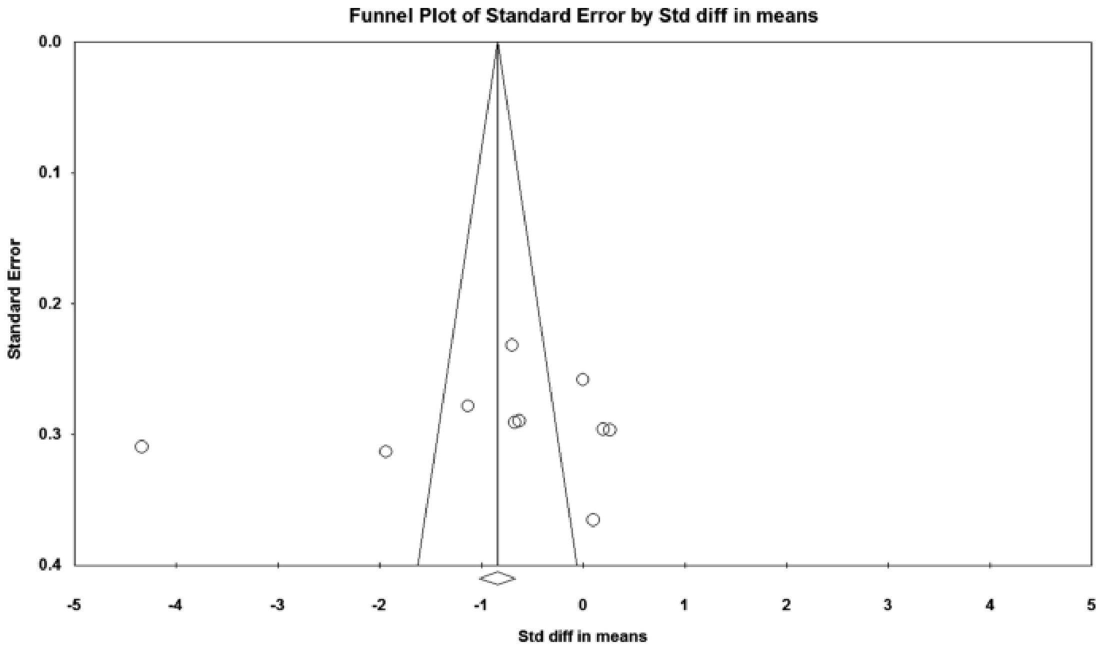


**Figure 2.** Publication bias: Publication bias was not detected. The funnel plot was symmetric and Egger's regression intercept (Intercept − 1.66, *t*-value 0.42, *p*-level 0.65) was not significant

*Quality Assessment*

The CMS score (Table 7) revealed that all five level I evidence studies were of good quality, with a mean score of 75 (range 71-78). All level II evidence studies were assessed as having fair quality, with a mean score of 65 (range 59-67). Five of the twenty-six LOE III studies [18,23,33,40,42] were assessed as having fair quality, and the other twenty-one studies had poor quality. The MINORS score (Table 8) for non-randomized studies assessed one study [40] as having good quality, twenty studies [16,18,21-23,25-27,29-31,33,35,37,39,42,43,46,47,49] as having fair quality, and five studies [19,34,36,44,45] as low quality.

There were obvious differences with regards to quality assessment between CMS and MINORS. These differences are explained by a slightly different focus of these quality scores. MINORS has a focus on an adequate and contemporary control group and

**Table 7: Modified Coleman Methodology Score**

| Authors | Total Points | Study Size | Mean Follow-Up | Percent of patients with follow-up | Number of Interventions | Type of Study | Diagnostic Certainty | Description Surgical Technique | Description Post Op Rehabilitation | Outcome Criteria | Procedures for Assessing Outcomes | Description of Subject Selection |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Zhu 2011 | **78** | **4** | **3** | **3** | **10** | **15** | **5** | **5** | **5** | **10** | **8** | **10** |
| Cai 2012 | **71** | **0** | **0** | **3** | **10** | **15** | **5** | **5** | **5** | **10** | **8** | **10** |
| Sebastia-Forcada 2014 | **75** | **4** | **0** | **3** | **10** | **15** | **5** | **5** | **5** | **10** | **8** | **10** |
| Gracitelli 2016 | **75** | **4** | **0** | **3** | **10** | **15** | **5** | **5** | **5** | **10** | **8** | **10** |
| Fraser 2020 | **75** | **7** | **0** | **0** | **10** | **15** | **5** | **5** | **5** | **10** | **8** | **10** |
| **Mean Score** | **75** | | | | | | | | | | | |
| Bastian 2009 | **67** | **7** | **0** | **0** | **10** | **15** | **5** | **5** | **5** | **10** | **5** | **5** |
| Cuff 2013 | **59** | **4** | **0** | **0** | **10** | **15** | **5** | **5** | **0** | **10** | **5** | **5** |
| Plath 2019 | **67** | **4** | **0** | **0** | **10** | **15** | **5** | **5** | **3** | **10** | **5** | **10** |
| Spross 2019 | **68** | **10** | **0** | **0** | **10** | **15** | **5** | **5** | **0** | **10** | **3** | **10** |
| **Mean Score** | **65** | | | | | | | | | | | |
| Dietrich 2008 | **52** | **7** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **5** | **5** |
| Gallinet 2009 | **33** | **0** | **0** | **3** | **10** | **0** | **5** | **5** | **0** | **10** | **0** | **0** |
| Gradl 2009 | **50** | **10** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **5** | **0** |
| Solberg 2009 | **60** | **7** | **0** | **3** | **10** | **0** | **5** | **5** | **5** | **10** | **5** | **10** |
| Matziolis 2010 | **38** | **0** | **3** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **0** |
| Young 2010 | **35** | **0** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **0** |
| Kim 2011 | **39** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **0** |
| Boyle 2012 | **50** | **4** | **3** | **0** | **10** | **0** | **5** | **5** | **0** | **10** | **8** | **5** |
| Garrigues 2012 | **35** | **0** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **0** |
| Konrad 2012 | **58** | **10** | **0** | **0** | **10** | **10** | **5** | **5** | **0** | **10** | **3** | **5** |
| Lekic 2012 | **40** | **0** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **5** |
| Spross 2012 | **45** | **0** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **5** | **5** |
| Trepat 2012 | **41** | **0** | **0** | **3** | **10** | **0** | **5** | **5** | **5** | **10** | **3** | **0** |
| Urda 2012 | **47** | **4** | **0** | **3** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **5** |
| Baudi 2014 | **44** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **5** | **0** |
| Chalmers 2014 | **51** | **0** | **3** | **5** | **10** | **0** | **5** | **5** | **5** | **10** | **8** | **0** |
| Tamimi 2015 | **53** | **10** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **3** | **5** |
| Ge 2017 | **63** | **10** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **8** | **10** |
| Lee 2017 | **44** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **0** | **5** |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Repetto 2017 | **47** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **3** | **5** |
| Thorsness 2017 | **56** | **0** | **3** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **8** | **10** |
| Klug 2020 | **50** | **4** | **3** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **3** | **5** |
| Bu 2021 | **57** | **4** | **0** | **5** | **10** | **0** | **5** | **5** | **5** | **10** | **8** | **5** |
| Erpala 2021 | **42** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **3** | **0** |
| Gurnani 2022 | **42** | **0** | **0** | **0** | **10** | **10** | **5** | **5** | **0** | **7** | **5** | **0** |
| Peker 2022 | **47** | **4** | **0** | **0** | **10** | **0** | **5** | **5** | **5** | **10** | **8** | **0** |
| | **47** | | | | | | | | | | | |

**Table 8: MINORS Quality Assessment for Non-Randomized LOE III Studies**

| Authors | Total Points | Clearly Stated Aim | Inclusion of Consecutive Patients | Prospective Collection of Data | Appropriate Endpoints | Unbiased Assessment | Follow-up Appropriate | Loss of Follow-up <5% | Sample Size Calculation | Control Group | Contemporary Groups | Baseline Group Equivalence | Adequate Statistical Analysis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dietrich 2008 | **12** | **2** | **2** | **0** | **2** | **0** | **0** | **0** | **0** | **1** | **1** | **2** | **2** |
| Gallinet 2009 | **11** | **1** | **0** | **0** | **2** | **0** | **1** | **1** | **0** | **1** | **2** | **1** | **2** |
| Gradl 2009 | **14** | **2** | **2** | **0** | **2** | **0** | **1** | **1** | **0** | **0** | **2** | **2** | **2** |
| Solberg 2009 | **15** | **2** | **2** | **0** | **2** | **0** | **2** | **1** | **0** | **0** | **2** | **2** | **2** |
| Matziolis 2010 | **11** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **0** | **2** |
| Young 2010 | **10** | **1** | **1** | **0** | **2** | **0** | **1** | **0** | **0** | **0** | **1** | **2** | **2** |
| Kim 2011 | **11** | **1** | **0** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Boyle 2012 | **13** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Garrigues 2012 | **10** | **1** | **2** | **0** | **2** | **0** | **1** | **0** | **0** | **0** | **2** | **0** | **2** |
| Konrad 2012 | **15** | **2** | **2** | **2** | **2** | **0** | **1** | **0** | **0** | **0** | **2** | **2** | **2** |
| Lekic 2012 | **9** | **1** | **1** | **0** | **2** | **0** | **0** | **0** | **0** | **0** | **2** | **2** | **1** |
| Spross 2012 | **13** | **2** | **2** | **0** | **2** | **0** | **1** | **0** | **0** | **0** | **2** | **2** | **2** |
| Trepat 2012 | **14** | **2** | **2** | **0** | **2** | **1** | **1** | **0** | **0** | **0** | **2** | **2** | **2** |
| Urda 2012 | **10** | **2** | **2** | **0** | **2** | **0** | **0** | **0** | **0** | **0** | **2** | **0** | **2** |
| Baudi 2014 | **11** | **1** | **2** | **0** | **2** | **0** | **0** | **0** | **0** | **0** | **2** | **2** | **2** |
| Chalmers 2014 | **12** | **1** | **1** | **0** | **2** | **0** | **2** | **2** | **0** | **0** | **0** | **2** | **2** |
| Tamimi 2015 | **14** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **1** | **2** |
| Ge 2017 | **16** | **2** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **2** | **2** | **2** | **2** |
| Lee 2017 | **15** | **2** | **2** | **0** | **2** | **0** | **2** | **1** | **0** | **0** | **2** | **2** | **2** |
| Repetto 2017 | **15** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **2** | **2** | **2** | **2** |
| Thorsness 2017 | **14** | **2** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Klug 2020 | **13** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Bu 2021 | **14** | **1** | **2** | **0** | **2** | **1** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Erpala 2021 | **13** | **1** | **2** | **0** | **2** | **0** | **2** | **0** | **0** | **0** | **2** | **2** | **2** |
| Gurnani 2022 | **11** | **1** | **2** | **2** | **2** | **0** | **0** | **0** | **0** | **0** | **2** | **1** | **1** |
| Peker 2022 | **12** | **1** | **2** | **0** | **2** | **1** | **0** | **0** | **0** | **0** | **2** | **2** | **2** |

baseline characteristics, whereas CMS has a focus on study size, description of surgical technique, rehabilitation, outcome criteria and their assessment. Despite the subtle differences in quality assessment between the two scores, the outcome of quality assessment for the included studies were relatively similar with no obvious outliers.

Applying the Grade criteria to the individual studies, all randomized controlled trials (RCT) were categorized with initial high quality of evidence and observational comparative studies were categorized with low quality of evidence (Table 9). All RCTs were downgraded. Seven studies had a moderate final level of certainty [17,28,32,38,41,48,50] and two studies [20,47] were downgraded to a low final level of certainty. All comparative observational studies were downgraded to a very low level of certainty. The overall Grade domains were assessed as follows: Limitations in study design was downgraded to moderate as most studies had either some or high risk of bias and these limitations are likely to lower the confidence in the estimate of the effect. The domain inconsistency of results was downgraded to low quality as the heterogeneity as the $I^2$ statistic for the majority (64%) of pooled estimates was above 70% and only four pooled estimates (28%) had an $I^2$ statistic of 0%. The indirectness of evidence domain was downgraded to low quality as it could not be established whether the comparative observational studies sufficiently accounted for potential demographic differences in patient population. The imprecision of results domain was downgraded to low quality as all included studies did not report the 95% confidence intervals. None of the studies included any factors that increased the quality of the evidence. The final GRADE quality assessment was downgraded to low level of certainty as all four domains were downgraded.

**Table 9: Quality Assessment using the Cochrane GRADE system**

| Authors | LOE | Initial Level of Certainty | **Final Level of Certainty** | Limitations in Study Design - Risk of Bias | Inconsistency of Results | Indirectness of evidence | Imprecision of Results | Increase Quality of Evidence | Opposing Plausible Residual Bias and Confounding (Upgrading) |
|---|---|---|---|---|---|---|---|---|---|
| Cai 2012 | I | High | **Low** | High | N/A | | 95% CI missing | | |
| Zhu 2011 | I | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Sebastia-Forcada 2014 | I | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Gracitelli 2016 | I | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Fraser 2020 | I | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| | | | | | N/A | | 95% CI missing | | |
| Bastian 2009 | II | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Cuff 2013 | II | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Plath 2019 | II | High | **Moderate** | Low | N/A | | 95% CI missing | | |
| Spross 2019 | II | High | **Low** | High | N/A | | 95% CI missing | | |
| | | | | | N/A | | 95% CI missing | | |
| Dietrich 2008 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Gallinet 2009 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Gradl 2009 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Solberg 2009 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Matziolis 2010 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Young 2010 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Kim 2011 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Boyle 2012 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Garrigues 2012 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Konrad 2012 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Lekic 2012 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Spross 2012 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Trepat 2012 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Urda 2012 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Baudi 2014 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Chalmers 2014 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Ge 2017 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Lee 2017 | III | Low | **Very Low** | Low | N/A | | 95% CI missing | | |
| Tamimi 2015 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Repetto 2017 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |
| Thorsness 2017 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Klug 2020 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Bu 2021 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Erpala 2021 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Gurnani 2022 | III | Low | **Very Low** | Some | N/A | | 95% CI missing | | |
| Peker 2022 | III | Low | **Very Low** | High | N/A | | 95% CI missing | | |

**Clinical outcomes and between implant comparisons**

The clinical outcomes for all studies are summarized in Tables 1-4. Thirty-three studies [16-33, 35-43,45-50] reported Constant and ASES scores and were included in the analysis. When comparing ORIF to hemiarthroplasty, the pooled estimate for these studies demonstrated significant differences between the two groups in favor of ORIF (SMD -0.556, 95% CI: 0.721 to -0.391, p=0.0001, $I^2$= 88%; Figure 3). According to Cohen, the magnitude effect is medium, suggesting that the SMD represents true differences between groups and that 69% of the hemiarthroplasty group outcomes are below the mean of the ORIF group [51]. When comparing ORIF to IM Nailing, the pooled estimate for these studies demonstrated no significant differences between the two groups. (SMD 0.167, 95% CI: -2.08 to 0.541, p=0.383, $I^2$= 89%; Figure 4). According to Cohen the magnitude effect is small, suggesting that the differences between groups are negligible. [51]. When comparing ORIF to Reverse Shoulder Arthroplasty (RTS) and RTS to hemiarthroplasty, the pooled estimate for both these groups demonstrated non-significant differences between the two groups, and with both comparisons in favor of RTS (SMD 1.037, 95% CI: -0.865 to 2.939, p=0.285, $I^2$= 98%; Figure 5; SMD 1.160, 95% CI: 0.696 to 1.625, p=0.0001, $I^2$= 70%; Figure 6). According to Cohen the magnitude effect is large in the ORIF vs RTS group, suggesting that the SMD represents true differences between groups [51]. However, the lack of statistical significance and overlapping 95% confidence intervals suggest that these differences are not relevant. Regarding the RTS vs hemiarthroplasty, again according to Cohen the magnitude effect is large, strongly suggesting that the SMD represents true differences between groups. In this case, specifically, 79% of the hemiarthroplasty group outcomes are below the mean of the RTS group [51].

## ORIF vs Hemi

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Dietrich 2008 | -1.988 | 0.232 | 0.054 | -2.443 | -1.532 | -8.555 | 0.000 | 13.13 |
| Bastian 2009 | -0.526 | 0.203 | 0.041 | -0.925 | -0.127 | -2.585 | 0.010 | 17.12 |
| Solberg 2009 | 0.986 | 0.230 | 0.053 | 0.536 | 1.436 | 4.291 | 0.000 | 13.42 |
| Kim 2011 | -0.727 | 0.263 | 0.069 | -1.242 | -0.213 | -2.770 | 0.006 | 10.28 |
| Cai 2012 | 0.764 | 0.358 | 0.128 | 0.063 | 1.465 | 2.136 | 0.033 | 5.54 |
| Spross 2012 | -0.803 | 0.313 | 0.098 | -1.418 | -0.189 | -2.563 | 0.010 | 7.21 |
| Chalmers 2014 | -0.370 | 0.475 | 0.226 | -1.301 | 0.562 | -0.777 | 0.437 | 3.14 |
| Repetto 2014 | -0.592 | 0.314 | 0.098 | -1.206 | 0.023 | -1.886 | 0.059 | 7.20 |
| Thorsness 2017 | -0.884 | 0.383 | 0.146 | -1.634 | -0.134 | -2.310 | 0.021 | 4.84 |
| Spross 2019 | -1.235 | 0.545 | 0.297 | -2.303 | -0.167 | -2.267 | 0.023 | 2.39 |
| Erpala 2021 | -1.523 | 0.397 | 0.157 | -2.300 | -0.745 | -3.839 | 0.000 | 4.50 |
| Peker 2022 | -0.436 | 0.301 | 0.091 | -1.026 | 0.155 | -1.445 | 0.148 | 7.80 |
| Gurnani 2022 | -0.534 | 0.455 | 0.207 | -1.426 | 0.358 | -1.173 | 0.241 | 3.42 |
| | -0.556 | 0.084 | 0.007 | -0.721 | -0.391 | -6.609 | 0.000 | |

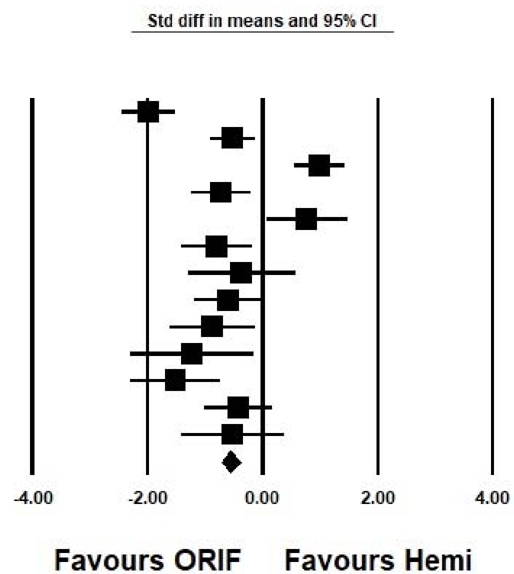


**Figure 3**. Forest plot comparing clinical outcomes of ORIF to hemiarthroplasty. The pooled estimate for all studies demonstrated significant differences in favor of ORIF ($p = 0.0001$)

## ORIF vs IM Nailing

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Gradl 2009 | -0.211 | 0.163 | 0.026 | -0.529 | 0.108 | -1.294 | 0.196 | 9.24 |
| Matziolis 2010 | 0.266 | 0.323 | 0.104 | -0.368 | 0.899 | 0.822 | 0.411 | 7.72 |
| Zhu 2011 | 0.185 | 0.283 | 0.080 | -0.371 | 0.740 | 0.651 | 0.515 | 8.13 |
| Konrad 2012 | -0.151 | 0.154 | 0.024 | -0.454 | 0.152 | -0.978 | 0.328 | 9.30 |
| Trepat 2012 | 0.013 | 0.410 | 0.168 | -0.790 | 0.816 | 0.031 | 0.975 | 6.81 |
| Urda 2012 | 0.588 | 0.343 | 0.118 | -0.084 | 1.261 | 1.715 | 0.086 | 7.51 |
| Tamini 2015 | -0.052 | 0.275 | 0.075 | -0.591 | 0.486 | -0.191 | 0.848 | 8.22 |
| Gracitelli 2016 | 0.083 | 0.236 | 0.056 | -0.379 | 0.546 | 0.354 | 0.723 | 8.61 |
| Lee 2017 | 0.230 | 0.243 | 0.059 | -0.246 | 0.706 | 0.946 | 0.344 | 8.54 |
| Ge 2017 | 1.716 | 0.197 | 0.039 | 1.330 | 2.102 | 8.708 | 0.000 | 8.96 |
| Plath 2019 | -0.149 | 0.270 | 0.073 | -0.679 | 0.380 | -0.553 | 0.581 | 8.27 |
| Bu 2021 | -0.510 | 0.228 | 0.052 | -0.956 | -0.064 | -2.239 | 0.025 | 8.68 |
| | 0.167 | 0.191 | 0.037 | -0.208 | 0.541 | 0.872 | 0.383 | |

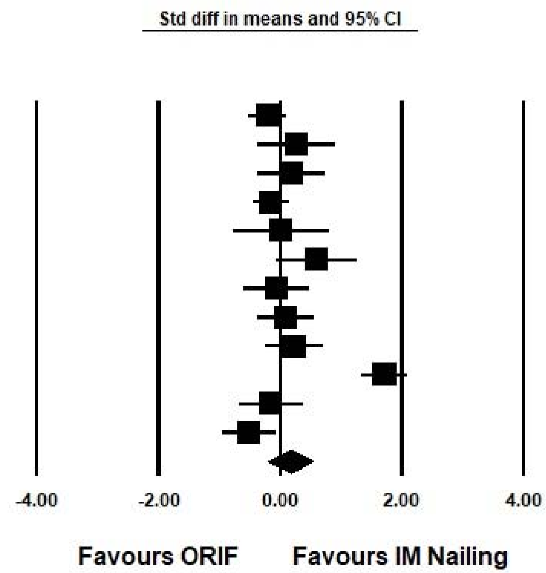


**Figure 4.** Forest plot comparing clinical outcomes of ORIF to intramedullary nailing. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.383$)

## ORIF vs Reverse

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Chalmers 2014 | 0.380 | 0.476 | 0.226 | -0.552 | 1.312 | 0.799 | 0.424 | 19.56 |
| Repeto 2017 | -0.288 | 0.301 | 0.091 | -0.878 | 0.302 | -0.958 | 0.338 | 20.13 |
| Spross 2019 | 0.367 | 0.281 | 0.079 | -0.184 | 0.918 | 1.305 | 0.192 | 20.18 |
| Fraser 2020 | 5.289 | 0.381 | 0.145 | 4.543 | 6.036 | 13.886 | 0.000 | 19.90 |
| Klug 2020 | -0.522 | 0.263 | 0.069 | -1.036 | -0.007 | -1.987 | 0.047 | 20.22 |
| | 1.037 | 0.970 | 0.942 | -0.865 | 2.939 | 1.069 | 0.285 | |

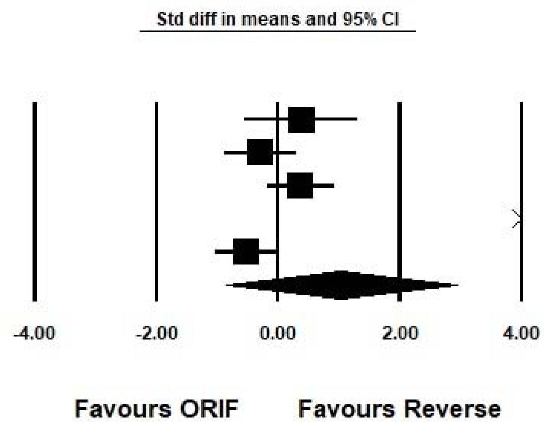


**Figure 5.** Forest plot comparing clinical outcomes of ORIF to reverse shoulder arthroplasty (RTS). The pooled estimate for all studies demonstrated no between-group differences ($p = 0.285$), but a large magnitude effect was observed in favor of RTS

# Hemi vs Reverse

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Gallinet 2009 | 1.333 | 0.385 | 0.148 | 0.579 | 2.088 | 3.463 | 0.001 | 11.76 |
| Garrigues 2009 | 3.722 | 0.759 | 0.576 | 2.235 | 5.209 | 4.905 | 0.000 | 6.21 |
| Young 2009 | -0.149 | 0.448 | 0.201 | -1.027 | 0.729 | -0.333 | 0.739 | 10.60 |
| Cuff 2013 | 1.380 | 0.325 | 0.105 | 0.744 | 2.016 | 4.250 | 0.000 | 12.92 |
| Baudi 2014 | 0.878 | 0.288 | 0.083 | 0.314 | 1.443 | 3.049 | 0.002 | 13.62 |
| Chalmers 2014 | 1.456 | 0.530 | 0.281 | 0.417 | 2.495 | 2.746 | 0.006 | 9.20 |
| Sebastia-Forcada 2014 | 1.054 | 0.273 | 0.075 | 0.518 | 1.589 | 3.855 | 0.000 | 13.90 |
| Repetto 2017 | 0.513 | 0.285 | 0.081 | -0.046 | 1.071 | 1.798 | 0.072 | 13.68 |
| Spross 2019 | 1.722 | 0.601 | 0.362 | 0.543 | 2.901 | 2.863 | 0.004 | 8.13 |
| | 1.160 | 0.237 | 0.056 | 0.696 | 1.625 | 4.895 | 0.000 | |

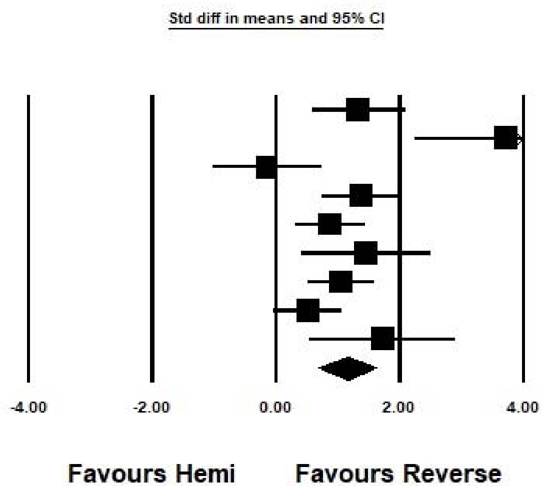


**Figure 6.** Forest plot comparing clinical outcomes of reverse shoulder arthroplasty (RTS) to hemiarthroplasty. The pooled estimate for all studies demonstrated significant ($p = 0.0001$) between-group differences in favor of RTS

**Range of motion**

Eighteen studies [21,22,25-29,32,34-36,40,41,43-45,48-50] included range of motion assessments and were included in the analysis.

*Abduction*

When comparing ORIF to hemiarthroplasty and reverse shoulder arthroplasty (RTS) to hemiarthroplasty, the pooled estimate for these studies demonstrated significant differences within each comparative group in favor of ORIF (SMD -1.002, 95% CI: -1.874 to -0.130, p=0.0024, $I^2$= 82%; Figure 7) and RTS (SMD 1.030, 95% CI: 0.023 to 1.328, p=0.0001, $I^2$= 0%; Figure 9), respectively. According to Cohen the magnitude effect is large, strongly suggesting that the SMD represents true differences between groups, and that 79% of the hemiarthroplasty group outcomes are below the mean of the ORIF and the RTS group [51]. When comparing ORIF to IM Nailing, the pooled estimate for these studies demonstrated no significant differences between the two groups. (SMD 0.256, 95% CI: -0.095 to 0.606, p=0.153, $I^2$= 0%; Figure 8). According to Cohen the magnitude effect is small, suggesting that the differences between groups are negligible. [51]. When comparing Reverse Shoulder Arthroplasty (RTS) to hemiarthroplasty, the pooled estimate for these studies demonstrated significant differences between the two groups in favor of RTS (SMD 1.030, 95% CI: 0.023 to 1.328, p=0.0001, $I^2$= 0%; Figure 9). According to Cohen the magnitude effect is large, strongly suggesting that the SMD represents true differences between groupsand that 79% of the hemiarthroplasty group outcomes are below the mean of the RTS group [51]. Only one study [22] compared abduction between ORIF and Reverse Shoulder Arthroplasty (RTS); pooled estimates were, therefore, not calculated.

## ORIF vs Hemi

| | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| | Statistics for each study | | | | | | | |
| Repetto 2017 | -0.597 | 0.314 | 0.098 | -1.212 | 0.018 | -1.904 | 0.057 | 26.78 |
| Erpala 2021 | -2.611 | 0.475 | 0.226 | -3.542 | -1.680 | -5.498 | 0.000 | 22.85 |
| Gurnani 2022 | -0.574 | 0.456 | 0.208 | -1.469 | 0.320 | -1.258 | 0.208 | 23.31 |
| Peker 2022 | -0.412 | 0.301 | 0.091 | -1.003 | 0.178 | -1.370 | 0.171 | 27.06 |
| | -1.002 | 0.445 | 0.198 | -1.874 | -0.130 | -2.253 | 0.024 | |

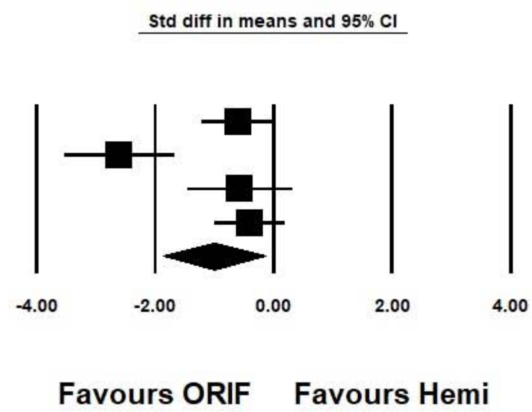


**Figure 7**. Forest plot comparing abduction of ORIF to hemiarthroplasty. The pooled estimate for all studies demonstrated significant differences in favor of ORIF ($p = 0.0024$)

## IM Nailing vs ORIF

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| | Statistics for each study | | | | | | | |
| Trepat 2012 | 0.167 | 0.410 | 0.168 | -0.638 | 0.971 | 0.406 | 0.685 | 19.01 |
| Urda 2012 | 0.559 | 0.294 | 0.087 | -0.018 | 1.135 | 1.898 | 0.058 | 36.98 |
| Plath 2019 | 0.040 | 0.270 | 0.073 | -0.489 | 0.568 | 0.147 | 0.883 | 44.00 |
| | 0.256 | 0.179 | 0.032 | -0.095 | 0.606 | 1.429 | 0.153 | |

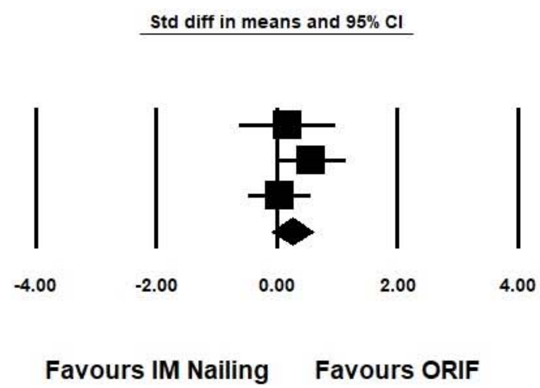


**Figure 8.** Forest plot comparing abduction of reverse shoulder arthroplasty (RTS) to hemiarthroplasty. The pooled estimate for all studies demonstrated significant between-group differences ($p$ = 0.0001) in favor of RTS

## Hemi vs Reverse

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Gallinet 2009 | 1.164 | 0.377 | 0.142 | 0.426 | 1.903 | 3.091 | 0.002 | 16.25 |
| Baudi 2014 | 1.205 | 0.299 | 0.089 | 0.619 | 1.791 | 4.030 | 0.000 | 25.78 |
| Sebastian-Forcada 2014 | 1.133 | 0.276 | 0.076 | 0.593 | 1.674 | 4.108 | 0.000 | 30.28 |
| Repetto 2017 | 0.677 | 0.288 | 0.083 | 0.112 | 1.242 | 2.347 | 0.019 | 27.70 |
| | 1.030 | 0.152 | 0.023 | 0.733 | 1.328 | 6.788 | 0.000 | |

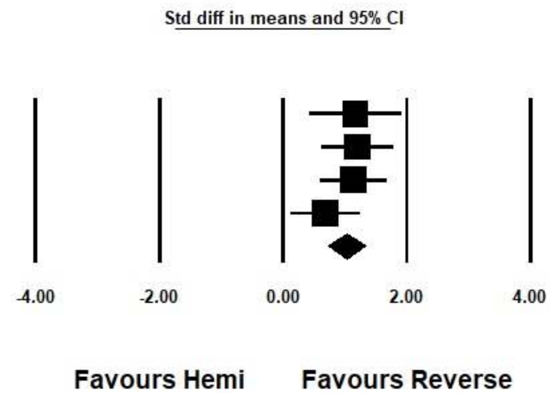


**Figure 9.** Forest plot comparing abduction of ORIF to intramedullary nailing. The pooled estimate for all studies demonstrated no between-group differences ($p$ = 0.153)

*Forward Flexion*

When comparing ORIF to hemiarthroplasty and reverse shoulder arthroplasty (RTS) to hemiarthroplasty, the pooled estimate for both these studies demonstrated significant differences within each comparative group in favor of ORIF (SMD -0.708, 95% CI: -1.143 to -0.272, p=0.001, $I^2$= 92%; Figure 10) and RTS (SMD 1.118, 95% CI: 0.873 to 1.362, p=0.0001, $I^2$= 0%; Figure 12), respectively. According to Cohen the magnitude effect is medium for the ORIF vs hemiarthroplasty comparison, and large for the RTS vs hemiarthroplasty comparison, suggesting that the SMD represents true differences between groups and that 69% and 79% of the hemiarthroplasty group outcomes are below the mean of the ORIF group and RTS group, respectively [51]. When comparing ORIF to IM Nailing, the pooled estimate for these studies demonstrated no significant differences between the two groups. (SMD 0.530, 95% CI: -0.090 to 1.150, p=0.094, $I^2$= 88%; Figure 11). According to Cohen the magnitude effect is large, suggesting that the SMD represents true differences between groups [51]. However, the lack of statistical significance and overlapping 95% confidence intervals suggests that these differences are not relevant. Only two studies [21.22] compared forward flexion between ORIF and Reverse Shoulder Arthroplasty (RTS); pooled estimates were, therefore, not calculated.

## ORIF vs Hemi

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight | Relative weight |
|---|---|---|---|---|---|---|---|---|---|
| Chalmers 2014 | -0.057 | 0.472 | 0.222 | -0.981 | 0.867 | -0.121 | 0.903 | 15.24 | |
| Repetto 2017 | -0.513 | 0.312 | 0.097 | -1.124 | 0.099 | -1.643 | 0.100 | 24.83 | |
| Erpala 2021 | -1.523 | 0.397 | 0.157 | -2.301 | -0.746 | -3.840 | 0.000 | 19.07 | |
| Gurnani 2022 | -0.883 | 0.469 | 0.220 | -1.802 | 0.035 | -1.886 | 0.059 | 15.38 | |
| Peker 2022 | -0.571 | 0.304 | 0.092 | -1.166 | 0.025 | -1.879 | 0.060 | 25.48 | |
| | -0.708 | 0.222 | 0.049 | -1.143 | -0.272 | -3.186 | 0.001 | | |

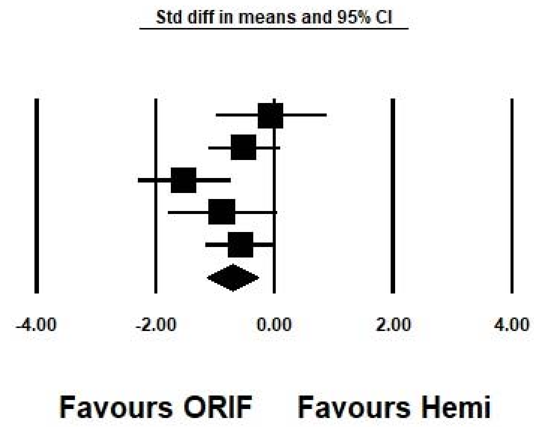


**Figure 10**. Forest plot comparing forward flexion of ORIF to hemiarthroplasty. The pooled estimate for all studies demonstrated significant differences in favor of ORIF ($p = 0.001$)

## IM Nailing vs ORIF

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Zhu 2011 | -0.257 | 0.284 | 0.081 | -0.814 | 0.299 | -0.907 | 0.365 | 18.14 |
| Lekic 2012 | 3.743 | 0.677 | 0.458 | 2.416 | 5.070 | 5.527 | 0.000 | 10.76 |
| Trepat 2012 | -0.033 | 0.410 | 0.168 | -0.836 | 0.770 | -0.081 | 0.935 | 15.66 |
| Urda 2012 | 0.586 | 0.343 | 0.118 | -0.086 | 1.259 | 1.710 | 0.087 | 17.00 |
| Ge 2017 | 0.272 | 0.169 | 0.029 | -0.060 | 0.603 | 1.605 | 0.109 | 20.03 |
| Plath 2019 | 0.136 | 0.270 | 0.073 | -0.394 | 0.665 | 0.502 | 0.616 | 18.40 |
| | 0.530 | 0.316 | 0.100 | -0.090 | 1.150 | 1.676 | 0.094 | |

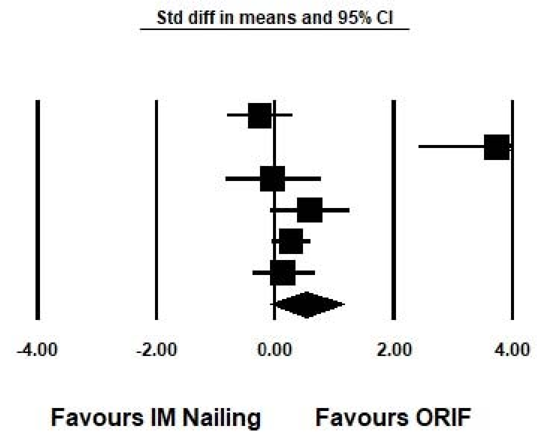


**Figure 11.** Forest plot comparing forward flexion of reverse shoulder arthroplasty (RTS) to hemiarthroplasty. The pooled estimate for all studies demonstrated significant between-group differences ($p = 0.0001$) in favor of RTS

## Hemi vs Reverse

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Gallinet 2009 | 1.596 | 0.400 | 0.160 | 0.812 | 2.380 | 3.991 | 0.000 | 9.71 |
| Garrigues 2009 | 1.380 | 0.511 | 0.261 | 0.378 | 2.382 | 2.700 | 0.007 | 5.94 |
| Young 2009 | 0.239 | 0.449 | 0.201 | -0.641 | 1.119 | 0.533 | 0.594 | 7.71 |
| Cuff 2013 | 1.497 | 0.330 | 0.109 | 0.850 | 2.144 | 4.534 | 0.000 | 14.24 |
| Baudi 2014 | 1.039 | 0.293 | 0.086 | 0.464 | 1.613 | 3.544 | 0.000 | 18.07 |
| Chalmers 2014 | 1.084 | 0.505 | 0.255 | 0.094 | 2.073 | 2.147 | 0.032 | 6.09 |
| Sebastian-Forcada 2014 | 1.134 | 0.276 | 0.076 | 0.593 | 1.675 | 4.110 | 0.000 | 20.39 |
| Repetto 2017 | 0.919 | 0.295 | 0.087 | 0.341 | 1.497 | 3.117 | 0.002 | 17.85 |
| | 1.118 | 0.125 | 0.016 | 0.873 | 1.362 | 8.970 | 0.000 | |

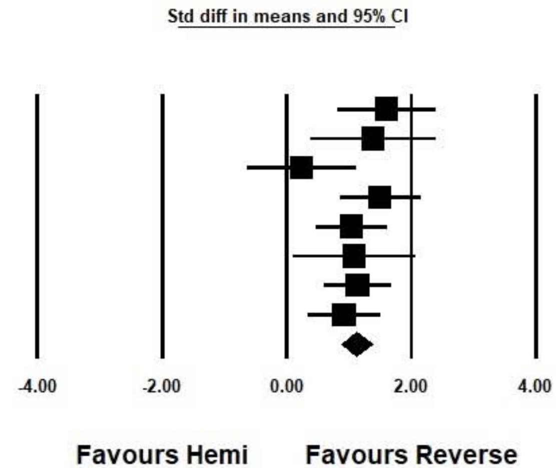


**Figure 12.** Forest plot comparing forward flexion of ORIF to intramedullary nailing. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.094$)

*External rotation*

When comparing ORIF to hemiarthroplasty and ORIF to Reverse Shoulder Arthroplasty (RTS), the pooled estimate for these studies demonstrated significant differences within each comparative group in favor of ORIF (SMD -1.048, 95% CI: -1.809 to -0.287, p=0.007, $I^2$= 71%; Figure 13) and RTS (SMD -0.406, 95% CI: -0.764 to -0.048, p=0.026, $I^2$= 0%; Figure 15), respectively. According to Cohen the magnitude effect for the ORIF vs hemiarthroplasty comparison is large, and medium for the ORIF vs RTS comparison, suggesting that the SMD represents true differences between groups; 79% of the hemiarthroplasty group outcomes are below the mean of the ORIF group, and 69% of the ORIF group outcomes are below the mean of the RTS group [51]. When comparing ORIF to IM Nailing and Reverse Shoulder Arthroplasty (RTS) to hemiarthroplasty, the pooled estimate for these studies demonstrated no significant differences within each comparative group (ORIF vs IM Nailing: SMD 0.122, 95% CI: -0.82 to 0.525, p=0.554, $I^2$= 56%; Figure 14; RTS vs hemiarthroplasty: SMD -0.278, 95% CI: -1.079 to 0.522, p=0.495, $I^2$= 89%; Figure 16). According to Cohen the magnitude effect within each comparative group is small, suggesting that the differences between groups are negligible. [51]. When comparing ORIF to Reverse Shoulder Arthroplasty (RTS), the pooled estimate for these studies demonstrated significant differences between the two groups in favor of RTS (SMD -0.406, 95% CI: -0.764 to -0.048, p=0.026, $I^2$= 0%; Figure 15). According to Cohen the magnitude effect is medium, suggesting that the SMD represent true differences between groups; 69% of the ORIF group outcomes are below the mean of the RTS group [51]. When comparing Reverse Shoulder Arthroplasty (RTS) to hemiarthroplasty, the pooled estimate for these studies demonstrated no significant differences between the two groups. (SMD -0.278, 95% CI: -1.079 to 0.522, p=0.495, $I^2$= 89%; Figure 16).

## ORIF vs Hemi

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Chalmers 2014 | -0.899 | 0.495 | 0.245 | -1.868 | 0.071 | -1.817 | 0.069 | 22.64 |
| Repetto 2017 | -0.503 | 0.312 | 0.097 | -1.114 | 0.108 | -1.612 | 0.107 | 29.07 |
| Erpala 2021 | -2.185 | 0.441 | 0.195 | -3.049 | -1.320 | -4.953 | 0.000 | 24.48 |
| Gurnani 2022 | -0.686 | 0.460 | 0.212 | -1.588 | 0.216 | -1.491 | 0.136 | 23.81 |
| | -1.048 | 0.388 | 0.151 | -1.809 | -0.287 | -2.699 | 0.007 | |

Statistics for each study

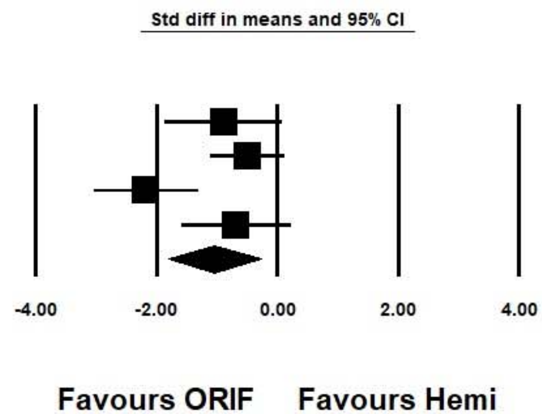


**Figure 13**. Forest plot comparing external rotation of ORIF to hemiarthroplasty. The pooled estimate for all studies demonstrated significant differences in favor of ORIF ($p = 0.007$)

## ORIF vs IM Nailing

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Zhu 2011 | -0.427 | 0.286 | 0.082 | -0.987 | 0.134 | -1.491 | 0.136 | 21.82 |
| Lekic 2012 | 0.624 | 0.418 | 0.175 | -0.196 | 1.443 | 1.492 | 0.136 | 14.75 |
| Trepat 2012 | -0.400 | 0.414 | 0.171 | -1.211 | 0.411 | -0.967 | 0.334 | 14.93 |
| Urda 2012 | 0.571 | 0.343 | 0.117 | -0.100 | 1.243 | 1.668 | 0.095 | 18.43 |
| Ge 2017 | 0.256 | 0.169 | 0.029 | -0.075 | 0.588 | 1.515 | 0.130 | 30.07 |
| | 0.122 | 0.206 | 0.042 | -0.282 | 0.525 | 0.591 | 0.554 | |

(Statistics for each study)

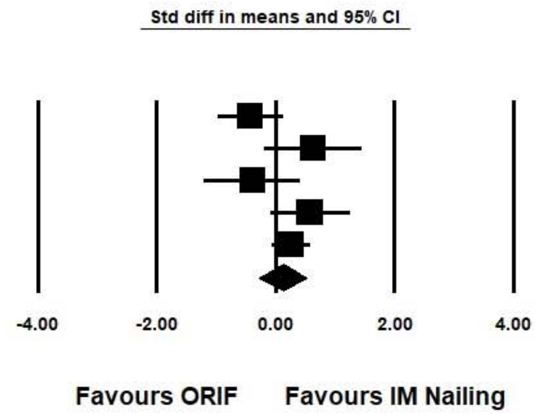


**Figure 14.** Forest plot external rotation of ORIF to reverse shoulder arthroplasty (RTS). The pooled estimate for all studies demonstrated significant between-group differences ($p = 0.026$) in favor of RTS

## Reverse vs ORIF

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Chalmers 2014 | -0.250 | 0.473 | 0.224 | -1.177 | 0.678 | -0.528 | 0.598 | 14.90 |
| Repetto 2017 | -0.293 | 0.301 | 0.091 | -0.883 | 0.297 | -0.973 | 0.331 | 36.82 |
| Klug 2020 | -0.541 | 0.263 | 0.069 | -1.056 | -0.026 | -2.059 | 0.040 | 48.28 |
| | -0.406 | 0.183 | 0.033 | -0.764 | -0.048 | -2.224 | 0.026 | |

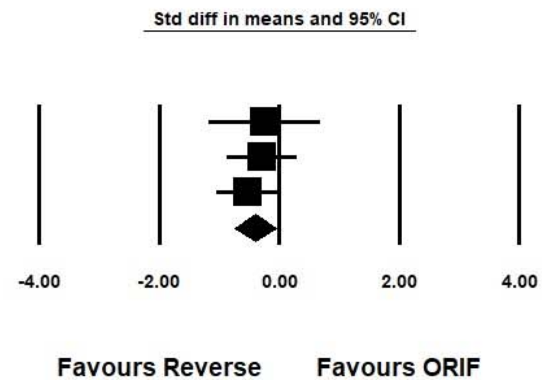


**Figure 15.** Forest plot comparing external rotation of ORIF to intramedullary nailing. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.554$)

## Hemi v Reverse

| Study name | Std diff in means | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|---|---|
| Gallinet 2009 | -0.302 | 0.350 | 0.123 | -0.988 | 0.385 | -0.861 | 0.389 | 12.90 |
| Garrigues 2009 | 0.258 | 0.461 | 0.213 | -0.646 | 1.162 | 0.559 | 0.576 | 12.06 |
| Young 2009 | 0.044 | 0.447 | 0.200 | -0.832 | 0.921 | 0.099 | 0.921 | 12.17 |
| Cuff 2013 | -0.095 | 0.292 | 0.085 | -0.667 | 0.478 | -0.324 | 0.746 | 13.29 |
| Baudi 2014 | -4.493 | 0.599 | 0.359 | -5.667 | -3.318 | -7.498 | 0.000 | 10.91 |
| Chalmers 2014 | 0.684 | 0.485 | 0.235 | -0.266 | 1.635 | 1.411 | 0.158 | 11.87 |
| Sebastian-Forcada 2014 | 0.800 | 0.266 | 0.071 | 0.278 | 1.322 | 3.006 | 0.003 | 13.44 |
| Repetto 2017 | 0.283 | 0.282 | 0.079 | -0.269 | 0.836 | 1.005 | 0.315 | 13.35 |
| | -0.278 | 0.408 | 0.167 | -1.079 | 0.522 | -0.682 | 0.495 | |

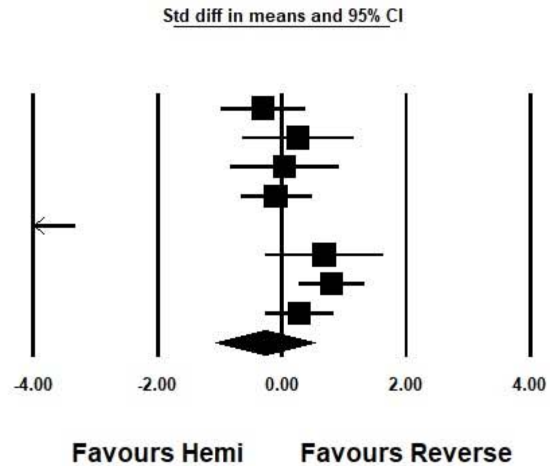


**Figure 16.** Forest plot comparing external rotation of reverse shoulder arthroplasty (RTS) to hemiarthroplasty. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.522$)

According to Cohen the magnitude effect is small, suggesting that the differences between groups are negligible. [51].

**Complications**

When comparing ORIF to hemiarthroplasty, IM Nailing, and Reverse Shoulder Arthroplasty (RTS), the pooled estimates in all cases demonstrated overall lower complication rates in the hemiarthroplasty group (Odds ratio 0.508, 95% confidence intervals 0.215-1.202, p=0.123; (Figure 17), the IM Nailing group (Odds ratio 1.215, 95% confidence intervals 0.849-1.738, p=0.287; Figure 18), and the RTS group (Odds ratio 0.424, 95% confidence intervals 0.192-0.937, p=0.034) (Figure 19), respectively. In the ORIF vs hemiarthroplasty and the ORIF vs IM Nailing groups these differences were not statistically significant, with overlapping 95% confidence intervals and small to medium odds ratios, and the between group differences are considered negligible. [52]. However, in the ORIF vs RTS group, these differences were statistically significant but overlapping 95% confidence intervals were observed, which indicates that the between group differences are negligible. When comparing Reverse Shoulder Arthroplasty (RTS) to hemiarthroplasty, the pooled estimates demonstrated overall lower complication rates in the RTS group (Odds ratio 0.383, 95% confidence intervals 0.190-0.770, p=0.007) (Figure 20). These differences were statistically significant, but overlapping 95% confidence intervals were observed indicating that the between group differences are negligible.

## Hemi vs ORIF

| Study name | Odds ratio | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|
| | Statistics for each study | | | | | |
| Bastian 2009 | 0.320 | 0.124 | 0.830 | -2.342 | 0.019 | 20.92 |
| Solberg 2009 | 0.404 | 0.156 | 1.047 | -1.866 | 0.062 | 20.91 |
| Cai 2012 | 0.750 | 0.128 | 4.389 | -0.319 | 0.750 | 12.87 |
| Spross 2012 | 0.027 | 0.003 | 0.242 | -3.231 | 0.001 | 9.99 |
| Chalmers 2014 | 1.000 | 0.053 | 18.915 | 0.000 | 1.000 | 6.57 |
| Spross 2019 | 1.400 | 0.177 | 11.083 | 0.319 | 0.750 | 10.71 |
| Peker 2022 | 1.867 | 0.554 | 6.286 | 1.008 | 0.314 | 18.03 |
| | 0.508 | 0.215 | 1.202 | -1.540 | 0.123 | |

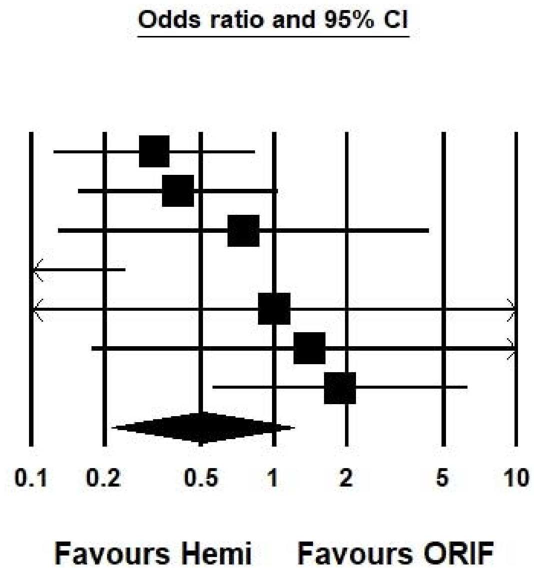


**Figure 17**. Forest plot comparing complication rates of ORIF to hemiarthroplasty. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.123$) but demonstrated overall lower complication rates in the hemiarthroplasty group

## ORIF vs IM Nailing

| Study name | Odds ratio | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|
| Gradl 2009 | 1.325 | 0.634 | 2.770 | 0.748 | 0.454 | 23.59 |
| Matziolis 2010 | 1.000 | 0.176 | 5.682 | 0.000 | 1.000 | 4.25 |
| Zhu 2011 | 11.294 | 1.290 | 98.889 | 2.190 | 0.029 | 2.73 |
| Konrad 2012 | 1.700 | 0.825 | 3.500 | 1.440 | 0.150 | 24.60 |
| Lekic 2012 | 0.700 | 0.133 | 3.684 | -0.421 | 0.674 | 4.65 |
| Trepat 2012 | 0.250 | 0.043 | 1.443 | -1.550 | 0.121 | 4.17 |
| Urda 2012 | 1.349 | 0.332 | 5.491 | 0.418 | 0.676 | 6.51 |
| Tamini 2015 | 4.000 | 0.464 | 34.493 | 1.261 | 0.207 | 2.76 |
| Gracitelli 2016 | 0.071 | 0.022 | 0.231 | -4.390 | 0.000 | 9.20 |
| Plath 2019 | 1.056 | 0.342 | 3.257 | 0.094 | 0.925 | 10.11 |
| Bu 2021 | 12.212 | 3.282 | 45.445 | 3.732 | 0.000 | 7.43 |
| | 1.215 | 0.849 | 1.738 | 1.064 | 0.287 | |

Odds ratio and 95% CI

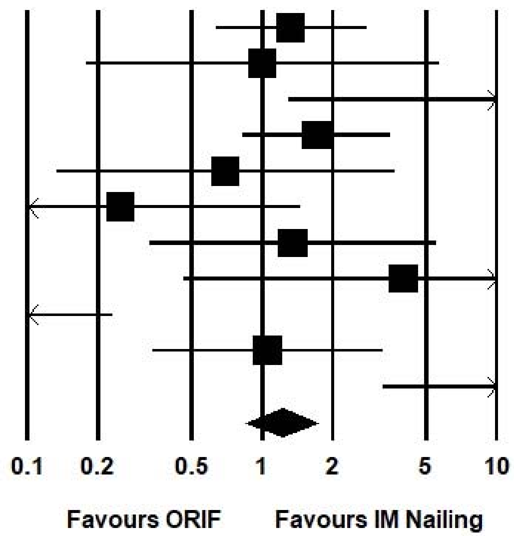


**Figure 18.** Forest plot comparing complication rates of ORIF to intramedullary nailing. The pooled estimate for all studies demonstrated no between-group differences ($p = 0.287$) but demonstrated overall lower complication rates in the IM nail group

# Reverse vs ORIF

| Study name | Odds ratio | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|
| Chalmers 2014 | 1.000 | 0.053 | 18.915 | 0.000 | 1.000 | 7.27 |
| Spross 2019 | 0.000 | 0.000 | 51933184824269001000000.000 | -0.306 | 0.760 | 0.02 |
| Fraser 2020 | 0.491 | 0.179 | 1.346 | -1.382 | 0.167 | 61.81 |
| Klug 2020 | 0.259 | 0.062 | 1.079 | -1.856 | 0.063 | 30.91 |
| | 0.424 | 0.192 | 0.937 | -2.122 | 0.034 | |

Statistics for each study

Odds ratio and 95% CI

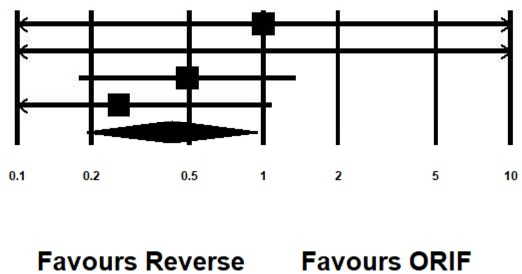


**Figure 19.** Forest plot comparing clinical outcomes of ORIF to reverse shoulder arthroplasty (RTS). The pooled estimate for all studies demonstrated significant between-group differences ($p = 0.034$) in favor of RTS, but overlapping 95% confidence intervals indicate that these differences were not relevant

## Reverse vs ORIF

| Study name | Odds ratio | Lower limit | Upper limit | Z-Value | p-Value | Relative weight |
|---|---|---|---|---|---|---|
| Gallinet 2009 | 1.077 | 0.184 | 6.319 | 0.082 | 0.935 | 15.60 |
| Garrigues 2009 | 0.139 | 0.012 | 1.608 | -1.580 | 0.114 | 8.14 |
| Young 2009 | 0.444 | 0.034 | 5.880 | -0.615 | 0.538 | 7.32 |
| Boyle 2012 | 1.036 | 0.223 | 4.807 | 0.045 | 0.964 | 20.74 |
| Cuff 2013 | 0.457 | 0.039 | 5.409 | -0.622 | 0.534 | 7.99 |
| Baudi 2014 | 0.125 | 0.014 | 1.101 | -1.873 | 0.061 | 10.32 |
| Chalmers 2014 | 1.000 | 0.053 | 18.915 | 0.000 | 1.000 | 5.65 |
| Sebastia-Forcada 2014 | 0.190 | 0.037 | 0.983 | -1.980 | 0.048 | 18.04 |
| Spross 2019 | 0.053 | 0.003 | 0.872 | -2.055 | 0.040 | 6.19 |
| | 0.383 | 0.190 | 0.770 | -2.694 | 0.007 | |

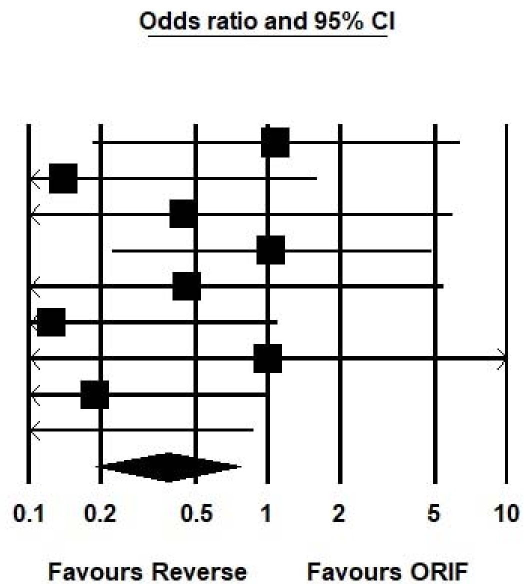


**Figure 20.** Forest plot comparing clinical outcomes of reverse shoulder arthroplasty (RTS) to hemiarthroplasty. The pooled estimate for all studies demonstrated significant ($p = 0.007$) between-group differences in favor of RTS, but overlapping 95% confidence intervals indicate that these differences were not relevant

## Discussion

The results of this meta-analysis had interesting and unexpected results. For three- and four-part fractures ORIF was clearly superior to hemiarthroplasty and resulted in better clinical outcomes and range of motion with similar complication rates. There were no available studies that also included two-part fractures and there is currently no accepted indication to treat two-part fractures with hemiarthroplasty. The data convincingly showed that 69-79% of hemiarthroplasty outcomes were below the mean results reported following ORIF. Although the complication rates were lower in the hemiarthroplasty group this difference was non-significant, with overlapping 95% confidence intervals suggesting that these between group differences were negligible. The meta data therefore advocates ORIF as a better option for 3- and 4-part fractures of the proximal humerus when compared to hemiarthroplasty. This confirms the results of a meta-analysis by Deng et al. [52], who included studies published until 2014 and concluded that patients treated with locked plating had better clinical outcomes with similar complication and reoperation rates compared to patients treated with hemiarthroplasty. Our meta-analysis included five additional studies and increased the sample size by 22%, further strengthening the conclusions of Deng et al. Surgeons should therefore carefully weigh their options, and when deciding whether to use hemiarthroplasty or locked plating for a specific patient they should strongly consider selecting locked plating as the better option.

Comparisons between ORIF and IM nailing did not reveal any significant between group differences. Meta-analysis demonstrated that clinical outcomes, range of motion, and complication rates were very similar for both surgical techniques indicating that either technique is a feasible option. Studies included for meta-analysis covered mainly treatment for 2- and 3-part fractures. Theoretically, locking plates have the advantage of better bending and torsion resistance [53]. In contrast, central fixation with IM nailing can resist greater varus loads [34] and possibly result in higher stability for eversion, flexion, and extension movements [54]. An earlier meta-analysis concluded that IM nailing is superior to locked plating in terms of blood loss, operative time, rates of fracture healing, and complications, but could not establish any differences for clinical outcomes, range of motion, and pain symptoms as these were not analyzed [55]. The study was also limited due to a narrow population group, comprising mainly Chinese studies. Although forest plots for intra-operative blood loss, operation time, fracture healing time, and overall complication rates were constructed, the above-mentioned flaws limit both the internal and external validity of their analysis. In 2017, Sun et al. performed a similar meta-analysis which included 10 studies. Their results were comparable to those of the current meta-analysis, in that they demonstrated that both locking plates and IM nails are comparably effective options for proximal humerus fractures [56]. For this reason, surgeons who consider either of the two options can choose between nailing or plating with equal confidence, and make the selection based on personal preference, available equipment, skills, and personal experience.

When comparing ORIF to reverse shoulder arthroplasty (RTS) there were no between group differences for clinical outcomes or complication rates. Although the meta data suggests that clinical outcomes and complication rates for 3- and 4-part fractures favored RTS with a large magnitude effect, the results were statistically non-significant and overlapping 95% confidence intervals were observed. This is somewhat surprising as there is a current trend towards the use of RTS as the preferred treatment for 3-part, 4-part, and head-splitting fractures [57]. Similar to studies comparing ORIF to hemiarthroplasty, there were no available studies that also included two-part fractures and there is currently no accepted indication to treat two-part fractures with RTS. Orman et al reported that RTS resulted in fewer adverse effects and better clinical outcomes when compared to hemiarthroplasty, although this was comparable to non-surgical treatment [58]. Unbehaun et al. performed a survival analysis for arthroplasty following proximal humerus fractures, deriving data from the Nordic Arthroplasty Register, and reported a 10-year survival rate of 86% with a median time to revision of 18 months [59]. In their opinion these figures constitute low survival rates and concluded that surgeons should be critical when it comes to initial fracture management [59]. However, the meta-data from the current study does not support RTS over ORIF, and surgeons should therefore carefully consider their options. Nevertheless, the large magnitude effect supports RTS, and future studies are required to increase sample sizes and provide further evidence. Certainly, non-operative treatment can also be considered, and Shu et al. reported the results of the members of the science of variation group meeting who suggested that surgeons should consider initial nonoperative treatment with an option for future conversion to reverse arthroplasty [60].

Comparisons between hemiarthroplasty and reverse shoulder arthroplasty (RTS) revealed significant between group differences in favor of RTS. For clinical outcomes, abduction, and flexion range, 79% of the mean outcomes for hemiarthroplasty are below the mean results of RTS for 3- and 4-part fractures. A significantly lower complication rate was also observed in the RTS group. However, despite an odds ratio of 0.38 in favor of the RTS group, overlapping 95% confidence intervals indicate that these results are not clinically relevant. As seen with the comparison between hemiarthroplasty and ORIF, treatment with hemiarthroplasty is clearly inferior and one could argue that hemiarthroplasty for the treatment of proximal humerus fractures should no longer be considered a feasible or reliable option.

The results of this meta-analysis will be helpful in decision-making for all surgeons who treat proximal humerus fractures. The recommendations based on the meta data of this analysis are that ORIF is clearly preferred over hemiarthroplasty, that ORIF and IM nailing produce comparable outcomes. While reverse shoulder arthroplasty is superior to hemiarthroplasty for 3- and 4-part fractures, it provides no obvious advantages over ORIF. However, many other factors should be taken into consideration when planning treatment including patient age, head trauma, concomitant femoral neck fractures, congestive heart failure, chronic alcoholism, obesity, dementia, pneumonia, and anemia, all of which may adversely affect outcomes [61]. Decision-making for these complex and controversial topics is inherently difficult. Internal fixation appears to be the preferred treatment for many surgeons, but there is still no current agreement on the main factors that influence their decisions [62]. Interestingly, 51% of surgeons consider patient-based factors, 51% fracture morphology, 42% surgeon factors, and 11% bone quality [62]. These difficulties in decision-making are also reflected by low agreement rates between surgeons. Experienced shoulder surgeons only agree in 63% of cases on the treatment that should be performed, and these uncertainties may also contribute to the potential for inferior outcomes [64].

**Limitations**

The limitations of this meta-analysis are inherently related to the limitations of the included studies. The Cochrane Handbook indicates that high risk of bias within and across trials may seriously alter the results, and is sufficient to affect the interpretation of results [15]. In this meta-analysis the across trial high risk of bias was 68%, and this is possibly sufficient to conclude that the results should be interpreted with caution. Unfortunately, 30 of the 35 studies had a very low final level of certainty, one study had a low final level of certainty, and only four studies had a moderate final level of certainty. According to the Cochrane Handbook [15], this reduces the confidence in the effect estimate and is sufficient to affect the interpretation of results with the true effect being different. The Grade Recommendations for clinicians are therefore considered to be weak [15]. For this reason, surgeons must recognize that different choices will be appropriate for different patients, and that surgeons must help each patient arrive at a management decision consistent with her or his values and preferences. Decision aids may well be useful in helping individuals making decisions consistent with their values and preferences. Clinicians should expect to spend more time with patients when working towards a shared decision [15]. An individual surgeons' experience and patient expectations will almost certainly continue to be the most important factors to be considered.

## Conclusions

This meta-analysis for surgical treatment of proximal humerus fractures demonstrated that ORIF is superior to hemiarthroplasty, ORIF is comparable to intramedullary nailing, and reverse shoulder arthroplasty is superior to hemiarthroplasty but comparable to ORIF with similar clinical outcomes, range of motion, and complication rates. However, the study validity remains compromised by high risk of bias and low level of certainty, and these results should therefore be interpreted with caution. At this time, shared decision-making should confidently reflect the treating surgeon's skills and experience balanced against patient demands and expectations, tempered by their comorbidities while respecting the realities of bone quality and injury severity.

## References

[1] Dauwe J, Danker C, Herteleer M, Vanhaecht K, Nijs S. (2020): Proximal humeral fracture osteosynthesis in Belgium: a retrospective population-based epidemiologic study. Eur J Trauma Emerg Surg; [online ahead of print]. https://doi 10.1007/s00068-020-01466-2

[2] Roux A, Decroocq S, El Batti N, Bonnevialle G, Trojani C, Boileua P, et al. (2012): Epidemiology of proximal humerus fractures managed in a trauma center. Orthop Traumatol Surg Res; 98 (6):715-719. https://doi 10.1016/j.otsr.2012.05.013

[3] Kannus P, Palvanen M, Niemi S, Pakkari J, Järvinnen M, Vuori I (2000): Osteoporotic fractures of the proximal humerus in elderly Finish persons: sharp increase in 1970-1998 and alarming projections for the new millennium. Acta Orthop Scand; 71 (5):465-470. https://doi 10.1080/000164700317381144

[4] Jawa A, Burnikel D (2016): Treatment of proximal humerus fractures: a critical analysis review. JBJS Rev; 12 (4):e2. https://doi 10.2106/JBJS.RVW.O.00003.

[5] Beks RB, Ochen Y, Frima H, Smeeing DPJ, van der Meijden O, Timmers TK, et al. (2018): Operative versus nonoperative treatment of proximal humerus fractures: a systematic review, meta-analysis, and comparison of observational and randomized controlled trials. J Shoulder Elbow Surg; 27 (8):1526-1534. https://doi 10.1016/j.jse.2018.03.009.

[6] Bhat SB, Secrist ES, Austin LS, Getz CL, Krieg JC, Mehta S, et al. (2016): Displaced proximal humerus fractures in older patients: shoulder surgeons versus traumatologists. Orthopedics; 39 (3):e509-e513. https://doi 10.3928/01477447-20160427-08.

[7] Ionnadis JPA (2016): The mass production of reduandant, misleading, and conflicted systematic reviews and meta-analysis. Milbank Q; 94 (3):485-514. https://doi 10.1111/1468-0009.12210.

[8] Handoll HHG, Brorson S (2015): Interventions for treating proximal humerus fractures in adults. Cochrane Database Syst Rev; 11:CD000434. https://doi 0.1002/14651858.CD000434.pub4.

[9] Fu BS, Jia HL, Zhou DS, Liu FX (2019): Surgical and non-surgical treatment for 3-part and 4-part fractures of the proximal humerus: a systematic review of overlapping meta-analyses. Orthop Surg; 11 (3):356-365. https://doi 10.1111/os.12486.

[10] Du S, Ye J, Chen H, Li X, Lin Q (2017): Interventions for treating 3- or 4-part proximal humerus fractures in elderly patient: A network meta-analysis of randomized controlled trials. Int J Surg; 48:240-246. https://doi 10.1016/j.ijsu.2017.09.002.

[11] Howick J, Koletsi D, Ionnadis JPA, Madigan C, Pandis N, Loef M, et al (2022): Most healthcare interventions tested in Cochrane reviews are not effective according to high quality evidence: a systematic review and meta-analysis. J Clin Epidemiol; SS0895-4356(22)00100-7. https://doi 10.1016/j.jclinepi.2022.04.017.

[12] Golder S, Loke YK, Bland M (2011): Meta-analysis of adverse effects data derived from randomized controlled trials as compared to observational studies: methodological overview. PLos Med; 8 (5):e1001026 https://doi 10.1371/journal.pmed.1001026.

[13] Bun RS, Scheer J, Guillo S, Tubach F, Dechartes A (2020): Meta-analysis frequently pooled different study types together: a meta-epidemiological study. J Clin Epidemiol; 118:18-28. https://doi10.1016/j.jclinepi.2019.10.013.

[14] Moher D, Liberati A, Tetzlaff J, Altman DG (2010): Preferred reporting items for systematic reviews and meta-analysis: the PRISMA statement. Int J Surg; 8 (5):336-34. https://doi.org/10.1016/j.ijsu.2010.02.007

[15] Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JPT, Thomas J (2019): Updated guidance for trusted symptomatic reviews of the Cochrane handbook for systematic reviews of interventions. Cochrane Database Syst Rev; 10:ED000142. https://doi.org/10.1002/14651858.ED000142

[16] Hozo SP, Djulbegovic B, Hozo I (2005): Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol; 5:13. https://doi.org10.1186/1471-2288-5-13.

[16] Dietrich M, Meier C, Lattmann T, Zingg U, Grüninger P, Platz A (2008): Complex fracture of the proximal humerus in the elderly. Locking plate vs hemiarthroplasty. Chirurg; 79 (3):231-240. https://doi.org10.1007/s00104-007-1436-z.

[17] Bastian JD, Hertel R (2009): Osteosynthesis and hemiarthoplasty of fractures of the proximal humerus: outcomes in a consecutive case series. J Shoulder Elbow Surg; 18 (2):216-219. https://doi.org0.1016/j.jse.2008.09.015.

[18] Solberg BD, Moon CN, Franco DP, Paiement GD (2009): Surgical treatment of three- and four-part proximal humeral fractures. J Bone Joint Surg Am; 91 (7):1689-1697. https://doi.org 10.2106/JBJS.H.00133.

[19] Kim DS, Lee DK, Yi CH, Park JH, Rah JH (2011): Comparison of results between internal plate fixation and hemiarthroplasty in comminuted proximal humerus fracture. J Korean Fract Soc; 24 (2):144-150. https://doi.org10.12671/jkfs.2011.24.2.144.

[20] Cai M, Tao K, Yang C, Li S (2012): Internal fixation versus shoulder hemiarthroplasty for displaced 4-part proximal humerus fractures in elderly patients. Orthopedics; 35 (9):e1340-1346. https://doi.org10.3928/01477447-20120822-19.

[21] Chalmers PN, Slikker 3rd W, Mall NA, Gupta AK, Rahman Z, Enriquez D, et al (2014): Reverse total shoulder arthroplasty for acute proximal humeral fracture: comparison to open reduction-internal fixation and hemiarthroplasty. J Shoulder Elbow Surg; 23 (2):197-204. https://doi.org10.1016/j.jse.2013.07.044.

[22] Repetto I, Alessio-Mazzola M, Cerruti P, Sanguineti F, Formica M, Felli L (2017): Surgical management of complex proximal humeral fractures: pinning, locked plate and arthroplasty: clinical results and functional outcome on retrospective series of patients. Muskuloskelet Surg; 101 (2):153-158. https://doi.org10.1007/s12306-017-0451-6.

[23] Thorsness R, Shields E, Chen RE, Owens K, Gorczyca J, Voloshin I (2017): Open reduction and internal fixation versus hemiarthoplasty in the management of complex articular fractures and fracture-dislocations of the proximal humerus. J Shoulder Elbow Arthroplasty; https://doi.org 10.1177/2471549217709364

[24] Spross C, Meester J, Mazzucchelli RA, Puskas GJ, Zdravkovic V, Jost B (2019): Evidence-based algorithm to treat patients with proximal humerus fractures – a prospective study with early clinical and overall performance results. J Shoulder Elbow Surg; 28 (6):1022-1032. https://doi.org 10.1016/j.jse.2019.02.015

[25] Erpela F, Tahta M, Oztürk T, Zengin C (2021): Comparison of treatment options of three- and four-part humerus proximal fractures over 50 years of age. Cureus; 13 (8):e17516. https://doi.org 10.7759/cureus.17516.

[26] Gurnani S, Pisal T, Phalak MO, Chaudari T, Patel S, Yadav P, Mizanur SK (2022): Assessment of surgical outcome in three- and four-part proximal humerus fracture treated with proximal humerus internal locking system (PHILOS) plate versus Neer's prosthesis in elderly patients. Cureus; 14 (2): e21857. https://doi.org 10.7759/cureus.21857.

[27] Peker B, Polat AE, Carkci E, Senel A, Soydan C, Tuzuner T (2022): Functional outcomes and complication analysis of plate osteosynthesis versus hemiarthroplasty in three-part and four-part proximal humerus fractures. J Pak Med Assoc; 72 (1):57-61. https://doi.org 10.47391/JPMA.1147.

[28] Fraser AN, Bjørdal J, Wagle TM, Karlberg AC, Lien OA, Eilertsen L, et al (2020): Reverse shoulder arthroplasty is superior to plate fixation at 2 years for displaced proximal humerus fractures in the elderly. J Bone Joint Surg Am; 102 (6):477-485. https://doi.org 10.2106/JBJS.19.01071.

[29] Klug A, Harth J, Hoffmann R, Gramlich Y (2020): Surgical treatment of complex proximal humerus fractures in elderly patients: a matched-pair analysis of angular-stable vs. reverse shoulder arthroplasty. J Shoulder Elbow Surg; 29 (9):1796-1803. https://doi.org 10.1016/j.jse.2020.01.073.

[30] Gradl G, Dietze A, Kääb M, Hopfenmüller W, Mittlmeier T (2009): Is locking nailing of humeral head fractures superior to locking plate fixation? Clin Orth Relat Res; 467:2986-2993. https://doi.org 10.1007/s11999-009-0916-5

[31] Matziolis D, Kaeaeb M, Zandi SS, Perka C, Greiner S (2010): Surgical treatment of two-part fractures of the proximal humerus: comparison of fixed-angle plate osteosynthesis and Zifko nails. Injury; 41 (10):1041-1046. https://doi.org 10.1016/j.injury.2010.04.017.

[32] Zhu Y, Lu Y, Shen J, Zhang J, Jiang C (2011): Locking intramedullary nails and locking plates in the treatment of two-part proximal humeral surgical neck fractures. J Bone Joint Surg; 93 (2):159-168. https://doi.org 10.2106/JBJS.J.00155.

[33] Konrad G, Audige L, Lambert S, Hertel R, Südkamp NP (2012): Similar outcomes for nail versus plate fixation of three-part proximal humerus fractures. Clin Orth Relat Res; 470:602-609. https://doi.org 10.1007/s11999-011-2056-y.

[34] Lekic N, Montero NM, Takemoto RC, Davidovitch RI, Egol KA (2012): Treatment of two-part proximal humerus fractures: intramedullary nail compared to locked plating. HSS J; 8 (2):86-91. https://doi.org 10.1007/s11420-012-9274-z.

[35] Trepat AD, Popescu D, Fernandez-Valencia JA, Cune J, Rios M, Prat S (2012): Comparative study between locking plates versus proximal humeral nail for the treatment of 2-part proximal humeral fractures. Eur J Orthop Traumatol; 22:373-379. https://doi.org 10.1007/s00590-011-0852-4

[36] Urda A, Gonzalez A, Colino A, Lopiz Y, Garcia-Fernandez C, Marco F (2012): Management of displaced surgical neck fractures of the humerus: health related quality

of life, functional and radiographic results. Injury; 43; Suppl 2:S12-19. https://doi.org 10.1016/S0020-383(13)70174-8.

[37] Tamimi I, Montesa G, Collado F, Gonzalez D, Carnero P, Rojas F, et al (2015): Displaced proximal humeral fractures: when is surgery necessary? Injury; 46 (10):1921-1929. https://doi.org 10.1016/j.injury.2015.05.049.

[38] Gracitelli MEC, Malavolta EA, Assuncao JH, Kojima KE, dos Reis P. Silva JS, Neto AAF, Hernandez AJ (2016): Locking intramedullary nails compared with locking plates for two- and three-part proximal humeral surgical neck fractures: a randomized controlled trial. J Shoulder Elbow Surg; 25 (5):695-703. https://doi.org 10.1016/j.jse.2016.02.003.

[39] Lee W, Park JY, Chun YM (2017): Operative treatment of two-part surgical neck fractures of the humerus: intramedullary nail versus locking compression plate with technical consideration. J Orthop Trauma; 31 (9):3270-e274. https://doi.org 10.1097/BOT.0000000000000916.

[40] Ge W, Sun Q, Li G, Lu G, Cai M, Li SH (2017): Efficacy comparison of intramedullary nails, locking plates and conservative treatment for displaced proximal humeral fractures in the elderly. Clin Interv Aging; 29 (12):2047-2054. https://doi.org 10.2147/CIA.S144084.

[41] Plath JE, Kerschbaum C, Seebauer T, Holz R, Henderson DJH, Förch S, Mayr E (2019): Locking nail versus locking plate for proximal humerus fracture fixation in an elderly population: a prospective randomized controlled trial. BMC Musculoskeletal Disord; 20 (1):20. https://doi.org 10.1186/s12891-019-2399-1.

[42] Bu G, Sun W, Li J, Yang T, Li M, Wei W (2021): Mutiloc nail versus Philos plate in treating proximal humeral fractures: a retrospective study among the elderly. Geriatr Orthop Surg Rehabil; 12:2151459321104961. https://doi.org 10.1177/21514593211043961.

[43] Gallinet D, Clappaz P, Garbuio P, Tropet Y, Obert L (2009): Three or four part complex proximal humerus fractures: hemiarthroplasty versus reverse prosthesis: a comparative study of 40 cases. Orthop Traumatol Surg Res; 95 (1):48-55. https://doi.org 10.1016/j.otsr.2008.09.002

[44] Garrigues GE, Johnston PS, Pepe MD, Tucker BS, Ramsey ML, Austin LS (2012): Hemiarthroplasty versus reverse total shoulder arthroplasty for acute proximal humerus fractures in elderly patients. Orthopedics; 35 (5):e703-708. https://doi.org 0.3928/01477447-20120426-25.

[45] Young SW, Segal BS, Turner PC, Poon PC (2010): Comparison of functional outcomes of reverse shoulder arthroplasty versus hemiarthroplasty in the primary treatment of acute proximal humerus fracture. ANZ J Surg; 80 (11):789-793. https://doi.org 10.1111/j.1445-2197.2010.05342.x.

[46] Boyle MJ, Youn SM, Frampton CMA, Ball CM (2012): Functional outcomes of reverse shoulder arthroplasty compared with hemiarthroplasty for acute proximal humeral fractures. J Shoulder Elbow Surg; 22 (1): 32-37. https://doi.org 10.1016/j.jse.2012.03.006.

[47] Spross C, Platz A, Erschbaumer M, Lattman T, Dietrich M (2012): Surgical treatment of Neer group IV proximal humerus fractures: retrospective comparison of PHILOS® and hemiarthroplasty. Clin Orth Relat Res; 470 (7):2035-2042. https://doi.org 10.1007/s11999-011-2207-1.

[48] Cuff DJ, Pupello DR (2013): Comparison of hemiarthroplasty and reverse shoulder arthroplasty for the treatment of proximal humeral fractures in elderly patients. J Bone Joint Surg Am; 95 (22):2050-2055. https://doi.org 10.2106/JBJS.L.01637.

[49] Baudi P, Campochiaro G, Serafini F, Gazzotti G, Matino G, Rovesta C, et al. (2014): Hemiarthroplasty versus reverse shoulder arthroplasty: comparative study of functional and radiological outcomes in the treatment of acute proximal humerus fracture. Musculoskelet Surg; 98 Suppl 1: 19-25. https://doi.org 10.1007/s12306-014-0322-3.

[50] Sebastia-Forcada E, Cebrian-Gomez R, Lizaur-Utrilla A, Gil-Guillen V (2014): Reverse shoulder arthroplasty for acute proximal humeral fractures. A blinded, randomized, controlled, prospective study. J Shoulder Elbow Surg; 23 (10):1419-1426. https://doi.org 10.1016/j.jse.2014.06.035.

[51] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Erlbaum; 1988

[52] Chen H, Cohen P, Chen S (2009): How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Communications in Statistics; 39 (4):860-864. https://doi.org 10.1080/03610911003650383

[52] Deng J, Zhang S, Yu Y, Zhang L, Jiang W, Yang K, et al. (2021): Efficacy of hemiarthroplasty vs. locking plate fixation for proximal humerus fractures: a meta-analysis. Front Surg; 8:651554. https://doi.org 10.3389/fsurg.2021.651554.

[53] Edwards SL, Wilson NA, Zhang LQ, Flores S, Merk BR (2006): Two-part surgical neck fractures of the proximal part of the humerus. A biomechanical evaluation of two

fixation techniques. J Bone Joint Surg Am; 88 (10):2258-2264. https://doi.org 10.2106/JBJS.E.00757.

[54] Kitson J, Booth G, Day R (2007): A biomechanical comparison of locking plate and locking nail implants used for fractures of the proximal humerus. J Shoulder Elbow Surg; 16 (3):362-366. https://doi.org 10.1016/j.jse.2006.01.019.

[55] Shi X, Liu H, Xing R, Mei W, Zhang L, Ding L, Huang Z, Wang P (2019): Effect of intramedullary nailing and locking plate of proximal humerus fracture: an update systematic review and meta-analysis. J Orthop Surg Res; 14 (1):285. https://doi.org 10.1186/s13018-019-1345-0.

[56] Sun Q, Ge W, Li G, Wu J, Lu G, Cai M, et al. (2018): Locking plates versus intramedullary nails in the management of displaced proximal humeral fractures: a systematic review and meta-analysis. Int Orthop; 42 (3):641-650. https://doi.org 10.1007/s00264-017-3683-z.

[57] Kelly BJ, Myeroff CM (2020): Reverse shoulder arthroplasty for proximal humerus fracture. Curr Rev Musculoskelet Med; 13 (2):186-199. https://doi.org 10.1007/s12178-020-09597-0.

[58] Orman S, Mohamadi A, Serino J, Murphy J, Hanna P, Weaver MJ, et al. (2020): Comparison of surgical and non-surgical treatments for 3- and 4-part proximal humerus fractures: a network meta-analysis. Shoulder Elbow; 12 (2):99-108. https://doi.org 10.1177/1758573219831506

[59] Unbehaun D, Rasmussen S, Hole R, Fenstad AM, Salomonsson B, Demir Y, Jensen SL, Brorson S, Äärimaa V, Mechlenburg I, Rasmussen JV (2020): Low arthroplasty survival after treatment for proximal humerus fracture sequelae: 3,245 shoulder replacement from the Nordic Arthroplasty Register Association. Acta Orthop; 17:1-6. https://doi.org 10.1080/17453674.2020.1793548.

[60] Shu R, Crijns T, Ring D, Fatehi A (2021): Surgeons consider nonoperative treatment with potential for future conversion to reverse arthroplasty a reasonbale option for older, relatively infirm and less active patients. J Orthop Trauma; https://doi.org 10.1097/BOT.0000000000002278.

[61] Neuhaus V, Swellengrebel CH, Bossen JKJ, Ring D (2013): What are the factors influencing outcomes among patients admitted to a hospital with a proximal humerus fracture? Clin Orth Relat Res; 471 (5):1698-1706. https://doi.org 10.1007/s11999-013-2876-z.

[62] Gradl G, Knobe M, Pape HC, Neuhaus PV, Ring D, Guitton T (2015): Decision making in displaced fractures of the proximal humerus: fracture or surgeon based? Int Orthop; 39 (2):329-334. https://doi.org10.1007/s00264-014-2630-5.

[63] LaMartina 2nd J, Christmas KN, Simon P, Streit JJ, Allert JW, Clark J, et al. (2018): Difficulty in decision making in the treatment of displaced proximal humerus fractures: the effect of uncertainty on surgical outcomes. J Shoulder Elbow Surg; 27 (3):470-477. https://doi.org 10.1016/j.jse.2017.09.033.