Shoulder & Elbow

EOR | VOLUME 6 | OCTOBER 2021
DOI: 10.1302/2058-5241.6.210049
www.efortopenreviews.org

EFORT open reviews

Reverse total shoulder arthroplasty (RTSA) versus open reduction and internal fixation (ORIF) for displaced three-part or four-part proximal humeral fractures: a systematic review and meta-analysis

Heri Suroto[1,5]
Brigita De Vega[2,3,5]
Fani Deapsari[3]
Tabita Prajasari[1]
Pramono Ari Wibowo[1]
Steven K. Samijo[4]

- Despite rapid medical technology development, various challenges exist in three- and four-part proximal humeral fracture (PHF) management. This condition has led to a notably increased use of the reverse total shoulder arthroplasty (RTSA); however, open reduction and internal fixation (ORIF) is still the most widely performed procedure. Thus, these two modalities are crucial and require further discussion. We aim to compare the outcomes of three- or four-part PHF surgeries using ORIF and RTSA based on direct/head-to-head comparative studies.
- We conducted a systematic review and meta-analysis based on the Cochrane handbook and PRISMA guidelines. We searched MEDLINE (PubMed), Embase (Ovid), and CENTRAL (Cochrane Library) from inception to October 2020. Our protocol was registered at PROSPERO (registration number CRD42020214681). We assessed the individual study risk of bias using ROB 2 and ROBINS-I tools, then appraised our evidence using the GRADE approach.
- Six head-to-head comparative studies were included, comprising one RCT and five retrospective case-control studies. We found that RTSA significantly improved forward flexion but was comparable to ORIF in abduction ($p = 0.03$ and $p = 0.47$, respectively) and more inferior in external rotation ($p < 0.0001$). Moreover, RTSA improved the overall Constant-Murley score, but the difference was not significant ($p = 0.22$). Interestingly, RTSA increased complications (by 42%) but reduced the revision surgery rates (by 63%) compared to ORIF ($p = 0.04$ and $p = 0.02$, respectively).
- RTSA is recommended to treat patients aged 65 years or older with a three- or four-part PHF. Compared to ORIF, RTSA resulted in better forward flexion and Constant-Murley score, equal abduction, less external rotation, increased complications but fewer revision surgeries.

Keywords: humeral fractures; internal fracture fixation; meta-analysis; shoulder fractures; shoulder replacement arthroplasty; systematic review

Cite this article: *EFORT Open Rev* 2021;6:941-955.
DOI: 10.1302/2058-5241.6.210049

Introduction

Proximal humeral fractures (PHF) account for approximately 5% of all fractures. This type of fracture is most often found in women older than 60 years, and its incidence has been reported to increase with age.[1–4] As the world population ages, the PHF incidence is also projected to increase threefold by 2030[2] since advanced age, osteoporosis, and falls are the main risk factors for such fractures.[5,6] In general, fractures in this age group represent a considerable burden to the patients in terms of pain, loss of function, and even mortality.[7] Physical disabilities caused by PHF also contribute to difficulties in self-care, independent living, and finally disrupt the quality of life.[8]

Moreover, complex fracture patterns such as displaced three- or four-part PHF cause their own challenges. Several surgical intervention options for the treatment of complex PHF are open reduction and internal fixation (ORIF) and arthroplasty/joint replacement (either hemiarthroplasty (HA), anatomic total shoulder arthroplasty (TSA), or reverse

total shoulder arthroplasty (RTSA)). A recent epidemiological study comprising 1,162 patients in a level-1 trauma centre reported that there was a substantial increase of complex PHF fractures (+9.6%) followed by increased ORIF and RTSA (+13.5% and +9.4%, respectively), while rates of other treatment options (including conservative, nail osteosynthesis, and TSA) were decreasing (–7.2%, –8.0%, and –4.4%, respectively).[9] Moreover, the National Joint Registry in the United Kingdom recorded that RTSA is the most commonly performed shoulder arthroplasty in 2012–2019. During this period, the trend is decreasing for HA and TSA but increasing for RTSA, with a total number of procedures of 19,300 for RTSA alone.[10] This fact shows that ORIF and RTSA are the two modalities that require further discussion.

Although many prior systematic reviews have compared various management strategies for PHF, these studies found no or limited direct/head-to-head comparative studies between ORIF and RTSA.[11–13] Previous literature has described methodological concerns of systematic reviews that include indirect comparison.[14] Therefore, the purpose of this systematic review is to compare the outcomes of three- or four-part PHF surgeries using ORIF and RTSA based on direct/head-to-head comparative studies.

Methods

This systematic review followed PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and the *Cochrane handbook for systematic reviews of interventions*. Our full protocol was registered in PROSPERO (registration number CRD42020214681).[15]

Eligibility criteria

We included clinical studies of a direct/head-to-head comparison of reverse total shoulder arthroplasty (RTSA) versus open reduction and internal fixation (ORIF) for the treatment of three- or four-part proximal humeral fractures. We did not set any restrictions on language or publication time. However, we excluded studies involving other types of fracture configuration or other surgical procedures in which the raw data could not be separated. We also excluded case series, case reports, reviews, systematic reviews, meta-analysis, editorials, letters, book chapters, study protocols, non-clinical/pre-clinical studies (in vitro, cadaver, animals) and conference abstracts in which a full report was unavailable.

Electronic search

We searched the following databases from inception to October 2020: MEDLINE (PubMed), Embase (Ovid), and CENTRAL (Cochrane Library). We developed the search strategy by combining free-texts and subject headings (MeSH for MEDLINE and Emtree for Embase) to enhance the sensitivity and specificity of our search. We used the following search strategy: (((((Three-part[Title/Abstract] OR three part*[Title/Abstract] OR 3-part[Title/Abstract] OR 3 part*[Title/Abstract]) AND ((Fracture*[Title/Abstract]) OR (bone fractures[MeSH Terms]))) AND (((Humeral head[Title/Abstract] OR humeral neck[Title/Abstract] OR proximal humerus[Title/Abstract] OR proximal humeral[Title/Abstract] OR humeral[Title/Abstract] OR humerus[Title/Abstract] OR shoulder[Title/Abstract] OR glenohumeral[Title/Abstract]) OR (Humerus[MeSH Terms])) OR (shoulder[MeSH Terms]) OR (shoulder joint[MeSH Terms]))) OR (((Four-part[Title/Abstract] OR four part*[Title/Abstract] OR 4-part[Title/Abstract] OR 4 part*[Title/Abstract]) AND ((Fracture*[Title/Abstract]) OR (bone fractures[MeSH Terms]))) AND (((Humeral head[Title/Abstract] OR humeral neck[Title/Abstract] OR proximal humerus[Title/Abstract] OR proximal humeral[Title/Abstract] OR humeral[Title/Abstract] OR humerus[Title/Abstract] OR shoulder[Title/Abstract] OR glenohumeral[Title/Abstract]) OR (Humerus[MeSH Terms])) OR (shoulder[MeSH Terms]) OR (shoulder joint[MeSH Terms])))) AND ((Inverse*[Title/Abstract] OR reverse*[Title/Abstract] OR invert*[Title/Abstract]) AND ((Arthroplast*[Title/Abstract] OR replacement*[Title/Abstract]) OR (Arthroplasty, replacement[MeSH Terms])))) AND (((Fixation*[Title/Abstract] OR fixator*[Title/Abstract] OR plate*[Title/Abstract] OR screw*[Title/Abstract] OR pins[Title/Abstract] OR pin[Title/Abstract] OR wire*[Title/Abstract] OR nail*[Title/Abstract]) OR (Internal fixators[MeSH Terms])) OR (Fracture fixation[MeSH Terms])).

Study selection and data extraction

Obtained references were exported to Endnote X9 (Clarivate, USA) for removal of duplicates. Two independent reviewers screened the study records by title and abstract using Rayyan.[16] Potentially eligible studies were included for full-text reading, and the reason for exclusion after full-text reading was recorded. Any discrepancies that arose in the selection process were resolved by a third reviewer. All of the references were also checked for additional relevant studies. Our workflow was presented with the PRISMA chart.

Selected studies were extracted with a standardized form using Microsoft Excel (Microsoft Corporation, USA). The following data were collected: author, year of publication, country, study design, patient demographic (sex, age), the indication of surgery, pre-operative comorbidity, surgical method, follow-up duration, range of motion (ROM) of the shoulder, Constant-Murley score, complications, revision surgery, author conclusion.

Risk of bias assessment

For randomized controlled trials (RCTs), we assessed study quality with the Cochrane RoB 2 tool.[17] As for non-RCTs, we used the ROBINS-I tool developed by the same group.[18] The visualization of the bias assessment's summary was

generated using the Robvis tool.[19] To ensure the robustness of our present evidence, the meta-analysis was performed using only moderate risk-of-bias or better studies. We planned to conduct a funnel-plot and Egger's test to assess the possibility of reporting bias across studies. However, the power of this test is too low when fewer than ten studies are included.[20]

Synthesis of results

For our primary outcome (ROM of the shoulder), we measured the mean difference (MD) of continuous data with 95% confidence intervals (CIs). As for the secondary outcomes (Constant-Murley score, complications, revision surgery rates), we calculated the risk ratio (RR) for the dichotomous data and MD for continuous data. Statistical analysis was performed using Review Manager (RevMan) 5.3 (The Nordic Cochrane Centre, Denmark).

The presence of heterogeneity was analysed using Cochran's Q and I^2 tests. A low p-value of < 0.1 in Cochran's Q along with $I^2 > 50\%$ showed substantial evidence of heterogeneity.[21] When this occurred, we used a random-effects model. Otherwise, we used the fixed-effects model to analyse the data. To explore the cause of high heterogeneity, we also planned a subgroup analysis based on the study design (RCT and non-RCT).

Evidence quality assessment

To ensure transparency and robustness, we also appraised our evidence quality by using the GRADE approach, as recommended by the Cochrane method. The GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach is utilized to assess the certainty (quality) of the analysed outcomes, comprising five domains: risk of bias of individual study is analysed with RoB 2 and ROBINS-I tools, inconsistency is assessed with a heterogeneity test, indirectness is assessed by determining whether the PICO elements (patients, intervention, comparison, and outcomes) can be easily spotted to answer our review question or not, imprecision is determined by judging whether the sample size is adequate, and other considerations.[22]

Results

Study selection

The electronic search of three main databases (MEDLINE, Embase, CENTRAL) yielded 156 records. After removal of duplicates, two independent reviewers screened 130 records by title and abstract, and thereafter selected 13 studies for full-text article assessment. Seven studies were discarded due to ineligibility (one conference abstract, four studies included two-part fractures in which the data could not be separated, one study used closed reduction internal fixation with humerus block (not ORIF), and one study was irrelevant). Finally, we included six studies in the qualitative synthesis (Fig. 1).

Study characteristics

Of the six selected studies, only one study was an RCT,[23] while the remaining studies were retrospective comparative studies. In total, our present review analysed 324 patients (165 patients in RTSA and 159 patients in ORIF) aged 44 to 91 years old, of whom the majority were female (89.86% in RTSA and 85.71% in ORIF). However, one study did not specify the male/female distribution.[24] Moreover, from two studies that provided the indication of surgery, the data showed that 91.55% of the procedures were indicated due to trauma.[23,25] Only one study provided the patients' pre-operative comorbidities.[23] The details of patients' characteristics are shown in Table 1.

We also recorded the details of surgical methods, including ORIF techniques and RTSA prosthesis types (Table 2). All ORIF procedures were carried out using a locking plate system. Tuberosity reconstruction in ORIF was performed either with sutures (in four studies[23–26]) or wires (in one study[27]). Ockert et al[28] did not specify their reconstruction method. Moreover, cerclage fixation was implemented in two studies,[23,25] intraosseous suture in one study,[26] trans-osseous suture in one study,[24] while two studies did not provide adequate information.[27,28] In case of bone defect problems, Repetto et al utilized allogeneic bone graft,[24] Fraser et al[23] utilized autologous bone graft or bone substitute, while the remaining studies did not provide adequate information.

As for the RTSA group, three studies[24,26,27] used the same prosthesis type, i.e. the cementless SMR™ Reverse modular shoulder system (Lima Corporate, Italy), two studies used a cemented RTSA prosthesis,[23,28] and one study used a hybrid of cementless/cemented prosthesis.[25] Regardless of the different designs, all included studies adopted the Grammont-typical medialized prosthesis principle with 155° inclination (neck-shaft angle) inlay stem, except in two studies that utilized a more lateralized prosthesis design with glenoid lateral offset[25,28] and 150° inclination stem.[25] Moreover, the glenosphere diameter ranged from 36 mm[26] to 44 mm,[27] but three studies did not report the glenosphere diameter.[24,25,28] Tuberosity reconstruction in the RTSA group was also performed in almost all studies, except in a study conducted by Chalmers et al,[25] who did not provide adequate information.

Risk of bias within individual studies

Three reviewers assessed the risk of bias within individual studies with the appropriate tools according to the study design. One RCT was assessed using RoB 2 and considered to have a low risk of bias (Fig. 2a).[23] Whereas, five non-RCT

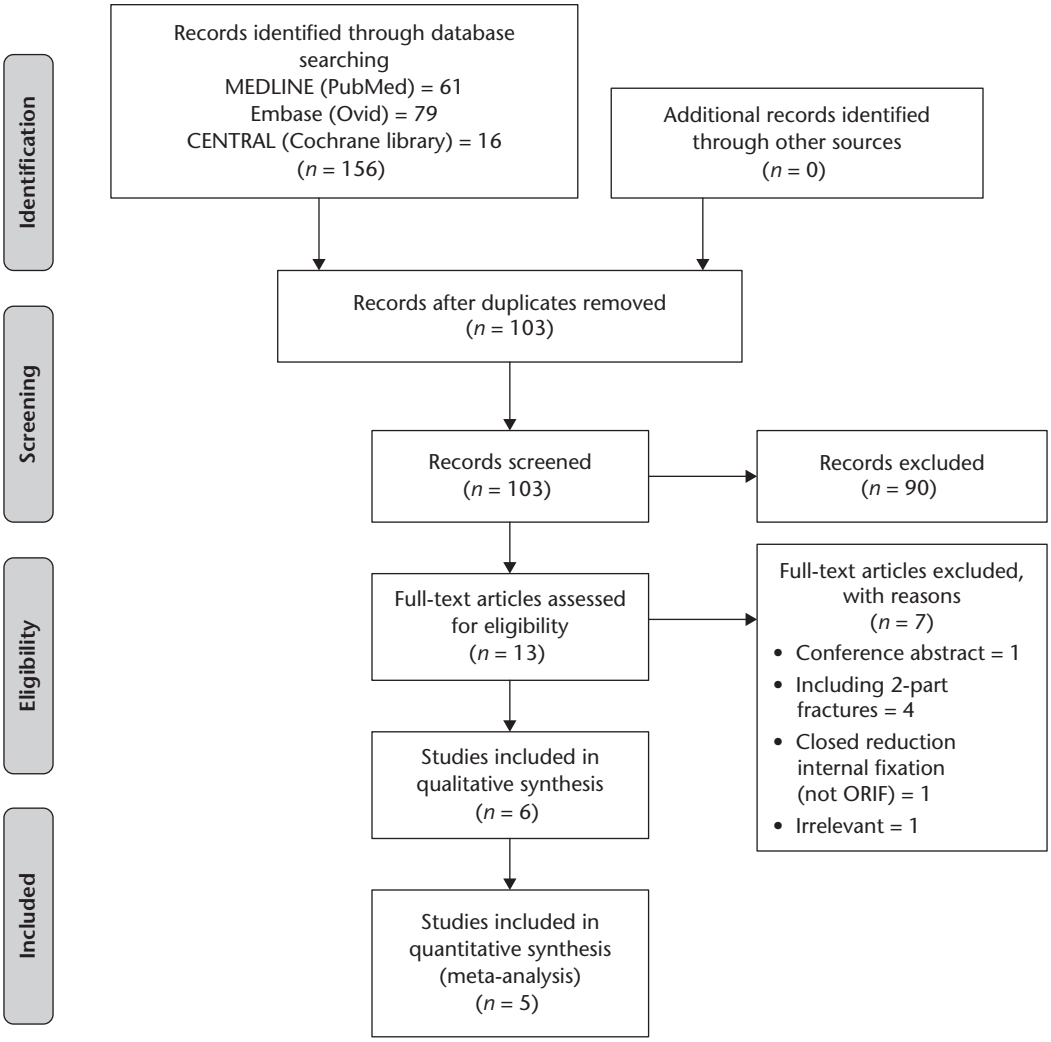


Fig. 1 Study selection.

Note. ORIF, open reduction and internal fixation.

studies were assessed using ROBINS-I (Fig. 2b), of which four studies indicated a moderate risk of bias,[25–28] and one study had a high risk of bias[24] due to confounding. The distinct age gap was determined as a confounding factor due to two reasons: arthroplasty is generally indicated in older patients caused by the concern of prosthesis 'wear and tear' and early revision surgery in younger patients (thus affecting surgical method choice),[29] and younger patients tended to report worse outcomes due to heightened demands and expectations for return to previous level of activity.[30] Repetto et al[24] included patients from 44 years old in the ORIF group, while the RTSA group in this study only included patients 65 years and older. This study carried a serious bias due to the difference in baseline characteristics (confounding). Whereas the remaining studies enrolled patients within the same age range in both groups. The inclusion of studies with a serious risk of bias would alter our estimate of effect, so we excluded this study from the meta-analysis.

Qualitative synthesis

The primary outcome of the present review was the ROM of the shoulder (forward flexion, abduction, external rotation, and internal rotation); whereas, the secondary outcomes were Constant-Murley score (CMS), complications, and revision surgery. The outcomes are reported in Table 3 and Table 4.

Quantitative synthesis (meta-analysis)

We performed meta-analyses for primary (Fig. 3) and secondary outcomes (Fig. 4). One study[23] provided the ROM in Constant score (not in degrees); therefore, it was not

Table 1. Study and patient characteristics

No	Author, publication year	Country	Study design	Patient demographic									Indication of surgery (RTSA/ORIF)	Pre-operative comorbidity		
				RTSA (*n* patients)			ORIF (*n* patients)			Age (years)				Items	RTSA	ORIF
				Total	Male	Female	Total	Male	Female	Range	RTSA	ORIF				
1	Chalmers et al, 2014[25]	USA	Prospective & retrospective comparative study	9	2	7	9	2	7	Over 65	77 ± 6	71 ± 7	Trauma: 9/9	NR	NR	NR
2	Fraser et al, 2020[23]	Norway	Multicentre RCT	64	5	59	60	8	52	65–85	75.7 ± 6.1	74.7 ± 6.5	Trauma: 56/56 Sport-related: 3/2 Not specified: 5/2	Diabetes Smoking ASA	8 2 2.2 ± 0.5	1 4 2.2 ± 0.7
3	Giardella et al, 2017[27]	Italy	Retrospective comparative study	21	3	18	23	7	16	65–91	77.2 ± 6.4	72.1 ± 6.0	NR	NR	NR	NR
4	Luciani et al, 2020[26]	Italy	Retrospective comparative study	22	2	20	26	1	25	65–85	75.5 ± 5.6	73 ± 7.1	NR	NR	NR	NR
5	Ockert et al, 2013[28]	Germany	Retrospective comparative study (matched-paired analysis)	22	2	20	22	2	20	Over 65	77.9 ± 9.1	77.9 ± 9.1	NR	NR	NR	NR
6	Repetto et al, 2017[24]	Italy	Retrospective comparative study	27	NR	NR	19	NR	NR	44–83 for ORIF, 65–80 for RTSA	71.2 ± 7.5	65.3 ± 12.4	NR	NR	NR	NR

Note. RCT, randomized controlled trial; RTSA, reverse total shoulder arthroplasty; ORIF; open reduction and internal fixation; NR, not reported; ASA: American Society of Anesthesiologists scoring system.

calculated in the meta-analysis. Moreover, none of the studies that reported their ROM in degrees stated whether they used a goniometer/inclinometer to measure the ROM, making it a potential bias source. The outcome of the study with a high risk of bias[24] was shown in the graph but was not included in all of the calculations as it would alter our effect estimate.

Primary outcome

The meta-analysis of forward flexion (Fig. 3a) shows that the patients operated with RTSA achieved a significantly better forward flexion compared to those who underwent ORIF (mean 122.23 ± 28.95 vs. 112.67 ± 31.6, MD 9.69; 95% CI [0.69, 18.7], p = 0.03, I^2 = 38%). As for the abduction (Fig. 3b), we found a relatively comparable result between the two modalities (mean 105.67 ± 19.79 vs. 102.73 ± 19.44, MD 2.93; 95% CI [–4.99, 10.85], p = 0.47, I^2 = 0%). In contrast, external rotation was significantly worse in RTSA compared to ORIF (Fig. 3c) (mean 27.46 ± 19.06 vs. 38.49 ± 19.97, MD –12.24; 95% CI [–18.24, –6.24], p < 0.0001, I^2 = 0%).

Secondary outcome

In the RCT study, the CMS was found to be significantly higher in the RTSA group (MD 13.40; 95% CI [5.91, 20.89], p = 0.0005) compared to ORIF (Fig. 4a). Whereas, in non-RCT studies, the two interventions showed no difference in terms of CMS (MD 2.68; 95% CI [–6.63, 11.90], p = 0.57, I^2 = 72%). The overall difference in CMS between the two interventions was not significant (mean 65.96 ± 15.9 vs. 57.53 ± 20.57, MD 5.50; 95% CI [–3.20, 14.20], p = 0.22, I^2 = 78%). However, we found substantial heterogeneity despite the subgroup analysis based on the study design. Therefore, this evidence should be interpreted with caution.

RTSA resulted in significantly more complication events in comparison to ORIF (32.73% vs. 27.04%, risk ratio (RR) 1.42; 95% CI [1.02, 1.98], p = 0.04, I^2 = 44%) (Fig. 4b). The most frequent complications (Table 4) were tuberosity lysis/resorption (15/54) and scapular notching (13/54) in the RTSA group, while in the ORIF group the most common was avascular necrosis (AVN) (17/43). Interestingly, the revision surgery rates were significantly lower in RTSA compared to ORIF (5.45% vs. 14.47%, RR 0.37; 95% CI [0.16, 0.85], p = 0.02, I^2 = 0%) (Fig. 4c). Moreover, the revision surgery in RTSA was mostly component exchange (5/9), while in ORIF it was conversion to RTSA (10/23).

Evidence quality assessment

Table 5 shows the assessment of our evidence quality based on the GRADE approach. Overall, our evidence quality is low, arising from individual study bias and imprecision (low sample size), except for Constant score, which is very low due to high heterogeneity (inconsistency). The low and very low evidence quality implies that our estimate of effect could be affected by further research. Nevertheless, we believe that our present work is the current best evidence.

Table 2. ORIF and RTSA prosthesis characteristics

No	Author, publication year	ORIF	RTSA		
			Prosthesis design	Prosthesis characteristics	Tuberosity reconstruction
1	Chalmers et al, 2014[25]	Anatomically contoured locking proximal humeral plate (Depuy Synthes, USA). Tuberosity reconstruction (+) with cerclage sutures.	Trabecular Metal™ Reverse (Zimmer Biomet, USA).	Glenoid component • Diameter: NI (36 and 40 mm are available) • CoR lateral offset: NI (2.5 and 4.0 mm are available) • Inferior tilt: NI Humeral component • Neck-shaft angle (inclination): 150° • Stem geometry (onlay/inlay): Inlay • Retroversion: NI • Spacer: NI (9 and 12 mm are available) Cementation: proximal (-), distal (+)	NI
2	Fraser et al, 2020[23]	PHILOS angular stable plate (DePuy Synthes, Switzerland). Tuberosity reconstruction (+) with cerclage suture. Bone graft was used when needed (Norian® or autologous bone graft from the iliac crest).	Delta Extend™ Reverse Shoulder System (DePuy Synthes, Johnson & Johnson, UK) for 52 patients and Promos Reverse Prosthesis (Smith & Nephew, US) for 12 patients.	Glenoid component • Diameter: 42 mm • CoR lateral offset: None • Inferior tilt: NI Humeral component • Neck-shaft angle (inclination): 155° • Stem geometry (onlay/inlay): Inlay • Retroversion: NI • Spacer: NI (9 mm is available) Cementation: (+)	Yes, with braided polyester suture-cerclages (no. 5) and wire-cerclages.
3	Giardella et al, 2017[27]	PHILOS angle-stable plate (Depuy Synthes, Switzerland). Tuberosity reconstruction (+) with non-absorbable wires.	SMR™ Reverse modular shoulder system (Lima Corporate, Italy).	Glenoid component • Diameter: 40 or 44 mm • CoR lateral offset: None • Inferior tilt: NI Humeral component • Neck-shaft angle (inclination): 155° • Stem geometry (onlay/inlay): Inlay • Retroversion: 20° • Spacer: NI (9 mm is available) Cementation: (-)	Yes, with no. 2 non-absorbable wires, except in 8 patients (impossible to repair).
4	Luciani et al, 2020[26]	PHILOS plate (Depuy Synthes, USA). Tuberosity reconstruction (+) with intraosseous non-absorbables suture (2/0 Ethibond Suture, Ethicon Inc., USA).	SMR™ Reverse modular shoulder system (Lima Corporate, Italy).	Glenoid component • Diameter: 36 or 40 mm • CoR lateral offset: None • Inferior tilt: 10° Humeral component • Neck-shaft angle (inclination): 155° • Stem geometry (onlay/inlay): Inlay • Retroversion: 0° • Spacer: NI (9 mm is available) Cementation: (-)	Yes, with non-absorbable suture (2/0 Ethibond Suture, Ethicon Inc., USA), except in 5 cases (poor bone stock or excessive bone fragmentation).
5	Ockert et al, 2013[28]	Angle-stable (locked) plate fixation. Tuberosity reconstruction (+).	Aequalis® Reversed Fracture (Tornier©, Germany).	Glenoid component • Diameter: NI (36 and 42 mm are available) • CoR lateral offset: 7 or 10 mm Bio RSA™ (Tornier-Wright, USA) • Inferior tilt: NI Humeral component • Neck-shaft angle (inclination): 155° • Stem geometry (onlay/inlay): Inlay • Retroversion: 10° • Spacer: NI (9 mm is available) Cementation: (+)	Yes, using Fiber Wire® suture no. 5 (Arthrex, USA).
6	Repetto et al, 2017[24]	PHILOS locked plate (DePuy Synthes, Italy). Tuberosity reconstruction (+) with trans-osseous non-absorbable sutures (2/0 Ethibond Suture, Ethicon Inc., USA). Fresh frozen fibular bone allograft was used in case of metaphyseal and calcar defects.	SMR™ Reverse modular shoulder system (Lima Corporate, Italy).	Glenoid component • Diameter: NI (36, 40, and 44 mm are available) • CoR lateral offset: None • Inferior tilt: NI Humeral component • Neck-shaft angle (inclination): 155° • Stem geometry (onlay/inlay): Inlay • Retroversion: NI • Spacer: NI (9 mm is available) Cementation: (-)	Yes, with non-absorbable suture (2/0 Ethibond Suture, Ethicon Inc., USA).

Note. RTSA, reverse total shoulder arthroplasty; ORIF; open reduction and internal fixation; NI, no information; CoR, centre of rotation.

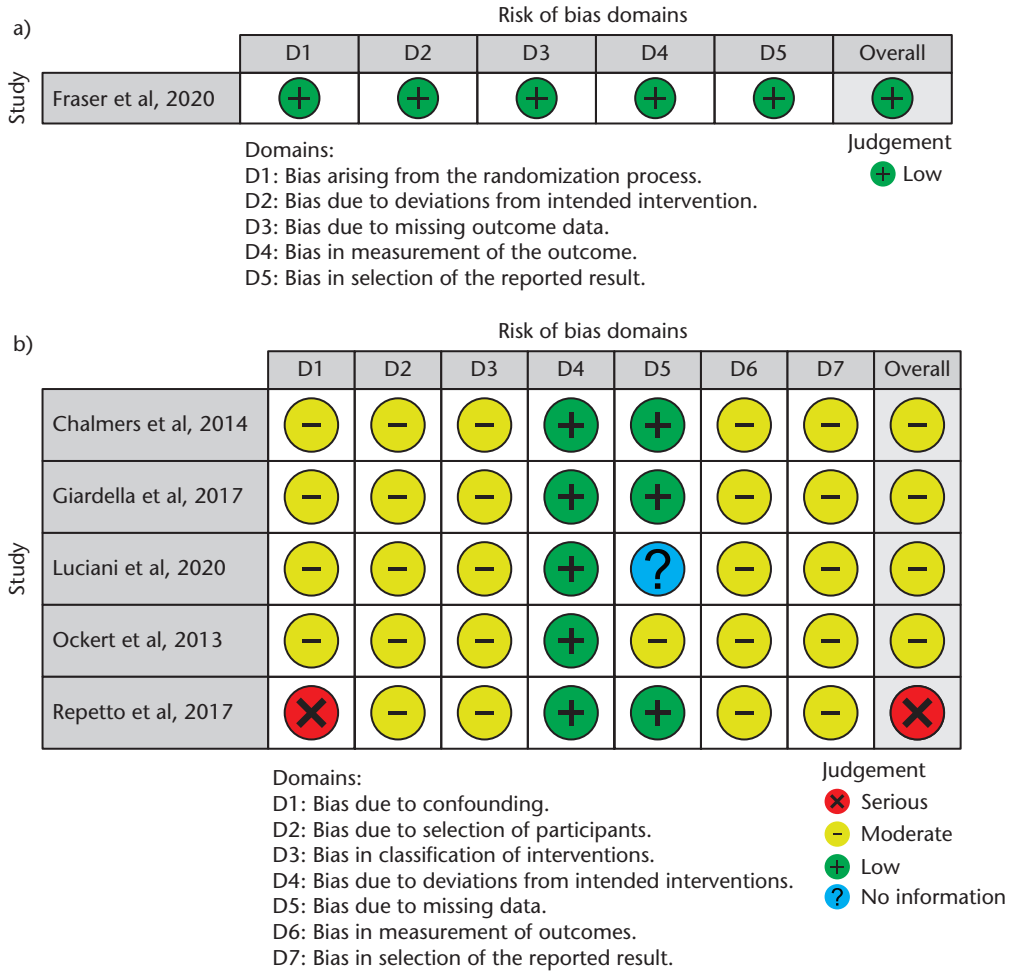


Fig. 2 Risk of bias assessed with (a) RoB 2 tool for randomized controlled trial (RCT) studies and (b) ROBINS-I tool for non-RCT studies.

Table 3. Range of motion (ROM) and Constant-Murley score (CMS) of the included studies

No	Author, publication year	Follow-up (months)		Range of motion/ROM (degrees)								Constant-Murley Score (CMS)	
		RTSA	ORIF	Forward flexion (degrees)		Abduction (degrees)		External rotation (degrees)		Internal rotation (degrees)			
				RTSA	ORIF	RTSA	ORIF	RTSA	ORIF	RTSA	ORIF	RTSA	ORIF
1	Chalmers et al, 2014[25]	14.4 ± 6.0	36 ± 18	133 ± 20	108 ± 40	NR	NR	41 ± 19	46 ± 21	46 ± 15	42 ± 13	NR	NR
2	Fraser et al, 2020[23]*	24		7	5.2	6.7	4.7	7	4.4	5.9	5.7	68.00 ± 17.76	54.60 ± 24.11
3	Giardella et al, 2017[27]	Median 24 (range 12 to 84)	Median 40 (range 12 to 66)	133.3 ± 32.6	112.8 ± 21.0	101.4 ± 19.0	99.6 ± 17.4	35.5 ± 15.6	47.4 ± 20.2	L3 level in 7 patients (33.3%)	SI joint level in 9 patients (39.1%)	65.9 ± 14.7	52.9 ± 17.0
4	Luciani et al, 2020[26]	33.4 ± 10.4	40.0 ± 25.4	124.50 ± 20.45	125.75 ± 33.13	109.75 ± 20.09	105.50 ± 21.02	14.25 ± 13.69	28.00 ± 14.18	L5-S1 (68%)	D7 (52%)	63.65 ± 12.14	65.85 ± 15.73
5	Ockert et al, 2013[28]	12	12	105 ± 29	NR	99 ± 31	NR	22 ± 23	NR	65 ± 26	NR	62.4 ± 14.5	64.3 ± 13.3
6	Repetto et al, 2017[24]	41.7 ± 17.1	36.5 ± 17.9	125.0 ± 44.6	130.6 ± 49.6	109.7 ± 32.1	104.4 ± 21.3	20.3 ± 10.6	23.2 ± 8.8	Gluteus level	L3	58.5 ± 8.5	61.8 ± 14.7

Note. RTSA, reverse total shoulder arthroplasty; ORIF; open reduction and internal fixation; NR, not reported; SI, sacroiliac.

*ROM was reported based on Constant score.

Table 4. Complications, revision surgery, and authors' conclusions of the included studies

No	Author, publication year	Complications (*n* patients and *n* events)		Revision surgery (*n* patients and *n* procedures)		Authors' conclusions
		RTSA	ORIF	RTSA	ORIF	
1	Chalmers et al, 2014[25]	1 patient (complex regional pain syndrome)	1 patient (arthrofibrosis, then in 2 years developed AVN)	0	1 patient (arthroscopic capsular release and subacromial decompression, probably will need RTSA conversion)	There were a significantly greater number of patients who achieved > 90° of active forward elevation after an RTSA procedure ($p = 0.012$). This study also showed evidence of the superiority of RTSA over ORIF in terms of cost savings.
2	Fraser et al, 2020[23]	7 patients (2 nerve injury, 2 deep wound infection, 2 periprosthetic fracture, 1 perioperative glenoid fracture)	11 patients (6 AVN, 2 screw penetration, 1 nonunion, 1 fracture distal to plate, 1 nonunion, 1 rotator cuff rupture)	4 patients (2 components change, 2 other revision surgery)	7 patients (4 procedures of RTSA conversion, 3 implant removal)	At 2 years, the mean Constant score resulted in a significant mean difference of 13.4 points (95% CI, 6.2 to 20.6 points; $p < 0.001$) between the two interventions, which favoured reverse total shoulder arthroplasty (RTSA). RTSA showed an advantage exceeding ORIF for the treatment of displaced OTA/AO type-B2 and C2 proximal humeral fractures in elderly patients.
3	Giardella et al, 2017[27]	5 patients (2 asymptomatic scapular notching, 1 partial lysis of the greater tuberosity, 1 partial lysis of the lesser tuberosity, 1 scapular spur)	4 patients (2 partial reabsorption of the greater tuberosity and 1 partial reabsorption of the lesser tuberosity, 1 partial necrosis of the humeral head)	0	0	Better range of motion (flexion and external rotation) and Constant score in the RTSA group was observed in comparison to the ORIF group (p-values of 0.008, 0.05, and 0.013 respectively). RTSA is one of the foremost treatments in proximal humeral fractures in the elderly patients, whose rotator cuff status is often poor or degenerating.
4	Luciani et al, 2020[26]	14 patients (1 instability, 1 deep infection, 7 scapular notching, 5 tuberosities malunion)	13 patients (5 AVN, 3 loss of reduction, 1 deep infection, 2 subacromial impingement, 2 tuberosities malunion)	2 patients (1 component replacement for instability, 1 two-step revision for deep infection)	9 patients (3 RTSA conversion due to AVN, 2 plate replacement, 2 implant removal, 1 two-step revision for deep infection)	Compared to RTSA patients, ORIF patients had significantly higher mean external rotation (28° versus 14°, $p = 0.0059$) and better results in modal internal rotation (hand at D7 versus L5-S1). However, DASH and Constant scores showed no significant differences. This study also reported a lower revision rate in the RTSA group compared to the ORIF (34.6% versus 9.1%), with avascular necrosis and loss of reduction as the most frequent causes of revision surgery in both groups.
5	Ockert et al, 2013[28]	18 patients (5 tuberosity dislocation > 5 mm, 13 resorption of the tuberosities)	7 patients experienced limited functional outcome due to secondary dislocation, screw cut-out or humeral head necrosis	0	3 patients required secondary/third intervention	The functional result in the patients treated with RTSA for complex multiple-fragment fracture of the proximal humerus after 1 year is comparable to the result after treatment with ORIF (62.4 ± 14.0 and 64.3 ± 13.3, $p = 0.360$). This modality is a suitable alternative for elderly patients, especially in the presence of an extensive rotator cuff tear and/or a humeral head that cannot be reconstructed.
6	Repetto et al, 2017[24]	9 patients (4 clinically silent scapular notching, 2 instability, 1 traumatic periprosthetic fracture, 1 postoperative haematoma, 1 deep infection)	7 patients (4 avascular necrosis, 1 transient circumflex nerve palsy, 2 acromion impingement)	3 patients (1 diaphysis plating for fracture, 1 two-stage revision for deep infection, and 2 component exchange)	3 patients (3 RTSA conversion due to persistent pain and restricted ROM)	The present study shows that locked plating procedure resulted in a significantly higher DASH and SST score compared to RTSA ($p < 0.01$) but not with Constant score ($p > 0.01$). The treatment of complex proximal humerus fracture is nowadays a challenge even for skilled shoulder surgeons. The range of reported complications from 18.2% to 37.5% remains concerning, but most of them did not affect clinical outcome.

Note. RTSA, reverse total shoulder arthroplasty; ORIF; open reduction and internal fixation; AVN, avascular necrosis; ROM, range of motion; DASH, disabilities of the arm, shoulder and hand; SST, simple shoulder test.

Discussion

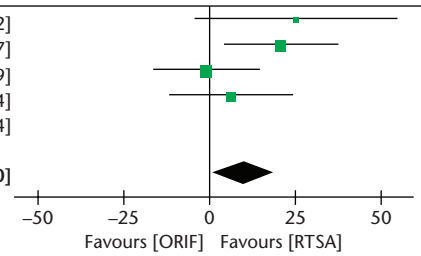
Patient and intervention characteristics

The present systematic review was conducted to compare the outcomes after RTSA and ORIF procedures to treat three- or four-part PHF. Overall, our study included predominantly elderly patients over 65 years old (except in one study[24]) with a reasonably similar male/female ratio across the two groups (1/9 and 1.5/8.5 in RTSA and ORIF, respectively). Higher incidence of complex PHF in elderly

a)

Study or Subgroup	RTSA Mean	RTSA SD	RTSA Total	ORIF Mean	ORIF SD	ORIF Total	Weight	Mean Difference IV, Fixed, 95% CI
Chalmers et al, 2014	133	20	9	108	40	9	9.5%	25.00 [–4.22, 54.22]
Giardella et al, 2017	133.3	32.6	21	112.8	21	23	30.3%	20.50 [4.13, 36.87]
Luciani et al, 2020	124.5	20.45	22	125.75	33.13	26	34.5%	–1.25 [–16.59, 14.09]
Ockert et al, 2013	105	29	22	99	31	22	25.8%	6.00 [–11.74, 23.74]
Repetto et al, 2017	125	44.6	27	130.6	49.6	19	0.0%	–5.60 [–33.54, 22.34]
Total (95% CI)			**74**			**80**	**100.0%**	**9.69 [0.69, 18.70]**

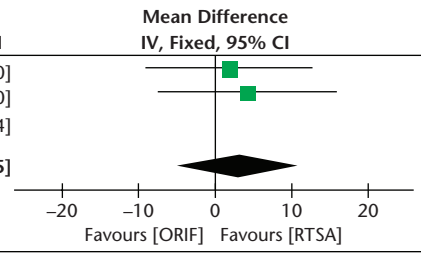
Heterogeneity: Chi² = 4.85, df = 3 (P = 0.18); I² = 38%
Test for overall effect: Z = 2.11 (P = 0.03)



b)

Study or Subgroup	RTSA Mean	RTSA SD	RTSA Total	ORIF Mean	ORIF SD	ORIF Total	Weight	Mean Difference IV, Fixed, 95% CI
Giardella et al, 2017	101.4	19	21	99.6	17.4	23	53.8%	1.80 [–9.00, 12.60]
Luciani et al, 2020	109.75	20.09	22	105.5	21.02	26	46.2%	4.25 [–7.40, 15.90]
Repetto et al, 2017	109.7	32.1	27	104.4	21.3	19	0.0%	5.30 [–10.14, 20.74]
Total (95% CI)			**43**			**49**	**100.0%**	**2.93 [–4.99, 10.85]**

Heterogeneity: Chi² = 0.09, df = 1 (P = 0.76); I² = 0%
Test for overall effect: Z = 0.73 (P = 0.47)



c)

Study or Subgroup	RTSA Mean	RTSA SD	RTSA Total	ORIF Mean	ORIF SD	ORIF Total	Weight	Mean Difference IV, Fixed, 95% CI
Chalmers et al, 2014	41	19	9	46	21	9	10.5%	–5.00 [–23.50, 13.50]
Giardella et al, 2017	35.5	15.6	21	47.4	20.2	23	31.9%	–11.90 [–22.51, –1.29]
Luciani et al, 2020	14.25	13.69	22	28	14.18	26	57.6%	–13.75 [–21.65, –5.85]
Repetto et al, 2017	20.3	10.6	27	23.2	8.8	19	0.0%	–2.90 [–8.53, 2.73]
Total (95% CI)			**52**			**58**	**100.0%**	**–12.24 [–18.24, –6.24]**

Heterogeneity: Chi² = 0.73, df = 2 (P = 0.69); I² = 0%
Test for overall effect: Z = 4.00 (P < 0.0001)

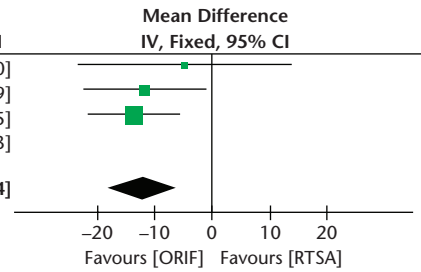


Fig. 3 Meta-analysis of the primary outcome (range of motion): (a) forward flexion, (b) abduction, (c) external rotation.

Note. ORIF, open reduction and internal fixation; RTSA, reverse total shoulder arthroplasty; SD, standard deviation.

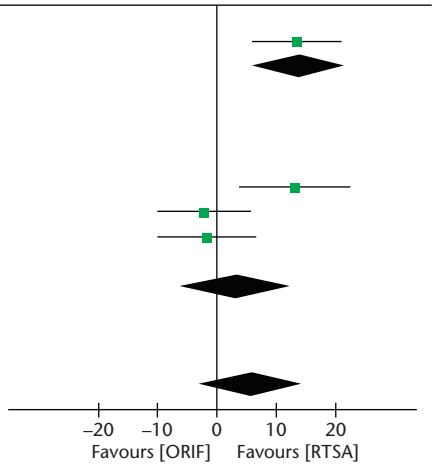
females has been significantly correlated to osteoporosis.[31,32] We also found that most of the PHF were caused by trauma. Although the included study did not specify the type of trauma, previous literature has reported that trauma mechanisms in females differ from those in males significantly, with females being more likely to experience low-energy trauma while male patients were more likely to experience high-energy trauma.[33]

We observed a wide variety of surgical techniques (Table 2), which may potentially induce a bias (not within an individual study but across studies as a whole). However, probably the more relevant question is to what extent this bias affects our estimate of effect. To anticipate this issue, we could look at the heterogeneity test, which assesses the variation across studies (in this case, surgical technique). In all our forest plots, the Cochrane Q test's p-values are all > 0.1 and the I^2 value is < 50% (except in the Constant score forest plot), suggesting that there is no substantial heterogeneity (i.e. the variation in surgical technique might less likely alter our estimate of effect)[21] (except for Constant score). As for the Constant score, we believe that there are other factors (explained in the next section) that may cause bias; thus, the result of Constant score meta-analysis should be interpreted with caution.

Bone graft was used in two studies.[23,24] A previous study that compares ORIF alone and ORIF with fibular allograft to treat three- or four-part PHF in the elderly found that there were significant differences in terms of Constant-Murley, DASH (Disability of Arm-Shoulder-Hand), ASES (American Shoulder and Elbow Surgeons) scores, and complications, favouring the usage of fibular allograft.[34] As we excluded Repetto et al's study from all analyses (due to the high risk of bias arising from the significant age gap), we can ensure that our result is not biased due to this study. As for Fraser et al, it is only included in the secondary outcomes, which may partly explain the heterogeneity seen in Constant score and complication analyses. Thus, the inclusion of the studies that used bone graft[23] may cause bias in our estimate effect in Constant score and complications outcome. However, the bone graft use seems does not affect the revision surgery rate (as seen with the 0% heterogeneity).

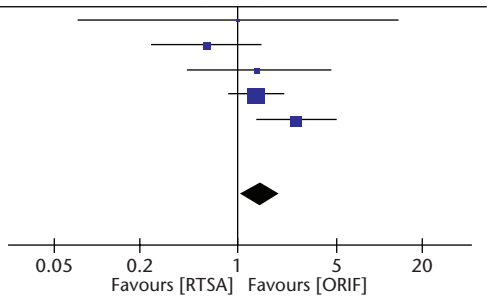
a)

Study or Subgroup	RTSA Mean	RTSA SD	RTSA Total	ORIF Mean	ORIF SD	ORIF Total	Weight	Mean Difference IV, Random, 95% CI
4.1.1 RCT								
Fraser et al, 2020	68	17.76	64	54.6	24.11	60	26.0%	13.40 [5.91, 20.89]
Subtotal (95% CI)			**64**			**60**	**26.0%**	**13.40 [5.91, 20.89]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.50 (P = 0.0005)								
4.1.2 non-RCT								
Giardella et al, 2017	65.9	14.7	21	52.9	17	23	23.5%	13.00 [3.63, 22.37]
Luciani et al, 2020	63.65	12.14	22	65.85	15.73	26	25.5%	–2.20 [–10.09, 5.69]
Ockert et al, 2013	62.4	14.5	22	64.3	13.3	22	25.0%	–1.90 [–10.12, 6.32]
Repetto et al, 2017	58.5	8.5	27	61.8	14.7	19	0.0%	–3.30 [–10.65, 4.05]
Subtotal (95% CI)			**65**			**71**	**74.0%**	**2.68 [–6.63, 11.99]**
Heterogeneity: $Tau^2 = 48.98$; $Chi^2 = 7.25$, df = 2 (P = 0.03); $I^2 = 72\%$								
Test for overall effect: Z = 0.56 (P = 0.57)								
Total (95% CI)			**129**			**131**	**100.0%**	**5.50 [–3.20, 14.20]**
Heterogeneity: $Tau^2 = 61.17$; $Chi^2 = 13.50$, df = 3 (P = 0.004); $I^2 = 78\%$								
Test for overall effect: Z = 1.24 (P = 0.22)								
Test for subgroup differences: $Chi^2 = 3.09$, df = 1 (P = 0.08), $I^2 = 67.6\%$								



b)

Study or Subgroup	RTSA Events	RTSA Total	ORIF Events	ORIF Total	Weight	Risk Ratio IV, Fixed, 95% CI
Chalmers et al, 2014	1	9	1	9	1.6%	1.00 [0.07, 13.64]
Fraser et al, 2020	7	64	11	60	14.3%	0.60 [0.25, 1.44]
Giardella et al, 2017	5	21	4	23	8.0%	1.37 [0.42, 4.43]
Luciani et al, 2020	14	20	13	25	49.3%	1.35 [0.84, 2.16]
Ockert et al, 2013	18	22	7	22	26.8%	2.57 [1.35, 4.89]
Repetto et al, 2017	9	27	7	19	0.0%	0.90 [0.41, 2.00]
Total (95% CI)		**136**		**139**	**100.0%**	**1.42 [1.02, 1.98]**
Total events	45		36			
Heterogeneity: $Chi^2 = 7.13$, df = 4 (P = 0.13); $I^2 = 44\%$						
Test for overall effect: Z = 2.07 (P = 0.04)						



c)

Study or Subgroup	RTSA Events	RTSA Total	ORIF Events	ORIF Total	Weight	Risk Ratio IV, Fixed, 95% CI
Chalmers et al, 2014	0	9	1	9	7.3%	0.33 [0.02, 7.24]
Fraser et al, 2020	4	64	7	60	49.9%	0.54 [0.17, 1.74]
Giardella et al, 2017	0	21	0	23		Not estimable
Luciani et al, 2020	2	20	9	25	34.6%	0.28 [0.07, 1.14]
Ockert et al, 2013	0	22	3	22	8.2%	0.14 [0.01, 2.61]
Repetto et al, 2017	3	27	3	19	0.0%	0.70 [0.16, 3.12]
Total (95% CI)		**136**		**139**	**100.0%**	**0.37 [0.16, 0.85]**
Total events	6		20			
Heterogeneity: $Chi^2 = 0.95$, df = 3 (P = 0.81); $I^2 = 0\%$						
Test for overall effect: Z = 2.34 (P = 0.02)						

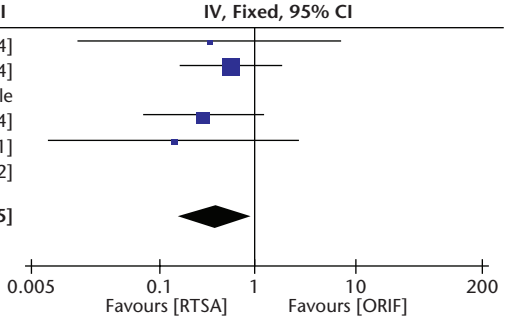


Fig. 4 Meta-analysis of the secondary outcomes: (a) Constant-Murley score (CMS), (b) complications, (c) revision surgery.

Note. ORIF, open reduction and internal fixation; RTSA, reverse total shoulder arthroplasty; RCT, randomized controlled trial; SD, standard deviation.

Moreover, the neck-shaft angle in our included studies is 150° and 155°, which we assume does not make enough lateralization difference to produce any significant differences in clinical evaluation. A recent systematic review of 21 studies comparing tuberosity healing and functional outcomes following RTSA of different inclination prostheses (135°, 145°, and 155°) found that the 135° showed statistically different abduction and tuberosity healing than 145° and 155° (for abduction: average 83°, 105°, and 108°, respectively; for tuberosity healing: average 83%, 69%, and 66%, respectively). As reflected in their average, the significant difference lies between 135° vs. 145° and 135° vs. 155°, but not with 145° vs. 155°. No significant differences were found in forward flexion, external rotation, or post-operative Constant score between groups.[35] Thus, we conclude that our included studies (comprising 150° and 155° inclination) are not biased due to the prosthesis inclination, although a more

Table 5. Evidence quality assessment

Outcomes	Total patients & studies analysed	GRADE assessment					
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Overall certainty of evidence
Forward flexion	154 patients (4 studies)	Serious [a]	Not serious	Not serious	Serious [b]	None	⊕⊕○○ LOW
Abduction	92 patients (2 studies)	Serious [a]	Not serious	Not serious	Serious [b]	None	⊕⊕○○ LOW
External rotation	110 patients (3 studies)	Serious [a]	Not serious	Not serious	Serious [b]	None	⊕⊕○○ LOW
Constant score	260 patients (4 studies)	Serious [a]	Serious [c]	Not serious	Serious [b]	None	⊕○○○ VERY LOW
Complications	275 patients (4 studies)	Serious [a]	Not serious	Not serious	Serious [b]	None	⊕⊕○○ LOW
Revision surgery	275 patients (4 studies)	Serious [a]	Not serious	Not serious	Serious [b]	None	⊕⊕○○ LOW

[a]Moderate risk of bias as assessed with ROBINS-I. [b]Sample size is relatively small (less than 400). [c]Substantial heterogeneity.

GRADE Working Group grades of evidence

High certainty: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate certainty: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low certainty: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low certainty: We are very uncertain about the estimate.

formal investigation (including calculating and comparing the total offset) should be conducted in the future to confirm our argument.

Primary outcome

Our results showed that RTSA significantly improved forward flexion but was equal to ORIF in abduction ($p = 0.03$ and $p = 0.47$, respectively). However, RTSA was inferior in external rotation compared to ORIF ($p < 0.0001$). The better result of RTSA in forward flexion (MD 9.69) is within the range of minimal clinically important difference (MCID) of shoulder ROM for forward flexion (12 ± 4).[36] Moreover, although the included study did not specify their methods of ROM measurement, a previous study showed that minimal detectable change (MDC) of shoulder mobility measurements using a digital inclinometer was 8°.[37] The MDC is crucial to ensure that the measured value was not due to interrater variability or measurement error. Thus, our result shows that the ROM difference observed is most likely consistent and clinically meaningful.

This difference is as expected, in accordance with the biomechanical advantage principle of RTSA shown in the original Grammont prosthesis design. The medialization of the centre of rotation (CoR) to the face of the glenoid (medialized glenoid) along with inlay humeral component (medialized humerus) will increase the deltoid muscle lever arm, thereby increasing the torque force of the deltoid, and finally improving deltoid efficiency to perform shoulder elevation with a relatively lower force needed.[38,39] Hence, it is unsurprising that forward flexion is superior in RTSA compared to ORIF.

Abduction, surprisingly, was not better in RTSA than ORIF. A previous study has claimed that RTSA also improved abduction.[40] Our overall mean data showed that the abduction range of motion was $105.67° \pm 19.79°$ and $102.73° \pm 19.44°$ for RTSA and ORIF, respectively. This range might explain our indifferent finding in abduction between the two interventions, as it suggests the maximal function of the glenohumeral joint and scapulothoracic movement in the frontal plane.[41]

Furthermore, external rotation was significantly higher in ORIF (MD 12.24). The MCID and MDC for external rotation are 3 ± 2[36] and 9 degrees,[37] respectively. Again, this finding implies its consistency and clinical importance. ORIF superiority in external rotation is most likely related to tuberosity reconstruction status. The primary goal of the ORIF procedure is to achieve anatomical healing (by repairing tuberosity), thus restoring external rotation function at the same time. In contrast, it is not evident from the RTSA group's data whether the external rotators are intact or an attempt has been made to repair the major tubercle or healing of the external rotators has been achieved.

In five studies included in this review, the surgical techniques included reconstruction and reattachment of tuberosities when possible (subject to bone stock quality).[23,24,26–28] However, in the study by Chalmers et al, there was no explanation about the tuberosity reconstruction attempts.[25] Literature has shown that better external rotation was reported in patients whose tuberosities were repaired in an RTSA procedure.[42,43] Moreover, a multicentre study comprising 420 PHF patients who were analysed according to their tuberosity healing status following RTSA (Group A: anatomical healing was achieved; Group B: tuberosity resorption, malunion, or nonunion; Group C: tuberosity was excised) showed that Group A achieved a significantly better external rotation and forward flexion.[44]

In addition, rotational movements in the glenohumeral joint have long been a subject of discussion in RTSA due to the reduced lever arm between muscle insertion sites of the remaining rotator cuff after RTSA,[45] in which

prosthesis redesign (glenoid and humeral lateralization)[46–48] as well as additive surgical solutions (latissimus dorsi transfer, latissimus dorsi–teres major transfer) have been proposed.[49–52] As for the internal rotation, our finding shows an equal range of motion between the two surgical modalities. However, it is difficult to give an exact average of the internal rotation as two studies reported this function in degrees,[25,28] and the remaining were in anatomical endpoints.[23,24,26,27]

Secondary outcome

The CMS outcome showed mixed findings. RTSA showed a significantly better CMS outcome in the RCT study, while no significant difference was found in the non-RCT studies ($p = 0.0005$ and $p = 0.57$, respectively). However, the non-RCT studies carried substantial heterogeneity ($I^2 = 72\%$), which affected the overall heterogeneity ($I^2 = 78\%$). Since clinical and methodological diversity always occurs in a meta-analysis, statistical heterogeneity is inevitable.[21] A possible explanation of the high heterogeneity is the subjective domain of CMS, which includes pain. Existing comorbidities such as osteoarthritis in advanced age may cause persistent pain after an ORIF procedure, while in an RTSA, there would be a significant pain reduction. In addition, the strength measurements in CMS should have been stratified by age and sex, since older and female patients may have lower strength.[53] We found that of the 324 patients included, almost two-thirds were female. The studies included in this review mostly did not account for these considerations; only one study performed age-stratified CMS measurement.[23] Hence, the result of the CMS outcome in this review needs to be interpreted with caution.

The patients' risks of developing complications were significantly higher (by 42%) following RTSA compared to ORIF. However, the confounding factor here is the follow-up time and prosthesis design. Complications such as notching of the humeral component can be commonly found after a certain period following RTSA procedures, which is not specific to RTSA in fractures only. Also, scapular notching is related to the prosthesis design, with the Grammont design resulting in more scapular notching than modern designs.[54] We could not detect clearly whether these complications were all clinically significant but, rather, radiographically important (notching, lysis, or malunion of the tubercles). It is also known that advanced age, severe osteoporosis, high-energy trauma, severe comminution of the fracture, and surgical experience represent the most important predicting factors for risk of complications.[55]

However, RTSA significantly reduced the risk of revision surgery by 63% compared to ORIF. Our finding was in accordance with a recent study which reported revision rates of 12.1% and 5.1% for ORIF and RTSA, respectively, after a follow-up duration of 46 months (range 2–10 years).[56] The higher revision rate seen in ORIF compared to the RTSA group is as expected because this technique lends itself to conversion to reversed arthroplasty (a total of ten patients in our findings). Moreover, there is an ongoing debate as to whether scapular notching is a complication or a normal phenomenon in RTSA, especially with the Grammont-type prosthesis. A recent systematic review of 8,258 analysed shoulders found that the Grammont design caused significantly higher scapular notching events compared to all non-Grammont designs combined (42.5% and 12.3%, respectively), with 79.9% of these events being lower-grade scapular notching (grade I or II).[54] Although higher-grade notching was related to glenoid plate instability,[57] reduced survivorship in long follow-up periods (83% at five years, 60% at 10 years, and 43% at 15 years), and increased glenoid loosening risk,[58] the clinical significance of lower-grade notching remains controversial and rarely requires revision.[54,59] This partly explains the distinct complication and revision rates between the two groups, as scapular notching was reported as the second most common RTSA complication (27%). Another reason is that surgeons may be more reluctant to perform revision for RTSA compared to revision for ORIF because the options for RTSA revision are limited, more technically demanding, and give controversial outcomes.[60–63]

Agreement and disagreement with other studies

From the data in this review, we can conclude that RTSA can be recommended for patients aged 65 years or older with a three- or four-part PHF. This is in accordance with the study by Du et al. In their study comparing four treatment modalities (non-operative, ORIF, HA, RTSA), they reported that ORIF was the worst and that RTSA was the best choice in the elderly.[13] Our meta-analysis was also in agreement, with a network meta-analysis comparing RTSA, HA, ORIF, intramedullary nail, and non-operative treatment, which showed that RTSA resulted in the highest probability of improving functional outcome as well as a reduction in total requirements of revision in complex PHF.[12]

Our finding contradicts a recent systematic review and meta-analysis comparing RTSA, HA, and ORIF, which revealed that better outcomes were seen in ORIF procedures compared to HA and RTSA. However, the comparison in the aforementioned study was indirect as they found no head-to-head comparative studies of RTSA and ORIF. Moreover, their results also carried a high risk of bias as the baseline characteristics of the patients included was heterogeneous, with a distinct age gap (mean 61.9 years in the ORIF group and 76.2 years in the RTSA group) and simpler fracture pattern in the ORIF group (two-part fracture) (selection bias).[11] Therefore, the result of that

study should be critically discussed. In contrast, our present review was conducted with a thorough bias analysis, ensuring the robustness of our results.

Strengths and limitations

The present review's limitations are the low number of RCTs included, variation in follow-up time and surgical technique, and limited information in some studies that led to bias. Also, we did not perform a meta-regression analysis to explore possible correlations of age, sex, and follow-up time to the CMS outcome, as conducting meta-regression using very few studies is of questionable value.[21] Further studies should explore this option. Nevertheless, our methodology is robust, as we performed thorough bias analysis and incorporated the GRADE approach to appraise our evidence quality. Thus, we have provided the most recent best evidence of direct/head-to-head comparative studies between RTSA and ORIF.

Conclusion

Compared to ORIF, RTSA resulted in a better forward flexion, comparable abduction, yet less external rotation in the operative treatment of three- or four-part proximal humeral fractures in patients aged 65 and older. Nevertheless, the overall shoulder function assessment using the Constant-Murley score showed no significant difference between the two interventions. Although RTSA was related to increased complications, the revision rate in RTSA was significantly lower than in ORIF. However, our evidence quality may be affected by either individual study bias, relatively small sample size, and/or high heterogeneity; thus, the results remain to be interpreted with caution. Given these facts, the choice of surgical procedure to treat three- or four-part proximal fractures should consider the outcomes, complications, revision rate, as well as the condition and circumstances (i.e. age) of the patients.

AUTHOR INFORMATION

[1]Department of Orthopaedics & Traumatology, Dr. Soetomo General Hospital / Universitas Airlangga, Surabaya, Indonesia.
[2]Institute of Orthopaedics and Musculoskeletal Science, University College London, London, UK.
[3]Cell & Tissue Bank-Regenerative Medicine Centre, Dr. Soetomo General Academic Hospital, Surabaya, Indonesia.
[4]Department of Orthopaedics and Traumatology Zuyderland Medisch Centrum, Heerlen, the Netherlands.
[5]These authors contributed equally to this work.

Correspondence should be sent to: Heri Suroto, Department of Orthopaedics & Traumatology, Dr. Soetomo General Hospital / Universitas Airlangga, Surabaya, Indonesia.
Email: heri-suroto@fk.unair.ac.id

ICMJE CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest in this study.

FUNDING STATEMENT

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

OPEN ACCESS

© 2021 The author(s)
This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) licence (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed.

REFERENCES

1. Court-Brown CM, McQueen MM. The relationship between fractures and increasing age with reference to the proximal humerus. *Curr Orthop* 2002;16:213–222.

2. Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures: *Clinical Orthopaedics and Related Research* (1976–2007). *Clin Orthop Relat Res* 2006;442:87–92.

3. Court-Brown CM, Garg A, McQueen MM. The epidemiology of proximal humeral fractures. *Acta Orthop Scand* 2001;72:365–371.

4. Bengnér U, Johnell O, Redlund-Johnell I. Changes in the incidence of fracture of the upper end of the humerus during a 30-year period: a study of 2125 fractures. *Clin Orthop Relat Res* 1988;231:179–182.

5. Palvanen M, Kannus P, Parkkari J, et al. The injury mechanisms of osteoporotic upper extremity fractures among older adults: a controlled study of 287 consecutive patients and their 108 controls. *Osteoporos Int* 2000;11:822–831.

6. Kannus P, Palvanen M, Niemi S, Parkkari J, Järvinen M, Vuori I. Osteoporotic fractures of the proximal humerus in elderly Finnish persons: sharp increase in 1970–1998 and alarming projections for the new millennium. *Acta Orthop Scand* 2000;71:465–470.

7. Fink HA, Ensrud KE, Nelson DB, et al. Disability after clinical fracture in postmenopausal women with low bone density: the fracture intervention trial (FIT). *Osteoporos Int* 2003;14:69–76.

8. Hodgson S. Proximal humerus fracture rehabilitation. *Clin Orthop Relat Res* 2006;442:131–138.

9. Dey Hazra RO, Blach RM, Ellwein A, Katthagen JC, Lill H, Jensen G. Latest trends in the current treatment of proximal humeral fractures: an analysis of 1162 cases at a level-1 trauma centre with a special focus on shoulder surgery. *Z Orthop Unfall* 2021. doi:10.1055/a-1333-3951 [Epub ahead of print].

10. NJR. *National Joint Registry 17th Annual Report 2020*. https://www.njrcentre.org.uk/njrcentre/Reports-Publications-and-Minutes (date last accessed January 18, 2021).

11. Pizzo RA, Gianakos AL, Haring RS, et al. Are arthroplasty procedures really better in the treatment of complex proximal humerus fractures? A comprehensive meta-analysis and systematic review. *J Orthop Trauma* 2021;35:111–119.

12. Chen L, Xing F, Xiang Z. Effectiveness and safety of interventions for treating adults with displaced proximal humeral fracture: a network meta-analysis and systematic review. *PLoS One* 2016;11:e0166801.

13. Du S, Ye J, Chen H, Li X, Lin Q. Interventions for treating 3- or 4-part proximal humeral fractures in elderly patient: a network meta-analysis of randomized controlled trials. *Int J Surg* 2017;48:240–246.

14. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:b1147.

15. Suroto H, De Vega B, Deapsari F, Prajasari T, Wibowo P, Samijo S. Reverse total shoulder arthroplasty (RTSA) versus open reduction internal fixation (ORIF) for 3-part or 4-part proximal humeral fractures: a systematic review. PROSPERO 2020. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020214681 (date last accessed February 5, 2021).

16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan: a web and mobile app for systematic reviews. *Syst Rev* 2016;5:210.

17. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.

18. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.

19. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 2021;12:55–61.

20. Page M, Higgins J, Sterne J. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane handbook for systematic reviews of interventions version 6.1 (updated September 2020)*. Cochrane, 2020. Available from www.training.cochrane.org/handbook (date last accessed January 18, 2021).

21. Deeks J, Higgins J, Altman D. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane handbook for systematic reviews of interventions version 6.1 (updated September 2020)*. Cochrane, 2020. Available from www.training.cochrane.org/handbook (date last accessed January 18, 2021).

22. Schünemann H, Higgins J, Vist G, et al. Chapter 14: Completing 'summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane handbook for systematic reviews of interventions version 6.1 (updated September 2020)*. Cochrane, 2020. Available from www.training.cochrane.org/handbook (date last accessed January 18, 2021).

23. Fraser AN, Bjørdal J, Wagle TM, et al. Reverse shoulder arthroplasty is superior to plate fixation at 2 years for displaced proximal humeral fractures in the elderly: a multicenter randomized controlled trial. *J Bone Joint Surg [Am]* 2020;102-A:477–485.

24. Repetto I, Alessio-Mazzola M, Cerruti P, Sanguineti F, Formica M, Felli L. Surgical management of complex proximal humeral fractures: pinning, locked plate and arthroplasty: clinical results and functional outcome on retrospective series of patients. *Musculoskelet Surg* 2017;101:153–158.

25. Chalmers PN, Slikker W III, Mall NA, et al. Reverse total shoulder arthroplasty for acute proximal humeral fracture: comparison to open reduction-internal fixation and hemiarthroplasty. *J Shoulder Elbow Surg* 2014;23:197–204.

26. Luciani P, Procaccini R, Rotini M, Pettinari F, Gigante A. Angular stable plate versus reverse shoulder arthroplasty for proximal humeral fractures in elderly patient. *Musculoskelet Surg* 2020;5.

27. Giardella A, Ascione F, Mocchi M, et al. Reverse total shoulder versus angular stable plate treatment for proximal humeral fractures in over 65 years old patients. *Muscles Ligaments Tendons J* 2017;7:271–278.

28. Ockert B, Biermann N, Haasters F, Mutschler W, Braunstein V. [Reverse shoulder arthroplasty for primary fracture treatment: displaced three- and four-part fractures of the proximal humerus in the elderly patient]. *Unfallchirurg* 2013;116:684–690.

29. Charette RS, Sloan M, DeAngelis RD, Lee GC. Higher rate of early revision following primary total knee arthroplasty in patients under age 55: a cautionary tale. *J Arthroplasty* 2019;34:2918–2924.

30. Matthews CJ, Wright TW, Farmer KW, Struk AM, Vasilopoulos T, King JJ. Outcomes of primary reverse total shoulder arthroplasty in patients younger than 65 years old. *J Hand Surg Am* 2019;44:104–111.

31. Taskesen A, Göçer A, Uzel K, Yaradılmış YU. Effect of osteoporosis on proximal humerus fractures. *Geriatr Orthop Surg Rehabil* 2020;11:2151459320985399.

32. Bahrs C, Stojicevic T, Blumenstock G, et al. Trends in epidemiology and patho-anatomical pattern of proximal humeral fractures. *Int Orthop* 2014;38:1697–1704.

33. Passaretti D, Candela V, Sessa P, Gumina S. Epidemiology of proximal humeral fractures: a detailed survey of 711 patients in a metropolitan area. *J Shoulder Elbow Surg* 2017;26:2117–2124.

34. Zhao L, Qi YM, Yang L, et al. Comparison of the effects of proximal humeral internal locking system (PHILOS) alone and PHILOS combined with fibular allograft in the treatment of neer three- or four-part proximal humerus fractures in the elderly. *Orthop Surg* 2019;11:1003–1012.

35. O'Sullivan J, Lädermann A, Parsons BO, et al. A systematic review of tuberosity healing and outcomes following reverse shoulder arthroplasty for fracture according to humeral inclination of the prosthesis. *J Shoulder Elbow Surg* 2020;29:1938–1949.

36. Simovitch RW, Roche CP, Jones RB, et al. Effect of tuberosity healing on clinical outcomes in elderly patients treated with a reverse shoulder arthroplasty for 3- and 4-part proximal humerus fractures. *J Orthop Trauma* 2019;33:e39–e45.

37. Kolber MJ, Vega F Jr, Widmayer K, Cheng MS. The reliability and minimal detectable change of shoulder mobility measurements using a digital inclinometer. *Physiother Theory Pract* 2011;27:176–184.

38. Boileau P, Watkinson DJ, Hatzidakis AM, Balg F. Grammont reverse prosthesis: design, rationale, and biomechanics. *J Shoulder Elbow Surg* 2005;14:147S–161S.

39. Grammont P, Trouilloud P, Laffay J, Deries X. Concept study and realization of a new total shoulder prosthesis. *Rhumatologie* 1987;39:407–418.

40. Gallinet D, Ohl X, Decroocq L, Dib C, Valenti P, Boileau P; French Society for Orthopaedic Surgery (SOFCOT). Is reverse total shoulder arthroplasty more effective than hemiarthroplasty for treating displaced proximal humerus fractures in older adults? A systematic review and meta-analysis. *Orthop Traumatol Surg Res* 2018;104:759–766.

41. Lee KW, Kim YI, Kim HY, Yang DS, Lee GS, Choy WS. Three-dimensional scapular kinematics in patients with reverse total shoulder arthroplasty during arm motion. *Clin Orthop Surg* 2016;8:316–324.

42. Schmalzl J, Jessen M, Holschen M, et al. Tuberosity healing improves functional outcome following primary reverse shoulder arthroplasty for proximal humeral fractures with a 135° prosthesis. *Eur J Orthop Surg Traumatol* 2020;30:909–916.

43. Jain NP, Mannan SS, Dharmarajan R, Rangan A. Tuberosity healing after reverse shoulder arthroplasty for complex proximal humeral fractures in elderly patients: does it improve outcomes? A systematic review and meta-analysis. *J Shoulder Elbow Surg* 2019;28:e78–e91.

44. Ohl X, Bonnevialle N, Gallinet D, et al; SOFCOT. How the greater tuberosity affects clinical outcomes after reverse shoulder arthroplasty for proximal humeral fractures. *J Shoulder Elbow Surg* 2018;27:2139–2144.

45. Herrmann S, König C, Heller M, Perka C, Greiner S. Reverse shoulder arthroplasty leads to significant biomechanical changes in the remaining rotator cuff. *J Orthop Surg Res* 2011;6:42.

46. Werthel JD, Walch G, Vegehan E, Deransart P, Sanchez-Sotelo J, Valenti P. Lateralization in reverse shoulder arthroplasty: a descriptive analysis of different implants in current practice. *Int Orthop* 2019;43:2349–2360.

47. Lädermann A, Tay E, Collin P, et al. Effect of critical shoulder angle, glenoid lateralization, and humeral inclination on range of movement in reverse shoulder arthroplasty. *Bone Joint Res* 2019;8:378–386.

48. Werner BS, Chaoui J, Walch G. The influence of humeral neck shaft angle and glenoid lateralization on range of motion in reverse shoulder arthroplasty. *J Shoulder Elbow Surg* 2017;26:1726–1731.

49. Wey A, Dunn JC, Kusnezov N, Waterman BR, Kilcoyne KG. Improved external rotation with concomitant reverse total shoulder arthroplasty and latissimus dorsi tendon transfer: a systematic review. *J Orthop Surg (Hong Kong)* 2017;25:2309499017718398.

50. Boileau P, Chuinard C, Roussanne Y, Neyton L, Trojani C. Modified latissimus dorsi and teres major transfer through a single delto-pectoral approach for external rotation deficit of the shoulder: as an isolated procedure or with a reverse arthroplasty. *J Shoulder Elbow Surg* 2007;16:671–682.

51. Popescu IA, Bihel T, Henderson D, Martin Becerra J, Agneskirchner J, Lafosse L. Functional improvements in active elevation, external rotation, and internal rotation after reverse total shoulder arthroplasty with isolated latissimus dorsi transfer: surgical technique and midterm follow-up. *J Shoulder Elbow Surg* 2019;28:2356–2363.

52. Ortmaier R, Hitzl W, Matis N, Mattiassich G, Hochreiter J, Resch H. Reverse shoulder arthroplasty combined with latissimus dorsi transfer: a systemic review. *Orthop Traumatol Surg Res* 2017;103:853–859.

53. Thomas M, Dieball O, Busse M. [Normal values of the shoulder strength in dependency on age and gender: comparison with the constant, UCLA, ASES scores and SF36 health survey]. *Z Orthop Ihre Grenzgeb* 2003;141:160–170.

54. Shah SS, Gaal BT, Roche AM, et al. The modern reverse shoulder arthroplasty and an updated systematic review for each complication: part I. *JSES Int* 2020;4:929–943.

55. La Verde L, Franceschetti E, Familiari F, Franceschi F. Complications of ORIF in proximal humeral fractures. In: Huri G, Familiari F, Moon YL, Doral MN, Marcheggiani Muccioli GM, eds. *Shoulder arthroplasty: the shoulder club guide*. Cham: Springer International Publishing, 2020:173–186.

56. Klug A, Wincheringer D, Harth J, Schmidt-Horlohé K, Hoffmann R, Gramlich Y. Complications after surgical treatment of proximal humerus fractures in the elderly: an analysis of complication patterns and risk factors for reverse shoulder arthroplasty and angular-stable plating. *J Shoulder Elbow Surg* 2019;28:1674–1684.

57. Roche CP, Stroud NJ, Martin BL, et al. The impact of scapular notching on reverse shoulder glenoid fixation. *J Shoulder Elbow Surg* 2013;22:963–970.

58. Spiry C, Berhouet J, Agout C, Bacle G, Favard L. Long-term impact of scapular notching after reverse shoulder arthroplasty. *Int Orthop* 2021;45:1559–1566.

59. Zumstein MA, Pinedo M, Old J, Boileau P. Problems, complications, reoperations, and revisions in reverse total shoulder arthroplasty: a systematic review. *J Shoulder Elbow Surg* 2011;20:146–157.

60. Black EM, Roberts SM, Siegel E, Yannopoulos P, Higgins LD, Warner JJ. Failure after reverse total shoulder arthroplasty: what is the success of component revision? *J Shoulder Elbow Surg* 2015;24:1908–1914.

61. Glanzmann MC, Kolling C, Schwyzer HK, Audigé L. Conversion to hemiarthroplasty as a salvage procedure for failed reverse shoulder arthroplasty. *J Shoulder Elbow Surg* 2016;25:1795–1802.

62. Wagner ER, Houdek MT, Hernandez NM, Cofield RH, Sánchez-Sotelo J, Sperling JW. Cement-within-cement technique in revision reverse shoulder arthroplasty. *J Shoulder Elbow Surg* 2017;26:1448–1453.

63. Chalmers PN, Boileau P, Romeo AA, Tashjian RZ. Revision reverse shoulder arthroplasty. *J Am Acad Orthop Surg* 2019;27:426–436.