This updated meta-analysis investigated whether operative treatment is superior to nonoperative treatment in complex proximal humeral fractures. The authors searched the Cochrane Central Register of Controlled Trials, PubMed, and EMBASE. Randomized controlled trials that evaluated operative vs nonoperative treatment for exclusively 3- or 4-part proximal humeral fractures were considered. Six studies with a total of 287 patients who had proximal humeral fractures were included. According to the meta-analysis, no statistically significant differences were found between operative and nonoperative treatment in Constant-Murley shoulder scores (Constant scores); Disabilities of the Arm, Shoulder, and Hand scores; total complication events; mortality; infection; nonunion; avascular necrosis; osteoarthritis; redisplacement of fractures; or dislocation or resorption of tuberosity. For health-related quality of life, EuroQol-5D (EQ-5D) favored operative treatment, but 15D scores showed no significant difference. Compared with nonoperative treatment, open reduction and internal fixation required significantly more additional surgeries (risk ratio, 6.50; 95% confidence interval, 1.54-27.50; \( P = .01 \)), and more penetrations into joint space occurred (risk ratio, 9.56; 95% confidence interval, 2.27-40.13; \( P = .002 \)). The limited evidence suggests that no convincing findings support the use of either open reduction and internal fixation or hemiarthroplasty for the treatment of complex proximal humeral fractures. The findings of the current study should be interpreted cautiously because of the modest sample size and the short follow-up period.
Proximal humeral fracture is a common skeletal injury that comprises 4% to 9% of all fractures of the human body.\(^1\)\(^-\)\(^5\) It is the 3rd most common fracture in patients aged more than 65 years, after hip and wrist fractures. The annual incidence of these fractures is approximately 66 per 10,000 persons, and the incidence is increasing by 15% every year.\(^6\)\(^-\)\(^8\) Fragile bones and a risk of falling are considered the 2 predominant risk factors.\(^8\)

Most proximal humeral fractures are not displaced and can be treated nonoperatively. However, nonoperative treatment has not shown consistently satisfactory results,\(^4\)\(^,\)\(^9\) especially with complex fractures, such as displaced 3- and 4-part proximal humeral fractures classified by Neer,\(^10\) which account for approximately 12.6% of all proximal humeral fractures.\(^11\) In addition to nonoperative treatment, a broad spectrum of operative treatments are available, including open reduction and internal fixation (ORIF) with a variety of devices, hemiarthroplasty, and percutaneous fixation.\(^5\) Technical advances in locking plate fixation have led to an increased interest in ORIF in patients with osteoporosis.\(^4\)\(^,\)\(^12\)\(^,\)\(^13\) According to a recent study, the rates of operative treatment and ORIF increased by 25.6% and 28.5%, respectively, from 1999 to 2005 in the United States; the rates of repeat surgery also increased significantly.\(^14\)

Several earlier systematic reviews summarized the results of treatment for 3- and 4-part proximal humeral fractures.\(^15\)\(^,\)\(^16\) However, they did not improve evidence-based decision making because of lack of randomized controlled trials.\(^15\) Recently, several randomized controlled trials have investigated whether operative treatment may provide greater benefits than nonoperative treatment.\(^17\)\(^-\)\(^22\) These studies had modest sample sizes and yielded conflicting results. Optimal treatment of these fractures remains controversial, and functional outcomes can be disappointing.\(^5\)\(^,\)\(^21\)\(^,\)\(^23\) Therefore, the authors updated previous systematic reviews and performed a meta-analysis of randomized controlled trials to evaluate the efficacy and safety of operative and nonoperative treatment for complex proximal humeral fractures.

**MATERIALS AND METHODS**

**Eligibility Criteria**

Studies were considered acceptable for inclusion in the systematic review if they met the following criteria: (1) displaced 3- and 4-part proximal humeral fractures;\(^15\) (2) operative treatment vs nonoperative treatment; (3) at least 1 of the following outcomes reported: functional scores, such as the Constant-Murley shoulder score (Constant score), Disabilities of the Arm, Shoulder, and Hand (DASH) score, American Shoulder and Elbow Surgeons (ASES) score, health-related quality of life (HRQoL) score, range of motion (ROM), or complications; and (4) randomized controlled trial study design.

Studies were excluded if they (1) were abstracts, letters, or meeting proceedings; (2) had repeated data; or (3) enrolled patients with pathologically or metabolically induced fractures or open fractures.
lists of the included studies were manually searched for potentially eligible studies.

Data Extraction and Quality Assessment

Two reviewers independently extracted relevant data from the included studies, and discrepancies were resolved by a 3rd reviewer. Data included demographic information, study design, interventions, allocation concealment process, blinding, time to last follow-up, rate of follow-up, intention to treat, functional scores, HRQoL score, ROM, and complications. Some corresponding authors were contacted by e-mail to obtain missing information.

The risk of bias of each eligible study was independently assessed by 2 authors, in accordance with the Cochrane risk of bias tool, which defines 9 aspects: (1) random sequence generation (selection bias); (2) allocation concealment (selection bias); (3) blinding of participants (performance bias); (4) blinding of treatment providers (performance bias); (5) blinding of outcome assessors (detection bias); (6) intention to treat (attrition bias); (7) selective reporting (reporting bias); (8) comparable study groups; and (9) other bias. A qualification of risk of bias, including low risk, unclear risk, or high risk, was provided. The final qualification for each study was determined by consensus among 3 authors.

Statistical Analysis

Statistical analysis was performed with Review Manager 5.1 software (Cochrane Collaboration, Oxford, United Kingdom) for outcome measures. \( P<.05 \) was considered statistically significant. Heterogeneity was evaluated by visual inspection of the forest plot along with consideration of the test for heterogeneity and the \( I^2 \) statistic. \( I^2 \) value of greater than 50% was considered substantial heterogeneity. A fixed effects model was used in the meta-analysis unless there was significant heterogeneity among studies. The random effects model of DerSimonian and Laird was used, regardless of heterogeneity. Continuous variables were presented as mean difference (MD) and dichotomous variables were presented as risk ratio (RR), both with a 95% confidence interval (CI). When the RR of outcomes was statistically significant, the authors calculated and reported numbers needed to treat or numbers needed to harm (NNH). To define sources of heterogeneity, subgroup analyses based on ORIF vs hemiarthroplasty were defined during the analysis design phase. Publication bias was tested by funnel plots when possible.

Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Operative Treatment</th>
<th>Nonoperative Treatment</th>
<th>Mean age, y (O vs N) (O vs N)</th>
<th>% Female</th>
<th>Follow-up, mo</th>
<th>Rate of Follow-up</th>
<th>Diagnosis Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boons et al 2012</td>
<td>Global 1 FX shoulder fracture endoprosthesis (DePuy, Leeds, UK)</td>
<td>Shoulder immobilizer</td>
<td>79.9 (7.7) vs 76.4 (5.6)</td>
<td>25 vs 25</td>
<td>92 vs 96</td>
<td>24</td>
<td>Displaced 4-part fractures</td>
</tr>
<tr>
<td>Fjalestad et al 2010</td>
<td>Angular stable locking plates (nonspecific LCPs of the AO basic type; Synthes, Bettlach/Solothurn, Switzerland)</td>
<td>Modified Velpeau bandages</td>
<td>72.2 (60-86) vs 73.1 (60-88)</td>
<td>25 vs 25</td>
<td>80 vs 94</td>
<td>12</td>
<td>Displaced 3- or 4-part fractures of OTA 11-B2 or 11-C2</td>
</tr>
<tr>
<td>Olerud et al 2011</td>
<td>Open reduction and locking plates (Philos plates)</td>
<td>Slings</td>
<td>72.9 (56-92) vs 74.9 (58-88)</td>
<td>30 vs 30</td>
<td>80 vs 83</td>
<td>24</td>
<td>Displaced 3-part fractures</td>
</tr>
<tr>
<td>Olerud et al 2011</td>
<td>Global FX prosthesis (DePuy, Sollentuna, Sweden)</td>
<td>Slings</td>
<td>75.8 (58-90) vs 77.5 (60-92)</td>
<td>27 vs 28</td>
<td>85 vs 86</td>
<td>24</td>
<td>Displaced 4-part fractures</td>
</tr>
<tr>
<td>Stableforth 1984</td>
<td>Neer prosthesis, uncemented</td>
<td>Closed manipulation</td>
<td>65.6 (52-88) vs 70.1 (60-85)</td>
<td>16 vs 16</td>
<td>75 vs 81.3</td>
<td>6-48</td>
<td>Displaced 4-part fractures</td>
</tr>
<tr>
<td>Zyto et al 1997</td>
<td>Cerclage wiring or surgical tension band</td>
<td>Slings</td>
<td>73 (7.5) vs 75 (6.7)</td>
<td>20 vs 20</td>
<td>90 vs 85</td>
<td>36-60</td>
<td>Displaced 3- or 4-part fractures</td>
</tr>
</tbody>
</table>

*Abbreviations: LCPs, locking compression plates; O vs N, operative vs nonoperative treatment; OTA, Orthopaedic Trauma Association.*
Study Selection and Characteristics

The initial search identified 422 references. After duplicate references were removed and the titles, abstracts, and contents of the full text were examined, 6 randomized controlled trials were included in the meta-analysis. Figure 1 shows the process of literature selection. Table 1 shows the characteristics and demographic data of each included study. A total of 287 patients with complex humeral fractures were included in this study. The 6 randomized controlled trials were performed in the Netherlands, Norway, Sweden, and the United Kingdom during the period between 1984 and 2012. The total number of patients in each study ranged from 32 to 60. Mean age ranged from 65.5 to 79.9 years. The percentage of female patients in the study populations ranged from 75% to 96%. The studies followed patients for periods of 6 to 60 months, and the rate of patient follow-up ranged from 72.5% to 98%. Two studies included displaced 3- or 4-part fractures. Three studies included only displaced proximal humeral 4-part fractures, and the remaining study included only displaced 3-part fractures. The studies investigated the following implants: Global 1 FX shoulder fracture endoprosthesis (DePuy, Leeds, United Kingdom); angular stable locking plates (nonspecific locking compression plates of the AO basic type; Synthes, Bettlach/Solothurn, Switzerland); locking plates (Philos plates); GlobalFx prostheses (DePuy, Sollentuna, Sweden); Neer prostheses, uncemented; and cerclage wiring or surgical tension band.

Study Quality

Figure 2 shows the methodologic quality of the randomized controlled trials as independently assessed by 2 authors. All of the studies were single-center studies. Four of the studies were judged as having used sufficient allocation concealment. One study used closed envelopes without reporting adequate safeguards. Only 1 study used closed envelopes without reporting adequate safeguards. Four studies reported a proper intention-to-treat analysis and clearly stated interventions. The comparability of baseline characteristics was generally acceptable.

Outcomes

There were no statistically significant differences in Constant scores, DASH scores, or ASES scores between operative and nonoperative treatment at either 12 or 24 months. For HRQoL, operative treatment yielded significantly better EuroQol-5D (EQ-5D) scores at both 12 and 24 months of follow-up compared with nonoperative treatment. However, no statistically significant difference was detected for 15D (Table 2).

Meta-analysis was considered inappropriate for ROM because of the obvious clinical heterogeneity and considerable variations in measurement of results. ROM did not differ significantly between operative and nonoperative treatment in the randomized controlled trials that examined this outcome, except that in

---

**Table 1**

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation (selection bias)</th>
<th>Blinding of participants (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Intention to treat (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Comparable study groups</th>
<th>Other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boons 2012</td>
<td>+</td>
<td>+</td>
<td>★</td>
<td>★</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fjalestad 2010</td>
<td>+</td>
<td>+</td>
<td>★</td>
<td>★</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Olerud (HA) 2011</td>
<td>+</td>
<td>+</td>
<td>★</td>
<td>★</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Olerud (ORIF) 2011</td>
<td>+</td>
<td>+</td>
<td>★</td>
<td>★</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
</tr>
</tbody>
</table>

**Figure 2:** Risk of bias summary. Abbreviations: +, low risk; ?, unclear risk; -, high risk; HA, hemiarthroplasty; ORIF, open reduction and internal fixation.
1 study with a small sample size\textsuperscript{22} from 1984 operative treatment with hemiarthroplasty was favored. \textbf{Table 3} shows the results for ROM.

The 6 randomized controlled trials\textsuperscript{17-22} included in this meta-analysis included a total of 287 fractures, and they all provided information on total complication events after at least 1 year of follow-up. Meta-analysis of total complication events showed no significant difference between operative and nonoperative treatment (RR, 1.81; 95% CI, 0.87-3.76; \( P=0.11 \)).

Because of the significant heterogeneity (\( P<0.0001; I^2=85\% \)), subgroup analysis of ORIF vs hemiarthroplasty was performed. Subgroup analysis did not show specific sources of heterogeneity when divided by type of intervention. Therefore, the random effects model was used (\textbf{Figure 3}).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|}
\hline
\textbf{Outcome} & \textbf{No. of Trials} & \textbf{No. of Patients (O vs N)} & \textbf{MD (95\% CI)} & \textbf{\( P \) for MD} & \textbf{I\(^2 \)} & \textbf{\( P \) for Heterogeneity} \\
\hline
Constant scores at 12 mo & 4\textsuperscript{18-21} & 200 (99/101) & 2.92 (-2.07 to 7.92) & 0.25 & 0\% & 0.93 \\
Constant scores after 24 mo & 3\textsuperscript{17,19,21} & 140 (70/70) & -1.36 (-7.64 to 4.92) & 0.67 & 0\% & 0.64 \\
DASH scores at 12 mo & 2\textsuperscript{19,21} & 105 (53/52) & -4.5 (-13.5 to 4.48) & 0.33 & 0\% & 0.74 \\
DASH scores at 24 mo & 2\textsuperscript{19,21} & 102 (51/51) & -7.48 (-16.1 to 1.15) & 0.09 & 0\% & 0.79 \\
ASES score & 1\textsuperscript{18} & 58 (23/25) & 0.20 (-0.71 to 1.11) & 0.67 & NA & NA \\
EQ-5D at 12 mo & 2\textsuperscript{19,21} & 105 (53/52) & 0.13 (0.04 to 0.22) & 0.004 & 0\% & 0.45 \\
EQ-5D at 24 mo & 2\textsuperscript{19,21} & 102 (51/51) & 0.15 (0.05 to 0.24) & 0.004 & 0\% & 0.65 \\
15D at 12 mo & 1\textsuperscript{18} & 58 (23/25) & 0.02 (-0.03 to 0.08) & 0.42 & NA & NA \\
\hline
\end{tabular}
\caption{Functional Status and Health-related Quality of Life}
\end{table}

Meta-analysis of the rate of additional surgical event (NNH, 8.7; 95% CI, 5.4-22.6). Furthermore, subgroup analysis according to the type of surgery showed that only the ORIF subgroup was significantly worse compared with nonoperative treatment (RR, 6.50; 95% CI, 1.54-27.50; \( P=0.01; I^2=0\% \)). The NNH was 5 (95% CI, 3.10-12.94). The hemiarthroplasty subgroup did not differ significantly from nonoperative treatment (RR, 2.24; 95% CI, 0.52-9.60; \( P=0.28; I^2=0\% \); \textbf{Figure 4}).

Four studies\textsuperscript{17-19,21} that included 205 fractures provided data on the incidence of penetration of the implant into joint spaces. The rate of penetration events was significantly higher in the operative group compared with the nonoperative group (RR, 9.56; 95% CI, 2.27-40.13; \( P=0.002; I^2=0\% \)). Given the penetration rate of 0% with nonoperative treatment, 6 fractures would have to be treated operatively to sustain 1 additional penetration event (NNH, 6; 95% CI, 4.18-10.6; \textbf{Figure 5}). All of the penetration events occurred in patients who underwent ORIF; 1 patient received a locking plate after randomization to the hemiarthroplasty group.\textsuperscript{21}

Five randomized controlled trials\textsuperscript{17-21} that reported a total of 255 fractures provided information on the rates of avascular necrosis. Meta-analysis of the rates of avascular necrosis found no significant difference between operative and nonoperative treatment (RR, 0.65; 95% CI, 0.36-1.19; \( P=0.16; I^2=58.6\% \)). The authors then performed subgroup analysis of ORIF vs shoulder hemiarthroplasty over nonoperative treatment, no statistically significant difference was detected (RR, 0.17; 95% CI, 0.02-1.38; \( P=0.1; I^2=0\% \)). Furthermore, ORIF did not significantly differ from nonoperative treatment (RR, 0.74; 95% CI, 0.39-1.37; \( P=0.37; I^2=0\% \); \textbf{Figure 6}).

No statistically significant difference was seen in the rates of mortality, infection, nonunion, osteoarthritis, redisplacement of fracture, or dislocation or resorption of tuberosity between operative and nonoperative treatment (\textbf{Table 4}).

\textbf{DISCUSSION}

The authors believe that this report is the first systematic review and meta-analysis of randomized controlled trials to compare operative and nonoperative treatment in complex proximal humeral fractures. Meta-analysis of randomized controlled trials is generally considered to provide the highest level of evidence for
Therefore, this report provides an update of previous systematic reviews focusing on 3- and 4-part fractures, which were limited because of lack of randomized controlled trials.\(^{15,16}\)

Meta-analysis yielded the following findings:

1. There were no statistically significant differences in Constant scores, DASH scores, ASES scores, or the rates of total complications, mortality, infection, nonunion, avascular necrosis, osteoarthritis, redisplacement of fracture, and dislocation or resorption of tuberosity between operative and nonoperative treatment.\(^{29}\)

2. Regarding HRQoL, operative treatment yielded significantly better EQ-5D scores compared with nonoperative treatment, but no statistically significant difference was seen for 15D.\(^{29}\)

3. For ROM, only 1 study from 1984 with a small sample size showed no statistically significant difference; the other studies showed no statistically significant differences in mean (range) values for forward flexion, abduction, and external rotation.\(^{29}\)

4. ORIF was associated with significantly more additional surgery and more penetration of implants into joint space compared with nonoperative treatment.\(^{29}\)

Table 3

<table>
<thead>
<tr>
<th>Outcome for ROM</th>
<th>Follow-up, mo</th>
<th>Boons et al(^{12})</th>
<th>Fjalestad et al(^{18})</th>
<th>Olerud et al(^{19})</th>
<th>Olerud et al(^{21})</th>
<th>Stableforth(^{22})</th>
</tr>
</thead>
<tbody>
<tr>
<td>CROM (O vs N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>12</td>
<td>20 (8.3) vs 18 (6.9)</td>
<td>17.0 vs 18.2</td>
<td>27.0 (8.7) vs 24.1 (7.1) (P=.11)</td>
<td>21.2 (7.2) vs 19.2 (7.8) (P=.46)</td>
<td>NA</td>
</tr>
<tr>
<td>Forward flexion (O vs N)</td>
<td>12</td>
<td>98 (45-165) vs 94 (45-165) (P=.86)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mean (range), °</td>
<td>24</td>
<td>120 vs 111 (P=.36)</td>
<td>93 vs 95 (P=.85)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>&gt;90°, No.</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>&lt;45°, No.</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Abduction (O vs N)</td>
<td>12</td>
<td>77 (45-165) vs 87 (30-130) (P=.36)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mean, °</td>
<td>24</td>
<td>NA</td>
<td>NA</td>
<td>114 vs 106 (P=.28)</td>
<td>86 vs 87 (P=.89)</td>
<td>NA</td>
</tr>
<tr>
<td>External rotation (O vs N)</td>
<td>12</td>
<td>17 (10-25) vs 19 (15-25) (P=.10)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mean (range), °</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>&gt;25°, No.</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>&lt;5°, No.</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Internal rotation (O vs N)</td>
<td>12</td>
<td>L3 vs L3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Thumb to lumbar level</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>No. thumb to T12 spinous process</td>
<td>48</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: CROM: Constant score ROM subscale (maximum, 40); NA, not available; O vs N, operative vs nonoperative treatment; ROM, range of motion.
A previous systematic review selected only studies dealing exclusively with 3- and 4-part fractures of the proximal humerus and found that nonoperative treatment had a poorer ROM than operative treatment. However, the authors indicated that their data from former published literature were inadequate for evidence-based decision making because of the lack of randomized controlled trials. Another systematic review including 3- and 4-part fractures of the proximal humerus also compared hemiarthroplasty and nonoperative treatment. The authors found that the mean Constant score in the nonoperative group was higher than in the hemiarthroplasty group, and they only included 1 randomized controlled trial. As an update, the current meta-analysis included 6 randomized controlled trials focusing on complex proximal humeral fractures.

In this meta-analysis, participants were not restricted to the elderly, but patients selected for trials were generally older than 52 years. Additionally, many surgical options were available, and they were separated into 2 subgroups: ORIF and hemiarthroplasty. Over the past decade, locking plate fixation to treat proximal humeral fractures has gained considerable popularity. In this meta-analysis, 2 recent randomized controlled trials used surgical treatment with locking plates.

The flaws of methodologic quality in the included randomized controlled trials were mainly in the area of blinding. However, this risk of bias was not considered critical because of the difficulty of performing blinding in surgical research. Inadequate allocation concealment leads to overestimation of treatment, whereas adequate allocation concealment can avoid selection bias. Fortunately, in this meta-analysis, 4 high-quality randomized controlled trials reported adequate allocation concealment, and they were published in the 2010s. This suggests an increasing awareness of the importance of high-quality randomized controlled trial design.

Regarding functional status, analyses of Constant scores, DASH scores, and ASES scores showed no statistically significant differences. These results contradict a previous systematic review. However, the current meta-analysis differs from the previously published review because it is based on randomized controlled trials and did not detect any significant heterogeneity ($I^2=0$).

In terms of HRQoL, the results of meta-analysis of EQ-5D showed that...
operative treatment led to better scores than nonoperative treatment. The latest diagnostic study reported that EQ-5D had good internal and external responsiveness in patients with proximal humeral fractures. Additionally, other recent studies of proximal fractures reported EQ-5D scores for HRQoL. However, other studies have suggested that 15D has better responsiveness than EQ-5D because 15D allows assessors to detect small changes. One randomized controlled trial that was included reported that the 15D data showed no differences between operative and nonoperative treatment.

Therefore, it is impossible to draw a clear conclusion about optimal treatment to yield a better HRQoL.

Regarding ROM, a meta-analysis could not be performed because of the variety of types of data. As a qualitative review, no clinical conclusions can be reached that are in agreement with previous systematic reviews. Another systematic review of hemiarthroplasty reported obvious functional limitations and ROM restrictions at 3.7 years of follow-up. However, the follow-up of most randomized controlled trials in the current study was less than 2 years. Thus, the results are inadequate for interpreting long-term ROM.

A recent systematic review of locking plates that included proximal humeral intra-articular fractures showed marked complication and additional surgery rates; this finding was also supported by the previous 2 systematic reviews. A higher additional surgery rate was confirmed in patients who underwent ORIF compared with those who received nonoperative treatment (23.6% vs 3.2%). This rate was also higher than in the previous systematic review (23.6% vs 14%).

No statistically significant differences were detected in rates of total complications, mortality, avascular necrosis, infection, nonunion, osteoarthritis, redisplacement of fracture, or dislocation or resorption of tuberosity. Infection was not clearly more prevalent in the surgical group. There were 2 possible explanations. One was that the borderline result may be caused by the limited study power, and the key weakness was the small sample size. Therefore, this result cannot be used as strong evidence. The other explanation was that the 2 randomized controlled trials that reported infection were performed before 2000, whereas the 3 randomized controlled trials that reported no infection were all performed after 2000.

Table 4

Other Complications Reported

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of Trials</th>
<th>No. of Patients</th>
<th>Operative Group</th>
<th>Nonoperative Group</th>
<th>RR (95% CI)</th>
<th>P for RR</th>
<th>I²</th>
<th>P for Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>2,18,22</td>
<td>82</td>
<td>3/41</td>
<td>1/41</td>
<td>2.33 (0.36-15.14)</td>
<td>.43</td>
<td>0%</td>
<td>.37</td>
</tr>
<tr>
<td>Infection</td>
<td>3,17,19,22</td>
<td>132</td>
<td>6/66</td>
<td>0/66</td>
<td>5.0 (0.9-27.89)</td>
<td>.07</td>
<td>0%</td>
<td>1</td>
</tr>
<tr>
<td>Nonunion</td>
<td>5,17-21</td>
<td>253</td>
<td>4/125</td>
<td>7/128</td>
<td>0.66 (0.23-1.88)</td>
<td>.44</td>
<td>0%</td>
<td>.79</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>2,17,21</td>
<td>95</td>
<td>2/47</td>
<td>7/48</td>
<td>0.39 (0.03-4.27)</td>
<td>.44</td>
<td>52%</td>
<td>.15</td>
</tr>
<tr>
<td>Redisplacement of fracture</td>
<td>3,18,20,21</td>
<td>155</td>
<td>1/77</td>
<td>3/78</td>
<td>0.51 (0.1-2.68)</td>
<td>.42</td>
<td>0%</td>
<td>.82</td>
</tr>
<tr>
<td>Dislocation or resorption of tuberosity</td>
<td>4,17,19,21</td>
<td>205</td>
<td>21/102</td>
<td>3/103</td>
<td>4.65 (0.49-44.58)</td>
<td>.06</td>
<td>60%</td>
<td>.18</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; RR, risk ratio.
published after 2010, which may suggest advances in surgical technique, operative condition, and perioperative management to prevent infection.

Clear heterogeneity was shown for total complications, the total complications subgroup (ORIF vs NO), osteoarthritis, and dislocation or resorption of tuberosity. This heterogeneity probably originated from conflicting data among trials. It is clear that there was clinical bias as well because locking plate fixation and tension bands were combined into the ORIF group. Subgroup analysis within the ORIF group was not conducted because of the small number of randomized controlled trials.

This analysis has certain limitations. First, only 6 randomized controlled trials, with a total of 287 fractures, could be included. Obviously, the sample size was modest. Publication bias analysis could not be conducted because of the small number of included trials. Second, there is no entirely reliable classification of proximal humeral fractures. The Neer classification has poor inter- and intraobserver reliability. For complex proximal humeral fractures, however, 3- and 4-part fractures classified by Neer were widely used in clinical trials and previous systematic reviews. 15-22 28 No other classification is an acceptable alternative to represent the complex proximal humeral fractures in a systematic review. Nonetheless, unreliable classification of the fractures may bias the results. Third, the follow-up of most randomized controlled trials in the current study was less than 2 years. Therefore, further large long-term follow-up randomized controlled trials of sound methodologic quality should be conducted and more reliable classification of fractures should be used. Fortunately, such randomized controlled trials have already been designed. 41-43

**CONCLUSION**

Current limited evidence suggests that no convincing findings support the use of either ORIF or hemiarthroplasty to treat complex proximal humeral fractures. The current findings should be interpreted cautiously because of the modest sample size and short-term follow-up. Future randomized controlled trials should be large, with a longer follow-up period, and they should be of sound methodologic quality and use a reliable classification of fractures.

**REFERENCES**


