Total Elbow Arthroplasty Versus Plate Fixation for Distal Humeral Fractures in Elderly Patients: A Systematic Review and Meta-Analysis
Introduction

Distal humerus fractures represent about 1-2% of adult fractures and about 10% of humeral fractures. [1] The injury has a bimodal distribution with a peak incidence in young males secondary to high energy trauma and a second peak in osteoporotic elderly (typically female) patients over the age of 60 years. [2] It is predicted that the annual rate of distal humeral fractures in the elderly population is likely to triple by 2030 due to an increasingly aging population structure. [3] These fractures typically require surgical treatment as non-operative treatment is associated with a high frequency of complications such as non-union, malunion, stiffness and pain - any of which can lead to severe functional deficit and a subsequent loss of independence due to an inability to perform activities of daily living. [4-9] It is for this reason that non-operative management is typically only advocated for those patients who are unfit for anaesthesia and surgery. [10]

The AO classification of distal humeral fractures defines Type C injuries as comminuted intra-articular fractures [11] and these injuries provide a significant surgical challenge. The choice of surgical intervention is controversial and forms the basis of this meta-analysis. Surgical options include open reduction and internal fixation (ORIF), which in the contemporary literature is most frequently performed with either orthogonal or parallel plating according to AO principles. [12, 13] However, osteoporotic bone and highly comminuted fracture patterns often preclude anatomic reduction and early mobilisation, and predispose to failure of fixation, revision surgery, stiffness and a high rate of functional limitation (though still preferable to non-operative treatment). An alternative surgical option is total elbow arthroplasty (TEA). However, this strategy is also associated with a risk of complications such as infection, dislocation, peri-prosthetic fracture, nerve palsy, skin complications and revision for other causes. [14-16]
Although the literature has previously compared ORIF versus TEA there remains a lack of consensus regarding the optimum treatment choice. The aim of this meta-analysis was to determine which procedure provided superior clinical outcomes for elderly patients with distal humeral fractures.

**Methods**

A systematic review of the literature was conducted in accordance with the PRISMA guidelines [17] using the online databases Medline and EMBASE. The searches were performed independently by two authors on the 1st September 2016 and repeated on the 12th September 2016 to ensure accuracy. The Medline search strategy is illustrated in Table I. We included only studies that were published in English. Both cases series and comparative studies reporting outcomes after TEA and ORIF in patients aged 60 years and above with an acute distal humerus fracture were included. Studies reporting outcomes of patients with a co-existent diagnosis of rheumatoid arthritis were included. The TEA could be of any design via any approach and the ORIF group could include any type of plate fixation via any approach. The study must have reported either a functional outcome measure or associated complications. Studies were excluded if participants included chronic injuries, non-unions or cases of failed plate fixation. In addition, only primary research was considered for review with any abstracts, comments, review articles and technique articles excluded.

The studies were appraised independently by two authors using a validated quality assessment scale for non-controlled study, [18] STROBE checklist [19] for comparative studies and the CONSORT statement for randomised controlled trials (RCT). [20]
Statistical methods

Results were pooled from different studies using meta-analysis techniques. If the required data was not provided in the original article, the corresponding author of the respective article was contacted to request these. Patient related outcome measures were only included in the meta-analysis if they were reported in at least two studies. Data regarding complications and revision surgery was also included.

Two main sets of analyses were performed. The first set of analyses compared TEA and ORIF using results from the comparative studies only. Four outcomes were used for comparisons; the Mayo Elbow Performance Score (MEPs), the Disability of Arm, Shoulder and Hand score (DASH), complications and re-interventions. MEPS and DASH scores are continuous and their differences in means were pooled assuming they are normally distributed. The complications and re-interventions reported as binary and percentage differences from identified studies were pooled. For the second set of analyses, TEA and ORIF were compared using results from both comparative studies and case series. In order to include data from case series as well from comparative studies, data from the TEA and ORIF arms were pooled separately. The means for MEPS and DASH scores were pooled assuming they were normally distributed. Complication and re-intervention rates were pooled assuming the number of complications and revisions were distributed binomially. The packages “meta” and “metaphor” in the R statistical program were used to perform these calculations. In both sets of analyses, a random-effects meta-analysis was used because it was believed that studies have inherent differences.
Results

The search strategy identified 27 studies eligible for inclusion; one randomised controlled trial,\[21\] four comparative studies,\[22-25\] 14 ORIF cases series\[26-39\] and 8 TEA case series.\[40-47\] A flow chart of the search strategy is shown in Figure I. The total number of participants in all studies was 1307; comparative studies (n=330), ORIF case series (n=777) and TEA case series (n=200). Concise details of included studies are given in Table II to V.

Comparative Studies

Of the five studies included, one was a randomised controlled trial providing level I evidence and the remaining four were level III retrospective comparative studies. The lack of randomisation in the retrospective comparative studies risks selection bias, and the failure to define a primary outcome measure or inclusion of a power calculation reduces the strength of these studies further. The study quality varied as demonstrated by the wide-ranging adherence to the STROBE checklist and CONSORT statement (Table VI and VII).

McKee et al.\[21\] performed a multi-centre randomised controlled trial of 40 patients which provided the highest level of evidence reviewed. The authors reported a statistically significant improvement in MEPS at every time point up to 2 years (p=0.015) and in the DASH (p=0.04) up until six months in the TEA group. However, the study failed to demonstrate a statistically significant difference in complication rate (p=0.40) which was the study’s defined primary outcome measure. A limitation of this study is the use of a combination of locking and non-locking plates in the ORIF treatment arm. It is clear from the contemporary literature that locking plates confer a significant biomechanical advantage and therefore the use of non-locking plates could be considered a possible confounder\[23, 24\] that may potentially have resulted in poorer outcomes in the ORIF group. The power calculation
was based on the intention to detect a 40% difference in reoperation rates, however a lower rate of reoperation would be an important clinical difference to distinguish, and therefore the high rate set has the potential to result in an under powered study. In addition, during the study 5 patients originally allocated to the ORIF group were transferred to the TEA group at the time of surgery as the surgeon deemed the fracture to be unfixable. This cross-over of these patients has the potential to unbalance the two groups by concentrating patients with more complex fractures within the TEA group. Patients with more complex fractures may have additional known and unknown confounding factors that may independently affect outcomes.

Ellwein et al. [23] retrospectively reviewed 29 patients, of whom 19 were in the subgroup over the age of 60 years included in the meta-analysis. The authors reported that those undergoing TEA had improved functional outcomes, DASH (p=0.023) and MEPS (p=0.078), and a 4.4 times lower risk of a major complication (95% CI 0.65-29.30). However, the study has limitations that include variations in characteristics between the two groups (ORIF group had a lower mean age, higher proportion of male patients and less severe fracture patterns), lack of details regarding reasons for treatment allocation and a variable length of follow up. The described difference in study populations may reflect true clinical practice as young male patients are deemed a relative contraindication to TEA due to the lifetime restrictions in function and concern regarding longevity of the implants.

Frankle et al. and Egol et al. [22, 24] performed retrospective comparative studies and demonstrated comparable results between the treatment modalities. Common limitations included the absence of randomisation, lack of clarity over the treatment allocation process and variation in patient characteristics. In addition, Frankle et al. [22] reported only female patients and 67% of patients in the TEA group suffered from rheumatoid arthritis compared to 0% in the ORIF group. Obert et al. [25] report a combination of a retrospective review
(n=410) and prospective study (n=87) focused on the complication rate associated with these procedures. Despite showing a higher complication rate after ORIF (44% versus 23%), the failure of the paper to describe further patient characteristics, surgical technique and functional outcome limits the information that can be obtained from this study.

Meta-analysis of comparative studies

Functional outcomes

There were three eligible studies that compared TEA to ORIF using the MEPS score, and the results of the meta-analysis for this outcome are presented in Figure II. In all studies, the mean MEPS score for TEA patients was greater than for ORIF patients. The pooled mean difference is 13.1 (95% CI 9 to 17) indicating TEA is associated with better outcomes with respect to MEPS. Two of the included studies compared TEA versus ORIF using the DASH score. The results of the meta-analysis are summarised in Figure III. In both studies, the mean DASH scores for TEA patients was superior to the mean DASH scores for ORIF patients. The pooled mean difference is 14.2 (95% CI 4 to 22) indicating TEA is statistically better than ORIF with respect to DASH.

Complications

Four of the included studies compared the complication profiles of TEA verses ORIF. Figure IV summarises the results for the meta-analysis of percentage complications differences. In all studies complications were higher in the ORIF group. The pooled percentage complications difference was 21 (95% CI 12 to 29) indicating TEA was associated with fewer complications than ORIF.
Re-interventions

Two studies compared the need for re-intervention after TEA and ORIF. The results are conflicting with McKee et al. \[^{21}\] reporting a lower re-intervention rate with TEA and Egol et al. \[^{24}\] the opposite (Figure V). Meta-analysis of the data from these two studies showed that the pooled percentage re-intervention difference is -8 (95% CI -29 to 13) suggesting that TEA is associated with lower risk of re-intervention than ORIF but the difference is not statistically significant.

To summarise, TEA is statistically superior to ORIF based on three outcomes (MEPS score, DASH score and complications) but not in terms of revision rate.

Case Series

In total 22 case series were reviewed, 14 ORIF cases series \[^{26-39}\] and 8 TEA case series. \[^{40-47}\] The size of the studies varied from 7 to 342 participants. 777 patients with a mean age of 77.8 years were analysed in the ORIF case series and 200 patients with a mean age of 75.7 years in the TEA case series. These studies provide only level IV evidence and hence have significant limitations that must be taken into account when interpreting the pooled data. The study quality varied as demonstrated by the wide-ranging adherence to the Rangel criteria (Table VIII and IX). Significant heterogeneity was encountered in study methodology that included treatment allocation, fracture pattern, surgical approach, type of implants and length of follow up.
**Single arm meta-analysis**

**Functional Outcomes**

Three comparative studies \(^{[21-23]}\) and seven TEA case series \(^{[40-42, 44-47]}\) reported MEPS score and the meta-analysis is shown in Figure VI. The pooled mean MEPS score from the ten TEA studies is 91.5 (95% CI = 88-95). Three comparative studies \(^{[21-23]}\) and three ORIF case series \(^{[28, 38, 39]}\) reported MEPS, the meta-analysis is shown in Figure VII. The pooled mean from the six studies is 82.8 (95% CI = 77-89). The 95% confidence intervals for MEPS in the TEA (88-95) and ORIF (77-89) groups overlap and therefore this difference is not statistically significant.

**Complications**

Eleven studies; seven case series \(^{[40-44, 46, 47]}\) and four comparative studies \(^{[40-42, 44]}\), reported complications for TEA (Figure VIII). The pooled percentage of complications in all eleven studies is 25% (95% CI = 19-32). Seventeen studies, thirteen case series \(^{[26-32, 34-39]}\) and four comparative studies, \(^{[21-23, 25]}\) reported complications for ORIF (Figure IX). The pooled percentage of complications in all seventeen studies is 34% (95% CI = 28-42). The 95% confidence intervals for complications in TEA (19-32%) and ORIF (28-42%) overlapped meaning the difference is not statistically significant.

**Re-intervention rates**

Two cases series \(^{[41, 42]}\) and two comparative studies \(^{[21, 24]}\) reported re-intervention rates for TEA. The pooled percentage of re-interventions for these four studies is 20% (95% CI = 14-28). Four cases series \(^{[26, 31-33]}\) and two comparative studies \(^{[21, 24]}\) reported re-intervention rates for ORIF. The pooled percentage of re-interventions for these six studies is 15% (95%
CI = 8-25). Overlapping of the confidence intervals suggests the difference in re-intervention rates for TEA and ORIF is not statistically significant.

Discussion

Meta-analysis of data from the included comparative studies has demonstrated that TEA is associated with superior outcomes with respect to MEPS, DASH and frequency of complications when compared to ORIF and that these findings are statistically significant. The pooled mean differences between TEA and ORIF of 13.1 for MEPS and 14.2 for DASH are higher than the recognised minimal clinically important differences for these metrics (10 for MEPS and 7-10 for DASH). [48-50] The only level 1 evidence from McKee et al. [21] also reported a statistically significant improvement in functional outcome after TEA with excellent or good outcomes according to the MEPS in 84% of TEA patients compared to 53% in the ORIF group. ORIF was associated with a 21% (95% CI 12 to 29) pooled increase in complications compared to TEA. These results demonstrate that TEA is associated with clinically superior outcomes with fewer complications when compared to ORIF.

Inclusion of five studies for the comparative meta-analysis resulted in 330 patients being available for analysis. Although combination of data increases the power of a meta-analysis, the low availability of studies still risks under powering. The RCT from McKee et al. [21] was conducting in keeping with the CONSORT statement providing high quality level I evidence. The four comparative studies were appraised against the STROBE statement which provides 22 criteria to assess the quality of the study against. The number of criteria met varied from 7 to 20 demonstrating that the quality of the evidence varied. Common weaknesses included the limited information provided on the methods of recording data, techniques used to
minimise bias, study limitations and statistical tests used. Obert et al. [25] only achieved 7 of the 22 criteria. Although the authors provided important data on risk of complications the limited information on sample selection, data collection, patient demographics and functional outcomes limits the strength of their results.

When single arm meta-analysis was performed for all studies, including comparative studies and case series, any differences in outcomes did not reach statistical significance. The failure of this part of the analysis to demonstrate significant results may be explained by the need to compute 95% confidence intervals for TEA and ORIF separately, which is more conservative estimate than calculating the confidence intervals for the difference in comparative studies. Furthermore, the inclusion of non-comparative studies increases the risk of bias and confounding.

Despite the meta-analysis of comparative studies demonstrating that TEA had a lower complication rate than ORIF, this analysis does not take into account the severity of the complication and the impact of the complications on patients. When performing arthroplasty, it is necessary to take into account the survivorship of the implant and the burden of any salvage procedures on the patient. The risk of TEA failure, in the form of component loosening, osteolysis or bushing failure, are all more likely after 5 years and this was not assessed in the included studies. The longevity of TEA in fracture patients has only recently been explored, Prasad et al. analysed 37 TEA non-rheumatoid trauma patients and reported implant survivorship of 89.5% at ten years. The study also showed that at ten years only 53% of the original cohort were alive, highlighting the typical patient demographic selected for TEA. [51] An attempt was made to compare the rate of re-intervention following the two procedures, however only two of the five included studies reported this outcome, their results were contradictory and the length of follow up in these studies was only 24 months. This
highlights the need for long-term comparative studies with explicit reporting of complications and re-operations.

Although this meta-analysis has shown an improved functional outcome and lower rate of complications after TEA in comparative studies, it is important to highlight that these findings should not be generalised to younger patients. Patients included in TEA groups had a trend to more complex fracture patterns as selection bias resulted in simpler fractures entering the ORIF groups. Therefore the role of TEA in a subgroup of younger patients (60-70 years) with less complex fractures perhaps provide the greatest clinical conundrum, where the survivorship and functional limitations of TEA must be balanced against the risks of ORIF. It is important to note that the lack of significant difference between the groups with respect to revision and complication rates should be interpreted with caution due to this selection bias and further randomised controlled study is appropriate to more clearly define the roles of each procedure in management of type C distal humerus fractures particularly in this subgroup of younger patients.

Alternative treatment options for distal humeral fractures are available that have not been included in this meta-analysis. Two recent retrospective studies report modest results when treating low-demand or medically unfit patients non-operatively.\(^{[10, 52]}\) Both papers conclude that that non-operative treatment can be considered in these patient categories in order to avoid the risks of surgery whilst TEA can still be used as a salvage procedure if non-operative treatment fails.\(^{[10, 52]}\) Hemiarthroplasty of the distal humerus is another surgical option that is gaining in popularity. A recent study reported at a mean of 3 years follow up that the functional outcomes were a mean MEPS score of 90 and a mean DASH score of 20, a complication rate of 19% and a 12% rate of revision surgery.\(^{[53]}\) Further work and research is required to fully delineate the role of these different surgical options in these fractures and
assess whether there are further subgroups who would particularly benefit from the differing surgical techniques.

The small number of comparative studies available meant case series were also evaluated to increase the data available for analysis. The inclusion of this lower quality evidence increases the risk of introducing bias into the results with the main limitations being the lack of a comparative group and randomisation. The case series were appraised against Rangel’s criteria which showed a wide variation in quality; this system includes 16 criteria to measure quality with scores ranging from 6 to 15. Common themes of study limitation included restricted information on the surgeons carrying out the procedure, the peri-operative care, the handling of missing data and the details regarding patient selection. These weaknesses were mitigated to some extent by performing statistical analyses on comparative and non-comparative studies separately. However, the failure of the non-comparative part of the meta-analysis to demonstrate any significant differences between groups may actually be as a result of the lower quality evidence and the variation in study quality. Therefore the ability of the comparative studies, which provide more robust evidence, to demonstrate statistically significant improvements after TEA form the basis for the studies conclusion.

Conclusion

Meta-analysis of comparative studies demonstrates that TEA is associated with statistically significant and clinically superior MEPS and DASH when compared to ORIF in elderly patients.
Conflicts of Interest

All authors confirm that they have no conflicts of interest related to this manuscript that might lead to bias or a conflict of interest. Professor Adnan Saithna has no conflicts of interest related to this manuscript but is a Consultant for Arthrex and has received expenses from Smith & Nephew.

Acknowledgment

The authors would like to acknowledge the help that authors of previous studies provided in contributing raw data for the meta-analysis.
References


Table I: Search strategy for Medline

Table II – Summary of the comparative studies included

Table III – Complications and revision rate in comparative studies

Table IV: Summary of the included ORIF case series

Table V: Summary of the included TEA case series

Table VI: STROBE Statement, checklist of items that should be included in reports of observational studies

Table VII: CONSORT 2010 checklist of information to include when reporting a randomised trial*

Table VIII: Adequacy of reporting of TEA case series based on criteria proposed by Rangel et al.

Table IX: Adequacy of reporting of ORIF case series based on criteria proposed by Rangel et al.
Figure I: Flow diagram of review process

Figure II: Forest plot comparing MEPS scores for TEA and ORIF from comparative studies

Figure III: Forest plot comparing DASH scores for TEA and ORIF from comparative studies

Figure IV: Forest plot comparing percentage complications for TEA and ORIF for comparative studies

Figure V: Forest plot comparing percentage revisions for TEA and ORIF from comparative studies

Figure VI: Forest plot for mean MEPS scores for TEA in comparative and case series

Figure VII: Forest plot for mean MEPS scores for ORIF in comparative and case series

Figure VIII: Forest plot for percentage of complications for TEA in case series and comparative studies
Figure IX: Forest plot for percentage of complications for ORIF in case series and comparative studies