Limited evidence is available on the optimal treatment of complex fractures of the proximal humerus. No randomized prospective studies have compared hemiarthroplasty, open reduction internal fixation and/or conservative treatment. Two systematic reviews are available but they do not include angle-stable plate osteosynthesis or third-generation prosthetic designs.

We conducted a systematic review of studies published in the last 10 years. The MOOSE guidelines were followed and we focused on functional outcome and its relation to age, gender and tuberosity healing after hemiarthroplasty.

Sixteen studies met the inclusion criteria. We found a correlation between increasing age and decreasing Constant score (r = -0.60, p = 0.0142). Tuberosity healing has influenced functional outcome in all series mentioning this parameter.

Hemiarthroplasty remains a valuable option for the treatment of proximal humerus fractures in elderly patients. However the quality of the available reports is poor. Large-scale, structured, prospective studies are needed to determine the current place of hemiarthroplasty of the shoulder in treatment of fractures of the proximal humerus.

Keywords: proximal humerus; fractures; hemiarthroplasty.

INTRODUCTION

Although the first shoulder arthroplasty dates back to the late 19th century, when Péan implanted a shoulder prosthesis to salvage a painful shoulder destroyed by tuberculosis, shoulder arthroplasty did not become popular until the 1970’s with Neer (18); he reported uniformly good or excellent results in acute fractures, but early enthusiasm rapidly waned when the results proved to be less reliable in other hands.

Misra et al (16) performed a systematic literature review (24 studies) on complex fractures of the proximal humerus in adults. They compared the clinical outcome following management of three- and four-part fractures of the proximal humerus with conservative treatment, internal fixation and hemiarthroplasty. They concluded that the range of motion was better in the hemiarthroplasty group and that there was no difference regarding pain and incidence of infection between fixation and arthroplasty. The risk of infection was of course lower in...
the non-operative group, but pain was significantly worse. However, the studies evaluated in this review were published between 1969 and 1999. The number of studies reporting on hemiarthroplasty is low (only 5) and there is a wide spread in results, with good and excellent results reported in 36 to 88% of cases. Lanting et al (12) recently published a systematic review of treatment modalities for proximal humerus fractures; they included 13 studies reporting on hemiarthroplasty. Compared to open reduction and internal fixation, the results of hemiarthroplasty in their review are less favourable regarding range of motion in three-part fractures and are comparable in four-part fractures. Arthroplasty resulted in significantly fewer complications. Meanwhile, technical advances have clearly occurred in both internal fixation (introduction of angle-stable plating and specific proximal humeral nails) and in hemiarthroplasty (fracture-specific prosthetic designs). Technical advances in the field of internal fixation resulted in some surgeons stating that hemiarthroplasty no longer has a place in the treatment of acute fractures. On the other hand, more recent reports still mention failure in up to one third of complex proximal humeral fractures treated by open reduction internal fixation (ORIF).

Inclusion criteria

- Prospective or retrospective study on fracture arthroplasty, published in English, German, French or Dutch
- Acute fracture care (i.e. within 30 days after trauma)
- Functional outcome scored with the Constant – Murley score
- Containing reports on at least 10 prostheses
- At least one year of follow-up
- Full articles published between January 1998 and December 2007

All studies were rated according to their level of evidence, which was rated from 1 to 4 according to Sackett et al (21) (randomized control trials, prospective trials, retrospective trials and case series) (12).

We decided not to include the abstracts of studies presented at the ESSSE (European Society for Surgery of Shoulder and Elbow) and ICSS (International Conference on Shoulder Surgery) as the format of reporting was too inconsistent.

Data extraction

All data were included in an Excel spreadsheet. Data to be included were:

- Number of prostheses included
- Outcome (absolute Constant score)
- Mean age
- Tuberosity healing (if available)
- Gender distribution (if available)
- Types of prostheses

Data analysis and reporting of results

Data were analysed and reported according to the MOOSE guidelines (23).

RESULTS

Sources

The Medline and Ovid search resulted in 405 different hits. The secondary search did not reveal any additional studies meeting all inclusion criteria. Forty-six studies meeting the inclusion criteria on the basis of title and abstract were reviewed. Overall 30 papers failed to meet the inclusion criteria for the following reasons: cadaver study (1),
Italian language (4), Chinese language (3), Czech language (3), Serb language (1), scoring system other than Constant score (6), review article (7), less than 10 (acute) cases (1), double publication of data (1), reversed prosthesis (3). Evaluation of the full text identified only 16 papers (2-6,8-11,13,14,17,19,22,24) fulfilling all inclusion criteria. An overview of all included studies, with their respective level of evidence (21) is presented in table I.

### Demographics (table II)

Six hundred and sixty-four patients were included in this review. Only 8 (1,6,9,10,13,17,19,22) of the 16 studies mentioned the initial number of prostheses implanted in the study period. Due the number of drop-outs in these studies, complete follow-up was available for only 272 of the 346 patients treated, i.e. 78.5%. The main reasons for high drop-out rates were patient’s inability to attend follow-up investigations because of poor general condition and patient death during the study period.

The mean age of the patients in the review was 66.8 years, with a large variation between studies: in the study by Demirhan et al (8), the mean age was only 58 years, versus 77.6 years in the study by Anjum and Butt (3).

Twelve studies (1-6,10,11,13,14,17,19,22) mentioned the gender distribution of the patients. The gender of 571 out of 690 (82.7%) patients was reported. There was an overall female predominance: 432/571 (75.7%). The gender distribution varied, with a proportion of female patients ranging from 59.3% in the study by Becker et al (4) to 87.8% in the study by Reuther et al (19).

Thirteen studies (1-5,9-11,13,14,17,19,22) mentioned the prostheses used. Eight different types were used. As the number per type was low and more than one type had been used in 5 out of 12 studies, without mention of the respective Constant score, no analysis of the influence of the type of prosthesis on outcome could be made based on the available data.

### Parameters predicting outcome (table II)

Age is often defined as one of the most important factors predicting outcome. All 16 studies provided data on age and Constant score. There was a negative correlation (Pearson) between age and Constant score: \( r = -0.60, p = 0.0142 \). A correlation between age and Constant score was reported in 7/16 studies, but in two studies no relation between age and Constant score could be demonstrated. In

---

**Table I. — Articles included in the systematic review**

<table>
<thead>
<tr>
<th>Author</th>
<th>Journal</th>
<th>Year</th>
<th>Level of evidence (21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agorastides et al (1)</td>
<td>JSES</td>
<td>2007</td>
<td>4</td>
</tr>
<tr>
<td>Loew et al (13)</td>
<td>JBJS (Br)</td>
<td>2006</td>
<td>4</td>
</tr>
<tr>
<td>Mehlhorn et al (14)</td>
<td>Acta Orthop Belg</td>
<td>2006</td>
<td>4</td>
</tr>
<tr>
<td>Gierer et al (9)</td>
<td>Orthopäde</td>
<td>2006</td>
<td>4</td>
</tr>
<tr>
<td>Anjum &amp; Butt (3)</td>
<td>Acta Orthop Belg</td>
<td>2005</td>
<td>4</td>
</tr>
<tr>
<td>Schmal et al (22)</td>
<td>Unfallchirurg</td>
<td>2004</td>
<td>4</td>
</tr>
<tr>
<td>Kralinger et al (11)</td>
<td>JBJS (Br)</td>
<td>2004</td>
<td>4</td>
</tr>
<tr>
<td>Kollig et al (10)</td>
<td>Zentralbl Chir</td>
<td>2003</td>
<td>4</td>
</tr>
<tr>
<td>Demirhan et al (8)</td>
<td>JOT</td>
<td>2003</td>
<td>4</td>
</tr>
<tr>
<td>Boileau et al (5)</td>
<td>JSES</td>
<td>2002</td>
<td>4</td>
</tr>
<tr>
<td>Ambacher et al (2)</td>
<td>Zentralbl Chir</td>
<td>2000</td>
<td>4</td>
</tr>
<tr>
<td>Bosch et al (6)</td>
<td>JSES</td>
<td>1998</td>
<td>4</td>
</tr>
<tr>
<td>Zyto et al (24)</td>
<td>JSES</td>
<td>1998</td>
<td>4</td>
</tr>
</tbody>
</table>
7/16 studies the correlation between age and Constant score was not mentioned or could not be calculated (fig 1).

Some authors state that females have a poorer prognosis than males; other studies demonstrate no such relationship. Considering the percentage of female patients in relation to the Constant score, we did not observe a significant influence (Pearson) in the 12 studies (1,3,4,5,10,11,13,14,17,19,22) in which patients’ gender was specified: \( r = -0.17, p = 0.6044 \) (fig 2).

Healing of the tuberosities is often mentioned as the most important factor predicting functional outcome. In all the individual studies presenting the relationship between tuberosity healing percentage and functional outcome (7/16) there was a significant difference between the group with healed tuberosities and the group without (1,5,9,11,13,19,22). However, in a review of the 6 studies in which healing of the tuberosities was reported, we could not find a significant correlation (Pearson) between the percentage of tuberosity healing and the Constant score: \( r = -0.41, p = 0.3556 \) (fig 3).

**DISCUSSION**

Shoulder hemiarthroplasty remains a valuable option in the treatment of complex proximal humeral fractures in the elderly. However, Neer’s (18) initial optimism regarding the results should be mitigated. The average Constant score for the entire review population (664 patients) was 53.9 points. Most authors report little pain after hemiarthroplasty for acute fractures, while mobility and strength remain limited. Somehow the low Constant score contrasts with the reported patient satisfaction, which is usually high. This discrepancy may be explained by the relatively low demands of the aged (female) population typically treated with hemiarthroplasty. The typical patient is satisfied if she is free of pain and can take care of her (limited) ADL activities.

<table>
<thead>
<tr>
<th>Author</th>
<th>N</th>
<th>N (init)</th>
<th>CS</th>
<th>Age (years)</th>
<th>% females</th>
<th>% tub heal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reuther et al (19)</td>
<td>57</td>
<td>66</td>
<td>50</td>
<td>74.9</td>
<td>87.72%</td>
<td>36</td>
</tr>
<tr>
<td>Agorastides et al (1)</td>
<td>49</td>
<td>59</td>
<td>48.4</td>
<td>70</td>
<td>79.59%</td>
<td>92</td>
</tr>
<tr>
<td>Loew et al (13)</td>
<td>39</td>
<td>47</td>
<td>53.2</td>
<td>72.2</td>
<td>76.92%</td>
<td>56.5</td>
</tr>
<tr>
<td>Mehlhorn et al (14)</td>
<td>26</td>
<td>26</td>
<td>52</td>
<td>70.3</td>
<td>73.08%</td>
<td>NA</td>
</tr>
<tr>
<td>Gierer et al (9)</td>
<td>18</td>
<td>24</td>
<td>56</td>
<td>75.6</td>
<td>NA</td>
<td>31</td>
</tr>
<tr>
<td>Anjum &amp; Butt (3)</td>
<td>20</td>
<td>20</td>
<td>47.5</td>
<td>77.6</td>
<td>75.00%</td>
<td>NA</td>
</tr>
<tr>
<td>Schmal et al (22)</td>
<td>17</td>
<td>20</td>
<td>51.7</td>
<td>70.2</td>
<td>80.00%</td>
<td>26.5</td>
</tr>
<tr>
<td>Kralinger et al (11)</td>
<td>167</td>
<td>167</td>
<td>55.4</td>
<td>70</td>
<td>76.05%</td>
<td>53.9</td>
</tr>
<tr>
<td>Kollig et al (10)</td>
<td>38</td>
<td>46</td>
<td>66</td>
<td>60</td>
<td>73.91%</td>
<td>NA</td>
</tr>
<tr>
<td>Demirhan et al (8)</td>
<td>32</td>
<td>32</td>
<td>68</td>
<td>58</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Boileau et al (5)</td>
<td>66</td>
<td>66</td>
<td>56</td>
<td>66</td>
<td>68.18%</td>
<td>50</td>
</tr>
<tr>
<td>Becker et al (4)</td>
<td>27</td>
<td>27</td>
<td>45</td>
<td>67</td>
<td>59.26%</td>
<td>NA</td>
</tr>
<tr>
<td>Ambacher et al (2)</td>
<td>27</td>
<td>27</td>
<td>65</td>
<td>69</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Bosch et al (6)</td>
<td>25</td>
<td>39</td>
<td>54.2</td>
<td>64.5</td>
<td>68.00%</td>
<td>NA</td>
</tr>
<tr>
<td>Movin et al (17)</td>
<td>29</td>
<td>45</td>
<td>38</td>
<td>71</td>
<td>82.76%</td>
<td>NA</td>
</tr>
<tr>
<td>Zyto et al (24)</td>
<td>27</td>
<td>27</td>
<td>46</td>
<td>71</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

N : number of patients with complete follow-up
N (init) : initial number of patients included in the study
CS : mean Constant score at completion of follow-up
Age : mean age at inclusion
% females : % of female patients with complete follow-up
% tub heal : % of patients with documented anatomical tuberosity healing in the study.
NA : data not available.
Age is the most constantly cited factor predicting outcome and the only one for which we could demonstrate a statistically significant correlation with functional outcome rated with the Constant score. Although tuberculosis healing correlated with function in every individual study in which this was calculated, the percentage of healed tuberosities did not correlate with the Constant score in our review. The lack of precise data in most studies made it impossible to calculate the precise influence of tuberculosis healing on functional outcome. Since a significant relationship could be demonstrated in all individual studies, tuberculosis healing should be considered as an important factor predicting outcome. The findings are less consistent regarding gender. Three authors report a significant relationship between female gender and poor outcome; three other studies do not. On the whole, no significant relationship is revealed in the study population. It was impossible to determine any relationship between the type of prosthesis and outcome, owing to the large number of different prosthesis types used in the studies and the lack of precise data. One study mentions a significant influence of the type of prosthesis on tuberculosis healing (11).

One weakness of the review is that some of the studies reporting on larger series (7,15,20) of hemiarthroplasty for fracture could not be included, since they used other scores than the absolute Constant score to report functional outcome. However, the overall conclusions of these studies

---

Scatter plot “Mean age versus Constant score” of the 16 studies included in this review. Note the statistical significant negative correlation between increasing age and functional outcome, as scored by the Constant score.

**Fig. 1.** — Linear regression and correlation Constant Score to Age.

Scatter plot “%female patients included in the study versus functional outcome” in the 12 studies indicating this parameter. We could not demonstrate a correlation between the percentage of female patients and the functional outcome.

**Fig. 2.** — Correlation Constant Score to gender.
on parameters influencing outcome are similar to those of our review. Of course, as with all reviews based upon published studies, there is a potential publication bias. Eleven potentially interesting studies had to be excluded for linguistic reasons (4 Italian, 3 Czech, 3 Chinese and 1 Serb). It was not possible to determine how many of them met all of the other inclusion criteria.

The quality of the studies included is low, as revealed by the level of evidence according to Sackett et al (21). In order to investigate the value of shoulder hemiarthroplasty in the treatment of complex acute fractures of the proximal humerus, and especially to investigate its value compared to conservative treatment or open reduction and internal fixation, high quality prospective (randomised) studies are needed. Considering the high numbers of patients which would be required to demonstrate possible benefits of specific treatments in prospective randomised studies, prospective cohort studies focusing on one parameter (such as tuberosity healing) could also give an indication on the value of the method or technique used.

REFERENCES