Radiographic and histopathologic analysis of osteolysis after total shoulder arthroplasty

Christopher K. Kepler, MD, MBA, Shane J. Nho, MD, MS, Manjula Bansal, MD, Owen L. Ala, MD, Edward V. Craig, MD, Timothy M. Wright, PhD, Russell F. Warren, MD

Hypothesis: This study analyzed clinical, radiographic, and histologic data from failed total shoulder arthroplasties (TSAs) to determine factors associated with osteolysis.

Materials and methods: From 1985 to 2005, 52 patients (mean age, 61.6 years) underwent revision TSA at a single institution at a mean of 4.3 years after their index surgery. Patients were retrospectively assigned to 2 cohorts based on the presence (n = 10) or absence (n = 42) of osteolysis around their implants on the last prerevision surgery radiographs. Clinical information, associated histopathology from tissues obtained at revision surgery, and polyethylene wear data from the retrieved glenoid components were compared between groups.

Results: In the osteolysis group, 20% had screw fixation compared with 2.5% without osteolysis (P = .039). The radiolucent score was significantly higher in the osteolysis group: 12.7 ± 2.0 vs 8.7 ± 3.7 (P = .003). Wear analysis of the osteolysis group demonstrated significant increases in third-body particles compared with those implants without osteolysis (P = .004). Histology available from retrieved implants demonstrated particulate debris in 62% of patients with osteolytic lesions vs 67% without osteolytic lesions (P > .05).

Discussion: Significant differences were found in patients with osteolytic lesions compared with patients undergoing TSA revision surgery without such lesions, specifically with regard to glenoids that used adjuvant screw fixation, the presence of increased radiolucent lines, and an abundance of third-body wear. No significant differences were found in particulate wear debris despite the prevailing notion that osteolysis is associated with particulate debris from implant wear.

Conclusion: Screw fixation and third-body wear were associated with osteolysis after TSA.

Level of evidence: Level III, Retrospective Comparative Study.

Keywords: Failed total shoulder arthroplasty; osteolysis; polyethylene wear; implant debris; retrieval analysis

Shoulder arthroplasty was introduced in 1955 by Dr Charles Neer for treatment of proximal humeral fractures with hemiarthroplasty. The indication for the procedure...
was extended to osteoarthritis, and eventually, development of a polyethylene component for resurfacing the glenoid allowed for total joint replacement. The success of total shoulder arthroplasty (TSA) in alleviating pain and improving range of motion is well documented.

Despite this success, the need for revision after TSA does arise: the first revision series was published only 8 years after the introduction of polyethylene glenoid replacement. The subsequent evolution of implant design and surgical technique has not erased the truth in Dr Neer's evaluation of the postrevisions shoulder: 'The function of revisions is often impaired by muscle damage, bone loss, and scar.'

The incidence of component revision after TSA ranges from 0% to 13% at 3 to 6 years. A recent meta-analysis of TSA outcomes by van de Sande et al. estimated average revision incidence at 8% over an average follow-up of 59 months. Few articles have been published that have average follow-up lengths of longer than 6 years. Sneppen et al. published a revision rate of 8% in 62 patients during an average of 92 months, and Torchia et al. recorded a 10% revision rate for 113 patients over 144 months. Average time to revision was 25 months for Amstutz et al., but was 7.3 years for Torchia et al., suggesting that medium-term follow-up studies underestimate revision rates.

Although long-term studies comment on TSA revision rates, few have specifically described the possible contribution of osteolysis to the need for revision surgery. Of the studies listed in the previous paragraph that reported results after TSA, none specifically commented on the presence of osteolysis. Osteolysis surrounding TSA components is a source of concern given the associated loss of bone stock and painful component loosening requiring revision surgery. Differences have been documented in the size, shape, and texture of particles generated by wear of shoulder components compared with those after hip replacement, but osteolysis in the shoulder likely occurs by much the same mechanism as in the hip—phagocytosis of particulate debris from component abrasive and adhesive wear triggers macrophage activation and cytokine release that in turn activates osteoclasts.

Although the incidence of radiolucent lines around the glenoid component ranges from 10% to 96%, the presence of incomplete radiolucency adjacent to the glenoid is not necessarily synonymous with a loose component, and care must be taken to distinguish radiolucent lines from osteolytic lesions. In addition, although a 2% to 12% incidence of glenoid loosening after TSA has been reported, these figures include many cases in which the only evidence of loosening is complete radiolucent lines surrounding components, without significant bone resorption. No prevalence data for osteolysis around glenoid components have been published, perhaps because far fewer TSAs are performed compared with total hip arthroplasty (THA) or total knee arthroplasty (TKA), although the annual percentage increase in TSA in recent years rivals those of THA and TKA.

Likely because of this lower incidence, factors associated with osteolysis surrounding TSA are much less well characterized compared with THA and TKA. Delineation of these factors may be important for prognostic evaluation of painful shoulder prostheses before the appearance of osteolysis as well as in directing implant design. The purposes of this study were to (1) report the prevalence of osteolysis in a population of patients requiring TSA revision surgery and (2) to define demographic, implant-related, radiographic, and histologic factors associated with the development of osteolysis requiring revision surgery.

**Materials and methods**

This study was approved by the Hospital for Special Surgery Institutional Review Board (Study No. 24097) in December 2007. From 1985 to 2005, 52 consecutive patients with a mean age of 61.6 ± 11.9 years underwent revision TSA at a single institution at a mean of 4.3 ± 4.7 years after the index surgery. The inclusion criteria were patients who underwent revision TSA with availability of prerevision surgery radiographs and availability of the explanted components for evaluation of wear patterns.

The prerevision surgery radiographs were reviewed and defined to demonstrate osteolysis as a well-defined scalloped area in the bone surrounding the glenoid component with relative radiolucency compared with the surrounding bone according to the Engh criteria. After radiographs were graded by 2 observers and a consensus was achieved regarding the presence or absence of osteolysis on each, the senior author (R.F.W.) reviewed the radiographs with suspected osteolytic lesions to confirm presence of a lesion. This radiographic criterion was used to retrospectively assign patients to 2 cohorts based on the presence (n = 10) or absence (n = 42) of osteolysis around their implants.

Clinical information was retrospectively obtained from medical records, including demographics, hand dominance, tobacco use, occupation, surgical dates, medical comorbidities, shoulder history, and prerevision clinical findings (range of motion, symptoms). Intraoperative findings, implant information, and postoperative complications were also obtained.

Tissue specimens were taken on removal of the glenoid component from soft tissue adjacent to the glenoid component and the surrounding cement mantle. The specimens were fixed in 10% buffered formalin, and representative samples were selected for analysis using light microscopy. For histologic examination, 5-μm-thick serial sections were made from paraffin-embedded specimens. All sections were studied by light and polarized-light microscopy, which allowed comparison of histopathology with respect to particulate wear debris and inflammatory reactions using a modification of the semiquantitative scoring system described by Willert et al. and Agins et al. This system assigns a grade of 1 to 5 based on the prevalence of each type of particulate debris (polyethylene, cement, and metal) and each proinflammatory cell type (lymphocytes, plasmacytes, histiocytes, and giant cells) after a review of all histologic specimens retrieved at the time of surgery.

All explanted implants were available for wear analysis. The polyethylene bearing surfaces of the components were examined microscopically using loupes with ×4 magnification and in a light stereomicroscope at magnifications up to ×31 for evidence of
burnishing, abrasion, scratching, pitting, delamination, focal wear, surface deformation, embedded third-body debris, and fracture. The surface was divided into anterior, posterior, superior, and inferior quadrants and given a subjective damage score of 0 to 3 for each damage mode in each quadrant plus a possible extra point for severely damaged quadrants using an established grading system.\textsuperscript{18} The quadrant scores were then summed to arrive at a composite score (maximum score, 12).

The most recent plain anteroposterior (AP) and axillary radiographs before revision surgery were examined. The extent and amount of radiolucency in the AP view was measured with digital calipers and scored according to the method of Moloč,\textsuperscript{21} in which the radiolucent lines around the prosthesis are assigned a numeric value depending on the thickness of the lucency. Measurements of lucency in all 6 zones adjacent to the glenoid prosthesis and keel or pegs (Fig. 1, A) provided a composite score that corresponded to no loosening (0 to 6 points), possible loosening (7 to 12 points), or definite loosening (13 to 18 points). We extended this methodology in scoring the axillary radiograph (“modified Moloč score”) for 3 zones corresponding to the anterior rim of the glenoid, the fixation keel or pegs, and the posterior rim of the glenoid (Fig. 1, B). Radiographs were additionally scored according to previously described classification systems for glenoid loosening\textsuperscript{32} and glenoid lucency.\textsuperscript{16} Subluxation of the humeral head\textsuperscript{32} was assessed by measuring translation as a percentage of the width of the head on both the AP and axillary radiographs for both groups. Finally, the glenoid version was measured as the angle formed by a line perpendicular to the scapular axis and a line drawn across the face of the glenoid cavity.\textsuperscript{27}

Data from the cohorts with and without osteolytic lesions were compared using the $t$ test with SPSS 14.0 software (SPSS Inc, Chicago, IL). Patient characteristics, clinical information, and radiographic and damage data were analyzed for between-cohort differences using a value of $P = .05$ to define significance.

**Results**

**Clinical information and demographics**

Overall, 10 of 52 patients (19%) had radiographic evidence of osteolysis surrounding the glenoid component. No significant differences existed between the patients with osteolytic lesions and those without significant lesions (Table I) with regard to age, concordance of operative side with dominant hand, smoking history, primary diagnosis, length of implantation, range of motion, concurrent rotator cuff tear, and subjective instability.

**Components**

The implant system used did not demonstrate significant variation by decade of index surgery. The 2 most commonly used implant systems used were Biomet Biomodular and variations of the Neer prosthesis with DePuy implants and custom implants used more rarely. The proportions of prostheses used in the 1980s, 1990s, and 2000s were 67%, 71%, and 80%, respectively, for Biomet and 26%, 16%, and 20%, respectively, for Neer. Cement fixation was used in 49 glenoid prostheses (Fig. 2), and screw fixation was used in 3 (Fig. 3). Radiographs that

<table>
<thead>
<tr>
<th>Table I</th>
<th>Characteristics of patients with and without osteolysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Osteolysis $*$</td>
</tr>
<tr>
<td>Patients, No.</td>
<td>10</td>
</tr>
<tr>
<td>Age, y</td>
<td>60.5 $+/-$ 12.9</td>
</tr>
<tr>
<td>Dominant/operative hand concordance, %</td>
<td>40</td>
</tr>
<tr>
<td>Smoking history, %</td>
<td>30</td>
</tr>
<tr>
<td>Primary diagnosis, %</td>
<td></td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>70</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>20</td>
</tr>
<tr>
<td>Posttraumatic</td>
<td>10</td>
</tr>
<tr>
<td>Avascular necrosis</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
<tr>
<td>Length of implantation, y</td>
<td>6.2 $+/-$ 5.7</td>
</tr>
</tbody>
</table>

$* P > .05$ for differences between groups for all characteristics.
demonstrated osteolysis were significantly more likely to have come from a TSA with a metal-backed glenoid component with adjuvant screw fixation compared with those without osteolysis (20% vs 2.5%, \(P = .039\); Fig. 3).

Radiographic measurements and scores

Radiolucent lines surrounding the glenoid component in the osteolysis group were thicker on prerevision radiographs compared with those in the cohort without osteolytic lesions in zones 1 (3.0 mm vs 1.6 mm; \(P = .037\)) and 2 (3.6 mm vs 1.3 mm; \(P = .019\)), corresponding to the areas adjacent to the superior portions of the glenoid.

The overall radiolucency score was 12.7 in the osteolysis group vs 8.7 in those without osteolysis, which was significantly higher (\(P = .003\)). Radiolucency subscores in areas 1, 2, and 5 corresponding to the superior and inferior portions of the glenoid and the interface between the superior aspect and keel were also significantly higher in the osteolysis group. The average score for the TSAs with osteolysis fell into the “definite loosening” (score of 12 to 18) category, whereas the nonosteolytic average score was in the “possible loosening” range on the grading scale. Modified scores based on the axillary radiograph were not significantly different between the osteolysis and non-osteolysis groups.

Measurements of coronal and sagittal plane subluxation were not significantly different between groups. Glenoid version in the osteolysis cohort was not different from the version in the nonosteolysis cohort, but demonstrated a trend toward increased retroversion in the osteolysis cohort (3.5° vs 17.6°, \(P = .14\)).

Wear analysis

No significant differences were found in the wear scores between those glenoids taken from patients with osteolysis vs those retrieved from patients without osteolytic lesions for burnishing, abrasion, scratching, pitting, delamination, focal wear, surface deformation, and fracture. However, glenoid components retrieved from the osteolysis group had significantly higher scores for embedded third-body particles (Fig. 4) compared with those glenoids from patients without evidence of osteolysis (\(P = .004\)), consistent with the presence of large amounts of metallic debris embedded into the bearing surface.

Histology

Grossly, the specimens taken at the time of revision surgery were generally tan-gray. They lacked the dark-gray or blackened areas seen in soft-tissue metallosis, with the exception of those specimens taken from patients who had metal-backed glenoid components.

Examination by light microscopy demonstrated varying degrees of particulate implant consisting of polyethylene, metal, and cement. The polyethylene appeared as cleft-like spaces containing large and small shards of transparent/translucent material (Fig. 5, A). Associated foreign body giant cells and histiocytic reaction to polyethylene was
noted. In other areas, the cytoplasm of histiocytes and foreign body giant cells contained dot-like black metallic debris measuring 1 to 2 \( \mu \)m in diameter (Fig. 5, B). In addition, large empty foci with giant cell reaction, scalloped edges, and containing buff-colored material (barium) were seen, representing the presence of methylmethacrylate (Fig. 5, C). Polarized microscopy confirmed the presence of polyethylene particles, which appear strongly birefringent. The small black metallic particles are refractile due to diffraction of light from their irregular edges, and the empty spaces considered to be dissolved cement do not show refractile material under polarized light.

Histologic specimens taken at revision surgery demonstrated no significant differences between patients with or without osteolysis with respect to the presence of debris particles (60% vs 67%). Further analysis of the specimens for differences with regard to the presence of specific particles from polyethylene, metal, or cement also yielded no significant differences between groups (Table II). Cells associated with inflammatory states (lymphocytes, plasma cells, histiocytes and giant cells) were found to a similar degree in both groups.

**Discussion**

Although factors associated with osteolysis in THA are well established, osteolysis after TSA has not been well characterized. With reported revision rates as high as 10% and an increasing number of TSAs being performed each year, osteolysis around glenoid components is an increasingly troubling occurrence, especially in light of the difficulties that extensive bone loss poses for revision surgery.

The rate of osteolysis after TSA in clinical series has been reported as high as 23%, but most series have not specifically described the incidence of osteolysis after TSA. The incidence of osteolysis in patients undergoing revision TSA

---

**Figure 4** In this photograph of a retrieved glenoid component, the circles designate areas of implanted third-body debris.

**Figure 5** (A) Photomicrograph shows large and small shards of polyethylene with histiocytic and giant cell reaction (hematoxylin and eosin [H&E] stain, original magnification \( \times 40 \)). (B) Photomicrograph shows particulate black metallic debris within histiocytes (H&E stain, original magnification \( \times 25 \)). (C) Photomicrograph shows irregular spaces with giant cell reaction to cement/granular material (barium; H&E stain, original magnification \( \times 25 \)).
in this series was 19% considering all implant types together. One would expect the osteolysis rate to be higher in this study compared with other follow-up studies after TSA because of the preselection for failed TSA in our series.

Of the 3 metal-backed glenoids with screw fixation, 2 demonstrated osteolysis (Fig. 3), corroborating the findings of Boileau et al.,\(^5\) who described the association of metal-demonstrated osteolysis, (Fig. 3), corroborating the findings for failed TSA in our series.

Compared with other follow-up studies after TSA because of the small number of patients involved.

The osteolysis cohort in our study had higher overall Molé scores as well as higher subscores for 3 different zones compared with patients without osteolysis, a finding that is intuitive because the width of the osteolytic lesions was measured for Molé scoring of the quadrant in which it was located. Although osteolysis is traditionally seen with particle disease, whereas radiolucent lines are associated with mechanical instability, we believed that osteolytic regions would likely contribute to mechanical loosening in the same way as radiolucent lines and so considered them to pose similar threats to stability from a functional if not etiologic perspective, but acknowledge that the non-weight-bearing nature of the shoulder will mitigate this effect to an extent.

### Table II
Comparison of histologic findings of patients with and without osteolysis

<table>
<thead>
<tr>
<th>Finding</th>
<th>Osteolysis* (%)</th>
<th>No osteolysis* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particulate debris</td>
<td>60</td>
<td>67</td>
</tr>
<tr>
<td>Metallic</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>Polyethylene</td>
<td>60</td>
<td>63</td>
</tr>
<tr>
<td>Cement</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Lymphocytes</td>
<td>40</td>
<td>58</td>
</tr>
<tr>
<td>Plasmacytes</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>Histocytes</td>
<td>60</td>
<td>66</td>
</tr>
<tr>
<td>Giant gels</td>
<td>60</td>
<td>73</td>
</tr>
</tbody>
</table>

* \( P > .05 \) for differences between groups for all characteristics evaluated.

### Table III
Characteristics of patients with and without screw fixation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cemented</th>
<th>Screw fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, No.</td>
<td>49</td>
<td>3</td>
</tr>
<tr>
<td>Age, y</td>
<td>60.8</td>
<td>69.4</td>
</tr>
<tr>
<td>Length of implantation, y</td>
<td>4.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Primary diagnosis, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>57</td>
<td>100</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Posttraumatic</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Avascular necrosis</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Average RLL thickness, mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zone 4</td>
<td>1.6*</td>
<td>3.8*</td>
</tr>
<tr>
<td>Other zones</td>
<td>1.6</td>
<td>2.1</td>
</tr>
<tr>
<td>Version, deg</td>
<td>-4</td>
<td>-23</td>
</tr>
</tbody>
</table>

RLL, radiolucent lines.  
* \( P < .05 \) for differences between groups.
Embedded third-body debris into the bearing surfaces of the retrieved polyethylene glenoid components was the only wear parameter that was significantly increased in the osteolysis group compared with the nonosteolysis group. This finding is consistent with reports published describing associations of osteolysis with implant debris, including reports after failure of TKA and THA. Although it is widely assumed that osteolysis represents the manifestation of macrophage activation by wear particles, no prior study has associated implant wear debris with osteolysis in a study involving patients both with and without radiographic signs of particle disease. Surprisingly, other damage modes were not correlated with osteolysis, because these wear modes also produce particulate debris. This finding may suggest that cement and metal particles contribute significantly to the generation of osteolytic lesions after TSA in addition to the debris from polyethylene bearing surfaces.

Although particles of both methacrylate and metal were seen embedded in retrieved polyethylene during retrieval analysis and in histologic specimens, no grading systems have yet been developed to quantify wear of metal or cement in TSA components. Interestingly, no increase in the number of debris particles was found on histologic examination in the osteolysis group, a result that may be due to sampling inconsistencies and the sensitivity of the quantification methods available to a retrospective study.

Limitations of the present study include its retrospective nature. One limitation of using a technique such as damage mapping is that it can only be practically done retrospectively. We only analyzed the polyethylene glenoid components for damage; wear of the metallic humeral heads from these retrievals could also have contributed to the development of osteolysis through altered mechanics or independent generation of particulate debris. Grading of osteolysis based on radiographs has been shown to have poor interobserver reliability but moderate to excellent intraobserver reliability—for this reason we used a 2-observer consensus method to evaluate the radiographs for osteolysis and vetted these findings with a third observer. Nevertheless, this is another limitation of our retrospective study.

In addition, the long time period spanned by our study means that different implant designs and cementing techniques were considered together. However, many confounding factors, including prosthesis type that could have affected our results, were not significantly different between groups (as described in our results).

Another study limitation involves the evaluation of histologic specimens taken at revision surgery and the method of intraoperative tissue selection. Our semi-quantitative analysis of these specimens found no significant differences between the groups with and without osteolysis, but this technique may be subject to sampling variations. Samples from multiple areas were examined to include as much of each specimen as possible in grading the specimen for debris and cellular reaction. Despite attempts to mitigate sampling error, the degree of specimen heterogeneity due to intraoperative sampling variation as well as within an individual gross specimen cannot be quantified. Development of rigorous criteria to guide tissue sampling in the operating room and use of techniques to analyze cellular activity in the implant-bone interface zones are being considered but will need to be studied in a prospective manner.

Conclusions

This study examined demographic, clinical, radiographic, and retrieval analysis characteristics associated with osteolysis around TSA components on prerevision radiographs. We found metal-backed glenoid components were associated with osteolysis, consistent with the findings of other studies. Third-body embedded metallic particles into the polyethylene surface and thicker radiolucent lines were also significantly associated with the presence of osteolysis. Further insight into associations between particulate debris found in tissue adjacent to explanted TSA and the resulting biologic response will likely come with more sophisticated analytic methods and prospective, immediate analysis of retrieved tissue in a preordained fashion.

Acknowledgments

Partial support from the Clark and Kirby Foundations is gratefully appreciated.

Disclaimer

The authors Christopher Kepler, Shane Nho, Manjula Bansal, Owen Ala, and Timothy Wright, their immediate families, and any research foundations with which they are affiliated, have not received any financial payments or other benefits from any commercial entity related to the subject of this article. Edward Craig has a royalty agreement with Biomet. Russell Warren has royalty agreements with Biomet and Smith & Nephew, stock considerations with Zimmer, is a paid consultant for KFx, and receives research support from Smith & Nephew.

References

Factors associated with osteolysis after TSA


