We tested the accuracy of MRI for the precise quantification of the volume of osteonecrosis in 30 hips (stage III). The values were compared with direct anatomical measurements of the femoral heads obtained after total hip replacement. When the area of osteonecrosis was determined visually, and manually outlined on each slice, the accuracy of the measurement of volume was satisfactory, and the mean absolute deviation between MRI and anatomical measurements was similar to that between two MRI data sets. For ten of the hips which were measured by MRI, both before and after collapse, the volume did not appear to change significantly. Our findings suggest that the volume of osteonecrosis can be determined with accuracy by MRI, both before and after collapse.

Received 12 June 2000; Accepted after revision 3 January 2001

The outcome of osteonecrosis of the femoral head is influenced by the size of the lesion. Many methods\(^1,2\) have been used to estimate the size of the lesion on radiographs or MR scans. These studies have also indicated that the extent of the necrotic lesion affects its evolution.\(^3,4\) Since, however, there are various sources of artifact and because geometrical distortions may occur with MRI,\(^5,6\) quantitative measurements of the infarct cannot be validated accurately on phantom models. To our knowledge, no method of measuring the volume of the osteonecrosis by MRI has been compared with the absolute volume estimated anatomically. Our aim therefore was to test the accuracy of MRI for the precise quantification of the volume of the infarct.

Patients and Methods

We used MRI to assess stage-III osteonecrosis in 20 patients (30 hips) at the time of diagnosis using a 1.5 T superconducting unit (Magneton: Siemens, Erlanden, Germany). Preoperative medical diagnoses included lupus erythematosus, asthma, glomerulonephritis, renal transplantation and haematological malignancy. Risk factors included the use of corticosteroids and alcohol abuse. All were treated by total hip arthroplasty (THA). Ten had had MRI before collapse at stage I or stage II. The data obtained before and after collapse were used to test the accuracy of MRI to measure the volume of the infarct.

Measurement of the volume of the osteonecrosis. Coronal T1-weighted spin-echo images were obtained using the following parameters; 500/20 (repetition (TR) ms/echo time (TE) ms), 256 \times 256 matrix, 0 mm gap and 16 cm field of view. The same frequencies were used to transmit and receive attenuations for each specimen in order to allow a quantitative comparison of signal intensities. The images were filmed in a standard manner using a window width and level settings appropriate to optimise subjective assessment of signal intensity, and particularly to identify the hypointense margin line of demarcation. In order to evaluate the influence of the thickness of each cut on the measurement of volume, images with cuts of differing thickness of between 2 and 5 mm slices were used.

The area of the osteonecrosis was controlled visually on each MRI slice. For each coronal slice, the area of necrosis was considered to be the sector demarcated by the serpiginous line corresponding to the band-like hypointense margin. The inner border of this low intensity was assumed to represent the edge of the necrotic area. Electronic data were transferred to a multiprocessing computer and the area of the osteonecrosis was outlined using an image-analysis program (public-domain National Institutes of Health Image program).\(^7,8\) Although the MR image is two-dimensional, it provides information from a three-dimensional slab. For each slice, the volume of necrosis was calculated by multiplying the area of necrosis by the thickness of the slice.\(^9\) The total volume of necrotic bone was the sum of the individual volume of each slice. In order to assess the validity of the method of measurement, hips were evaluated independently by both authors. Each observer measured...
each volume twice, with an interval of at least one month between the first and second measurements.

Anatomical measurement of the volume of the osteonecrosis. After THA the femoral heads were cut into slices 5 mm thick (Fig. 1). In each cut, the separated osteonecrotic material and the cartilage were removed, and each observer measured the volume of necrotic material using fluid displacement.

We correlated the MR scans and the macroscopic appearance of the area of osteonecrosis with the histological features of the femoral head. At microscopy, four different structures were identified in the medullary space: bone marrow, fibrovascular tissue, disrupted fibrovascular tissue and amorphous necrotic material. We compared the histological appearance of each slide with the corresponding area of osteonecrosis outlined on the MR slice, and the histological appearance of the rest of the femoral head.

Reproducibility and accuracy of the method of measurement of the volume of osteonecrosis in stage-I and stage-II disease. There were no specimens available for anatomical measurement in the precollapse stage I or stage II. We assumed that the method of MRI was as good a means of measurement at stage II as at stage I. We therefore used the ten hips which had been evaluated at stage I or stage II and then at stage III to assess the validity of the method of measurement for stages I and II. The volume of the osteonecrosis obtained by MRI before collapse was compared with that after collapse by MRI and by anatomical methods.

Statistical analysis. Comparative relationships between different methods of measuring the volumes were examined using the Spearman rank-correlation test. When the relationship was significant ($p < 0.05$), a linear regression analysis was used to explore the relationship and the result given by the value of the regression coefficient with a 95% confidence interval (CI).

Assuming that the data from the anatomical measurements were the true values, the accuracy of the MRI measurements was verified by examination of the error and the error ratio. For example, the error for the MRI method

\[
\text{error} = (\text{volume with MRI}) - (\text{volume measured directly by anatomical measurement})
\]

and the error ratio of the MRI method (%)

\[
\text{error ratio} = \left( \frac{\text{error of MRI method}}{\text{volume measured directly on anatomical specimens}} \right) \times 100
\]

The difference, the percentage deviation and the systematic percentage deviation, between the two methods of measurement, MRI measurement and anatomical measurement, or between each of the MRI estimations, was calculated and assessed using the Wilcoxon signed-rank test with 95% CI. The Wilcoxon signed-rank test was also used to assess whether the deviation between two different methods of measurement was significantly greater than that between the first and second MRI estimations made by each author.

Results

The histological study confirmed that the serpiginous line, corresponding to the band-like hypointense line on MRI, consisted of fibrovascular tissue, disrupted fibrovascular tissue and amorphous necrotic material. This line is considered to be the boundary of the repair process, confirming that the material removed was indeed osteonecrotic. No osteonecrotic bone was found in the rest of the femoral head.

The lesion was bilateral in ten patients. No differences were found between the volumes of osteonecrosis in the right and left hips of these patients. For analysis of the variation between the two sides, and to be certain that each hip could be analysed as if it was from a different patient, a scatterplot of the mean values of the volume of osteonecrosis and the absolute difference between the two sides was examined visually. This did not reveal an association between the mean volume of each pair and the difference between the two volumes of the pair. Thus, these hips were analysed as if they were independent hips from different patients.

The variability between the observers was determined by defining a tolerance interval using an analysis of variance. With the MRI method it was $0.25 \text{ cm}^3$ (95% CI -2.5 to
The mean difference between the first and second observation by each observer was measured as 0.05 cm³ (95% CI -0.05 to +0.15). We concluded that there were no effective differences, and estimated the interobserver and intraobserver variance components to be 0.157 and 0.065, respectively. With this assumption, the difference between two different readings made by the same observer had a variance of $2 \times (0.157 \pm 0.065) = 0.444$, giving an SD of 0.666 cm³. The difference between two different readings made by different observers had a variance of $2 \times (0.157 \pm 0.065) = 0.444$, giving an SD of 0.666 cm³. We therefore expect that 95% of intraobserver differences will have a magnitude of no more than 1.7 cm³, whereas 95% of interobserver differences will have a magnitude of no more than 1.7 cm³. Thus, it is 95% certain that for 95% of the time one observer’s reading is no more than 2 cm³ greater or less than another observer’s reading, attributable to observer error alone.

The same calculations were made using the anatomical method of measurement, and the results were similar. It is therefore 95% certain that 95% of the time one observer’s reading is no more than 1 cm³ greater or less than another observer’s reading because of observer error alone.

Based on the anatomical method, the mean volume of the 30 specimens of osteonecrosis was 27.2 cm³ (13 to 45) (Table I). The volume of osteonecrosis obtained by the MRI method with 2 mm slices was only 0.98% lower than that from anatomical measurements when the values of all the necrotic volumes were added (Table I). These values were not significantly different (p = 0.059). Using the Spearman rank test the relationship between volumes measured by the MRI method (2 mm slices) and by the anatomical method, was close (r = 0.998). Regression analysis relating the osteonecrotic volumes calculated by the MRI method to the volumes obtained by direct anatomical measurement, showed a strong linear relationship (r = 0.98, CI 0.95 to 1.01).

Comparison of the volumes obtained with coronal slices on the first reading using the MRI method with 2 mm slices and with those obtained on the second reading with 5 mm slices showed a close relationship (r = 0.975 and p = 0.000 001) between these volumes, and a strong linear relationship was obtained by regression analysis (r = 1.04, CI 0.98 to 1.10). With the Wilcoxon matched-pair test, the level of statistical significance (Table I) was 0.072. The mean volume difference between the anatomical and the MRI measurements (2 mm slices) was 0.32 cm³ (95% CI -0.23 to +1.87) and that between the first and the second estimations with MRI was 1.01 cm³ (95% CI -0.15 to +1.87) (Table I). The errors between the two MRI estimations (5 and 2 mm slices) were not significantly greater than those between MRI (2 mm slices) and anatomical measurements (Table I). The level of statistical significance of these comparisons was 0.124 for the difference, 0.069 for the absolute percentage deviation and 0.074 for the systematic deviation (Table I).

When comparing the volume obtained for the same ten hips at different stages (before and after collapse), the values showed no significant difference when the volumes obtained with MRI before and after collapse were compared with the anatomical measurement (Wilcoxon signed-rank test) (p = 0.084 and p = 0.075. Regression analysis relating the volumes calculated by the MRI method before collapse to those obtained after collapse, showed a strong linear relationship (r = 1.08, 95% CI 0.98 to 1.18). The mean volume difference between MRI after collapse and MRI before collapse was 0.92 cm³ (95% CI -0.12 to +1.96). This difference was not significant (p = 0.064).

### Discussion

Several authors have used MRI to assess the extent of osteonecrosis by different methods, for example, a visual estimation of the percentage of osteonecrosis in quartiles and volume estimation using 2 cm slices. Although these methods are useful and allow indirect estimation of osteonecrosis, they are not a precise measurement of volume. These authors have shown that collapse is closely related to the extent of the lesion in osteonecrosis of the femoral head. We have tested the accuracy of MRI in quantifying the volume of osteonecrosis with 2 and 5 mm slices.

Because of subtle variations in the magnetic field, geometrical distortions may occur in MRI and the limits of the lesion may be difficult to outline precisely. We also measured the volume of necrosis macroscopically on surgical specimens. Using anatomical sections and several different techniques, it was possible to measure the volume of the infarct. A potential source of error in the measurement of the surgically retrieved femoral heads was that destructive methods (anatomical slices) were used and it is possible that some changes in the volume of necrosis occurred after sectioning and during freezing and thawing of the specimens. The histological study confirmed that the lesion measured was necrotic and corresponded to the MRI image. Even if no necrotic bone was found in the remaining part of each head, this did not mean that cell death was absent in areas where it did not appear on MRI. Our aim was only to verify that the volume of the lesion seen on MRI could be accurately measured. We used only the MRI signal given by the T1 relaxation time. Theoretically, T1

<table>
<thead>
<tr>
<th>Table I. Comparison of the mean (± sd) volumes of osteonecrosis on 30 femoral heads (stage III) determined anatomically and by MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume measured (cm³)</strong></td>
</tr>
<tr>
<td>Vol 2a, MRI 2 (2 mm slices)</td>
</tr>
<tr>
<td>Vol 2b, MRI 2 (5 mm slices)</td>
</tr>
<tr>
<td>Vol 1, anatomical measurement</td>
</tr>
<tr>
<td><strong>Absolute percentage deviation</strong></td>
</tr>
<tr>
<td>(Vol 2a - Vol 1): Vol 1</td>
</tr>
<tr>
<td>(Vol 2b - Vol 2a): Vol 2a</td>
</tr>
<tr>
<td><strong>Systematic deviation</strong></td>
</tr>
<tr>
<td>Vol 2a: Vol 1</td>
</tr>
<tr>
<td>Vol 2b: Vol 2a</td>
</tr>
</tbody>
</table>

THE JOURNAL OF BONE AND JOINT SURGERY
images do not distinguish between dead and living bone. The characteristic MR signal behaviour of active repair tissue was used to differentiate repair tissue from necrotic marrow, and was compared with the histological findings. According to our histological examination, the volume measured anatomically and the volumes obtained with MRI represented unrepaired necrotic marrow. Measurement of the volume of the osteonecrosis, before or after collapse, did not change significantly.

The volume of the osteonecrosis may theoretically be greater before collapse. In our series, the mean volume measured on MRI after collapse was 1.08% greater than that measured on MRI before collapse, which may reflect further damage during separation of the infarcted fragment. An adequate imaging protocol with sufficient resolution is required, and the accuracy of measurement is critically dependent on image processing. Our results suggest that the volume of osteonecrosis can be accurately determined with 2 or 5 mm slices by MRI performed before or after collapse.

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

References