99mTc-HMPAO-leukocyte scintigraphy in patients with symptomatic total hip or knee arthroplasty: improved diagnostic accuracy by means of semiquantitative evaluation

This is the author's manuscript

Original Citation:

Availability:
This version is available http://hdl.handle.net/2318/38221 since

Terms of use:

Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.
The aim of this study was to evaluate the diagnostic value, in suspected infectious prostheses, of 99mTc-labeled hexamethylpropyleneamine oxime (99mTc-HMPAO) leukocyte scintigraphy interpreted with the addition of a semiquantitative analysis.

Methods: By means of a retrospective review, we included a group of 78 consecutive patients with suspected hip or knee prosthesis infection. We performed 91 99mTc-HMPAO-leukocyte scintographies and examined 95 localizations that were suspect. Images were acquired at 3 different time points after the injection of the labeled leukocytes: 50 min (early images), 4 h, and 24 h (late images). The scintigraphic examinations were independently evaluated by 3 observers; qualitative and semiquantitative analyses were performed. The final diagnosis of infection was based on surgical, histologic, and bacteriologic data and follow-up.

Results: On qualitative analysis, sensitivity, specificity, and accuracy were 80.4%–87%, 65.3%–71.4%, and 75.8%–77.9%, respectively. On semiquantitative analysis, sensitivity, specificity, and accuracy were 95.6%, 95.8%, and 95.8%, respectively. The analysis of 95% confidential intervals showed statistically significant differences in specificity and accuracy between semiquantitative and qualitative analyses.

Conclusion: In those patients who underwent 99mTc-HMPAO-leukocyte scintigraphy for suspected hip or knee prosthesis infection, the addition of a semiquantitative evaluation to the qualitative analysis of early and late images leads to a significant improvement in both specificity and accuracy.

Key Words: infection radionuclide imaging; hip prosthesis; knee prosthesis; 99mTc-labeled leukocytes

alone, interpreted with the addition of a semiquantitative analysis.

MATERIALS AND METHODS

Study Population

In this retrospective study, we included only patients with symptomatic total hip or knee replacement who underwent $^{99m}$Tc-HMPAO-leukocyte scintigraphy in our institute between January 1999 and January 2002. All patients were studied after written informed consent was obtained. We studied 78 patients (42 females, 36 males; mean age $\pm$ SD, 69.5 $\pm$ 10.7 y; age range, 30–87 y). Of these patients, 4 had bilateral symptomatic prostheses; thus, a total of 82 prostheses were included in the analysis (47 total hip replacements, 35 total knee replacements). Some patients (13) were evaluated twice; therefore, a total of 95 prosthesis sites were analyzed. Antibiotic therapy was withdrawn at least 20 d before the study. We evaluated laboratory parameters (white blood cell [WBC] count, percentage of neutrophils, and erythrocyte sedimentation rate [ESR]) in all patients before scanning. Patients were divided in 2 groups: (a) hip group, including patients with total hip replacement (55 sites); and (b) knee group, including patients with total knee replacement (40 sites). The reasons for the original joint replacement were arthritis, rheumatic arthritis, trauma, and congenital hip abnormalities. Symptoms—mainly pain on motion or at rest, leading to a suspicion of either infection or prosthesis loosening—were the typical reason for the original admission. In 47 prostheses, further surgery was performed after nuclear imaging, which allowed comparison of the images with the intraoperative samples obtained for bacterial culture. Prostheses were considered infected if tissue cultures grew bacteria or in case of a grossly purulent wound at surgery. Prostheses were considered uninfected in the absence of purulence at surgery and negative operative samples. In those cases in which no further surgery was performed (35 prostheses), the clinical outcome was carefully monitored for at least 1 y after the imaging procedure (range, 1–3 y). On the basis of clinical parameters, we considered as not infected those patients for whom antibiotic treatment was not required at the end of follow-up. On the other hand, we considered as infected those patients who required a justified antibiotic therapy during follow-up.

$^{99m}$Tc-HMPAO-Leukocyte Scintigraphy

Blood samples were collected for leukocyte labeling. The leukocyte fractions were labeled with $^{99m}$Tc-HMPAO (Ceretec; Amersham) using a dose of 740 MBq $^{99m}$Tc and standard techniques (5,17). The labeling yield of the WBCs was 60%–80%. The leukocytes were reinjected intravenously into each patient with a delay of no more than 2 h 30 min; the administered activity ranged was 430–600 MBq.

The $^{99m}$Tc-labeled WBCs were acquired in anterior and posterior views of the pelvic and thigh area (patients with hip prostheses) or in anterior and posterior views of the pelvic and knee area (patients with knee prostheses) at 3 different time points after injection: 50 min (early images), 4 h, and 24 h (late images).

The images were acquired using an imaging time of 7 min in the early and 4-h images, while acquisition time was 15 min for the late images. Matrix size was $256 \times 256$; a high-resolution collimator was used. Imaging was performed on a double-head $\gamma$-camera (Millennium VG; General Electric Medical Systems).

Image Interpretation

Qualitative Analysis. Three experienced nuclear medicine physicians (designated as observers 1, 2, and 3) independently assessed the studies, unaware of the clinical data. Evaluation criteria were as follows: The scans were considered to be positive when abnormal localized activity in the region of interest (ROI) increased in intensity or in extension in comparison with the contralateral region or with the ipsilateral adjacent bone segment. A scan was considered to be negative when the local activity was identical to that present in the contralateral bone segment or the ipsilateral adjacent bone segment or when abnormal localized activity in the ROI decreased or did not increase with time in intensity or extension in comparison with the contralateral region or the ipsilateral adjacent bone segment (5,17). The studies were considered to be ambiguous when there was no concordance among the 3 observers (ambiguous subgroup); on the other hand, we considered the studies as unambiguous when complete agreement among the observers was present (unambiguous subgroup).

Semiquantitative Analysis. To semiquantitatively evaluate the activity of labeled WBCs in the area of the symptomatic prosthesis, 2 circular ROIs were drawn: (a) over the prosthesis regions suspected of infection (suspected region [SR]) and (b) over the left iliac crest, on the anterior projection (reference region [RR]; region representative of the bone marrow). SR and RR ROIs were drawn both in early and in late images. Mean counts per pixel in these ROIs were recorded and used to determine the SR-to-RR ratio ($K = SR/RR$) for the early and late images ($K_{\text{early}}$ and $K_{\text{late}}$, respectively). Because patient position is not exactly the same in different imaging sessions, to semiquantitatively compare early and late images, we used the SR-to-RR ratios, not the mean counts per pixel in the SR, corrected for time decay and acquisition time (Fig. 1). When $K$ increased with time ($K_{\text{late}} > K_{\text{early}}$ by at least 10%), we considered the scan to be indicative of infection (Fig. 2); the scan was considered to be negative when $K_{\text{late}}$ was similar or decreased with respect to $K_{\text{early}}$ (Fig. 1). As reported by others (11,12,18), because leukocytes are usually taken up by the bone marrow, the detection of infection is complicated by normal bone marrow activity. The early and late SR-to-RR ratios led us to evaluate SR with respect to bone marrow behavior and to differentiate between infection and ectopic bone marrow. When $K_{\text{early}}$ was similar to $K_{\text{late}}$, we related the periprosthetic uptake to the presence of bone marrow; on the other hand, when $K_{\text{late}}$ was higher than $K_{\text{early}}$ by at least 10%, infection was strongly suspected.

Statistical Analysis

In patients with infected and uninfected prostheses, we used the Student $t$ test for unpaired data to compare the mean values of the WBC count, percentage of neutrophils, and ESR. The evaluation of qualitative and semiquantitative analysis of leukocyte scintigraphy in comparison with the clinical or instrumental follow-up was performed by traditional methods based on sensitivity, specificity, accuracy, and the relative 95% confidence intervals (CIs). The concordance among observers was evaluated with the $\kappa$-statistic and reported with the relative 95% CI. We computed the 95% CI associated to each indicator, with the exact Wilson method (19). A $\chi^2$ test was performed to compare the prevalence of ambiguous and unambiguous scans in the hip and knee groups.

RESULTS

A total of 91 $^{99m}$Tc-HMPAO-leukocyte scintigrapghies were performed and 95 prosthesis sites were analyzed (55
symptomatic total hip replacements, 40 symptomatic total knee replacements). In 47 of the 95 cases (49.5%), the clinical or microbiologic investigations confirmed the presence of an infection. The WBC count (3,500–13,500/mL), percentage of neutrophils (35%–85%), and ESR (5–125 mm/h), evaluated in all cases within 3 days of the examination, did not significantly differ in patients with infected and uninfected prostheses (P > 0.05; Student t test).

Qualitative Analysis

Table 1 reports the sensitivity, specificity, and accuracy and the relative 95% CI of each observer. The accuracy in reporting was nearly the same in all observers (77.9%, 75.8%, and 76.8%). Table 2 reports the Cohen κ-measure-

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>True-positive</td>
<td>39</td>
<td>40</td>
<td>37</td>
</tr>
<tr>
<td>True-negative</td>
<td>35</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>False-positive</td>
<td>14</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>False-negative</td>
<td>7</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>84.8</td>
<td>87.0</td>
<td>80.4</td>
</tr>
<tr>
<td>95% CI</td>
<td>72–92</td>
<td>74–94</td>
<td>67–89</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>71.4</td>
<td>65.3</td>
<td>73.5</td>
</tr>
<tr>
<td>95% CI</td>
<td>58–82</td>
<td>51–77</td>
<td>60–84</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>77.9</td>
<td>75.8</td>
<td>76.8</td>
</tr>
<tr>
<td>95% CI</td>
<td>69–85</td>
<td>66–83</td>
<td>67–84</td>
</tr>
</tbody>
</table>
ment of interobserver agreement with the relative 95% CI. In all cases, there was moderate to substantial agreement. In 69 of the 95 cases (72.6%), a complete concordance was present among the 3 observers (unambiguous subgroup). Figure 3 shows the imaging results of a typical infected hip prosthesis. Periprosthetic activity is evident around the hip replacement (Fig. 3, arrows). The sites of infection are clearly visible, particularly on images obtained at 24 h. We did not have complete concordance among the 3 observers in 26 cases (27.3%; ambiguous subgroup). Figure 4 shows the imaging results of an ambiguous case. The intertrochanteric activity of the hip replacement does not clearly change with time.

**Semiquantitative Analysis**

The sensitivity, specificity, and accuracy and the relative 95% CIs obtained with semiquantitative analysis in our group of patients and in the ambiguous and unambiguous subgroups are reported in Table 3. Analysis of the 95% CIs showed no statistically significant differences of each parameter in both subgroups (Table 3). In all cases, semiquantitative analysis improved the level of sensitivity, specificity, and accuracy of each observer on qualitative analysis. Evaluation of the 95% CIs showed statistically significant differences in specificity and accuracy between semiquantitative and qualitative evaluation. Figure 1 presents the semiquantitative analysis of early and late images of Figure 4. $K_{late} = K_{early}$; therefore, the scan is considered to be negative; on follow-up, this was proven to be a true-negative case. Figure 2 presents an example of an infected knee prosthesis that is a true-positive case.

**Hip and Knee Groups**

Similar results were observed in both hip and knee groups. Tables 4 and 5 report the sensitivity, specificity, and accuracy obtained with semiquantitative and qualitative analysis. The prevalence of ambiguous cases due to the incomplete concordance among the 3 observers in the hip group (23.6%; 13/55) and in the knee group (32.5%; 13/40) was not statistically different ($\chi^2$, 0.246; $P = N/S$). In both hip and knee groups, semiquantitative evaluation improved the level of sensitivity, specificity, and accuracy that each observer had reached on qualitative analysis.

**DISCUSSION**

Several authors have reported that the rate of infection after primary implantation of either total hip or total knee replacement ranges from 1% to 4% and that the rate is higher after revision procedures (20–22). Treatment of infection after total hip and total knee arthroplasty is costly because of the subsequent need for further surgery and the prolonged hospitalization often required to eradicate the infection. Because of an aging population that will need an increasing number of arthroplasties, methods to prevent, diagnose, and treat infection must be perfected to reduce the social costs of total hip and total knee arthroplasties. How-

<table>
<thead>
<tr>
<th>Statistic</th>
<th>1 vs. 2</th>
<th>1 vs. 3</th>
<th>2 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>0.741</td>
<td>0.598</td>
<td>0.553</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.71–0.77</td>
<td>0.56–0.63</td>
<td>0.52–0.59</td>
</tr>
</tbody>
</table>
ever, infection after total hip and total knee arthroplasty is often a diagnostic challenge. No test is 100% sensitive and 100% specific. Thus, the diagnosis of infection relies on the physician’s judgment of clinical and instrumental aspects. The consequences of misdiagnosis are considerable. Reimplantation of a prosthesis into an infected tissue bed, without appropriate debridement, is likely to result in persistent infection.

$^{99m}$Tc-Methylene diphosphonate (MDP) bone scintigraphy has been the most extensively studied and undoubtedly the most frequently performed radionuclide study for approaching a painful joint replacement. For both total hip and total knee prostheses, a bone scan is defined as normal when the periprosthetic uptake is indistinguishable from adjacent nonarticular, presumably normal, bone. However, difficulties associated with $^{99m}$Tc-MDP bone scans include multiple conditions—such as fractures, tumors, heterotopic ossification, and inflammatory disorders—that can result in an increased uptake in the periprosthetic tissue. Moreover, scans can remain positive for as long as 1 y after an uncomplicated hip replacement and for >2 y after insertion of a prosthesis without cement; most importantly, these scans cannot be used to differentiate between infection and an aseptic loosening (23–25).

$^{67}$Ga scintigraphy has proven to be a useful addition to bone scintigraphy in patients with painful joint replacements; the accuracy of combined $^{99m}$Tc-MDP bone/gallium scintigraphy for detecting joint replacement infection is nearly 80% (26).

$^{111}$In- or $^{99m}$Tc-HMPAO-leukocyte scintigraphy has also been evaluated alone and in combination with other radiotracers. Generally, when used alone, interpretation of the study has been based on either the intensity of periprosthetic labeled leukocyte activity in comparison with various reference points (26) or a grading system of periprosthetic uptake (27). The reported sensitivity of the procedure using these criteria was 50%–100% and the specificity was 45%–100%; in his review, Schauwecker reported an average $^{99m}$Tc-HMPAO sensitivity of 87% and an average specificity of 81% (14). These values are similar to those obtained in our work at qualitative evaluation.

As reported by others, the relatively low levels of sensitivity, specificity, and accuracy of labeled WBC scintigraphy are mainly a consequence of the accumulation of leukocytes in the reticuloendothelial bone marrow (12,18). In fact, although the normal distribution of bone marrow in adults up to 70 y old generally includes the axial skeleton, humeral heads, femoral trochanters, and, frequently, the femoral shafts, individual variation is considerable (28). Moreover, the implantation of a prosthetic device may produce additional variation. Changes in the normal distribution of bone marrow through surgical manipulation has been demonstrated in animals (11,28).

### TABLE 3

<table>
<thead>
<tr>
<th>Analysis</th>
<th>All studies $(n = 95)$</th>
<th>Unambiguous $(n = 69)$</th>
<th>Ambiguous $(n = 26)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True-positive</td>
<td>44</td>
<td>36</td>
<td>8</td>
</tr>
<tr>
<td>True-negative</td>
<td>47</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>False-positive</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>False-negative</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>95.7</td>
<td>97.3</td>
<td>88.9</td>
</tr>
<tr>
<td>95% CI</td>
<td>85–99</td>
<td>86–100</td>
<td>57–98</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>95.9</td>
<td>93.8</td>
<td>100</td>
</tr>
<tr>
<td>95% CI</td>
<td>86–99</td>
<td>80–98</td>
<td>82–100</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>95.8</td>
<td>95.7</td>
<td>96.2</td>
</tr>
<tr>
<td>95% CI</td>
<td>90–98</td>
<td>88–99</td>
<td>81–99</td>
</tr>
</tbody>
</table>

FIGURE 4. $^{99m}$Tc-HMPAO-leukocyte scintigraphy in suspected left hip prosthesis infection: ambiguous case of infection (same as Fig. 1). The intertrochanteric activity of the hip replacement does not clearly change with time (arrows).
To improve labeled leukocytes’ diagnostic accuracy in prosthesis infection, $^{99m}$Tc-labeled sulfur colloid bone marrow imaging addition was suggested for the first time in 1986 by Lavender et al. (29). The physiologic distribution of $^{99m}$Tc-labeled sulfur colloid and labeled leukocytes in bone marrow is similar. However, infection exerts opposite effects on leukocytes and sulfur colloid. While stimulating leukocyte accumulation, bone infection has been reported to decrease sulfur colloid accumulation; these inverse effects result in incongruent images and allow discrimination between infection and unusual, but not abnormal, periprosthetic bone marrow distribution (10–12). The reported sensitivity, specificity, and accuracy of labeled leukocytes plus sulfur colloid bone marrow imaging is 96%, 97%, and 97%, respectively (11)—that is, superior to imaging with labeled leukocytes alone or in combination with routine bone scintigraphy (11,12).

However, although combined labeled leukocyte/bone marrow imaging has greatly improved the accuracy of leukocyte scintigraphy alone in the detection of osteomyelitis adjacent to a loose or painful joint arthroplasty and it is still considered to be the gold standard technique for the diagnosis of infected joint prosthesis, too much time is required to complete the diagnostic process and the costs are high. Therefore, in our work, we studied the possible improvement in accuracy of $^{99m}$Tc-HMPAO-leukocyte scintigraphy alone with the addition of a semiquantitative analysis. The mean counts per pixel in the area of suspected infection and in the left iliac crest (i.e., an area representative of bone marrow behavior) were recorded and SR-to-RR ratios were calculated. As reported by Peters, because bone marrow margiates leukocytes to a considerable degree, comparing the early images after labeled leukocyte injection with late images could be an alternative to bone marrow scanning (18). The early images still do not reflect significant leukocyte uptake at the site of infection and can be considered the expression of the presence of bone marrow (11,18); thus, the incongruence between the early and late images indicates the presence of infection and can be considered analogous to that between sulfur colloid and leukocyte scans (18).

In our study, we evaluated the changes in the ratio between early and late imaging, assuming that when the periprosthetic uptake of labeled leukocytes is related to the presence of eiterotopic bone marrow, its behavior is similar to that present in the iliac crest and the ratio does not change with time. On the other hand, in the presence of prosthesis infection, the ratio tends to increase ($K_{\text{late}} > K_{\text{early}}$). This leads us to a final diagnosis with a sensitivity, specificity, and accuracy of 95.6%, 95.9%, and 95.8%, respectively. Our results are similar to those presented by others with combined labeled leukocyte/$^{99m}$Tc-labeled bone marrow imaging (4,11).

**CONCLUSION**

In $^{99m}$Tc-labeled leukocyte scintigraphy, the addition of a semiquantitative evaluation of early and late images to the qualitative analysis leads to a significant improvement of specificity and accuracy. Furthermore, in our study, semiquantitative evaluation of late and early images allows one to obtain levels of sensitivity, specificity, and accuracy similar to those obtained by others using both labeled leukocytes and sulfur colloid bone marrow imaging.

**ACKNOWLEDGMENTS**

We thank Vincenzo Arena, MD, and Stefania Bergese, MD, for their precious contribution and Francesca Giunta, MD, for help with the use of English in the revision of the manuscript.
REFERENCES