Reproducibility of Image Interpretation in Immunoscintigraphy Performed with Indium-111- and Iodine-131-Labeled OC125 F(ab')$_2$ Antibody Injected into the Same Patients

Jean-Philippe Vuillez, Patrick Peltier, Jean-Claude Mayer, Francis Dubois, Alain Chetanneau, Pierre Fumoleau, Alain Bertrand, and Jean-François Chatal

Centre René Gauducheau, Nantes, France; Centre Hospitalier Régional, Nancy, France; and Centre Hospitalier Régional, Grenoble, France

An important criterion for the clinical use of a new imaging technique is the correct reproducibility of interpretation. Forty-six paired immunoscintigraphic examinations were performed on 43 patients with suspected ovarian carcinoma recurrence using F(ab')$_2$ fragments of OC125 antibody labeled first with indium-111 and then with iodine-131. Planar scintigraphy (PS) and emission computed tomography (ECT) images were interpreted blindly and separately by three observers, and reproducibility was evaluated by a kappa concordance index. Intra- and interobserver reproducibility were generally satisfactory (κ values of 0.6 and 0.7, respectively). Binomial analysis of κ values for ECT showed the superiority of indium-111 for intraobserver ($p = 0.035$) and interobserver ($p = 0.0039$) study. However, for PS there was no significant difference in reproducibility with the two radionuclides.

J Nucl Med 1991; 32:221–226

The advantages of immunoscintigraphy, an approach based on specific recognition of a defined antigen target, have been demonstrated in numerous studies, particularly in oncology (*1,2*). As with any new imaging technique, problems have been encountered in the development phase, not all of which have yet been definitively solved. These concern immunology, hemodynamics, pharmacokinetics, and methodology (*3,4*). Among the methodologic problems, the choice of the radionuclide used in labeling antibodies or their fragments has an effect on the pattern of the images obtained and the results of the examination (*5–7*). Iodine-131 (^{131}I) was initially used for labeling (*8*), but there are now various theoretical and experimental reasons for preferring indium-111 (^{111}In) (*7,9–14*). However, the superiority of immunoscintigraphy performed with an antibody labeled with ^{111}In rather than ^{131}I has not yet been clearly demonstrated in clinical practice.

One limitation in routine clinical application of immunoscintigraphy is the difficulty in interpreting images when tumor-to-nontumor contrast is moderate. An important criterion for the clinical use of an examination is the correct reproducibility of interpretation by the same observer or different observers. The purpose of this study was to evaluate and compare the reproducibility of interpretation of immunoscintigraphic images obtained with F(ab')$_2$ fragments of OC125 antibody (*15*) labeled first with ^{111}In and then (immediately afterward) with ^{131}I.

MATERIALS AND METHODS

Patients

Forty-six paired immunoscintigraphic examinations were performed on 43 patients: 7 times retrospectively (known recurrence site) and 39 times prospectively. For the latter, the indication of the examination was 31 times isolated elevation of circulating CA125 concentration. Mean age was 59.6 ± 8.8 yr (range 34–82). Mean CA125 concentrations on the day of antibody injection was 1.106 ± 2.451 U/ml (range 10–14,500).

Radioantibodies

All examinations were performed with F(ab')$_2$ fragments of OC125 monoclonal antibody labeled with ^{131}I or ^{111}In. The labeling methods with both radionuclides have been previously reported (*8,9*). Briefly, an iodogen method was used with ^{131}I and the method of Hnatowich et al. (*16*) employing the bicyclic anhydride of DTPA with ^{111}In. Specific activity was on average 111 MBq/mg (3 mCi/mg) with both radionuclides.

Immunoreactivity was tested by affinity chromatography using a sandwich assay (*8*). The percentage of immunoreactivity with intact monoclonal antibody was 55%–60% after both ^{131}I and ^{111}In labeling. It was 50%–55% with ^{131}I- and ^{111}In-OC125 F(ab')$_2$ fragments. The percentage of immunoreactivity tested by cell-binding assay using the NIH OVCAR

Received Feb. 28, 1990; revision accepted Aug. 14, 1990.

For reprints contact: Jean François Chatal, Service de Médecine Nucléaire, Centre René Gauducheau, Quai Moncousu, 44000 Nantes, France.

3 serous ovarian adenocarcinoma cell line was 50% in each case.

The radiochemical purity (tested with gel-filtration) ranged from 80% to 90% with ^{131}I-OC125 F(ab')$_2$ and was always >90% with ^{111}In-OC125 F(ab')$_2$.

Immunoscintigraphic Protocol

The pelvis and abdomen were explored in all patients 3 days after injection of radiolabeled OC125 F(ab')$_2$. Planar scintigraphy (PS) in anterior and posterior views and elliptical 360° emission computed tomography (ECT) were performed. Acquisition times were 10 min in the planar mode and 40 min in the ECT mode. All data were acquired with a 64 × 64 word-mode matrix. From ECT acquisitions, 6-mm thick contiguous sections were reconstructed by retroprojection using a Wiener filter in the frontal, sagittal, and transverse planes.

To obtain anatomical landmarks (visualization of bone structures, kidneys and the urinary bladder), 185 MBq (5 mCi) of technetium-99m-labeled hydroxymethylene diphosphonate (^{99m}Tc-HMDP) were injected 2 hr before recordings were done. The examinations were performed with a single-headed tomographic Sophy camera (Sopha Médical, France), allowing simultaneous acquisition (with satisfactory anatomical superpositioning) of the landmark and radioantibody distribution images. A 20% window and a high-energy collimator were used for images of ^{131}I-labeled antibody distribution. For images of ^{111}In-labeled antibody distribution, both energy peaks (173 and 247 keV) were taken into account (20% and 10% windows), and a medium-energy collimator was used. For anatomical landmark images obtained at the same time (with the same collimation as for antibody distribution images), a symmetrical 20% window of the ^{99m}Tc gamma peak was used.

After informed consent of patients was obtained, each examination was performed with antibody labeled with ^{111}In and then with ^{131}I according to the following protocol:

Day 0: Thirty-minute i.v. injection of ^{111}In-DTPA-OC125 F(ab')$_2$ fragments diluted in 100 ml of NaCl 99%. Mean injected activity was 118.4 MBq (3.2 mCi) (range: 92–167 MBq).

Day 3: Scintigraphic recording of anatomical landmark (^{99m}Tc-HMDP) and ^{111}In-labeled antibody distribution images, followed by injection of ^{131}I-OC 125 F(ab')$_2$ fragments, also diluted in 100 ml of NaCl 99%. Mean injected activity was 109 MBq (2.94 mCi) (range: 63–122 MBq). Thyroid uptake of free ^{131}I was blocked by oral administration of a Lugol's solution (100 mg/day) or of perchlorate in case of iodine intolerance, during 8 days beginning 2 days before injection of the iodinated antibody.

Day 6: Scintigraphic recording of anatomical landmark (^{99m}Tc-HMDP) and ^{131}I-labeled antibody distribution images.

Image Interpretation

All images were interpreted blindly (without knowledge of clinical history or results of other examinations) by three observers (O1, O2, O3) working separately. O1 and O2 performed two independent interpretations at an interval of at least 30 days. O1 and O3 had good experience in interpreting immunoscintigraphic images, whereas O2 was being trained in this methodology.

All images obtained after injection of ^{111}In-OC125 were interpreted first, and then, after a mean interval of 8 days, all ^{131}I-OC125 images. For each examination, interpretation of planar images and reconstructed ECT sections was done separately for each radionuclide and for each topographic area (abdomen and pelvis). The limit between the abdomen and pelvis was empirically defined by a horizontal plane passing through the iliac crests visible on ^{99m}Tc-HMDP images. Within these areas (abdomen and pelvis), each observer had to score results using the following scale: 0 = clearly negative, 1 = probably negative, 2 = doubtful, 3 = probably positive, and 4 = clearly positive.

For visualization of images, a scale of 16 colors with 256 levels was used. In images obtained with ^{111}In-labeled antibody, it was generally necessary to saturate the liver area by lowering the upper threshold to obtain good contrast between the pathologic focus and background. This adjustment was performed empirically on the basis of experience. The threshold was set so as to provide approximate reproducibility (same color) in the imaging of bone structures (particularly the sacroiliac joints) from one examination to another.

In the ECT mode, the criterion for considering an uptake focus to be pathologic was its appearance in at least three successive sections and in at least two section planes.

Reproducibility Study

Inter- and intraobserver reproducibility studies were performed for each recording mode (PS and ECT): (a) for each of the anatomical areas (abdomen, pelvis); (b) in consideration of all of the 92 sites explored (by site); and (c) without taking individual sites into account but rather the total area explored (global). For this last analysis, the highest of the two scores (abdomen, pelvis) was retained.

Inter- and intraobserver reproducibility was evaluated by the weighted kappa (κ) concordance index of Landis and Koch (*17*) as modified by Kramer and Feinstein (*18*). This index (between 0 and 1) indicates a strength of agreement proportional to κ value (the empirical scale given by Landis and Koch is reproduced in Table 1) for which statistical significance can be calculated.

For interobserver reproducibility, the κ index was calculated between O1 and O2 and between O1 and O3 (O2 × O3 was not taken into account since it could not be made completely independent of O1 × O2 and O1 × O3). Global κ values were also calculated (three-rater generalization of kappa) (*19*).

Statistical Comparison Tests

The difference in reproducibility as a function of the radionuclide used was evaluated by applying the properties of binomial law. The percentage P of κ values of one radionuclide greater than those of the other was taken into account for the test. Three-rater kappa values were compared by means of a customary Z-test.

RESULTS

Interobserver Reproducibility (Table 2)

Kappa values were almost always very significant for both ^{131}I-labeled and ^{111}In-labeled antibody, which indicates that concordance of interpretation was not a

TABLE 1
Correspondence Between Kappa Values and Strength of Agreement Between Interpretations (According to Landis and Koch)

Value of κ	Strength of agreement
<0	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

matter of chance. Concordance was of variable intensity depending on acquisition mode, site and radionuclide but was generally satisfactory (κ values around 0.60). Kappa values for examinations performed with ^{111}In were generally higher than those with ^{131}I.

Planar Scintigraphy. Two-rater interpretation reproducibility was similar (κ around 0.50) for ^{111}In and ^{131}I images, although the only $\kappa > 0.60$ was obtained with ^{131}I. Iodine-131 κ values were better than those of ^{111}In 5/8 times (nonsignificant according to binomial law: $p = 0.363$). Three-rater κ values were significant but rather low and the difference between ^{111}In and ^{131}I was never significant. There was, thus, no difference in reproducibility in planar mode with the two radionuclides.

ECT. Regardless of analysis mode (pelvis, abdomen, by site, global), interpretation reproducibility of ^{111}In ECT images was superior to that of ^{131}I ECT images.

Two-rater ^{111}In κ values were better than those of ^{131}I 8/8 times (highly significant: $p = 0.0039$). Three-rater κ values were significant and higher for ^{111}In images; the difference was significant for the abdomen ($p < 0.001$), for all sites combined ($p < 0.05$), and for global analysis ($p < 0.0001$)

TABLE 2
Interobserver Reproducibility

		^{131}I		^{111}In	
		kappa ± s.d.	*p*	kappa ± s.d.	*p*
Planar					
Pelvis	O1 × O2	0.715 ± 0.144	<0.00001	0.506 ± 0.133	0.00007
	O1 × O3	0.526 ± 0.146	0.00015	0.520 ± 0.137	0.00007
	O1 × O2 × O3	0.381 ± 0.052	<0.00001	0.353 ± 0.064	<0.00001
Abdomen	O1 × O2	0.234 ± 0.133	0.0393	0.482 ± 0.174	0.00288
	O1 × O3	0.541 ± 0.135	0.00003	0.433 ± 0.17	0.00554
	O1 × O2 × O3	0.097 ± 0.061	>0.05	0.104 ± 0.092	>0.05
Global	O1 × O2	0.499 ± 0.058	0.00083	0.582 ± 0.127	<0.00001
	O1 × O3	0.468 ± 0.160	0.00171	0.540 ± 0.132	0.00002
	O1 × O2 × O3	0.287 ± 0.054	<0.00001	0.305 ± 0.052	<0.00001
By site	O1 × O2	0.506 ± 0.079	<0.00001	0.466 ± 0.094	<0.00001
	O1 × O3	0.593 ± 0.082	<0.00001	0.565 ± 0.096	<0.00001
	O1 × O2 × O3	0.288 ± 0.034	<0.00001	0.249 ± 0.054	<0.00001
ECT					
Pelvis	O1 × O2	0.547 ± 0.138	0.00004	0.623 ± 0.143	0.00001
	O1 × O3	0.545 ± 0.128	0.00001	0.593 ± 0.147	0.00003
	O1 × O2 × O3	0.356 ± 0.047	<0.00001	0.434 ± 0.059	<0.00001
Abdomen	O1 × O2	0.085 ± 0.162	>0.05	0.233 ± 0.124	0.02993
	O1 × O3	0.239 ± 0.185	>0.05	0.548 ± 0.135	<0.00001
	O1 × O2 × O3	0.075 ± 0.089	>0.05	0.455 ± 0.053	<0.00001
Global	O1 × O2	0.373 ± 0.134	0.00264	0.498 ± 0.161	0.00101
	O1 × O3	0.487 ± 0.126	0.00005	0.564 ± 0.156	0.00015
	O1 × O2 × O3	0.300 ± 0.035	<0.00001	0.696 ± 0.063	<0.00001
By site	O1 × O2	0.420 ± 0.084	<0.00001	0.477 ± 0.082	<0.00001
	O1 × O3	0.506 ± 0.086	<0.00001	0.606 ± 0.085	<0.00001
	O1 × O2 × O3	0.285 ± 0.042	<0.00001	0.405 ± 0.039	<0.00001

p = degree of kappa significance.
O1 × O2 and O1 × O3: weighted kappa values according to Landis and Koch (*17*).
O1 × O2 × O3: three-rater kappa values according to Fleiss (*19*).

Thus, the best interobserver reproducibility was obtained with ^{111}In ECT, and the other three recording modes (^{131}I PS, ^{111}In PS, and ^{131}I ECT) had lower and similar interpretation reproducibility. Using three-rater κ values, the comparison of ^{111}In ECT with ^{111}In PS and ^{131}I PS showed significant differences in reproducibility for the abdomen ($p < 0.001$), for all sites combined ($p < 0.05$), and for global analysis ($p < 0.001$).

Reproducibility was not as good between observers O1 and O2 as between O1 and O3. In particular, κ values between O1 and O2 for the abdomen were <0.03 for ^{131}I PS, ^{131}I ECT, and ^{111}In ECT.

Intraobserver Reproducibility (Table 3)

The different κ values for intraobserver reproducibility were also satisfactory at around 0.70. Out of 16 κ values calculated for ^{111}In, none was <0.50 (0.44) and 13 were >0.60 (one >0.80), whereas for ^{131}I, 3 values were <0.50 (one <0.20) and 8 > 0.60 (one >0.80). For ^{111}In, interpretation reproducibility was similar regardless of acquisition mode both in the pelvis and abdomen, whereas for ^{131}I, it was better for PS (1 κ value < 0.50 and 7 > 0.60) than for ECT (2 κ values < 0.40 and 1 > 0.60), especially for the abdomen.

In planar mode, ^{131}I κ values were better than those of ^{111}In 6/8 times (nonsignificant). In ECT, κ values of ^{111}In were better than those of ^{131}I 7/8 times (significant: $p = 0.035$). Indium-111 thus provided better intraobserver reproducibility in ECT. O2 × O2 reproducibility was less good overall than that of O1 × O1; this difference was more evident for examinations using ^{131}I.

DISCUSSION

The purpose of this study was to determine whether the theoretical advantages of ^{111}In over ^{131}I provide better reproducibility in interpretation of ^{111}In immunoscintigraphic images. It was logical to carry out the study in cases of suspected recurrence of ovarian cancer, which is currently the best indication of immunoscintigraphy. This technique contributes to earlier diagnosis of a recurrence, as has been shown in several studies using antibodies labeled with ^{131}I (*8,20–22*), ^{123}I (*23–30*), and more recently ^{111}In (*9,31,32*).

To provide valid comparison, the only solution not requiring a very large number of cases was to perform immunoscintigraphy twice successively in the same patients using the same antibody labeled with ^{111}In and then with ^{131}I and to interpret the results blindly. The second injection was performed less than a week (3 days) after the first in order to avoid interference from human antimouse antibodies, which could have biased comparative interpretation of results with the two radionuclides.

TABLE 3
Intraobserver Reproducibility

		^{131}I		^{111}In	
		kappa ± s.d.	*p*	kappa ± s.d.	*p*
Planar					
Pelvis	O1 × O1	0.747 ± 0.138	<0.00001	0.639 ± 0.134	<0.00001
	O2 × O2	0.667 ± 0.144	<0.00001	0.555 ± 0.142	0.00005
Abdomen	O1 × O1	0.699 ± 0.136	<0.00001	0.697 ± 0.151	<0.00001
	O2 × O2	0.473 ± 0.154	0.00106	0.723 ± 0.220	0.0005
Global	O1 × O1	0.758 ± 0.211	0.00016	0.823 ± 0.134	<0.00001
	O2 × O2	0.663 ± 0.150	<0.00001	0.617 ± 0.131	<0.00001
By site	O1 × O1	0.733 ± 0.079	<0.00001	0.683 ± 0.033	<0.00001
	O2 × O2	0.637 ± 0.079	<0.00001	0.617 ± 0.104	<0.00001
ECT					
Pelvis	O1 × O1	0.576 ± 0.136	0.00001	0.701 ± 0.147	<0.00001
	O2 × O2	0.693 ± 0.133	<0.00001	0.501 ± 0.137	0.00012
Abdomen	O1 × O1	0.343 ± 0.165	0.01913	0.731 ± 0.133	<0.00001
	O2 × O2	0.004 ± 0.162	0.48898	0.620 ± 0.127	<0.00001
Global	O1 × O1	0.580 ± 0.138	0.00001	0.628 ± 0.158	0.00004
	O2 × O2	0.515 ± 0.131	0.00004	0.545 ± 0.167	0.00057
By site	O1 × O1	0.569 ± 0.083	<0.00001	0.713 ± 0.085	<0.00001
	O2 × O2	0.506 ± 0.086	<0.00001	0.613 ± 0.082	<0.00001

p = degree of kappa significance.

Immunoscintigraphic reproducibility in our study was satisfactory overall; close to 0.60 for interobserver and 0.70 for intraobserver results. In terms of both inter- and intraobserver results, reproducibility was better for ^{111}In ECT than for ^{131}I ECT. Several factors contributed to this finding. First, the physical characteristics of ^{111}In are superior to those of ^{131}I for gamma camera detection. Second, higher tumor uptake is obtained with ^{111}In-labeled antibodies and background is lower since clearance from the bloodstream is faster than that with ^{131}I-labeled antibodies (*12*). With ^{131}I-labeled antibody, the patchy distribution of nonspecific activity is a cause of false images that can be considered either as abnormal foci or artifacts. However, ^{111}In labeling, based on coupling a chelating agent with the antibodies (*16*), has proved relatively stable. Tumor-to-background contrast is thus generally superior with ^{111}In, resulting in higher tumor-to-normal tissue image contrast (although with the inconvenience of high liver activity). With ^{111}In, the greater intensity and higher contrast of abnormal foci provide better and easier interpretation of ECT sections, which are therefore more reproducible (Fig. 1).

Although interpretation of ^{131}I images can be hindered by nonspecific gastrointestinal activity, little or no image processing is required. However, with ^{111}In, intense hepatic uptake requires saturation of the liver area in order to visualize a lesion. The level selected for the upper image threshold can cause artifacts. Although these have little effect on ECT images, they can lead to mistakes in interpreting PS and result in poor reproducibility. This explains why reproducibility of planar ^{131}I and ^{111}In images was paradoxically similar despite better tumor-to-tissue ratios with ^{111}In-labeled antibody. (Fig. 2).

Two particular types of discordance suggest that these reproducibility results could be improved for both radionuclides, especially for ^{111}In. First, when the image of an uptake focus is situated in the border region between the pelvic and abdominal areas, it can be assigned to one of these areas by a first observer and to the other by a second observer. The result is a strong discordance in opposite directions for the pelvis and abdomen. This discordance disappears in global analysis, which explains why κ values are then generally higher. Second, when involvement is diffuse, radioantibody uptake is distributed over the image with no localized focus. This diffuse image, as a function of the visualization parameters (threshold saturation), can be interpreted as positive or negative. This occurs especially with ^{111}In since liver activity requires modification of the upper image threshold. This situation clearly indicates the need to formulate standardized interpretation criteria.

Intraobserver reproducibility was less good for observer O2 than for O1, and interobserver reproducibility of O1 × O2 was less good than that of O1 × O3. These differences can be explained by the fact that at the time of the study O1 and O3, but not O2, already had a lot of experience in interpreting immunoscintigraphics images, pinpointing the need for experience and training, independently of the radionuclide and acquisition characteristics involved, if good interpretation is to be achieved. These differences with respect to observers were more apparent when ^{131}I was used, which suggests that even greater experience is required when interpreting images acquired with this radionuclide.

The greater or lesser experience of each observer explains in part why intraobserver reproducibility was poorer with ^{131}I. This difference in experience may also have contributed to poorer interobserver reproducibility, although it was not the main reason since κ values for O1 and O3, who had similar experience, were also not as good for the ^{131}I-labeled antibody.

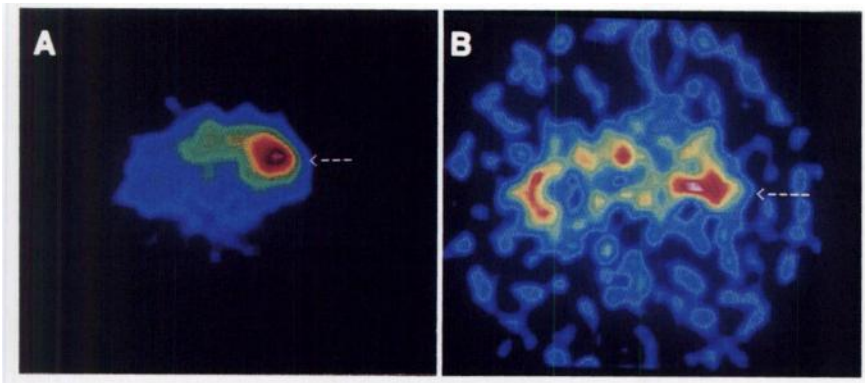

FIGURE 1
This case illustrates the better reproducibility of ^{111}In as compared to ^{131}I ECT. For ^{111}In ECT (transversal section) (A), the three observers were in definite agreement about the pelvic focus (arrow), which did not correspond to any nonspecific uptake site and was very likely of tumor origin (as confirmed by surgery). For ^{131}I ECT (B), this same focus appeared among other areas of increased activity, which made identification difficult. Observers either considered that there were several foci or interpreted everything as nonspecific activity. In fact, O1 gave a score of 3 at first reading and a score of 4 at second reading; O2 considered the examination to be negative at both readings (score 0); and O3 gave a score of 3.

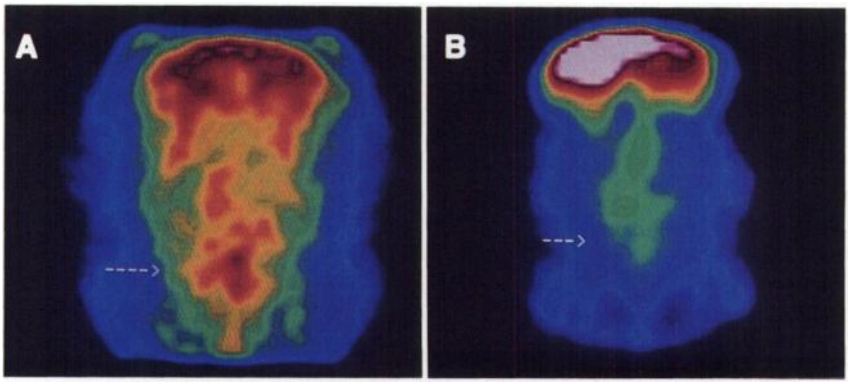

FIGURE 2
This case illustrates that reproducibility is sometimes better with ^{131}I in the planar mode. All three observers noted a pelvic focus (arrow) (score 4) in planar ^{131}I (A) images, whereas there was discordance for planar ^{111}In (B): score 0 for O1 at both readings; score 3 and then 0 for O2 and score 2 for O3. In the ^{111}In PS image, intense liver activity made tumor contrast much less apparent; the color of the tumor area is the same as that of overlying nonspecific bone marrow.

Reproducibility is an important criterion (*33*) because of its role in guiding therapeutic strategy. However, this problem is rarely considered in the literature. Surprisingly, the few studies done indicate that the reproducibility of widely used and well-codified examinations is far from perfect. In diagnosis of chronic pyelonephritis on i.v. pyelograms, a κ value of only 0.425 has been found (*34*). Kappa value can be calculated as 0.471 for detection of pneumoconiosis on chest X-rays (*35*), and 0.601 for auscultation of heart valve diseases (*36*). Hoey et al. (*37*) have shown the low reproducibility of lung scans for segmental and subsegmental defects ($\kappa < 0.5$). There have not been many studies of the reproducibility of interpretation of computed tomography for evaluation of lesions. This reproducibility is not always satisfactory (*38–40*). To our knowledge, there have been no studies of the reproducibility of computed tomography or magnetic resonance imaging interpretation for suspected recurrence of ovarian cancer. Warde et al. have examined the reliability of computed tomography in the assessment of patients with advanced ovarian cancer (*40*). Intraobserver agreement was moderate to excellent (weighted kappa values between 0.52 and 0.84), while interobserver agreement was lower (weighted kappa values between 0.36 and 0.79).

It may be concluded that the present study shows satisfactory intra- and interobserver reproducibility for immunoscintigraphy, which is an important factor in validating the utility of this imaging method. Reproducibility was better when ^{111}In rather than ^{131}I was used for antibody labeling, and with ECT rather than PS. This superiority was based on rigorous comparison since the radionuclide was the only variable factor between the two recordings. Moreover, it would seem that the period of training for interpretation of immunoscintigraphic images is shorter when ^{111}In is used.

The technology is progressing rapidly, and it is likely that ^{99m}Tc will be increasingly used for antibody labeling—a tendency encouraged by the first clinical results (*41*). Similar studies to ours should be carried out to compare the reproducibility of ^{111}In-labeled antibodies and ^{99m}Tc-labeled antibodies.

ACKNOWLEDGMENTS

The authors thank Pr. M. Comet for critically reviewing the manuscript, J. Gray for assistance in manuscript translation, and C. Arnaud for secretarial assistance. The figures have been reproduced with the kind permission of CRC Press.

REFERENCES

1. Britton KE, Granowska M. Radioimmunoscintigraphy in tumour identification. *Cancer Surveys* 1987;6:247–267.
2. Chatal JF, Saccavini JC, Fumoleau P, et al: Immunoscintigraphy of colon carcinoma. *J Nucl Med* 1984;25:307–14.
3. Epenetos AA, Kosmas C. Monoclonal antibodies for imaging and therapy. *Br J Cancer* 1989;59:152–55.
4. Halpern SE, Dillman RO. Problems associated with radioimmunodetection and possibilities for future solutions. *J Biol Response Mod* 1987;6:235–262.
5. Goldenberg DM. Targeting of cancer with radiolabeled antibodies. Prospects for imaging and therapy. *Arch Pathol Lab Med* 1988;112:580–587.
6. Halpern SE. The advantage and limits of indium-111 labeling of antibodies: experimental studies and clinical applications. *Nucl Med Biol* 1986;13:195–201.
7. Murray JL, Lamki LM, Shanken LJ, et al. Immunospecific saturable clearance mechanisms for indium-111-labeled antimelanoma monoclonal antibody 96.5 in humans. *Cancer Res* 1988;48:4417–4422.
8. Chatal JF, Fumoleau P, Saccavini JC, et al. Immunoscintigraphy of recurrences of gynecological carcinomas. *J Nucl Med* 1987;28:1807–1819.
9. Peltier P, Fumoleau P, Kremer M, Thedrez P, Chetanneau A, Chatal JF. Immunoscintigraphy of ovarian cancers using In-111-DTPA-OC125 monoclonal antibody. In: Chatal JF, ed. *Monoclonal antibodies in immunoscintigraphy*. Boca Raton, Florida: CRC Press; 1989:187–209.
10. Carney PL, Rogers PE, Johnson DK. Dual-isotope study of iodine-125- and indium-111-labeled antibody in athymic mice. *J Nucl Med* 1989;30:374–384.
11. Fairweather DS, Bradwell AR, Dykes PW, et al. Improved tumour localisation using indium-111-labeled antibodies. *Brit Med J* 1983;187:167.
12. Yokoyama K, Carrasquillo JA, Chang AE, et al. Differences in biodistribution of indium-111- and iodine-131-labeled B72-3 monoclonal antibodies in patients with colorectal cancer. *J Nucl Med* 1989;30:320–327.
13. Halpern SE, Haindl W, Beauregard J, et al. Scintigraphy with In-111-labeled monoclonal antitumor antibodies: kinetics, biodistribution, and tumor detection. *Radiology* 1988;168: 529–536.
14. Perkins AC, Pimm MV. Differences in tumour and normal

tissue concentrations of iodine- and indium-labelled monoclonal antibody. I. The effect on image contrast in clinical studies. *Eur J Nucl Med* 1985;11:295–299.
15. Bast RC, Feeney M, Lazarus H, et al. Reactivity of a monoclonal antibody with human ovarian carcinoma. *J Clin Invest* 1981;68:1331–1337.
16. Hnatowich DJ, Griffin TW, Koscivczyk C, et al. Pharmacokinetics of an indium 111-labelled monoclonal antibody in cancer patients. *J Nucl Med* 1985;26:849–858.
17. Landis RJ, Koch GG. The measurement of observer agreement for categorial data. *Biometrics* 1977;33:159–174.
18. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 1981;29:111–123.
19. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
20. Epenetos AA, Carr D, Johnson PM, Bodmer WF, Lavender JP. Antibody-guided radiolocalisation of tumours in patients with testicular or ovarian cancer using two radioiodinated monoclonal antibodies to placental alkaline phosphatase. *Brit J Radiol* 1986;59:117–125.
21. Haisma HJ, Moseley KR, Battaile AI, Griffiths TC, Zurawski VR, Knapp RC. Biodistribution, pharmacokinetics and imaging of ^{131}I-labelled OC125 in ovarian cancer. *Int J Cancer* 1988;(suppl 2):109–113.
22. Symonds EM, Perkins AC, Pimm MV, Baldwin RW, Hardy JG, Williams DA. Clinical implications for immunoscintigraphy in patients with ovarian malignancy: a preliminary study using monoclonal antibody 791T/36. *Br J Obstet Gynæcol* 1985;92:270–276.
23. Britton KE, Granowska M. Localization of cancer of the ovary and metastases using ^{123}I-labeled monoclonal antibody HMFG-2 compared to surgical findings. In: Baldwin RSW, Byers VS, eds *Monoclonal antibodies for cancer detection and therapy*. Orlando, FL: Academic Press; 1985:201–221.
24. Critchley M, Brownless S, Patten M, et al. Radionuclide imaging of epithelial ovarian tumours with ^{123}I-labelled monoclonal antibody (H317) specific for placental-type alkaline phosphatase. *Clin Radiol* 1986;37:107–112.
25. Davies JO, Davies ER, Howe K, et al. Radionuclide imaging of ovarian tumours with ^{123}I-labeled monoclonal antibody ($NDOG_2$) directed against placental alkaline phosphatase. *Brit J Obstet Gynecol* 1985;92:277–286.
26. Epenetos AA, Sheperd J, Britton KE, et al. I-123 radioiodinated antibody imaging of occult ovarian cancer. *Cancer* 1985;55:984–987.
27. Granowska M, Britton KE, Sheperd JH, et al. A prospective study of ^{123}I-labeled monoclonal antibody imaging in ovarian cancer. *J Clin Oncol* 1986;4:730–736.
28. Pateiski N, Philipp K, Skodler WD, Czerwenka K, Hamilton G, Burchell J. Radioimmunodetection in patients with suspected ovarian cancer. *J Nucl Med* 1985;26:1369–1676.
29. Pateiski N, Philipp K, Sevelda P, et al. Radioimmunoscintigraphy using monoclonal antibodies before second-look surgery in patients suffering from ovarian cancer. *Gynecol Obstet Invest* 1987;24:212–216.
30. Sheperd J, Granowska M, Britton KE, et al. Tumour-associated monoclonal antibodies for the diagnosis and assessment of ovarian cancer. *Br J Obstet Gynæcol* 1987;94:160–167.
31. Hunter RE, Doherty P, Griffin TW, et al. Use of indium-111-labeled OC-125 monoclonal antibody in the detection of ovarian cancer. *Gynecol Oncol* 1987;27:325–337.
32. Powell MC, Perkins AC, Pimm MV, et al. Diagnostic imaging of gynecologic tumors with the monoclonal antibody 791T/36. *Am J Obstet Gynecol* 1987;157:28–34.
33. Koran LM. The reliability of clinical methods, data and judgments. *N Engl J Med* 1975;293:642–695.
34. Norden C, Philipps E, Levy P, Kass E. Variation in interpretation of intravenous pyelograms. *Amer J Epid*, 1970;91:155–160.
35. Felson B, Morgan WKC, Bristol LJ, et al. Observations on the results of multiple readings of chest films in coal miner's pneumoconiosis. *Radiology* 1973;109:19–23.
36. Taranta A, Spagnuolo M, Snyder R, Gerbarg DS, Hofler JJ. Auscultation of the heart by physicians and by computer. In: Enslein K, ed. *Data acquisition and processing in biology and medicine, volume 3*. New York: Macmillan; 1964;23–52.
37. Hoey JR, Farrer PA, Rosenthall LJ, Spengler RF. Interobserver and intraobserver variability in lung scan reading in suspected pulmonary embolism. *Clin Nucl Med* 1980;5:508–513.
38. Kolmannskog F, Swensen T, Larsen S, Larssen T. Reproducibility and observer variation at computed tomography and ultrasound of the normal pancreas. *Acta Radiol Diagn* 1983;24:21–25.
39. Shinar D, Gross CR, Hier DB, et al. Interobserver reliability in the interpretation of computed tomographic scans of stroke patients. *Arch Neurol* 1987;44:149–155.
40. Warde P, Rideout DF, Herman S, et al. Computed tomography in advanced ovarian cancer. Inter- and intraobserver reliability. *Invest Radiol* 1986;21:31–33.
41. Lind P, Langsteger W, Költringer P, et al. Tc-99m-labeled monoclonal anti-carcinoembryonic antigen antibody (BW 431/26). Clinical results in the detection of colorectal carcinomas and recurrences. *Scand J Gastroenterol* 1989;24:1205–1211.