ROC and Localization ROC Analyses of Lesion Detection in Whole-Body FDG PET: Effects of Acquisition Mode, Attenuation Correction and Reconstruction Algorithm

Thomas H. Farquhar, Jorge Llacer, Carl K. Hoh, Johannes Czernin, Sanjiv S. Gambhir, Marc A. Seltzer, Daniel H.S. Silverman, Jinyi Qi, Chinghan Hsu and Edward J. Hoffman

Department of Molecular and Medical Pharmacology, Division of Nuclear Medicine and Biophysics, UCLA School of Medicine, Los Angeles; Ahmanson Biological Imaging Clinic, UCLA School of Medicine, Los Angeles; EC Engineering Consultants, Los Gatos; Department of Electrical Engineering-Systems, Signal and Image Processing Institute, University of Southern California, Los Angeles, California; and School of Medical Technology, Chang Gong University, Tao-Yuan, Taiwan

Receiver operating characteristic (ROC) and localization ROC (LROC) studies were performed to compare lesion detection at the borderline of detectability on images reconstructed with two-dimensional filtered backprojection (FBP) without attenuation correction (a common clinical protocol), three-dimensional FBP without attenuation correction, two-dimensional FBP with segmented attenuation correction and a two-dimensional iterative maximum a posteriori (MAP) algorithm using attenuation correction. Lung cancer was the model for the study because of the prominent role of ^{18}F-fluorodeoxyglucose PET in the staging of lung cancer and the importance of lesion detection for staging. **Methods:** Simulated lung cancer lesions were added to two-dimensional and three-dimensional PET data from healthy volunteers. Data were reconstructed using the four methods. Four nuclear medicine physicians evaluated the images. Detection performance with each method was compared using ROC and LROC analysis. Jackknife analysis provided estimates of statistical significance for differences across all readers for the ROC results. **Results:** ROC and LROC results indicated statistically significant degradation in detection performance with three-dimensional acquisition (average area under ROC curves $[A_z]$ 0.51; average area under LROC curves $[A_{z,LROC}]$ 0.13) and segmented attenuation correction (average A_z 0.59; average $A_{z,LROC}$ 0.29) compared with two-dimensional FBP without attenuation correction (average A_z 0.79; average $A_{z,LROC}$ 0.54). ROC and LROC results indicated an improvement in detection performance with iterative MAP reconstruction (average A_z 0.83; average $A_{z,LROC}$ 0.64) compared with two-dimensional FBP reconstruction; this improvement was not statistically significant. **Conclusion:** Use of segmented attenuation correction or three-dimensional acquisition with FBP reconstruction is not expected to improve detection of lung lesions on whole-body PET images compared with images with two-dimensional FBP without attenuation correction. The potential improvement in detection obtained with an iterative MAP reconstruction method is small compared with that obtained with two-dimensional FBP without attenuation correction.

Key Words: PET; receiver operating characteristic analysis; lesion detection; lung cancer

J Nucl Med 1999; 40:2043–2052

Whole-body ^{18}F-fluorodeoxyglucose (FDG) PET has shown clinical usefulness in the detection of a wide variety of primary and metastatic malignancies (*1*). Particularly, whole-body FDG PET has gained acceptance for the staging of lung cancer (*2*) and the characterization of solitary pulmonary nodules (*3–8*). Work by many researchers has focused on methods to improve imaging with whole-body PET. Proposed methods for the acquisition, correction and reconstruction of PET images are typically evaluated in terms of objective, quantitative measures such as resolution, contrast recovery or noise variance. Three widely researched methods are used for improving whole-body PET images. The first method includes emission studies with three-dimensional acquisition for count-limited applications, such as whole-body PET, because of the improvement in sensitivity (*9–11*). Although some objective measures have shown advantages with three-dimensional acquisition compared with two-dimensional acquisition, improvement with three-dimensional acquisition in a realistic, clinical setting has not been reliably shown for whole-body PET (*12–14*). The second method includes the development of practical, accurate techniques for attenuation correction with the goal of providing an image that reflects the actual tracer distribution with greater quantitative accuracy (*15–21*). However, efforts to show an improvement in lesion contrast have been inconclusive (*13,14*). The third includes statistical, iterative

methods as an alternative to filtered backprojection (FBP) for reconstruction of PET data from whole-body studies and other applications. Numerous studies have shown objective improvements such as improved resolution and contrast recovery with equivalent noise variance (22–27). However, only one study has shown an improvement in an appropriate clinical task (28), and no studies have shown an improvement in lesion detection by human observers on clinically relevant whole-body images. The identification of metastases and involved lymph nodes is crucial for the interpretation of whole-body FDG PET scans for staging lung cancer and evaluating solitary pulmonary nodules. As such, detection of solitary foci of elevated uptake was deemed an appropriate measure of algorithm performance. Receiver operating characteristic (ROC) methodology is well established as a reliable method of statistically quantifying differences between the detection performance of human observers for different imaging modalities. This study compares observer performance using four modalities: (a) two-dimensional whole-body acquisition reconstructed by two-dimensional FBP without attenuation correction (a widely used clinical protocol), (b) two-dimensional whole-body acquisition reconstructed by two-dimensional FBP with segmented attenuation correction, (c) two-dimensional whole-body acquisition reconstructed by an iterative, Bayesian, maximum a posteriori (MAP) algorithm with MAP-based attenuation correction and (d) three-dimensional whole-body acquisition reconstructed by three-dimensional FBP (the projection–reprojection algorithm [29]) without attenuation correction.

MATERIALS AND METHODS

An ROC study with a matched-pair design was used to facilitate accurate, meaningful comparisons between modalities. Issues in the proper experimental design of an ROC study were addressed to ensure that the final results had adequate statistical validity (30–32).

Data Acquisition

The emission data for both true-positive and true-negative datasets in the study were created by adding simulated lesions to emission data from healthy volunteers. This approach avoids many of the difficulties of using patient data while preserving the realism of the imaging environment.

Twenty-five healthy volunteers were imaged from December 2, 1996, to May 23, 1997, in accordance with the guidelines of the Human Subjects Protection Committee at our institution. A summary of the relevant characteristics of each volunteer is shown in Table 1. After fasting for 12 h, all volunteers were imaged on an ECAT HR+ 962 PET scanner (CTI/Siemens, Inc., Knoxville, TN). All acquisitions consisted of three bed positions from the neck to the lower abdomen. The protocol began with transmission scans acquired for 3 min per bed position using three rotating [67]Ge rod sources, each with approximately 185 MBq [67]Ge. Next, approximately 296 MBq FDG were injected intravenously. The dose for each volunteer is recorded in Table 1. Acquisition of two-dimensional emission data began after a 30-min delay, which was then followed by acquisition of three-dimensional emission data.

TABLE 1
Whole-Body Volunteer Information

Sex	Dose (MBq)	Age (y)	Comments
M	277	22	
M	277	29	Excessive subject motion
M	292	24	
F	292	19	
M	296	19	
M	303	22	Excessive subject motion
M	300	20	
M	296	24	
F	311	18	Excessive subject motion
M	300	33	Did not complete protocol
M	311	21	
F	270	20	
M	311	20	
M	289	30	
M	292	23	
M	296	33	
M	300	29	
F	289	19	
F	303	19	
F	303	40	
M	314	21	Excessive subject motion
F	300	19	Physiologic uptake on scan
F	303	31	
F	270	18	Excessive subject motion
F	303	23	
Average	296	24.2	
SD	13	6.1	

The clinical protocol at this institution uses an injection dose of 555 MBq FDG with frame durations of 6 min each. To obtain an equivalent number of true events per bed position (10–14 million counts over 63 two-dimensional planes in the thorax) from the volunteer protocol, a frame duration of 9 min per bed position was used for the two-dimensional acquisitions. The reduced, 296-MBq injected dose was intended to decrease radiation dose to the volunteer and to mitigate dead time in the three-dimensional acquisitions. The three-dimensional frames were acquired for half of the two-dimensional frame duration (4.5 min). Furthermore, the three-dimensional acquisitions began approximately 70 min after the two-dimensional acquisitions.

The emission data from 6 patients were removed from the study because of excessive subject motion or failure to complete the entire protocol, as indicated in Table 1. The images of the remaining 19 volunteers were reviewed by a nuclear medicine physician. Eighteen volunteers were established as healthy volunteers with no suspicious or anomalous uptake. One volunteer was excluded from the study because of diffuse uptake in the right lung, consistent with the volunteer's symptomatic upper respiratory tract infection.

Image Preparation

The datasets for the study and training session were created by adding simulated lesions to the emission data. Subsequently, the datasets were reconstructed by each of the reconstruction modalities into image volumes. Coronal slices were taken from the reconstructed image volumes to provide true-positive and true-negative images for interpretation. Each true-positive image contained a single simulated lesion. True-negative images were coronal planes that did not contain a simulated lesion. The steps in

the selection, simulation and addition of lesions to the scan data from healthy volunteers are described. Twenty-two distinct lesion shapes were obtained from CT scans of patients with biopsy-proven lung cancer. The lesion locations were chosen to be anatomically accurate and pathologically relevant, with an anatomic distribution between central and peripheral lung structures, and distant sites (i.e., suprarenal), representative of the proportions observed in epidemiologic studies (33). The realism of the chosen lesion shapes and locations were also verified by a nuclear medicine physician. Each lesion was simulated by forward projection using the transition matrix for an ECAT HR+ 962 scanner. The system model used in computation of the transition matrix has been described (26). In addition to the fundamental geometric factors, the model includes the physical effects of photon pair noncolinearity, intercrystal scatter and crystal penetration as well as the axial rebinning and angular mashing performed on the projection data.

Because the forward projection did not incorporate the effects of photon attenuation or detector normalization, the projection datum for each line of response was divided by the appropriate attenuation correction factor and normalization factor. The attenuation correction factor for each line of response was obtained from the segmented attenuation correction algorithm (19), whereas the normalization data for each scanner and acquisition mode provided the detector efficiencies for each line of response. A global scale factor was applied to each set of projection data to adjust the contrast of the lesion in the final reconstructed image. This was determined empirically for each lesion to provide marginal detectability and verified in a preliminary, small-scale ROC study with a single reader. Lastly, Poisson deviates of the fully simulated lesion data were added to the sinogram data of a healthy volunteer.

To compare two-dimensional reconstruction modalities with a three-dimensional reconstruction modality, an equivalent three-dimensional lesion dataset was constructed for each two-dimensional lesion dataset. The three-dimensional attenuation correction factors were obtained by forward projection of a reconstructed two-dimensional attenuation image. The appropriate three-dimensional normalization factors were calculated using a three-dimensional normalization protocol with the septa retracted. More importantly, several scale factors, in addition to the contrast scale factor, were incorporated into the three-dimensional simulation. These additional scale factors corrected for differences in frame duration and radioactive decay, dead-time efficiency and sensitivity of the direct segment in two-dimensional and three-dimensional studies (i.e., with and without interplane septa).

Image Reconstruction

Because of the matched study design, readers interpreted four reconstructed images for each of 50 true-positive and 50 true-negative datasets. The lesion datasets were reconstructed using four methods:

1. Two-dimensional FBP without attenuation correction. This method is the current clinical reconstruction protocol for whole-body studies at this institution. A Hann filter with cutoff at 0.40 of Nyquist (0.89 cycle/cm, full width at half maximum [FWHM] 12.6 mm) was applied to the ramp reconstruction filter; axial filtering was also performed using a Hann filter with cutoff at 0.42 of Nyquist (0.87 cycle/cm, FWHM 12 mm). An implementation of FBP provided with the ECAT HR+ 962 system was used to perform the reconstructions.

2. Two-dimensional FBP with segmented attenuation correction. The same radial and axial filters described in the first method were used. The particular implementation of a segmented attenuation correction algorithm has been described (21) and was available with the Clinical Applications Programming Package (CAPP) software on the ECAT HR+ 962 system. In addition to the transmission scans, calculation of the attenuation correction factors used a blank scan of 430 million counts from a ^{67}Ge cylinder collected daily for each scanner.

3. Two-dimensional iterative MAP algorithm with a MAP-based attenuation correction. The MAP algorithm has been described (26). Twenty iterations with a three-dimensional quadratic prior and smoothing parameter (beta) of 2000 were performed. These parameters were chosen empirically to provide adequate smoothing for whole-body data. Attenuation correction was performed using correction factors reprojected from an attenuation image also reconstructed with an iterative MAP algorithm on the original transmission and blank scan data.

4. Three-dimensional FBP without attenuation correction. The reconstructions were performed with an implementation of the three-dimensional reprojection algorithm with the variable axial rebinning approximation available on the ECAT HR+ 962 scanner. The same radial and axial filters were used as in methods 1 and 2. A Hann filter with cutoff at Nyquist (2.22 cycles/cm, FWHM 6 mm) was used in the two-dimensional FBP reconstruction of the estimated image volume. The scanner implementation of the projection–reprojection algorithm also incorporated a simulation-based scatter correction as described (34).

For all reconstruction methods, each bed position was reconstructed and then rebinned into a single image volume with 2.25-mm pixels in all dimensions. Three consecutive coronal planes were extracted from the image volume. The center coronal slice included the center of a simulated lesion for true-positive images. The other two adjacent planes provided anatomic context for the observers. For comparison, Figure 1 shows a coronal slice from reconstructions of the same data using each of the four methods.

Image Presentation and Evaluation

A training session preceded the study. The nature of the study was explained and detailed instructions on the use of the ROC evaluation software were provided. Four nuclear medicine physicians with expertise reading whole-body FDG PET scans participated as readers. Image evaluation was performed on a Sun workstation (Sun Microsystems, Mountain View, CA) that is used daily for clinical rounds. Readers were able to modify the color scale of the display as desired for each image. The readers were instructed to consider the following clinical scenario:

"For each image, you are to assume that you are viewing the study of a patient with a presentation and history suspicious for lung cancer with any possible associated metastases. In rating each image, you are to consider the possible presence of a lesion or abnormality. A lesion or abnormality is to include BOTH malignant disease AND benign, inflammatory processes."

Readers rated the central image of each image triplet using a continuous rating scale in answer to the question "Does the image contain an abnormality?" Also, the reader indicated the most probable lesion location with a mouse. Each reader rated the entire
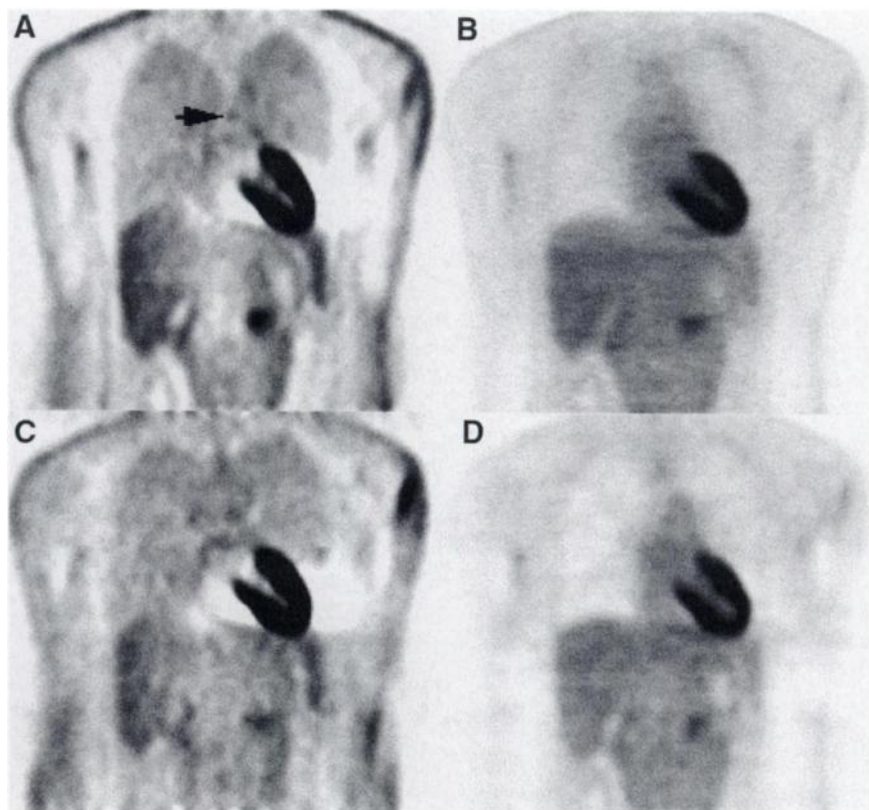
FIGURE 1. Sample coronal slice reconstructed by four reconstruction methods. (A) Two-dimensional FBP without attenuation correction. Arrow indicates simulated lesion. (B) Two-dimensional iterative MAP reconstruction with MAP-based attenuation correction. (C) Three-dimensional FBP without attenuation correction. (D) Two-dimensional FBP with segmented attenuation correction.

dataset of 400 images over a period of several weeks. The images were stratified and then randomized within strata to minimize the likelihood that a reader would interpret the same study reconstructed with two modalities within a short period of time.

Receiver Operating Characteristic Analysis

The ratings for each observer were analyzed using the binormal curve-fitting routine of the CORROC2 program developed at the University of Chicago by Metz et al. (35). The CORROC2 analysis calculated a binormal curve fit, the area under the fitted ROC curve (A_z) and the SE of the estimate of A_z for the two modalities being compared. In addition, a probability value was computed to indicate the likelihood of statistical significance for the observed difference in A_z values between methods. The null hypothesis assumed that all four reconstruction methods were equivalent. A 95% confidence threshold was used to determine statistical significance. Four intermodality comparisons were made: first, two-dimensional versus three-dimensional acquisition, both with FBP reconstruction and no attenuation correction; second, no attenuation correction versus use of a segmented attenuation correction algorithm, both with two-dimensional FBP reconstruction; third, FBP reconstruction without attenuation correction versus MAP reconstruction with attenuation correction, with two-dimensional acquisition; and fourth, FBP versus iterative MAP reconstruction, both with two-dimensional acquisition and attenuation correction. A Bonferroni correction was applied to correct for multiple comparisons on the same data. For the first, second and third comparisons, a P of <0.0167 was considered significant; for the fourth, a P of <0.025 was considered significant. Pooled reader ROC data for each of the four intermodality comparisons were analyzed with a jackknife technique (32,36). Results of the

jackknife analysis show reduced bias associated with between-case and between-reader correlations and a reduction in SE of the estimated A_z (36).

Localization Receiver Operating Characteristic Analysis

The ratings and suspected lesion location for each observer were analyzed using a localization (L/ROC) analysis program developed by Swensson (37). In this analysis, the responses of each reader for each modality were analyzed separately, yielding a fitted LROC, the area under the fitted LROC curve, and an estimate of the SE of the measured area. A lesion was recognized as being properly localized when the pixel specified by the reader was within a four-pixel radius (9.0 mm) of the center of the true lesion location. To ascertain the statistical significance of differences between modalities observed across readers, a Student t test for paired data was performed. The t statistic was converted to a probability value to facilitate comparison with the ROC results.

RESULTS

Receiver Operating Characteristic Results for Individual Readers

The results of the CORROC2 analysis are shown in Table 2 and Figure 2. Table 2 shows the area under the fitted binormal curve for each modality, the SE estimate for this area and the probability values of the calculated difference for each of the four intermodality comparisons. In a separate plot for each reader, the fitted curves for each modality are illustrated in Figure 2. The A_z values and SEs of Table 2 and

TABLE 2
ROC Results for Four Nuclear Medicine Physicians

Modality	A_z Reader 1	Reader 2	Reader 3	Reader 4	Average
2-D FBP, no AC	0.79	0.74	0.79	0.83	0.79
	(0.05)	(0.05)	(0.04)	(0.04)	
2-D MAP, MAP AC	0.81	0.83	0.84	0.85	0.83
	(0.04)	(0.04)	(0.04)	(0.04)	
2-D FBP, seg AC	0.54	0.67	0.63	0.50	0.59
	(0.06)	(0.05)	(0.06)	(0.06)	
3-D FBP, no AC	0.54	0.49	0.45	0.57	0.51
	(0.06)	(0.06)	(0.06)	(0.06)	
P					
2-D FBP, no AC					
vs. 3-D FBP, no AC	<0.001	<0.001	<0.001	<0.001	
vs. 2-D MAP, MAP AC	0.623	0.154	0.378	0.653	
vs. 2-D FBP, seg AC	<0.001	0.317	0.016	<0.001	
2-D MAP, MAP AC					
vs. 2-D FBP, seg AC	<0.001	0.011	<0.001	<0.001	

ROC = receiver operating characteristic; A_z = area under fitted, binormal ROC curve; 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; MAP = maximum a posteriori; seg = segmented; 3-D = three-dimensional.
Values in parentheses are estimates of SE. Also shown are two-sided probability values for each intermodality comparison.

the fitted curves of Figure 2 are averages of the repeated measurements for modalities in multiple comparisons.

For all readers, a statistically significant degradation in detection performance (from an average A_z of 0.79 for two-dimensional acquisition to 0.51 for three-dimensional acquisition) was measured with three-dimensional acquisition and reconstruction compared with the two-dimensional protocol. For all readers, the clinical protocol of two-dimensional FBP without attenuation correction gave slightly depressed detection performance (average A_z of 0.79 versus 0.83) compared with the iterative MAP technique. However, this improvement was not considered statistically significant for any of the readers. All readers showed reduced detection performance (to an average A_z of 0.59 down from 0.79) with segmented attenuation correction. This degradation was statistically significant for readers 1 and 3. Lastly, use of MAP image reconstruction and attenuation correction was seen to dramatically improve detection performance (from an average A_z of 0.59 to 0.83) for all readers compared with FBP reconstruction using segmented attenuation correction. Again, this improvement was statistically significant for each individual reader.

### Jackknife Analysis of Pooled Receiver Operating Characteristic Data

The results of the jackknife analysis are shown in Table 3. The probability values for the pooled data are consistent with the trends observed in the individual ROC results reported above. The measured improvement using iterative MAP reconstruction with MAP-based attenuation correction compared with the common clinical reconstruction protocol of two-dimensional FBP without attenuation correction was

not statistically significant at a confidence level of 95%. For the three other comparisons, the results of the pooled data comparisons were statistically significant, with $P < 0.01$.

Localization Receiver Operating Characteristic Results

The results of the LROC analysis for each reader and each modality are summarized in Table 4 and shown in Figure 3. Table 4 reports the areas under the fitted LROC curve $(A_{z,LROC})$ and the estimated SE of the measured area for each modality and each observer. As with the ROC results, four intermodality comparisons were made, and probability values indicating the statistical significance of the observed differences are given. In a separate plot for each reader, the fitted LROC curves for each modality are illustrated in Figure 3. The results of the Student t test are reported in Table 5 in a manner similar to that of the jackknife ROC results.

The trends observed in the ROC analysis were largely confirmed by the LROC analysis. For example, the improved detection performance of the clinical two-dimensional whole-body protocol (average $A_{z,LROC} = 0.54$) over three-dimensional FBP without attenuation correction (average $A_{z,LROC} = 0.13$) was noted for all four readers $(P \le 0.01)$ and in the pooled data. Likewise, the average LROC A_z for two-dimensional FBP without attenuation correction (0.54) was higher than that for two-dimensional FBP with segmented attenuation correction (0.29). This improvement was significant for two of four readers with $P \le 0.02$ and three of four readers with $P \le 0.06$. The probability value for the pooled reader results bordered on statistical significance $(P = 0.06)$ at a 95% confidence level. Comparison of the detection performance with the iterative MAP reconstruction with MAP-based attenuation correction (average
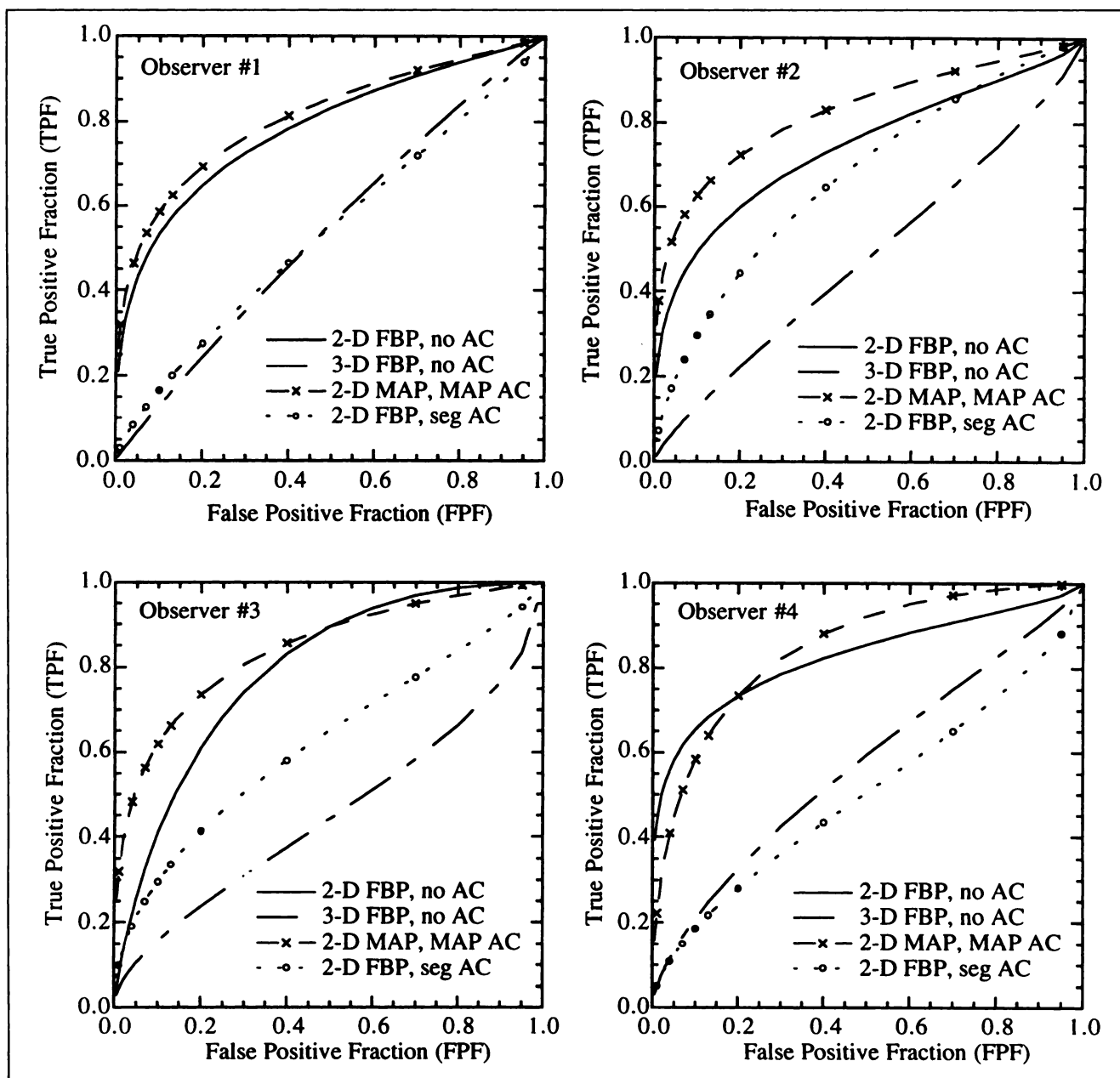
FIGURE 2. ROC curves measured for four reconstruction modalities by four nuclear medicine physicians. 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; 3-D = three-dimensional; MAP = maximum a posteriori; seg = segmented.

$A_{z,LROC}$ = 0.64) with two-dimensional FBP with segmented attenuation correction (average $A_{z,LROC}$ = 0.29) was even more definitive; statistically significant improvement ($P \le$ 0.02) was found for all four readers individually and the pooled reader results. Lastly, as with the results of the CORROC2 analysis, an increase in average $A_{z,LROC}$ was noted for all four readers with the iterative MAP reconstruction (average $A_{z,LROC}$ = 0.64) compared with the two-dimensional whole-body protocol (average $A_{z,LROC}$ = 0.54). Again, however, this difference failed to have statistical significance for any of the four readers with P = 0.34 for the t test of the pooled reader LROC results. In conclusion, the

results of the LROC analysis suggest that localization performance is consistent with the detection performance measured with the CORROC2 ROC analysis.

DISCUSSION

The results of this study suggest that lesion detection performance for small, low-contrast lesions with increased uptake in the lung is not improved by the use three-dimensional acquisition or segmented attenuation correction compared with the common clinical protocol of two-dimensional FBP reconstruction without attenuation correc-

TABLE 3
Jackknife Analysis of Pooled, Multireader ROC Results

Modality 1	Modality 2	A_z 1	A_z 2	95% CI	P
2-D FBP, no AC	3-D FBP, no AC	0.76 (0.03)	0.49 (0.03)	[0.20, 0.34]	<0.001
2-D FBP, no AC*	2-D MAP, MAP AC	0.67 (0.03)	0.71 (0.03)	[−0.09, 0.02]	0.230
2-D FBP, no AC*	2-D FBP, seg AC	0.76 (0.03)	0.60 (0.03)	[0.07, 0.23]	<0.001
2-D FBP, seg AC	2-D MAP, MAP AC	0.56 (0.03)	0.81 (0.02)	[−0.31, −0.19]	<0.001

*Ratings of reader 4 were omitted from pooled analysis of two comparisons because jackknife software returned a degenerate result for analysis of reader 4.

ROC = receiver operating characteristic; A_z = area under ROC curve of jackknife analysis; 95% CI = 95% confidence interval; 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; 3-D = three-dimensional; MAP = maximum a posteriori; seg = segmented.

Results from four intermodality comparisons are shown. A_z is reported for each modality (A_z 1 and A_z 2). Values in parentheses are estimates of SE. The 95% CIs are given for estimated difference in A_z (A_z of modality 1 − A_z of modality 2). Two-sided probability value for each comparison is also shown.

tion. Although previous studies (12–14) have compared these methods with objective, image-based criteria, the comparison of detection performance using ROC and LROC analyses in this study is more directly relevant to the task of clinical image interpretation. Thus, the additional burdens of acquiring transmission scans and reconstructing three-dimensional data are not likely to be worthwhile for this application of whole-body PET. Similarly, the difference in detection performance between the iterative MAP reconstruction technique and two-dimensional FBP without attenuation correction was not statistically significant. Although improvements in objective measures such as contrast recov-

ery and noise variance have been reported for the iterative MAP reconstruction method compared with FBP (22,26), until an improvement in lesion detection can be shown, two-dimensional FBP without attenuation correction will be preferred because it does not require the transmission scans or computation complexity of iterative reconstruction.

The inferior performance of three-dimensional acquisition is speculated to result from decreased image contrast associated with unsatisfactory corrections for scattered and random events (9,10). For the acquisition protocol in this study, the improved sensitivity of the three-dimensional acquisitions is not sufficient to result in improved detection

TABLE 4
LROC Results for Four Nuclear Medicine Physicians

Modality	$A_{z,LROC}$				
	Reader 1	Reader 2	Reader 3	Reader 4	Average
2-D FBP, no AC	0.58 (0.06)	0.45 (0.06)	0.53 (0.06)	0.61 (0.06)	0.54
2-D MAP, MAP AC	0.62 (0.06)	0.59 (0.06)	0.65 (0.06)	0.70 (0.06)	0.64
2-D FBP, seg AC	0.29 (0.06)	0.31 (0.06)	0.32 (0.06)	0.23 (0.05)	0.29
3-D FBP, no AC	0.12 (0.04)	0.14 (0.04)	0.13 (0.04)	0.14 (0.05)	0.13
P					
2-D FBP, no AC					
vs. 3-D FBP, no AC	0.002	0.011	0.004	0.002	
vs. 2-D MAP, MAP AC	0.642	0.166	0.232	0.331	
vs. 2-D FBP, seg AC	0.020	0.180	0.063	0.006	
2-D MAP, MAP AC					
vs. 2-D FBP, seg AC	0.011	0.023	0.013	0.002	

LROC = localization receiver operating characteristic; $A_{z,LROC}$ = area under fitted LROC curve; 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; MAP = maximum a posteriori; seg = segmented; 3-D = three-dimensional.

Values in parentheses are estimates of SE. Also shown are two-sided probability values for each intermodality comparison.
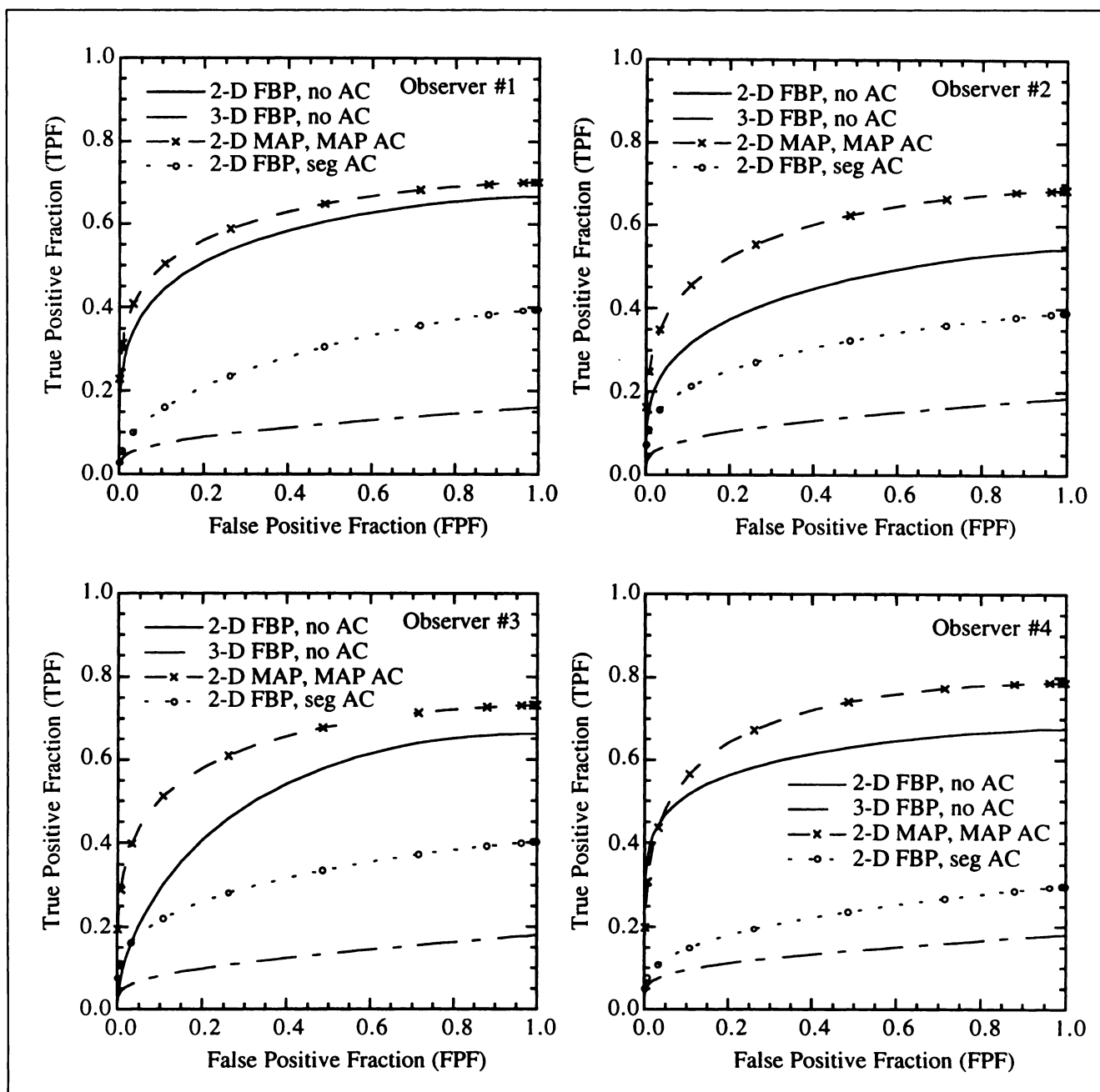
FIGURE 3. Localization ROC curves measured for four reconstruction modalities by four nuclear medicine physicians. 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; 3-D = three-dimensional; MAP = maximum a posteriori; seg = segmented.

performance after the deleterious effects of scattered and random events. Compared with the two-dimensional acquisitions in this study, the acquisition time for three-dimensional scans in this study was reduced by 50% and the dose at scan time was decreased by ~33%. The decreased dose is well justified because the use of a full two-dimensional dose would likely result in unacceptable dead-time losses. Furthermore, decreased dose for serial research protocols or pediatric studies is often cited as an advantage of three-dimensional acquisition. The possibility of decreasing scan time is also offered as a benefit of the sensitivity improve-

ment with three-dimensional acquisition; this possibility motivated the use of frames with half the duration of the two-dimensional studies. If the three-dimensional scans in this study had been acquired with the same frame duration as the two-dimensional scans, an improvement in the signal-to-noise ratio of the three-dimensional images would be expected, possibly to a level similar to or even better than that of two-dimensional studies. Establishing this improvement would require an additional study.

Compared with FBP reconstruction without attenuation correction, images reconstructed with FBP using segmented

TABLE 5
Analysis of Pooled, Multireader LROC Results

Modality 1	Modality 2	$A_{z,LROC}$ 1	$A_{z,LROC}$ 2	95% CI	P
2-D FBP, no AC	3-D FBP, no AC	0.54	0.13	[0.18, 0.64]	0.013
2-D FBP, no AC	2-D MAP, MAP AC	0.54	0.64	[−0.16, 0.36]	0.338
2-D FBP, no AC	2-D FBP, seg AC	0.54	0.29	[0.00, 0.51]	0.060
2-D FBP, seg AC	2-D MAP, MAP AC	0.29	0.64	[0.10, 0.60]	0.025

LROC = localization receiver operating characteristic; $A_{z,LROC}$ = average area under LROC curve; 95% CI = 95% confidence interval; 2-D = two-dimensional; FBP = filtered backprojection; AC = attenuation correction; 3-D = three-dimensional; MAP = maximum a posteriori; seg = segmented.

Results from four intermodality comparisons are shown. $A_{z,LROC}$ is reported for each modality ($A_{z,LROC}$ 1 and $A_{z,LROC}$ 2). The 95% CIs are given for estimated difference in $A_{z,LROC}$ ($A_{z,LROC}$ of modality 1 − $A_{z,LROC}$ of modality 2). Two-sided probability value from Student t test of each comparison is also shown.

attenuation correction were shown to have markedly degraded detection performance. This degradation in performance is not attributed to a high degree of noise or error in the attenuation correction factors but to the noise amplification that results from applying large attenuation factors to noisy, low-count emission data. The use of a segmented attenuation correction algorithm should greatly reduce the random noise in the attenuation correction factors compared with a measured attenuation correction technique, and accurate implementation should minimize systematic error associated with improperly chosen attenuation values or poor definition of boundaries. However, even with noise-free and low-error correction factors, noise in the emission data will be amplified by the correction.

Because the MAP method has shown improvement in objective measures of image quality compared with accelerated iterative methods such as ordered-subject expectation maximization (26,38) and iterative methods without such an accurate system model (39), the fact that a statistically significant improvement in lesion detection performance was not shown casts doubt on the potential benefit from these accelerated iterative reconstruction algorithms. The presumed improvement with an iterative technique, even if shown to be statistically significant, is likely to be small, as indicated by the 95% confidence interval for the difference calculated using the pooled data. Such a modest, conjectured improvement, measured only for borderline lesions, would not merit the additional effort of transmission scans and lengthy iterative reconstruction. However, a particularly promising result is the substantial improvement in detection performance of the MAP reconstruction method compared with FBP with segmented attenuation correction. This result suggests that iterative reconstruction may be favorable, if not necessary, when attenuation correction must be used. It is possible that the lack of experience of the physicians in interpreting images from iterative reconstruction methods depressed the detection performance of the MAP algorithm. The noise patterns and other aspects of image quality are dramatically different in images reconstructed with an iterative technique compared with FBP. An additional study after extensive training with images from iterative methods would be required to establish this conjectured improvement in detection performance. Currently, side-by-side interpretation of FBP images without attenuation correction and MAP-reconstructed images with attenuation correction would be recommended to provide possible improvement in detection performance and to train readers in the interpretation of MAP reconstructed images.

It should be acknowledged that these results are dependent on the specific parameters of the scanner, acquisition protocols and reconstruction methods used. However, each parameter in the study was chosen to approximate the current clinical standards. Thus, the general trends that have been described should not be discounted unless contradicted by an observer performance study. An additional objection is that the detection paradigm used here does not perfectly represent the process of clinical interpretation. Although this paradigm is indeed artificial, it is not likely to invalidate the study as a measure of comparative detection performance.

CONCLUSION

Lesion detection performance in images of whole-body FDG PET scans simulating possible lung cancer lesions was quantified by ROC and LROC analysis. The results tend to suggest that neither three-dimensional acquisition nor the use of attenuation correction with FBP reconstruction will improve lesion detection in whole-body FDG PET scans for evaluation of suspected lung cancer. Thus, the additional burdens of acquiring transmission scans and reconstructing three-dimensional data are not deemed to be worthwhile. Likewise, until the iterative MAP algorithm can be proven to provide substantially improved task-based performance, it is unlikely that the method will be preferred over the current standard of two-dimensional FBP without attenuation correction, especially because two-dimensional FBP does not require the acquisition of transmission scans or the increased computation time of iterative MAP reconstruction.

ACKNOWLEDGMENTS

REFERENCES

1. Hoh CK, Hawkins RA, Glaspy JA, et al. Cancer detection with whole- body PET using 2-[18F]fluoro-2-deoxy-D-glucose. *J Comput Assist Tomogr.* 1993;17:582–589.
2. Sasaki M, Ichiya Y, Kuwabara Y, et al. The usefulness of FDG positron emission tomography for the detection of mediastinal lymph node metastases in patients with non-small cell lung cancer: a comparative study with x-ray computed tomography. *Eur J Nucl Med.* 1996;23:741–747.
3. Dewan NA, Gupta NC, Redepenning LS, Phalen JJ, Frick MP. Diagnostic efficacy of PET-FDG imaging in solitary pulmonary nodules: potential role in evaluation and management. *Chest.* 1993;104:997–1002.
4. Gupta NC, Frank AR, Dewan NA, et al. Solitary pulmonary nodules: detection of malignancy with PET with 2-[F-18]-fluoro-2-deoxy-D-glucose. *Radiology.* 1992;184:441–444.
5. Kubota K, Matsuzawa T, Fujiwara T, et al. Differential diagnosis of lung tumor with positron emission tomography: a prospective study. *J Nucl Med.* 1990;31:1927–1932.
6. Knight SB, Delbeke D, Stewart JR, Sandler MP. Evaluation of pulmonary lesions with FDG-PET: comparison of findings in patients with and without a history of prior malignancy. *Chest.* 1996;109:982–988.
7. Patz EF Jr, Lowe VJ, Hoffman JM, et al. Focal pulmonary abnormalities: evaluation with F-18 fluorodeoxyglucose PET scanning. *Radiology.* 1993;188:487–490.
8. Rege SD, Hoh CK, Glaspy JA, et al. Imaging of pulmonary mass lesions with whole-body positron emission tomography and fluorodeoxyglucose. *Cancer.* 1993;72:82–90.
9. Cherry SR, Dahlbom M, Hoffman EJ. 3D PET using a conventional multislice tomograph without septa. *J Comput Assist Tomogr.* 1991;15:655–668.
10. Dahlbom M, Eriksson L, Rosenqvist G, Bohm C. A study of the possibility of using multi-slice PET systems for 3D imaging. *IEEE Trans Nucl Sci.* 1989;36:1066–1071.
11. Townsend DW, Geissbuhler A, Defrise M, et al. Fully three-dimensional reconstruction for a PET camera with retractable septa. *IEEE Trans Med Imaging.* 1991;10:505–512.
12. Kohlmyer SG, Mankoff DA, Lewellen TK, Kaplan MS. Comparison of 2D and 3D qualitative whole body positron emission tomography (PET) without attenuation or scatter correction. In: Moonier PA, ed. *1995 IEEE Nuclear Science Symposium and Medical Imaging Conference Record.* Piscataway, NJ: IEEE; 1996;2:1367–1371.
13. Gupta NC, Raylman RR, Graeber GM, Bishop MA. Rapid 3D acquisition of whole body PET images for cancer detection: pros and pits [abstract]. *J Nucl Med.* 1997;38(suppl):197P.
14. Bengel F, Ziegler S, Avril N, Laubenbacher C, Schwaiger M. Whole-body-PET in clinical oncology: comparison of attenuation-corrected and non-corrected images [abstract]. *J Nucl Med.* 1997;38(suppl):198P.
15. Carson RE, Daube-Witherspoon ME, Green MV. A method for postinjection PET transmission measurements with a rotating source. *J Nucl Med.* 1988;29:1558–1567.
16. Chatziioannou A, Dahlbom M, Hoh CK. Study on the use of transmission scans for whole body PET attenuation correction. *IEEE Trans Nucl Sci.* 1994;41:1545–1550.
17. deKemp RA, Nahmias C. Attenuation correction in PET using single photon transmission measurement. *Med Phys.* 1994;21:771–778.
18. Huang SC, Hoffman EJ, Phelps ME, Kuhl DE. Quantitation in positron emission computed tomography: 2. Effects of inaccurate attenuation correction. *J Comput Assist Tomogr.* 1979;3:804–814.
19. Meikle SR, Dahlbom M, Cherry SR. Attenuation correction using count-limited transmission data in positron emission tomography. *J Nucl Med.* 1993;34:143–150.
20. Meikle SR, Bailey DL, Hooper PK, et al. Simultaneous emission and transmission measurements for attenuation correction in whole-body PET. *J Nucl Med.* 1995;36:1680–1688.
21. Xu M, Luk WK, Cutler PD, Digby WM. Local threshold for segmented attenuation correction of PET imaging of the thorax. *IEEE Trans Nucl Sci.* 1994;41:1532–1537.
22. Mumcuoglu EU, Leahy RM, Cherry SR. Bayesian reconstruction of PET images: methodology and performance analysis. *Phys Med Biol.* 1996;41:1777–1807.
23. Matej S, Herman GT, Narayan TK, Furuie SS, Lewitt RM, Kinahan PE. Evaluation of task-oriented performance of several fully 3D PET reconstruction algorithms. *Phys Med Biol.* 1994;39:355–367.
24. Kinahan PE, Matej S, Karp JS, Herman GT, Lewitt RM. A comparison of transform and iterative reconstruction techniques for a volume-imaging PET scanner with a large axial acceptance angle. *IEEE Trans Nucl Sci* 1995;42:2281–2287.
25. Herman G, Odhner D. Performance evaluation of an iterative image reconstruction algorithm for positron emission tomography. *IEEE Trans Med Imaging.* 1991;10:336–346.
26. Qi J, Leahy RM, Hsu C, Farquhar TH, Cherry SR. Fully 3D Bayesian image reconstruction for the ECAT EXACT HR+. *IEEE Trans Nucl Sci.* 1998;45:1096–1103.
27. Reader AJ, Visvikis D, Erlandsson K, Ott RJ, Flower MA. Intercomparison of four reconstruction techniques for positron volume imaging with rotating planar detectors. *Phys Med Biol.* 1998;43:823–834.
28. Llacer J, Veklerov E, Baxter LR, et al. Results of a clinical receiver operating characteristic study comparing filtered backprojection and maximum likelihood estimator images in FDG PET studies. *J Nucl Med.* 1993;34:1198–1203.
29. Kinahan PE, Rogers JG. Analytic 3D image reconstruction using all detected events. *IEEE Trans Nucl Sci.* 1989;36:964–968.
30. Metz CE. ROC methodology in radiologic imaging. *Invest Radiol.* 1986;21:720–733.
31. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol.* 1989;24:234–245.
32. Berbaum KS, Dorfman DD, Franken EA Jr. Measuring observer performance by ROC analysis: indications and complications. *Invest Radiol.* 1989;24:228–233.
33. Saccomanno G, Auerbach O, Kuschner M, et al. A comparison between the localization of lung tumors in uranium miners and in nonminers from 1947 to 1991. *Cancer.* 1996;77:1278–1283.
34. Watson CC, Newport D, Casey ME, deKemp RA, Beanlands RS, Schmand M. Evaluation of simulation-based scatter correction for 3-D PET cardiac imaging. *IEEE Trans Nucl Sci.* 1997;44:90–97.
35. Metz CE, Wang PL, Kronman HB. A new approach for testing the significance of differences between ROC curves measured from correlated data. In: Deconinck F, ed. *Information Processing in Medical Imaging.* The Hague, The Netherlands: Martinus Nijhoff; 1984:432–445.
36. Dorfman DD, Berbaum KS, Metz CE. Generalization to the population of readers and patients with the jackknife method. *Invest Radiol.* 1992;27:723–731.
37. Swensson RG. Unified measurement of observer performance in detecting and localizing target objects on images. *Med Phys.* 1996;23:1709–1725.
38. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging.* 1994;13:601–609.
39. Mumcuoglu EU, Leahy RM, Cherry SR, Hoffman E. Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET. In: Del Guerra A, ed. *1996 IEEE Nuclear Science Symposium and Medical Imaging Conference Record.* Piscataway, NJ: IEEE; 1996;3:1569–1573.