Comparative Evaluation of Lesion Detectability for 6 PET Imaging Platforms Using a Highly Reproducible Whole-Body Phantom with ^{22}Na Lesions and Localization ROC Analysis

Dan J. Kadrmas, PhD[1]; and Paul E. Christian, BS[2]

[1]Medical Imaging Research Laboratory, Department of Radiology, University of Utah, Salt City, Utah; and [2]Huntsman Cancer Institute and Department of Radiology, University of Utah, Salt Lake City, Utah

The lesion detectability performance of 6 PET imaging platforms has been compared using a highly reproducible whole-body phantom and localization receiver operating characteristic (LROC) analysis. **Methods:** A realistic whole-body phantom consisting of brain, thorax with lungs and liver, and pelvis with bladder was assembled and outfitted with 27 semipermanent ^{22}Na lesions of various sizes and activity concentrations. The background compartments were reproducibly filled with ^{18}F solutions. The phantom was imaged under the condition of equal emission scan time on 7 PET platforms: Advance, HR+, HR961, C-PET, IRIX, MCD, and AXIS. Imaging data were processed using manufacturer-supplied software and defaults, and LROC evaluation was performed using 11 human observers. **Results:** Near-nominal counting rates were obtained for the NaI systems, and the bismuth germanate (BGO) systems were operated well below nominal counting rates. The BGO systems provided the highest lesion detection performance, followed by the large-area dedicated NaI system, and hybrid PET gamma cameras. Lesion detectability was highly dependent on lesion size, with all systems exhibiting similar performance for 16-mm lesions but differentiated performance for lesions $\leq$12 mm. **Conclusion:** Reconstruction methodology can have a significant effect on lesion detectability. PET lesion detectability performance is correlated with system cost and imaging characteristics. For a particular imaging task, care should be taken to ensure that the scanner being used is appropriate and that the scan time is adjusted accordingly to ensure good lesion detectability.

Key Words: PET; lesion detectability; hybrid PET; localization receiver operating characteristics

J Nucl Med 2002; 43:1545–1554

$\mathbf{A}$ period of tremendous growth is being experienced by PET, which has become widely accepted for a range of cancer imaging applications. Past reimbursement policies have generally been set for the PET modality as a whole, generally based on data obtained from higher-end dedicated PET cameras. Manufacturers presently provide a wide range of dedicated and hybrid PET systems, and reimbursement policymakers are now taking into account platform-specific performance issues.

Several groups have investigated the differences in imaging performance for various PET and hybrid PET cameras (*1–11*), and numerous recent abstracts have compared hybrid PET with dedicated PET in particular (*12–20*). Most of these studies compare pairs of systems, and it is difficult to extend the conclusions to a larger number of systems. It is also difficult to perform clinical comparisons under identical physiologic conditions with controlled activity levels, and the truth regarding the presence or absence of disease is difficult to establish. Phantom work overcomes some of these limitations; however, the relatively simplistic phantoms generally available fail to provide a representative whole-body activity distribution, which is imperative if evaluations of fully 3-dimensional (3D) or counting rate–limited systems are to be performed.

The objective of this investigation was to comparatively evaluate the lesion detectability performance of a variety of PET-capable systems in as objective and reproducible a manner as possible. Acquisition times and whole-body activity distributions were held constant except where noted otherwise. A specialized whole-body phantom was assembled that used semipermanent ^{22}Na focal lesions ($t_{1/2}$ = 2.6 y) affixed within large background compartments filled with ^{18}F ($t_{1/2}$ = 110 min). The phantom was imaged under carefully controlled conditions on 7 different PET-capable imaging systems. The lesion detectability performance of 6 of the systems was then quantified and compared by performing a human observer study with localization receiver operating characteristic (LROC) analysis (*21–24*).

MATERIALS AND METHODS

Whole-Body Phantom

The whole-body phantom consisted of 3 main components as shown in Figure 1: a 3D Hoffman brain phantom (Data Spectrum
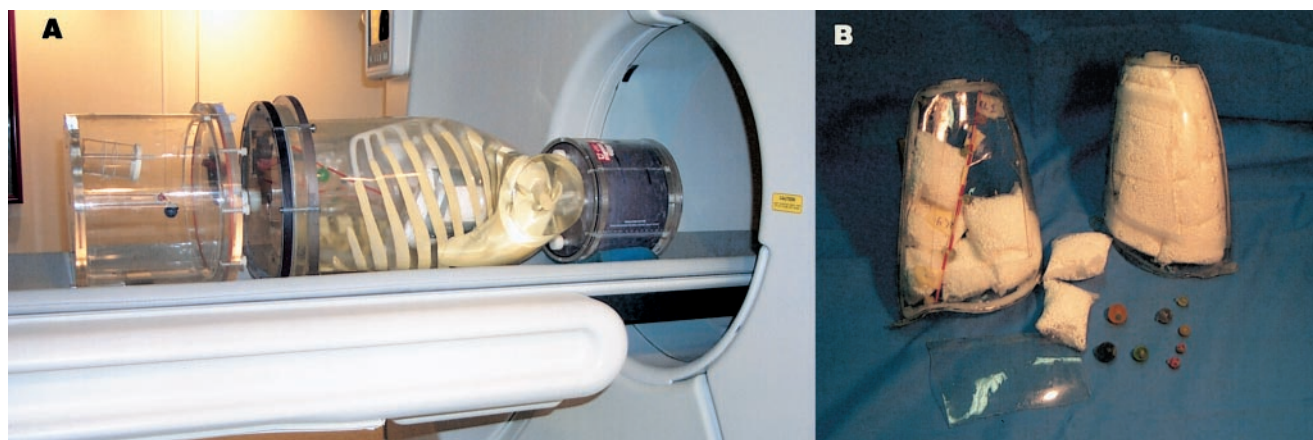
FIGURE 1. (A) Whole-body phantom consisted of head, thorax, and pelvis compartments and contained 27 semipermanent ^{22}Na lesions. (B) Lungs were custom-modified to permit passive filling. Styrofoam beads were contained using small nylon mesh bags, and lesions were permanently mounted using monofilament line.

Corp., Hillsborough, NC) (*25*), an anthropomorphic thorax phantom (Radiology Support Devices Inc., Long Beach, CA), and a $31.8 \times 23.4 \times 20.0$ cm elliptic cylinder pelvis phantom (Data Spectrum). The thorax phantom contained lungs and a liver compartment, and a 110-mL plastic bottle was placed in the pelvis to simulate the bladder. The cardiac insert and central mounting assembly of the thorax were removed so that lesions could be placed in the mediastinum. The lungs were modified to be self-filling, using nylon mesh bags to hold the Styrofoam (The Dow Chemical Co., Midland, MI) beads in place. With this setup, the lungs passively filled as the thorax compartment was filled, thereby improving reproducibility and easing phantom preparation. Activity concentration in the lungs was measured to be 0.37 times that of the soft-tissue background using this setup, and lung density averaged 0.40 g/cm^3.

Background Activity Concentrations

The activity concentrations used in the background compartments were based on 12 ^{18}F-FDG patient studies performed at our institution and were similar to representative values determined by Coleman et al. (*4*) and Turkington et al. (*26*). The values used for the standard activity phantom are listed in Table 1. The total activity present in the phantom was 111 MBq at the start of the scan. This value was chosen on the basis of estimated counting rate capabilities of the NaI-based PET cameras studied and is representative of a modest 185-MBq injected dose with 1-h delay before imaging.

^{22}Na Lesions

Lucite spheres were machined to inner diameters of 7, 8, 12, and 16 mm (measured volumes of 0.17, 0.27, 0.91, and 2.10 mL, respectively) with a uniform wall thickness of 1 mm. The lesions were filled with an epoxy mixture containing dissolved ^{22}Na activity and allowed to harden. The ^{22}Na concentrations were accurately measured using Eppendorf pipettes (Brinkmann Instruments, Inc., Westbury, NY), and food coloring was used to help ensure homogeneous mixing and to color-code activity concentrations. Wipe tests and ^{99m}Tc-labeled controls were used to monitor radioactive containment.

The lesions were designed to have activity concentrations of 4, 6, 10, and 16 times that of soft-tissue background. Because measurable decay of the ^{18}F background occurs over the duration of the scan, the semipermanent lesion activities were calibrated for the midpoint of the scans (24 min after scan start), and precise acquisition timing was maintained. Additionally, the ^{22}Na activities were set to account for the difference in relative abundance for positron emission between ^{22}Na (89.8%) and ^{18}F (100%).

Twenty-seven lesions were mounted in the phantom. Rather than using nylon rods or other mounting hardware, which displace background activity, the lesions were mounted on monofilament line using small plastic spacers. Lesion sizes and activities were mixed to cover a wide range, some of which were easily detectable on all systems and others which were undetectable even on the highest-end systems studied. The lesions were arrayed throughout the phantom as follows: 4 in the brain, 5 in the mediastinum, 4 in each lung, 4 in the liver, 4 in the abdomen, and 2 in the pelvis.

Systems Studied

The phantom was imaged on a total of 7 PET-capable systems, including 3 dedicated bismuth germanate (BGO) systems (unit cost, approximately $1.5–2.5 million), 1 dedicated NaI system (cost, approximately $1 million), and 3 hybrid PET/SPECT gamma cameras (cost, approximately $0.5–0.75 million):

- Advance (General Electric Medical Systems, Milwaukee, WI)
- ECAT EXACT HR+ (CTI PET Systems, Knoxville, TN)
- ECAT EXACT HR961 (CTI PET Systems)
- C-PET (ADAC Laboratories, Milpitas, CA)
- IRIX (Marconi Medical Systems, Cleveland, OH)
- VERTEX MCD (ADAC Laboratories)
- AXIS (Marconi Medical Systems)

TABLE 1
Whole-Body Phantom Activity Distribution

Compartment	Volume (mL)	Activity concentration (kBq/mL)	Relative concentration
Brain	~1,400	Average 21.8 (gray:white = 4:1)	6.6:1
Bladder	110	76.6	23:1
Body	17,200	3.3	1.0:1
Liver	1,025	6.7	2.0:1
Lungs	875 (L), 1,152 (R)	13.7	0.37:1

Table 2 lists the imaging performance characteristics of each of the systems according to the National Electrical Manufacturers Association (NEMA) NU2-1994 standard as reported (27–33), the Reilly Communications Group Research Report 319 (September 2000; www.itnonline.net), and the manufacturers' product data sheets. Ranges are provided where different sources were not in agreement. The Advance and HR+ systems were each <2 y old, but the HR961 was an older system installed in 1992. The C-PET is a 6-detector dedicated PET camera using curved NaI crystals and operates in 3D mode only. The MCD is a 2-head hybrid PET/SPECT gamma camera with an attenuation correction transmission system. The IRIX and AXIS are 3-head and 2-head hybrid PET cameras, respectively, and both were equipped with the γPET^{Az} electronics upgrade. The IRIX camera had a β-test Beacon transmission device, and the AXIS camera had no transmission scanning capability.

Data Acquisition

The comparative experiment was performed under the condition of identical emission scanning times for all cameras, and, unless otherwise indicated, the standard 111-MBq phantom and lesion set was used. The background compartments were precisely filled with ^{18}F water solutions calibrated for the target start time. Because the ^{22}Na lesion activities remain essentially fixed but the ^{18}F background decays measurably over the duration of the scan, the scan times were precisely controlled to ensure that the correct target-to-background ratios were achieved for each study.

The experiments were performed over a period of several weeks, during which time the ^{22}Na lesions decayed by <3%. A second C-PET experiment was also performed about a year later (22.4% lesion decay). For each experiment, the amount of ^{22}Na decay was computed and the ^{18}F background activity levels were reduced accordingly to exactly reproduce the desired target-to-background ratios. The emission scans were targeted to cover an axial extent of 86.7 cm in 48 min, and the scan time for each experiment (except the second C-PET experiment) was increased by the same factor as the ^{18}F activities were decreased so that the product of the activity present times the emission scan time was constant across all studies.

Because the axial field of view (FOV) and degree of overlap differed among the cameras studied, the emission scans were setup as follows. The number of bed positions required to scan an axial extent of at least 86.7 cm was determined, and the actual axial coverage for this number of beds, ΔZ, was computed. Note that $\Delta Z \geq 86.7$ cm. The total emission scan time was then set as closely as possible to $(\Delta Z/86.7)\cdot 48$ min. When ΔZ was >86.7 cm, the scan start time and starting position were adjusted by half of their respective disparities so that the time after injection, at which each individual portion of the phantom was imaged, was reproduced as closely as possible across all systems.

The emission scans were acquired in oncologic whole-body mode starting at the pelvis and moving toward the head. All systems were operated in 2-dimensional (2D) mode with septa or slat collimators in place, except for the C-PET, which operates exclusively in 3D mode. After emission scanning was complete, transmission scanning was performed according to manufacturer-suggested guidelines. Thus, whereas emission scan times were held constant, transmission scan times were allowed to vary between systems.

Data Processing

The data were reconstructed using manufacturer-supplied software with manufacturer-suggested default processing parameters. It is important to note that the processing parameters were not specifically optimized for the lesion detection task being studied. As listed in Table 3, each camera used iterative ordered-subsets expectation maximization (OSEM) except for the C-PET, which used Fourier rebinning (34,35) followed by the row action maximum likelihood algorithm (RAMLA) (36). Note that while the C-PET data were acquired in fully 3D mode, 2D reconstruction methods were applied; future development of fully 3D reconstruction methods would likely lead to improved image quality for this system. Randoms correction and segmented attenuation correction were performed when available.

The reconstructed images for each camera were downloaded for local display and analysis. Each image was interpolated onto a $256 \times 256 \times 512$ image matrix with cubic 2-mm voxels. The precise location of each lesion was determined using interactive coregistration techniques, accounting for slight differences in positioning between experiments. For each lesion, 4-mm-thick trans-axial images were formed by summing the 2 adjacent 2-mm-thick slices closest to the lesion center. Maximum intensity projections (MIPs) of the whole-body images were also computed.

LROC Study

An LROC study was performed using 11 human observers. Each observer was presented with 20 training images followed by 46 test images for each camera as described below. Each image consisted of a single transaxial slice that either was normal or

TABLE 2
Reported NEMA NU2-1994 Performance Characteristics for Cameras Studied

Characteristic	Advance	HR+	HR961	C-PET	IRIX	MCD	AXIS
Scintillator	BGO	BGO	BGO	1-in (2.54 cm) NaI	¾-in (1.91 cm) NaI	⅝-in (1.59 cm) NaI	¾-in (1.91 cm) NaI
Crystals (no.)	12,096	18,432	18,816	6 (curved)	3	2	2
Axial FOV (cm)	15.2	15.5	15.0	25.6	38.0	38.0	38.0
Sensitivity (kcps/μCi/cc)	228	211	177	444 (3D)	~25	72	~20
Transaxial FWHM (mm)*	3.8–4.8	4.3–4.6	3.7	4.6–5.0	4.8	4.5–5.2	4.8
Axial FWHM (mm)	4.2–6.0	4.2	4.0	5.5–5.7	6.0	4.9–5.0	6.0
Scatter fraction (%)	10	17	14	25 (35)†	24	30–37	20

*Spatial resolution measures are at center of FOV.
†Scatter fraction in 3D mode is more appropriately measured according to NU2-2001 standard and was 35% for C-PET.

TABLE 3
Reconstruction Parameters for Each Camera

Parameter	Advance	HR+	HR961	C-PET	IRIX	MCD	AXIS
Prj. matrix	283×35×336	288×63×144	196×47×336	128×64×96×7 tilts	128×83×90	128×96×96	128×83×90
Algorithm	OSEM	OSEM	OSEM	FORE > RAMLA	OSEM	OSEM	OSEM
Recon.	2 iterations	2 iterations	6 iterations	1 iteration	4 iterations	2 iterations	4 iterations
Parameters	28 subsets	8 subsets	16 subsets	$\lambda = 0.1$	15 subsets	8 subsets	15 subsets
Randoms correction	Delays	Delays	Delays	Background subtraction	No	Background subtraction	No
Attenuation correction	Segmented	Segmented	Segmented	Segmented	Yes	Segmented	No
Scatter correction	Deconvolution	Deconvolution	Deconvolution	Background subtraction	No	Background subtraction	No
Postfilter	8-mm gaussian	8-mm gaussian	8-mm gaussian	None	6-mm gaussian	Weiner 0.67	Butterworth 10, 0.30/cm
Voxel width (mm)	4.30	5.1	4.53	4.0	4.67	4.0	4.67
Slice thickness (mm)	4.25	5.1	3.13	4.0	4.67	4.0	4.67

Prj. = projection; OSEM = ordered-subsets expectation maximization; FORE = Fourier rebinning; RAMLA = row action maximum likelihood algorithm; Recon. = reconstruction.

contained exactly 1 lesion. In cases where a second lesion overlapped or intruded on the slice of interest, the intruding lesion was masked by blacking out a 24 × 24 pixel region containing the lesion. Additional masked regions were randomly placed on a similar proportion of the normal images so that the presence of a masked region did not provide any information as to whether or not there was a lesion present.

To use all 27 lesions for test purposes, simulated lesions were added to normal slices for training purposes. This was done by grabbing a true lesion and its immediate local area from a positive slice, translating it a random distance, and then blending it onto a normal slice. Care was taken to ensure that there was no visual indication on the final training images that the lesions were simulated.

Eleven observers participated in the LROC study, including 3 trained nuclear medicine physicians, 5 PhD researchers, and 3 graduate students in the Medical Imaging Research Laboratory. The images were read over 2 sessions approximately 1 wk apart. For each camera, the reader was presented with 20 training images followed by 46 test images. Approximately half of the images were normal and half were abnormal. The observer was asked to provide 2 responses for each image: a discrete 6-point confidence rating that a lesion was either present or absent, and the location that the observer felt was most likely to contain a lesion. The location data were entered by flagging crosshairs controlled by a computer mouse, and the most likely location was entered regardless of which confidence rating was chosen.

To minimize the effects of reading order, the order of the images presented for each camera was randomized, and the order in which the sets of images for each camera were presented to each reader was also randomized. The readings took place in a darkened room, and the display monitor settings were calibrated so that the gray scale provided a logarithmic–linear relationship (37). No restraint was placed on reading time, and each reading session averaged about 1 h.

The observer data were analyzed using the LROCFIT program (21) to obtain LROC curves and the probability of correct local-ization for each camera. The LROC results were analyzed by a 2-way ANOVA, and differences between individual cameras were tested using the Scheffé multiple comparisons test (38). No significant difference in performance between observers was measured. The curves were averages over all observers to obtain the average area under the LROC curve and the average probability of correct localization for each camera. These values were used as figures of merit for the analysis.

RESULTS

Statistics and Counting Rates

The lesion detectability performance of a particular system will depend on its imaging characteristics, including detection efficiency, spatial resolution, 2D versus 3D acquisition mode, reconstruction methodology, and compensation for attenuation, randoms, and scatter. The reported performance characteristics of the systems studied are listed in Table 2, and the imaging statistics for our series of experiments are given below. The count levels for each experiment correlated well with the corresponding NEMA sensitivities. The Advance and HR+ acquired 35 and 36 million coincidence events (after randoms correction), respectively, and the HR961 acquired 28 million. Note that these systems were each operated well below their nominal counting rates, representative of a clinical center using low-to-moderate doses.

Because of lower detection efficiency and dead-time limitations, the 2D NaI systems acquired considerably fewer events. The MCD dual-head system had a singles rate of 2.2 Mcps (1.1 Mcps per head) at the start of the scan, with a nominal singles rate of ~2.0 Mcps. This system acquired a total of 7.2 million coincidence events (trues + scatter + randoms). The initial singles rate for the AXIS dual-head system was 2.0 Mcps, and 6.5 million coincidence events

were acquired. However, the AXIS had γPET^{Az} electronics with a nominal singles rate of ~4 Mcps. Because the AXIS data were acquired well below the nominal counting rate, and because it had no transmission scanning capability, the AXIS results were excluded from the quantitative analysis of lesion detectability.

The 3-head IRIX system also had the γPET^{Az} electronics upgrade. To operate at its nominal counting rate, a second set of lesions was constructed with twice the ^{22}Na activity as the standard set, and the experiment was performed with twice the activity in the background compartments as well. With this setup, the IRIX had an initial singles rates of 5.9 Mcps (~2.0 Mcps per head), and 13 million coincidence events were acquired. The nominal counting rate was ~5.5

Mcps; hence, the IRIX was slightly above its recommended counting rate at the start of the scan.

The 6 detector C-PET camera had a singles rate of 7.0 Mcps in its first experiment. Picchio et al. (20) recently showed that the C-PET lesion detectability performance drops when the camera is operated outside the range of 3–5 Mcps. Thus, a second C-PET experiment was performed 1 y later, allowing the ^{22}Na lesions to decay to 77.6% of their initial activities. The background activity was reduced from 111 to 85.1 MBq, accordingly, and this experiment yielded an initial singles rate of 5.0 Mcps (~800 kcps per head). Because the counting rates for this second experiment (2.9–5.0 Mcps) match the reported nominal range, these data were used for all subsequent analysis. The C-PET scan
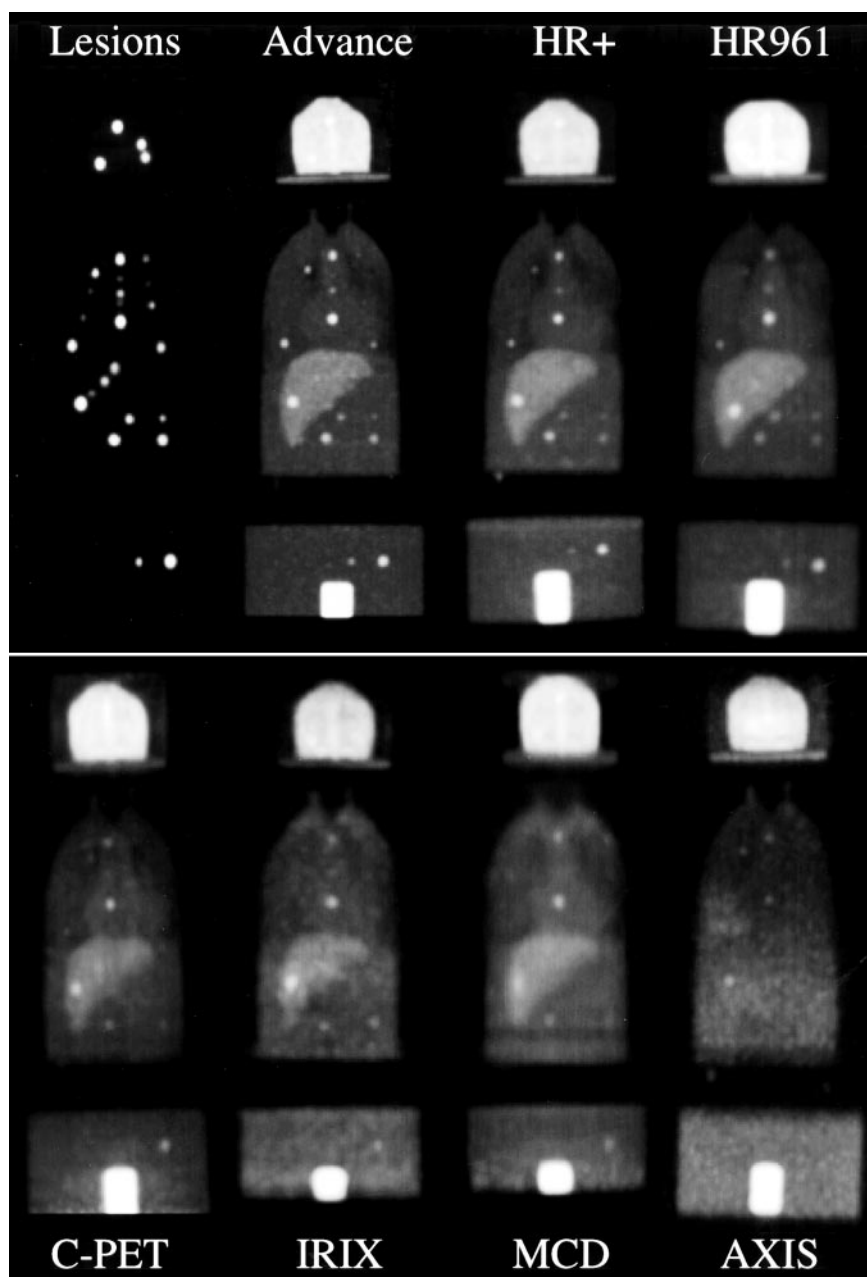

FIGURE 2. Coronal MIPs of whole-body phantom for each camera studied. Note that BGO systems (top row) were operated well below nominal counting rates, and NaI systems (bottom row) were operated at near-nominal counting rates. AXIS image is shown without attenuation correction because no transmission source was available on that system.

acquired a total of 103 million coincidence events; however, because this system was operated in 3D mode, a higher proportion of these events included scatters and randoms (actual numbers not available). Thus, a direct comparison between the number of coincidence events obtained by this system with the other 2D systems may not be appropriate.

Image Quality

The imaging results are summarized in Figure 2, which shows coronal MIPs for each camera studied. Significant differences in visual image quality can be observed for the different cameras, and there is a reduction in the number of lesions that can be easily visualized as one moves from the dedicated to the hybrid systems. There was also a subtle artifactual background in the shoulder region visible on the images from the hybrid PET systems that did not appear for any of the dedicated systems. The shoulder region of the thorax phantom consisted of a thick plastic wall with embedded bone structures; hence, it had significant attenuating medium but no background activity. We postulate that the artifactual background in the shoulder region for the hybrid PET systems arises from incomplete compensation for scatter or randoms.

Figure 3 show transaxial images—for example, lung, mediastinal, and liver lesions. These images demonstrate marked differences in lesion visualization for the different cameras, and differences in noise texture, contrast, and streak artifacts between systems can also be seen. The differences in image quality are due in part to the varying resolution and imaging statistics achieved by the various hardware used, but differences in data processing also play a major role.

LROC Study

The radius of correct localization for the analysis was chosen on the basis of Figure 4, which plots the fraction of lesions that were correctly localized versus the radius threshold. The threshold value used was 5 pixels (10 mm), which intersects the curves in regions of low variability. Thus, a small change in the choice of localization threshold would have little effect on the results.

Figure 5 depicts the LROC curves—correctly localized true-positive fraction (TPF) plotted versus false-positive fraction (FPF)—for each camera averaged over all observers. The results of the analysis are also shown quantitatively in Table 4, where the areas under the LROC curves and the probabilities of correct localization are listed. These results show a progression of improved lesion detectability performance when moving from hybrid cameras toward high-end dedicated systems, as expected. Although a significant difference in lesion detectability performance was measured between the Advance and the HR+, such a difference would not be expected given their similarities in technical specifications. This difference may be due in large part to the different processing parameters applied on the different systems, which is discussed further in the Discussion. Although the HR961 had the best-reported NEMA spatial
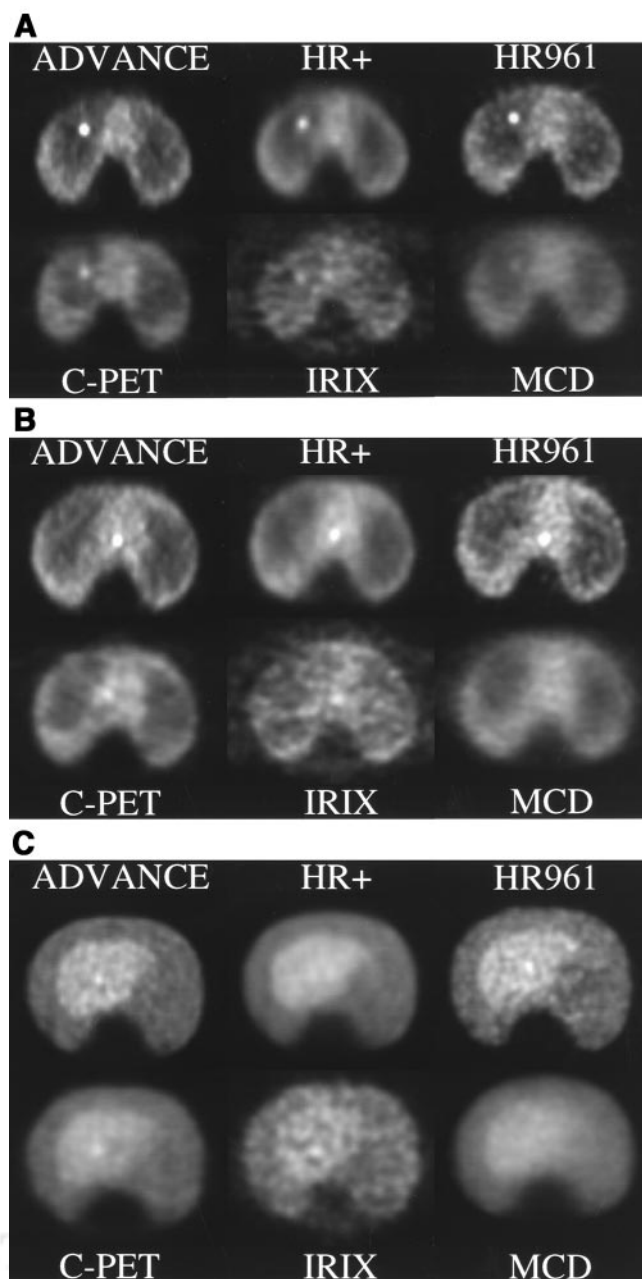

FIGURE 3. Transaxial images of 8-mm-diameter lesion in right lung at target-to-background ratio of 43:1 (A), 7-mm-diameter lesion in mediastinum at 16:1 (B), and 12-mm-diameter lesion in liver at 3:1 (C). Images demonstrate significant differences in lesion visualization for various systems. Marked differences in speckle noise patterns are also seen, a direct result of various count levels and data-processing parameters of each system.

resolution, it was outperformed by both the Advance and the HR+. This was partly a result of its lower sensitivity, although data processing may have played a role as well.

Effect of Lesion Size and Concentration

The percentage of correctly localized lesions for each camera is plotted versus lesion size and activity concentra-
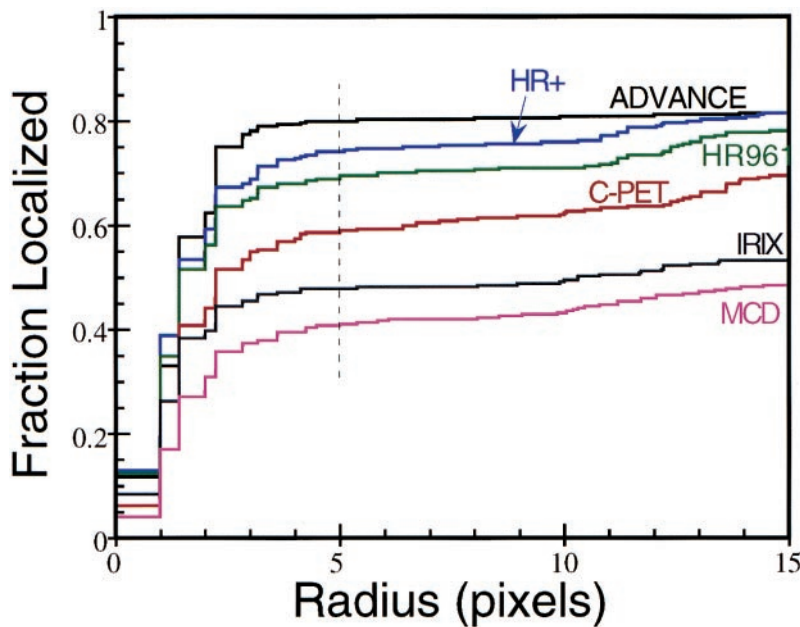
FIGURE 4. Fraction of correct tumor localizations plotted as function of radius of correct localization. Dashed vertical line indicates threshold chosen for LROC analysis (5 pixels = 10 mm), which marks point of low variability for each curve.

tion in Figure 6. Note that all cameras performed similarly for the largest (16 mm) lesions, and individual performance varied for smaller lesions. There was an overall trend of improved performance for hotter lesions, as expected, but Figure 6 also shows differences in performance at each particular lesion activity concentration. These results demonstrate that the relative performance of the individual systems studied depends on both the size and the activity concentration of smaller lesions. Given these results, care should be taken to ensure that each instrument is used appropriately in a clinical setting. For example, all of the systems studied may perform adequately for evaluating single pulmonary nodules that are ≥16 mm in size. However, for a whole-body task such as staging melanoma, in

which small distant metastases may be clinically significant, we recommend using the highest-end systems possible.

DISCUSSION

The results of this study show a clear trend in the rankings of the various systems studied, from block-design BGO scanners to the large-area NaI scanner and multihead coincidence-detection gamma cameras. The system rankings correlate with price and with the performance characteristics listed in Table 2 (except that the sensitivity of the 3D mode C-PET doesn't directly compare with those of the 2D systems). There are few papers in the literature comparing the lesion detectability performance of dedicated PET cam-
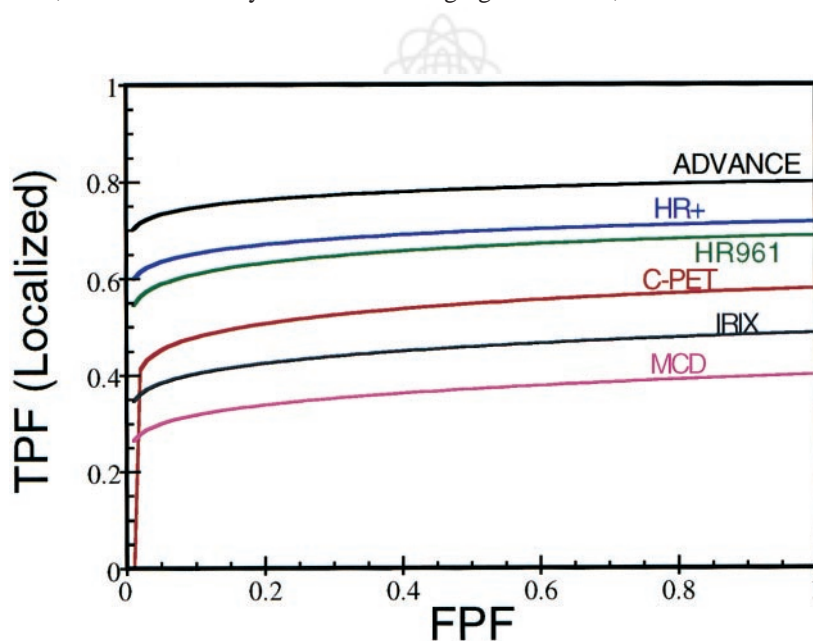

FIGURE 5. LROC curves for each camera averaged over all observers. Area under curve measures lesion detectability performance, and intersection at false-positive fraction (FPF) = 1.0 indicates probability of correct localization for each system. TPF = true-positive fraction.

TABLE 4
LROC Results Averaged Over All Observers

Camera	P_{CL}	A_{LROC}	P
Advance	0.799 ± 0.083	0.777 ± 0.084	≤ 0.0001
HR+	0.716 ± 0.091	0.689 ± 0.089	NS
HR961	0.688 ± 0.094	0.655 ± 0.091	< 0.0001
C-PET	0.578 ± 0.103	0.537 ± 0.101	
IRIX	0.487 ± 0.107	0.449 ± 0.104	0.0005
MCD	0.400 ± 0.110	0.363 ± 0.102	0.0002

P_{CL} = probability of correct localization; A_{LROC} = area under LROC curve; NS = not significant.

eras. Picchio et al. (20) compared the Advance and the C-PET in 28 patients, reporting the same lesion detection sensitivity when the C-PET counting rate was in its nominal range. A single ^{18}F-FDG injection was used, and the C-PET scans were acquired hours later when activity levels had decayed to appropriate levels. Because lesion uptake and contrast are higher 2–3 h after injection than at 1 h (39,40), physiologic conditions benefited the C-PET in that study. Using reproducible phantom methodology to obtain identical imaging conditions, we identified a difference in lesion detectability for these 2 cameras.

Numerous clinical studies have compared coincidence-detection hybrid gamma cameras with dedicated PET (1,3,5–18). The reported sensitivity of hybrid PET ranged from 55% to 95% (average, 78%) of that of dedicated PET, and there is general agreement that hybrid PET performs well for large lesions, but sensitivity drops significantly for lesions ≤ 10–15 mm in diameter. Our results are consistent with these findings. Phantom comparison work has also been done by Coleman et al. (4) and Yutani et al. (8). Again, their results were consistent with our findings with regard to both comparative sensitivity and size-related detectability issues.

Although the design of our study was set up to be as objective as possible, logistic and practical issues placed subjective constraints on data processing. Recall that the processing parameters for each camera were chosen on the basis of manufacturer-determined defaults. Because the processing parameters were not specifically optimized for each individual camera, this study did not explicitly compare camera hardware; rather, the combined hardware/software packages were compared. Different software parameters could therefore affect the results of the study.

Consider, in particular, the Advance and the HR+ data. These systems have similar specifications, but we measured a significant difference in lesion detectability performance. Figure 7 shows transaxial images of an 8-mm lesion in the left lung. The HR+ image (Fig. 7, left) was reconstructed using 2 iterations of OSEM with 8 subsets, roughly equivalent to 16 maximum-likelihood expectation maximization

(MLEM) iterations. This may be insufficiently converged to recover the resolution and definition necessary to detect small focal lesions. The HR+ data were reprocessed using 7 iterations OSEM with 8 subsets (~ 56 MLEM iterations), shown in the middle image in Figure 7. This resulted in much improved lesion definition and visualization and was comparable with the Advance image (Fig. 7, right). Given the similarity in hardware specifications, we postulate that the performance difference between the Advance and the HR+ measured in this study was a result of suboptimal data processing and is not representative of the HR+ hardware capabilities. Care should be taken in differentiating between hardware and software issues when drawing conclusions from this work.

Our study held the acquisition time constant for all systems, in effect evaluating performance under fixed throughput conditions. In practice, longer scan times are generally recommended for the lower-end systems to make up for detection efficiency or counting rate limitations. Although our results provide a generalized overall comparison, the actual performance of a particular camera could be improved by increasing the scan time at the expense of reduced throughput. A full cost comparison and discussion of issues related to patient discomfort and motion artifacts is beyond the scope of this paper. Along similar lines, the performance of the BGO systems could be improved by increasing the dose, whereas the large-area NaI systems were operated at nominal counting rates. For example, we repeated the Advance experiment using the $2\times$ activity phantom, which increased the area under the LROC curve from 0.777 to 0.874 ($P < 0.001$). All of these issues need to be considered in determining which instrument is appropriate for particular imaging centers and clinical applications.

CONCLUSION

To our knowledge, this work represents the first detailed comparison of a large number of PET-capable systems using a highly reproducible whole-body phantom. Such a comparison is complex, and the particular details of the study should be considered when evaluating the results. Note that this study evaluated combined hardware/software platforms using clinically representative processing parameters. Differences in performance were measured between the dedicated BGO systems, but this was likely due to differences in data processing and may not represent a disparity in hardware performance. A significant difference in lesion detectability performance was measured between the dedicated BGO scanners and the current NaI-based systems, which have lower detection efficiencies and more limited counting rate capabilities. Increased scan times, or technologic advances, would be required for large-area NaI-based systems to become more competitive. The next generation of PET scanners will include pixelated NaI systems, lutetium oxyorthosilicate and cerium-doped gadolinium orthosilicate (GSO) scintillators, and advanced fully 3D ge-
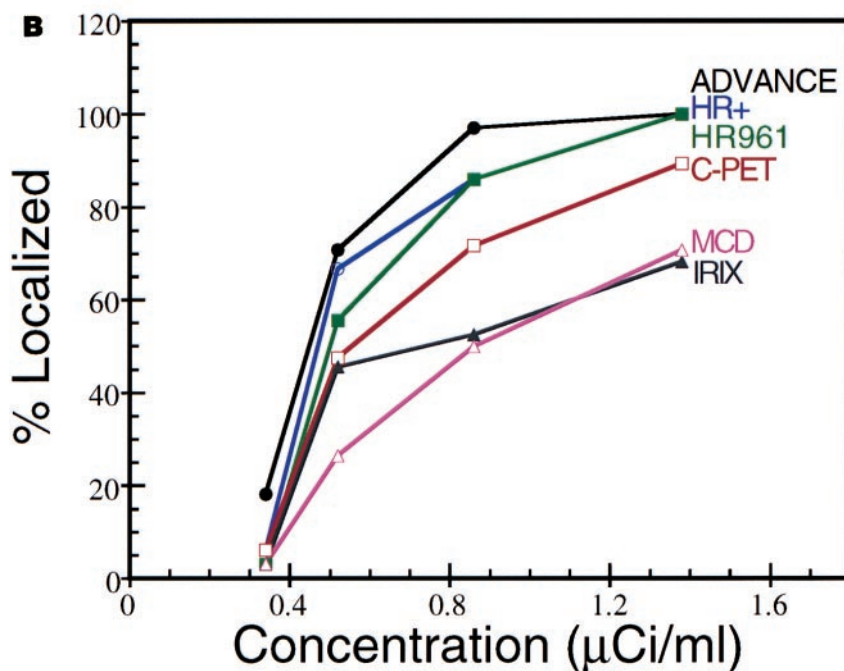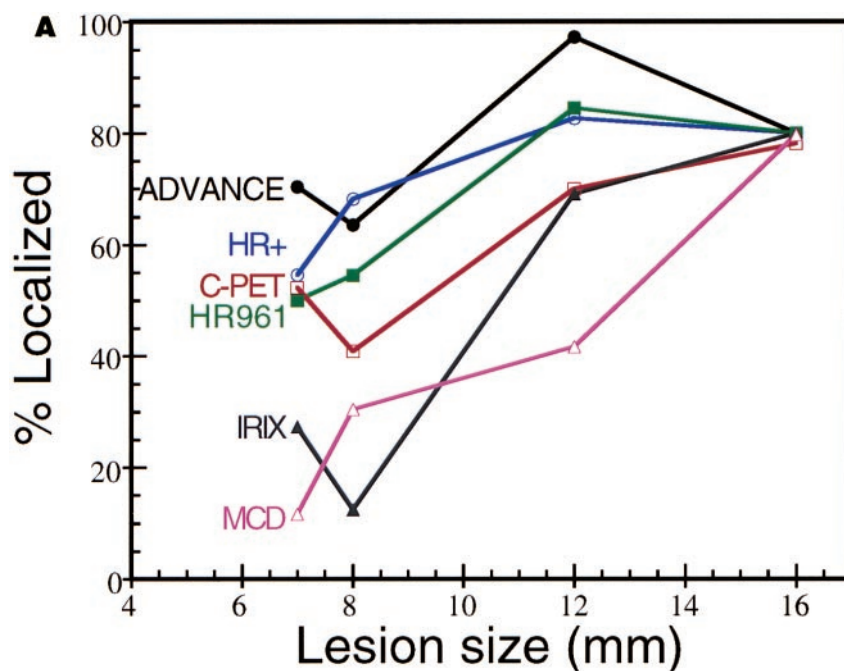
FIGURE 6. Percentage of lesions correctly localized in observer study for each camera plotted as function of lesion size (A) and lesion activity concentration (B). Each camera studied gave similar performance for largest (16 mm) lesions, and differences in individual performance were observed for smaller lesions. Each camera also obeyed trend of improved performance for higher activity concentrations, as expected.
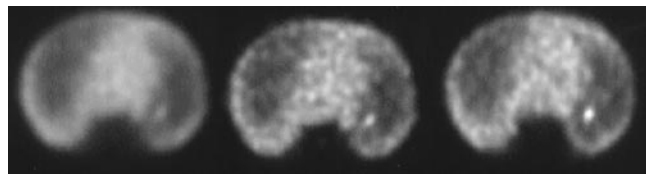

FIGURE 7. Images of 8-mm-diameter lesion in left lung for HR+ processed with default 2 iterations (left) and with 7 iterations (middle) compared with Advance image (right). Reprocessed HR+ image shows much better lesion definition and visualization. Results suggest that HR+ performance would have been comparable with Advance performance if improved data-processing parameters had been used.

ometries that enhance and refine features of the scanners studied here to improve performance. It is hoped that this work will guide future development of PET hardware and also serve to highlight the crucial role that reconstruction algorithms and data-processing parameters play in determining lesion detectability performance.

ACKNOWLEDGMENTS

REFERENCES

1. Shreve PD, Steventon RS, Deters EC, Kison PV, Gross MD, Wahl RL. Oncologic diagnosis with 2-[fluorine-18]fluoro-2-deoxy-d-glucose imaging: dual-head coincidence gamma camera versus positron emission tomographic scanner. *Radiology.* 1998;207:431–437.

2. Adam LE, Karp JS, Smith RJ. PET camera performance measurements: a comparison between three PET cameras [abstract]. *J Nucl Med.* 1999;40(suppl):76P.

3. Boren EL Jr, Delbeke D, Patton JA, Sandler MP. Comparison of FDG PET and positron coincidence detection imaging using a dual-head gamma camera with 5/8-inch NaI(Tl) crystals in patients with suspected body malignancies. *Eur J Nucl Med.* 1999;26:379–387.

4. Coleman RE, Laymon CM, Turkington TG. FDG imaging of lung nodules: a phantom study comparing SPECT, camera-based PET, and dedicated PET. *Radiology.* 1999;210:823–828.

5. Delbeke D, Patton JA, Martin WH, Sandler MP. FDG PET and dual-head gamma camera positron coincidence detection imaging of suspected malignancies and brain disorders. *J Nucl Med.* 1999;40:110–117.

6. Landoni C, Gianolli L, Lucignani G, et al. Comparison of dual-head coincidence PET versus ring PET in tumor patients. *J Nucl Med.* 1999;40:1617–1622.

7. Tatsumi M, Yutani K, Watanabe Y, et al. Feasibility of fluorodeoxyglucose dual-head gamma camera coincidence imaging in the evaluation of lung cancer: comparison with FDG PET. *J Nucl Med.* 1999;40:566–573.

8. Yutani K, Tatsumi M, Shiba E, Kusuoka H, Nishimura T. Comparison of dual-head coincidence gamma camera FDG imaging with FDG PET in detection of breast cancer and axillary lymph node metastasis. *J Nucl Med.* 1999;40:1003–1008.

9. Zimny M, Kaiser HJ, Cremerius U, et al. F-18-FDG positron imaging in oncological patients: gamma camera coincidence detection versus dedicated PET. *Nuklearmedizin.* 1999;38:108–114.

10. Tiepolt C, Beuthien-Baumann B, Hliscs R, et al. ^{18}F-FDG for the staging of patients with differentiated thyroid cancer: comparison of a dual-head coincidence gamma camera with dedicated PET. *Ann Nucl Med.* 2000;14:339–345.

11. Inoue T, Oriuchi N, Koyama K, et al. Usefulness of dual-head coincidence gamma camera with thick NaI crystals for nuclear oncology: comparison with dedicated PET camera and conventional gamma camera with thin NaI crystals. *Ann Nucl Med.* 2001;15:141–148.

12. Al-Sugair A, Hamblem SM, Laymon CE, Turkington TG, Coleman RE. Lesion detection using nonattenuation corrected dedicated and camera-based PET compared with attenuation corrected dedicated PET [abstract]. *J Nucl Med.* 1998;39(suppl):109P.

13. Shreve P, Steventon RS, Deters E, Gross MD, Wahl RL. Lesion detection in oncologic diagnosis: comparison of dual head coincidence with dedicated PET FDG imaging [abstract]. *J Nucl Med.* 1998;39(suppl):109P.

14. Segall GM, Carlisle M, Bocher M. Prospective comparison of coincidence imaging versus dedicated PET in patients with cancer [abstract]. *J Nucl Med.* 1999;40(suppl):135P.

15. Tiepolt C, Beuthien-Baumann B, Kuhne A, et al. Diagnostic accuracy of F-18 FDG in differentiated thyroid carcinoma: comparison of a dedicated PET scanner with a dual head coincidence camera [abstract]. *J Nucl Med.* 1999;40(suppl):135P.

16. Kropp J, Naumann R, Schirmer L, et al. Comparison of dedicated PET with hybrid PET in patients with malignant lymphomas [abstract]. *J Nucl Med.* 2000;41(suppl):71P.

17. Rakotonirina H, Delahaye N, Crestani B, et al. Preoperative staging of patients with non-small cell lung cancer using F-18-FDG imaging: comparison of dual-head coincidence gamma-camera versus dedicated PET [abstract]. *J Nucl Med.* 2000;41(suppl):292P.

18. Tatsumi M, Tokita N, Nishimura T. Whole-body hybrid positron emission tomography with F-18-FDG in the staging of non-Hodgkin's lymphoma [abstract]. *J Nucl Med.* 2000;41(suppl):292P.

19. Caldwell CB, Ehrlich LE, Ganguli SN. A comparison of dual and triple-detector hybrid PET [abstract]. *J Nucl Med.* 2001;42(suppl):98P.

20. Picchio M, Savi A, Landoni C, et al. Effect of count rate on clinical performance of a large NaI crystal PET scanner: comparison with a multi-ring PET [abstract]. *J Nucl Med.* 2001;42(suppl):205P.

21. Swensson RG. Unified measurement of observer performance in detecting and localizing target objects on images. *Med Phys.* 1996;23:1709–1725.

22. Farquhar TH, Llacer J, Hoh CK, et al. ROC and localization ROC analyses of lesion detection in whole-body FDG PET: effects of acquisition mode, attenuation correction and reconstruction algorithm. *J Nucl Med.* 1999;40:2043–2052.

23. Gifford HC, King MA, Wells RG, Hawkins WG, Narayanan MV, Pretorius PH. LROC analysis of detector-response compensation in SPECT. *IEEE Trans Med Imaging.* 2000;19:463–473.

24. Farquhar TH, Llacer J, Sayre J, Tai YC, Hoffman EJ. ROC and LROC analyses of the effects of lesion contrast, size, and signal-to-noise ratio on detectability in PET images. *J Nucl Med.* 2000;41:745–754.

25. Hoffman EJ, Cutler PD, Digby WM, Mazziotta JC. 3-D phantom to simulate cerebral blood flow and metabolic images for PET. *IEEE Trans Nucl Sci.* 1990;37:616–620.

26. Turkington TG, Williams NE, Hamblen SM. Regional FDG uptake, attenuation, and geometries measurements for whole body phantom design [abstract]. *J Nucl Med.* 1999;40(suppl):281P.

27. Sossi V, Bruckbauer T, Wienhard K. Comparison of the performance of the SSRB technique on the Siemens/CTI ECAT 921 and 961 PET tomographs with the ECAT 953B. In: *Proceedings of the 1993 IEEE Nuclear Science Symposium and Medical Imaging Conference.* San Francisco, CA: IEEE Press; 1993;2:1085–1087.

28. De Grado TR, Turkington TG, Williams JJ, et al. Performance characteristics of a whole-body PET scanner. *J Nucl Med.* 1994;35:1398–1406.

29. Levin CS, Tai Y-C, Hoffman EJ, Dahlbom M, Farquhar TH. Removal of the effect of Compton scattering in 3-D whole body positron emission tomography by Monte Carlo. In: *Proceedings of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference.* San Francisco, CA: IEEE Press; 1995;2:1050–1054.

30. Brix G, Zaers J, Adam LE, et al. Performance evaluation of a whole-body PET scanner using the NEMA protocol. *J Nucl Med.* 1997;38:1614–1623.

31. Adam LE, Karp JS, Daube-Witherspoon ME, Smith RJ. Performance of a whole-body PET scanner using curve-plate NaI(Tl) detectors. *J Nucl Med.* 2001;42:1821–1830.

32. Sossi V, Pointon B, Boudoux C, et al. NEMA NU 2–2000+ performance measurements on an ADAC MCD camera. *IEEE Trans Nucl Sci.* 2001;48:1518–1523.

33. Wienhard K, Dahlbom M, Eriksson L, et al. The ECAT EXACT HR: performance of a new high resolution positron scanner. *J Comput Assist Tomogr.* 1994;18:110–118.

34. Defrise M, Kinahan PE, Townsend DW, Michel C, Sibomana M, Newport DF. Exact and approximate rebinning algorithms for 3D PET data. *IEEE Trans Med Imaging.* 1997;16:145–158.

35. Defrise M, Liu X. A fast rebinning algorithm for 3D positron emission tomography using John's equation. *Inv Prob.* 1999;15:1047–1065.

36. Browne J, DePierro AR. A row-acting alternative to the EM algorithm for maximizing likelihoods in emission tomography. *IEEE Trans Med Imaging.* 1996;15:687–699.

37. Nawfel RD, Chan KH, Wagenaar DJ, Judy PF. Evaluation of video gray-scale display. *Med Phys.* 1992;19:561–567.

38. Zar JH. *Biostatistical Analysis.* Upper Saddle River, NJ: Prentice-Hall, Inc.; 1999.

39. Kubota K, Itoh M, Ozaki K, et al. Advantage of delayed whole-body FDG-PET imaging for tumour detection. *Eur J Nucl Med.* 2001;28:696–703.

40. Boerner AR, Weckesser M, Herzog H, et al. Optimal scan time for fluorine-18 fluorodeoxyglucose positron emission tomography in breast cancer. *Eur J Nucl Med.* 1999;26:226–230.