

---

---

# Impact of Acquisition Geometry, Image Processing, and Patient Size on Lesion Detection in Whole-Body $^{18}\text{F}$ -FDG PET

Georges El Fakhri<sup>1</sup>, Paula A. Santos<sup>2,1</sup>, Ramsey D. Badawi<sup>3</sup>, Clay H. Holdsworth<sup>4</sup>, Annick D. Van Den Abbeele<sup>5</sup>, and Marie Foley Kijewski<sup>1</sup>

<sup>1</sup>Division of Nuclear Medicine, Department of Radiology, Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts; <sup>2</sup>Biophysics and Biomedical Engineering Institute, Faculty of Sciences, University of Lisbon, Lisbon, Portugal; <sup>3</sup>Department of Radiology, University of California Davis Medical Center, Sacramento, California; <sup>4</sup>Massachusetts College of Pharmacy and Health Sciences, Boston, Massachusetts; and <sup>5</sup>Department of Radiology, Dana Farber Cancer Institute, Boston, Massachusetts

The aim of this work was to develop a rigorous evaluation methodology to assess performance of different acquisition and processing methods for variable patient sizes in the context of lesion detection in whole-body  $^{18}\text{F}$ -FDG PET. **Methods:** Fifty-nine bed positions were acquired in 32 patients in 2-dimensional (2D) and 3-dimensional (3D) modes 1–4 h after  $^{18}\text{F}$ -FDG injection (740 MBq) using a BGO PET scanner. Three spheres (1.0-, 1.3-, and 1.6-cm diameter) containing  $^{68}\text{Ge}$  were also imaged separately in air, at locations corresponding to possible lesion sites in 2D and 3D (590 targets per condition). Each bed position was acquired for 7 min in 2D and 6 min in 3D and corrected for randoms using delayed window randoms subtraction (DWS) or randoms variance reduction (RVR). Sphere sinograms were attenuated using the 2D or 3D attenuation map derived from the transmission scan of the patient, after scaling 2D and 3D sinograms with identical factors to ensure marginal detectability. Resulting 2D sinograms were reconstructed with filtered backprojection (FBP) and ordered-subsets expectation maximization (OSEM) without any scatter or attenuation correction (FBP-NATS and OSEM-NATS) or corrected for scatter and attenuation and reconstructed using FBP (FBP-ATT) or attenuation-weighted OSEM (AWOSEM). 3D sinograms were processed identically after Fourier rebinning. Next, reconstructed volumes were compared on the basis of performance of a 3-channel Hotelling observer (CHO-SNR [SNR is signal-to-noise ratio]) in detecting the presence of a sphere of unknown size on an anatomic background while modeling observer noise. The noise equivalent count (NEC) rate was computed in 2D and 3D for 3 different phantom sizes (40, 60, and 95 kg) and compared with lesion detection SNR. **Results:** 3D imaging yielded better lesion detectability than 2D ( $P < 0.025$ , 2-tailed paired  $t$  test) in patients of normal size (body mass index [BMI]  $\leq 31$ ). However, 2D imaging yielded better lesion detectability than 3D in large patients (BMI  $> 31$ ), as 3D performance deteriorated in large patients ( $P < 0.05$ ). 2D and 3D yielded similar results for different lesion sizes. CHO-SNR were 40% greater for AWOSEM, FBP-ATT, and

FBPNAT than for OSEM ( $P < 0.05$ ), and AWOSEM yielded significantly better lesion detectability than did FBP. In all patients, RVR yielded a systematic improvement in CHO-SNR over DWS in both 2D and 3D.  $\sqrt{\text{NEC}}$  was characterized by a behavior similar to that of  $\text{SNR}_{\text{CHO}}$  for the 3 different phantom sizes considered in this study.

**Key Words:** whole-body  $^{18}\text{F}$ -FDG PET; 2D/3D acquisition; lesion detection; numeric observer

**J Nucl Med 2007; 48:1951–1960**  
DOI: 10.2967/jnumed.108.007369

---

**T**hree-dimensional (3D) “septumless” whole-body (WB)  $^{18}\text{F}$ -FDG PET yields increased sensitivity to true coincidences at the expense of increased sensitivity to scattered and random coincidences. Several groups have assessed the relative merits of 3D and 2-dimensional (2D) WB-PET for BGO-based PET scanners using numeric simulations, phantom studies, or patient studies, applying various metrics. Raylman et al. (1), in a single-observer subjective visibility phantom study, reported equal performance for 2D and 3D PET, whereas Moore et al. (2) reported higher nonprewhitening signal-to-noise ratio for 3D than for 2D in a study using an anthropomorphic torso phantom. On the other hand, Farquhar et al. (3) reported significantly better performance in detecting simulated lesions on normal patient background images with 2D than with 3D WB  $^{18}\text{F}$ -FDG PET, with a 3D performance close to the line of chance. Lartizien et al. (4,5) reported, on the basis of human and numeric observer detectability studies using numeric simulations of the MCAT torso phantom (Mathematical Cardiac Torso phantom), better performance with 2D PET and a 740-MBq injection of  $^{18}\text{F}$ -FDG than with 3D PET with a 444-MBq injection. They found, however, that detectability with 3D PET was greater than or similar to that of 2D PET if a 444-MBq injection was simulated for both modes. Visvikis et al. (6) reported that human observers, asked to subjectively assess WB scans, rated

---

Received Jul. 29, 2006; revision accepted Aug. 24, 2007.  
For correspondence or reprints contact: Georges El Fakhri, PhD, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115.  
E-mail: elfakhri@bwh.harvard.edu  
COPYRIGHT © 2007 by the Society of Nuclear Medicine, Inc.

2D and 3D images similarly when overall acquisition time was held constant. LSO scanners are generally expected to perform better than BGO scanners in 3D mode due to higher count-rate capability, better timing resolution, and, therefore, better randoms rejection, as well as better energy resolution, and higher energy threshold, and, therefore, better scatter rejection. For LSO scanners, Lodge et al. found reduced image variability in patient images for 3D mode, under conditions of matched target-to-background ratio (7).

The aim of this work was to develop a rigorous evaluation methodology that allows objective assessment of the performance of different PET scanner geometries and processing schemes for variable patient sizes and that is based on a clinically meaningful task—that is, lesion detection in WB-PET. Actual patient data, rather than phantom or experimental data, acquired in multiple anatomic positions were used so that a large range of patient sizes was represented, as our hypothesis was that the relative merits of 2D and 3D imaging depend on patient size and that patient data, in contrast with phantom experiments, was crucial to capture the physiologic variability of  $^{18}\text{F}$ -FDG uptake in different parts of the body (8). Objective, rather than subjective or visual, assessment of image quality was performed. In addition to assessing the relative merits of 2D and 3D imaging, we assessed the impact of the processing scheme (randoms and attenuation correction, tomographic reconstruction) on image quality for different lesion sizes and patient sizes.

## MATERIALS AND METHODS

### Patient Studies

Thirty-two patients were included in this study, which was approved by the institutional review board of the Dana-Farber Cancer Institute, after having given informed consent consistent with institutional review board guidelines. Forty-four bed positions were acquired in 22 patients with various types of cancer randomly selected from the clinical population. In addition, 15 bed positions in both 2D and 3D modes were acquired in 10 patients participating in a pilot study of an antitumor agent for prostate cancer, which included  $^{18}\text{F}$ -FDG PET. Therefore, the total number of bed positions included in this study was 59, obtained from 32 patients. The body mass index ( $\text{BMI} = \text{weight (kg)}/\text{height}^2 (\text{m}^2)$ ), as defined by the National Institutes of Health publication nos. 04-5283 and 98-4083 (9), was used to quantify patient size. In our population, BMI ranged from 18 to 50, distributed as follows: 15 bed positions were associated with patients with normal BMI ( $18.5 < \text{BMI} < 24.9$ ), 21 bed positions were associated with overweight patients ( $25 < \text{BMI} < 29.9$ ), 19 bed positions were associated with obese patients ( $30 < \text{BMI} < 39.9$ ), and 4 bed positions were associated with extremely obese patients ( $\text{BMI} > 40$ ). A weighted estimate of total attenuation volume in each bed position was also computed for each patient by summing the linear attenuation coefficients (in  $\text{cm}^{-1}$ ) over every voxel in the reconstructed transmission image; the correlation between these values and BMI was measured.

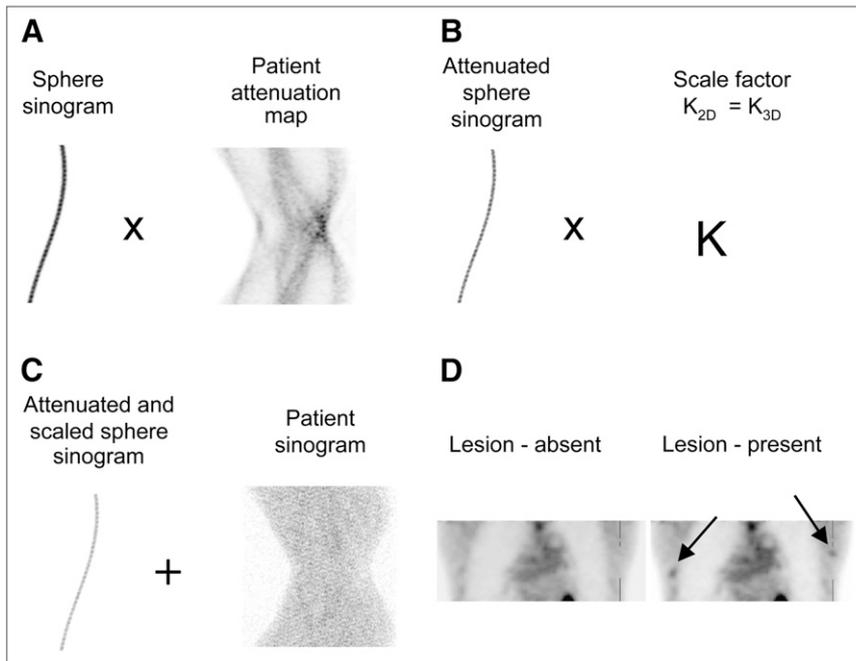
Patients were injected with 740 MBq of  $^{18}\text{F}$ -FDG 45 min before data were acquired in 2D mode for 60–70 min on an ECAT HR+

BGO-based WB-PET scanner (Siemens Medical Solutions, Inc.) using the WB-PET protocol used routinely in the clinic at the Dana-Farber Cancer Institute. An experienced nuclear medicine physician reviewed the reconstructed volumes and identified 1 or 2 disease-free bed positions to be subsequently imaged in both 2D and 3D. The patients were then removed from the scanner and allowed to rest for 1 h. At the end of this time (~3 h after injection), they were encouraged to empty their bladders and were repositioned on the scanner with the first previously identified bed-position located in the scanner's FOV. 2D and 3D emission scans were interleaved (3 min-3D/3 min-2D/3 min-3D/3 min transmission/4 min-2D). A lower-level discriminator of 350 keV was used in 2D and 3D modes. After compensating for randoms and scatter (Material and Methods; Reconstruction Schemes), the time frames were summed into a 6-min 3D scan and a 7-min 2D scan for each bed position. The difference in imaging times compensates for the greater degree of bed overlap required in 3D WB-PET compared with 2D WB-PET.

### Image Generation

*Lesion Generation.* To model in a realistic manner “lesion-present” clinical studies with lesions present—ensuring perfect knowledge of the presence and location of each lesion—spheric lesions were added to disease-free bed positions, yielding synthetic “lesion-present” studies. To avoid errors associated with numeric simulation of lesions to be added (e.g., normalization errors, incomplete modeling of the scanner components, etc.), and to ensure that spheric lesions were added in a realistic manner to the normal images in both 2D and 3D, 3 plastic spheres filled with 74 kBq/mL of  $^{68}\text{Ge}$  were acquired separately in 2D (for 300 s) and 3D (for 258 s) modes at the exact location in the patient volume where a simulated lesion was to be added. To accomplish this, a grid was manufactured with 1,300 threaded holes that allowed positioning of the spheres at any desired location in the PET scanner FOV with 5-mm accuracy in all 3 directions. The spheres of 1.0-, 1.3-, and 1.6-cm internal diameter were individually acquired in air, in 2D and 3D, in 20 locations corresponding to patient anatomic regions of interest (e.g., lungs, soft tissues, bones) within the central third of the FOV (i.e., slices 21–40 of 63 slices) to ensure that the axial sensitivity profiles in 2D and 3D modes were flat (an overlap of 22 slices was used in 3D). We assumed that scatter associated with the sphere acquired in air was negligible as compared with the scatter associated with the patient. For each bed position, the most likely 10 lesion locations within the volume, out of the 20 possible ones, were identified by an experienced nuclear medicine physician. The 3D- (2D) sphere sinograms were attenuated by the 3D (2D) attenuation map generated for each patient, corrected for randoms and scaled to a contrast value that would yield the desired marginal detectability, as assessed subjectively by iterative trial and error of the whole process by a human observer, using the same factor (derived from the 2D images) in 2D and 3D. Because we determined the contrast in the 2D image, and then scaled the sphere data equivalently for 3D, the signal strength was the same in 3D—but the image contrast might well have been different (due to, for example, differences in scatter correction). Poisson noise deviates were then generated from the sphere sinograms and added to the patient sinograms to ensure marginal detectability. Figure 1 illustrates the schema used to generate the lesion-present studies.

To validate the realism of the lesion generation strategy in 2D and 3D modes, a phantom study was undertaken. An elliptic



**FIGURE 1.** Generation of lesion-present studies. (A) Sphere sinograms were attenuated with the corresponding 2D or 3D attenuation map of the patient. (B) Attenuated sinograms of spheres were corrected for randoms and scaled to ensure marginal detectability, using the same scale factor for 2D and 3D modes. (C) These attenuated and scaled sphere sinograms were added to the patient's sinograms. (D) Reconstructed volumes of the patient's sinograms with and without the spheres added were, respectively, the signal-present and signal-absent data used in the observer study.

phantom (long axis = 30 cm, short axis = 22 cm) was filled with a solution of 4.63 MBq/mL of  $^{18}\text{F}$ -FDG, and a 1.3-cm-diameter sphere was filled with a solution of 9.0 MBq/mL of  $^{18}\text{F}$ -FDG. First, the elliptic phantom was positioned in the scanner and emission sinograms were acquired in 2D and 3D modes; a transmission scan was acquired as well. Next, the sphere was positioned inside the phantom at an off-centered position on the long axis of the elliptic cylinder, and a second acquisition (with the sphere inside the phantom) was performed in 2D and 3D under conditions that were identical to those of the first acquisition. Finally, the sphere was mounted on the positioning grid at a location that corresponded to its location inside the cylindrical phantom and acquired alone in air. This allowed us to compare profiles as well as mean pixel noise and sphere-to-background contrast ratios in 2D and 3D in the reconstructed volumes of the spheric lesion present in the cylindrical phantom with the synthetic 2D and 3D volumes obtained by adding the lesion's sinogram to that of the cylindrical phantom using the same approach as that used in the patient studies. The mean pixel noise in 2D and 3D was computed as the ratio of the SD of reconstructed counts in a rectangular region of interest to the mean counts in the same region in 3 locations within 3 slices. The sphere-to-background ratio was computed in 2 steps: first, the maximum within the sphere in the central slice through the sphere was determined for each acquisition ( $\max_{\text{sphere}}$ ) as well as the mean reconstructed counts in a background region of interest in the sphere's vicinity (bkg); next, the sphere-to-background ratio was computed as:  $(\max_{\text{sphere}} - \text{bkg}) / (\max_{\text{sphere}} + \text{bkg})$ .

**Reconstruction Schemes.** The acquisition (2D, 3D) and processing protocols (randoms correction, attenuation correction, reconstruction algorithm) were evaluated in 2 steps. First, 2D and 3D acquisition protocols were assessed for lesion detection using clinical reconstruction protocols to determine relative performance for several patient size ranges. Lesion detection SNR was calculated from each bed position, based on 30 samples per bed position (3 sphere sizes, 10 locations selected from 20

possible locations). Next, several randoms and attenuation compensation and reconstruction approaches were assessed using the same task-based figure of merit with all bed positions from all patient sizes pooled together to reduce the number of conditions explored.

In the first evaluation step, the acquired sinograms were reconstructed using the clinical reconstruction protocol. Prompt and random coincidence data were acquired separately, and random coincidences in each bed position were compensated using the delayed window randoms subtraction (DWS) approach (10). The 2D sinograms were corrected for scatter using a deconvolution with a double-sided exponential, whereas the 3D sinograms were corrected for scatter using a single-scatter, Klein-Nishina modeling technique (11). 3D data were reduced to 2D sinograms using Fourier rebinning (FORE) (12), and then 3D and 2D data were reconstructed iteratively using the same reconstruction algorithm and parameters (attenuation-weighted OSEM, 8 subsets, 2 iterations) recommended by the manufacturer at installation time for WB  $^{18}\text{F}$ -FDG studies performed in clinical routine on the scanner. Regularization was achieved using a 3D gaussian-kernel sieve of 6-mm full width at half maximum. The reconstructed pixel size in the  $128 \times 128 \times 63$  image volume was  $0.51 \times 0.51 \text{ cm}^2$  in the transverse plane; the plane separation was 0.24 cm. The 2D and 3D volumes for each bed position were also reconstructed in the same manner as that described without adding any lesions, to yield the background "lesion-absent" volumes used in the observer study. In summary, 590 lesion-present images and 590 lesion-absent images were generated for each of 3 sphere sizes for both 2D and 3D, by combining 59 bed positions (normal patient images) with images of spheres at each of 10 locations (selected from 20 possible locations based on the clinical likelihood of tumor presence in each location) in patients with BMIs ranging from 18 to 50. The lesion-absent image at a given location was the same for the 3 sphere sizes. These images were used to evaluate the performance of 2D and 3D acquisition modes as a function of patient size. In the second evaluation step, 16 reconstruction/processing schemes were

evaluated using all datasets (1,770 lesion-present datasets (59 bed positions  $\times$  10 locations  $\times$  3 sphere sizes), and 590 lesion-absent datasets (59 bed positions  $\times$  10 locations)):

Two random coincidence correction methods were considered: the standard delayed window subtraction approach (DWS) (10) and the randoms variance reduction approach (RVR) (13,14). In the latter, random sinograms were unmasked by interpolation, and Casey smoothing was performed within each plane (the solution to the randoms variance reduction problem is approximate in this case but yields very similar results to the exact solution (14)). Next, the data were remashed and subtracted from the prompts to yield true coincidence estimates. After random coincidence correction, the time frames were summed into a 6-min 3D scan and a 7-min 2D scan for each bed position.

The scatter correction methods used were the deconvolution using a double-sided exponential in 2D and the single-scatter, Klein–Nishina in 3D (11). Attenuation correction was performed either at the sinogram level or during iterative reconstruction. Therefore, sinograms obtained after summing the data into 6-min 3D or 7-min 2D scans were either reconstructed without scatter or attenuation correction using filtered backprojection (FBP-NATS) or ordered-subsets expectation maximization (OSEM-NATS, 8 subsets, 2 iterations), or corrected for scatter and reconstructed using FBP, after correcting for attenuation (FBP-ATT), or while modeling attenuation in the projector/backprojector of an attenuation-weighted iterative algorithm (AWOSEM). Postreconstruction smoothing was performed using a 3D gaussian-kernel of 6-mm full width at half maximum. In summary, 1,770 lesion-present and 590 lesion-absent datasets were considered for each of the 16 conditions evaluated that corresponded to different acquisition and processing schemes (2D, 3D)  $\times$  (DWS, RVR)  $\times$  (FBP-NATS, OSEM-NATS, FBP-ATT, AWOSEM).

### Channelized Hotelling Observer

The acquisition and processing schemes were assessed on the basis of performance of a model observer in detecting the presence of a spheric lesion of unknown size on an anatomic background. The model observer was a 3-channel Hotelling observer (CHO) (15), by which the  $32 \times 32$  pixel subimage data were processed through the frequency channels that are believed to exist in the human visual system. The  $32 \times 32$  image corresponded to the slice with the greatest signal from a lesion located in pixel (16,16) at the center of the image (only 1 lesion was present in any  $32 \times 32$  image). The CHO-SNR is given by:

$$\text{SNR}_{\text{CHO}}^2 = (\Delta f)^T \cdot S_2^{-1} \cdot (\Delta f), \quad \text{Eq. 1}$$

where  $\Delta f$  is the mean interclass channel output difference vector and  $S_2$  is the intraclass scatter matrix, calculated from the (channelized) covariance matrices  $M_1$  and  $M_2$  of the 2 classes (lesion-present and lesion-absent) being discriminated by  $S_2 = \frac{M_1 + M_2}{2}$ . We used a 3-channel difference-of-gaussians Hotelling observer (15) with (radially symmetric) channel profiles  $C_0$ ,  $C_1$ , and  $C_2$  given by:

$$C_j(\rho) = e^{-\frac{1}{2}\left(\frac{\rho}{2\sigma_j}\right)^2} - e^{-\frac{1}{2}\left(\frac{\rho}{\sigma_j}\right)^2}, \quad \text{Eq. 2}$$

where  $\rho$  is the spatial frequency,  $j = 1, 3$  indices the channels,  $\sigma_j = \sigma_0 2^{j-1}$ , and  $\sigma_0 = 0.052$ . The parameters of this observer model are appropriate for a viewing distance of 60 cm and a displayed pixel

size of 0.51 mm; the effect of the observer noise on task performance was incorporated by doubling the diagonal elements of the channel covariance matrix (15). The lesion detectability performance estimated with the CHO,  $d_A$ , can be related to the corresponding area under the receiving-operator-characteristic curve ( $A_z$ ) by:

$$d_A = 2 \operatorname{erf}^{-1}(2A_z - 1), \quad \text{Eq. 3}$$

where  $\operatorname{erf}^{-1}$  is the inverse error function. Figure 2 shows the difference-of-gaussian channel profiles used in this study. The 590 lesion-present and 590 background (lesion-absent) images were used to compute the CHO detection SNR for the 3 sphere sizes, in each  $32 \times 32$  subimage, yielding 1,770 lesion-present images for each of the 16 conditions considered in this study. Because the CHO-SNR were averaged over sets of subimages that were derived from various body locations, we were concerned that differences in background means among the locations might lead to spurious low-frequency covariance components in the pooled intraclass covariance matrices. To ensure that our results were not affected by this phenomenon, we examined the individual channel output values for evidence of clustering.

### Noise Equivalent Count (NEC) Rate

The performance of 2D and 3D acquisition modes in the task of lesion detection as determined by CHO-SNR was compared with a widely used global metric, the NEC. This metric has the advantage that is easy to compute; it does not, however, account for the scanner's spatial resolution, variable sensitivity in the FOV, or noise correlation in the reconstructed tomographic volume. Three phantoms were considered: The first phantom was the National Electrical Manufacturers Association (NEMA) NU-2 standard phantom that consisted of a 70-cm-long polyethylene cylinder of 20-cm diameter with an off-center line source (16). The second phantom was obtained by fitting an external collar over the  $20 \times 70$  cm cylinder, yielding an effective phantom diameter of 27 cm. The third phantom was obtained by fitting a third external collar yielding an effective phantom diameter of 35 cm (17). The 3

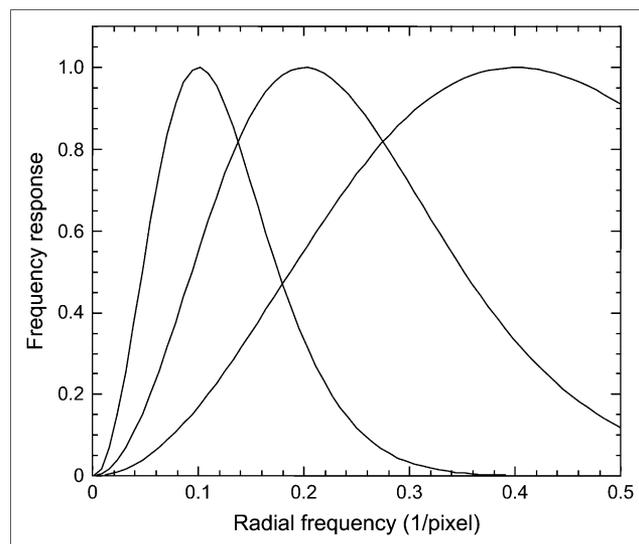


FIGURE 2. Three difference-of-gaussian channels used for the channelized Hotelling observer study.

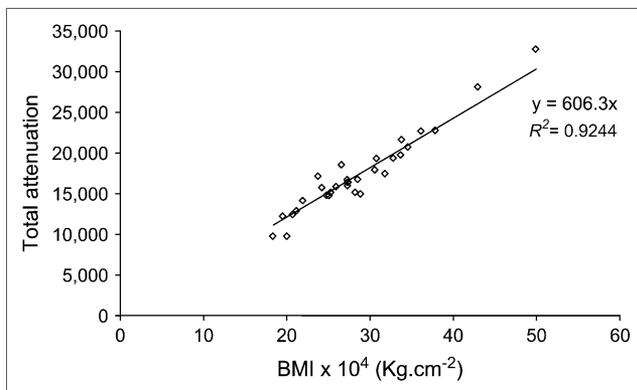
phantoms corresponded to patients weighing 40, 60, and 95 kg, respectively, with assumed BMIs of 18.5, 23.5, and 31, respectively. The line source inside the phantom was filled with 740 MBq of  $^{18}\text{F}$ -FDG, and the phantoms were scanned in 2D and 3D modes 3 h after filling to mimic the clinical conditions under which the detection study was performed in the patient population described in this work. The ratio of the square-root values of NEC in 2D and 3D for the 3 different phantom sizes was plotted along with the lesion detection  $\text{SNR}_{\text{CHO}}$  results obtained in the patient population for reference.

## RESULTS

### Validation of Image Generation

Thirty-eight (64%) bed positions were obtained in 21 normal-weight and overweight subjects ( $\text{BMI} < 30$ ), whereas 21 bed positions (36%) were obtained from 11 obese or extremely obese subjects ( $\text{BMI} \geq 30$ ). Because different bed positions in the same patient (same BMI) can correspond to different levels of attenuation, we correlated the BMI to the total attenuation present in each bed position. Figure 3 shows that BMI and total attenuation correlated very well in the population studied in this work ( $r^2 = 0.924$ ). Therefore, we used BMI as a surrogate of total attenuation.

Adding synthetic lesions to real background images yielded realistic studies. We found very good agreement between the synthetic lesion-present study (sphere added to uniform cylinder) and the true lesion-present study (sphere acquired inside the cylinder), both in 2D and 3D based on the profiles shown in Figure 4 and the mean pixel noise and lesion-to-background contrast values. The mean pixel noise in the synthetic lesion-present study was  $9.72\% \pm 0.41\%$  in 2D ( $4.64\% \pm 0.24\%$  in 3D) as compared with  $8.03\% \pm 0.34\%$  in the case of the 2D true lesion-present study ( $3.19\% \pm 0.21\%$  in 3D) (not statistically significant Student paired  $t$  test). The sphere-to-background contrast was 0.264 in 2D in the case of the synthetic lesion-present study (0.305 in 3D) as compared with 0.276 in the case of the 2D true lesion-present study (0.293 in 3D).



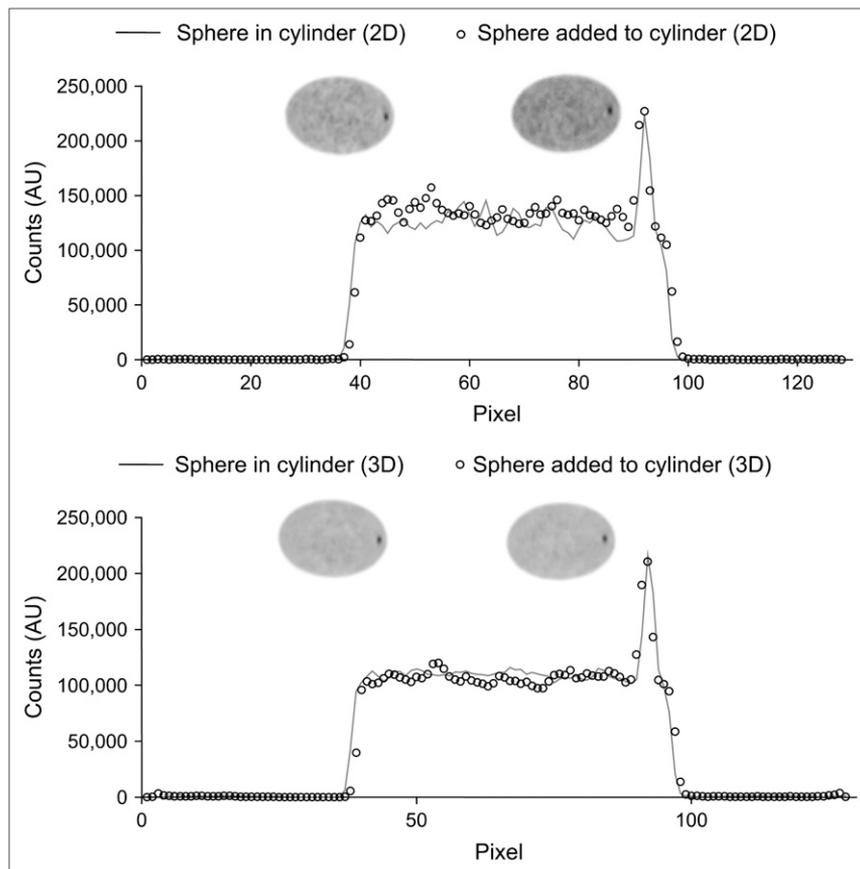
**FIGURE 3.** Correlation of BMI with weighted total attenuation volume in each bed position computed by summing the linear attenuation coefficients (in  $\text{cm}^{-1}$ ) over every voxel (cm) in the patient's reconstructed transmission image.

### Lesion Detection as a Function of Acquisition Mode

Figure 5 shows lesion-present (arrow) coronal slices through studies of patient with normal BMI ( $\text{BMI} = 18$ , 2 lesions), patients who are obese ( $\text{BMI} = 33$ ), and extremely obese ( $\text{BMI} = 50$ ) in 2D and 3D modes reconstructed using the clinical protocol (AWOSEM). In the patient with normal BMI, the lesion is clearly more visible in 3D than in 2D, whereas the lesion is more easily discernable in 2D than in 3D in the extremely obese patient. The visual assessment is confirmed by the objective assessment of lesion detectability. The channelized Hotelling SNR, computed over all lesion sizes and locations (30 conditions), is shown as a function of BMI in 2D and 3D (Fig. 6A). The ratio of  $\text{SNR}_{\text{CHO}}$  (2D) to that in 3D (Fig. 6B) is  $< 1$  for low values of BMI but becomes  $> 1$  for high values of BMI. The best straight-line fit between the ratio of  $\text{SNR}_{\text{CHO}}$  (2D) to  $\text{SNR}_{\text{CHO}}$  (3D) and BMI intercepts the unity line at  $\text{BMI} \approx 31$ . The mean  $\text{SNR}_{\text{CHO}}$ , computed by pooling the 10 lesion locations and 3 sphere sizes and patients with  $\text{BMI} \leq 31$  was significantly higher in 3D mode ( $2.77 \pm 0.89$ ) compared with 2D mode ( $2.18 \pm 0.74$ ),  $P < 0.025$  (Student paired  $t$  test). However, the mean  $\text{SNR}_{\text{CHO}}$  pooled over all lesion sizes and locations and patients with  $\text{BMI} > 31$  was significantly higher in 2D mode ( $2.65 \pm 0.61$ ) than in 3D mode ( $2.25 \pm 0.53$ ),  $P < 0.025$  (Student paired  $t$  test). When all patients were pooled, the mean  $\text{SNR}_{\text{CHO}}$  was not significantly different between 2D and 3D modes. The ratio of square-root NEC measured in 2D and 3D at 3 different BMI values was characterized by a behavior similar to that of  $\text{SNR}_{\text{CHO}}$  (2D)/ $\text{SNR}_{\text{CHO}}$  (3D), with values of 0.70, 0.80, and 0.88 for  $\sqrt{\text{NEC}}$  (2D)/ $\sqrt{\text{NEC}}$  (3D) for the  $20 \times 70$ ,  $27 \times 70$ , and  $35 \times 70$  cm cylinders, respectively. These values were well within the patient variability observed for  $\text{SNR}_{\text{CHO}}$  (2D)/ $\text{SNR}_{\text{CHO}}$  (3D) in our patient population.

Table 1 shows the  $\text{SNR}_{\text{CHO}}$  calculated for the pooled population of 32 patients (59 bed positions) and 10 locations (10 locations per bed position) in 2D and 3D modes for 1.0-, 1.3-, and 1.6-cm lesions. Our results show that no statistically significant difference was seen between 2D and 3D as a function of lesion size. As expected,  $\text{SNR}_{\text{CHO}}$  significantly increased as the lesion size increased from 1-cm diameter to 1.3 cm (17%) and 1.6 cm (37%),  $P < 0.05$ .

Figure 7 is a 3D scatter plot of the individual CHO output values. Channels 1, 2, and 3 are the spatial frequency channels, defined in Equation 2, through which the image data are processed, yielding the 3 channel outputs. It is evident that if values are clustered around different means for various body locations, the differences between such means are small compared with the overall dispersion. This gives us confidence that our approach, which involves averaging over disparate body locations, is valid. If there had been clustering, then we would have observed distinct clouds of points. The fact that we did not implies that our pooled covariance matrix was not affected by any differences in background means among various body locations.



**FIGURE 4.** Transverse slices of 2D and 3D reconstructed images of the elliptical phantom acquired with the sphere inside the cylinder, as well as with the sphere synthetically added to the uniform cylinder without the sphere along with the corresponding profiles through the acquired and added spheres. AU = arbitrary units.

#### Lesion Detection as a Function of Processing Scheme

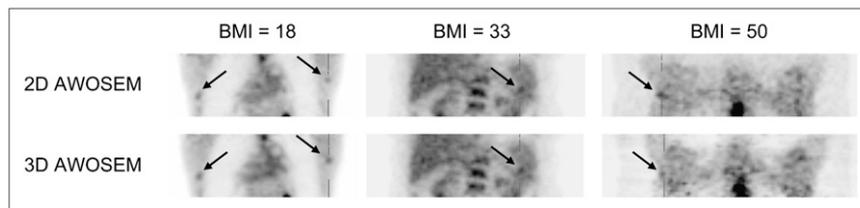
Figure 8 shows coronal slices of a patient with normal BMI (BMI = 18) and an obese patient (BMI = 38) in 2D and 3D reconstructed with FBP with and without attenuation and scatter corrections, OSEM without attenuation and scatter corrections, and AWOSEM. The arrows point to synthetic lesions added in the mediastinal area (2 lesions are shown for BMI = 18, 1 lesion is shown for BMI = 38). The differences between the different processing protocols are clearly visible in images of the obese patient, where the lesion becomes extremely hard to see with OSEM-NATS compared with AWOSEM. Table 2 shows the  $SNR_{CHO}$  computed over all 59 bed positions and 10 sphere locations and 3 sizes (1,770 conditions) for the 8 processing schemes evaluated in 2D and 3D. AWOSEM, FBP-ATT, and FBP-NATS yielded significantly better lesion detection  $SNR_{CHO}$  (40%) than OSEM ( $P < 0.05$ ) in both 2D and 3D. AWOSEM yielded also a significant improvement in

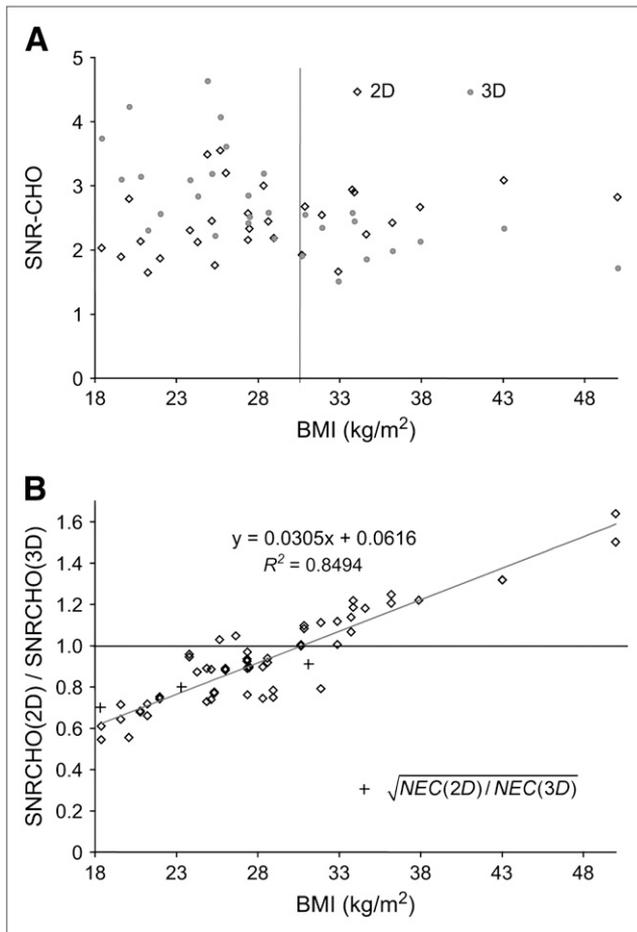
$SNR_{CHO}$  compared with FBP-ATT and FBP-NATS (14% with RVR and 7% with DWS) ( $P < 0.05$ ). In all patients, RVR yielded a systematic improvement in  $SNR_{CHO}$  over DWS in both 2D (6%, not statistically significant [NS]) and 3D (13%,  $P < 0.05$ ) acquisition modes. Furthermore, in patients of typical size (BMI < 33),  $SNR_{CHO}$  was 20% higher with DWS in 3D than with DWS in 2D ( $P < 0.05$ ) and 15% higher with RVR in 3D than with RVR in 2D. On the other hand, 2D significantly outperformed 3D in large patients for lesion detectability, both with DWS (17% higher  $SNR_{CHO}$ ) and RVR (13% higher  $SNR_{CHO}$ ) ( $P < 0.05$ ).

#### DISCUSSION

The aim of this study was to develop an evaluation methodology for the assessment of different acquisition and processing schemes in WB  $^{18}F$ -FDG PET using a task-based metric—that is, lesion detection as measured by a channelized

**FIGURE 5.** Lesion-present (arrow) coronal slices through normal-weight (BMI = 18, 2 lesions shown), obese (BMI = 33, 1 lesion shown), and extremely obese (BMI = 50, 1 lesion shown) patient studies in 2D and 3D modes reconstructed using the clinical protocol (AWOSEM).





**FIGURE 6.** 2D and 3D channelized Hotelling SNR computed over all lesion sizes and locations (30 conditions) as a function of BMI (A) as well as SNR ratio of 2D and 3D (B).  $\sqrt{\text{NEC}}(2\text{D})/\sqrt{\text{NEC}}(3\text{D})$  is plotted for 3 phantoms ( $20 \times 70$ ,  $27 \times 70$ , and  $35 \times 70$  cm) corresponding to patients weighing 40, 60, and 95 kg, respectively.

Hotelling observer. Our approach provides clinically realistic images in which ground truth regarding the presence or absence of a lesion is known. Because the aim of this work was to assess lesion detection and not the determination of lesion morphology or density, only spheric lesions were considered. This is a reasonable model for small metastatic lesions as seen routinely, for example, in lung, liver, or breast

**TABLE 1**

Channelized Hotelling SNR Computed Over All 59 Bed Positions and 10 Locations (590 Conditions) as Function of Lesion Size

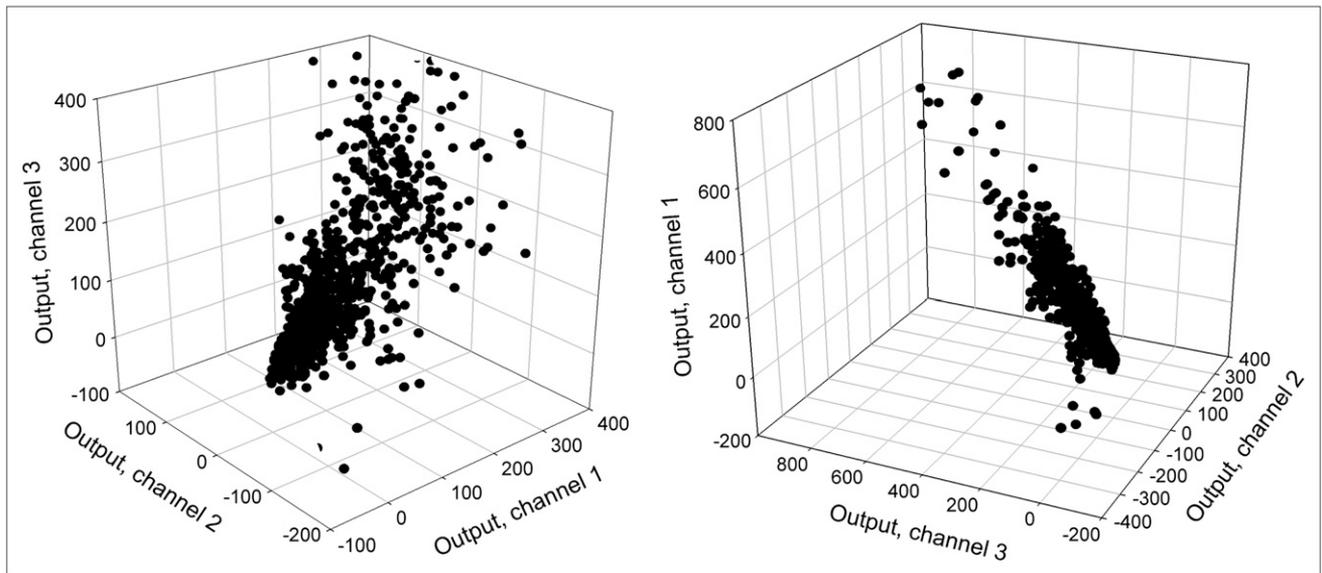
Mode	Lesion size = 10 mm	Lesion size = 13 mm	Lesion size = 16 mm
2D	$1.668 \pm 0.792$ *	$2.001 \pm 0.667$ *	$2.669 \pm 0.625$ *
3D	$1.752 \pm 0.683$ *	$2.127 \pm 0.654$ *	$2.836 \pm 0.598$ *

\*Not statistically significant.

cancers. This approach was used to assess the performance of 2D and 3D acquisition modes, as well as different processing schemes, including random event variance reduction and delayed window subtraction, FBP, and OSEM reconstruction algorithms, with and without scatter and attenuation corrections. In contrast with several published studies that reported comparison of 2D and 3D for a given phantom size (*1,2,5,18*), this study included a population of patients with a large range of BMI to make possible the assessment of the impact of patient size. This is crucial, as our results show that the relative performance of 2D and 3D modes correlated with BMI. Furthermore, the use of clinical data, in contrast with numerically simulated data, ensured that all physical factors—such as scatter, attenuation, randoms, and dead time—were correctly modeled for all body sizes and physiologic activity distributions. This gives us confidence that any conclusions drawn from the data can be extrapolated to the clinical setting.

One of the major difficulties of using clinical data is the absence of a gold standard. Unlike previous studies (*6,7*), this study had the advantage of a gold standard, because the lesion-present cases were generated from artificial lesions that were added to the raw sinograms. Our methodology for adding the lesions was validated in cylindrical phantom experiment, and the realism was verified by ensuring that likely locations of disease were chosen by an experienced physician for each bed position. Performing 2D and 3D imaging at the same time ensured imaging under identical physiologic (redistribution of the tracer, specific uptake) as well as physical (counting rate, scatter and randoms fractions, radioactive decay) conditions. The limitation of this approach, however, is that 2D and 3D imaging was performed at one time point that might not be optimal for both 2D and for 3D. Previous reports on phantom data have shown that optimal 2D and 3D imaging using BGO scanners correspond to different injected doses (*18*), implying that a given imaging time cannot be optimal for both acquisition geometries. In fact, it is most likely that the imaging time was close to optimal for 3D (injection of 740 MBq and 3-h delay) but not for 2D. As a result, we would expect that under optimal 2D imaging conditions, lesion detection performance would improve for 2D, yielding better lesion detection SNR than does 3D at lower BMI values.

In PET, the linear attenuation coefficient (LAC) at every point in the patient is proportional to the density of the tissue at that point. Therefore, one would expect the integral value of the LACs in the transmission images to be perfectly correlated to “partial patient weight”—that is, the weight of tissue in the FOV. In this work, we have considered BMI instead of partial patient weight and reported an  $r^2$  of 0.9244 ( $r = 0.961$ ) between BMI and total attenuation. The reason for using BMI instead of partial patient weight is a practical one: when deciding whether to perform 2D or 3D imaging, one needs to know before starting the scan what protocol to use, and one



**FIGURE 7.** Two views of the 3D scatter plots of the individual CHO output values.

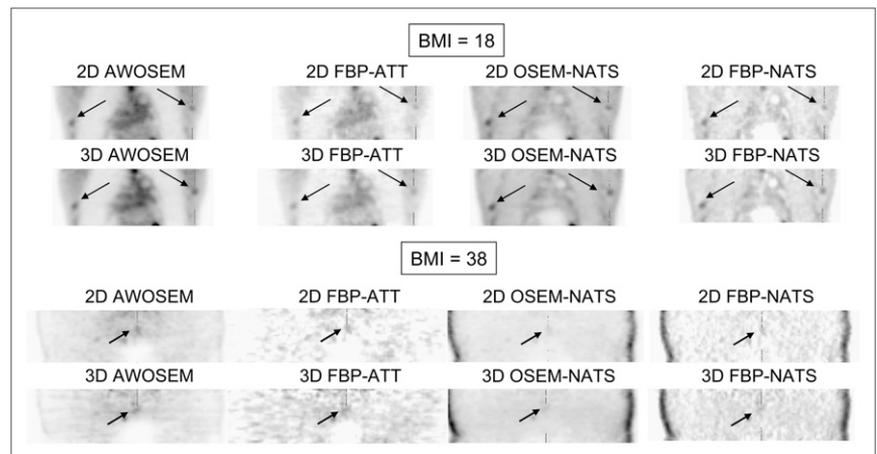
cannot compute the sum of LACs before scanning the patient but can easily compute the patient BMI.

The task considered in this work was lesion detection, measured by  $SNR_{CHO}$  (15,19,20). Although lesion detection performance is best assessed using human observer studies (e.g., receiver operating characteristic [ROC] (3), Localization ROC [LROC] (3), or Alternative Free-response ROC [AFROC] (5)), the numeric observer approach allows more rapid assessment. The numeric observer we used, the CHO, has been shown to be a good surrogate of human observers under a wide range of conditions (15,21). Because the 30 samples were 3 sphere sizes and 10 locations, these numbers multiply as each sphere size was placed at each of the 10 locations. Although every effort was made to ensure that the 3 sphere center planes with the highest intensity corresponded to at least 2 different slices, 3D gaussian regularization of the reconstructed volume will have introduced correlations between the backgrounds used for the 3 sphere sizes that were neglected in this work.

The background images were portions of the patient anatomy that did not overlap and, therefore, each bed position's CHO-SNR was based on a sample of 30 independent noise backgrounds. A good strategy for evaluating system performance would be to use a numeric observer to identify the optimal schemes of acquisition or processing and then validate these findings in a smaller study using human observers. Because we used the entire set of images both to estimate the scatter matrices and to obtain the performance estimates, it is possible that our results are affected by resubstitution bias. However, even if this is true, both modalities should be affected equally, and our comparisons should be valid.

In this work we found that AWOSEM yielded a significant improvement compared with FBP and that AWOSEM, FBP-ATT, and FBP-NAT yielded significantly better  $SNR_{CHO}$  than did OSEM. These results are consistent with previously published results by Lartizien et al. (22), who reported region-dependent performance of the reconstruc-

**FIGURE 8.** Coronal slices through normal-weight patient (BMI = 18) and obese patient (BMI = 38) in 2D and 3D corresponding to the following image processing protocols: FBP with (FBP-ATT) and without (FBP-NATS) attenuation and scatter corrections, OSEM without attenuation and scatter corrections (OSEM-NATS) and attenuation-weighted OSEM (AWOSEM). Arrows point to lesions added in mediastinum. Randoms were compensated for using the RVR technique.



**TABLE 2**  
Channelized Hotelling SNR Computed Over All 59 Bed Positions and 10 Sphere Locations and 3 Sizes  
(1,770 Conditions) for the 8 Processing Schemes Evaluated in 2D and 3D\*

Protocol	2D + RVR	3D + RVR	2D + DWS	3D + DWS
AWOSEM	2.621 ± 0.305	2.759 ± 0.316	2.207 ± 0.312	2.279 ± 0.323
OSEM-NATS	1.451 ± 0.198	1.513 ± 0.258	1.394 ± 0.195	1.482 ± 0.242
FBP-ATT	2.307 ± 0.269	2.399 ± 0.293	2.014 ± 0.266	2.142 ± 0.285
FBP-NATS	2.304 ± 0.260	2.309 ± 0.279	2.004 ± 0.260	2.103 ± 0.283

\*FBP with (FBP-ATT) and without (FBP-NATS) attenuation and scatter corrections, OSEM without attenuation and scatter corrections (OSEM-NATS), and attenuation-weighted OSEM (AWOSEM) along with randoms variance reduction (RVR) or delayed window subtraction (DWS).

tion algorithms with lower performance in lesion detection with OSEM than with AWOSEM and FBP in non-Monte Carlo simulated projections of the MCAT torso phantom that were corrected for scatter and attenuation in the hepatic region, and similar performance of FBP and AWOSEM in that region. The poor performance of OSEM, especially for large patients, is probably caused by inconsistencies between the model and the physics of the data acquisition. We are presently performing detailed comparisons of the different processing schemes as a function of the lesion location. We also reported significantly better values of  $SNR_{CHO}$  in 2D mode for patients with a BMI > 31, but higher  $SNR_{CHO}$  in 3D mode for patients with a BMI ≤ 31, whereas no statistical difference was observed between 2D and 3D when all patients were pooled. The latter finding could explain the results reported recently by Visvikis et al. (6), who found similar image quality in a qualitative assessment of 2D and 3D scans.

We found that random event variance reduction yielded a systematic improvement in  $SNR_{CHO}$  over the delayed window subtraction of random events in 2D (6%, NS) and 3D (13%,  $P < 0.05$ ) acquisition modes. Because the delayed channel in DWS is also subject to Poisson noise, a direct subtraction of the delayed coincidence measurement results in an increase of the effective noise contribution of random coincidences to the data, whereas a variance reduction approach similar to that currently used in normalization of PET cameras does not.

This work pertains to a PET scanner equipped with BGO crystal detectors, and our findings may not translate to scanners equipped with faster cerium-doped LSO, LYSO, GSO, or LaBr<sub>3</sub> detectors. These scanners may have reduced sensitivity to random coincidences and, depending on the detector and electronics design, less dead time and even time-of-flight capability. In addition, such scanners may have improved energy resolution and, thus, reduced scatter fraction. These factors would all act to improve 3D imaging over 2D imaging, perhaps especially for obese patients. Our results may also not be applicable if other reconstruction techniques are used, as it is not clear that FORE followed by OSEM reconstruction represents the optimal treatment of 3D PET data. Our methodology can be easily applied to

studies performed on any scanner and is not limited to comparisons of 2D and 3D mode imaging—it could also be used to compare other acquisition parameter choices or different reconstruction techniques. We are presently pursuing this avenue of research.

## CONCLUSION

We have compared the relative merits of different acquisition (2D vs. 3D) and processing (data correction and tomographic reconstruction) schemes in 59 PET bed positions acquired in 32 patients with BMI ranging from 18 to 50. Our results show that 3D imaging yielded better lesion detectability than 2D in patients of typical size (BMI ≤ 31) on our BGO scanner. However, 2D imaging yielded better lesion detectability than 3D in large patients (BMI > 31) as 3D performance deteriorated in large patients. 2D and 3D yielded similar results for different lesion sizes. Furthermore, attenuation-weighted OSEM yielded significantly better lesion detectability than did filtered backprojection and OSEM without scatter and attenuation corrections.

## ACKNOWLEDGMENTS

This work was supported by a grant from the Society of Nuclear Medicine Education and Research Foundation and in part by grants RO1-EB005876, RO1-EB000802, and RO1-EB001989 from the National Institutes of Health (NIH). The contents of this article are solely the responsibility of the authors and do not represent the official views of the NIH. The authors thank Dr. Suleman Surti for providing the phantom collars and Dr. Craig Abbey for useful discussions.

## REFERENCES

1. Raylman RR, Kison PV, Wahl RL. Capabilities of two- and three-dimensional FDG-PET for detecting small lesions and lymph nodes in the upper torso: a dynamic phantom study. *Eur J Nucl Med.* 1999;26:39–45.
2. Moore SC, El Fakhri G, Badawi R, Van Den Abbeele AD, Zimmerman RE. Relative lesion detectability in 3D vs 2D dedicated multi-ring PET. Paper presented at: IEEE Transactions on Nuclear Science Medical Imaging Conference 2001; November 4–10, 2001; San Diego, California.

3. 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.
4. Lartzien C, Kinahan PE, Swensson R, Comtat C, Trebossen R. Numerical and human observer comparisons of tumor detection for 2D versus 3D whole-body PET imaging protocols. Paper presented at: IEEE Transactions on Nuclear Science Medical Imaging Conference 2001; November 4–10, 2001; San Diego, California.
5. Lartzien C, Kinahan PE, Comtat C. A lesion detection observer study comparing 2-dimensional versus fully 3-dimensional whole-body PET imaging protocols. *J Nucl Med.* 2004;45:714–723.
6. Visvikis D, Griffiths D, Costa DC, Bomanji J, Ell PJ. Clinical evaluation of 2D versus 3D whole-body PET image quality using a dedicated BGO PET scanner. *Eur J Nucl Med Mol Imaging.* 2005;32:1050–1056.
7. Lodge MA, Badawi RD, Gilbert R, Dibos PE, Line BR. Comparison of 2-dimensional and 3-dimensional acquisition for <sup>18</sup>F-FDG PET oncology studies performed on an LSO-based scanner. *J Nucl Med.* 2006;47:23–31.
8. El Fakhri G, Badawi RD, Santos PA, Moore SC, Van den Abbeele AD, Kijewski MF. Impact of acquisition geometry and patient habitus on tumor detectability in whole-body FDG-PET: a channelized Hotelling observer study. *IEEE Nuclear Science Symposium and Medical Imaging Conference.* November 11–15, 2002; Norfolk, Virginia. IEEE. 2002;3:1402–1405.
9. National Institutes of Health. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report. NIH Publication No. 98-4083. 1998.
10. Hoffman EJ, Huang SC, Phelps ME, Kuhl DE. Quantitation in positron emission computed tomography. 4. Effect of accidental coincidences. *J Comput Assist Tomogr.* 1981;5:391–400.
11. 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.
12. Defrise M, Kinahan PE, Townsend DW, Michel C, Sibomana M, Newport D. Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans Nucl Sci.* 1997;16:145–158.
13. Casey ME, Hoffman EJ. Quantitation in positron emission computed tomography. 7. A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration. *J Comput Assist Tomogr.* 1986;10:845–850.
14. Badawi RD, Miller MP, Bailey DL, Marsden PK. Randoms variance reduction in 3D PET. *Phys Med Biol.* 1999;44:941–954.
15. Abbey CK, Barrett HH. Human- and model-observer performance in ramp-spectrum noise; effects of regularization and object variability. *J Opt Soc Am A Opt Image Sci Vis.* 2001;18:473–488.
16. Daube-Witherspoon ME, Karp JS, Casey ME, et al. PET performance measurements using the NEMA NU 2-2001 standard. *J Nucl Med.* 2002;43:1398–1409.
17. Surti S, Badawi RD, Holdsworth CH, El Fakhri G, Kinahan PE, Karp JS. A multiscanner evaluation of PET image quality using phantom studies. IEEE Nuclear Science Symposium and Medical Imaging Conference. October 19–25, 2003; Portland, Oregon. 2003;4:2425–2427.
18. Lartzien C, Comtat C, Kinahan P, Ferreira N, Bendriem B, Trebossen R. Optimization of injected dose based on noise equivalent count rates for 2- and 3-dimensional whole-body PET. *J Nucl Med.* 2002;43:1268–1278.
19. Fiete RD, Barrett HH, Smith WE, Myers KJ. Hotelling trace criterion and its correlation with human-observer performance. *J Opt Soc Am A.* 1987;4:945–953.
20. Yao J, Barrett HH. Predicting human performance by a channelized Hotelling observer model. *Proc SPIE.* 1992;1768:161–168.
21. Lartzien C, Kinahan PE, Comtat P. Volumetric model and human observer comparisons of tumor detection for whole-body positron emission tomography. *Acad Radiol.* 2004;11:637–648.
22. Lartzien C, Kinahan PE, Swensson R, et al. Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. *J Nucl Med.* 2003;44:276–290.