HHS Public Access

Author manuscript

*Phys Med Biol*. Author manuscript; available in PMC 2021 January 24.

Published in final edited form as:

*Phys Med Biol*. ; 65(3): 035002. doi:10.1088/1361-6560/ab6011.

# Numerical observer study of lesion detectability for a long axial field-of-view whole-body PET imager using the PennPET Explorer

**Varsha Viswanath**[1], **Margaret E. Daube Witherspoon**[2], **Joel S. Karp**[2], **Suleman Surti**[2]

[1]Department of BioEngineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA 19104

[2]Radiology Department, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

## Abstract

This work uses lesion detectability to characterize the performance of long axial field of view (AFOV) PET scanners which have increased sensitivity compared to clinical scanners. Studies were performed using the PennPET Explorer, a 70-cm long AFOV scanner built at the University of Pennsylvania, for small lesions distributed in a uniform water-filled cylinder (simulations and measurements), an anthropomorphic torso phantom (measurement), and a human subject (measurement). The lesion localization and detection task was quantified numerically using a generalized scan statistics methodology. Detectability was studied as a function of background activity distribution, scan duration for a single bed position, and axial location of the lesions. For the cylindrical phantom, the areas under the localization receiver operating curve (ALROCs) of lesions placed at various axial locations in the scanner were greater than 0.8 – a value considered to be clinically acceptable (i.e., 80% probability of detecting lesion) – for scan times of 60 s or longer for standard-of-care (SoC) clinical dose levels. 10-mm diameter lesions placed in the anthropomorphic phantom and human subject resulted in ALROCs of 0.8 or greater for scan times longer than 30 s in the lung region and 60 s in the liver region, also for SoC doses. ALROC results from all three activity distributions show similar trends as a function of counts detected per axial location. These results will be used to guide decisions on imaging parameters, such as scan time and patient dose, when imaging patients in a single bed position on long AFOV systems and can also be applied to clinical scanners with consideration of the sensitivity differences.

## 1. Introduction:

The current generation of whole-body, time-of-flight (TOF) PET scanners currently has a 15-26 cm axial field-of-view (AFOV) and a coincidence timing resolution (CTR) between 200-500 ps (Kolthammer et al., 2014; Rausch et al., 2019; van Sluis et al., 2019; Jakoby et al., 2011; Ilan et al., 2017; Bettinardi et al., 2011). Routine clinical [$^{18}$F]-fluorodeoxyglucose (FDG) scans on these systems are as short as 60-180 s per bed position depending on patient body mass index due to the high system sensitivities, which range from 5.5-20 cps/kBq and depend primarily on the scanner AFOV, since they all use lutetium scintillators with a thickness of roughly 18-25 mm.

In recent years, there have been efforts to develop total-body PET scanners with longer AFOVs in order to significantly increase sensitivity and to enable whole-body dynamic imaging (Cherry et al., 2018). Current efforts to build long AFOV scanners include the 70-cm long PennPET Explorer that we have developed at the University of Pennsylvania (Karp et al., 2019) and the 195-cm long uExplorer, developed by United Imaging in collaboration with the University of California, Davis (Badawi et al., 2019).

Oncologic FDG PET scans represent the majority of clinical PET studies, where lesion detection and quantification is a primary clinical task. We expect that the improved sensitivity of long AFOV systems will improve detectability of lesions, which is particularly important for small, low-uptake lesions that can represent inflamed lymph nodes, typically 1 cm in size, or foci of cancer, where lesions can be as small as 3 mm. Lesions smaller than 1 cm suffer large partial volume effects of 50% or greater that prove challenging for both detectability and quantitation tasks, even with the improved TOF and spatial resolution of newer clinical systems. The increased sensitivity of long AFOV systems may provide the boost in performance necessary for small lesion detection and localization. The improved lesion detectability will allow for better detection of low uptake cancer lesions and more accurate monitoring of disease progression and recurrence. Alternatively, the higher sensitivity of long AFOV systems can be leveraged to reduce clinical scan times in an effort to improve clinical workflow, reduce patient motion artifacts, and allow for breath-hold imaging while maintaining acceptable detectability for clinically relevant lesions.

The gold standard for quantifying lesion detectability is via a reader study, where trained physicians are requested to read studies and detect and localize lesions; however, these studies are difficult to organize, require hours of clinicians' time, and are subject to variations in reader training. Therefore, numerical observer methodologies have been developed to efficiently perform this task. The most commonly used numerical observer methodology in PET imaging is the channelized hotelling observer (CHO) (Abbey and Barrett, 2001). This technique as used in nuclear medicine has been shown to correlate with human observer performance, given a proper choice of channels. However, CHO only performs a detectability task, while the generalized scan statistics methodology, the numerical observer that we chose to use in this work, performs a detectability and localization task that more closely approximates the task performed by a physician. The generalized scan statistics methodology scans an image volume to determine the distribution of true (lesions) and false (background noise) signals from which lesion detection and localization can be directly estimated. Thus far, the generalized scan statistics methodology has only been implemented it in uniform activity distributions, whereas this work extends application of the technique in non-uniform activity distributions (e.g. human subjects) and in non-uniform organ regions, such as the lung (Popescu and Lewitt, 2006; Surti and Karp, 2015; Surti and Karp, 2010; Surti et al., 2013a; Surti et al., 2013b).

The generalized scan statistics methodology was previously used to perform lesion detectability studies of long AFOV scanners from Monte Carlo simulations with simple cylindrical phantoms (Surti and Karp, 2015). These studies demonstrated the benefit of longer AFOV and improved TOF resolution, indicating a potential factor of ten reduction in total scan time with a combination of increasing the AFOV by 4x from 18 cm to 72 cm and

improving the TOF resolution from 600 ps to 300 ps (Surti and Karp, 2015). This work presents a more extensive and systematic study of lesion detectability using the long AFOV PennPET Explorer that was developed in the last several years (Karp et al., 2019). Importantly, the studies include more realistic activity distributions including a torso phantom and human subject backgrounds. Both simulations and measurements on the PennPET Explorer are used to perform the numerical detection studies.

## 2. Methods

### 2.1. Scanner Geometry

**2.1.1 Simulation—**Simulations were done using Geant4 Application for Tomographic Emission (GATE) V8.0 (Strulab et al., 2003), where true, random, and scattered events were all modeled. The macros and conversion scripts were based on those modeling the Philips Vereos (Perkins, 2017). Prior validation of the macros used for our simulations for the Vereos was performed by Trindade et al. (Trindade et al., 2015). The simulation macros and list-mode conversion scripts model the scanner geometry, line-of-response (LOR) sensitivity, and light sharing between crystals, which are coupled in a 1:1 arrangement on the digital SiPM device. The Vereos model was extended from 5 rows of tiles (16.4 cm axially) to 21 (70 cm axially) to create the 70-cm scanner simulations, to reflect the configuration of the PennPET Explorer. Based on the Philips Vereos parameters, all simulations were run with a 10.9% energy resolution (450-613 keV window) and at 100-ps timing resolution (4.02-ns coincidence window) that was later degraded to a 250-ps timing resolution to match the performance of the PennPET Explorer. The simulated scanner sensitivity, determined using the NEMA NU-2 methodology (2001), was 90 kcps/MBq.

**2.1.2. Measurement—**All measured data were collected on the PennPET Explorer, a multi-ring design based on digital SiPM detectors coupled to $3.86 \times 3.86 \times 19$ $mm^3$ LYSO crystals. The PennPET Explorer in its prototype form is composed of three 23-cm ring-segments, each with 6.6-cm inactive regions due to current data readout limitations, and a 1-cm physical gap between ring-segments, thereby resulting in 64.4 cm AFOV (labelled as 64-cm AFOV here). Once readout from the inactive regions has been established, the PennPET Explorer will have the same geometry as the simulated scanner. The measured scanner sensitivity in this prototype configuration using the standard 70-cm NEMA line source is 55 kcps/MBq, timing resolution is 250 ps and spatial resolution is 4.0 mm at the center of the scanner. The maximum axial acceptance angle is ±40°, which covers the axial extent of the scanner. Further details on the scanner geometry and performance can be found in Karp, et al. (Karp et al., 2019).

**2.1.3. Axial sensitivity profile.—**Axial sensitivity profiles were generated for a line source in a 20-cm attenuating cylinder to understand how the extension of the AFOV and the inactive axial regions, resulting from the data readout limitations, affect the roughly triangular axial sensitivity profile seen with clinical scanners. A 140-cm line source filled with 37 MBq of activity was simulated in a 20-cm diameter x 140-cm long cold water-filled cylinder on the 70-cm scanner. List-mode events from a NEMA sensitivity line measured in air on the PennPET Explorer were statistically attenuated using the probability of

attenuation along each LOR through a 20×70 cm water-filled cylinder to calculate the axial sensitivity profile for the PennPET Explorer. Each list-mode event was then placed into the axial location from which that event most likely originated based on the TOF difference of the two annihilation photons and scanner geometry (Figure 1a (Snyder and Politte, 1983)). This distribution reflects the relative sensitivity to activity at each axial location while including the effects of attenuation. We hypothesized that lesion detectability and localization would depend on the axial location due to changes in the relative sensitivity with position as well as any degrading effects of axial parallax errors due to the large acceptance angle.

## 2.2 Data Acquisition and Generation.

**2.2.1. Simulations.—**Simulations were run for a uniform cylinder and a cylinder with spherical lesions placed at various axial locations to characterize the relationship between lesion detectability and slice sensitivity for single-bed-position imaging with a long AFOV scanner. This study was also designed to provide insight into the performance of the PennPET Explorer without axial inactive regions. A 20-cm diameter x 70-cm long water-filled cylinder was filled with 81 MBq of FDG (3.7 kBq/cc), positioned in the center of the scanner, and imaged for 3 min. A second cylinder with spheres at four different configurations was similarly simulated, and replicate GATE simulations were run to generate ten statistically independent replicates of each cylinder configuration.

Four different lesion uptakes and sizes were investigated: 6-mm spheres with 4:1 and 6:1-uptakes and 10-mm spheres with 2:1 and 3:1-uptakes. The sphere sizes and uptakes were chosen to be challenging based on visual detectability of reconstructed lesion images. Sphere sizes were also chosen to reflect clinically relevant sizes in oncologic PET imaging. Additionally, prior simulations showed good detectability for a 10-mm sphere with 3:1 uptake on a 72-cm AFOV scanner (Surti and Karp, 2015); therefore, we chose to study a similar sphere along with a lower uptake as a more challenging case. Corresponding uptakes were selected for 6-mm spheres based on visual detectability.

For each configuration eight lesions were placed −30, −21, −9, −5, 0, 3, 11, 15, 23, and 27 cm from the axial center of the FOV, positions that correspond to the maximum variations in axial sensitivity for the PennPET Explorer (see Section 3.1). In each slice, eight lesions were placed at a radius of 7 cm for a total of 80 lesions per image. Figure 1b shows the positioning of the lesions. All datasets were also subsampled to shorter time frames, ranging from 5 s to 180 s prior to reconstruction.

**2.2.2. Measurements.—**Data were separately acquired as background data and sphere-in-air data, which were then combined using the lesion embedding methodology to create list-mode data containing lesions of known uptake. The distributions ranged from a simple cylindrical phantom to measurements of more complex backgrounds.

**2.2.2a Lesion Embedding:** Lesion embedding synthetically adds measured sphere data to complex geometries, allowing for the study of quantitative metrics such as contrast recovery or detectability using known sphere size and uptake information. Sphere-in-air data are acquired at the desired locations within the scanner, and the number of sphere list-mode

counts to embed in the background list-mode data are selected based on the desired uptake and activity in the background image at the sphere location. Sphere list-mode events are statistically attenuated using the probability of attenuation by the phantom/body along the LOR of each sphere event. The attenuated sphere and phantom list-mode events are then merged and jointly reconstructed to create a dataset with spheres embedded with known uptake at the desired locations. A more thorough description of this methodology can be found in Daube-Witherspoon, et al. (Daube-Witherspoon et al., 2014). Details describing sphere sizes, uptakes, and background data specifics vary for each generated dataset, so those details are included below for each background distribution.

**2.2.2b Spheres in air.:** Data were collected with 6- and 10-mm diameter spheres filled with 37-74 MBq of activity at many locations on a pre-defined grid in order to overlap with the volumes of both liver and lung of a human subject placed within the scanner. The previously described lesion embedding methodology was used to add spheres to background data to create lesion present datasets.

**2.2.2c Cylindrical phantom.:** A 20-cm diameter x 120-cm long water-filled cylinder, constructed using an 8-inch diameter polyvinyl chloride pipe, was filled with 151 MBq of $^{18}$F, a concentration of 4.0 kBq/cc, and imaged for 21 min. The 21-min dataset was separated into seven independent 180-s datasets that were reconstructed into background images.

Sphere data were embedded in the uniform cylinder data set in the same four size/uptake configurations as in the GATE simulations (6-mm spheres with 4:1 and 6:1-uptake and 10-mm spheres with 2:1 and 3:1-uptake). For each configuration, the axial and transverse locations were also similar to those in the GATE simulations (eight lesions each at positions −30, −21, −9, −5, 0, 3, 11, 15, 23, and 27 cm from the axial center of the FOV and at an average radial distance of 6 cm (range: 5.0 - 6.5 cm). The differences in sphere location axially between simulation and measurement (Figure 1c) show how the axial sensitivity varies across the various locations in which lesions were embedded. Due to small measurement and positioning differences, embedded measured lesions were not in the exact same location as they were in the simulation. Similar to the simulation, data were reconstructed for durations of 5, 10, 15, 30, 60, 90, 120, and 180 s to characterize how measured lesion localization and detectability changes with scan time.

**2.2.2d CTN torso phantom.:** We extended our measurements to anthropomorphic geometries with non-uniform regions to study the impact on lesion detectability in a more realistic but still well-defined distribution. Lesions were embedded in the Society of Nuclear Medicine and Molecular Imaging (SNMMI) Clinical Trials Network (CTN) torso phantom, a 30-cm-long torso phantom with both a uniform background region and two lung regions filled with Styrofoam to emulate the attenuation and activity of human lung (Ulrich et al., 2018). To extend the axial extent of the activity distribution, a 20-cm diameter x 30-cm long cylindrical phantom, within which we could easily embed spheres, was placed directly inferior to the CTN phantom to act as a uniform liver region. Both phantoms were centered in the transverse FOV; the CTN phantom was centered in the axial center of the FOV with the uniform cylinder placed directly inferior; this positioning resulted in only 16 cm of the

30-cm long cylinder being inside the AFOV. Because the CTN phantom has a 2-inch lip, there is a gap between the two phantoms in reconstructed images. The described phantom positioning was meant to roughly emulate the positioning of human subjects in the PennPET Explorer, given differences in body structure, with the vertex aligned with the top edge of the scanner, lungs within the center of the AFOV, and the liver above the bottom edge of the scanner.

The CTN phantom was scanned 60 min after filling according to standard instructions (Ulrich et al., 2018). 85 MBq of $^{18}$F were mixed into the 9.5-L water-filled background region of the CTN phantom for a concentration of 6.3 kBq/cc when imaging began, and the spheres included in the CTN phantom were filled with a 4.1:1-uptake. While liver FDG uptake in humans is typically twice that of background, for this study the uniform cylinder was filled with the same activity concentration as the CTN phantom background. The phantom setup was imaged for 60 min on the PennPET Explorer, and the full dataset was split into ten 360-s background datasets after randomly reordering the list-mode events to eliminate effects of radioactive decay and changing randoms fraction. The physical spheres in the CTN phantom (7-37 mm diameter) were not used in this study.

Spheres were embedded into the lung and the liver regions in two configurations: 6-mm lesions were embedded into the lung at an 8:1-uptake and into the liver at a 4:1-uptake, and 10-mm lesions were embedded into the lung at a 6:1-uptake and into the liver at a 3:1-uptake. Uptake ratios were chosen to be challenging based on visual detectability for both liver and lung lesions. Twenty-two lesions were embedded into the lungs (12 right lung, 10 left lung) and 17 lesions were embedded into the liver region. There are slight positioning discrepancies between the 10-mm and 6-mm spheres that are related to the sphere data acquisitions. Each of the 10 background datasets and 20 lesion datasets – 10 replicates for each sphere size including both the liver and lung – were subsampled and reconstructed at nine scan times ranging from 5 s to 6 min.

**2.2.2e Human subject.:** Ten healthy human volunteers have been imaged thus far on the PennPET Explorer (Pantel et al., 2019). The protocol was approved by the University of Pennsylvania Institutional Review Board, and all subjects signed informed consent. For this work, we used one representative scan with a scan duration long enough to generate multiple independent realizations. We embedded lesions in the liver and lung of one human subject dataset. This builds on work done using the CTN phantom but also introduces some new challenges, such as locally non-uniform activity and attenuation distributions.

The subject was a normal 79-year-old male who was imaged 102 min following injection of 551 MBq of FDG. The subject was positioned such that his vertex was aligned with the top edge of the AFOV, his lungs were in the center of the AFOV, and his liver was towards the bottom of the AFOV. Twenty min of data were collected, and the list-mode file was parsed into six 180-s background datasets, following random reordering of the list-mode events.

Both 6- and 10-mm lesions were embedded into the lung and liver using the same uptakes as described for the CTN phantom. Twenty-seven lesions were embedded into the lungs (18 right lung, 9 left lung) and 16 lesions were embedded into the liver. Each of the 6

background datasets and 12 lesion datasets – 6 replicates for each sphere size including both the liver and lung – were subsampled and reconstructed at eight scan durations ranging from 5 s to 3 min.

### 2.3 Image Reconstruction and Analysis

**2.3.1 Reconstruction.—**All datasets were reconstructed using a list-mode time-of-flight ordered subset expectation maximization (TOF-OSEM) (5 iterations x 25 subsets) algorithm (Popescu et al., 2004) using blob basis functions with a 5.7-mm full-width at half-maximum (FWHM) and 6-mm spacing on a body-centered cubic grid (Matej and Lewitt, 1996). The reconstructed image was re-sampled into 2×2×2 $mm^3$ image voxels. Scatter was estimated using three iterations of time-of-flight single scatter simulation (TOF-SSS) (Werner et al., 2006), and randoms were estimated from a smoothed delayed coincidences sinogram. Because the PennPET Explorer currently does not have an on-board CT, attenuation correction sinograms were generated from CT images acquired on our clinical PET/CT system and aligned to non-attenuation corrected (NAC) PET images from the PennPET Explorer using rigid body registration (MIM Software, Inc.). To aid in this registration, immobilizers commonly used in radiation oncology were used for the human subject scan to maintain body positioning between the PennPET Explorer PET scan and the clinical CT.

**2.3.2 Numerical lesion detection.—**A generalized scan statistics methodology was used to quantify lesion localization and detectability (Popescu and Lewitt, 2006). The scan statistics methodology uses background images to measure local contrast over the background region and generate the corresponding lesion-absent probability distribution function (PDF), which reflects the probability of false positive detection. Contrast is calculated as H/B, where H is the mean counts within a spherical VOI with a 6- or 10-mm diameter, corresponding to the size of the lesion of interest, and B is the mean counts in a spherical shell with an inner diameter twice the embedded sphere diameter and a thickness of 1 cm. Background regions for the CTN phantom and human subject were defined using region masks drawn in MIM (MIM Software, Inc.). For lesion present images, local contrast is measured for each sphere to generate the lesion PDF. The area under the localization receiver operating curve (ALROC) is then calculated statistically from first principles making use of the lesion-present and lesion-absent contrast PDFs (Swensson, 1996). All results are shown for the image reconstruction iteration number that achieves maximum ALROC. The calculated ALROC value reflects the probability of correctly localizing and detecting the lesion.

## 3. Results

### 3.1 Axial Sensitivity Profiles

Axial sensitivity profiles for a line source in a 20×70-cm cylinder are shown in Figure 1c for the simulated 70-cm scanner and the PennPET Explorer. The profile for the simulated scanner is roughly triangular with a slight decrease in expected axial sensitivity in the center of the AFOV due to attenuation effects. The inactive regions in the PennPET Explorer result in a roughly 2x decrease in total sensitivity with peaks and valleys seen in its axial

sensitivity profile. The peaks and valleys reflect axial positions with the greatest and fewest LORs passing through that position, respectively, as illustrated in Figure 1d.

### 3.2 Cylindrical Phantom

Contrast measures for the 180-s cylindrical phantom dataset from both the simulation and PennPET Explorer measurement are shown as a function of axial position in Figure 2. Average contrast values at 180 s over all axial positions are shown for all sphere configurations in Table I. The mean contrast values are consistent across simulation and measurement and do not vary over the AFOV; the standard deviation is systematically larger on the PennPET Explorer due to the lower sensitivity. Contrast results for the spheres are also consistent with values calculated using the NEMA image quality phantom (2001; Ulrich et al., 2018).

Simulated (left) and measured (right) ALROCs for the 6-mm, 6:1-uptake spheres are shown in Figure 3 as a function of axial location for scan times ranging from 5 s to 180 s. These results demonstrate the trends seen with all four sphere configurations studied. For the simulated scanner, ALROC saturates at 1.0 in the center of the scanner for scan durations of 60 s or longer; for measured data, an ALROC > 0.75 is maintained at most axial locations for scan times longer than 90 s.

Because the results in Figure 3 suggest that ALROC tracks with axial sensitivity for shorter scan times, we plotted ALROC as a function of counts originating from different axial locations for all sphere configurations in Figure 4. Counts originating from different axial locations were estimated using most likely positioning of events using TOF information along the LOR of each event, as was done for the axial sensitivity profile determination (Figure 1a). Results show a strong relationship between the measured counts, which varies axially for long AFOV scanners, and ALROC for a given sphere size/uptake combination, as well as good agreement between measurement and simulation. Since there is no change in measured contrast axially (Figure 2), the ALROC only varies as a function of measured counts, which does vary axially. The lowest number of counts at which the ALROC is 1.0 is 0.85 Mcts and 0.30 Mcts for the 4:1- and 6:1-uptake 6-mm spheres, respectively, and 0.80 Mcts and 0.25 Mcts for the 2:1- and 3:1-uptake 10-mm spheres, respectively.

### 3.3 CTN Phantom

Results from the CTN phantom lung region show the first known implementation of the generalized scan statistics methodology in a non-uniform background. Representative coronal slices of data embedded with both 6-mm and 10-mm spheres at all nine scan times as well as corresponding ALROC curves as a function of scan time are shown in Figure 5. In the plots shown in Figure 5b, ALROC plateaus at 0.90 for 10-mm lung lesions with scan durations greater than 30 s, and above 0.95 for 10-mm liver lesions with scan durations longer than 90 s. Results for the 6-mm liver lesion show the ALROC steadily increasing with scan time up to an ALROC of 0.94 at 360 s. ALROC curves for the 6-mm lung lesion plateau at 0.4 (0.38-0.44) for scan durations of 90 s and longer and are unlikely to reach an ALROC of 1.0 with longer imaging times. Corresponding visual lesion detectability can be seen in the representative coronal image slices (Figure 5a).

### 3.4 Human Subject.

Representative coronal slices of the human subject datasets with 6-mm and 10-mm embedded lesions and the corresponding quantitative ALROC results are shown in Figure 6. The 10-mm sphere results show ALROCs of 0.8 or greater at scan times longer than 30 s and 60 s for the lung and liver spheres, respectively. The 6-mm lung and liver spheres have ALROCs that increase steadily up to values of 0.66 and 0.39 at 180 s, suggesting that the ALROC will continue to increase at longer scan times.

Figure 7 compares ALROC results as a function of measured axial counts from all three distributions studied – the cylinder (faded squares), the CTN phantom (yellow diamonds), and the human subject (green diamonds) – and shows that they follow similar trends. The 6-mm sphere with 4:1-uptake (hard to detect) and the 10-mm sphere with 3:1-uptake (easy to detect) in the liver were specifically selected for Figure 7 to allow direct comparison of both the CTN phantom and human subject results to results from the cylindrical phantom. Trend lines were drawn through the cylinder data (solid) along with the CTN and human data (dashed). Measured axial counts for the CTN phantom and human subject were calculated using a weighted average over slices in which spheres were embedded, based on the number of spheres in each slice.

## 4. Discussion

For scan times of 30 s or less, the curves of ALROC as a function of axial position for the simulated cylindrical phantom (Figure 3a) roughly follow the shape of the axial sensitivity profile (Figure 1c), which peaks in the center of the scanner and decreases towards the edges of the AFOV, confirming that ALROC is closely related to the number of counts collected per axial location. This correspondence is especially pronounced for the 30-s scan duration. The relationship of ALROC with axial sensitivity is also apparent for the measured cylinder results (Figure 3b) where dips in the ALROC correspond directly to dips in the axial sensitivity profile of the PennPET Explorer. These results demonstrate the compromise of imaging with a single bed position, which can provide coverage of the organs of interest, but not with uniform sensitivity. This variation is highlighted with the 70-cm PennPET Explorer. The axial sensitivity, and therefore ALROC will become more uniform with axial location as the scanner is extended to a longer AFOV with additional ring segments.

The cylindrical phantom results are in agreement with past Monte Carlo simulation work done by Surti and Karp on a long AFOV scanner (Surti and Karp, 2015). They simulated a 72-cm AFOV scanner with 4×4×20 $mm^3$ LYSO crystals and a 300-ps timing resolution with a 35-cm cylindrical phantom, meant to represent a heavy patient, with 10-mm 3:1-uptake lesions in the center of the AFOV. After adjusting for both scan time and background distribution, our ALROC results are consistent with those results: for adjusted scan times at 60 and 30 s, Surti and Karp report ALROCs of 0.92 and 0.73 while our ALROCs are 0.98 and 0.74 at scan durations of 30 and 15 s. The factor of 2x difference in scan times is likely due to the lower attenuation in the current work due to the use of a 20-cm diameter cylinder as well as differences in simulated scanner sensitivity (higher in previous work). This result is also consistent with the 2x higher total counts measured here, ensuring that our results are consistent with past work.

The PennPET Explorer measurements were taken in its prototype 3-ring configuration with two axially inactive regions between ring segments that make up 23% of the axial detector extent. Because long AFOV scanners will prove costly to build, axial gaps can be part of the scanner design to maximize the AFOV for a given number of detectors (Yamaya et al., 2009), but it is important to consider the implications of introducing such gaps for quantitative metrics such as lesion detectability. First, the total sensitivity is reduced by about a factor of 2, and the axial sensitivity profiles, measured using a line source attenuated in a 20-cm cylinder (Figure 1c), show that there are significant variations in sensitivity corresponding to the different numbers of cross-coincidences accepted at these axial positions (Figure 1d). The normalization correction leads to uniform images and the lesion contrast values are not affected by the inactive regions or the low points of the sensitivity measurement, although the measurement uncertainty (error bars) shows a slight increase at the axial locations related to the decrease in collected counts. Since ALROC is dependent on measured axial counts (Figure 4), decreased axial sensitivity does translate to poorer lesion localization and detectability (Figure 3).

The plots of ALROC vs. axial counts in Figure 4 show similarities between the two sphere sizes, due to the similarities in measured contrast of these spheres. For the higher uptake spheres – 10-mm with 3:1-uptake and 6-mm with 6:1-uptake spheres – the curves plateau at a similar point (0.25 vs. 0.30 Mcts, see section 3.2) where the average contrast measured for these spheres is similar, (1.88 ± 0.09 vs. 2.02 ± 0.12, see Table Ia). This is also true for the two lower uptake spheres – the 10-mm with 2:1-uptake and 6-mm with 4:1-uptake spheres – with contrasts of 1.43 ± 0.06 and 1.60 ± 0.10 where ALROC equals 1.0 at 0.80 and 0.85 Mcts, respectively.

As can be seen in the images in Figures 5 and 6, images with ALROCs ≥ 0.80 have good visual detectability of the lesions. Note that a single replicate is displayed and is only representative of the many replicates used to calculate the ALROC. For the CTN phantom and human subject, 10-mm liver spheres have an ALROC of 0.8 or better for scan times greater than 60 s, whereas for the 10-mm lung spheres, ALROC is greater than 0.8 for scan times longer than 30 s for both the CTN phantom and human subject (Table II). In comparison, for the smaller 6-mm sphere the ALROC exceeds 0.8 only in the liver of the CTN phantom at 360 s. While not being equivalent to a trained clinician reading the images, a visual observation of images indicates that, qualitatively, images with ALROC values of 0.80 or higher would generally lead to a very high confidence in the lesion detection and localization task. Using this value (ALROC=0.8) as a cut-off, we see in Figure 4 that the measured axial counts for a 6-mm diameter sphere with 6:1 uptake in a uniform background can be lowered from the measured 0.69 Mcts at the axial center of the PennPET Explorer for a 180-s scan to 0.25 Mcts. This means, that either the scan time or patient dose can be lowered by a factor of 2.8 on the 70-cm PennPET Explorer in its current configuration while maintaining “good” detectability for lesions of this size in the central plane.

To extend the current results to longer scanners, Figure 8 shows axial sensitivity profiles of the PennPET Explorer scanner geometry extended to 4 ring segments (AFOV = 96 cm) and 6 ring segments (AFOV = 144 cm). The maximum sensitivity in the center of the AFOV increases by a factor of 2.2, 2.5, and 2.7 for the fully active 70-cm, 96-cm, and 144-cm

AFOVs, respectively, relative to the center of the AFOV of the PennPET Explorer in its current configuration (Figure 1c). When imaging with these longer AFOVs, “good” detectability can, therefore, be maintained while imaging with an additional reduction in scan time or dose by a factor of 2.2, 2.5, or 2.7 for the three longer scanners, respectively. Note that the advantage of the longer AFOVs is not a drastic increase in sensitivity, but a larger axial range over which this improved sensitivity is maintained. Due to attenuation of oblique lines-of-response (LORs) in the body, the point sensitivity in the center of AFOV does not increase significantly for scanners longer than 70 cm. Thus, our detectability measures near the center of the AFOV will be representative of detectability achieved on scanners with even longer AFOV such as the uExplorer from United Imaging. Additionally, using their axial sensitivity profiles, similar analysis could be done for other long AFOV scanners with varying axial lengths, potentially with axial gaps of varying sizes. This will be examined further as the AFOV of the PennPET Explorer is expanded from 3 to 6 ring segments.

In Figure 7, all datasets – the cylindrical phantom, the CTN phantom, and the human subject – follow similar trends; however, the CTN phantom ALROC values at a given count level are systematically higher than those for the cylindrical phantom, while the human subject results are systematically lower than the cylindrical phantom data, although the differences are less pronounced for the 10-mm sphere. The differences for the 6-mm sphere curves could be due to differences in measured contrast. While contrast measures for the 10-mm 3:1-uptake spheres are relatively similar for the cylinder, CTN phantom, and human subject (see Table I), the contrasts for the 6-mm 4:1-uptake spheres do vary. The average contrast for the 6-mm 4:1-uptake spheres in the CTN phantom is slightly higher than the cylinder, both simulated and measured, and the contrast measures for the human subject is slightly lower than for the cylinder. Additionally, the spread of these contrast measures is larger for both the CTN phantom and human subject when compared to the cylindrical phantom. While further study will be necessary to more systematically quantify these differences, the results offer insight into the performance of long AFOV systems for varying activity distributions.

A few limitations of this work are noted. One limitation is the design of the lung in the CTN phantom, which makes the detectability results of small spheres in the lung of this phantom difficult to interpret. The ALROC for the 6-mm 8:1-uptake lung lesions in the CTN phantom reaches a plateau at an ALROC of 0.4 (0.38-0.44) for 90 – 360 s scan times, implying that increasing the number of counts collected would not improve the ALROC. We believe these results are confounded by the design of the lungs in the CTN phantom, constructed using Styrofoam peanuts, which does not represent the structure of human lung parenchyma on FDG PET scans even though it does approximate the typical attenuation and uptake in the lung. This is further supported by ALROC results from 6-mm lung lesions in the human subject that continue increasing from 90 to 180 s with a trajectory that should reach an ALROC of 1.0 at longer scan times. The human results demonstrate that the scan statistics methodology can work in a non-uniform region such as the lung, whereas the non-uniformity seen in the lung in the CTN phantom is clinically unrealistic for detectability studies. Also note that our work confines the background region to a specific organ in the CTN phantom and human subject studies. This correspondingly constrains the localization

and detection task to a specific organ, while the traditional clinical task requires the reader to scan the full body.

For the studies in this work, the embedded lesions did not include respiratory motion and therefore the results for short scans for lung lesions may be relevant for breath-hold studies, which need to be 30 s or shorter. ALROC results for the 10-mm 6:1-uptake lung lesions in both the CTN phantom and the human subject can be reduced to 30 s while maintaining detectability with ALROC > 0.8 (ALROC = 0.86 ± 0.03 and 0.82 ± 0.03, respectively). However, for smaller lesions such as the 6-mm 8:1-uptake lesions, ALROC at 30 s is quite poor: 0.28 ± 0.04 and 0.27 ± 0.03 for the CTN and human subject, respectively. Overall, this implies that the scan duration could be reduced to 30 s with minimal detectability losses for lesions 10 mm and larger, but future studies would be needed to consider the feasibility and benefits of such studies. To maintain good image quality, current clinical scans are only as short as 60 s per bed position for low BMI patients, a time span potentially unreasonable for breath-hold imaging on patients with lung disease for whom even a 30-s scan may be too long.

## 5. Conclusion

In this work, we quantified the lesion detectability and localization task using the generalized scan statistics methodology in a cylindrical phantom, the anthropomorphic CTN phantom, and a human subject on the PennPET Explorer, a long AFOV system. ALROC results showed that scan times on our long AFOV system can be reduced to 60-90 s for a single bed position static image with “good” detectability, ALROC > 0.8. ALROC results on both the simulated 70-cm scanner and the PennPET Explorer correspond with variations in axial sensitivity due to axial location. The relationship between ALROC and axial counts for spheres of a given size can be used to design imaging protocols for the PennPET Explorer for single-bed acquisitions that lower dose or scan time, depending on the application. Compared to the PennPET Explorer in its prototype 3-ring configuration, the extended 6-ring (1.4-m) system will lead to a high sensitivity over a large axial range, extending beyond the liver and lungs, that is greater or equal to the maximum sensitivity for the shorter AFOV system studied in this work. Thus, the results of ALROC of this work near the central axial locations can be used to guide the design of imaging protocols for the larger AFOV PET scanner. Moreover, once differences in sensitivity are accounted for these results can also be applied to clinical scanners.

## Acknowledgments

This work was supported in part by NIH Grants R33-CA-225310, R01-CA113941, R01-CA206187, R01-CA196528. The clinical studies were made possible by contributions from Erin Schubert, Michael Parma, and Janet Reddin, Ph.D.

## References

2001 NEMA Standards Publication NU-2-2001: performance measurements of positron emission tomography (Rosslyn, VA: National Electrical Manufacturers Association)

Abbey CK and Barrett HH 2001 Human-and model-observer performance in ramp-spectrum noise: effects of regularization and object variability JOSA A 18 473–88 [PubMed: 11265678]

Badawi RD, Shi H, Hu P, Chen S, Xu T, Price PM, Ding Y, Spencer BA, Nardo L and Liu W 2019 First human imaging studies with the EXPLORER total-body PET scanner J Nucl Med 60 299–303 [PubMed: 30733314]

Bettinardi V, Presotto L, Rapisarda E, Picchio M, Gianolli L and Gilardi M 2011 Physical performance of the new hybrid PET/CT Discovery-690 Med Phys 38 5394–411 [PubMed: 21992359]

Cherry SR, Jones T, Karp JS, Qi J, Moses WW and Badawi RD 2018 Total-body PET: Maximizing sensitivity to create new opportunities for clinical research and patient care J Nucl Med 59 3–12 [PubMed: 28935835]

Daube-Witherspoon ME, Surti S, Perkins AE and Karp JS 2014 Determination of accuracy and precision of lesion uptake measurements in human subjects with time-of-flight PET J Nucl Med 55 602–7 [PubMed: 24604909]

Ilan E, Deller T, Kjellberg F, Peterson W and Lubberink M 2017 Performance comparison of three commercially available PET systems: SIGNA PET/MR, Discovery IQ and Discovery MI J Nucl Med 58

Jakoby B, Bercier Y, Conti M, Casey M, Bendriem B and Townsend D 2011 Physical and clinical performance of the mCT time-of-flight PET/CT scanner Phys Med Biol 56 2375 [PubMed: 21427485]

Karp JS, Viswanath V, Geagan M, Muehllehner G, Pantel A, Parma M, Perkins A, Schmall J, Werner M and Daube-Witherspoon ME 2019 PennPET Explorer: Design and preliminary performance of a whole-body imager J Nucl Med jnumed. 119229997

Kolthammer JA, Su K-H, Grover A, Narayanan M, Jordan DW and Muzic RF 2014 Performance evaluation of the Ingenuity TF PET/CT scanner with a focus on high count-rate conditions Phys Med Biol 59 3843 [PubMed: 24955921]

Matej S and Lewitt RM 1996 Practical considerations for 3-D image reconstruction using spherically symmetric volume elements IEEE Trans Med Img 15 68–78

Pantel AR, Viswanath V, Daube-Witherspoon ME, Dubroff JG, Muehllehner G, Parma MJ, Pryma DA, Schubert EK, Mankoff DA and Karp JS 2019 PennPET Explorer: Human imaging on a whole-body imager J Nucl Med jnumed. 119231845

Perkins A 2017 Personal communication.

Popescu LM and Lewitt RM 2006 Small nodule detectability evaluation using a generalized scan-statistic model Phys Med Biol 51 6225–44 [PubMed: 17110782]

Popescu LM, Matej S and Lewitt RM Nuc Sci Symp, 2004), vol. Series 6): IEEE) pp 3536–40

Rausch I, Ruiz A, Valverde-Pascual I, Cal-González J, Beyer T and Carrio I 2019 Performance evaluation of the Vereos PET/CT system according to the NEMA NU2–2012 standard J Nucl Med 60 561–7 [PubMed: 30361382]

Snyder DL and Politte DG 1983 Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements IEEE Trans Nuc Sci 30 1843–9

Strulab D, Santin G, Lazaro D, Breton V and Morel C 2003 GATE (Geant4 application for tomographic emission): a PET/SPECT general-purpose simulation platform Nucl Phys B 125 75–9

Surti S and Karp J 2015 Impact of detector design on imaging performance of a long axial field-of-view, whole-body PET scanner Phys Med Biol 60 5343 [PubMed: 26108352]

Surti S and Karp JS 2010 Application of a generalized scan statistic model to evaluate TOF PET images IEEE Trans Nuc Sci 58 99–104

Surti S, Shore AR and Karp JS 2013a Design study of a whole-body PET scanner with improved spatial and timing resolution IEEE Trans Nuc Sci 60 3220–6

Surti S, Werner M and Karp J 2013b Study of PET scanner designs using clinical metrics to optimize the scanner axial FOV and crystal thickness Physics in Medicine & Biology 58 3995 [PubMed: 23685783]

Swensson RG 1996 Unified measurement of observer performance in detecting and localizing target objects on images Med Phys 23 1709–25 [PubMed: 8946368]

Trindade PRA, Andreyev A, Perkins AE, Miller M, Griesmer J, Ye J and Laurence T IEEE NSS/MIC, (San Diego, California, 2015), vol. Series): IEEE)

Ulrich EJ, Sunderland JJ, Smith BJ, Mohiuddin I, Parkhurst J, Plichta KA, Buatti JM and Beichel RR 2018 Automated model-based quantitative analysis of phantoms with spherical inserts in FDG PET scans Med Phys 45 258–76 [PubMed: 29091269]

van Sluis JJ, de Jong J, Schaar J, Noordzij W, van Snick P, Dierckx R, Borra R, Willemsen A and Boellaard R 2019 Performance characteristics of the digital Biograph Vision PET/CT system J Nucl Med jnumed. 118215418

Werner ME, Surti S and Karp JS Nuc Sci Symp, 2006), vol. Series 3>): IEEE) pp 1768–73

Yamaya T, Yoshida E, Inadama N, Nishikido F, Shibuya K, Higuchi M and Murayama H 2009 A multiplex “OpenPET” geometry to extend axial FOV without increasing the number of detectors IEEE Trans Nuc Sci 56 2644–50

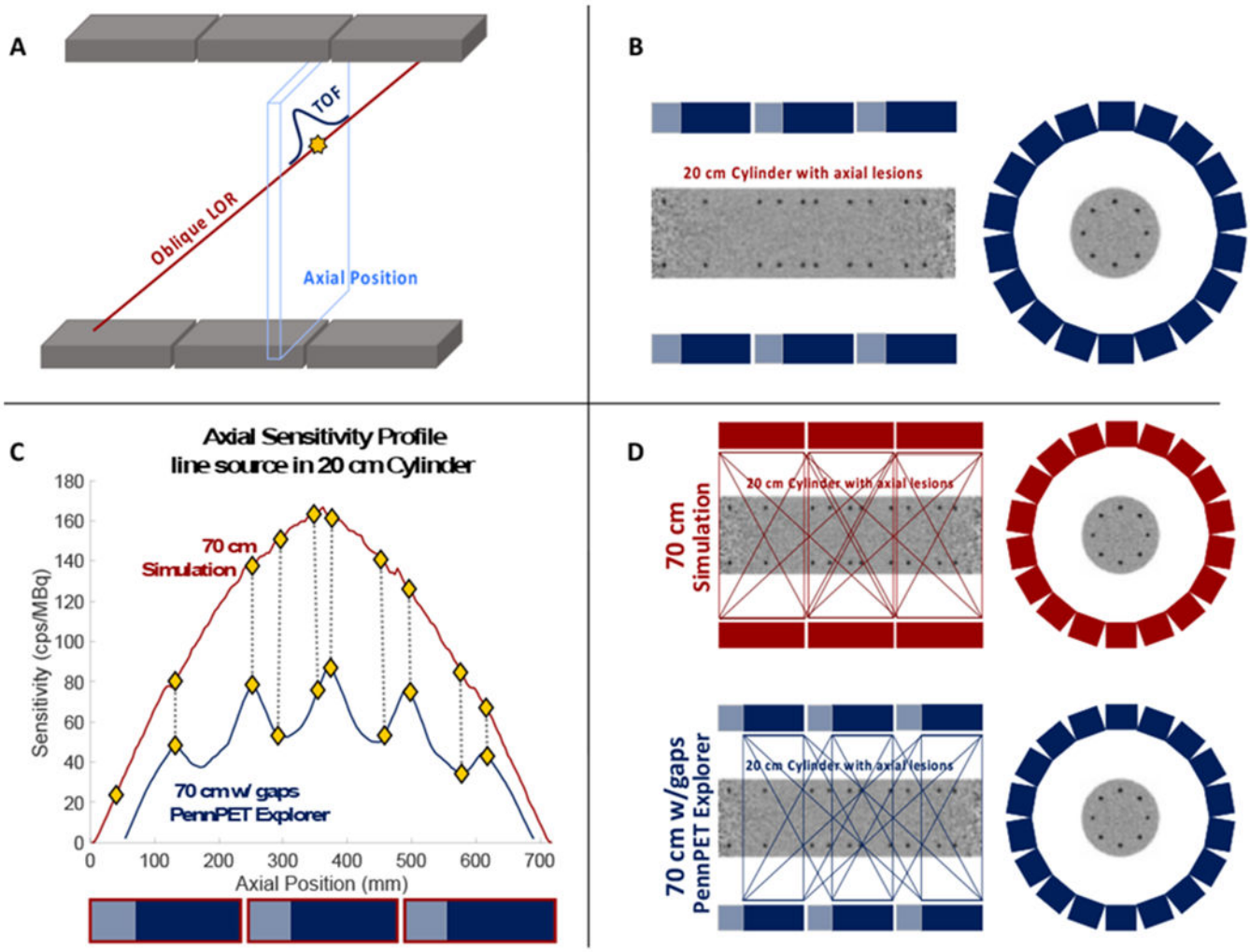


**Figure 1.**

(a) Schematic of the most likely positioning algorithm used to calculate the axial sensitivity for each axial location. Events are deposited into their corresponding axial position, based on the TOF difference of arrival of the two annihilation photons. (b) Schematic illustrating the axial placement of spheres with respect to the inactive regions in the PennPET Explorer. (c) Plot of the axial sensitivity profiles for both simulated and measured scanners along with the axial locations at which spheres were placed (diamonds). Shaded boxes below the plot indicate axial coverage of data active (blue) and inactive (gray) regions of the PennPET Explorer, while the red outlines indicate the axial coverage of the 70-cm simulated scanner. (d) Schematic of the simulated scanner and PennPET Explorer where shaded regions reflect axial LORs that pass through that region. The number of LORs that pass through any given axial location reflects the axial sensitivity at that location.

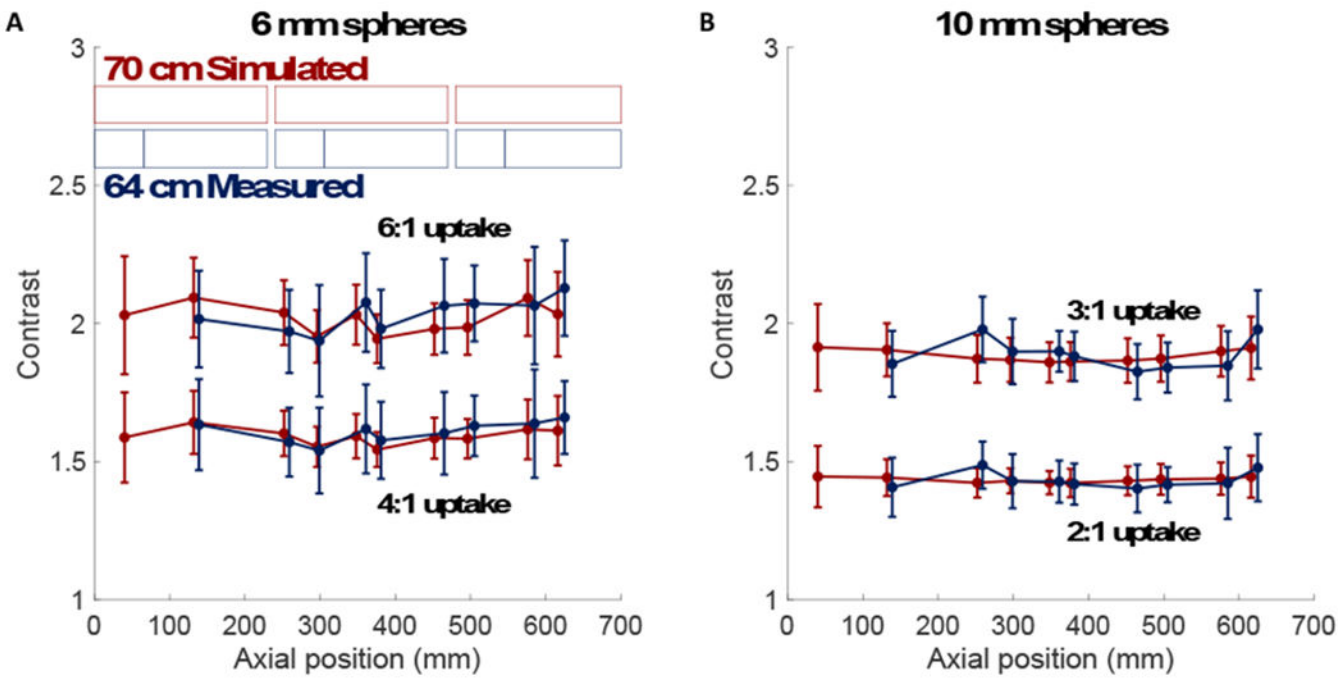


**Figure 2.**

Contrast vs. axial position for (a) 6-mm and (b) 10-mm spheres embedded in a 20-cm cylindrical phantom. (Top) Higher uptake and (bottom) lower uptake spheres are shown for both the simulation in red and measurement in blue. Scanner detector geometries for the 3 ring segments are shown in shaded boxes; simulation (red) and measurement (blue). The measurements cover 64 cm AFOV due to inter-ring data gaps indicated with lighter shading.

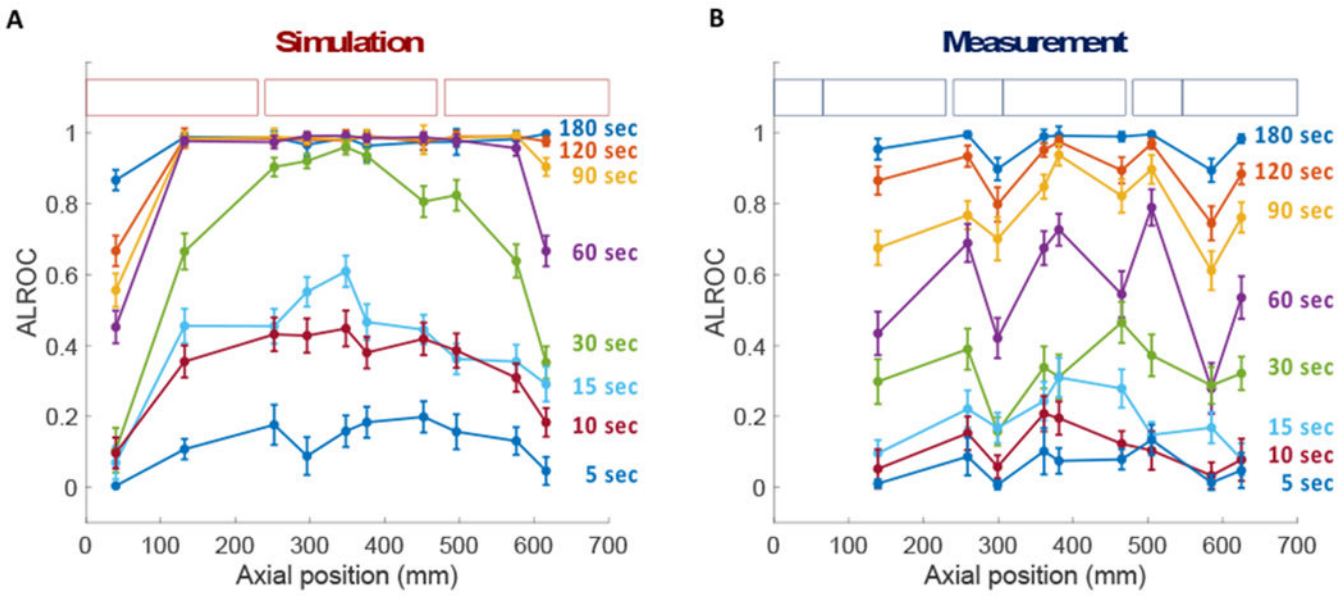


**Figure 3.**

ALROC vs. axial position for 6-mm spheres embedded at 6:1 uptake in the cylindrical phantom at scan times ranging from 5-180 s. Scanner detector geometries for (a) simulation and (b) measurement are shown in shaded boxes.

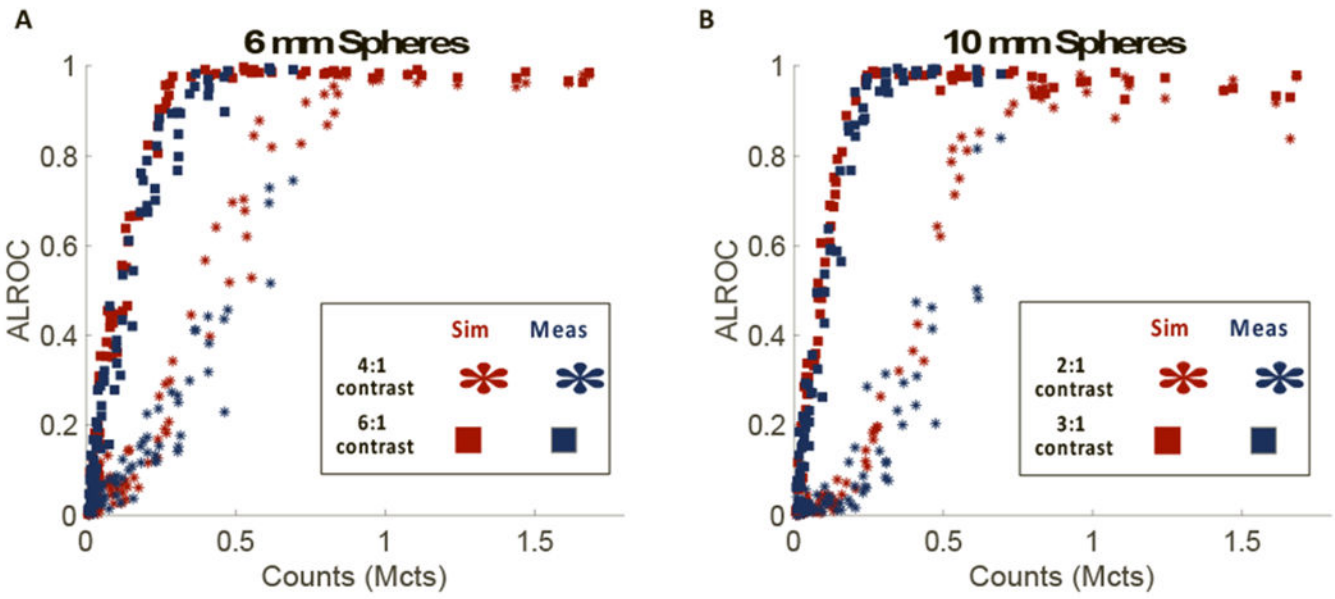


**Figure 4.**

ALROC for (a) 6-mm and (b) 10-mm spheres in the cylindrical phantom as a function of counts. Note that each data point corresponds to a different axial location and counts were determined by the most likely positioning algorithm. Data from the simulation (red) and measurement (blue) are included for both the higher uptake (square) and lower uptake (star) spheres.

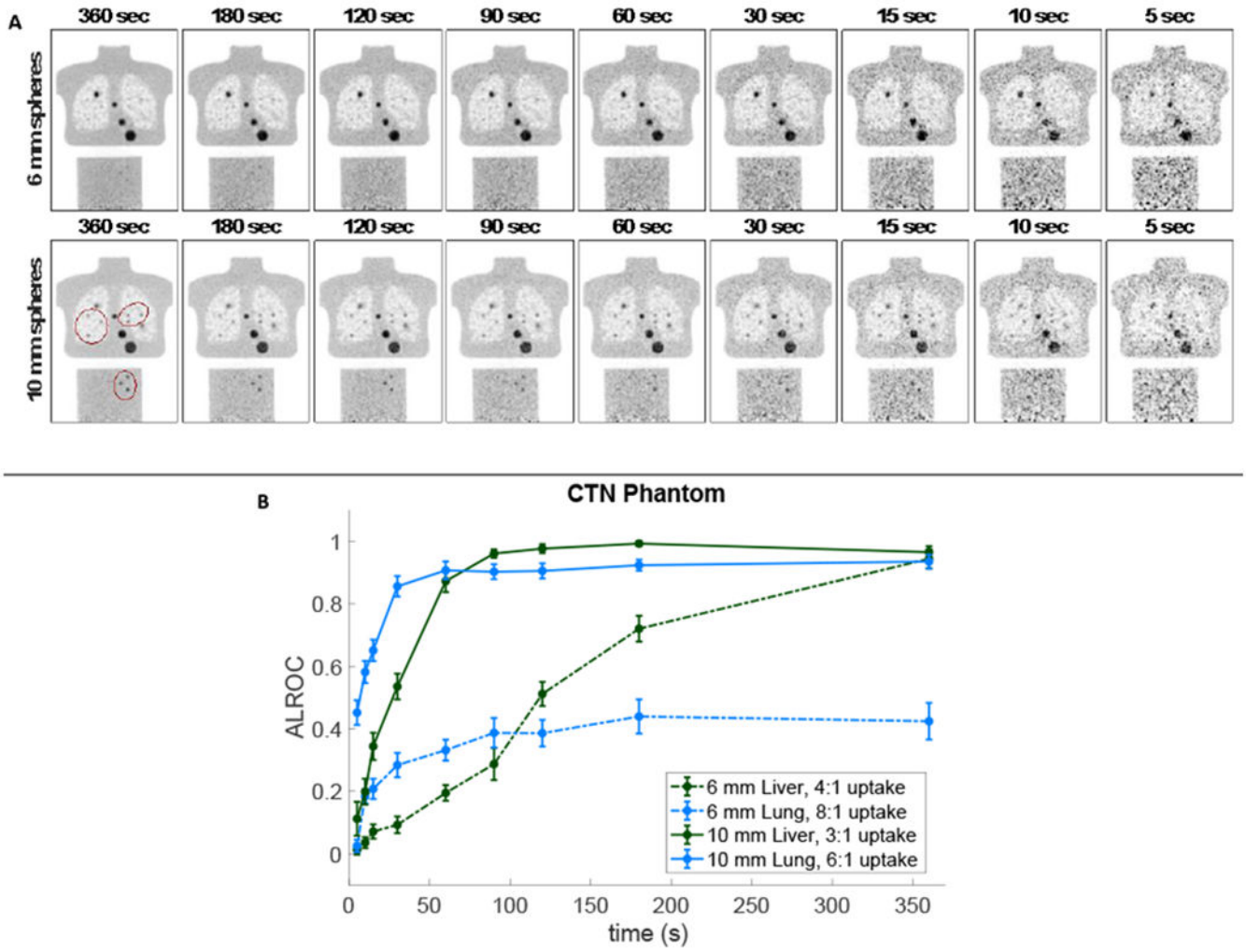


**Figure 5.**

(a) Representative images from the CTN phantom with (top) 6-mm and (bottom) 10-mm embedded spheres for a scan times ranging from 360 s (left) to 5 s (right) and (b) the corresponding ALROC plot as a function of scan duration including 6-mm (dashed) and 10-mm (solid) lesions in both the liver (green) and lung (blue). There were 22 spheres embedded in the lung (6 visible in the 2-mm thick coronal image shown) and 17 spheres in the liver (3 visible) distributed volumetrically throughout each organ region. The embedded lesions are circled to identify them from the physical CTN lesions of varying size.

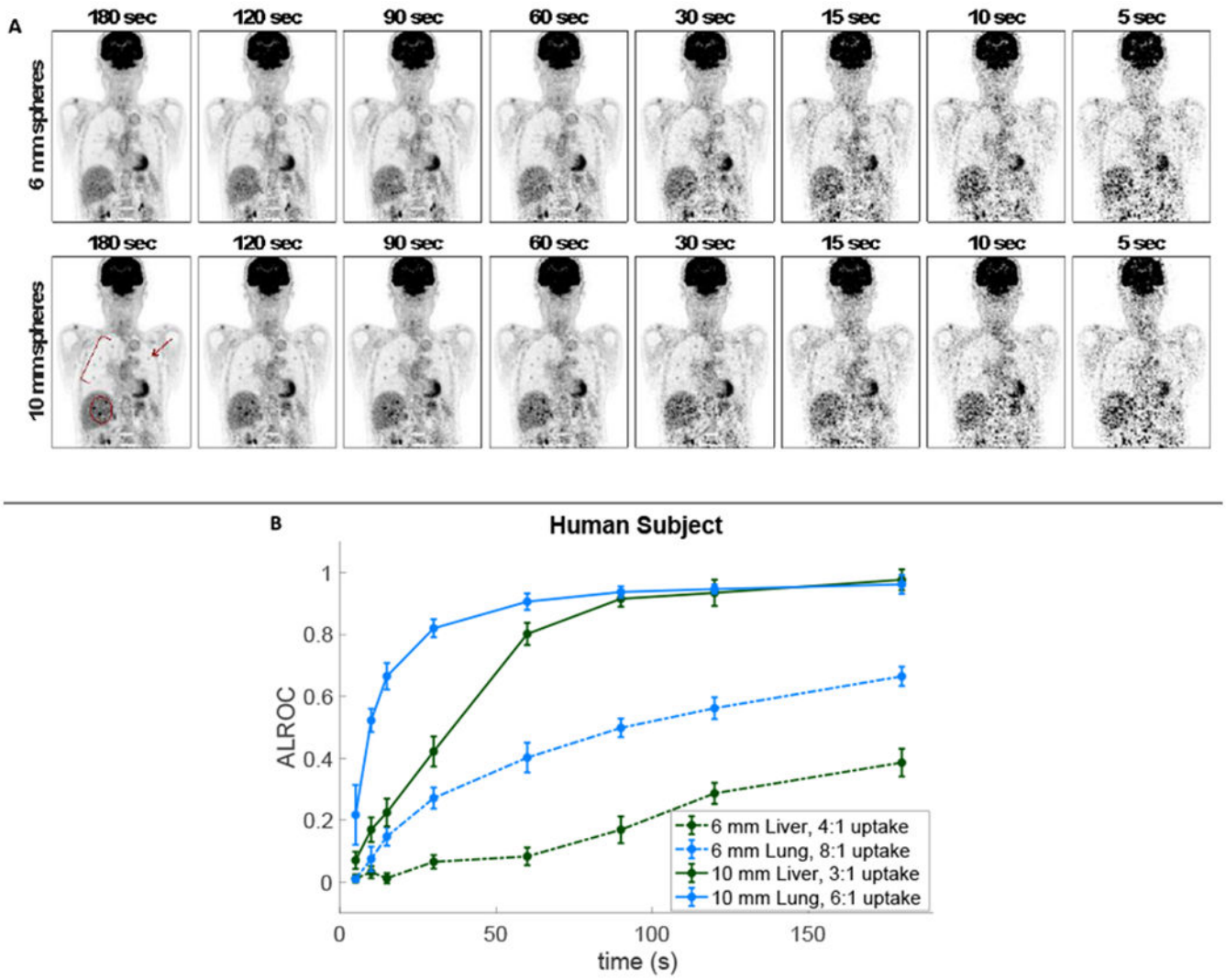


**Figure 6.**

(a) Representative images from the human subject with (top) 6-mm and (bottom) 10-mm embedded spheres for a scan times ranging from 180 s (left) to 5 s (right) and (b) the corresponding ALROC plot as a function of scan duration including 6-mm (dashed) and 10-mm (solid) lesions in both the liver (green) and lung (blue). There were 27 spheres embedded in the lung (5 visible in the 2-mm thick coronal image shown) and 16 spheres in the liver (3 visible) distributed volumetrically throughout each organ. The embedded lesions are highlighted.

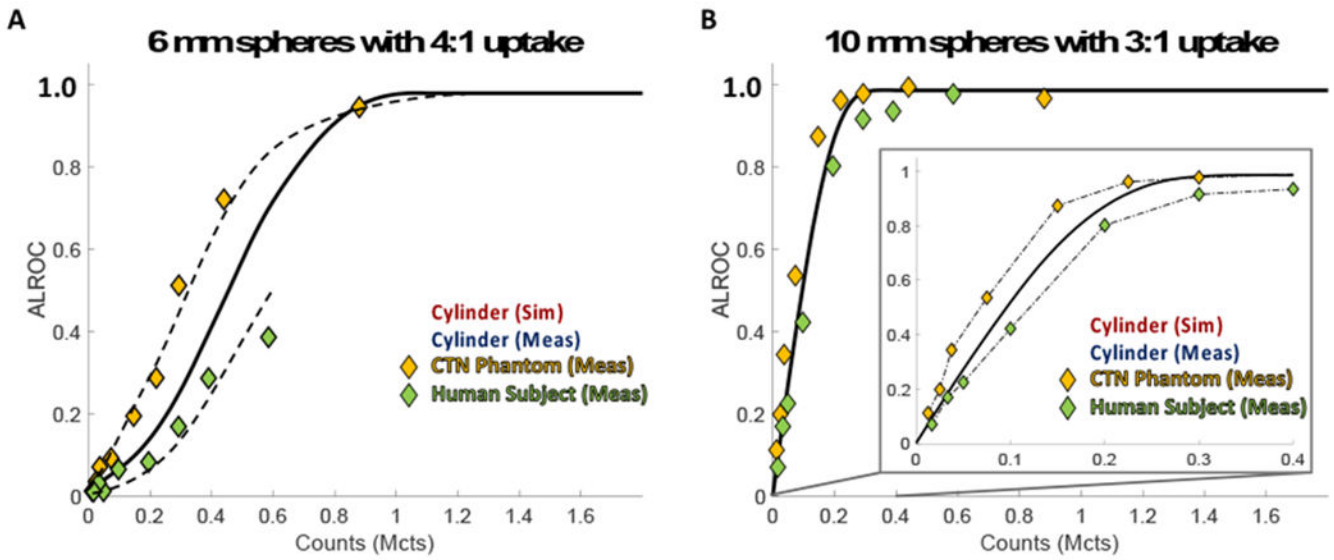


**Figure 7.**

ALROC for the (a) 4:1-uptake 6-mm spheres and (b) 3:1-uptake 10-mm spheres as a function of measured axial counts for the cylindrical phantom (red and blue squares), CTN phantom (yellow diamonds), and human subject (green diamonds) with corresponding trend curves.

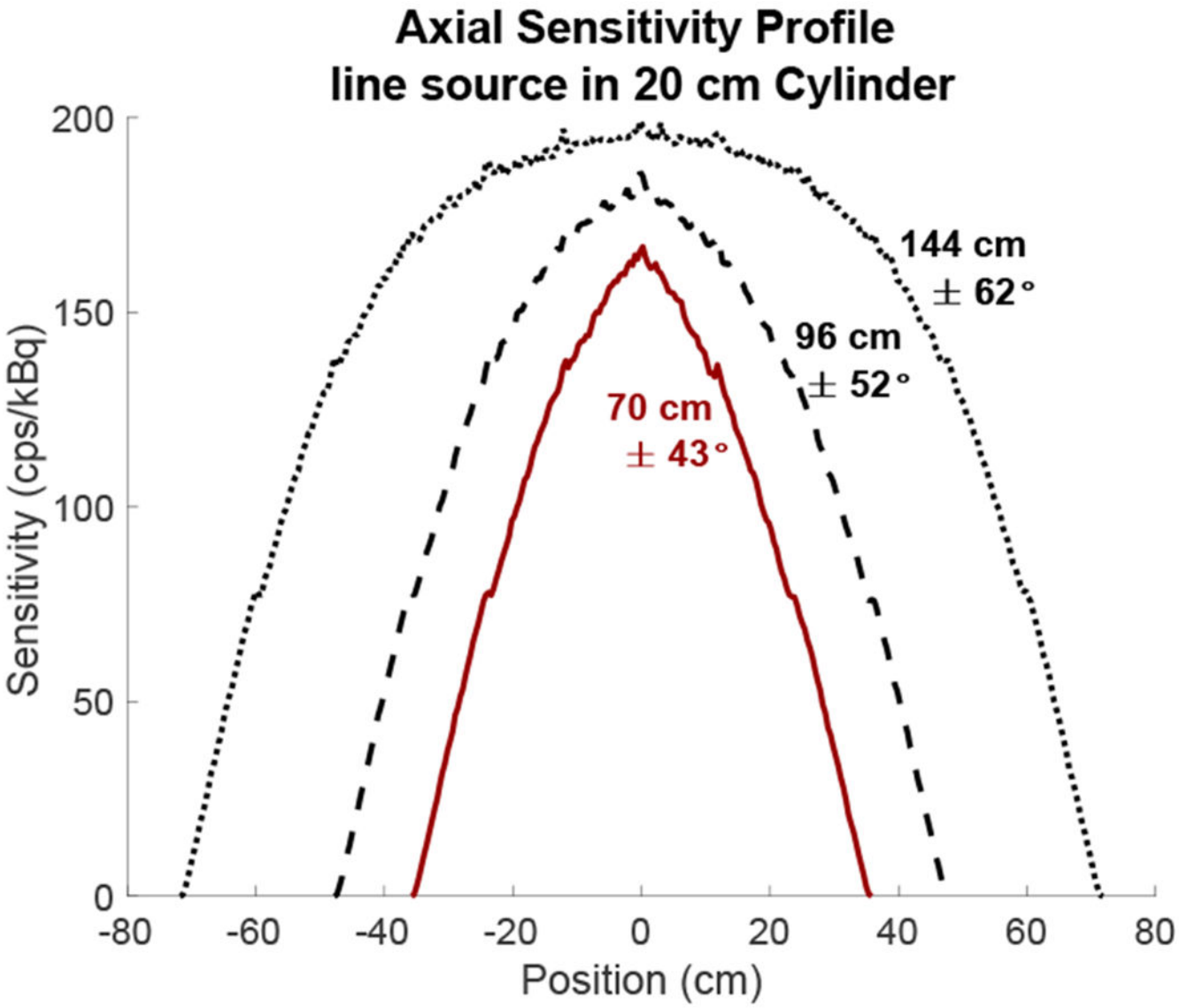


**Figure 8.**

Axial sensitivity profiles for the PennPET Explorer with 70-cm (3 ring segments, red line), 96-cm (4 ring segments, black dashed line), and 144-cm (6 ring segments, black dotted line) AFOV, calculated using a line source in a 20-cm attenuating cylinder. Acceptance angles were not restricted for these profiles and the maximum acceptance angle is noted for each scanner.

**Table I.**

Average measured contrast over all axial positions

| | a. Cylindrical phantom | | | |
|---|---|---|---|---|
| | 6 mm Spheres | | 10 mm Spheres | |
| | ***4:1 uptake*** | ***6:1 uptake*** | ***2:1 uptake*** | ***3:1 uptake*** |
| *Simulation* | 1.59 ± 0.10 | 2.02 ± 0.12 | 1.43 ± 0.06 | 1.88 ± 0.09 |
| *Measurement* | 1.61 ± 0.15 | 2.03 ± 0.17 | 1.43 ± 0.09 | 1.89 ± 0.11 |
| | **b. CTN phantom and human subject** | | | |
| | 6 mm Spheres | | 10 mm Spheres | |
| | ***Liver (4:1)*** | ***Lung (8:1)*** | ***Liver (3:1)*** | ***Lung (6:1)*** |
| *CTN Phantom* | 1.62 ± 0.18 | 2.45 ± 0.53 | 1.92 ± 0.12 | 2.97 ± 0.60 |
| *Human Subject* | 1.54 ± 0.15 | 2.15 ± 0.34 | 1.92 ± 0.20 | 2.84 ± 0.54 |

**Table II.**

Minimum scan duration* with ALROC ≥ 0.80

| | | 6 mm Spheres | 10 mm Spheres |
|---|---|---|---|
| **CTN Phantom** | *Lung* | (360 s)** | 26 s |
| | *Liver* | 239 s | 53 s |
| **Human Subject** | *Lung* | (180 s)§ | 28 s |
| | *Liver* | (180 s)§ | 60 s |

*Scan time estimated by linear interpolation from curves in Figures 5b and 6b.

**ALROC curves reached a plateau at 0.42 and would not be expected to reach 0.80 at any scan duration.

§ALROC curves had not reached 0.80 by 180 s but were continuing to rise and would be expected to reach 0.80 at scan times longer than 180 s