Published in final edited form as:

IEEE Trans Nucl Sci. 2016 June ; 63(3): 1359–1366. doi:10.1109/TNS.2016.2518177.

Effect of Using 2mm Voxels on Observer Performance for PET Lesion Detection

A. Michael Morey,

Utah Center for Advanced Imaging Research, Department of Radiology, and Department of Bioengineering, University of Utah, Salt Lake City, UT 84108 USA

Frédéric Noo [Member, IEEE], and

Utah Center for Advanced Imaging Research, Department of Radiology, the University of Utah, Salt Lake City, UT 84108 USA

Dan J. Kadrmas [Member, IEEE]

Utah Center for Advanced Imaging Research, Department of Radiology, Department of Bioengineering, and Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84108 USA (telephone: 801-581-5937)

A. Michael Morey: ammorey@uair.med.utah.edu; Frédéric Noo: noo@uair.med.utah.edu; Dan J. Kadrmas: kadrmas@uair.med.utah.edu

Abstract

Positron emission tomography (PET) images are typically reconstructed with an in-plane pixel size of approximately 4mm for cancer imaging. The objective of this work was to evaluate the effect of using smaller pixels on general oncologic lesion-detection. A series of observer studies was performed using experimental phantom data from the Utah PET Lesion Detection Database, which modeled whole-body FDG PET cancer imaging of a 92kg patient. The data comprised 24 scans over 4 days on a Biograph mCT time-of-flight (TOF) PET/CT scanner, with up to 23 lesions (diam. 6–16mm) distributed throughout the phantom each day. Images were reconstructed with 2.036mm and 4.073mm pixels using ordered-subsets expectation-maximization (OSEM) both with and without point spread function (PSF) modeling and TOF. Detection performance was assessed using the channelized non-prewhitened numerical observer with localization receiver operating characteristic (LROC) analysis. Tumor localization performance and the area under the LROC curve were then analyzed as functions of the pixel size. In all cases, the images with ~2mm pixels provided higher detection performance than those with ~4mm pixels. The degree of improvement from the smaller pixels was larger than that offered by PSF modeling for these data, and provided roughly half the benefit of using TOF. Key results were confirmed by two human observers, who read subsets of the test data. This study suggests that a significant improvement in tumor detection performance for PET can be attained by using smaller voxel sizes than commonly used at many centers. The primary drawback is a 4-fold increase in reconstruction time and data storage requirements.

Index Terms

image quality assessment; image reconstruction; PET/CT; PET/CT reconstruction

I. Introduction

Positron emission tomography (PET) images are typically reconstructed with an in-plane pixel size of ~4mm for many general oncologic imaging applications, and ~2mm for brain imaging. Reconstruction with smaller pixels has been found to improve spatial resolution and contrast recovery in reconstructed PET images, however, it also affects image noise properties [1–3] and is more computationally expensive. Advances in reconstruction algorithms, computer processing speeds, and storage media have made routine reconstruction with smaller pixel sizes feasible for routine use in the clinical setting.

Although image spatial resolution, contrast, and noise are affected by changing the voxel size, these measures of image fidelity are not necessarily predictive of performance for clinical tasks such as detection and staging of cancer. Image quality can be objectively evaluated using task-based assessments which quantify an observer's ability to perform a task such as detecting a focal, hot lesion on a structured noisy background. This task, relevant to general oncologic PET imaging, includes both determining whether a lesion is actually present (sensitivity) and correctly ruling out noise blobs that are not lesions (specificity) [4]. Receiver operating characteristics (ROC) analysis, and variants thereof, can be used to quantify the observer's performance for lesion-detectability tasks in meaningful measures.

Our group has developed techniques for evaluating general oncologic lesion-detectability in PET using whole-body phantom experiments [5–8], and these methodologies and data have been combined into the Utah PET Lesion Detection Database Resource [9]. The resource contains experimental data useful for performing localization receiver operating characteristics (LROC) studies [10–12] with both the channelized nonprewhitened (CNPW) mathematical observer [13] and human observers. These data and LROC methods have been previously used to evaluate the effects of modeling the point spread function (PSF) [5], TOF [6], and varying the number of OSEM subsets used for iterative reconstruction [8].

The objective of this work was to evaluate the effect of reconstructing with smaller voxel sizes (*e.g.* ~2mm in-plane, as compared to ~4mm) on lesion-detection performance for general oncologic PET imaging. Lesion-detection performance was assessed using standardized metrics, and the effect of decreasing pixel size was evaluated for images reconstructed both with and without PSF modeling and TOF. The following sections describe the experimental data, reconstruction and processing techniques, LROC study methods and results. The effect of using smaller pixels on lesion-detection is then analyzed, and conclusions are drawn based on these data.

II. Methods

A. Experimental Phantom Data

The lesion-detection study used experimental data from the Utah PET Lesion Detection Database Resource [9] for a custom large whole-body phantom scanned on a Biograph mCT TOF PET/CT scanner (Siemens Medical Solutions) with timing resolution 527.5 ± 4.9 ps [14]. The phantom, shown in Fig. 1, has three main components: a 3-dimensional (3D) brain

phantom; anthropomorphic thorax phantom containing liver, lungs, and rib cage; and a pelvis with bladder compartment. The approximate dimensions of the phantom are 43.0×28.0 cm at the widest points, and the total length is approximately 83.1 cm. Accounting for the missing mass of the arms and legs, this phantom models a patient of approximately 92 Kg. The phantom also has a number of custom modifications designed to increase realism for modeling whole-body general oncologic imaging with ^{18}F -fluorodeoxyglucose (FDG).

The experimental data consisted of four days of experiment, with six back-to-back whole-body scans acquired each day. The overall activity levels for the six scans broadly covered the full range of activity levels representative of sites administering 3.4–10.6 mCi FDG with uptake times ranging from 60 to 120 min. Each scan acquired listmode data for four minutes per bed over six bed positions. Three of the four days had 21–23 “shell-less” ^{68}Ge ($T_{1/2} = 270.8\text{d}$) sources [15], diameters 6–16 mm, distributed throughout the phantom lungs, liver, and soft tissue compartments (mediastinum, abdomen, pelvis). These sources modeled tumors with focal FDG uptake with tumor:background ratios ranging from 1.9 to 5.9 in the various phantom compartments and scans. The scans performed on the final day had no lesions present, providing true-negative images for the observer studies. This multi-scan protocol provided numerous images with varying count levels and lesion contrasts, as demonstrated by maximum intensity projection images shown in Fig. 2. Lesion locations and activities were designed to cluster near the verge of detectability, maximizing statistical power for differentiating the test algorithms in the observer studies.

B. Image Reconstruction

The raw scan data, including listmode files, attenuation maps, scanner calibrations, and scatter and randoms estimates, were reconstructed offline using manufacturer-provided reconstruction software. Images were reconstructed with line-of-response (LOR) ordered-subsets expectation-maximization (OSEM) with 14 subsets out to 12 iterations, both with and without spatially variant PSF modeling [16] and TOF. The algorithms analyzed throughout this study are referred to as LOR-OSEM (baseline), PSF, TOF, and PSF+TOF. Corrections for scanner normalization, deadtime, attenuation, scatter, and randoms were applied using the manufacturer-provided reconstruction software.

The reconstructions with each algorithm were repeated with two in-plane pixel sizes: 4.073mm and 2.036mm, referred to as “4mm” and “2mm” throughout this paper. The corresponding reconstructed image matrix sizes were 168×168 and 336×336 , respectively. In all cases, the slice thickness was 2.027 mm. The images for all algorithms and both pixel sizes, including the intermediate iterations, were stored for subsequent processing and analysis.

The true location of each lesion in the phantom was determined from phantom setup coordinate grids and confirmed on the CT scans. A total of 402 lesion-present test cases ($21 - 23$ lesions $\times$ 6 scans/day $\times$ 3 days with lesions present) were used for each algorithm and pixel size, along with the corresponding 402 lesion-absent test image slices taken from the scans acquired without lesions.

C. LROC Studies with CNPW Observer

The CNPW observer, as developed by Gifford *et al.* [13, 17], was used with the LROC study to compute a perception rating and most-likely lesion location for each test image. Additional details regarding the CNPW observer and its training and application to our experimental phantom data can be found in [5–8, 13, 17]. Of note, the observer template was the same size (44mm $\times$ 44mm) for both the 2mm and 4mm pixel images, and hence contained twice as many pixels for the 2mm case. Two versions of the CNPW observer were used in this work: a “2D” observer and a “3D” observer. Here, the 2D observer read single slices of the image (centered at lesion-center), computing the test statistic at every pixel and selecting the location with the highest statistic. The 3D observer was presented with a 7-slice image volume, and searched for lesion locations across the central slice. As such, both the 2D and 3D observers searched the same possible lesion locations, but the 2D observer only had in-slice information whereas the 3D observer had volumetric information (coming from the 3 neighboring slices on each side). The two observers were identical in all other respects.

Figure 3 shows the probability of correct lesion localization (P_{LOC}) plotted as a function of the localization radius acceptance threshold. The P_{LOC} calculation ignores the CNPW observer rating information, and thus corresponds to the false positive fraction (FPF) = 1.0 intercept of the corresponding LROC curve. A radius threshold value of 10.182mm was found to correctly identify ‘hits’ while minimizing random localizations, and this threshold was used throughout the study. Given the size of the phantom and searchable area for each slice, this radius threshold results in less than a 1% chance of randomly locating a lesion. Two figures-of-merit were used for quantifying lesion-detection performance: P_{LOC} and the area under the LROC curve (A_{LROC}). Here, P_{LOC} is the fraction of lesions correctly localized within the 10.182mm threshold, or more simply the fraction of lesions found by the observer. A_{LROC} is the area under the LROC curve, which plots the correctly-localized true positive fraction vs. the false positive fraction, and is computed from the observer rating data and known truth. Higher values for these metrics indicate higher lesion-detection performance. Since the A_{LROC} metric utilizes both location and rating information, and the chance of random localization is less than 1%, a purely-random observer would result in A_{LROC} values less than 0.01.

Preliminary LROC studies using the CNPW mathematical observer were performed in order to select near-optimal parameters. There were 21 post-reconstruction 3-dimensional (3D) Gaussian filters applied to the images for each iteration, with filter width (standard deviation, SD) ranging from 0.0 (no filter) to 8.15mm in 0.41mm increments. The A_{LROC} was computed for each iteration-filter combination. Since changing the pixel size affects both the rate of iterative convergence and noise properties, the iteration number and post-reconstruction filter were optimized for all cases in order to ensure that each test case was evaluated at near-optimal performance. Figure 4 shows example data for TOF images demonstrating how A_{LROC} changed as a function of iteration and filter for both pixel sizes. In order to ensure that the images for both pixel sizes were being evaluated with near-optimal processing parameters, the iteration number and 3D Gaussian filter strength combination that maximized A_{LROC} for each algorithm and pixel size was identified and used for the LROC study with the CNPW observer. This empirical optimization of the

number of iterations and filter strengths required reading 1,620,864 test images to cover 402 lesion-present and lesion-absent test cases for each algorithm, iteration and filter. The resultant parameter values used for the CNPW observer study are listed in Table I.

D. Effect of Changing Pixel Size

The CNPW observer results for the four algorithms and optimal reconstruction parameters were compared across both pixel sizes in order to determine the effect of reducing the pixel size upon P_{LOC} and A_{LROC} for each of the four algorithms. The uncertainty in each figure-of-merit was estimated as the standard deviation over 10,000 bootstrap estimates, where each bootstrap sampled the 402 lesion-present and lesion-absent test images with replacement. The paired-sample Tukey HSD multiple comparison test was then used to test the null hypotheses that the A_{LROC} for 2mm and 4mm pixels were the same, versus the alternative hypothesis that 2mm pixels had higher A_{LROC} than 4mm pixels, for each of the four reconstruction algorithms. These tests were performed with significance level $\alpha = 0.05$. Note that statistical tests comparing the different reconstruction algorithms (LOR vs. PSF vs. TOF) were not performed, as these algorithms have previously been compared and the objective of the current work is to evaluate the effect of using 2mm vs. 4mm pixels.

Key results from the CNPW observer were then confirmed using human observers (as in [17–20]), where manageable-sized subsets of the test data for the TOF and PSF+TOF algorithms at both pixel sizes were read by 2 human observers. Notably, the number of iterations and filter strengths used for the human observer studies, shown in Table I, were heuristically selected to be representative of clinical practices at our institution, rather than the rather high ‘optimal’ number of iterations identified by the CNPW observer (see Table I for comparison).

The CNPW results were first used to identify subgroups of the test images to be read by the human observers that provided challenging detection tasks across all reconstruction algorithms studied. As in previous work [7], scans 2–5 of the 6 scans acquired each day of experiment were found to provide count levels most representative of clinical scans. Of the 268 lesion-present test cases in scans 2–5, the CNPW observer was used to exclude lesions that were either always missed or always found, resulting in a set of 200 test images (100 lesion-present plus 100 corresponding lesion-absent test cases). This process was intended to maximize statistical power of the human observer study by maximizing the number of informative test cases without including such a large number of test cases that observer fatigue became significant. These images were randomly divided in 40 training images and 160 test images for the TOF and PSF+TOF algorithms at each pixel size. Two experienced medical physicists acted as observers for the studies. Note that these observers were not trained PET clinicians; however, previous work demonstrated that such observers are appropriate for the lesion-detection task studied herein [6, 21].

The human observers were blinded to which algorithm was presented, and both the ordering of the test cases and images presented were randomized. The study was performed in 2D, evaluating a single image slice at a time. The observers performed two tasks on each image. First, the location determined to be the most likely to contain a lesion was selected by a mouse click. Second, a confidence rating was selected on a 6-point scale ranging from 1

(high confidence lesion absent) to 6 (high confidence lesion present). The observers were informed that approximately half of the test images would contain lesions, and that each image would have exactly 1 or 0 lesions present, but there could be many noise blobs present. For each test case, the observers first underwent a training session by reading 40 training images. Here, the observers were immediately provided with the truth regarding lesion presence and location after reading each image. Each training session was immediately followed by the test session for the same algorithm and pixel size, where no feedback was provided after reading each image.

The LROC curves for each observer were computed using the non-parametric approach of Popescu [12] with Epanechnikov kernel. The fraction of lesions found (P_{LROC}) and area under the LROC curve (A_{LROC}) for each observer were also computed for each observer, and then averaged to obtain the final results. The methods used were the same as those previously developed and used for the Utah PET Lesion Detection Database Resource, and additional details can be found in [5, 6]. As with the CNPW observer results, a paired-sample Tukey HSD multiple-comparison test was performed to determine statistically-significant differences in A_{LROC} for the different pixel sizes.

III. Results

Figure 5 shows example reconstructed images at both pixel sizes for the TOF and PSF+TOF algorithms. Each image contains one lesion, and no filter was applied to these images so that the differences in noise texture can be visually assessed. Visual differences in spatial resolution, contrast, and background noise can be observed in the images. These differences display different image characteristics for the two pixel sizes, in particular differences in noise texture and lesion contrast.

Figure 6 provides a more detailed example of images with an 8mm lesion in the left lung, reconstructed with the parameters determined for this study (Table I). The contrast of the lesion was markedly higher for the image with 2mm pixels than for the image with 4mm pixels; however, the background noise was also somewhat higher for the smaller pixels. The LROC studies performed in this work objectively assess how these differences in image characteristics affect observer performance for the lesion-detection task. In other words, the LROC studies objectively assess whether or not the improved contrast and noise differences result in improved observer performance for detecting lesions.

The results of the study are presented in Table II, with key results highlighted in Fig. 7. Lesion-detection performance for all reconstruction algorithms improved when using 2mm pixels as compared to 4mm pixels for both the 2D and 3D CNPW observers for all algorithms. The difference was statistically-significant ($p < 0.05$) for all cases except the PSF algorithm with 3D CNPW observer ($p = 0.07$). The magnitude of the differences in A_{LROC} for 2mm vs. 4mm pixels was somewhat lower for the 3D CNPW observer as compared to the 2D CNPW observer, but remained statistically significant for 3 of the 4 algorithms studied. Since improvement using smaller pixels was measured for all algorithms, these results suggest that the use of smaller pixels bring value regardless of whether or not PSF

modeling and/or TOF is used. The magnitude and significance of these improvements are discussed in more detail in the Discussion section.

The main results from the CNPW mathematical observer study were further evaluated with two human observers, who read manageable-sized subsets of the test data as described in the Methods section. These results are shown in Table III and Fig. 8. The human observer results were consistent with the numerical observer results, finding statistically-significant improvement in A_{LROC} when using 2mm vs. 4mm pixels for both algorithms studied. Note that the absolute value of the human and CNPW results should not be compared with each other, as the human observers read only a subset of the images that the CNPW observer did. However, the results show the same trend and similar magnitude of improvement when using 2mm as compared to 4mm voxels, demonstrating consistent results comparing these pixel sizes.

IV. Discussion

When LROC studies are performed, it is important to provide a context for interpreting the magnitude of differences in the figures-of-merit (i.e., in P_{LOC} and A_{LROC}) in clinically relevant terms. Comparison of the different algorithms provides a context for evaluating the degree of improvement attained by using smaller pixels. Reconstructing with 2mm pixels provided a greater degree of improvement in these data than that provided by PSF modeling, and it provided approximately half the degree of improvement as measured for TOF. Notably, the degree of improvement offered by using the smaller pixel size was similar regardless of whether or not PSF modeling was used, and all three reconstruction differences (pixel size, PSF, and TOF) provided cumulative improvements. This suggests that reconstructing with smaller pixels, PSF modeling, and TOF all utilize fundamentally different mechanisms for improving image quality for lesion-detectability.

To provide an additional context for interpreting the results, we repeated the PSF+TOF reconstructions and CNPW LROC studies as a function of scan time. Here, the raw list-mode PET data files were statistically pruned from 240 seconds per bed position to 180, 120 and 90 seconds per bed position (corresponding to whole-body scan times of 24, 18, 12 and 9 min., respectively). This technique for contextualizing the magnitude of LROC study results has been previously established [7]. By repeating the LROC analysis for these images, we computed the change in A_{LROC} as a function of scan time for both 2mm and 4mm pixels. The results are shown in Fig. 9. These data demonstrate that decreasing the pixel size from 4mm to 2mm provides a degree of improvement in lesion-detection performance similar to increasing the scan time by approximately 25–33% per bed position.

It is also important when evaluating the results of this work to take into account the limitations of the study. While the phantom experiments provide a large anthropomorphic object with range of lesion sizes, locations, and count levels, they do not fully represent the full variability that would be encountered in clinical practice. The lesion localization and detection task likewise provides a reasonable but not exhaustive model of the relevant clinical task. Similarly, the CNPW observer is well established for lesion-detection studies of this type; however, it is a model numerical observer that does not fully match human

observer performance. Other model observers, such as the Hotelling observer, which include a pre-whitening component may react differently to the different noise characteristics present in the images reconstructed with 2mm versus 4mm voxels. These limitations are offset in part by the human observer results provided in this work, which confirm certain aspects of the numerical observer studies. Altogether, these data provide objective evidence that reconstruction with 2mm voxels can improve lesion-detection performance, but additional work is necessary to confirm this result and determine the full extent for which such improvement may apply to clinical practice.

Summary and Conclusion

This study evaluated how the use of smaller pixels affects lesion-detection performance in general oncologic PET imaging. The results demonstrate that reconstructing with smaller pixel sizes (*i.e.* ~2mm instead of ~4mm) can significantly improve detection performance for focal lesions in a noisy background. The degree of improvement observed here was greater than that offered by PSF modeling, and was approximately half of that offered by TOF; however, the relative magnitude of these differences depends in part on the phantom size used and may differ in broader situations. Improved performance when using smaller pixels was observed regardless of whether PSF modeling or TOF was used, suggesting that each utilizes different mechanisms to improve detection performance. The degree in improvement by using 2mm pixels was also similar to that observed by increasing the scan time approximately 25–33% per bed position. The primary drawbacks of using 2mm instead of 4mm pixels were approximately four-fold increases in reconstruction time and image storage requirements. These results that reconstructing with smaller voxel sizes may provide an important benefit for general PET cancer imaging applications.

Acknowledgments

This work was supported by grants R03EB014454 and R01EB007236 from the National Institute of Biomedical Imaging and Bioengineering (NIBIB) of the National Institutes of Health (NIH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIBIB or the NIH. Support for the experimental data was also provided in part grant R01CA107353 from the National Cancer Institute, by Siemens Medical Solutions, and by the Ben B. and Iris M. Margolis Foundation.

References

1. Cherry, SR.; Sorenson, JA.; Phelps, ME. Physics in Nuclear Medicine. Elsevier Health Sciences; 2012.
2. Epley, W.; Jacene, H.; Lodge, M.; Wahl, R. J Nucl Med Meeting Abstracts. 2009. The impact of PET pixel size on contrast recovery for small lesions; p. 2018
3. Shakirin, G.; Crespo, P.; Fiedler, F.; Wagner, A.; Enghardt, W. Optimum voxel size for reconstruction of in-beam PET data. Conference Record of the 2008 IEEE Nuclear Science Symposium and Medical Imaging Conference; 2008. p. 5066-5069.
4. Barrett, HH.; Myers, KJ. Foundations of Image Science. Hoboken, New Jersey: John Wiley & Sons, Inc; 2004.
5. Kadmas DJ, Casey ME, Black NF, Hamill JJ, Panin VY, Conti M. Experimental comparison of lesion detectability for four fully-3D PET reconstruction schemes. IEEE Trans Med Imag. 2009; 28:523–34.
6. Kadmas DJ, Casey ME, Conti M, Jakoby BW, Lois C, Townsend DW. Impact of time-of-flight on PET tumor detection. J Nucl Med. 2009; 50:1315–23. [PubMed: 19617317]

7. Kadrmas DJ, Oktay MB, Casey ME, Hamill JJ. Effect of scan time on oncologic lesion detection in whole-body PET. *IEEE Trans Nucl Sci.* 2012; 59:1940–1947. [PubMed: 23293380]
8. Morey AM, Kadrmas DJ. Effect of varying number of OSEM subsets on PET lesion detectability. *J Nucl Med Technol.* 2013; 41:268–73. [PubMed: 24221921]
9. Kadrmas, DJ. The Utah PET lesion detection database. *Conference Record of the 2009 IEEE Nuclear Science Symposium and Medical Imaging Conference*; 2009. p. 3693-3699.
10. Khurd P, Gindi G. Rapid computation of LROC figures of merit using numerical observers (for SPECT/PET reconstruction). *IEEE Trans Nucl Sci.* 2003; 4:2516–2520. [PubMed: 20442799]
11. Khurd P, Gindi G. Fast LROC analysis of Bayesian reconstructed emission tomographic images using model observers. *Phys Med Biol.* 2005; 50:1519–32. [PubMed: 15798341]
12. Popescu LM. Nonparametric ROC and LROC analysis. *Med Phys.* 2007; 34:1556–64. [PubMed: 17555237]
13. Gifford HC, Kinahan PE, Lartzien C, King MA. Evaluation of multiclass model observers in PET LROC studies. *IEEE Trans Nucl Sci.* 2007; 54:116–123. [PubMed: 18392119]
14. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol.* 2011; 56:2375–89. [PubMed: 21427485]
15. Hamill, JJ.; Arnsdorff, CE.; Casey, ME.; Xinli, L.; Raulston, WJ. A Ge-68 PET hot-sphere phantom with no cold shells. *Conference Record of the 2005 IEEE Nuclear Science Symposium and Medical Imaging Conference*; 2005. p. 1609-1613.
16. Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging.* 2006; 25:907–21. [PubMed: 16827491]
17. Gifford HC, King MA, Pretorius PH, Wells RG. A comparison of human and model observers in multislice LROC studies. *IEEE Trans Med Imag.* 2005; 24:160–9.
18. Fiete RD, Barrett HH, Smith WE, Myers KJ. Hotelling trace criterion and its correlation with human-observer performance. *J Opt Soc Am A.* 1987; 4:945–53. [PubMed: 3598746]
19. Gifford HC, King MA, de Vries DJ, Soares EJ. Channelized hotelling and human observer correlation for lesion detection in hepatic SPECT imaging. *J Nucl Med.* 2000; 41:514–21. [PubMed: 10716327]
20. Lartzien C, Kinahan PE, Comtat C. Volumetric model and human observer comparisons of tumor detection for whole-body positron emission tomography. *Acad Radiol.* 2004; 11:637–48. [PubMed: 15172366]
21. Kadrmas DJ, Christian PE. Comparative evaluation of lesion detectability for 6 PET imaging platforms using a highly reproducible whole-body phantom with (^{22}Na) lesions and localization ROC analysis. *J Nucl Med.* 2002; 43:1545–54. [PubMed: 12411558]

Fig. 1.

The whole body phantom (top) consists of a brain compartment; thorax with liver, lungs, and rib cage/spine; and pelvis with bladder. It models a patient of approximately 92 Kg. Coronal PET (bottom left) and CT (bottom right) images show the main phantom compartments and structures. Example lesions can also be seen in the PET image in the lungs, the mediastinum, and pelvis regions.

Fig. 2.

Example maximum intensity projection (MIP) images with lesions present. The upper limit of the grayscale was lowered to enhance visualization of the body compartments. The images show increasing lesion contrast for each successive scan as the ^{18}F background decayed while the ^{68}Ge lesions remained ~constant. This provided multiple contrasts for each lesion, ranging from undetectable to borderline to easily visible, increasing statistical power for the observer studies.

Fig. 3.

The fraction of lesions correctly localized by the 2D and 3D CNPW observers, plotted for each pixel size as a function of radius of correct localization for the PSF+TOF algorithm. A radius threshold of 10.182mm was used in this work to determine correct localization.

(a)

(b)

Fig. 4.

Example analysis TOF results used for selecting number of iterations and filter strength. Top plot (a) shows A_{LROC} vs. iteration for both pixels studied. Here, the data are shown for filters that maximized A_{LROC} at each iteration. Analogous plot below (b) shows A_{LROC} vs. filter SD, where each datum is shown for number of iterations that maximized A_{LROC} for that filter strength. These data represent a portion of multidimensional sampling used to optimize the number of iterations and filter strength for phantom data used in this work.

Fig. 5.

Example unfiltered reconstructed images with 2mm and 4mm pixels for both TOF (a) and PSF+TOF (b) reconstructions. Each image contains exactly one lesion (from left-to-right: right lung, mediastinum, left lung, and abdomen). Visual differences in noise textures and spatial resolution properties are evident for the two pixel sizes.

Fig. 6.

Example reconstructed images with 4mm and 2mm pixels and optimum iteration and filter as determined by this study, demonstrating potential effects on lesion-detection. Focus in the left lung (white arrow) is a true 8-mm hot lesion. Horizontal profiles showing relative intensity (arbitrary units) demonstrate that using 2mm pixels resulted in an increase in lesion contrast (black arrow), as compared to 4mm pixels.

Fig. 7.

Results for the 3D CNPW observer showing lesion-detection performance, as measured by A_{LROC} , shown for 2mm and 4mm pixels and each of the reconstruction algorithms studied. Lesion-detection performance was significantly higher using 2mm pixel size for all cases studied.

Fig. 8. LROC curves for the human observer study for 2mm and 4mm pixels for the TOF (a) and PSF+TOF (b) algorithms. The plots show the correctly localized true positive fraction (TPF) as a function of the false positive fraction (FPF). These human observer results demonstrate improved lesion-detection performance when using 2mm pixels, confirming the results for the CNPW mathematical observer.

Fig. 9.

Comparison differences in lesion-detection performance for changing pixel sizes versus changing scan time for PSF+TOF reconstructions. These data provide a context for interpreting the significance of changes in A_{LROC} observed in this work. For example, use of 2mm pixels instead of 4mm pixels provided an improvement in performance roughly equivalent to increasing the scan time 25–33% per bed position.

TABLE I

Selected Reconstruction Parameters

Model	4mm		2mm	
	No. of iterations	Filter SD (mm)	No. of iterations	Filter SD (mm)
2D CNPW Observer Study				
LOR-OSEM	12	5.29	7	2.85
PSF	12	5.29	12	4.07
TOF	6	3.67	4	2.44
PSF+TOF	6	4.07	9	2.85
3D CNPW Observer Study				
LOR-OSEM	11	2.44	6	2.44
PSF	11	2.44	12	3.67
TOF	5	0.41	6	1.63
PSF+TOF	12	2.04	9	2.04
Human Observer Study				
TOF	4	3.67	4	4.07
PSF+TOF	5	3.26	5	3.67

TABLE II

CNPW Observer Results

Model	P _{LOC} ± SD		A _{LROC} ± SD		Tukey HSD Test (P) ^a
	4mm	2mm	4mm	2mm	
2D CNPW Observer					
LOR-OSEM	0.629 ± 0.024	0.669 ± 0.024	0.566 ± 0.027	0.592 ± 0.023	0.001
PSF	0.639 ± 0.024	0.667 ± 0.024	0.572 ± 0.023	0.598 ± 0.023	0.005
TOF	0.677 ± 0.023	0.711 ± 0.023	0.614 ± 0.023	0.646 ± 0.022	<0.001
PSF+TOF	0.684 ± 0.023	0.724 ± 0.022	0.628 ± 0.023	0.663 ± 0.022	<0.001
3D CNPW Observer					
LOR-OSEM	0.649 ± 0.024	0.679 ± 0.023	0.574 ± 0.023	0.592 ± 0.023	0.041
PSF	0.654 ± 0.024	0.669 ± 0.023	0.581 ± 0.023	0.596 ± 0.023	0.072 (NS) ^b
TOF	0.692 ± 0.023	0.724 ± 0.022	0.626 ± 0.023	0.656 ± 0.022	0.002
PSF+TOF	0.706 ± 0.023	0.739 ± 0.022	0.645 ± 0.023	0.666 ± 0.022	0.008

^aTukey HSD multiple comparisons test performed on ALROC figure-of-merit^bNS – not significant

TABLE III

Human Observer Validation Study Results

Model	P _{Loc} ± SD		A _{LROC} ± SD		Tukey HSD Test (P) ^a
	4mm	2mm	4mm	2mm	
Human Observer Average					
TOF	0.750 ± 0.035	0.838 ± 0.029	0.656 ± 0.034	0.734 ± 0.031	0.022
PSF+TOF	0.757 ± 0.034	0.832 ± 0.029	0.662 ± 0.034	0.739 ± 0.030	0.015

^aTukey HSD multiple comparisons test performed on A_{LROC} figure-of-merit for human observer average only