Quantitative Accuracy and Lesion Detectability of Low-Dose FDG-PET for Lung Cancer Screening

Joshua D. Schaefferkoetter1,2, Jianhua Yan3, Therese Sjöholm1, David W. Townsend1,2, Maurizio Conti4, John Kit Chung Tam 5, Ross A. Soo6,7, Ivan Tham1,8

1 A*STAR-NUS, Clinical Imaging Research Centre, Singapore
2 Department of Diagnostic Radiology, National University Hospital, Singapore
3 Department of Nuclear Medicine, First Hospital of Shanxi Medical University, China
4 Siemens Healthcare Molecular Imaging, Knoxville, TN, USA 37919
5 Division of Surgery, National University of Singapore, Singapore
6 Department of Haematology-Oncology, National University Cancer Institute, Singapore
7 Cancer Science Institute of Singapore, National University of Singapore, Singapore
8 Department of Radiation Oncology, National University Cancer Institute, Singapore

For correspondence or reprints contact: Joshua D. Schaefferkoetter, A*STAR-NUS Clinical Imaging Research Centre, Centre for Translational Medicine (MD6), 14 Medical Drive, #B1-01, Singapore 117599.
E-mail: dnrjds@nus.edu.sg
ABSTRACT

Lung cancer remains responsible for more deaths worldwide than any other cancer, but recently there has been a significant shift in the clinical paradigm regarding the initial management of subjects at high risk for this disease. Low dose computed tomography (CT) has demonstrated significant improvements over planar X-ray screening for patient prognoses and is now performed in the U.S. Specificity of this modality, however, is poor and the additional information from positron emission tomography (PET) has the potential to improve its accuracy. Routine screening requires consideration of the effective dose delivered to the patient, and this work investigates image quality of PET for low-dose conditions, in the context of lung lesion detectability. Reduced radiotracer doses were simulated by randomly discarding counts from clinical lung cancer scans acquired in list-mode. Bias and reproducibility of lesion activity values were relatively stable even at low total counts of around 5 million trues. Additionally, numerical observer models were developed and trained with the results of 2 physicians and 3 postdoctoral researchers with PET experience in a detection task; detection sensitivity of the observers was well correlated with lesion signal-to-noise ratio (SNR). The models were used prospectively to survey detectability of lung cancer lesions, and the findings suggested a lower limit around 10 million true counts for maximizing performance. Under the acquisition parameters used in this study, this translates to an effective patient dose less than 0.4 mSv, potentially allowing a complete low dose PET/CT lung screening scan to be performed under 1 mSv.

Keywords: Lung cancer, PET/CT, low dose, screening

Running Title: Low dose FDG-PET/CT for lung screening

Word Count: 5,757
INTRODUCTION

Lung cancer is still the cancer with one of the worst prognoses and is a major source of mortality and morbidity. There were about 1.8 million new cases (12.9% of all cancers) and 1.5 million deaths from lung cancer worldwide in 2012 (1). Approximately 158,040 Americans died from the disease in 2015, accounting for 27 percent of all cancer deaths in the United States (2). Data from the Surveillance, Epidemiology and End Results program of the National Cancer Institute show that about 80% of patients are diagnosed with regional or distant disease, which is strongly associated with a poor overall survival rate (3).

CT lung cancer screening has been shown to improve survival over chest radiography. The National Lung Screening Trial recruited 53,454 current or former heavy smokers aged 55 to 74 in multiple centers over a 21-month period (4). The study showed a relative reduction in mortality from lung cancer with low-dose CT screening of 20% relative to chest X-ray screening (5). Following a systemic review, including this and several smaller studies (6), the US Preventive Services Task Force now recommends screening with helical CT instead of chest X-ray radiography. The Centers for Medicare and Medicaid Services now provides coverage for this procedure annually, making this the first time that lung cancer screening has been covered.

Despite the recognized benefits of CT screening, there are important considerations. The diagnostic advantage of CT over chest X-ray is attributed to its high sensitivity for detecting small cancerous lesions, but this can potentially lead to misclassification of non-malignant nodules and patient overdiagnosis. The National Lung Screening Trial researchers reported that CT screening yielded a false positive rate around 96% (5). The addition of metabolic information from PET has been shown to improve accuracy for detecting lung cancer compared with CT alone (7). Hence, recent work has focused on potential screening applications of the combined modality PET/CT in high risk groups (8). Currently, PET/CT is used clinically for staging (9,10), monitoring treatment response (11-13), and long term surveillance (14,15). The application of this hybrid modality as a screening tool, whereby subjects are scanned annually, would
need low-dose alternatives to current scanning protocols (16-19). For PET/CT, the CT component is typically responsible for a larger effective radiation dose delivered to the patient, relative to its PET counterpart, but while significant work has focused on developing appropriate low-dose CT protocols, far less attention has been paid to reducing the PET tracer dose. This would require efforts to understand these effects on image quality, in the medical setting.

PET is an intrinsically noisy modality, and the quality of the reconstructed images depends greatly on the number of acquired coincident counts. Many previous works have focused on the behavior of statistical PET reconstructions in low-count conditions (18,20-25) and typically characterized image quality in terms of noise and bias. However, more work is needed to translate the implications of these findings for specific clinical tasks. A platform has been presented to evaluate the noise effects of reducing the amount of PET data on the performance in the specific task of detecting small isolated lung nodules. The analysis tools were developed for patients with infectious lung lesions, scanned with simultaneous PET/magnetic resonance (MR) (26).

The present work applies these analysis methods to a cohort of patients with malignant lung nodules scanned with PET/CT. It aims to investigate the lowest practical limits for accurate image quantification while maximizing sensitivity for detecting small, isolated lung lesions.

**METHODS**

Twenty patients (range 37.2-91 kg) were enrolled in this Institutional Review Board approved study after written informed consent was obtained. Inclusion criteria were patients with biopsy-proven primary lung cancer or patients with suspicious radiological abnormalities planned for definitive lung surgery. All scans were performed on the Biograph mCT (Siemens Healthcare Molecular Imaging) after an uptake period of 60 min, following injection of 218.3±5.18 MBq 2-deoxy-2-(18F) fluoro-D-glucose (FDG). PET data were acquired in list-mode, and the true scan counts were found by subtracting the smoothed delayed counts from the total prompts (throughout this manuscript, “true counts” refers to prompts minus randoms,
which are actually the true and scattered events). All subjects were scanned with 2 bed positions covering
the lungs at 10 min each, resulting in 120±25 million mean true coincident counts per bed position.

Reduced effective doses were simulated by randomly discarding events in the PET list-mode according
to 9 predefined true count levels, defined as prompts minus delayed: 0.25×10^6, 0.5×10^6, 1×10^6, 2×10^6,
5×10^6, 7.5×10^6, 10×10^6, 15×10^6, and 20×10^6. For each patient-true count combination, the highest possible
number of independent realizations was generated and reconstructed, up to a maximum of 50. The
reconstruction algorithm was Ordinary Poisson Ordered Subset Expectation Maximization (OP-OSEM)
(27,28), using Time of Flight (TOF) and Point Spread Function (PSF), with 2 iterations, 21 subsets, and 3
mm smoothing, producing 400 x 400 image matrices with voxel size 2.04 x 2.04 x 2.03 mm. The CT
images were 512 x 512 with voxel size 1.52 x 1.52 x 5 mm. All PET reconstructions included attenuation
and scatter corrections.

This work focused on PET image quality in the context of lung lesion detection. The images
reconstructed from the original full statistical set, i.e. all acquired events, were used to identify isolated lung
lesions of various sizes and contrasts, consistent with those expected of early-stage, subclinical lung cancer.
Cubic volumes of interest (VOIs) with dimensions 32.6 × 32.6 × 32.5 mm, centered on each selected lesion,
were delineated and stored. Additionally, for each patient, one cubic VOI including only healthy lung tissue
was also stored, generating a population of matched, lesion-absent test samples. Hence, each patient
contained one background and at least one signal sub-volumes. The bias and stability of the lesion activity
measurements were evaluated across all simulated dose levels, and detectability was determined by various
human-trained, numerical observer models.

At each simulated dose, the selected lesions were classified by 6 parameters: metabolic PET lesion
volume, FDG signal-to-background contrast, mean lesion activity concentration, lesion VOI standard
deviation, lesion-to-background SNR, and channelized Hotelling observer (CHO) SNR (29). Masks
delineating lesion VOIs were generated in the full-count images by selecting all voxels within the cubic
volume with values equal to or greater than 40% of the maximum value. The VOI segmentations are illustrated in Figure 1, which shows coronal slices containing lesion and background regions.

The metabolic PET volume is the 1st parameter and was defined as the number of voxels included in the lesion VOI multiplied by the individual voxel volume, in this case 8.45 mm$^3$. The mean activity concentrations in the lesion and background VOIs were used to calculate the 2nd parameter lesion contrast $C_{les}$, given by

$$C_{les} = \frac{\mu_{les} - \mu_{back}}{\mu_{back}}$$

where $\mu_{les}$ and $\mu_{back}$ are the means of the lesion and background ROIs in the full count images. These first 2 parameters were measured only in the full-count images, providing stable representations of size and contrast for each lesion.

The mean activity concentration in the lesion VOI was measured for all the independent realizations at each simulated dose; the mean of these realization measurements was recorded as the 3rd parameter lesion activity mean, and the standard deviation in these measurements was the 4th parameter lesion VOI standard deviation, providing a way to assess the reproducibility of the lesion. The 5th parameter was lesion SNR and was defined as the lesion activity mean divided the mean of the noise, i.e. the standard deviation in the background VOI, averaged across realizations. The last parameter used to characterize the lesions was included to reflect observer detectability. A CHO was implemented to provide a surrogate estimate of performance in a lesion detection task. For every lesion, at each dose, the target and background sub-volumes extracted from the independent noise realizations were used to train the model and test the observer SNR (30,31).

As an example, we demonstrate in Figure 2 the analyses for a sample lesion (the same lesion shown in Figure 1).
Defining each lesion by the 6-dimensional point comprising its parameter values provided a simplified and convenient way to represent all lesion samples in a space spanning their morphological, physiological and distinguishable characteristics.

A lesion detection task to survey performance for distinguishing lesions from healthy lung tissue was previously administered to 2 radiologists, board certified in nuclear medicine, and 3 postdoctoral researchers experienced with PET (26). Five hundred and fifty images of patients (range 45-79 kg) containing hypermetabolic lung lesions were randomly presented to the observers, with 3 orthogonal slices intersecting an outlined volume of interest. The observers were instructed to report if there was a lesion at the center of the VOI or not. Each observer then rated the confidence of the decision on a scale from 1 to 5, with 5 meaning 100% confidence of a lesion, 1 meaning 100% confidence of no lesion, and 3 meaning an equivocal interpretation. A lesion with a score of 4 or higher was assumed to be detectable by the observer. For each count category, sensitivity was calculated as

\[
Sensitivity = \frac{\text{Number of Detected Lesions}}{\text{Total Number of Lesion Test Samples}}
\]

where each sample point was classified as detected or not detected, according to a human-trained linear observer model.

The human observer decision data were used to classify the retrospective lesion data into 2 classes, detectable or not-detectable; these were the training data for the linear observer models. The task of the numerical observer was to determine to which class a given sample belonged. The discriminant function \( w_{obs} \) was calculated by

\[
w_{obs} = \frac{\mu_1 - \mu_2}{K}
\]

Here \( \mu_i \) is the mean of the \( i^{th} \) class and \( K \) is the covariance derived from the combination of the individual class covariances according to
where $N_i$ is the number of samples belonging to the $i^{th}$ class $C_i$. This approach, i.e. linear discriminant analysis (LDA) for discrete samples, ensured that SNR for class discrimination was maximized.

By representing the retrospective lesion data (and defining the corresponding linear discriminants) in the same vector space as the prospective lesions, i.e. comprising combinations of the 6 measurement parameters, the observer models could be directly applied to the new lesion data. By choosing the binary decision threshold that resulted in overall accuracy most closely matching that of the original observer, the numerical model could thus allow trained predictions for classifying the new lesions as detectable or not-detectable. The scatter plots in Figure 3 show this approach. The lesion samples, are represented as points comprising their respective parameters, along with the corresponding linear discriminant function.

RESULTS

In all 20 prospective patients, only isolated and relatively small lesions were chosen to simulate a typical screening situation; twelve lesions were selected as having suitable morphological and physiological characteristics for inclusion in the analyses. The mean lesion volume was $1.25\pm1.21$ cm$^3$, range 0.18-3.80, as measured by PET, and the mean lesion activity concentration was $13.9\pm11.4$ kBq/mL, range 3.3-37.5, (standardized uptake value (SUV) $6.2\pm5.5$, range 1.31-18.63). The mean volume was $2.61\pm2.86$ cm$^3$, range 0.15-7.15, as measured by CT, using the same thresholding approach used to delineate the lesions in the PET volumes. The lesion characteristics are described in Table 1.

For all comparative analyses reported here, the images generated from the full count data were used to draw volumes of interest and taken to be the “gold standard” for evaluating the quantification accuracy of the reconstructions at reduced counts.

The lesion metrics (aside from volume and contrast, which were constant across count levels) are shown as a function of true count level in Figure 4. Lesion SUV was relatively stable until the count level
approached 1 million, and only when the counts were decreased to 5 million did the average standard deviation become greater than 10%. Lesion SNR and CHO SNR exhibited continuous increase with count level.

Noisy data lead to bias in statistical reconstructions, and this effect on lesion SUV was investigated. The SUV mean and max measurements were relatively stable at all count levels, and as seen in Figure 5, it was only when the true counts approached 1 million that measurement bias was observed. As expected, in low-count conditions, SUV max showed positive bias since increases in the variability within a VOI yield maximum voxel values further from the mean. Negative bias was observed for SUV mean under these extreme conditions due to two phenomena, the backprojection of the sparse data within the OSEM framework (24) and, to a greater extent, the failure of the scatter simulation, namely the scaling part (32), which greatly overestimates the scatter at extremely low counts. Both of these phenomena are shown in Figure 6 for the same patient dataset used in Figure 2.

Validations of the best numerical models to accurately predict the respective human observers’ decisions were first performed using the retrospective data from the detection task. Once the linear discriminant was defined (and corresponding “accuracy-matched” decision threshold calculated) in each case, the model was used to classify the lesion data as detectable or not-detectable. Once the predictive utility of the linear discriminators was established, the observer models were applied to the new lesions from the prospective lung cancer patients. The models predicted different performance for each observer, but for all 5 human observers, as seen in Figure 7, performance was optimized near 10 million true counts.

**DISCUSSION**

This work was primarily motivated by the need for improvements to current lung cancer screening protocols. The poor specificity of CT lung screening, with a false positive rate around 96%, leads to potential over-diagnoses and, in some cases, unnecessary invasive procedures that are not without risk. The addition of complementary and/or confirmatory information needs to be incorporated into the screening process so that the referring physician can make an informed decision. Hybrid PET/CT has consistently
proved superior to CT alone since PET can provide information that improves the poor specificity of current screening methods. This has the potential to considerably improve the clinical management of high-risk patients by increasing overall diagnostic accuracy. If PET is to become a routine screening tool for patients at risk however, the dose should be kept as low as practically possible. Investigative efforts must focus on defining the lowest reliable limits of PET. Here, we present a task-based evaluation, specific to detection of small focal lesions in lung cancer patients. This work combines objective and subjective analyses to provide a comprehensive understanding of the behavior of clinical PET and lesion detectability in noisy conditions.

All experimental findings presented here are based on reduced PET tracer doses, emulated by randomly discarding count events from larger sets of PET list-mode data. This approach is predicated on the assumption that the ratio between the prompt and delayed events is constant throughout a given dataset, regardless of the randomly emulated count rates. However, although the prompts and singles rates scale linearly with the in-field activity, the randoms rates scale by the square of the singles. Hence, in this work, we essentially create a worst-case situation, where we emulate low count levels but with the higher randoms rates of the original data, which was 38.8% (range 34.8-44.4%) in this study. With actual lower injected activities, the randoms fraction and deadtime would be lower, and so image quality would be improved.

Generating smaller data sets from a larger one has the advantage that several random realizations of the same activity distribution can be realized, permitting the characterization of the noise associated with the reconstruction process. For this work, only independent realizations were created in order to eliminate possible correlations between test images. Of course, for any set of data, one is limited to the number of fully independent subsets that can be realized. For the current work, the authors felt that there was an appropriate balance between count level and maximum realization number. High count levels yielded fewer realizations but the reconstructions were supported by high statistics and low variability, and lower count groups produced greater numbers of realizations to compensate instability of the increasing noise levels.
This was not an issue for the lowest count levels (≤ 2 million), as 50 independent realizations could be extracted.

Analyses of bias and reproducibility in the lesion activity values showed that the measurements were stable until the count levels approached extreme conditions; in fact, bias in the lesion VOI mean and max SUV appeared relatively negligible until the true count level was decreased to 1 million. Variance on the reproducibility of lesion values showed a more dramatic trend, but the standard deviation was still around 10% at 5 million counts. It could easily be argued that this level is acceptable since PET is generally associated with a test-retest error on this order (33,34).

As seen in Figure 6, instability in the reconstruction at extremely low counts is caused by 2 phenomena, estimation of the scatter fraction and handling of the sparse projection data within the OSEM framework. At extremely low counts, the scatter simulation, and in particular the associated tail-fitted scaling, fails due to very noisy data in the sinogram tails. This results in an overestimation of the scatter: we verified that below 1 million true counts, the scatter fraction soars to unrealistic high values, and it is the main reason for the strong negative bias in SUV at the lower limits of our statistical range. In addition, ordering the subsets in iterative PET reconstruction, which is designed to accelerate reconstruction convergence by using a limited number of projections for each image update, increases the probability of back-projecting zeros into the image at low statistics. This, along with the multiplicative nature of the updates in the expectation maximization framework, can essentially “trap” image voxels at zero, resulting in negative bias. This problem is somewhat reduced by using pure Maximum Likelihood EM, i.e. using all projections for each update, but this would not be clinically practical.

Although, the objective analyses of lesion metrics suggested that image quality may be acceptable at very low count levels, the results of the subjective observer study were more conservative and likely better represent the actual limits to which clinical protocols should adhere. Translational research, applying technical analyses in the context of the medical setting, is often challenging and requires the involvement
of a multidisciplinary team. In this case, the input of certified physicians was essential to the development of the numerical observer model used to predict lesion detectability. The performance of this model was rigorously evaluated within the lesion population, using various combinations of the lesion parameters, to find the characteristics which best allowed it to accurately predict the human observers’ decisions. From these experiments, we found that observer detection was indeed well correlated with lesion SNR, as this metric yielded the smallest residual differences between the human and model observers in the sensitivity curves. Models using only the volume, contrast, and measurement variance of the lesions yielded relatively poorer agreement. The human-trained numerical models showed good predictive utility and may also provide a convenient surrogate for realistic interpretation, in future lesion detection studies.

Many works investigating lesion detection performance use the SNR of a numerical observer, e.g. the CHO, as the figure of merit since it has been shown to be a good predictor of human performance in certain tasks. However, in this work, we chose a discrimination model using a more comprehensive set of lesion parameters, intended to characterize each lesion objectively and subjectively. Channelized observer models use spatial frequencies to reduce dimensionality while extracting the important features of each class, but here we reduce the dimensionality, while retaining important discriminatory characteristics, through 6 lesion parameters. Hence, instead of depicting each lesion by only spatial frequency responses, we thought that the combination of many different properties (including CHO SNR) would better characterize the lesion population and improve discrimination. Although we found the model using lesion SNR to provide the discrimination performance most closely matching that of the original observers, the use of CHO SNR alone provided a close 2nd best. This might not be unexpected since both metrics use similar lesion properties, i.e. lesion mean, contrast, and noise (though the CHO uses channelized versions of these).

The findings of the observer studies pointed to a practical lower limit around 10 million true counts, which is certainly supported by the results of the bias and reproducibility analyses. It was at this level that the sensitivity for detecting lesions matched that in the full count images. Of course the detection limits are set by the smallest lesions, which is clear from Figure 7 where the sensitivity is poor for the small lesions
in the low count levels but is optimized at the same point, in both populations. It is not the intent of this work to propose changes to current clinical protocols, only to investigate the limits of clinical utility, in a highly constrained lesion detection task, under low-count conditions. In this regard, for this system, acquisition, reconstruction, patient population (62.4±13.7 kg), and in this specific task, the 10 million true count level was the lowest limit for accurate image quantification and reliable detection performance; this would be equivalent to a patient injected with 18.5 MBq and scanned with 2 bed positions for 10 minutes each. With this dose and scan time, the effective patient dose from PET would be less than 0.4 mSv.

Ultralow-dose CT protocols combined with iterative reconstruction techniques, which can further reduce the patient radiation dose by 80% (35,36), have produced promising results for the detection of pulmonary nodules (37,38). Hence, this could potentially allow a complete PET/CT lung screening scan to be performed at a total dose under 1 mSv, which is roughly equivalent to 30% of one year of average natural background. Notwithstanding this, it is likely that the best dose for routine PET screening would be higher than the lower limits presented here, as the potential health risks associated with these level of radiation are negligible compared to the factors which already designated the population as high risk for lung cancer (e.g. smoking). In other words, the benefit of improved accuracy from PET would far outweigh the small risk associated with the radiation.

**CONCLUSION**

Image quality was investigated in the context of lesion detectability using objective image metrics and subjective observer models. The analyses in this study suggested that accurate image quantification may be preserved at levels around 2-5 million trues, but detection sensitivity, which is more important for a screening task, was acceptable at trues levels around 10 million. This would mean an effective PET patient dose of less than 0.4 mSv using the acquisition parameters used in this work. Detection sensitivity of the observers was found to be well correlated with lesion SNR.
ACKNOWLEDGEMENTS

This study was funded by the National University Cancer Institute, Singapore Centre Grant Seed Funding Program. One author (RAS) is supported by the National Research Foundation, Singapore and the Singapore Ministry of Education under its Research Centers of Excellence initiative.
REFERENCES


**Figure 1.** Coronal slices containing signal and background regions. An isolated nodule is seen in the CT volume (yellow arrow). The PET analyses of each lesion consisted of 2 segmented signal regions and one background region. The lesion VOI (red) was used to calculate volume and other metrics, and the target cubic sub-volume (blue) was used in the CHO. The background segmentation required only the cubic sub-volume (green) to calculate contrast and noise metrics and for input into the CHO.
Figure 2. Typical example of SUV reproducibility (A) and SNR analyses (B) for one lesion over all count levels. Plot A shows the mean SUV as a function of true counts in the scan for 50 independent noise realizations, represented by the individual colored bars. Plot B shows the calculated lesion SNR at each point. For the higher count levels, when fifty independent realizations were not possible, the actual number of realizations at each point were 26, 17, 13, 8, and 6 for 5, 7.5, 10, 15, and 20 million true counts, respectively. The last point with only one realization corresponds to the full data set.
**Figure 3.** The training of a numerical observer model by human detection responses is shown here (for illustration purposes, the 6D lesion data are represented in the 3D space defined by the principal eigenvectors). The linear discriminant function was defined by maximizing classification SNR within the projected lesion sample points. The discrimination threshold point on the linear function was set so as to best match the overall accuracy between the model and the original observer (grey arrow). This point defined the orthogonal discrimination hyperplane, and the “trained” model was then used to classify a different set of lesion data.
**Figure 4.** Lesion metrics averaged across realizations, as a function of count level: lesion SUV (A), standard deviation in lesion SUV (B), lesion SNR (C), and channelized Hotelling observer SNR (D). The error bars show the standard deviation over all lesion measurements.
Figure 5. SUV measurements at various count levels, relative to the true SUV values as measured in the full-count images: from left to right, $20 \times 10^6$, $10 \times 10^6$, $5 \times 10^6$, $1 \times 10^6$, and $0.25 \times 10^6$ net true counts. Lesion SUV was relatively stable until the count level approached 1 million, in which case the SUV max (row A) demonstrated positive bias and SUV mean (row B) demonstrated slightly negative bias. The unity line is shown in each plot.
Figure 6. Two phenomena cause instability in the PET reconstruction at true count levels around 1 million. First, the scatter fraction estimates based on simulation models are overestimated in noisy conditions (A), which causes negative image bias from over-subtraction. Second, the ordered subset framework can cause negative bias in limited statistics. This bias is reduced when using pure MLEM, i.e. using all projections for each update (B); the OSEM reconstruction used 2 iterations and 21 subsets and MLEM used 40 iterations. Error bars denote the standard deviation across measurements made in 10 independent realizations at each true count level, except for 15 and 20 million which used 8 and 6, respectively.
Figure 7. Predicted detection sensitivity of the prospective data in all 12 lesions (column A) and in just the 8 lesions smaller than 1 cm$^3$ on PET (column B). Predicted performance varied across the 5 human observers, but was maximized for all observers around 10 million counts.
**Table 1. Characteristics of 12 lesions**

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Subject</th>
<th>Age (y)</th>
<th>Histology</th>
<th>TNM stage</th>
<th>Lesion size (cm³ on PET)</th>
<th>Lesion size (cm³ on CT)</th>
<th>Mean SUV*</th>
<th>Max SUV*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>52</td>
<td>Adenocarcinoma</td>
<td>IV</td>
<td>0.52</td>
<td>1.15</td>
<td>3.25</td>
<td>4.33</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>80</td>
<td>Adenocarcinoma</td>
<td>IV</td>
<td>0.94</td>
<td>0.43</td>
<td>1.31</td>
<td>1.79</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>68</td>
<td>SCC</td>
<td>IV</td>
<td>2.72</td>
<td>4.31</td>
<td>8.01</td>
<td>11.17</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>74</td>
<td>Adenocarcinoma</td>
<td>IV</td>
<td>0.18</td>
<td>0.73</td>
<td>6.37</td>
<td>8.29</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
<td>0.15</td>
<td>1.95</td>
<td>2.53</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>71</td>
<td>Adenocarcinoma</td>
<td>IV</td>
<td>1.80</td>
<td>7.15</td>
<td>12.33</td>
<td>16.93</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.73</td>
<td>6.60</td>
<td>11.14</td>
<td>15.69</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>64</td>
<td>Presumed NSCLC†</td>
<td>I</td>
<td>0.53</td>
<td>2.70</td>
<td>6.43</td>
<td>8.63</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>58</td>
<td>SCC</td>
<td>IIIB</td>
<td>3.80</td>
<td>7.00</td>
<td>18.63</td>
<td>26.44</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>68</td>
<td>Adenocarcinoma</td>
<td>IV</td>
<td>0.24</td>
<td>0.18</td>
<td>1.81</td>
<td>2.41</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.52</td>
<td>0.62</td>
<td>1.72</td>
<td>2.35</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.79</td>
<td>0.34</td>
<td>1.42</td>
<td>1.95</td>
</tr>
</tbody>
</table>

* Mean and max SUV as measured in the full-count images
† Serial CT images showed spiculated solitary pulmonary nodule increasing in size; patient declined biopsy or surgery

Multiple lesions within the same subject are grouped in consecutive rows of the same color.
Quantitative Accuracy and Lesion Detectability of Low-Dose FDG-PET for Lung Cancer Screening

Joshua D Schaefferkoetter, Jianhua Yan, Therese Anna Soderlund, David W. Townsend, Maurizio Conti, John Kit Chung Tam, Ross Soo and Ivan Tham

J Nucl Med.
Published online: September 29, 2016.
Doi: 10.2967/jnumed.116.177592

This article and updated information are available at:
http://jnm.snmjournals.org/content/early/2016/09/28/jnumed.116.177592

Information about reproducing figures, tables, or other portions of this article can be found online at:
http://jnm.snmjournals.org/site/misc/permission.xhtml

Information about subscriptions to JNM can be found at:
http://jnm.snmjournals.org/site/subscriptions/online.xhtml

JNM ahead of print articles have been peer reviewed and accepted for publication in JNM. They have not been copyedited, nor have they appeared in a print or online issue of the journal. Once the accepted manuscripts appear in the JNM ahead of print area, they will be prepared for print and online publication, which includes copyediting, typesetting, proofreading, and author review. This process may lead to differences between the accepted version of the manuscript and the final, published version.