Cancer Imaging

CrossMark

A method to assess image quality for Low-dose PET: analysis of SNR, CNR, bias and image noise

Jianhua Yan[1,2,3*], Josh Schaefferkoetter[3,4], Maurizio Conti[5] and David Townsend[3,4]

Abstract

Background: Lowering injected dose will have an effect on PET image quality. In this article, we aim to investigate this effect in terms of signal-to-noise ratio (SNR) in the liver, contrast-to-noise ratio (CNR) in the lesion, bias and ensemble image noise.

Methods: We present here our method and preliminary results using tuberculosis (TB) cases. Sixteen patients who underwent ^{18}F-FDG PET/MR scans covering the whole lung and portion of the liver were selected for the study. Reduced doses were simulated by randomly discarding events in the PET list mode data stream, and ten realizations at each simulated dose were generated and reconstructed. The volumes of interest (VOI) were delineated on the image reconstructed from the original full statistics data for each patient. Four thresholds (20, 40, 60 and 80 % of SUVmax) were used to quantify the effect of the threshold on CNR at the different count level. Image metrics were calculated for each VOI. This experiment allowed us to quantify the loss of SNR and CNR as a function of the counts in the scan, in turn related to dose injected. Reproducibility of mean and maximum standardized uptake value (SUVmean and SUVmax) measurement in the lesions was studied as standard deviation across 10 realizations.

Results: At 5×10^6 counts in the scan, the average SNR in the liver in the observed samples is about 3, and the CNR is reduced to 60 % of the full statistics value. The CNR in the lesion and SNR in the liver decreased with reducing count data. The variation of CNR across the four thresholds does not significantly change until the count level of 5×10^6. After correcting the factor related to subject's weight, the square of the SNR in the liver was found to have a very good linear relationship with detected counts. Some quantitative bias appears with count reduction. At the count level of 5×10^6, bias and noise in terms of SUVmean and SUVmax are up to 10 and 20 %, respectively. To keep both bias and noise less than 10 %, 5×10^6 counts and 20×10^6 counts were required for SUVmean and SUVmax, respectively.

Conclusions: Initial results with the given data of 16 patients diagnosed as TB demonstrated that 5×10^6 counts in the scan could be sufficient to yield good images in terms of SNR, CNR, bias and noise. In the future, more work needs to be done to validate the proposed method with a larger population and lung cancer patient data.

Keywords: Low dose, PET/MR, PET/CT, Lung, Image quality

* Correspondence: jianhua.yan@gmail.com
The original version of this article was revised. The correct name of the author (Josh Schaefferkoetter) has been updated in the authorship list.
[1]Department of Nuclear Medicine, First Hospital of Shanxi Medical University, 85 Jiefang S Rd, Yingze, Taiyuan, Shanxi 030001, China
[2]Molecular Imaging Precision Medicine Collaborative Innovation Center, Shanxi Medical University, 85 Jiefang S Rd, Yingze, Taiyuan, Shanxi 030001, China
Full list of author information is available at the end of the article

Yan *et al. Cancer Imaging* (2016) 16:26

Page 2 of 12

Background

PET is an important tool for in-vivo study of quantitative measurements of physiological, biochemical, or pharmacological processes. In thoracic oncology, ^{18}F-FDG PET currently plays a major role in clinical diagnosis, staging, prognosis and assessment of response to treatment [1]. Recent advances in PET and CT technology have improved image quality while reducing radiation exposure to patients [2, 3], which open new avenues for cancer screening with PET and CT. Several studies showed that low dose CT is superior to traditional chest radiography for lung cancer screening and follow-up by detecting more nodules and lung cancers including early-stage cancers [4, 5]. In addition, low dose CT screening for subjects at high risk could reduce lung cancer mortality [6]. However, due to its limited specificity, low dose CT screening also detected more than 18 % of all lung cancers which were indolent and led to overdiagnosis in screening for lung cancer [7] although computer-aided diagnosis could improve performance of CT screening [8]. A large clinical study on CT screening of patients at risk has shown that even if cancer mortality is reduced by low dose CT screening, however, 24.2 % of the patients were tested positive, but 96.4 % of these were false positives [6]. This large number of false positives calls for imaging techniques with higher specificity, in order to avoid unneeded invasive biopsy. Additional metabolic information from ^{18}F-FDG PET has been shown to be more specific than CT in detecting lung cancer [9]. Moreover, the combination of CT and PET demonstrated better performance in classifying solitary pulmonary nodules as benign or malignant than either PET or CT alone [8]. Thus, the synergetic effect of PET and CT could potentially improve the accuracy of screening for lung cancer [10]. Like low dose CT screening, the radiation exposure due to injected isotope should be minimized without compromising image quality of PET. The effective dose associated with ^{18}F-FDG PET exam in this study was computed based on the reported ICRP values of 0.019 mSv/MBq for a 70 kg adult. For example, the effective dose is about 7 mSv for typical administration of 10 mCi ^{18}F-FDG, which is much higher than that (1.5 mSv) of low dose CT protocol used in the National Lung Screening Trial [11]. The continual improvement of PET imaging, such as introduction of point spread function and time of flight technologies, could allow for lower injected activities while minimizing impact on image quality [12].

A number of studies to investigate the effect of different count levels on PET image quality with phantom have been reported [13, 14]. Our previous study demonstrated count statistics as low as 5×10^6 counts could achieve a fairly high detectability level using a data set of ^{18}F-FDG PET images of tuberculosis (TB) patients acquired on a PET/MR scanner [15]. In this work, we aimed to assess the relationship between numbers of counts in PET scan and image quality with these data, based on image metrics such as liver signal-to-noise ratio (SNR), lesion contrast-to-noise ratio (CNR), bias relative to the "true value", and ensemble noise in the image (lesion and normal tissue).

Methods

Data acquisitions

Sixteen patients with TB (male: 12, female: 4) having a mean age of 45 years (range: [24–67]), a mean weight of 58.53 kg (range: [45–79]) and a mean BMI of 19.78 (range: [15.21–26.70]) underwent ^{18}F-FDG PET/MR scans at A*STAR-NUS, Clinical Imaging Research Center using a Siemens mMR PET/MR scanner. The scanning usually started after the FDG uptake time of 60 min but were subject to the availability of the scanners. In addition, four patients had one PET/CT scan and one PET/MR scan but with one ^{18}F-FDG injection. The interval between the two scans was around 60 min and the order of these two PET scans was random. However, only the PET/MR scan was included in this study. The PET spatial resolution of mMR has been measured, at 1 and 10 cm from the center of the field of view (FOV), as 4.3 mm and 5.2 mm transaxially, and 4.3 mm and 6.6 axially [16]. The coincidence timing window is 5.9 ns and the energy window is 435–650 keV. The detector system includes 8 rings of 56 detector blocks, each comprising an 8×8 matrix of lutetium oxyorthosilicate crystals ($4 \times 4 \times 20$ mm), coupled to an array of 3×3 APDs. The axial PET FOV is 25.2 cm, and all emission data were acquired in 3D mode. The data were organized into separate prompts and delayed sinograms. The delayed events were smoothed and used to estimate the overall random coincidences. All patients fasted at least for 8 h with serum glucose level less than 10 mmol/L. The time difference between injection and acquisition was 80.1 ± 26.2 min (range: [36.5–126.4] min) after injection of 168.6 ± 50.0 MBq (range: [118.0–260.5] MBq) ^{18}F-FDG, with 15 min for one bed position covering the whole lung and portion of the liver. The true coincident events, after random subtraction, were $1.32 \times 10^8 \pm 3.91 \times 10^7$ (range: [8.33×10^7–2.00×10^8]).

Data reconstructions

The reduced doses were simulated by randomly discarding events in each list mode stream according to 10 predefined fractions of original net true counts: 5×10^{-1}, 2.5×10^{-1}, 1.25×10^{-1}, 6.25×10^{-2}, 3.33×10^{-2}, 1.67×10^{-2}, 5×10^{-3}, 3.33×10^{-3}, 1.67×10^{-3}, and 5×10^{-4}. The random events in the list mode data stream were discarded using the same procedure as resampling

Yan *et al. Cancer Imaging* (2016) 16:26

Page 3 of 12

true events. Ten independent realizations at each simulated dose were generated. Each realization was reconstructed with ordinary Poisson ordered subsets expectation maximization (OP-OSEM) with a system point spread function (PSF) incorporated in the projection matrix [17]. Corrections including attenuation, randoms, and scatter were carried out for each realization. Three iterations and 21 subsets were used to produce image matrices of $172 \times 172 \times 127$ with voxel sizes of $4.17 \times 4.17 \times 2.03$ mm. The image volumes were then smoothed with a 5 mm Gaussian filter.

VOI delineation

In this study, the mean standardized uptake value (SUVmean) was calculated as ^{18}F-FDG uptake normalized to injected dose and patient body weight and activity was decay corrected. Volume of interest (VOI) around solitary lung lesions were delineated using fixed threshold set to 40 % of the maximum standardized uptake value (SUVmax) in the lesion and followed by a manual adjustment to exclude neighboring nodes for each VOI if there was any in the VOI. This thresholding was used to investigate the effect of count level on CNR in the lesion, lesion bias with reference to SUV calculated in the images at the full statistical count level, noise across realizations and reproducibility of the SUVmean and SUVmax in the lesion. Twenty small lesions were obtained from these 16 subjects and the volume was 6.38 ± 4.75 ml (range: [1.2–17.58 ml]). The VOIs in the normal lung background and liver background were obtained by drawing spheres with diameter of 3 cm in these two regions. All of the VOIs were delineated on the images reconstructed with the original full statistics data.

Image analysis

For the evaluation of image quality, five metrics were used 1) SNR in the liver, 2) CNR in the lesion, 3) bias with reference to SUV calculated in the images at the full statistics count level, 4) noise in the image expressed as percentage coefficient of variation (COV) in VOIs including lesion, normal lung and liver, across 10 realizations, and 5) error or reproducibility of the SUVmean and SUVmax in the VOIs including lesion, normal lung and liver, across 10 realizations.

1) The signal to noise ratio (SNR) was calculated as the ratio of mean value to standard deviation (SD) in the VOI:

$$Mean(VOI) = \frac{\sum_{j \in VOI} I_{mean}(j)}{N_{VOI}} \tag{1}$$

$$SD(VOI) = \sqrt{\frac{\sum_{j \in VOI}(I_{mean}(j) - Mean(VOI))^2}{N_{VOI}}} \tag{2}$$

$$SNR(VOI) = \frac{Mean(VOI)}{SD(VOI)} \tag{3}$$

where $Mean(VOI)$ and $SD(VOI)$ are mean value and SD value of the VOI in the mean image (I_{mean}) across the realizations, respectively, N_{VOI} is the number of voxels in the VOI. VOI could be in the lesion, clear lung background and liver. The SNR in the liver was widely used to quantify ^{18}F-FDG PET image quality due to its relatively homogeneous uptake. Due to the Poisson statistics of nuclear positron emission, the SNR of PET images depends on the injected activity, the acquisition time, and the attenuation [18, 19], and it can be expressed as follows:

$$SNR_{liver}^2 \approx N \approx K_i \bullet D \bullet t \tag{4}$$

$$K_i = k \bullet \left(g \bullet \eta \bullet \frac{1}{a(m)}\right) \tag{5}$$

where N is the number of detected counts, K_i is a composite sensitivity factor, D is the injected activity and t is the scan time, k is a proportionality constant, g is a noise reduction factor due to reconstruction techniques such as PSF and time-of-flight, η is the scanner sensitivity, and $a(m)$ is the attenuation factor, a function of subject's weight m. The corrected SNR^2 in the liver can be made independent of patient dividing by K_i.

2) The contrast to noise ratio (CNR) is a measure of the signal level in the presence of noise given by [20]:

$$CNR = \frac{Mean(lesion) - Mean(background)}{SD(background)} \tag{6}$$

where background is measured in the neighboring normal lung tissue, $Mean(lesion)$ and $Mean(background)$ are the mean value of the lesion and background region in the mean image of the SUV across the realizations calculated with Eq. 1, respectively and $SD(background)$ is the SD value of the background region in the mean image across the realizations calculated with Eq. 2. The background mask was obtained by performing morphological operation on the lesion mask:

$$background\,mask = dilation(lesion\,mask, 2) - lesion\,mask \tag{7}$$

where $dilation(lesion\,mask, 2)$ means dilating the lesion mask by 2 voxels.

Yan *et al. Cancer Imaging* (2016) 16:26

Page 4 of 12

3) Bias describes the difference of estimated SUV from the true value, which is unknown in the clinic. In this work, the percentage difference relative to the activity uptake at full statistics count level is used to estimate bias.

$$Bias_{frac}(\%) = \frac{(Mean_{frac}(VOI) - Mean_{full}(VOI))}{Mean_{full}(VOI)} * 100 \quad (8)$$

where $Bias_{frac}$ represents the bias in percent at the predefined fraction of counts, $Mean_{frac}$ is the mean value of the VOI in the mean image (I_{mean}^{frac}) across the realizations at the predefined fraction count level. The $Mean_{full}(VOI)$ is the mean value of the VOI in the image at the full count level (I_{mean}^{full}), typically greater than 100×10^6 counts. Both $Mean_{frac}$ and $Mean_{full}(VOI)$ were calculated with Eq.1.

4) The coefficient of variation (COV) is a metric for describing ensemble noise or statistical noise in the image, and it can influence the detectability of the lesion.

$$COV_{frac}(VOI) = \frac{Mean_{mean}^{frac}(VOI)}{Mean_{SD}^{frac}(VOI)} * 100 \quad (9)$$

where $Mean_{mean}^{frac}(VOI)$ and $Mean_{SD}^{frac}(VOI)$ are the mean value of the VOI in the mean image and the SD image across the realizations at the predefined fraction count level calculated with Eq.1, respectively.

5) Standard error (STE) is a metric for describing error on the measurement of SUV, and it can represent the reproducibility of the measurement.

$$STE_{frac}(\%) = \frac{SD_{frac}(Mean_{frac}(VOI))}{Mean_{frac}(VOI)} * 100 \quad (10)$$

where SD_{frac} is the standard deviation (across the realizations) of the SUV in the VOI at the predefined fraction count level, normalized to the SUV averaged across realizations.

6) The effect of different thresholding (20, 40, 60 and 80 % of SUVmax) on CNR at different count level was investigated. In this comparison, the surrounding background mask delineated with 20 % of SUVmax thresholding was applied to the other three thresholding. Each CNR was calculated from the mean image across realizations with Eq. 6. COV of each CNR was calculated across the four thresholding to demonstrate the sensitivity of the CNR to the thresholding.
The bias, COV, and STE were studied as a function of the counts in the scan, in the lung lesions, in the lung background, and in the liver, in the pooled

sample of patients and lesions. For each count level we histogrammed the number of data point that pass a given threshold (in percent), which refers to either bias, COV, or STE for SUVmean and SUVmax. We used the mean of the histogram distribution to define the number of counts associated with such percent average bias (or COV, or STE). We finally plotted counts vs. percent variation for SUVmean or SUVmax bias (or COV, or STE) for TB lesions, background lungs and uniform liver region. This would allow a quantitative assessment of the number of counts needed for a given acceptable error in the measurement, systematic or statistical.
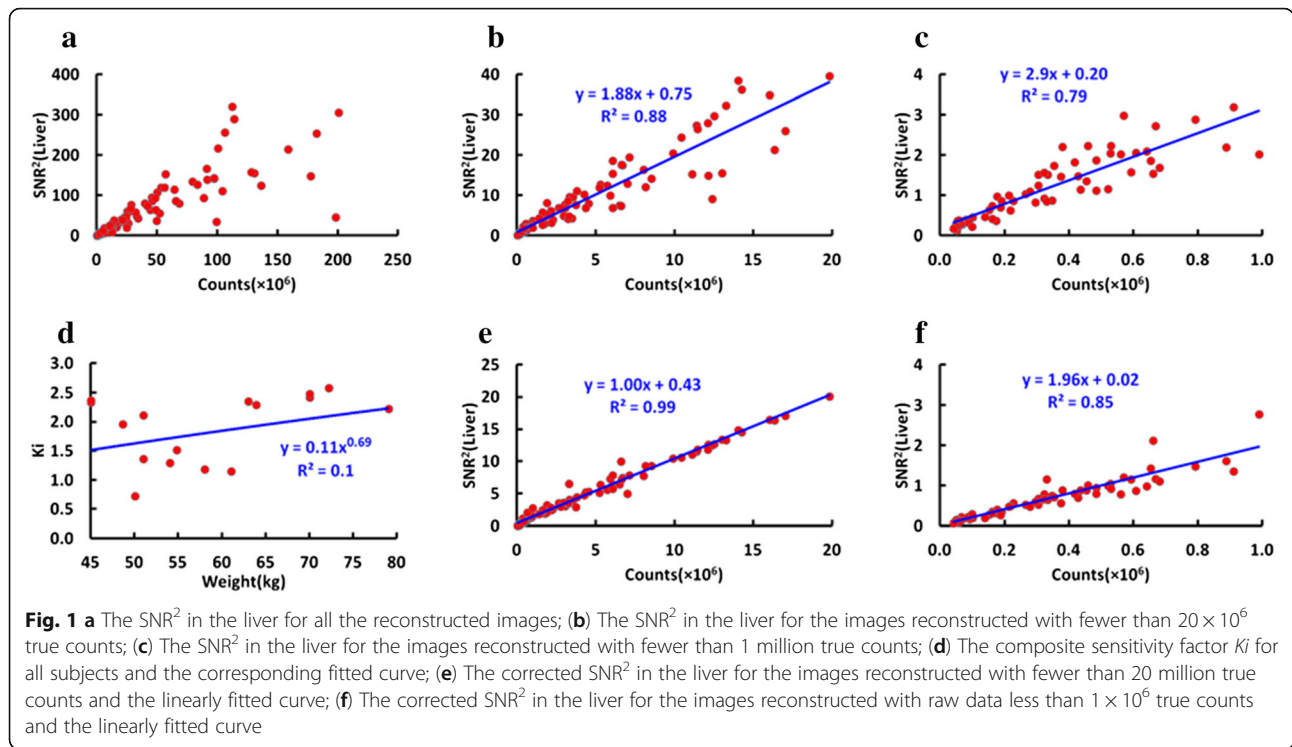
Results

Effect of counts on SNR in the liver

The average SUV_{mean} in the liver for the images reconstructed with the original full statistics data was 1.57 ± 0.40 (range: [0.54–2.12]). The SNR^2 in the liver for the images with full statistics, for the images with fewer than 20×10^6 true counts, and for the images with fewer than 1×10^6 true counts are shown in Fig. 1(a, b and c), respectively. The K_i for each subject was obtained by fitting the SNR^2 with the number of counts ($y = 0.11x^{0.69}$, $R^2 = 0.1$), according to Eq. (4), and is shown in Fig. 1(d). The SNR^2 corrected by the K_i are shown in Fig. 1(e and f) with true counts less than 20×10^6 and 1×10^6, respectively. One can see that the corrected SNR^2 has better linear relationship with detected counts than the original SNR^2.

In PET imaging, photons registered in the detectors follow a Poisson distribution and the standard deviation on the number of counts N is proportional to the square root of N, therefore, SNR^2 of the counts registered in the detectors is linear with N. Iterative reconstruction method based on maximum likelihood estimate is the most popular method in the most of PET scanner, which is a nonlinear method and leads to the non-linear relationship of the SNR^2 measured with the reconstructed image with N. However, in the case of high statistics, the nonlinear transformation can be approximate to a linear one. As expected, the proportionality coefficient in the mid-range of $1–20 \times 10^6$ counts is 1 (Fig. 1(e)), but at extremely low counts such proportionality does not hold anymore, and applying a linear fit the coefficient is now 1.96 (Fig. 1(f)). This represents a loss of proportionality at low counts, when Poisson statistics does not represent the data well anymore, and the relationship between counts and SNR^2 is not linear.

Effect of counts on CNR in the lesion

Twenty six lesions were delineated by 40 % of SUV_{max} thresholding method and the lesions with volume less

Fig. 1 a The SNR^2 in the liver for all the reconstructed images; (**b**) The SNR^2 in the liver for the images reconstructed with fewer than 20×10^6 true counts; (**c**) The SNR^2 in the liver for the images reconstructed with fewer than 1 million true counts; (**d**) The composite sensitivity factor Ki for all subjects and the corresponding fitted curve; (**e**) The corrected SNR^2 in the liver for the images reconstructed with fewer than 20 million true counts and the linearly fitted curve; (**f**) The corrected SNR^2 in the liver for the images reconstructed with raw data less than 1×10^6 true counts and the linearly fitted curve

than 20 ml, corresponding to approximate 3.37 cm of diameter for spherical lesion, were found in these 16 subjects with average SUV_{mean} of 1.92 ± 0.95 (range: [0.92–4.81]). Typical PET images with five realizations with a lesion of CNR of 3.17 and volume of 1.66 ml at the different count level are shown in Fig. 2. As expected, image quality decreases with decreasing counts. The lesion can be visually detected in the image at the count level of more than 3×10^6. Since it is challenging to reliably identify lesion with low CNR, especially at low count level, lesions with CNR less than 2 were excluded from this study. Twenty lesions were found to satisfy the requirement of volume ($\leq$20 ml) and CNR ($\geq$2).

Figure 3(a) shows CNR of all lesions at different count levels and Fig. 3(b) displays CNR at the count level range of $0–20 \times 10^6$. CNR decreases with the lowering of count. All CNR were normalized to the value at the full count level and are shown in Fig. 3(c and d). Normalized CNR can be fitted with count by a function ($y = 1/(1 + 2.41x^{-0.8})$, $R^2 = 0.91$). This curve allows to predict the amount of counts required to obtain a desired normalized CNR.

Effect of thresholding on CNR

Figure 4(a) shows the mean CNR of all lesions at different count level using different thresholding method. It can be seen that higher thresholding method produces higher CNR. In addition, all lesions' CNR decrease with

lowering counts. Figure 4(b) demonstrates the mean COV of all lesions' CNR across the realizations, which indicates that the mean COV do not significantly vary until a count level of 5×10^6.
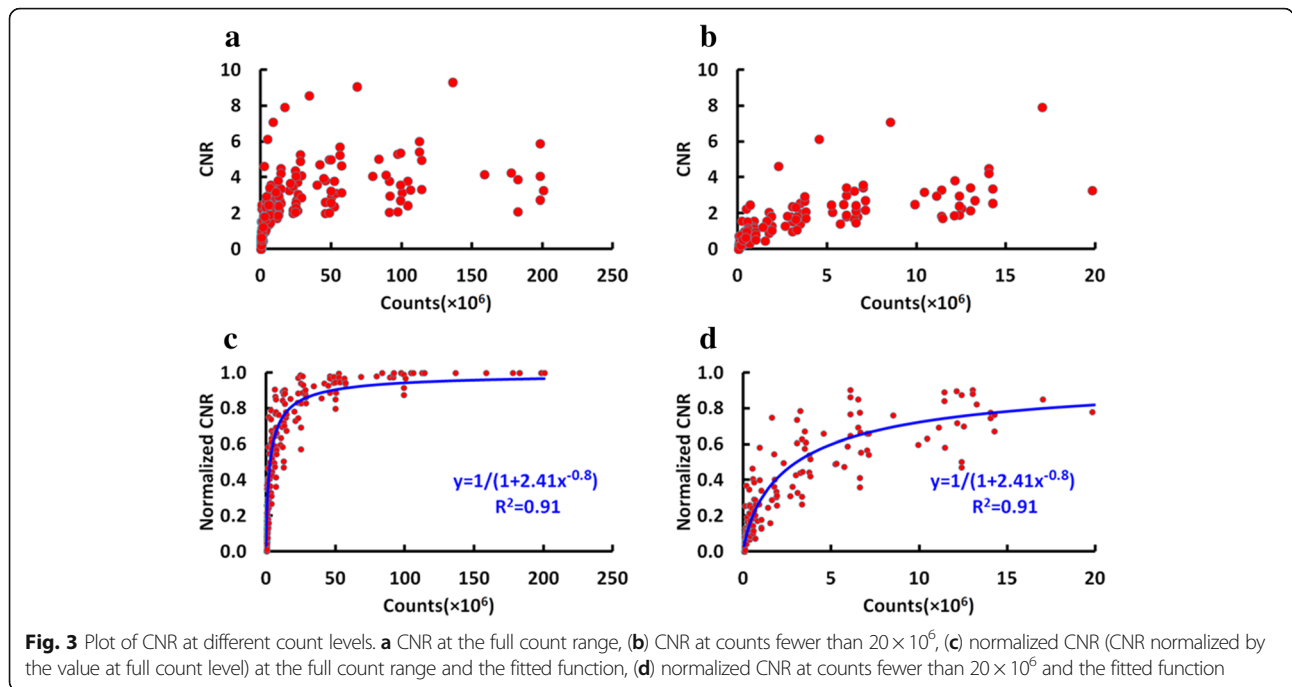
Effect of counts on bias and COV

Figure 5(a and b) shows the mean SUV_{mean} and SUV-max in the normal lung, liver and lesion, respectively. Figure 5(c) exhibits the CNR in the same lesion at variable count levels for the same patient. One can see that the SUV_{mean} in these three regions for this subject do not significantly change, relative to the full count level, down to a count level of 1×10^6. For the SUVmax, the count level is around 8×10^6.

The variation of the SUV_{mean} and SUVmax over the realizations increases with reducing counts. All patients and lesions were pooled and analyzed in order to study relationship between counts in the scan and two key parameters: measurement bias and error. In Fig. 6, an example of histogram representing the frequency of cases with COV over 10 % at given counts in the scan, in different regions based on SUVmean and SUVmax. Each bin represents the sum of patients (Fig. 6(a, d, b and e)) or lesions (Fig. 6(c and f)) at each count level with COV larger than 10 %. Histograms were created for variable error levels, and the mean and maximum value of the distribution was computed and plotted in Fig. 7(a and d). Similarly, in Fig. 7(b and e), the STE on SUV_{mean} and SUV_{max} is plotted vs. counts level. In Fig. 7(c and f), the

Yan *et al. Cancer Imaging* (2016) 16:26

Page 6 of 12

Fig. 2 Typical PET images (SUV: 0–6) with a lesion of CNR (=3.17) and volume (1.66 ml) at the different count level (5 realizations are displayed). Each column represents one realization. The count level corresponding to each row is 0.048×10^6, 0.16×10^6, 0.32×10^6, 0.48×10^6, 1.6×10^6, 3.2×10^6, 6×10^6, 1.2×10^7, 2.4×10^7, 4.8×10^7

Yan *et al. Cancer Imaging* (2016) 16:26

Page 7 of 12

Fig. 3 Plot of CNR at different count levels. **a** CNR at the full count range, (**b**) CNR at counts fewer than 20×10^6, (**c**) normalized CNR (CNR normalized by the value at full count level) at the full count range and the fitted function, (**d**) normalized CNR at counts fewer than 20×10^6 and the fitted function
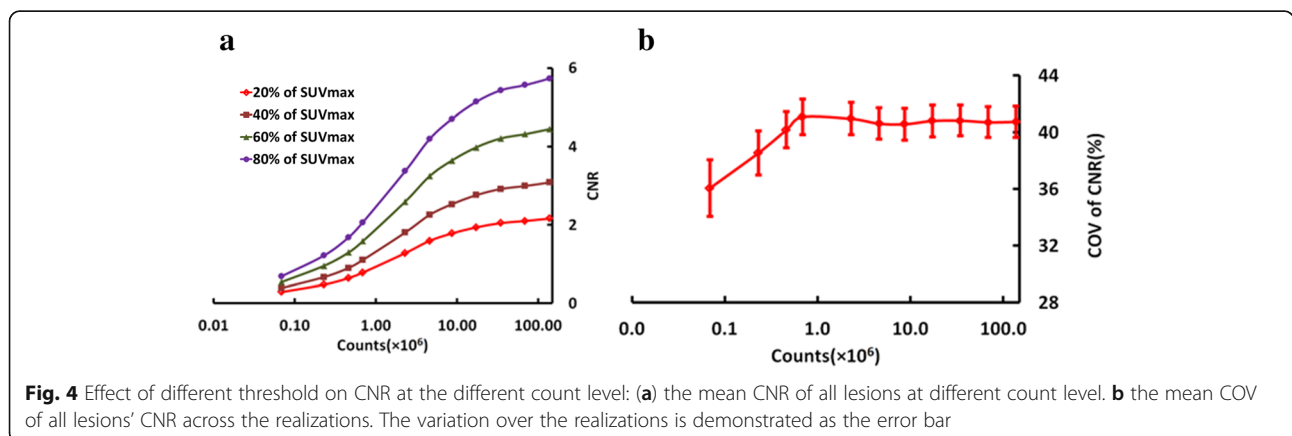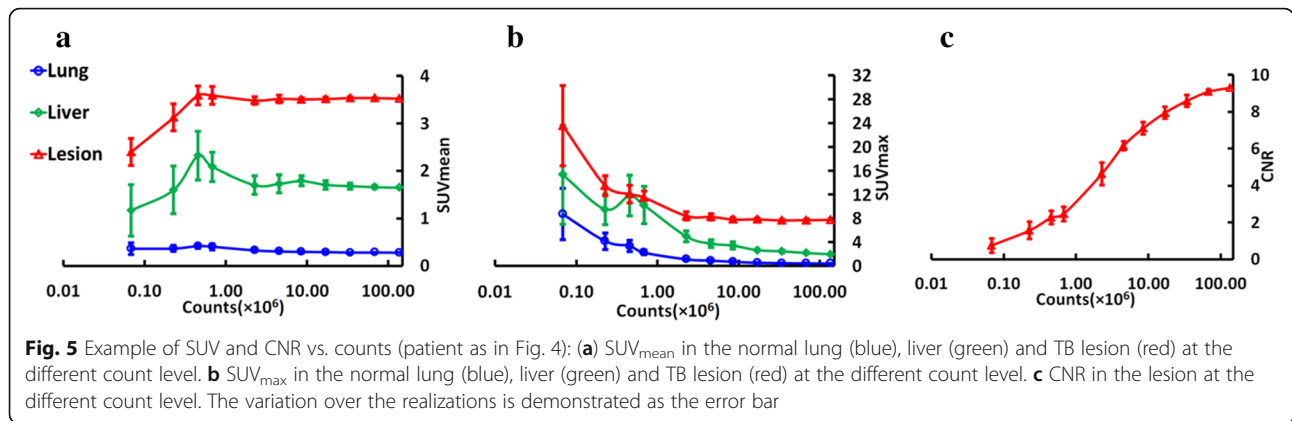
measurement bias, defined as difference relative to the full statistics SUV_{mean} and SUV_{max}, is plotted vs. counts level.

From Fig. 7(a), the counts required to keep the COV for SUVmean below 10 % for half of the cases in the sample are about 25×10^6, 25×10^6 and 5×10^6, respectively for the liver, lung, and TB lesions. The corresponding counts to keep the COV for SUVmax less than 10 % for half of the cases are about 56×10^6, 52×10^6 and 20×10^6, respectively for the liver, lung, and TB lesions from Fig. 7(d). From Fig. 7(b and e), only 1×10^6, 2.7×10^6 and 0.4×10^6 counts for SUVmean and 6.6×10^6, 8.5×10^6 and 2.8×10^6 counts for SUVmax are needed to reach the STE of 10 % for the liver, normal lung and lesion, respectively. From Fig. 7(c and f), only 2.2×10^6, 1×10^6 and 0.2×10^6 counts for SUVmean and 28.5×10^6, 24.2×10^6 and 5.2×10^6 counts for SUVmax are

needed to reach the percentage bias of 10 % for the liver, normal lung and lesion, respectively. At 5×10^6 counts in the scan, bias and noise in terms of SUVmean and SUVmax are up to 10 and 20 %, respectively.

The 20 lesions were categorized into two groups based on volume threshold of 5 ml and each group had 10 lesions. One subgroup has mean volume 2.71 ml [1.20–4.81 ml] and another subgroup has mean volume 9.54 ml [5.48–17.58 ml], approximately 17.3 mm [13.2 mm-21 mm] and 26.3 mm [22 mm-32.2 mm] in diameter, respectively. The counts needed to reach the same bias, STE and COV percentage for large lesions (volume ≥ 5 ml) are fewer than that for small lesions (volume ≤ 5 ml) for SUVmean, although the difference between the two groups is small. In addition, the bias and STE for SUVmax were the same for the two groups, as shown in Fig. 8.
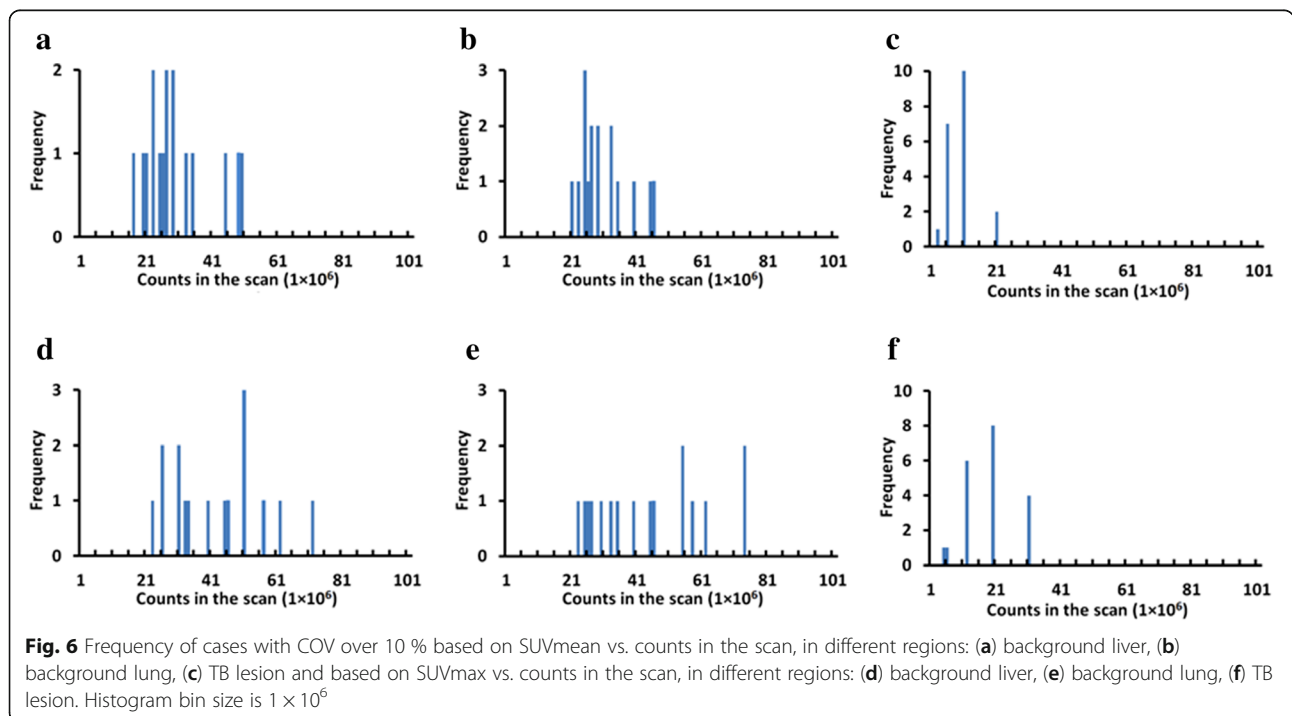

Fig. 4 Effect of different threshold on CNR at the different count level: (**a**) the mean CNR of all lesions at different count level. **b** the mean COV of all lesions' CNR across the realizations. The variation over the realizations is demonstrated as the error bar

Yan *et al. Cancer Imaging* (2016) 16:26

Page 8 of 12

Fig. 5 Example of SUV and CNR vs. counts (patient as in Fig. 4): (**a**) SUV$_{mean}$ in the normal lung (blue), liver (green) and TB lesion (red) at the different count level. **b** SUV$_{max}$ in the normal lung (blue), liver (green) and TB lesion (red) at the different count level. **c** CNR in the lesion at the different count level. The variation over the realizations is demonstrated as the error bar
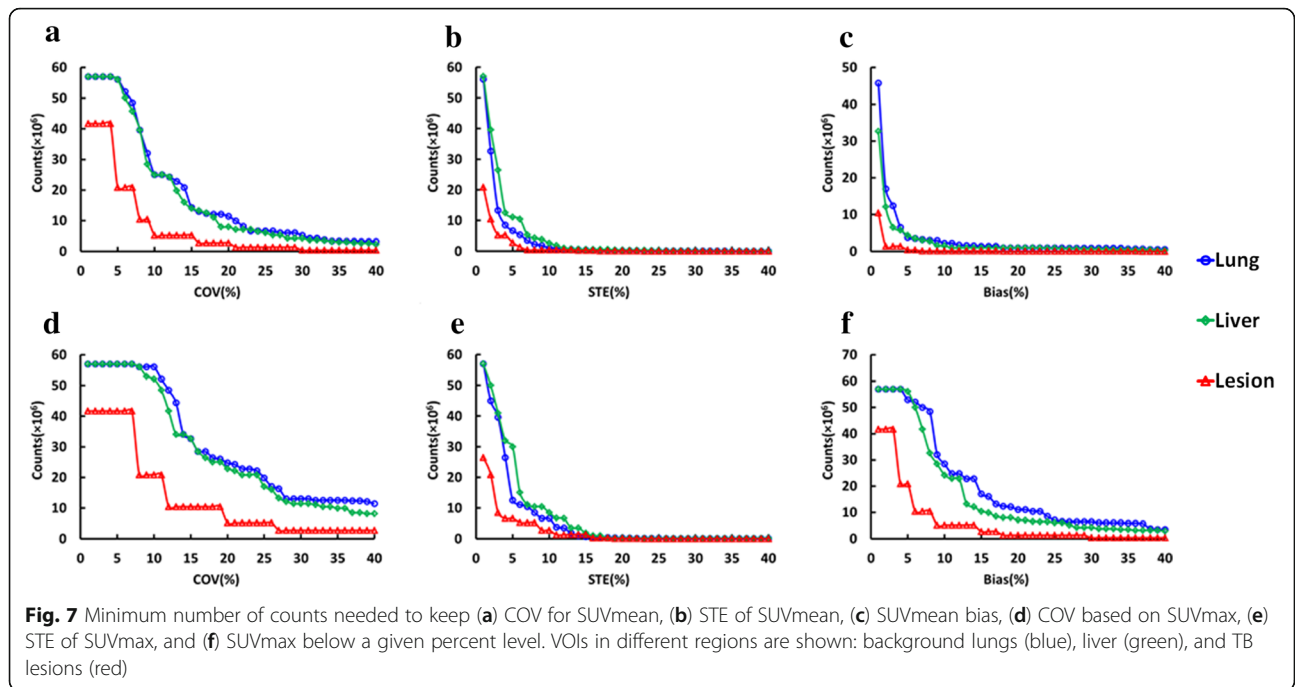
Discussion

We evaluated image quality with objective metrics including SNR in the liver, CNR in the lesions, bias and noise in the liver, normal lung and lesions, at simulated reduced doses, using ^{18}F-FDG PET data at various count levels from TB patients. The underlying biology of TB is different from that of lung cancer, but for a technical study such as this, the uptake levels in TB lesions will be more representative of early stage lung cancer lesions than those in more advanced lung cancer. This work will lay the foundation to determine the appropriate dose or scan time for a future prospective study with lung cancer patients PET/CT scanning.

Accurate delineation of lesions is a prerequisite for quantification of FDG uptake. Although a large number of approaches have been proposed to segment tumors in PET images including threshold based, gradient based [21], and fuzzy Bayesian based methods [22], accurate tumor segmentation is still a challenging task. This is due to limited spatial resolution and the relatively high noise level in PET images, and this process evidently becomes more challenging with fewer counts (Zaidi and El Naqa, [23]). A simple thresholding method was employed here to segment the solitary lesions in the lung using the full statistics images and the resulting VOIs were copied to the images at the lower count levels. This simple thresholding method may lead
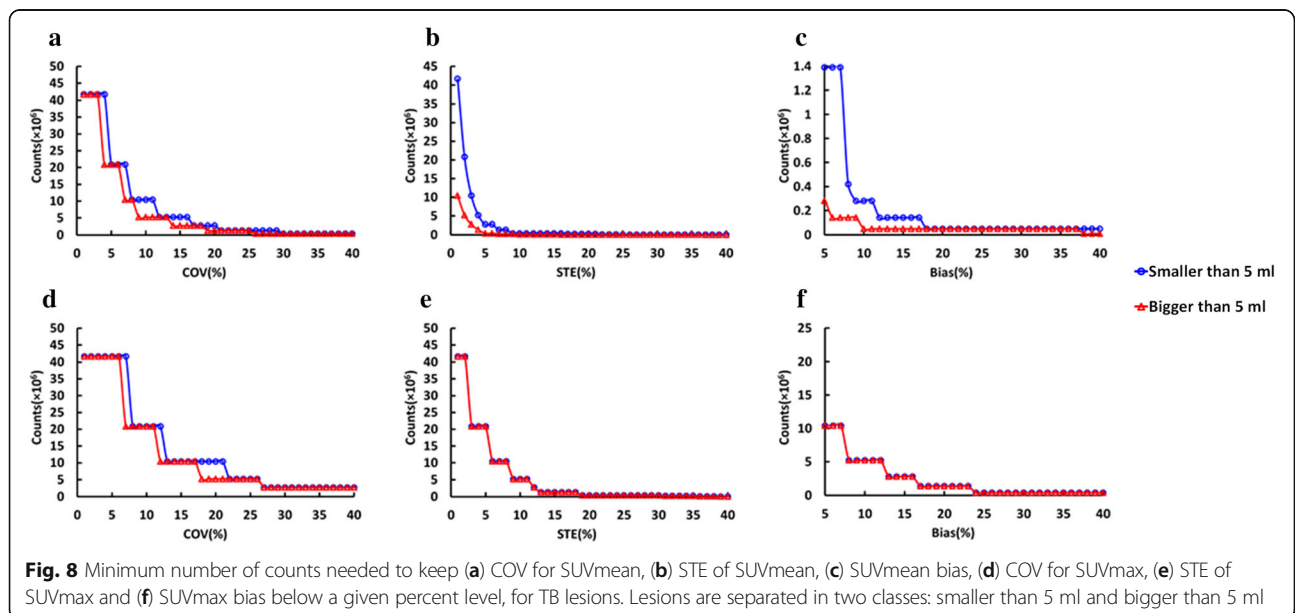

Fig. 6 Frequency of cases with COV over 10 % based on SUVmean vs. counts in the scan, in different regions: (**a**) background liver, (**b**) background lung, (**c**) TB lesion and based on SUVmax vs. counts in the scan, in different regions: (**d**) background liver, (**e**) background lung, (**f**) TB lesion. Histogram bin size is 1×10^6

Yan *et al. Cancer Imaging* (2016) 16:26

Page 9 of 12

Fig. 7 Minimum number of counts needed to keep (**a**) COV for SUVmean, (**b**) STE of SUVmean, (**c**) SUVmean bias, (**d**) COV based on SUVmax, (**e**) STE of SUVmax, and (**f**) SUVmax below a given percent level. VOIs in different regions are shown: background lungs (blue), liver (green), and TB lesions (red)

to imperfect delineation of the tumor. In addition, the spill-out from the tumor to the surrounding background can lead to lower CNR. However, the inaccurate delineation will not change the behavior of the image metrics since the error will have the same effect at the different count levels, which is partially supported by the result of CNR varying with different threshold (20, 40, 60 and 80 % of SUVmax). Since we work towards low dose PET imaging for those patients at high risk who have indefinite findings with low dose CT

screening, VOIs can be delineated on the CT image and copied onto registered PET images.

Figure 1(a) is the plot of SNR^2 in the liver for all patients at all count levels and Fig. 1(d) shows the K_i for all patients derived from the SNR^2 in the liver for each patient at all count level. The SNR in the liver depends on many factors including scanner sensitivity, administered dose, scan time and patient-dependent parameter such as weight. In this study, after adjusting the effect of patient's weight, the SNR^2 was found to have a very good

Fig. 8 Minimum number of counts needed to keep (**a**) COV for SUVmean, (**b**) STE of SUVmean, (**c**) SUVmean bias, (**d**) COV for SUVmax, (**e**) STE of SUVmax and (**f**) SUVmax bias below a given percent level, for TB lesions. Lesions are separated in two classes: smaller than 5 ml and bigger than 5 ml

Yan *et al. Cancer Imaging* (2016) 16:26

Page 10 of 12

linear relationship with detected counts ($y = 1.00 \times + 0.43$, $R^2 = 0.99$), which fits with Poisson statistics of nuclear positron emission [18, 19]. At very low counts, the Poisson statistics approximation does not hold anymore, and SNR^2 acquires a sharper decrease as shown in Fig. 1(b and c), and in Fig. 1(e and f). Since the SNR^2 in the liver will become nonlinear at very low count levels, the slope of the fit for the count range of $0–1 \times 10^6$ will be different from that of the fit at the count range of $0–20 \times 10^6$. As expected, also the CNR in the lesion decreases with decreasing counts. After being normalized to CNR at full count data, a function can be used to predict CNR at different count level. A heuristic curve that best fits the CNR was used. In this phase, no interpretative model is proposed for the behavior of the curve, but we applied the simplest function that could fit the experimental data and allow for the prediction of CNR at low counts.

A consequence of dose or count reduction is a possible bias in SUVmean or SUVmax measurement, and a larger error in the measurement, or degradation of measurement reliability, and an increase of noise in the image, which affects detectability of small lesions. This effect has been studied and a bias has been observed, as well as an increase of noise as COV, and an increase of STE of the SUVmean and SUVmax measurement in different regions of the patient. Several factors influence this including a positive bias in the cold background and negative bias in the hot regions associated with the positivity constraint of the OSEM reconstruction for SUVmean. The SUVmax is easily impacted by count reduction than SUVmean. In Fig. 5, one can observe larger error bars and SUV instability at very low counts.

As shown in Fig. 7, fewer counts in the scan correspond to higher COV, STE or error of the radioactivity measurement for both SUVmean and SUVmax, as well as a higher bias relative to high count rate. In comparison with SUVmean, more counts are needed to keep the same acceptable level for COV, bias and STE for SUVmax. In terms of counts required to obtain the same percent error, fewer counts are needed for lesions than for liver and lung background. For example, if the acceptable level for COV for SUVmean is 10 %, the corresponding required counts are 5×10^6, 25×10^6, and 25×10^6, respectively for TB lesions, liver and normal lungs (Fig. 7(a)). This can be explained by the inherent "real" local variations of uptake values in large VOIs in lungs and liver, which are clearly not uniform. In addition, the local variations can explain the different appearance of histograms in the Fig. 6. The number of counts to maintain the same COV for SUVmax are 20×10^6, 56×10^6 and 52×10^6 for TB lesions, liver and normal lungs (Fig. 7(d)). The actual variations add to the statistical noise. Apparently, bias for SUVmean is less

sensitive to count statistics, and minimum bias can be reached even below 1×10^6 counts. In addition, the same acceptable level for STE regarding to SUVmean can be maintained with fewer counts as compared to COV, which is demonstrated by 0.4×10^6, 2.7×10^6 and 1×10^6 counts required to reach STE threshold of 10 % for TB lesions, liver and normal lungs, respectively (Fig. 7(b)).

Preliminary investigations show that the behavior of SUVmax is similar, but SUVmax is much more sensitive to noise, and comparable levels can be reached only with much higher number of counts in the scan. An additional study was done by splitting the lesion group into two subgroups: small lesions, with volume smaller than 5 ml; and large lesions, with volumes greater than 5 ml. Each subgroup has 10 lesions. As demonstrated in Fig. 7, for a given number of counts in the scan, bias, COV and STE are larger for smaller volume lesions. Regarding to SUVmean, choosing a target number of counts of about 5×10^6 counts, a noise level (COV) of 9 % is obtained for large lesions, and of 12 % for small lesions. A STE of 4 and 2 % are reached at 5×10^6 counts for small lesions and large lesions, respectively. Finally, the bias at 5×10^6 counts for both large lesions and small lesions are less than 2 %. At the same count level of 5×10^6 counts, lesion CNR is about 60 % of value at the full statistics level (data in Fig. 3(d)) and SNR in the liver is about 3 (data from Fig. 1(b)).

The independent realizations at different low dose were obtained by randomly discarding the events in the list mode data stream, based on the desired count level and thus these realizations are not fully independent. In addition, we got very similar results with bootstrap resampling [24, 25], which was considered as a better method to produce independent realizations. For example, the fitted function for the SNR^2 in the liver for the images reconstructed with fewer than 1 million true counts $y = 2.85x + 0.19$, $R^2 = 0.81$, which is close to that function ($y = 2.9x + 0.20$, $R^2 = 0.79$) with the current simulation strategy. This could be due to the fact that most of the image metrics in this work were based on the mean value across these realizations. Therefore, in order to keep in line with our previous study [15], the results with the simulations by randomly discarding the events in the list mode data stream were presented in this work.

In the earlier works by Budinger, TF, et al. [26], Hoffman, EJ, et al. [27] and Strother, SC, et al. [28], the relationship between image counts and noise (root-mean-square) had been investigated. These earlier works used a different reconstruction scheme and our findings cannot be compared directly. Notwithstanding this, the statistical noise (COV) in the liver for each subject at the different count level in this study was close to the root-mean square calculated the equation found in the earlier studies up to the count level of 5 million.

Yan *et al. Cancer Imaging* (2016) 16:26

Page 11 of 12

In this work, we chose OSEM reconstruction with 3 iterations, 21 subsets and post reconstruction Gaussian smoothing with FWHM of 5 mm which are commonly used parameters in the clinical settings. In future, we will investigate the optimization of the reconstruction parameters with prospective lung cancer patient data once the impact on image quality of reducing counts is generally understood. In addition, the impact of inaccurate attenuation map on quantitative PET lung imaging due to respiratory motion will be thoroughly explored with the data at different counts levels.

Conclusions

We developed a method and the tools for objectively evaluating PET images at various count levels with ^{18}F-FDG PET data of TB patients acquired on a combined PET/MR scanner. This work allowed us to quantify the loss of SNR and CNR as a function of the counts in the scan, in turn related to dose injected. At 5×10^6 counts in the scan, the average SNR in the liver in the observed samples is about 3, and the CNR is reduced to 60 % of value at the full statistics level. At the count level of 5×10^6, bias and noise in terms of SUVmean and SUVmax are up to 10 and 20 %, respectively. This initial investigation presents the first step in a comprehensive image analysis of low dose PET imaging and lays the foundation for future study on low dose PET imaging for lung cancer, which will extend analysis to a larger and lung cancer patient sample, and assess if image quality at reduced counts is sufficient for lung cancer screening.

Abbreviations

CNR: Contrast-to-noise ratio; COV: Coefficient of variation; OP-OSEM: Ordinary poisson ordered subsets expectation maximization; PSF: Point spread function; SD: Standard deviation; SNR: Signal-to-noise ratio; SUV: Standardized uptake value; SUVmaxm: Maximum SUV; SUVmean: Mean SUV; TB: Tuberculosis; VOI: Volumes of interest

Availability of data and materials

None.

Authors' contributions

All authors read and approved the final manuscript. Study concept and design: JY, JS and MC. Acquisition of data: JY and JS. Analysis and interpretation of data: JY. Drafting of the manuscript: JY and JS. Critical revision of the manuscript: JY, JS, MC and DT.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

All authors read the final manuscript and consent for publication.

Ethics approval and consent to participate

This study was approved by the Domain Specific Review Board and Institutional Review Board of National Healthcare Group, Singapore and written consent form was obtained from each patient.

Author details

[1]Department of Nuclear Medicine, First Hospital of Shanxi Medical University, 85 Jiefang S Rd, Yingze, Taiyuan, Shanxi 030001, China. [2]Molecular Imaging Precision Medicine Collaborative Innovation Center, Shanxi Medical University, 85 Jiefang S Rd, Yingze, Taiyuan, Shanxi 030001, China. [3]A*STAR-NUS, Clinical Imaging Research Center, Center for translational medicine, 14 medical drive, #B1-01, 17599 Singapore, Singapore. [4]Department of Diagnostic Radiology, National University Hospital, Main Building, 5 Lower Kent Ridge Road, Level 3, 119074 Singapore, Singapore. [5]Siemens Healthcare Molecular Imaging, 810 Innovation Drive, Knoxville TN37932, USA.

References

1. Ambrosini V, Nicolini S, Caroli P, et al. PET/CT imaging in different types of lung cancer: an overview. Eur J Radiol. 2012;81:988–1001.
2. Willowson KP, Bailey EA, Bailey DL. A retrospective evaluation of radiation dose associated with low dose FDG protocols in whole-body PET/CT. Australas Phys Eng Sci Med. 2012;35:49–53.
3. Yanagawa M, Gyobu T, Leung AN, et al. Ultra-low-dose CT of the lung: effect of iterative reconstruction techniques on image quality. Acad Radiol. 2014;21:695–703.
4. Bach PB, Mirkin JN, Oliver TK, et al. Benefits and harms of CT screening for lung cancer: a systematic review. JAMA. 2012;307:2418–29.
5. Hanna WC, Paul NS, Darling GE, et al. Minimal-dose computed tomography is superior to chest x-ray for the follow-up and treatment of patients with resected lung cancer. J Thorac Cardiovasc Surg. 2014;147:30–3.
6. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2012;365:395–409.
7. Patz Jr EF, Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med. 2014;174:269–74.
8. Kim SK, Allen-Auerbach M, Goldin J, et al. Accuracy of PET/CT in characterization of solitary pulmonary lesions. J Nucl Med. 2007;48:214–20.
9. Schrevens L, Lorent N, Dooms C, Vansteenkiste J. The role of PET scan in diagnosis, staging, and management of non-small cell lung cancer. Oncologist. 2004;9:633–43.
10. Veronesi G, Bellomi M, Veronesi U, et al. Role of positron emission tomography scanning in the management of lung nodules detected at baseline computed tomography screening. Ann Thorac Surg. 2007;84:959–65. discussion 965–956.
11. Kramer BS, Berg CD, Aberle DR, Prorok PC. Lung cancer screening with low-dose helical CT: results from the National Lung Screening Trial (NLST). J Med Screen. 2011;18:109–11.
12. Conti M. Focus on time-of-flight PET: the benefits of improved time resolution. Eur J Nucl Med Mol Imaging. 2011;38:1147–57.
13. Doot RK, Scheuermann JS, Christian PE, Karp JS, Kinahan PE. Instrumentation factors affecting variance and bias of quantifying tracer uptake with PET/CT. Med Phys. 2010;37:6035–46.
14. Oehmigen M, Ziegler S, Jakoby BW, Georgi JC, Paulus DH, Quick HH. Radiotracer Dose Reduction in Integrated PET/MR: Implications from National Electrical Manufacturers Association Phantom Studies. J Nucl Med. 2014;55:1361–7.
15. Schaefferkoetter JD, Yan J, Townsend DW, Conti M. Initial assessment of image quality for low-dose PET: evaluation of lesion detectability. Phys Med Biol. 2015;60:5543–56.
16. Delso G, Furst S, Jakoby B, et al. Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner. J Nucl Med. 2011;52:1914–22.
17. 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.

18. Watson CC, Casey ME, Bendriem B, et al. Optimizing injected dose in clinical PET by accurately modeling the counting-rate response functions specific to individual patient scans. J Nucl Med. 2005;46:1825–34.
19. de Groot EH, Post N, Boellaard R, Wagenaar NR, Willemsen AT, van Dalen JA. Optimized dose regimen for whole-body FDG-PET imaging. EJNMMI Res. 2013;3:63.
20. Le Meunier L, Slomka PJ, Dey D, et al. Enhanced definition PET for cardiac imaging. J Nucl Cardiol. 2010;17:414–26.
21. Geets X, Lee JA, Bol A, Lonneux M, Gregoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur J Nucl Med Mol Imaging. 2007;34:1427–38.
22. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging. 2009;28:881–93.
23. Zaidi H, El Naqa I. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. Eur J Nucl Med Mol Imaging. 2010;37:2165–87.
24. Lartizien C, Aubin JB, Buvat I. Comparison of bootstrap resampling methods for 3-D PET imaging. IEEE Trans Med Imaging. 2010;29:1442–54.
25. Markiewicz PJ, Reader AJ, Matthews JC. Assessment of bootstrap resampling performance for PET data. Phys Med Biol. 2015;60:279–99.
26. Budinger TF, Derenzo SE, Greenberg WL, Gullberg GT, Huesman RH. Quantitative potentials of dynamic emission computed tomography. J Nucl Med. 1978;19:309–15.
27. Hoffman EJ, Huang SC, Phelps ME. Quantitation in positron emission computed tomography: 1. Effect of object size. J Comput Assist Tomogr. 1979;3:299–308.
28. Strother SC, Casey ME, Hoffman EJ. Measuring PET scanner sensitivity: relating counrates to image signal-to-noise ratios using noise equivalent counts. IEEE Trans Nucl Sci. 1990;37:783–8.