

PAPER

## Generalized PSF modeling for optimized quantitation in PET imaging

To cite this article: Saeed Ashrafinia *et al* 2017 *Phys. Med. Biol.* **62** 5149

View the [article online](#) for updates and enhancements.

### Related content

- [Single scan parameterization of space-variant point spread functions](#)  
F A Kotasidis, J C Matthews, G I Angelis et al.
- [Noise and signal properties in PSF-based fully 3D PET image reconstruction](#)  
S Tong, A M Alessio and P E Kinahan
- [Non-Gaussian space-variant resolution modelling for list-mode reconstruction](#)  
C Cloquet, F C Sureau, M Defrise et al.



 **myQA**<sup>®</sup>

### INTEGRATED QUALITY ASSURANCE

myQA provides one central platform and database integrating all your QA applications.

All-in-One. All Connected. All Secure.



ibar-dosimetry.com

# Generalized PSF modeling for optimized quantitation in PET imaging

Saeed Ashrafinia<sup>1,2</sup>, Hassan Mohy-ud-Din<sup>3</sup>,  
Nicolas A Karakatsanis<sup>4</sup>, Abhinav K Jha<sup>2</sup>, Michael E Casey<sup>5</sup>,  
Dan J Kadrmaz<sup>6</sup> and Arman Rahmim<sup>1,2</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States of America

<sup>2</sup> Department of Radiology and Radiological Sciences, School of Medicine, Johns Hopkins University, Baltimore, MD, United States of America

<sup>3</sup> Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, United States of America

<sup>4</sup> Translational and Molecular Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America

<sup>5</sup> Siemens Medical Systems, Knoxville, TN, United States of America

<sup>6</sup> Department of Radiology, University of Utah, Salt Lake City, UT, United States of America

E-mail: [ashrafinia@jhu.edu](mailto:ashrafinia@jhu.edu) and [rahmim1@jhmi.edu](mailto:rahmim1@jhmi.edu)

Received 7 September 2016, revised 11 February 2017

Accepted for publication 24 March 2017

Published 26 May 2017



CrossMark

## Abstract

Point-spread function (PSF) modeling offers the ability to account for resolution degrading phenomena within the PET image generation framework. PSF modeling improves resolution and enhances contrast, but at the same time significantly alters image noise properties and induces edge overshoot effect. Thus, studying the effect of PSF modeling on quantitation task performance can be very important. Frameworks explored in the past involved a dichotomy of PSF versus no-PSF modeling. By contrast, the present work focuses on quantitative performance evaluation of standard uptake value (SUV) PET images, while incorporating a wide spectrum of PSF models, including those that under- and over-estimate the true PSF, for the potential of enhanced quantitation of SUVs. The developed framework first analytically models the true PSF, considering a range of resolution degradation phenomena (including photon non-collinearity, inter-crystal penetration and scattering) as present in data acquisitions with modern commercial PET systems. In the context of oncologic liver FDG PET imaging, we generated 200 noisy datasets per image-set (with clinically realistic noise levels) using an XCAT anthropomorphic phantom with liver tumours of varying sizes. These were subsequently reconstructed using the OS-EM algorithm with

varying PSF modelled kernels. We focused on quantitation of both  $SUV_{\text{mean}}$  and  $SUV_{\text{max}}$ , including assessment of contrast recovery coefficients, as well as noise-bias characteristics (including both image roughness and coefficient of-variability), for different tumours/iterations/PSF kernels. It was observed that overestimated PSF yielded more accurate contrast recovery for a range of tumours, and typically improved quantitative performance. For a clinically reasonable number of iterations, edge enhancement due to PSF modeling (especially due to over-estimated PSF) was in fact seen to lower  $SUV_{\text{mean}}$  bias in small tumours. Overall, the results indicate that exactly matched PSF modeling does not offer optimized PET quantitation, and that PSF overestimation may provide enhanced SUV quantitation. Furthermore, generalized PSF modeling may provide a valuable approach for quantitative tasks such as treatment-response assessment and prognostication.

**Keywords:** PET, quantitation, standard uptake value (SUV), PSF modeling, partial volume correction

(Some figures may appear in colour only in the online journal)

## 1. Introduction

PET imaging is affected by a number of resolution degrading factors that result in undesirable cross-contamination between adjacent voxel regions, referred to as the partial volume effect (PVE) (Soret *et al* 2007, Rahmim and Zaidi 2008). This issue has been tackled via a number of partial volume correction (PVC) methods that aim to improve image quantitation (Erlandsson *et al* 2012). However, these techniques often involve simplifying assumptions, require access to well-registered anatomical (e.g. MRI, CT) images, and can amplify noise levels (Muller-Gartner *et al* 1992, Rousset *et al* 1998, 2007, 2008). A distinct approach to this problem is point-spread function (PSF) modeling, also referred to as resolution modeling (RM). PSF modeling aims to capture and correct for object-domain and/or detection-domain resolution degrading effects to facilitate more accurate modeling of the measurement (Iriarte *et al* 2016). It therefore reduces image degradations caused by model-mismatch and yields improvements in reconstructed image quality as it compensates for some partial volume effects. As such, PSF modeling has attracted considerable interest in PET over the past decade (Rahmim *et al* 2013, Iriarte *et al* 2016), and has been adopted by major PET vendors in their state-of-the-art PET scanners (Rapisarda *et al* 2010, Jakoby *et al* 2011, Miller *et al* 2015, Qi *et al* 2015, Slomka *et al* 2016).

PSF modeling implementations are commonly divided into (i) image-space, and (ii) projection-space methods. Theoretical analysis of these two approaches has been provided (Cloquet *et al* 2010, Rahmim *et al* 2013), and a few studies have performed preliminary comparison between the two (Sureau *et al* 2008, Kotasidis *et al* 2011). Image-based methods attempt to incorporate resolution blurring effects entirely in the image-space, based on the idea that the reconstructed image can be considered as the blurring of the true image by a point-spread function. This can be performed within the image reconstruction (PSF modeling in the image-space component of the system matrix, thus applied before forward projection and after back projection operations), or performed post-reconstruction in the form of iterative deconvolution (Reader *et al* 2003, Rahmim *et al* 2003, Antich *et al* 2005, Teo *et al* 2007, Cloquet *et al* 2010, Rapisarda *et al* 2010, Kotasidis *et al* 2011, Iriarte *et al* 2016). These methods are straightforward to implement, do not impose a significant computational burden, and produce images

with high quality, so that certain vendor PET scanners have already providing this option (Miller *et al* 2015).

On the other hand, projection-space methods model degradation effects within the projection-space component of the system matrix. Recently, Iriarte *et al* published a review on system models for statistical reconstruction of PET data. They indicate that the most popular approach of combining models of physical degradation factors is to factor the system matrix as a product of independent matrices, each one describing one or a collection of effects (Iriarte *et al* 2016). This arrangement is a well-established approach that has led to high quality efficient reconstructions and yields substantial enhancements in storage preserving and computational time. Also, some studies on image-based PSF modeling considered anisotropy (Rahmim *et al* 2003, Antich *et al* 2005, Sureau *et al* 2008, Rapisarda *et al* 2010), using more complex functions, e.g. mixtures of Gaussians and exponential (Antich *et al* 2005, Sureau *et al* 2008, Cloquet *et al* 2010), and PSF estimation methods based on measurements of arrays of point sources (Sureau *et al* 2008, Kotasidis *et al* 2011, Kotasidis *et al* 2014).

The projection-based methods can be categorized into: (a) empirical methods utilizing measured data points (Panin *et al* 2006a, 2006b), (b) Monte Carlo simulations (Mumcuoglu *et al* 1996, Qi *et al* 1998a, 1998b, Alessio *et al* 2006, Iriarte *et al* 2016), (c) analytical models (Lecomte *et al* 1984, Schmitt *et al* 1988, Selivanov *et al* 2000, Strul *et al* 2003, Rahmim *et al* 2008c) including additional modeling for positron range (Bai *et al* 2003, Bai *et al* 2005, Ruangma *et al* 2006, Rahmim *et al* 2008b, Alessio and MacDonald 2008), and (d) hybrid approaches incorporating combinations of these methodologies; e.g. starting with a simple geometrical calculation, and then imposing additional effects (Selivanov *et al* 2000, Carson *et al* 2003, Yamaya *et al* 2005, Takahashi *et al* 2007). These studies focus on benefits of different methodologies and exploit their synergies to compute the system matrix (Iriarte *et al* 2016).

PSF modeling provides a number of advantages: (i) improved spatial resolution and contrast recovery (Reader *et al* 2003, Rahmim *et al* 2008b, Sureau *et al* 2008, Alessio *et al* 2010, Rapisarda *et al* 2010, Kotasidis *et al* 2011); (ii) reduced spatial noise or image roughness (IR), resulting in a visually smoother image (Vercher-Conejero *et al* 2013), and (iii) improved focal lesion detectability performance (Tong *et al* 2010, Rahmim and Tang 2013a, Karakatsanis *et al* 2014, 2015).

At the same time, PSF modeling poses two concerns (Alessio *et al* 2013): First, it impacts noise characteristics of the reconstructed images as they appear smoother. Moreover, the resulting noise power spectrum (NPS) of the PSF modelled reconstructed image is seen to be amplified in the mid-frequency domain, while exhibiting smaller values at higher frequencies (Rahmim and Tang 2013b). Some efforts have been devoted to the analysis of the resulting noise properties that have important implications for quantitation and lesion detectability performance in PET imaging. These studies performed experimental evaluation of noise characteristics on real datasets (Boellaard *et al* 2004, Bernardi *et al* 2007, Kadrmas *et al* 2009, Rahmim and Tang 2013b), or through Monte Carlo simulations (Soares *et al* 2000, Soares *et al* 2005), and subsequently analysed the impact of reconstruction parameters by adopting a variety of figures-of-merit (FOMs). PSF modeling has different effects on different noise metrics (Tong *et al* 2010, Rahmim and Tang 2013b). Rahmim *et al* used analytic models of noise propagation (Barrett *et al* 1994, Wilson *et al* 1994) to investigate the impact with and without PSF modeling. In a PSF modelled system matrix, more lines-of-response (LORs) contribute to a single voxel, as each voxel is related to more measurement locations compared to a non-PSF modelled system. This results in a more ill-conditioned inverse problem that suffers from slow convergence (Tong *et al* 2011). Moreover, at matched iterations, voxels in a PSF reconstruction depict lower voxel variance and higher inter-voxel correlations versus no-PSF (Rahmim *et al* 2005, Tong *et al* 2010). As a result, PSF modeling noticeably alters the

noise texture. Tong *et al* derived analytical expressions relating image roughness and ensemble noise to voxel variance and inter-voxel correlations (Tong *et al* 2010). Due to PSF modeling, images become smoother, but the ensemble standard deviation of ROI mean uptake (a measure of reproducibility) may remain unchanged (Tong *et al* 2010) or even be amplified for smaller ROIs (Hoetjes *et al* 2010, Blinder *et al* 2012). We elaborate more on how this impacts different noise metrics in section 2.

The second issue concerning PSF modeling is its susceptibility to produce edge overshoot effect—a reminiscence of Gibbs ringing patterns at the edges of a region (Snyder *et al* 1987, Politte and Snyder 1988, Tong *et al* 2011), manifesting as overshoots in smaller regions. This issue may compromise the accuracy of signal quantitation in small regions (Rahmim *et al* 2013). Snyder suggested using a less blurred (i.e. underestimated) version of the true PSF in the reconstruction (Snyder *et al* 1987, Politte and Snyder 1988). This method was shown to be effective at suppressing the edge overshoot effect (Tong *et al* 2011, Blinder *et al* 2012). Snyder (1987) also suggested that a possible reason for the appearance of edge overshoots is the mismatch between estimated PSF and true PSF, and that the small mismatch can be amplified due to the instability of deconvolution process. Nonetheless, it was shown (Reader *et al* 2003, Tong *et al* 2011, Ashrafinia *et al* 2014a, 2014b)—as we also demonstrate in this work—that even reconstruction with the true PSF results in the edge overshoot effect. Furthermore, for specific detection or quantitation tasks, it is plausible that such an effect may even enhance task performance, as we show in this work for quantitation.

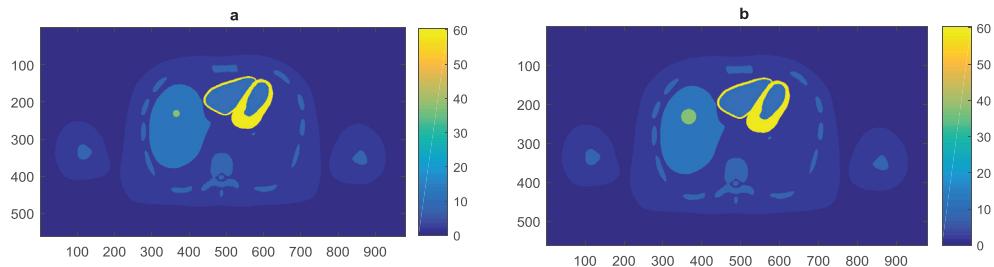
Finally, it is worth noting that clinical tasks vary from pure detection-related tasks (e.g. diagnosis and staging) to quantitation-related tasks (e.g. therapy response assessment and prognostication). The features that make a PET image suitable for detection task differ from those that make an image effective for tumour quantitation. In fact, noise contributes differently to these two general tasks, and PSF-induced noise propagation can result in improved quantitation but may reduce lesion detectability performance, or vice versa (Rahmim and Tang 2013b). However, such analyses and optimizations have only been performed for the case of no-PSF versus PSF modeling. It is possible to generalize PSF modeling to include overestimated and underestimated PSF kernels. This provides a wider range of options to study the impact of PSF modeling on quantitation tasks, thus facilitating task-based optimization for quantitation. Furthermore, given the aforementioned challenges with reproducibility as well as edge overshoots in PSF modeling, it is plausible that generalized PSF modeling may provide kernels that perform optimally, properly balancing different effects. The present work pursues such a generalized PSF framework in the context of tumour quantitation, which in future efforts can be thoroughly evaluated for lesion detectability tasks as well.

## 2. Methods

In this section, we describe our proposed approach to incorporate and assess generalized PSF kernels in the image reconstruction framework. First, we describe simulation configuration, followed by the image reconstruction method incorporating the true PSF kernel. We then explain the methodology of generating a spectrum of PSF kernels from the true PSF. Finally, we define the FOMs for assessing and analysing the results.

### 2.1. Simulation and phantom configuration

We used the 4D anthropomorphic XCAT phantom (Segars *et al* 2010) to generate dynamic FDG-PET images of different tumours, as well as the corresponding attenuation map. In this study, we chose to implement six liver tumours of different diameters (10, 13, 17, 22, 28 and 37 mm),



**Figure 1.** XCAT generated phantom as reference images with liver tumour sizes of (left) 10 mm, and (right) 22 mm. (a) Reference image, tumor diam.: 10 mm. (b) Reference image, tumor diam.: 22 mm.

which was in agreement with the NEMA NU-2 image quality phantom (NEMA 2012). The position of the centre of the tumour spheres is fixed across all six images for consistency. Figure 1 depicts two of these six reference images. The original reference image has a transaxial dimension of  $1024 \times 1024$  and a voxel size of  $0.5856\text{ mm}$ , and 2D OS-EM reconstructions with seven subsets are performed into  $256 \times 256$  images with voxel dimensions  $3.47 \times 3.47 \times 3.27\text{ mm}^3$ . Starting from a higher resolution image is more realistic to better capture the spatial continuity of the actual object (patient) being scanned, albeit its contribution to the reconstruction time.

We simulated a set of 60 min post-injection SUV PET images for a scan duration of 3 min. The FDG tracer kinetics were modelled based on a patient-derived input function (Feng *et al* 1993), a set of realistic kinetic parameters reported in the literature (table 1), and the standard two-tissue compartment kinetic model for FDG. Thus, a respective set of time-activity curves (TACs) were generated for each tissue and tumour to allow calculation of the activity concentration levels at 60-min post-injection. Lesion spheres were also modelled based on rate parameters in the liver region. Additionally, based on evaluation of multiple-patient F-18 FDG PET scans, we found out that the activity concentration values of the soft tissue background outside the liver was about 21% of the corresponding value in the liver. We set the liver rate constants to derive the background activity TAC accordingly. The dynamic acquisition protocol consisted of 9 passes (from 30 to 90 min, 45 s for each bed position). The uptake activities were then calculated by temporal integration for the duration of the scan.

## 2.2. Image reconstruction

The simulations were performed using an in-house validated reconstruction software (Rahmim and Tang 2013b). First, noise-free emission images were generated by assigning the tracer kinetic modelled values to the respective regions of the voxelized XCAT human torso digital phantom. Then, forward projection of the emission images was performed (Rahmim *et al* 2008a, Rahmim *et al* 2010) based on the geometry of a GE Discovery RX PET/CT (Kemp *et al* 2006). The generated sinograms were subsequently attenuated according to the XCAT attenuation factors, and scaled based on the reported sensitivity of the scanner (normalization).

Our reconstruction software performs an analytic OS-EM projection-space based PSF modeling reconstruction and models positron range, geometric projection, photon non-collinearity, inter-crystal scattering, crystal penetration, and corrects for attenuation and detector deficiencies. A detailed modelling of our analytic reconstruction is provided in appendix A.

Analytic simulations were performed for the images reconstructed using 367 radial bins (60 cm radial field of view) and 581 angular samples covering  $180^\circ$  with. Poisson noise was subsequently simulated to generate 200 independent noise-realizations. Finally, the generated

**Table 1.** Kinetic parameters used in the simulation of the anthropomorphic phantom for F-18 FDG tracer. References: myocardium and normal lung (Torizuka *et al* 2000), normal liver (Okazumi *et al* 1992), liver tumour (Okazumi *et al* 1992) and bone (Dimitrakopoulou-Strauss *et al* 2002).

Tissue compartment	$k_1$ (ml min <sup>-1</sup> g <sup>-1</sup> )	$k_2$ (min <sup>-1</sup> )	$k_3$ (min <sup>-1</sup> )	$k_4$ (min <sup>-1</sup> )	$V_b$
Lung	0.301	0.864	0.097	0.001	0.168
Liver	0.864	0.981	0.005	0.016	0.00
Bone	0.091	0.469	0.0023	0.067	0.00
Myocardium	0.6	1.2	0.1	0.001	0.00
Liver tumour	0.243	0.78	0.1	0.00	0.00
Background activity	0.183	0.981	0.005	0.016	0.00

PET projection data were reconstructed, using the proposed methods to produce PET images, as described in next subsection.

Artefacts in the reconstructed PET images are location specific; so ideally the location of the masked ROI has to be the same for both the tumour and the background to perform more precise quantitative analysis. Therefore, we ran our simulation once with tumours present (each of the six tumours) and once with the tumour absent, and then used the mask from the tumour-present reconstructed image to mask out the background region in the tumour-absent reconstructed image. To add more accuracy to our quantitative analysis, we assured that for every PSF-kernel and every iteration of the noise-free reconstructed images, the ROI location in the tumour-absent matches the actual location in the tumour-present.

### 2.3. Generalized PSF modeling

In this section, we describe how to generate a spectrum of PSF (generalized PSF) kernels from the true PSF kernel. We propose an analytical approach to generate a wide spectrum of PSF kernels that portray both *underestimations* and *overestimations* of the true PSF, in addition to no-PSF and true PSF. The ‘no-PSF’ kernel assumes the incoming rays are solely detected at their incident detector, whereas the true PSF kernel matches exactly with the forward-projector based on scanner parameters and mathematical models of blurring explained in appendix A.

In order to implement image reconstruction via a spectrum of PSF kernels that has a smooth transition from no-PSF kernel (identity matrix) to the analytically-modelled ‘true PSF’ and beyond, we observed that we have to simultaneously vary the outputs of three equations that model photon non-collinearity, inter-crystal scattering and penetration. We constructed a series of generalized PSF kernels that included under- and overestimation of the true PSF by applying a line-space of incremental scaling factors to these three modelled terms. More specifically, three sequences of numbers are multiplied by (i) the mass attenuation coefficient for the crystal (LYSO in this case) in equation (A.4) that models inter-crystal penetration, and (ii, iii) the FWHMs of  $D_{\text{non-col}}^\theta$  (equation (A.3)) and  $D_{\text{scatter}}$  that model non-collinearity and inter-crystal scattering, respectively. More details on generating 20 PSF modelled kernels along with a table of scaling factors is presented in appendix B.

### 2.4. FOMs for quantitative analysis

We defined and used multiple metrics for quantitative task-performance analysis. This included four types of noise (image roughness (IR),  $\text{SUV}_{\text{mean}}$  coefficient of variation (CoV),  $\text{SUV}_{\text{max}}$

CoV, and average max-min difference) and two types of bias ( $SUV_{\text{mean}}$  bias and  $SUV_{\text{max}}$  bias). In addition, mean-squared error (MSE) of each voxel and MSE of  $SUV_{\text{mean}}$  were computed. Definitions of these metrics is provided in details in appendix C.

### 3. Results

#### 3.1. Generalized PSF modelled kernels

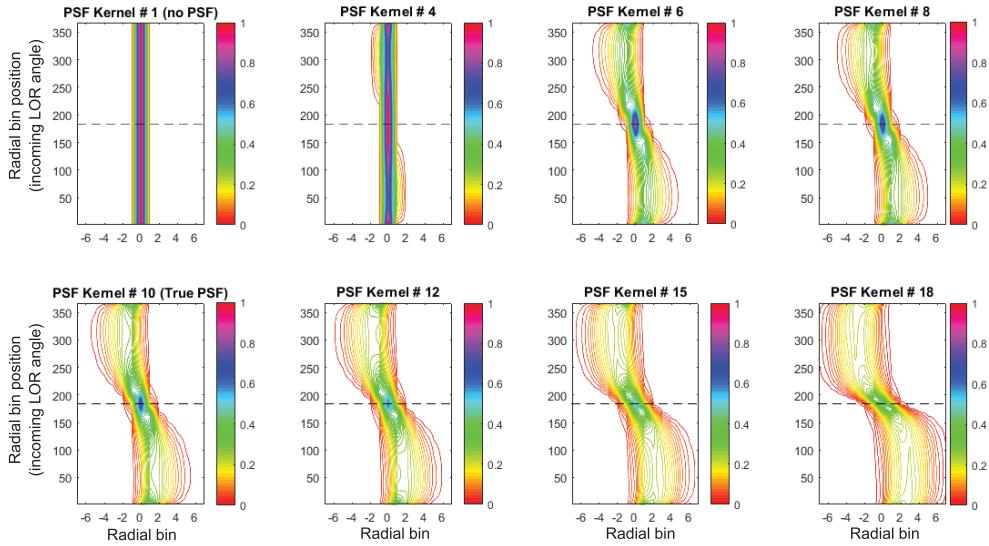
Figure 2 depicts a selection of eight out of 20 generalized PSF isocontours, including ‘true PSF’ and ‘no-PSF’ plots, in addition to three underestimated and three overestimated kernels. Here we follow the convention of Alessio *et al* to demonstrate these kernels by plotting isocontours of their radial profiles (Alessio *et al* 2010). The dashed horizontal line represents the LOR perpendicular to the detector element, in which the maximum probability of detection happens.

It is worth breaking down how each of the three degradations phenomenon affect the PSF kernel. Inter-crystal scattering symmetrically blurs the neighbouring crystals of the incident detector. Equation (A.3) in appendix A addressing photon non-collinearity also yields a symmetric blur. However, in inter-crystal penetration, photons penetrate the neighbouring crystals and cause the parallax effect. This skews the PSF with respect to the true LOR, thereby inducing a symmetry. The generalized PSF modeling kernels presented here has an advantage over the underestimated PSFs performed in image-space in previous studies that characterize PSF kernels by varying the FWHM of the measured PSF (Tong *et al* 2011, Ashrafinia *et al* 2014a, Niu *et al* 2015). Those approaches overlooked two issues with the realistic PSF kernels that we can observe in figure 2. First, realistic PSF kernels are anisotropic, so their FWHM varies with the angle of LOR. Second, under- and overestimating the true PSF not only changes its FWHM, but also shifts its angular-dependant peak location. This can be observed in figure 2, where the peak of radial bins corresponding to LORs entering with an oblique angle (radial bins 1–150 in figure 2) drifts from 1 to 4 with increasing PSF kernels width.

#### 3.2. Reconstructed images

**3.2.1. Noise-free reconstruction.** Figure 3 shows images of the noise-free reconstruction with 10 iterations and 7 subsets. PSF modeling is known to improve resolution and enhance contrast. This can be observed by comparing the no-PSF reconstructed images in the left column with the columns representing kernel #7 (slight underestimation) and beyond. The two major drawbacks of PSF modeling can also be addressed here, as we point out in some observations from this figure. (i) With sufficient iterations, edge ringing phenomenon—a staple aftermath of PSF modeling—starts to appear from kernel 6 (not shown in this figure—an intermediate underestimation of the true PSF—in all tumour sizes), and intensifies as the PSF kernel index—i.e. its deblurring effect—increases. This observation challenges the idea of using underestimated PSF kernels as a solution to eliminate edge overshoot effects (Snyder *et al* 1987, Politte and Snyder 1988). (ii) The other prominent issue with PSF modeling is that it causes higher inter-correlation between voxels. The background (normal liver) regions in the first two kernels (no-PSF and kernel #3) are observed to have a different texture in comparison to the subsequent kernels, as the noise loses some of its high frequency content. This difference becomes more conspicuous when comparing the noise texture of the noisy reconstructed images.

There are three more interesting observations in figure 3. (iii) The edge overshoot in PSF reconstructed images of tumours larger than 17 mm is not uniform across its ring; i.e. the edge is more

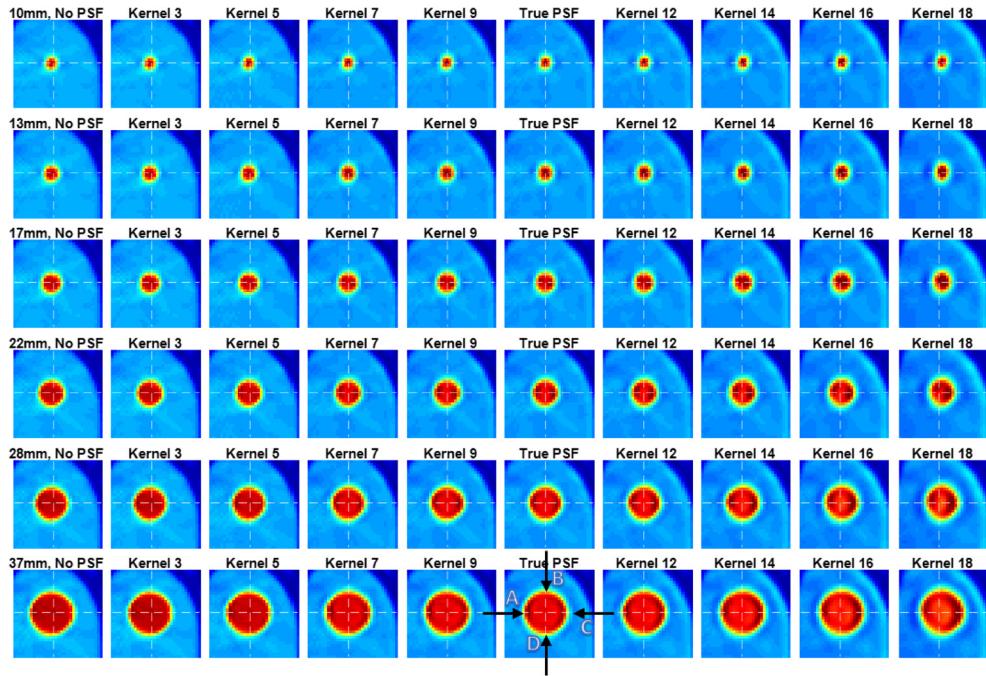


**Figure 2.** Isocontours of selected PSF modelled radial profiles: radial bins positions versus radial bins. The intensity of contours is the probability of an incoming radial bin (LOR) from different angles (vertical axis) to a particular bin and its seven neighbour bins (zero for centred bin and  $\pm 7$  bins in the horizontal axis). The dashed line represents the LOR perpendicular to the detector element. Kernels 4, 6 and 8 are examples of underestimated and kernels 12, 15 and 18 are examples of overestimated PSF kernels.

pronounced in the left and right, compare to the top and bottom. This can be observed in bottom-middle reconstructed image in figure 3 by comparing the regions pointed to by ‘A’ and ‘C’ having a darker red colour with ‘B’ and ‘D’. The reason is closely related to the parallax effect. Photons from annihilation events away from the centre of the FOV may experience significant inter-crystal penetration. Thus, the apparent LOR may not exactly match the true LOR and would be closer to the centre of the FOV. In no-PSF modeling reconstruction, this LOR mismatch resulting in skewed lesions towards the centre of the FOV will not be ‘deblurred’, whereas it will be deblurred by incorporating a true PSF modeling kernel. The edge overshoot appears as an aftermath of this deblurring. The overshoot would be more pronounced in the direction of the parallax effect that skews the regions towards the centre. In this figure the centre of the FOV is located approximately in the left side of the tumour, so the left and right edges of tumour undergo more deblurring compared to the top and the bottom (‘A’ and ‘C’ directions compare to ‘B’ and ‘D’), thus exhibiting more edge overshoot.

Furthermore, (iv) the overshoot on the right side of the ring (pointer ‘C’) is longer than the one on the left side (pointer ‘A’). The reason is the partial ring section in the right is farther with respect to the centre of the FOV than the left. Therefore, the amount of deblurring and edge overshoot is larger, and subsequently an asymmetric edge overshoots will appear on the left and the right of the region.

The final observation is that (iv) the apparent tumour location manifested in the reconstructed image actually drifts *away* from the centre of the FOV as we apply higher kernels. This movement can be tracked using the white dashed lines representing the centre of the tumour in each image. The reconstructed ROI with the 10th kernel (true PSF) is in a perfect position; while it slightly shifts towards the centre of the FOV for underestimated kernels including no-PSF, and slightly shifts away from the centre for overestimated ones. These effects result from



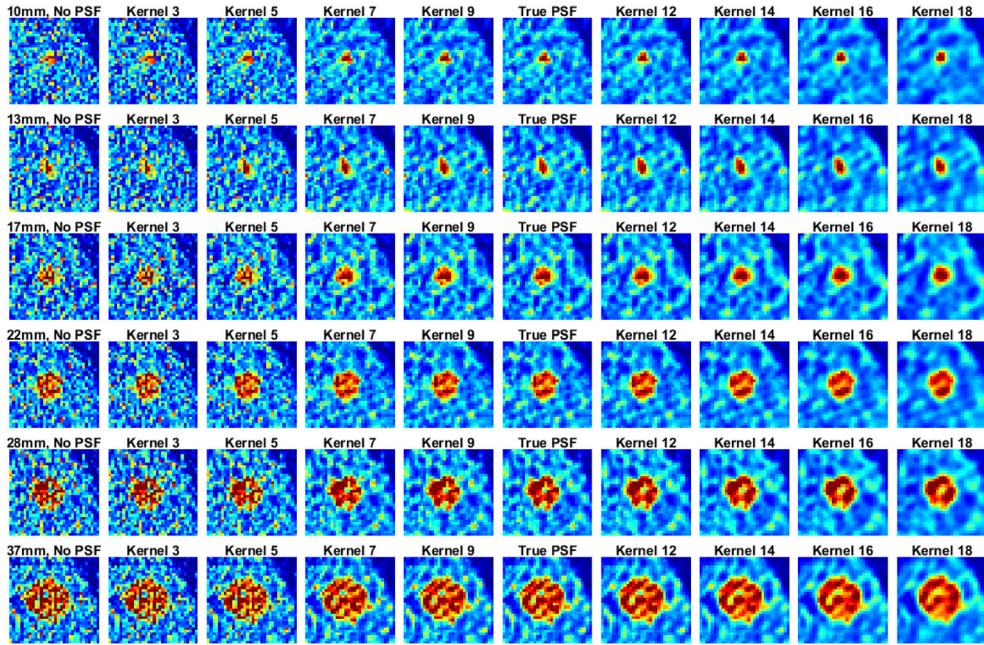
**Figure 3.** Noise-free reconstruction images of liver tumour and background (cropped to include liver tumour and its background tissue) after 10 iterations and 7 subsets. Rows represent different tumour sizes. Columns starting from the left indicate no-PSF reconstruction, four under estimating PSF kernels (#3, #5, #7 and #9), true PSF, and four overestimating PSF kernels (#12, #14, #16 and #18). The intersection of white dashed lines indicates the centre of the tumour in the true object. The centre of the FOV is located at the left-hand side of the tumour, and hence the tumour edges in its left and right sides pointed at by A and C arrows are more pronounced than top and bottom indicated by B and D.

under-/over-correcting for the parallax effect by various PSF kernels. By applying the under-estimated kernels, the full correction (i.e. deblurring) is not yet accomplished, thus the apparent position of the ROI is not back in its initial location; whereas the overestimated the kernels are actually over-correcting (over-deblurring) the region in the reconstruction.

**3.2.2. Noisy reconstruction.** Figure 4 shows noisy reconstructed images. These images display the edge overshoot in the reconstructed ROIs of kernels 7 and above, in addition to its asymmetry, as explained in section 3.2.1. However, they also demonstrate another principle of PSF modeling: modified noise texture. As we increase the deblurring kernel, images look smoother, voxel variance reduces in both the tumour and the background, and the noise becomes more correlated and blobby. The inter-voxel correlation increases as we apply wider PSF kernels, thus the images look smoother with a blobby noise-texture.

### 3.3. Contrast recovery analysis

Figure 5 shows plots of contrast recovery for  $SUV_{\text{mean}}$  and  $SUV_{\text{max}}$  ( $CRC_{\text{mean}}$  and  $CRC_{\text{max}}$ , respectively) of the tumour reconstructed with 20 PSF kernels. The first six images show that neither PSF nor no-PSF kernels can yield a CRC of one. PSF overestimation, however, yields

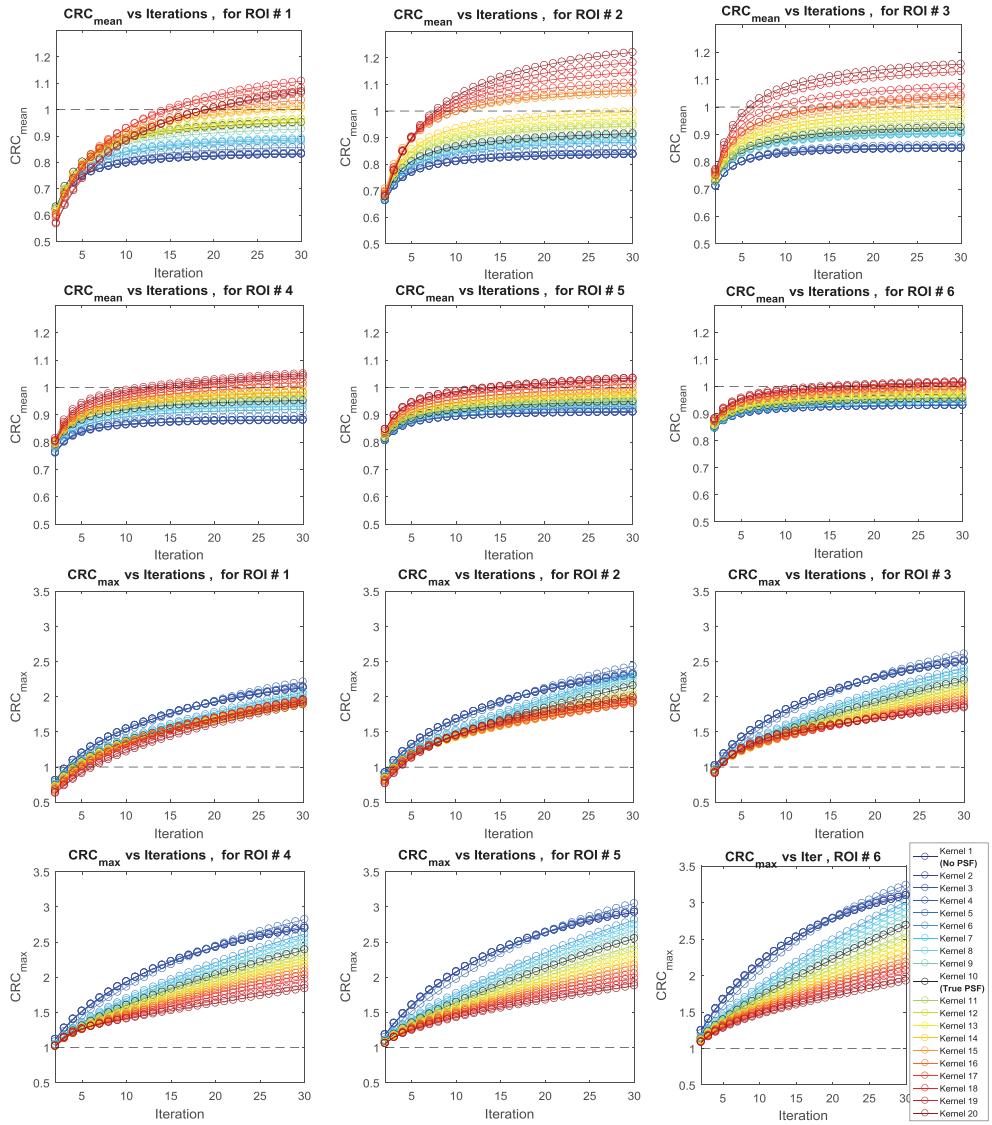


**Figure 4.** Noisy reconstruction images of liver tumour and background (cropped to include liver tumour) for iteration #10 iterations with 7 OS-EM subsets and no post-smoothing. Rows represent different tumour sizes. Columns starting from the left indicate no-PSF reconstruction four under estimating PSF kernels (#3, #5, #7 and #9), true PSF, and four overestimating PSF kernels (#12, #14, #16 and #18).

a CRC value closer to one. Yet in most cases, extreme overestimation (kernels 15 and above), results in  $\text{CRC}_{\text{mean}}$  higher than one, which is as undesirable as  $\text{CRC} < 1$  for underestimated kernels (mostly due to PVE at the edges) and induces an overestimation bias in the reconstructed region that alters quantification.

The  $\text{CRC}_{\text{mean}}$  curves have a smooth and monotonic transition with respect to increasing PSF kernel width. PSF modeling corrects for PVE and thus reduces blurring at the edges, therefore no-PSF modeling yields the maximum PVE and lowest  $\text{CRC}_{\text{mean}}$ .

Another observation from this figure is that the true PSF reconstruction yields  $\text{CRC}_{\text{mean}}$  less than one. The reason is that the overshoot at the edge of the ROI does not actually involve the outmost set of voxels of the region. The set of voxels undergoing overshoot are encircled by another ring(s) of voxels that (i) contain more voxels than the overshoot ring, and (ii) have less uptake than the reference truth. This is because the edge has not been completely recovered until after 40–60 iterations, and the algorithm has not yet perfectly converged. These surrounding voxels at the very edge of the ROI decrease  $\text{CRC}_{\text{mean}}$  to less than one—even in the presence of the overshoot edge—and consequently induce negative bias, as can be seen in the results of section 3.4. The EM algorithm is known to improve with every iteration towards convergence that eventually reconstructs edges perfectly after massive number of iterations. However, this is impractical in PET reconstruction due to the presence of noise and its severe amplification. Therefore, the PVE at the edges of the region impacts  $\text{CRC}_{\text{mean}}$  and causes it to be suboptimal. Thus, the observed underperformance is an attribute of the EM algorithm in PET reconstruction, and it disturbs image reconstruction with any degree of PSF modeling.



**Figure 5.** Averaged CRC of  $SUV_{mean}$  and averaged CRC of  $SUV_{max}$  versus OS-EM iterations for six tumours over 200 noise realizations. The dashed line highlights  $CRC = 1$ .  $CRC_{mean}$  plots have a fixed vertical axis range of [0.5, 1.3], and the range for  $CRC_{max}$  plots is fixed to [0.5, 3.5].

But it can be observed from figure 5 that overestimated PSF modeling kernels tend to mitigate this deficiency.

$CRC_{max}$  plots interestingly follow a reverse pattern, where, in contrary to  $CRC_{mean}$ , underestimated and no-PSF attain higher  $CRC_{max}$  values and the curves decline as kernels' widths increase. We observe that most of the curves lay above one, which is due to (i) the presence of noise and (ii) not performing any post-smoothing on the images that is shown to reduce CRC (Tong *et al* 2011). Moreover, iterative reconstruction algorithms, including OS-EM, are known to intensify the noise as they iterate. Therefore, this produces a monotonic increase

regardless of the generalized PSF kernel. However, PSF reconstruction with wider kernels yields more correlation between the voxels. As a result, voxels cannot oscillate freely in the presence of the noise and their fluctuation decrease as the PSF kernel width increases. This inter-voxel correlation not only constrains the oscillation of each voxel due to noise, but also limits the overshoot from rising too much in reconstruction with overestimated PSF kernels. The combination of these two effects contributes to the reduction of  $SUV_{\max}$  thus  $CRC_{\max}$ .

The  $CRC_{\max}$  curves also show that the smaller tumours (first three ROIs) have smaller  $CRC$  values; even the first few iterations may generate a  $CRC_{\max}$  of one. The reason is again due to the fact that the region is not fully recovered within 30 iterations. Moreover, in the small tumours, PVE lowers the values of the voxels from their reference truth, even when reconstructed with wide PSF kernels. Subsequently, the noise will then be added on top of this PVE and cause the  $CRC_{\max}$  to become closer or even exceed one. More importantly, edge overshoot has not yet developed in the early iterations of smaller tumours because the very few voxels across the ROI have not created sufficient extent for the overshoot to rise. Larger tumours on the other hand have an ample amount of space for multiple overshoot and undershoot rings to appear. Therefore, as the diameter of the region grows,  $CRC_{\max}$  increases, but it decreases with wider PSF kernels.

### 3.4. Noise-bias performance comparison

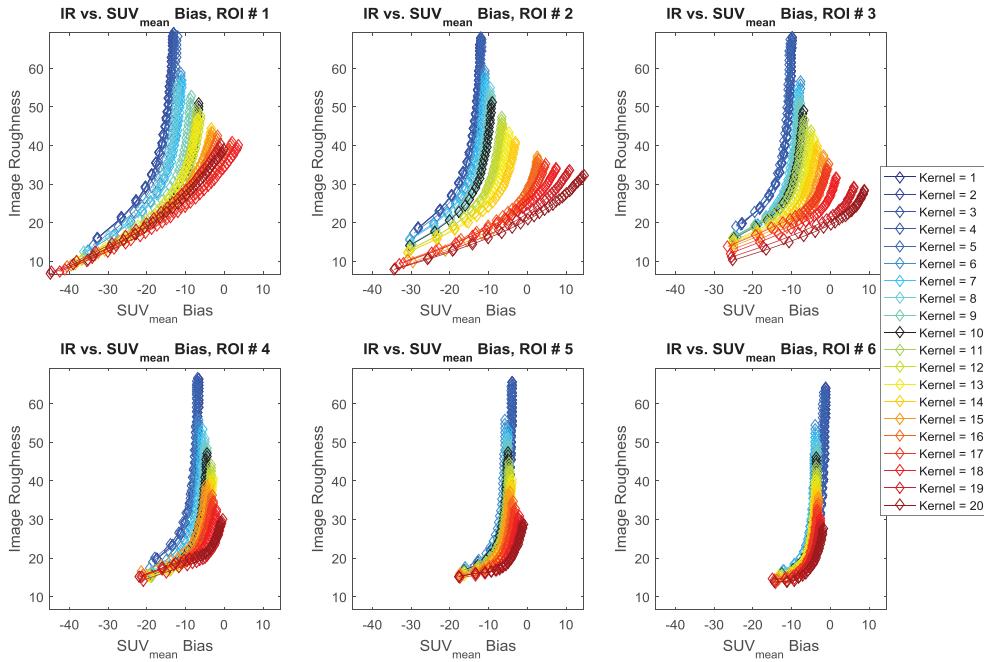
We defined three measures of bias, six measures of noise and two types of MSE in appendix C, and are presenting their plots in this section. Note that the curves representing reconstruction with PSF kernels in all plots follow the same legend as figure 6. In all figures, the starting iteration for plotting is two.

**3.4.1. Image roughness (IR) versus  $SUV_{\text{mean}}$  bias.** Figure 6 shows image roughness versus  $SUV_{\text{mean}}$  bias for six tumours. The range on all six plots is fixed for a better comparison unless otherwise stated.

IR decreases as the kernel index increases, which is consistent with PSF modeling reducing spatial voxel variation and yielding a smoother image. Comparing no-PSF with true PSF (kernel #10) and a medium overestimated PSF kernel #15 (orange curve) at matched iterations shows a range of 25%–35% and 38%–45% less noise for kernels #10 and #15 within all six tumours, respectively. At matched noise,  $SUV_{\text{mean}}$  biases of these two kernels compared to no-PSF degrade significantly for first four tumours (25%–45% less bias for kernel #10 and 60%~94% for kernel #15), while its variation with respect to different kernels drops for two larger tumours (–16% to 2% difference in bias for kernel #10 and 6%–16% less bias for overestimated kernel #15). Excessive overestimation, such as for kernel indices over 16, usually leads to a positive bias.

The behaviour of IR curves can be explained by equation (C.15). No-PSF and under-estimated kernels have higher  $\sigma_0$  and lower covariance value; both of which contribute to amplify the spatial noise. As the kernels width increases towards true PSF and its overestimation,  $\sigma_0$  degrades and voxels exhibit more covariance, and both yield lower image roughness. Moreover, IR has a small increase for larger ROIs. These regions consist of more voxels that result in less weight of the second term on the right-hand side of equation (C.15); which also results in a lower range of IR for larger tumours.

**3.4.2.  $SUV_{\text{mean}}$  CoV versus  $SUV_{\text{mean}}$  Bias.** The plot of  $SUV_{\text{mean}}$  noise versus bias is shown in figure 7. Unlike the other figures displayed earlier, each of the six plots has different axis ranges for more clear differentiation of the curves.

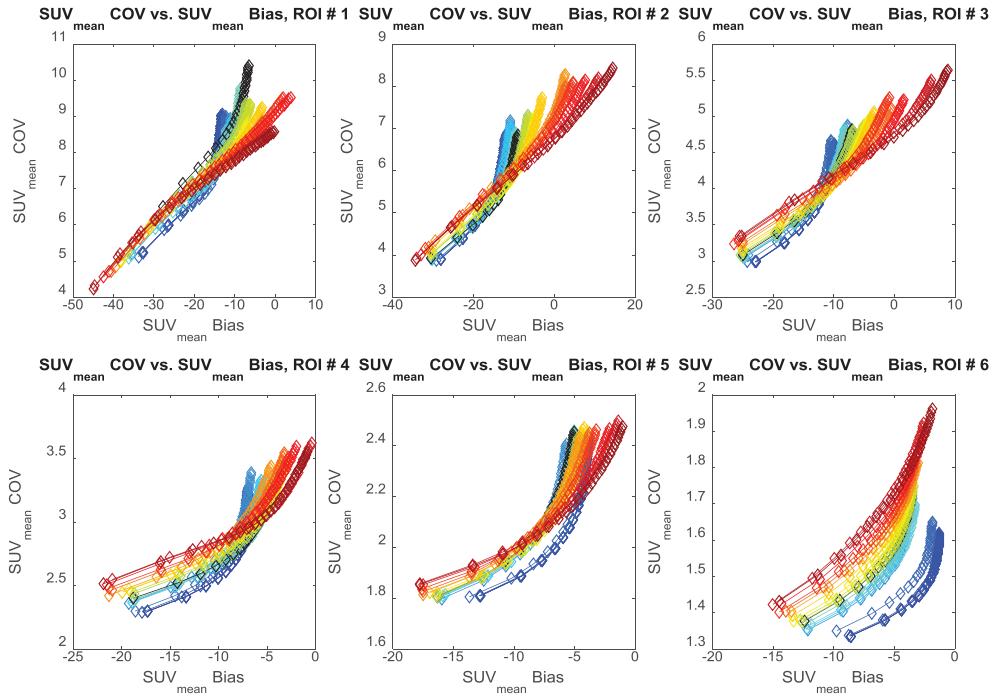


**Figure 6.** Image roughness versus  $SUV_{\text{mean}}$  bias for six tumours. Each point in the curves represents the results for a single OSEM iteration.

Compared to the previous noise versus bias performance curves with noise extending over 60%, the three plots for larger tumours in figure 7 have much lower values and a smaller range of  $SUV_{\text{mean}}$  CoV. Observing lower values for  $SUV_{\text{mean}}$  is predictable, because not only it is an averaging process, but also the PVE at the edges contribute largely to the negative bias. As the PSF kernel's width increases, some undershoots also start to appear and contribute to maintaining lower  $SUV_{\text{mean}}$  values. For matched iterations (7th iteration; same for the past two figures), no-PSF reconstruction demonstrates 0–5% and –1% to 10% less noise compared to kernels #10 and #15, respectively. One immediate reason for such a small  $SUV_{\text{mean}}$  CoV level for larger ROIs is higher number of voxels. Nonetheless, in the three smaller ROIs, the  $SUV_{\text{mean}}$  CoV values and ranges are slightly higher: no-PSF shows (mostly) improved noise performance compared to kernels #10 and #15 by –2% to 15% and 7%–11%, respectively. Some relevant discussions are provided in the next section explaining the observed patterns.

**3.4.3.  $SUV_{\text{max}}$  CoV versus  $SUV_{\text{max}}$  bias.** The plot is shown in figure 8, and displays an approximately linear relationship between noise and bias.

$SUV_{\text{max}}$  noise versus bias curves of the overestimated PSF kernels typically demonstrate lowered noise, lowered bias and thus a higher quantitation performance compared to other earlier PSF kernels. In terms of  $SUV_{\text{max}}$  CoV, kernels #10 (true PSF) and #15 show 43%–57% and 58%–72% less noise, respectively, as compared to no-PSF. In terms of  $SUV_{\text{max}}$  bias, the numbers are 18%–42% and 43%–55%, respectively. This is mainly due to increased inter-voxel correlation explained for  $CRC_{\text{max}}$  in section 3.3. The first three curves in this figure also show that smaller ROIs exhibit a negative  $SUV_{\text{max}}$  bias for the first few iterations (maximum of 7) for the overestimated PSF curves, whereas the three larger ROIs do not show a negative bias, due to reasons explained in 3.2 about the smaller ROIs greatly suffering from PVE.



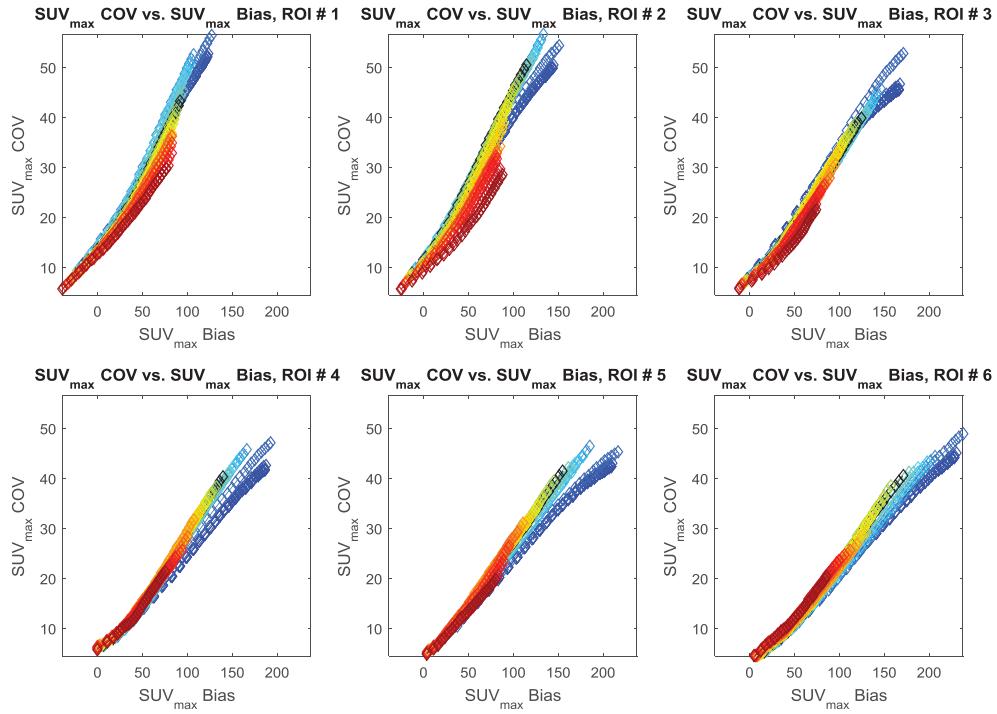
**Figure 7.**  $\text{SUV}_{\text{mean}}$  CoV versus  $\text{SUV}_{\text{mean}}$  bias for six tumours. Note that axes ranges are *not* the same for the plots in this figure.

**3.4.4. Average max-min difference versus  $\text{SUV}_{\text{mean}}$  bias.** The noise versus bias plot is presented in figure 9. It is mentioned in section C.7 that this noise metric essentially is a measure of shape, and quantifies regions' non-uniformities due to PVE and the edge overshoot effect.

The significance of this metric is its ability to assess edge overshoot effect and PVE. Comparing this shape metric for no-PSF with kernels #10 and #15 with matched iterations shows 15%–20% and 21%–28% less shape variability, respectively. Underestimated kernels are more prone to noise and post-smoothing was not performed here, so the majority of higher variability measured in the smaller blurring kernels is due to the noise. Nonetheless, it is interesting to observe the monotonic decrease of this shape variability metric that mostly depicts the effect of higher correlation as a result of more deblurring with wider PSF kernels.

**3.4.5. MSE versus PSF modelled kernels indices.** Figure 10 depicts plots of MSE versus PSF kernels, in which MSE effectively combines noise and bias within a single metric. Every line in the plot corresponds to an OS-EM iteration, and the results are shown for different ROIs. Interesting observations can be made. Except for first few iterations in the smallest ROI, we see generally decreasing MSE values with increasing PSF widths for each given iteration. Also, plotting MSE and performing minimization while allowing for iteration number to vary—i.e. looking at the bottom of each plot—we see that the least MSE is obtained, in all six ROIs, for overestimated kernels.

**3.4.6. MSE of  $\text{SUV}_{\text{mean}}$  versus PSF modelled kernels indices.** We calculated the MSE values for  $\text{SUV}_{\text{mean}}$  using equation (C.12) and plotted them versus PSF kernel indices, as shown in figure 11 (note that each plot has a different horizontal axis range). We indicated that the MSE

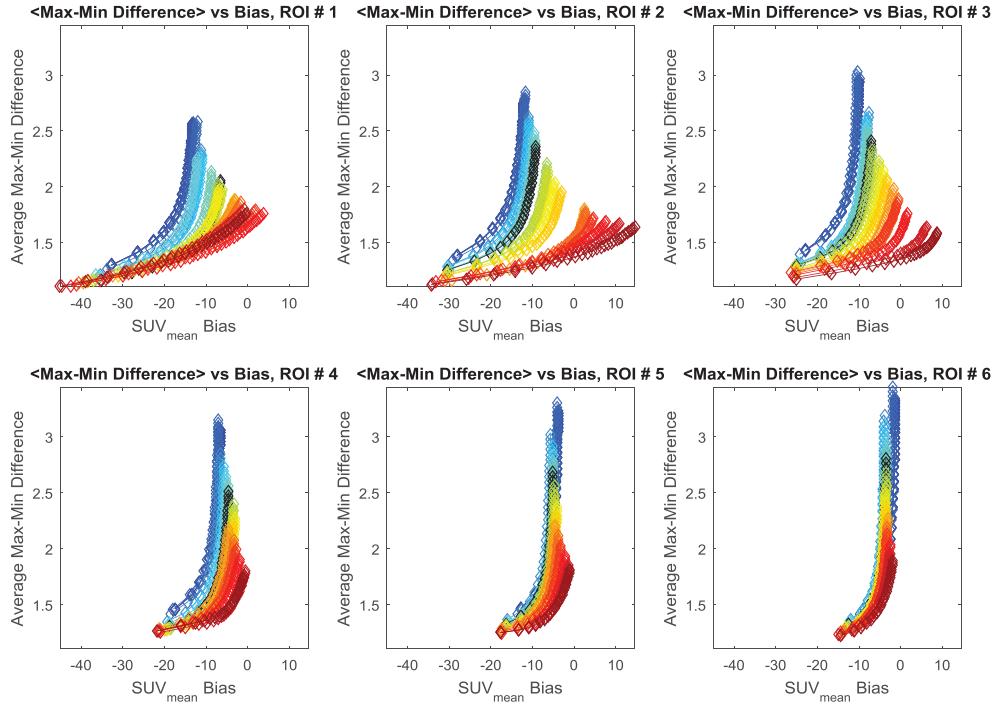


**Figure 8.** SUV<sub>max</sub> CoV versus SUV<sub>max</sub> bias for all six tumours.

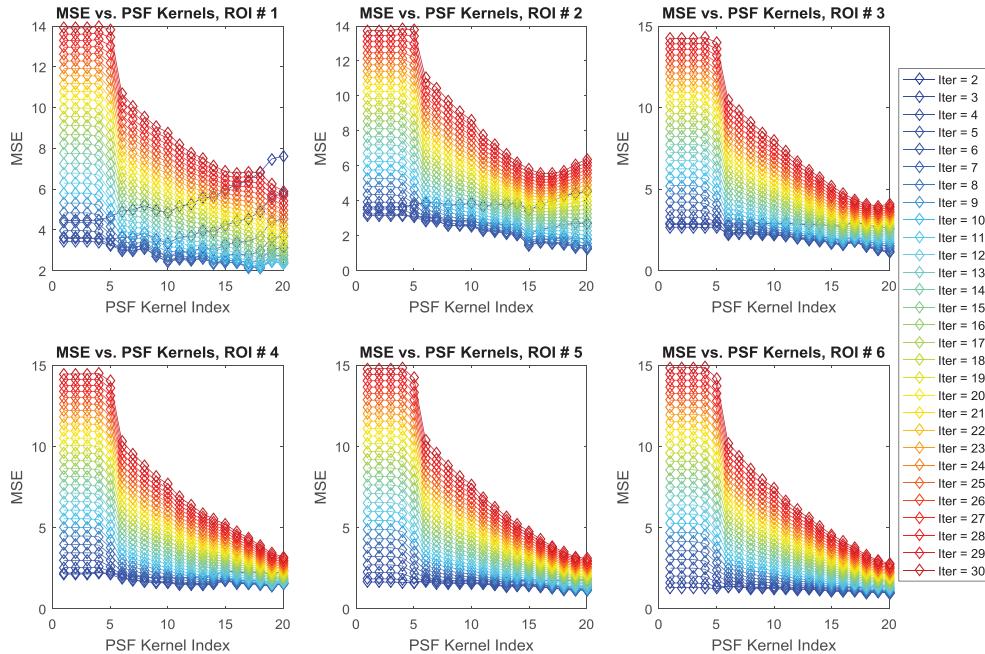
of SUV<sub>mean</sub> captures the effects of both its noise and bias, thus providing us with an indicator of the overall performance of SUV<sub>mean</sub>, which itself is robust to spatial noise and can be used to quantify PVE. In these plots following the trends of increasing iterations implies wider PSF kernels require more iterations to converge than narrower PSFs. Furthermore, we observe that the best performance, in terms of the minimum MSE of SUV<sub>mean</sub>, was obtained for overestimated PSF kernels. More specifically, it is seen that for small tumours, minimum overall MSE is obtained with medium PSF overestimation, whereas for the two largest regions, the slight decrease in SUV<sub>mean</sub> CoV (higher SUV<sub>mean</sub> reproducibility, as explained in section 3.4.2) boosts the performance of underestimated PSF kernels. However, the MSE improvement between underestimated PSFs and true PSF for the two largest regions are 0.10 and 0.11, while the improvements for the overestimated PSFs (kernel 14 for instance) versus the true PSF are 0.56, 0.61, 0.36 and 0.08 for ROIs 1 to 4, respectively. Therefore, overestimated PSF achieve more significant improvement.

**3.4.7. Plots versus sphere diameters.** In the following figure 12, six of the FOMs are plotted versus ROI sphere diameters. The iteration is fixed at 10 for plotting.

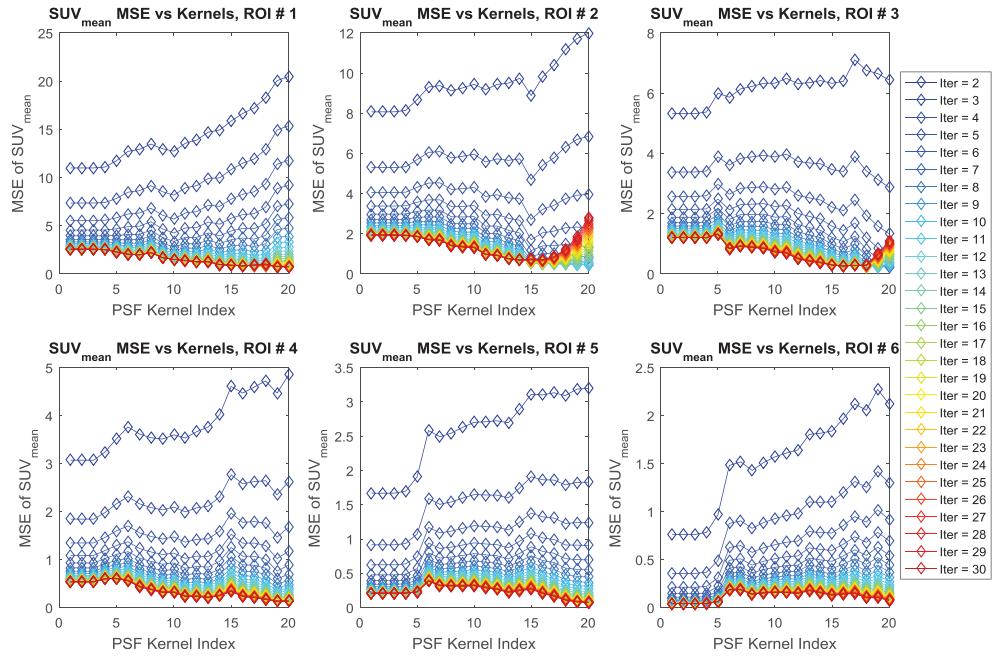
These six plots provide better intuition about the impact of PSF kernels for ROIs with different sizes and help make interesting observations. Figure 12(a) shows a small reduction in average voxel variation as the diameter increases. This is related to the fact that with larger ROIs, a smaller fraction of voxels is impacted by edge overshoot effect, which itself amplifies voxel variability (Rahmim and Tang 2013b). Also as seen in figure 12(b), image roughness decreases with increasing ROI size. This is also related to above effect, as well as the fact that second term in equation (C.15) decreases with increasing voxels (M). Figure 13(c) shows an important trend. As mentioned in section 3.4.2, the only plots that show inferior performance



**Figure 9.** Averaged max-min difference versus  $SUV_{\text{mean}}$  bias for each of the six tumours studied.



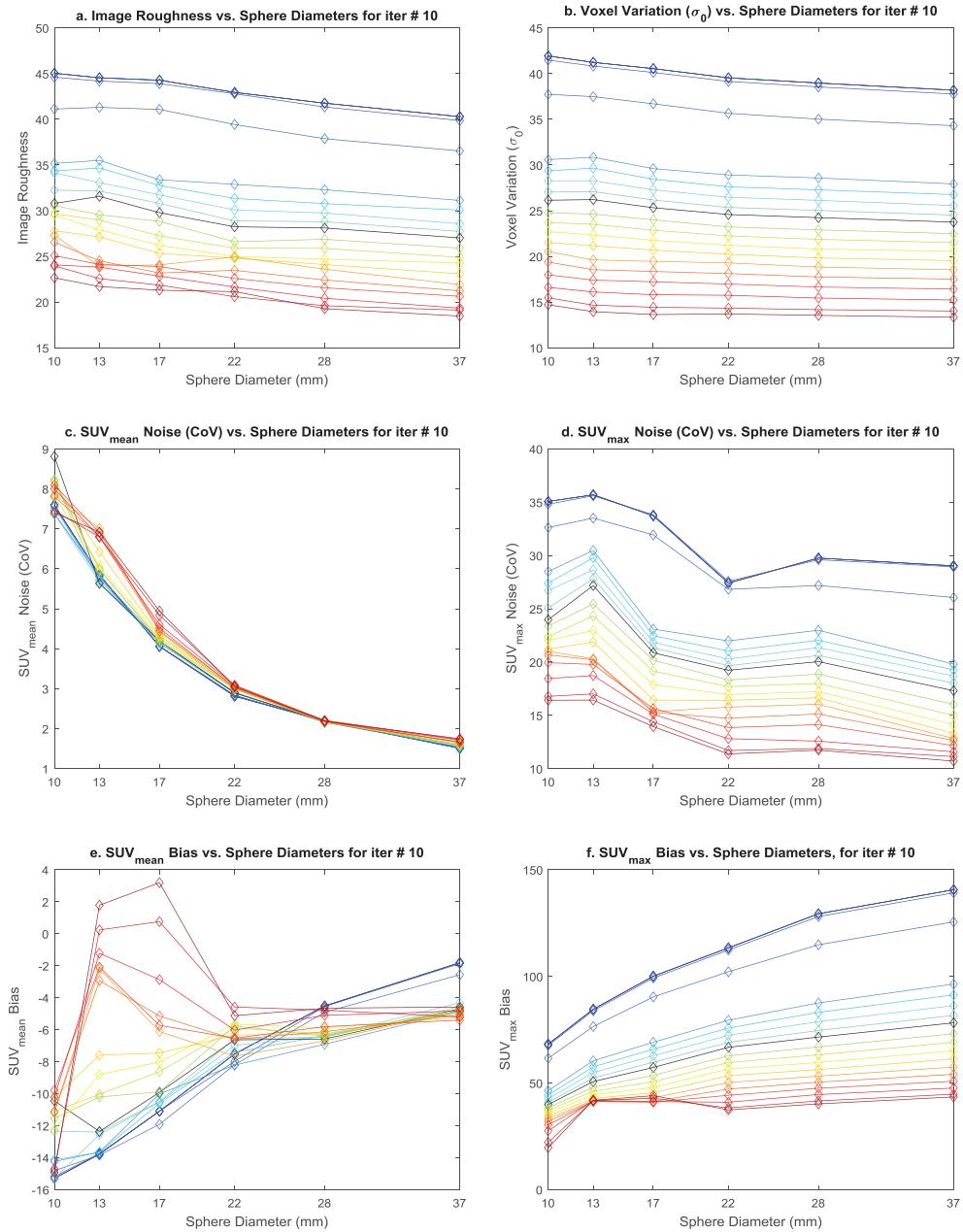
**Figure 10.** MSE versus PSF kernels. Each colour indicates an OS-EM iteration.



**Figure 11.** MSE of  $SUV_{\text{mean}}$  versus PSF kernels. Note that each plot has a different scale.

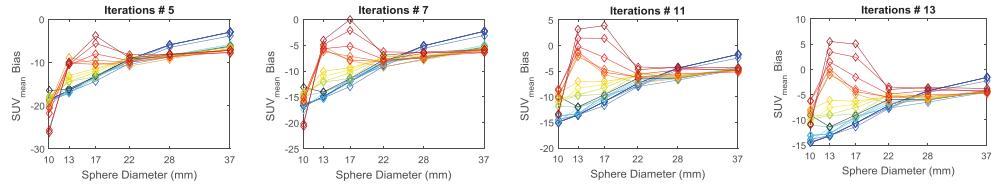
in overestimated PSF kernels belong to  $SUV_{\text{mean}}$  CoV. But as we explained, both figures 7 and 12(c) indicate that this inferior reproducibility is not substantial: in the worst case, it varies within a range of 1.5% for the 10 mm and 13 mm ROIs. This figure also shows a considerable decline in CoV with an increase in number of voxels in an ROI, as expected, because the mean attribute becomes substantially more robust.

An analogous trend observed in figures 12(d)–(f) is a peak in  $SUV_{\text{mean}}$  CoV,  $SUV_{\text{mean}}$  bias and  $SUV_{\text{max}}$  bias curves, respectively for 13–17 mm tumours, especially with PSF overestimation. This pattern is mostly generated because of edge overshoot effect. We explained in section 3.3 that although this aftermath is referred to as *edge* overshoot, it does not involve voxels exactly at the edge of the ROI. The outermost ring of voxels always undergoes PVE (even at extreme PSF overestimation) where they have not yet recovered their true value. Excluding this outermost ring, four ring of voxels can contribute to the first overshoot ring as can be observed in large ROIs in the last row of figure 3 for kernel #10 (true PSF), and even more rings in the case of extreme overestimation. In larger ROIs the distance between one end of the region to the other is much larger than 10 voxels, therefore a complete doughnut-shape overshoot can arise in the edge of the region. The centre of this doughnut accommodates smaller undershoots and overshoots in mid-sized ROIs, as well as a flat region in larger ROIs. However, in smaller lesions, such as the 13 mm and 17 mm tumours, the extent from one edge to the other is less than what is needed for these two overshoots to remain separate (thus generate a doughnut). Hence these two overshoots *merge* together and create a single overshoot with a value higher than each of the individual edges. This results in higher  $SUV_{\text{max}}$  values, including higher  $SUV_{\text{max}}$  noise (figure 12(d)) and bias (figure 12(f)). Due to this phenomenon in smaller regions, most of the voxels inside the reconstructed ROI undergo an overshoot and exhibit a positive bias, resulting in a more positive  $SUV_{\text{mean}}$  bias. This is the main reason for



**Figure 12.** Various quantification metrics versus ROI diameters (in mm) at iteration 10 from different kernel sizes: (a) image roughness, (b) voxel variation, (c) SUV<sub>mean</sub> CoV, (d) SUV<sub>max</sub> CoV, (e) SUV<sub>mean</sub> bias, and (f) SUV<sub>max</sub> bias. Legends are similar to figure 6.

the peak in 13 mm and 17 mm regions in figure 12(e). In larger tumours, the outermost ring of voxels contains more voxels than the overshoot rings exhibiting overshoot. Moreover, the negative bias that these outermost voxels experience due to PVE is much larger than the positive bias that inner ring of voxels exhibits due to the edge overshoot, which eventually causes the SUV<sub>mean</sub> bias to become negative.



**Figure 13.**  $SUV_{\text{mean}}$  bias versus sphere diameters for iterations number (a) 5, (b) 7, and (c) 12. Legend is the same as figure 6.

## 4. Discussion

### 4.1. Noise versus bias analysis

In the current work, we plotted various noise versus bias curves to assess the quantitation performance of the generalized PSF modeling kernels. We performed a comprehensive analysis of generalized PSF modeling reconstruction for assessing quantitative task performance, including noise versus bias analysis between four types of noise (IR,  $SUV_{\text{mean}}$  CoV,  $SUV_{\text{max}}$  CoV, and average max-min difference) and two biases ( $SUV_{\text{mean}}$  bias and  $SUV_{\text{max}}$  bias). Past efforts have focused on quantitation performance comparison between PSF versus no-PSF modeling reconstructions. These included bias (and/or contrast) versus noise trade-off curves, commonly illustrating outperformance of PSF when defining noise as IR (Reader *et al* 2003, Le Meunier *et al* 2010, Rapisarda *et al* 2010, Kotasidis *et al* 2011), or  $SUV_{\text{mean}}$  CoV (Alessio *et al* 2006, Sureau *et al* 2008, Alessio *et al* 2010, Cloquet *et al* 2010). Our present study shows improvement in IR,  $SUV_{\text{max}}$  CoV, and average max-min difference for the true PSF kernel versus no PSF, complying with previous reports, and demonstrates medium overestimated PSF kernels outperforming the true PSF. At the same time, reproducibility in terms of  $SUV_{\text{mean}}$  CoV between true PSF versus no PSF shows less significant improvement (for small tumours with the same number of OS-EM iterations) or even a slight degradation (for small regions for PSF versus no PSF or which complies with the previous reports; both of which complies with previous reports (Tong *et al* 2010, Hoetjes *et al* 2010, Blinder *et al* 2012)). The same behaviour is observed for overestimated PSF versus true PSF. However, since  $SUV_{\text{mean}}$  metric involves averaging voxels yielding a smoother value, its coefficient of variability is generally very low. Therefore, in practice, reproducibility degradation of overestimated PSF versus no-PSF is negligible (<10% variability for a CoV of 3%).

We also notice that more thorough analysis of reproducibility versus bias can be performed for many other metrics (beyond  $SUV_{\text{mean}}$  and  $SUV_{\text{max}}$ ) in the emerging area of radiomics and heterogeneity quantification (Eary *et al* 2008, El Naqa *et al* 2009, Tixier *et al* 2011, van Velden *et al* 2011, Asselin *et al* 2012, Lambin *et al* 2012, Kumar *et al* 2012, Vriens *et al* 2012, Chicklore *et al* 2013, Aerts *et al* 2014, Tixier *et al* 2014, Hatt *et al* 2015, van Velden *et al* 2015, Ashrafinia *et al* 2016a, Hatt *et al* 2016, Lu *et al* 2016, Rahmim *et al* 2016, Soufi *et al* 2016). This requires a distinct effort which can be pursued in the context of varying PSF kernels.

### 4.2. Important factors in quantitation analysis

In the present work, it was seen that two essential and determining parameters need to be carefully tuned for task-performance optimization: (i) number of iterations and (ii) segmentation of the target region or thresholding the ROI.

**4.2.1. Number of iterations in reconstruction.** It can be seen from the plots in sections 3.3 and 3.4 that matched iterations cause different levels of contrast, noise or bias in reconstructions

with different PSF kernels. This is because the degree of convergence in true PSF versus no-PSF reconstructions differs at the same number of iterations. This is very important when assessing and analysing quantitation task performance. As an example, consider the four plots of  $SUV_{\text{mean}}$  bias versus sphere diameters in figure 13 for iterations 5, 7, 11 and 13. We observe that these plots follow a comparable pattern, where curves peak at 13 mm and 17 mm ROIs. However, the values vary considerably. None of the curves demonstrate a positive bias in iterations 5 and 7; whereas in iterations 11 and 13 the maximum value of the plot increases and creates a positive bias. As such, merely comparing images reconstructed with different PSF kernels at the same number of iterations for PSF and no-PSF would not be sufficient (although this practice commonly appears in the literature (Niu *et al* 2015, Iriarte *et al* 2016)). Therefore, it is necessary to observe the *trend* of quantification metrics with increasing iterations in order to obtain an accurate and meaningful comparison.

**4.2.2. ROI segmentation.** ROI segmentation was performed using thresholding, and the level was set to 55% in the present work, which we explored and assured visually to correspond very closely to the reference truth region. In any case, we observed that ROI segmentation has a significant impact on quantitation. This is because the uptake values of the outermost voxels at the boundaries of a reconstructed ROI suffer from PVE and hence have relatively lower values compared to the reference truth, which eventually impacts the noise and bias metrics. Correspondingly, the maximum voxel in a noise-free reconstructed ROI may represent either the overshoot due to the edge overshoot effect in mild underestimated, full, and overestimated PSF modeling, or the true value of the ROI in no-PSF and heavily underestimated PSF modeling. On the other hand, the minimum voxel in such an ROI may represent a voxel in an undershoot of the ringing effect in mild underestimated, full, and overestimated PSF modeling, while it may also refer to a low uptake in the very edge of the region due to PVE. In this case, the thresholding should be defined properly so that the masked ROI excludes low uptakes at the edges. Otherwise, in the case of mild underestimated, full, and overestimated PSF modeling, it would be hard to determine whether the minimum uptake is due to an undershoot of the edge effect that occurs inside the first overshoot ring, or it is a blurring due to PVE at the border of the ROI. To prevent this confusion, we optimized the thresholding ratio to preserve the region shape, while excluding blurred voxels due to PVE in the noise-free reconstruction, and then apply that thresholding ratio to the noisy reconstructed images.

#### 4.3. Noise metrics

From the analysis provided in section C.6 of appendix C, it can be deduced that noise, when measured spatially, can be significantly reduced by PSF modeling (given the same iteration number) whereas  $SUV_{\text{mean}}$  CoV actually could be unchanged (Tong *et al* 2010), increase (Blinder *et al* 2012), even multi-fold (Hoetjes *et al* 2010). This potential increase in  $SUV_{\text{mean}}$  CoV means a reduction in  $SUV_{\text{mean}}$  reproducibility. The present work explored how generalized PSF modeling (under or over estimated PSF kernel) impacts the quantitative performance of different FOMs in addition to the impact of iteration numbers.

We observed in section 3.4.2 that  $SUV_{\text{mean}}$  CoV versus bias curves of underestimated PSF kernels demonstrated slightly better performance for small regions. Visual assessment of these plots reveals that the underestimated PSF kernels outperform true and overestimated PSF kernels only within the first few iterations. However, if iterations exceed 40 (not depicted), the noise versus bias curves of underestimated PSF kernels for the largest ROI in figure 7 would eventually be outperformed by overestimated and true PSF curves, as observed in figures 6,

8 and 9. The reason for this behaviour can be partially explained by equations (C.15) and (C.16) in appendix C; especially equation (C.16), in which the two terms of  $\sigma_0$  and covariance act against each other. Although more iterations increase  $\sigma_0$  and decrease covariance, the effect of  $1/M$  factor in the first term of equation (C.16) further diminishes the effect of  $\sigma_0$  compared to the first term in equation (C.15), thus the first term cannot impact  $SUV_{\text{mean}}$  CoV in equation (C.16) as much as it impacts IR in equation (C.15). Furthermore, kernels with higher PSF width cause higher covariance values in equation (C.16), and increased covariance values at earlier iterations contribute to  $SUV_{\text{mean}}$  CoV in smaller regions (small  $M$ ), and this increase eventually cause a decrease in the reproducibility of the PSF kernel with larger indices. Nonetheless, this reproducibility degradation in overestimated kernels, particularly the aforementioned kernel #15, was lower (<10%) compared to improvements (reductions) in IR (~50%) and  $\sigma_0$  (30%–50%), while increasing CRC (a  $CRC_{\text{mean}}$  of 0.95 for kernel #15 versus 0.85 for kernel #10 and 0.80 for no-PSF).

In the present work, we did not include any noise-suppression or control in the current analysis, but surely studies should be pursued to analyse methods such as post-reconstruction smoothing, penalized EM, and other credible methods and study their effects along with various degrees of PSF modeling in quantitation. Also, the present work was focusing on a single tumour contrast derived from clinically derived kinetic parameters for a liver tumour. Future studies need to be performed that includes tumours with higher contrasts. The present work proposes and expands on a particularly neglected dimension of imaging, namely on the extent of PSF kernels. But in future studies it should be integrated along with various filtering regimes and various numbers of TOF kernels to find the optimal combination for optimizing the overall PET quantification.

#### 4.4. Comparison with past efforts

A preliminary study of projection-space generalized PSF modeling was performed involving a spectrum of underestimated PSF kernels (Ashrafinia *et al* 2014b). In that study, we quantitatively analysed projection-space reconstructions with a spectrum of PSF kernels generated from the true PSF using a convex combination approach; i.e.  $(1 - \alpha) \times \text{'no-PSF kernel'} + \alpha \times \text{'true PSF kernel'}$ ,  $\alpha \in [0, 1]$ , where the ‘no-PSF’ kernel is a delta function that assumes the incoming rays are solely detected at their incident detector. Those intermediate kernels are considered *underestimations* of the true PSF kernels. The convex combination method used to generate PSF kernels for that study cannot be extended to generate overestimated kernels, thus we chose to rescale reconstruction parameters as we explained in appendix B. Therefore, although some of the underestimated PSF kernels on that study do not perfectly match the PSF kernels used in this work, the results follow the same pattern: the  $SUV_{\text{mean}}$  CoV of slightly underestimated kernels showed a 12% increase compared to true PSF.

The present work pursues such a generalized PSF framework in the context of quantitation (Ashrafinia *et al* 2016b), which in future efforts can be thoroughly evaluated for detection tasks. Additional work is being pursued in this area on the front of heterogeneity analysis, including studies on the effect of PSF modeling on heterogeneity quantification task performance, with the aim of adopting more reproducible and robust shape and textural features and optimizing them for enhanced prediction and prognostic tasks (Ashrafinia *et al* 2016a).

Positron range is another PET degradation that induces more blur to the system matrix, as it increases the FWHM of the PSF. Some vendor PET scanners use Ge-68 point sources to characterize the PSF as modelled within PSF-reconstruction (Panin *et al* 2006a). However, in comparison, the most popular isotope, F-18, has a relatively small positron range. Therefore, utilizing a PSF kernel obtained from Ge-68 point sources that has a significant positron range

to reconstruct F-18 PET data is approximately equivalent to an overestimated PSF kernel in the reconstruction, because both increase the FWHM of the PSF. It is interesting to note, based on our observations with overestimated PSF kernels, that this may not be a problem in fact, and may effectively lead to improvements in quantitative performance, though further analysis is required to implement the exact model of radiotracers with higher positron range, that is left for future study.

## 5. Conclusion

PSF modeling is an increasingly employed partial volume correction method. We studied the impact of an array of projection-space-based PSF models on PET reconstructed images for optimized quantitative task performance. The system PSF was constructed using models of photon non-collinearity, inter-crystal scattering, and inter-crystal penetration. Using these models, we generated 20 generalized PSF modeling kernels—including no and true PSF, as well as 8 underestimated and 10 overestimated PSF kernels. We used an XCAT anthropomorphic phantom with 6 different liver tumour sizes and kinetically derived F-18 FDG time-activity curves to reconstruct noise-free, as well as 200 noisy images using the OS-EM algorithm. The quantitative figures of merit included contrast recovery, mean-squared error, and various noise metrics (image roughness, voxel variation,  $SUV_{mean}$  and  $SUV_{max}$  coefficient of variability (CoV), averaged max-min difference) and biases ( $SUV_{mean}$ ,  $SUV_{max}$ ). We evaluated these metrics for different tumour sizes/iterations/PSF kernels/noise realization.

Our results show that for the standard range of iterations employed in clinic (not excessive), edge enhancement due to overestimation counter-intuitively lowered SUV bias in small tumours, while inter-voxel correlations suppressed image roughness and enhanced uniformity in all tumours, only slightly degrading  $SUV_{mean}$  reproducibility in the smallest tumours. One may at first imagine that overestimating the PSF would lead to higher overshoots at the edges. However, we only observed this at higher iterations. In fact, using an overestimated PSF resulted in increased contrast and limited edge overshoot effect at lower iterations, in turn enabling enhanced SUV quantitation. Overall, this work suggests that one ought not to necessarily utilize an exactly matched system PSF for enhanced image reconstruction performance, and that slightly overestimated PSF modeling can improve task-based quantitation. This framework can as such be pursued as a powerful and viable approach in quantitative task-based optimization including prognostication and treatment response assessment.

## Acknowledgments

This work was supported in part by Siemens Medical Solutions. The authors would also like to thank Doctors Martin Lodge and Adam Alessio for fruitful discussions.

## Appendix A. Image reconstruction model

This appendix contains the analytical modeling of the true PSF kernel and the image reconstruction used in this study. The PSF model is further expanded to a generalized PSF modeling framework containing a spectrum of PSF kernels. Suppose  $p(i,j)$  is the element of the detection probability matrix  $P \in \mathbb{R}^{M \times N}$  that represents the probability of detecting an emission from pixel  $j$  ( $j = 1, \dots, N$ ), at detector pair  $i$  ( $i = 1, \dots, M$ ). Currently factorized schemes for the system matrix are based on the proposed works of Mumcuoglu *et al* (1996) and Qi *et al* (1998a) for 2D and 3D acquisitions respectively:

$$P = P_{\text{det. sens}} P_{\text{det. blur}} P_{\text{atten}} P_{\text{geom}} P_{\text{im.blur}}. \quad (\text{A.1})$$

$P_{\text{im.blur}}$  accounts for image-based blurring effects, particularly the positron range. A detailed discussion on analytically modeling positron range effects in statistical image reconstruction can be found in Qi *et al* (1998a). As the current work focuses on F-18 FDG scanning known to exhibit short mean positron range (0.64 mm), the effect of image-based blurring can be safely ignored in the PSF model.

$P_{\text{geom}}$  is the geometric projection matrix where the  $(i,j)$  element defines the probability that a photon pair produced in voxel  $j$  reaches the front faces of the LOR  $i$  in the absence of attenuation and assuming perfect photon-pair collinearity. We used the built-in Radon transform command in Matlab® to perform the geometric projection. The number of projection bins of this Matlab® function is sufficient to compute the projection at unit intervals, even along the diagonal of the projection matrix.

$P_{\text{det.blur}}$  accounts for blurring in the sinogram space and consists of photon pair non-collinearity, inter-crystal scattering, and crystal penetration (Cherry *et al* 2012). Technically, modeling radial, angular and inter-sinogram blurring requires a 3D blurring scheme. However, in the present work, we assume a small axial acceptance angle, and these blurring effects are confined to a single sinogram using a 2D blurring model. Each of these effects can be analytically modelled and eventually combined. We briefly discuss modeling of resolution degrading effects in sinogram space and later we show how to exploit these analytical expressions to create adaptive generalized PSF kernels.

Photon non-collinearity has to be technically modelled in the geometric component  $P_{\text{geom}}$  in equation (A.1). However, as an approximation to considerably simplify the system matrix computation, one can assume photon non-collinearity is depth independent and model it in the projection-space component  $P_{\text{det.blur}}$  of the system matrix (Mumcuoglu *et al* 1996, Rahmim and Tang 2013b). The FWHM of the Gaussian approximation that models the effect of blurring on the spatial resolution using simple geometric calculations can be given as:

$$\text{FWHM} \cong \left( 0.25 \times \frac{\pi}{180} \right) \frac{L}{2} = 0.0022 \times L \quad (\text{A.2})$$

where  $L$  is the detector separation and can be related to the scanner diameter  $D$  for different angles of incidence  $\theta$  via  $L = 2D \cos(\theta)$ . Therefore, equation (A.2) becomes:

$$\text{FWHM} \cong 0.0022 \times 2D \cos(\theta). \quad (\text{A.3})$$

This results in an angular dependent Gaussian blurring kernel  $D_{\text{non-col}}^\theta(x, z)$ , which models blurring due to photon non-collinearity along the  $(x, z)$  in radial and axial directions of the sinogram, respectively.

Although the remaining two effects—inter-crystal scattering and penetration—are often not distinguished from one another, it would be very beneficial to conceptually separate them for proper modeling: inter-crystal penetration occurs when a photon penetrates the incident detector element and is detected in the adjacent crystal; whereas inter-crystal scattering of photons can occur even when the angle of incidence is 90°.

The penetration effect can be modelled using our knowledge of the 511 keV attenuation coefficient of crystals,  $\mu$ , as well as the angle of incidence  $\theta$ . If we model the individual detectors' penetration by  $p^\theta(x)$ , then we can calculate the resulting penetration distribution  $D_{\text{penet}}^\theta(x)$  for the coincident pair as (Lecomte *et al* 1984, Schmitt *et al* 1988):

$$D_{\text{penet}}^\theta = \int_x p^\theta(X) p^\theta(2x - X) dX \quad (\text{A.4})$$

where the 1D attenuation distribution  $p^\theta(x)$  can be described by the angular-dependent exponential function:

$$p^\theta(x) = e^{-\mu x'/\sin(\theta)} \quad (\text{A.5})$$

where  $x'$  denotes the tangential direction along the detector surface, such that  $x = x' \cos(\theta)$ .

For the scattering component, which can be effectively decoupled from the penetration effect (Rahmim *et al* 2008c), we used real measured projection data of a point source at the centre of the field-of-view (FOV) and determined the average radial and axial scatter blurring in the sinogram space. After correcting for the non-collinearity effect through subtraction in squares, we arrive at an estimate for  $D_{\text{scatter}}(x, z)$ . An important point is that, since the non-collinearity effect is independent of crystal blurring, it can be convolved with the corresponding crystal blurring kernels once they are created, as we briefly discuss next. Consequently, combining the above analytical models, we arrive at the overall projection-space blurring kernel  $D_{\text{proj}}^\theta(x, z)$  through the following convolution expression:

$$D_{\text{proj}}^\theta(x, z) = D_{\text{non-coil}}^\theta(x, z) * D_{\text{penet}}^\theta(x) * D_{\text{scatter}}(x, z). \quad (\text{A.6})$$

$P_{\text{atten}}$  in equation (A.1) is a diagonal matrix containing the attenuation coefficients. The attenuation image in current work was derived by forward projecting the attenuation map created by the XCAT anthropomorphic phantom for a typical 80 kVp CT scan. The sinogram was then corrected for 511 keV  $\gamma$ -rays using a three-step correction method proposed by Abella in Schmitt *et al* (1988). Finally,  $P_{\text{det. sens}}$  is also a diagonal matrix that contains the detector efficiencies for normalization. Here, we obtained a 3D normalization sinogram of the GE scanner and incorporated it into our model.

## Appendix B. Scaling factors used to generate generalized PSF modelled kernels

This appendix contains table B1, which includes the three scaling factors multiplied by the attenuation coefficient of crystals, as well as FWHMs of non-collinearity and crystal scattering used in generating 20 generalized PSF modelled kernels presented in this paper. The column on the right contains the factor we multiplied by the attenuation coefficient of LYSO crystal that is 0.087. The second to the right column includes the factor we used to rescale the FWHM of photon non-collinearity in equation (A.3). The third column contains the FWHM of inter-crystal scattering effect being used in equation (A.6).

## Appendix C. Signal and noise FOMs for quantitative analysis

We used a number of figures of merit (FOMs) to quantify the performance of generalized PSF modeling kernels in our study, and in this appendix, we define these metrics and elaborate on why each FOM is required for studying different aspects of image quantitation.

We denote the total number of independent noise realizations by  $R$ , the total number of image voxels by  $N$ , and the total number of voxels in a (tumour) region-of-interest (ROI) by  $M$ . If  $\hat{v}_i$  and  $v_i^r$  refer to the true and reconstructed uptake values of the  $i$ th voxel ( $i = 1, \dots, N$ ) for the  $r$ th noise realization, respectively, then: (i)  $\bar{v}^r = \text{SUV}_{\text{mean}}^r = \frac{1}{M} \sum_{i=1}^M v_i^r$  represents the spatial mean of voxels across an ROI at noise realization  $r$ , (ii)  $\bar{v}_i = \frac{1}{R} \sum_{r=1}^R v_i^r$  represents the mean of a given voxel  $i$ , averaged temporally across all noise realizations, and (iii)

**Table B1.** List of scaling factors used to generate 20 PSF modelled kernel.

	Kernel index	FWHM of inter-crystal scattering	Factors multiplied by the FWHM of photon non-collinearity	Factors multiplied by the attenuation coefficient of crystals
No-PSF	1	0.04	0	111.1
Underestimated PSF	2	0.392	0.091	44.534
	3	0.785	0.182	18.746
	4	1.438	0.409	6.983
	5	1.962	0.545	2.826
	6	2.354	0.682	1.319
	7	2.616	0.773	1.240
	8	2.877	0.818	1.156
	9	3.139	0.909	1.080
	10	3.27	1	1
Overestimated PSF	11	3.662	1.091	0.919
	12	3.924	1.182	0.840
	13	4.185	1.227	0.76
	14	4.447	1.318	0.68
	15	4.709	1.409	0.6
	16	4.970	1.5	0.519
	17	5.232	1.545	0.440
	18	5.493	1.636	0.360
	19	5.755	1.727	0.280
	20	5.886	1.773	0.240

$\bar{v} = \frac{1}{R} \sum_{r=1}^R \bar{v}^r$  is the ensemble mean of all  $SUV_{\text{mean}}^r$  values across  $R$  noise realizations. Since each ROI is set to have a uniform uptake in the true image, we denote  $\hat{v}_i = \hat{v}$  to represent the value of all the voxels inside that region. Several noise and bias metrics can be derived, as in the subsections below.

### C.1. Contrast recovery curves (CRC)

We define two types of CRC:  $SUV_{\text{mean}}$  CRC and  $SUV_{\text{max}}$  CRC. The former is defined as the ratio between the contrasts of the ROI (defined as the mean of the region) in the reconstructed image and the true object in terms of the ROI average. The latter would be the same ratio but in terms of the ROI maximum voxel.

### C.2. Bias in $SUV_{\text{mean}}$ and $SUV_{\text{max}}$

Biases in activity uptake quantitation were defined in terms of  $SUV_{\text{mean}}$  and  $SUV_{\text{max}}$ , calculated as follows (Tong *et al* 2010, Rahmim and Tang 2013b):

$$SUV_{\text{mean}} \text{ Bias} = \frac{1}{R} \sum_{r=1}^R (SUV_{\text{mean}}^r - \hat{v}) \quad (\text{C.7})$$

$$SUV_{\text{max}} \text{ Bias} = \frac{1}{R} \sum_{r=1}^R (SUV_{\text{max}}^r - \hat{v}) \quad (\text{C.8})$$

where  $\text{SUV}_{\max}^r = \max_{i \in \{1 \dots M\}} \{v_i^r\}$  represents the voxel with the maximum uptake value inside the tumour ROI of realization  $r$ .

Equations (C.7) and (C.8) show that  $\text{SUV}_{\text{mean}}$  and  $\text{SUV}_{\max}$  biases are basically defined as the deviation of the reconstructed image from the *true* value of the object. Both biases are then normalized to the true value for plots in section 3.4. It is worth noting that studies of quantitative task performance using clinical patient data would not have this level of rigor in determining the bias (given typical lack of access to the true value of every voxel).

### C.3. Coefficient of variability (CoV) of $\text{SUV}_{\text{mean}}$ and $\text{SUV}_{\max}$

We characterized the CoV for both  $\text{SUV}_{\text{mean}}$  and  $\text{SUV}_{\max}$ . The former was defined as follows:

$$\text{SUV}_{\text{mean}} \text{ CoV} = \left( \frac{1}{R-1} \sum_{r=1}^R (\text{SUV}_{\text{mean}}^r - \hat{v})^2 \right)^{\frac{1}{2}} \quad (\text{C.9})$$

where  $\hat{v}$  was defined at the beginning of this appendix. Similarly, the maximum uptake CoV was defined as the variability of the maximum voxel of selected ROI across all realizations, and can be calculated as:

$$\text{SUV}_{\max} \text{ CoV} = \left( \frac{1}{R-1} \sum_{r=1}^R (\text{SUV}_{\max}^r - \overline{\text{SUV}_{\max}})^2 \right)^{\frac{1}{2}} \quad (\text{C.10})$$

where  $\overline{\text{SUV}_{\max}}$  denotes the  $\text{SUV}_{\max}$  values averaged across the noise realizations. Both equations will be normalized to  $\hat{v}$  in plots of section 3.4 for more proper comparison.

### C.4. Mean-squared error (MSE)

MSE is a more general metric that combines noise and bias of voxels within a single quantity. It can be averaged over all voxels within an ROI to represent the MSE of the ROI, and then for the  $R$  realizations, the mean of all MSEs would be calculated:

$$\text{MSE} = \frac{1}{RM} \sum_{r=1}^R \sum_{i=1}^M (v_i^r - \hat{v})^2. \quad (\text{C.11})$$

Due to the importance of both the accuracy and precision of  $\text{SUV}_{\text{mean}}$  value in clinical practices and to better study the effect of different reconstruction kernels to the mean uptake, we defined the next FOM as the MSE of the mean uptake. MSE can also be calculated by summing the squared noise and squared bias (Tong *et al* 2010). The MSE of  $\text{SUV}_{\text{mean}}$  was defined as:

$$\text{MSE of } \text{SUV}_{\text{mean}} = \text{SUV}_{\text{mean}} \text{ Bias}^2 + \text{SUV}_{\text{mean}} \text{ CoV}^2. \quad (\text{C.12})$$

### C.5. Image roughness (spatial noise)

Image roughness (IR) measures the voxel by voxel variability in the image and can be calculated even for a single realization. Within a given ROI containing  $M$  voxels, image roughness was defined as the variability of the voxel values with respect to  $\text{SUV}_{\text{mean}}$ . This was then averaged over  $R$  noise realizations:

$$\sigma_{\text{spatial}} = \frac{1}{R} \sum_{r=1}^R \left( \frac{1}{M-1} \sum_{i=1}^M (v_i^r - \bar{v}^r)^2 \right)^{\frac{1}{2}} \quad (\text{C.13})$$

where  $\bar{v}^r$  is the mean of all voxels,  $v_i^r$ 's, inside the given ROI of realization  $r$ . The noise values plotted in the results section were normalized to  $\bar{v}^r$ .

#### C.6. Voxel variation ( $\sigma_0$ )

This metric provides a measure of the variability of individual voxels over multiple noise realizations:

$$\sigma_0 = \frac{1}{M} \sum_{i=1}^M \left( \frac{1}{R-1} \sum_{r=1}^R (v_i^r - \bar{v}_i)^2 \right)^{\frac{1}{2}}. \quad (\text{C.14})$$

Voxel variation impacts both image roughness and  $\text{SUV}_{\text{mean}}$  CoV (Tong *et al* 2010). Assume a uniform region consisting of  $M$  voxels with voxel variance  $\sigma_0^2$  for each voxel and inter-voxel covariance  $\text{cov}(i,j)$  between two voxels  $i$  and  $j$ . Tong *et al* showed that the expectations of image roughness (equation (C.13)) and  $\text{SUV}_{\text{mean}}$  CoV (equation (C.9)) is given by:

$$E[\sigma_{\text{spatial}}^2] = \sigma_0^2 - \frac{1}{(M-1)M} \sum_{i \neq j} \text{cov}(i,j) = \sigma_0^2 - \frac{2}{(M-1)M} \sum_{i > j} \text{cov}(i,j) \quad (\text{C.15})$$

and

$$E[\{\text{SUV}_{\text{mean}} \text{CoV}\}^2] = \frac{\sigma_0^2}{M} + \frac{1}{M^2} \sum_{i \neq j} \text{cov}(i,j) = \frac{\sigma_0^2}{M} + \frac{2}{M^2} \sum_{i > j} \text{cov}(i,j) \quad (\text{C.16})$$

In equation (C.15), it is seen that the reduced voxel variance and increased inter-voxel covariance due to PSF modeling result in overall reduction in image roughness. In equation (C.16) however, these two work against one another and the increasing (positive) covariance actually contributes positively to  $\text{SUV}_{\text{mean}}$  CoV.

#### C.7. Averaged differences of max and min uptake

To better quantify edge effects, we assessed the range of uptake within the ROI after reconstruction. For a total of  $R$  realizations, the average max-min difference was calculated as follows:

$$\text{Averaged Max - Min difference} = \frac{1}{R} \sum_{r=1}^R (\text{SUV}_{\text{max}}^r - \text{SUV}_{\text{min}}^r) \quad (\text{C.17})$$

where  $\text{SUV}_{\text{min}}^r = \min_m \{v_r\}$ , referring to the lowest uptake values within the ROI of the  $r$ th realization. The result was then averaged over all realizations. For plotting purposes in Results section 3.4.4, we normalized this measure to  $\bar{v}$ . This subtraction of the minimum undershoot from the maximum overshoot of every reconstructed ROI evaluates the deviation from the true flat uptake due to noise or false texture (including edge effects). Then, averaging its value over all noise realizations essentially provides a FOM for *shape*.

## References

- Aerts H J W L *et al* 2014 Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach *Nat. Commun.* **5** 4006
- Alessio A and MacDonald L 2008 Spatially variant positron range modeling derived from CT for PET image reconstruction *Nuclear Science Symp. Conf. Record* (IEEE) pp 3637–40
- Alessio A, Rahmim A and Orton C G 2013 Resolution modeling enhances PET imaging (point/counterpoint) *Med. Phys.* **40** 120601
- Alessio A M, Kinahan P E and Lewellen T K 2006 Modeling and incorporation of system response functions in 3D whole body PET *IEEE Trans. Med. Imaging* **25** 828–37
- Alessio A M, Stearns C W, Shan T, Ross S G, Kohlmyer S, Ganin A and Kinahan P E 2010 Application and evaluation of a measured spatially variant system model for PET image reconstruction *IEEE Trans. Med. Imaging* **29** 938–49
- Antich P, Parkey R, Selouanine S, Slavine N, Tsyanov E and Zinchenko A 2005 Application of expectation maximization algorithms for image resolution improvement in a small animal PET system *IEEE Trans. Nucl. Sci.* **52** 684–90
- Ashrafinia S, Karakatsanis N, Mohy-ud-Din H and Rahmim A 2014a Towards continualized task-based resolution modeling in PET imaging *Proc. SPIE* **9033** 903327
- Ashrafinia S, Mohy-ud-Din H, Karakatsanis N, Kadrmas D and Rahmim A 2014b Enhanced quantitative PET imaging utilizing adaptive partial resolution modeling *J. Nucl. Med.* **55** (Suppl. 1) 371
- Ashrafinia S, Mena E, Mohy-ud-Din H, Jha A, Subramaniam R and Rahmim A 2016a Adaptive PSF modeling for enhanced heterogeneity quantification in oncologic PET imaging *J. Nucl. Med.* **57** (Suppl. 2) 479
- Ashrafinia S, Mohy-ud-Din H, Karakatsanis N, Casey M, Lodge M and Rahmim A 2016b PSF overestimation improves PET image SUV quantification *J. Nucl. Med.* **57** 1965
- Asselin M C, O'Connor J P B, Boellaard R, Thacker N A and Jackson A 2012 Quantifying heterogeneity in human tumours using MRI and PET *Eur. J. Cancer* **48** 447–55
- Bai B, Laforest R, Smith A M and Leahy R M 2005 Evaluation of MAP image reconstruction with positron range modeling for 3D PET *Nuclear Science Symp. Conf. Record* (IEEE) pp 2686–9
- Bai B, Ruangma A, Laforest R, Tai Y C and Leahy R M 2003 Positron range modeling for statistical PET image reconstruction *Nuclear Science Symp. Conf. Record* (IEEE) pp 2501–5
- Barrett H H, Wilson D W and Tsui B M W 1994 Noise properties of the EM algorithm. I. Theory *Phys. Med. Biol.* **39** 833–46
- Bernardi E D, Mazzoli M, Zito F and Baselli G 2007 Resolution recovery in PET during AWOSEM reconstruction: a performance evaluation study *IEEE Trans. Nucl. Sci.* **54** 1626–38
- Blinder S A, Dinelle K and Sossi V 2012 Scanning rats on the high resolution research tomograph (HRRT): a comparison study with a dedicated micro-PET *Med. Phys.* **39** 5073–83
- Boellaard R, Krak N C, Hoekstra O S and Lammertsma A A 2004 Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study *J. Nucl. Med.* **45** 1519–27
- Carson R E, Barker W C, Liow J-S and Johnson C A 2003 Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT *IEEE Nuclear Science Symp. Conf. Record (Portland, OR)* pp 3281–5
- Cherry S R, Sorenson J A and Phelps M E 2012 *Physics in Nuclear Medicine* (Amsterdam: Elsevier)
- Chicklore S, Goh V, Siddique M, Roy A, Marsden P K and Cook G J R 2013 Quantifying tumour heterogeneity in F-18-FDG PET/CT imaging by texture analysis *Eur. J. Nucl. Med. Mol. I* **40** 133–40
- Cloquet C, Sureau F C, Defrise M, Van Simaeys G, Trotta N and Goldman S 2010 Non-Gaussian space-variant resolution modelling for list-mode reconstruction *Phys. Med. Biol.* **55** 5045–66
- Dimitrakopoulou-Strauss A, Strauss L G, Heichel T, Wu H, Burger C, Bernd L and Ewerbeck V 2002 The role of quantitative <sup>18</sup>F-FDG PET studies for the differentiation of malignant and benign bone lesions *J. Nucl. Med.* **43** 510–8
- Eary J F, O'Sullivan F, O'Sullivan J and Conrad E U 2008 Spatial heterogeneity in sarcoma <sup>18</sup>F-FDG uptake as a predictor of patient outcome *J. Nucl. Med.* **49** 1973–9
- El Naqa I *et al* 2009 Exploring feature-based approaches in PET images for predicting cancer treatment outcomes *Pattern Recognit.* **42** 1162–71
- Erlandsson K, Buvat I, Pretorius P H, Thomas B A and Hutton B F 2012 A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology *Phys. Med. Biol.* **57** R119–59

- Feng D, Huang S C and Wang X 1993 Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography *Int. J. Biomed. Comput.* **32** 95–110
- Hatt M D *et al* 2015 F-18-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort *J. Nucl. Med.* **56** 38–44
- Hatt M, Tixier F, Pierce L, Kinahan P E, Le Rest C C and Visvikis D 2016 Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur. J. Nucl. Med. Mol. I.* **44** 151–65
- Hoetjes N J, van Velden F H P, Hoekstra O S, Hoekstra C J, Krak N C, Lammertsma A A and Boellaard R 2010 Partial volume correction strategies for quantitative FDG PET in oncology *Eur. J. Nucl. Med. Mol. I.* **37** 1679–87
- Iriarte A, Marabini R, Matej S, Sorzano C O S and Lewitt R M 2016 System models for PET statistical iterative reconstruction: a review *Comput. Med. Imaging Graph.* **48** 30–48
- Jakoby B W, Bercier Y, Conti M, Casey M E, Bendriem B and Townsend D W 2011 Physical and clinical performance of the mCT time-of-flight PET/CT scanner *Phys. Med. Biol.* **56** 2375–89
- Kadrimas D J, Casey M E, Black N F, Hamill J J, Panin V Y and Conti M 2009 Experimental comparison of lesion detectability for four fully-3D PET reconstruction schemes *IEEE Trans. Med. Imaging* **28** 523–34
- Karakatsanis N A, Lodge M A, Rahmim A and Zaidi H 2014 Introducing time-of-flight and resolution recovery image reconstruction to clinical whole-body PET parametric imaging *IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSS/MIC)* (<https://doi.org/10.1109/NSSMIC.2014.7430771>)
- Karakatsanis N, Mehranian A, Lodge M, Casey M, Rahmim A and Zaidi H 2015 Clinical evaluation of direct 4D whole-body PET parametric imaging with time-of-flight and resolution modeling capabilities *IEEE Nuclear Science Symp. Conf. Record*
- Kemp B J, Kim C, Williams J J, Ganin A and Lowe V J 2006 NEMA NU 2-2001 performance measurements of an LYSO-based PET/CT system in 2D and 3D acquisition modes *J. Nucl. Med.* **47** 1960–7
- Kotasidis F A, Angelis G I, Anton-Rodriguez J, Markiewicz P, Lionheart W R, Reader A J and Matthews J C 2014 Image-based spatially variant and count rate dependent point spread function on the HRRT *IEEE Trans. Nucl. Sci.* **61** 1192–202
- Kotasidis F A, Matthews J C, Angelis G I, Noonan P J, Jackson A, Price P, Lionheart W R and Reader A J 2011 Single scan parameterization of space-variant point spread functions in image space via a printed array: the impact for two PET/CT scanners *Phys. Med. Biol.* **56** 2917–42
- Kumar V *et al* 2012 Radiomics: the process and the challenges *Magn. Reson. Imaging* **30** 1234–48
- Lambin P *et al* 2012 Radiomics: extracting more information from medical images using advanced feature analysis *Eur. J. Cancer* **48** 441–6
- Le Meunier L, Slomka P J, Dey D, Ramesh A, Thomson L E J, Hayes S W, Friedman J D, Cheng V, Germano G and Berman D S 2010 Enhanced definition PET for cardiac imaging *J. Nucl. Cardiol.* **17** 414–26
- Lecomte R, Schmitt D and Lamoureux G 1984 Geometry study of a high resolution PET detection system using small detectors *IEEE Trans. Nucl. Sci.* **31** 556–61
- Lu L, Lv W, Jiang J, Ma J, Feng Q, Rahmim A and Chen W 2016 Robustness of radiomic features in [11C]Choline and [18F]FDG PET/CT imaging of nasopharyngeal carcinoma: impact of segmentation and discretization *Mol. Imaging Biol.* **18** 935–45
- Miller M, Zhang J, Binzel K, Griesmer J, Laurence T, Narayanan M, Natarajamani D, Wang S and Knopp M 2015 Characterization of the vereos digital photon counting PET system *J. Nucl. Med.* **56** 434
- Muller-Gartner H W, Links J M, Prince J L, Bryan R N, McVeigh T, Leal J P, Davatzikos C and Frost J J 1992 Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects *J. Cereb. Blood Flow Metab.* **12** 571–83
- Mumcuoglu E U, Leahy R M, Cherry S R and Hoffman E 1996 Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET *Conf. Record IEEE Nuclear Science Symp.* vol 3 pp 1569–73
- NEMA 2012 NEMA NU 2-2012 Performance measurement of positron emission tomographs *National Electrical Manufacturers Association Standards Publication*
- Niu X, Asma E, Ye H, Wang W and Gagnon D 2015 *IEEE Nuclear Science Symp. & Medical Imaging Conf., Conf. Record (San Diego, CA, 31 October–7 November 2015)*

- Okazumi S, Isono K, Enomoto K, Kikuchi T, Ozaki M, Yamamoto H, Hayashi H, Asano T and Ryu M 1992 Evaluation of liver tumors using fluorine-18-fluorodeoxyglucose PET: characterization of tumor and assessment of effect of treatment *J. Nucl. Med.* **33** 333–9
- Panin V Y, Kehren F, Michel C and Casey M 2006a Fully 3D PET reconstruction with system matrix derived from point source measurements *IEEE Trans. Med. Imaging* **25** 907–21
- Panin V Y, Kehren F, Rothfuss H, Hu D, Michel C and Casey M E 2006b PET reconstruction with system matrix derived from point source measurements *IEEE Trans. Nucl. Sci.* **53** 152–9
- Politte D G and Snyder D L 1988 The use of constraints to eliminate artifacts in maximum-likelihood image estimation for emission tomography *IEEE Trans. Nucl. Sci.* **35** 608–10
- Qi J, Leahy R M, Cherry S R, Chatzioannou A and Farquhar T H 1998a High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner *Phys. Med. Biol.* **43** 1001–13
- Qi J, Leahy R M, Chinghan H, Farquhar T H and Cherry S R 1998b Fully 3D Bayesian image reconstruction for the ECAT EXACT HR + *IEEE Trans. Nucl. Sci.* **45** 1096–103
- Qi J, Zhou J, Ye H and Wang W 2015 Image-based point-spread-function modelling in time-of-flight positron-emission-tomography iterative list-mode reconstruction *Google Patents WO 2014066629 A1* “/patents/US20150199302?cl=en”
- Rahmim A, Cheng J-C and Sossi V 2005 Improved noise propagation in statistical image reconstruction with resolution modeling *Nuclear Science Symp. Conf. Record* (IEEE) pp 2576–8
- Rahmim A, Lenox M, Michel C, Reader A J and Sossi V 2003 Space-variant and anisotropic resolution modeling in list-mode EM reconstruction *IEEE Nuclear Science Symp. Conf. Record* vol 5 pp 3074–7
- Rahmim A, Qi J and Sossi V 2013 Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls *Med. Phys.* **40** 064301
- Rahmim A, Schmidlein C R, Jackson A, Sheikbahaei S, Marcus C, Ashrafinia S, Soltani M and Subramaniam R M 2016 A novel metric for quantification of homogeneous and heterogeneous tumors in PET for enhanced clinical outcome prediction *Phys. Med. Biol.* **61** 227–42
- Rahmim A and Tang J 2013a Impact of resolution modeling on detectability: analysis in terms of noise-equivalent quanta *J. Nucl. Med.* **54** 2119
- Rahmim A and Tang J 2013b Noise propagation in resolution modeled PET imaging and its impact on detectability *Phys. Med. Biol.* **58** 6945
- Rahmim A, Tang J, Ay M R and Bengel F 2010 4D Respiratory motion-corrected Rb-82 myocardial perfusion PET imaging *IEEE Nuclear Science Symp. Conf. Record* pp 3312–6
- Rahmim A, Tang J, Lodge M, Lashkari S, Ay M, Lautamaki R, Tsui B and Bengel F 2008a Improved clinical and quantitative dynamic Rb-82 cardiac imaging with resolution modeling *J. Nucl. Med.* **49** 62P
- Rahmim A, Tang J, Lodge M A, Lashkari S, Ay M R and Bengel F M 2008b Resolution modeled PET image reconstruction incorporating space-variance of positron range: rubidium-82 cardiac PET imaging *IEEE Nuclear Science Symp. Conf. Record* pp 3643–50
- Rahmim A, Tang J, Lodge M A, Lashkari S, Ay M R, Lautamaki R, Tsui B M and Bengel F M 2008c Analytic system matrix resolution modeling in PET: an application to Rb-82 cardiac imaging *Phys. Med. Biol.* **53** 5947–65
- Rahmim A and Zaidi H 2008 PET versus SPECT: strengths, limitations and challenges *Nucl. Med. Commun.* **29** 193–207
- Rapisarda E, Bettinardi V, Thielemans K and Gilardi M C 2010 Image-based point spread function implementation in a fully 3D OSEM reconstruction algorithm for PET *Phys. Med. Biol.* **55** 4131–51
- Reader A J, Julyan P J, Williams H, Hastings D L and Zweit J 2003 EM algorithm system modeling by image-space techniques for PET reconstruction *IEEE Trans. Nucl. Sci.* **50** 1392–7
- Rousset O G, Collins D L, Rahmim A and Wong D F 2008 Design and implementation of an automated partial volume correction in PET: application to dopamine receptor quantification in the normal human striatum *J. Nucl. Med.* **49** 1097–106
- Rousset O G, Ma Y and Evans A C 1998 Correction for partial volume effects in PET: principle and validation *J. Nucl. Med.* **39** 904–11
- Rousset O G, Rahmim A, Alavi A and Zaidi H 2007 Partial volume correction strategies in PET *PET Clin.* **2** 235–49
- Ruangma A, Bai B, Lewis J S, Sun X, Welch M J, Leahy R and Laforest R 2006 Three-dimensional maximum *a posteriori* (MAP) imaging with radiopharmaceuticals labeled with three Cu radionuclides *Nucl. Med. Biol.* **33** 217–26
- Schmitt D, Karuta B, Carrier C and Lecomte R 1988 Fast point spread function computation from aperture functions in high-resolution positron emission tomography *IEEE Trans. Med. Imaging* **7** 2–12

- Segars W P, Sturgeon G, Mendonca S, Grimes J and Tsui B M W 2010 4D XCAT phantom for multimodality imaging research *Med. Phys.* **37** 4902–15
- Selivanov V V, Picard Y, Cadorette J, Rodrigue S and Lecomte R 2000 Detector response models for statistical iterative image reconstruction in high resolution PET *IEEE Trans. Nucl. Sci.* **47** 1168–75
- Slomka P J, Pan T and Germano G 2016 Recent advances and future progress in PET instrumentation *Semin. Nucl. Med.* **46** 5–19
- Snyder D L, Miller M I, Thomas L J and Politte D G 1987 Noise and edge artifacts in maximum-likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **MI-6** 313–9
- Soares E J, Byrne C L and Glick S J 2000 Noise characterization of block-iterative reconstruction algorithms: I. Theory *IEEE Trans. Med. Imaging* **19** 261–70
- Soares E J, Glick S J and Hoppin J W 2005 Noise characterization of block-iterative reconstruction algorithms: II. Monte Carlo simulations *IEEE Trans. Med. Imaging* **24** 112–21
- Soret M, Bacharach S L and Buvat I 2007 Partial-volume effect in PET tumor imaging *J. Nucl. Med.* **48** 932–45
- Soufi M, Kamali-Asl A, Geramifar P and Rahmim A 2016 A novel framework for automated segmentation and labeling of homogeneous versus heterogeneous lung tumors in [<sup>18</sup>F]FDG PET imaging *Mol. Imaging Biol.* **2016** 1–13
- Strul D, Slates R B, Dahlbom M, Cherry S R and Marsden P K 2003 An improved analytical detector response function model for multilayer small-diameter PET scanners *Phys. Med. Biol.* **48** 979–94
- Sureau F C, Reader A J, Comtat C, Leroy C, Ribeiro M-J, Buvat I and Trebossen R 2008 Impact of image-space resolution modeling for studies with the high-resolution research tomograph *J. Nucl. Med.* **49** 1000–8
- Takahashi H, Yamaya T, Kobayashi T, Kitamura K, Hasegawa T, Murayama H and Suga M 2007 *IEEE Nuclear Science Symp. and Medical Imaging Conf. (Honolulu, HI, 2007)* series 5 pp 3478–81
- Teo B K, Seo Y, Bacharach S L, Carrasquillo J A, Libutti S K, Shukla H, Hasegawa B H, Hawkins R A and Franc B L 2007 Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data *J. Nucl. Med.* **48** 802–10
- Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, Ingrand P, Perdrisot R, Visvikis D and Le Rest C C 2014 Visual versus quantitative assessment of intratumor <sup>18</sup>F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer *J. Nucl. Med.* **55** 1235–41
- Tixier F, Le Rest C C, Hatt M, Albarghach N, Pradier O, Metges J P, Corcos L and Visvikis D 2011 Intratumor heterogeneity characterized by textural features on baseline <sup>18</sup>F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer *J. Nucl. Med.* **52** 369–78
- Tong S, Alessio A M and Kinahan P E 2010 Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation *Phys. Med. Biol.* **55** 1453–73
- Tong S, Alessio A M, Thielemans K, Stearns C, Ross S and Kinahan P E 2011 Properties and mitigation of edge artifacts in PSF-based PET reconstruction *IEEE Trans. Nucl. Sci.* **58** 2264–75
- Torizuka T, Nobezawa S, Momiki S, Kasamatsu N, Kanno T, Yoshikawa E, Futatsubashi M, Okada H and Ouchi Y 2000 Short dynamic FDG-PET imaging protocol for patients with lung cancer *Eur. J. Nucl. Med.* **27** 1538–42
- van Velden F, Kramer G, Frings V, Nissen I, Mulder E, Hoekstra O, Smit E and Boellaard R 2015 Impact of reconstruction algorithms on the repeatability of various PET/CT radiomics features in lung cancer patient *J. Nucl. Med.* **56** 428
- van Velden F H P, Cheebsumon P, Yaqub M, Smit E F, Hoekstra O S, Lammertsma A A and Boellaard R 2011 Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies *Eur. J. Nucl. Med. Mol. Biol.* **38** 1636–47
- Vercher-Conejero J, Sher A, Kolthammer J, Kohan A, Rubbert C, Wojtylak P, Partovi S, Gaeta M and O'Donnell J 2013 Contribution of the modeling of point spread function and time-of-flight technology on the image quality of a PET/CT system *J. Nucl. Med.* **54** 101
- Vriens D, Disselhorst J A, Oyen W J G, de Geus-Oei L F and Visser E P 2012 Quantitative assessment of heterogeneity in tumor metabolism using FDG-PET *Int. J. Radiat. Oncol.* **82** E725–31
- Wilson D W, Tsui B M W and Barrett H H 1994 Noise properties of the EM algorithm. II. Monte Carlo simulations *Phys. Med. Biol.* **39** 847–71
- Yamaya T et al 2005 Transaxial system models for jPET-D4 image reconstruction *Phys. Med. Biol.* **50** 5339–55