

---

---

# Impact of Image Reconstruction Settings on Texture Features in $^{18}\text{F}$ -FDG PET

Jianhua Yan<sup>1</sup>, Jason Lim Chu-Shern<sup>1</sup>, Hoi Yin Lo<sup>2</sup>, Lih Kin Khor<sup>2</sup>, Arvind K. Sinha<sup>2</sup>, Swee Tian Quek<sup>2</sup>, Ivan W.K. Tham<sup>1,3</sup>, and David Townsend<sup>1,2</sup>

<sup>1</sup>A\*STAR-NUS, Clinical Imaging Research Center, Singapore; <sup>2</sup>Department of Diagnostic Radiology, National University Hospital, Singapore; and <sup>3</sup>Department of Radiation Oncology, National University Cancer Institute, Singapore

Evaluation of tumor heterogeneity based on texture parameters has recently attracted much interest in the PET imaging community. However, the impact of reconstruction settings on texture parameters is unclear, especially relating to time-of-flight and point-spread function modeling. Their effects on 55 texture features (TFs) and 6 features based on first-order statistics (FOS) were investigated. Standardized uptake value (SUV) measures were also evaluated as peak SUV ( $\text{SUV}_{\text{peak}}$ ), maximum SUV, and mean SUV ( $\text{SUV}_{\text{mean}}$ ). **Methods:** This study retrospectively recruited 20 patients with lesions in the lung who underwent whole-body  $^{18}\text{F}$ -FDG PET/CT. The coefficient of variation (COV) of each feature across different reconstructions was calculated. **Results:**  $\text{SUV}_{\text{peak}}$ ,  $\text{SUV}_{\text{mean}}$ , 18 TFs, and 1 FOS were the most robust ( $\text{COV} \leq 5\%$ ) whereas skewness, cluster shade, and zone percentage were the least robust ( $\text{COV} > 20\%$ ) with respect to reconstruction algorithms using default settings. Heterogeneity parameters had different sensitivities to iteration number. Twenty-four parameters including  $\text{SUV}_{\text{peak}}$  and  $\text{SUV}_{\text{mean}}$  exhibited variation with a COV less than 5%; 28 parameters, including maximum SUV, showed variation with a COV in the range of 5%–10%. In addition, skewness, cluster shade, and zone percentage were the most sensitive to iteration number. In terms of sensitivity to full width at half maximum (FWHM), 15 TFs and 1 FOS had the best performance with a COV less than 5%, whereas  $\text{SUV}_{\text{peak}}$  and  $\text{SUV}_{\text{mean}}$  had a COV between 5% and 10%. Grid size had the largest impact on image features, which was demonstrated by only 11 features, including  $\text{SUV}_{\text{peak}}$  and  $\text{SUV}_{\text{mean}}$ , having a COV less than 10%. **Conclusion:** Different image features have different sensitivities to reconstruction settings. Iteration number and FWHM of the gaussian filter have a similar impact on the image features. Grid size has a larger impact on the features than iteration number and FWHM. The features that exhibited large variations such as skewness in FOS, cluster shade, and zone percentage should be used with caution. The entropy in FOS, difference entropy, inverse difference normalized, inverse difference moment normalized, low gray-level run emphasis, high gray-level run emphasis, and low gray-level zone emphasis are the most robust features.

**Key Words:** tumor texture; time-of-flight; point-spread function;  $^{18}\text{F}$ -FDG PET

**J Nucl Med 2015; 56:1667–1673**

DOI: 10.2967/jnumed.115.156927

**R**ecent studies have shown that tumors often display startling intratumoral heterogeneity, which is often associated with adverse tumor biology (1,2). Unfortunately, it is difficult to assess intratumoral heterogeneity with random sampling or biopsy as this does not represent the full extent of phenotypic or genetic variation within a tumor. Given the limitations of current biopsy strategies, there is an important potential for medical imaging, which has the ability to capture intratumoral heterogeneity in a noninvasive way.

In oncology,  $^{18}\text{F}$ -FDG PET is currently playing a major role in clinical diagnosis, staging, prognosis, and assessment of response to treatment (3). The use of standardized uptake value (SUV) is now routine in clinical  $^{18}\text{F}$ -FDG PET/CT oncology imaging for quantifying glucose metabolic activity in tissues. Maximum SUV ( $\text{SUV}_{\text{max}}$ ), mean SUV ( $\text{SUV}_{\text{mean}}$ ), and peak SUV ( $\text{SUV}_{\text{peak}}$ ) are the commonly used metrics (4). However, none of these SUV measures reflect the underlying spatial distribution of  $^{18}\text{F}$ -FDG. There is now growing interest in using texture analysis to assess tumor heterogeneity in the field of oncologic imaging.

The first attempt to use the texture features (TFs) in PET imaging was to predict treatment outcome in cervical and head and neck cancer (5). Subsequently, there are increasing numbers of publications related to the application of TFs in PET tumor imaging, mostly with  $^{18}\text{F}$ -FDG (5–21). These previous studies mainly focused on the predictive and prognostic value (5–10,13–25) of TF and also the potential use in radiotherapy planning (20,21). The test–retest reproducibility of TF from  $^{18}\text{F}$ -FDG PET was shown to be comparable to SUV (16). In addition, Galavis et al. reported the variability of TF as a function of different acquisition methods, reconstruction parameters, and postreconstruction filtering parameters (11). However, only filtered back-projection and ordered-subset expectation maximization (OSEM) were investigated. The techniques of time-of-flight (TOF) and point-spread function (PSF) modeling have been shown to clearly improve image quality in terms of signal-to-noise ratio (SNR), lesion contrast, and coefficient of variation (COV) on background and are now standard in most scanners (22–26). TOF improves SNR and thus reduces heterogeneity due to noise. PSF modeling typically leads to improved resolution by modeling in the system

---

Received Mar. 2, 2015; revision accepted Jul. 6, 2015.

For correspondence or reprints contact: Jianhua Yan, National University of Singapore, Center for Translation Medicine, 14 Medical Dr. #B1-01, Singapore 117599.

E-mail: yan\_jianhua@circ.a-star.edu.sg

Published online Jul. 30, 2015.

COPYRIGHT © 2015 by the Society of Nuclear Medicine and Molecular Imaging, Inc.

matrix physical processes that degrade image resolution and thus improves evaluation of tumor heterogeneity by delineating higher definition structure within a lesion.

In this context, the objective of this study was to investigate variability of TF in  $^{18}\text{F}$ -FDG PET images due to different reconstruction settings including postreconstruction gaussian filters with different full width at half maximum (FWHM), different iteration numbers, and voxel size. The filtered backprojection algorithm was not included in this study, because it has been adequately addressed in a previous study (11).

## MATERIALS AND METHODS

### Patients

This retrospective study was approved by the Domain Specific Review Board and Institutional Review Board of National Health Care Group, Singapore, and the requirement to obtain informed consent was waived. Patients with lesions in the lung who underwent whole-body  $^{18}\text{F}$ -FDG PET/CT at the National University Hospital, Singapore, from February 2013 to May 2014 were retrospectively enrolled while excluding those patients with a lesion volume less than  $5\text{ cm}^3$  (14). Only the lesions located in the lung were included for analysis. Finally, this study included 17 patients with non-small cell lung cancer (mean age, 70 y; age range, 50–82 y; 12 men, 5 women), 1 female patient with nasopharyngeal cancer (age, 71 y), and 2 female patients with lymphoma (age, 60 and 75 y). Clinical staging for the non-small cell lung cancer cohort was as follows: stage 1A ( $n = 1$ ), stage 2B ( $n = 1$ ), stage 3A ( $n = 1$ ), stage 3B ( $n = 4$ ), and stage 4 ( $n = 10$ ). The patients with nasopharyngeal cancer and lymphoma were found to be stage 4.

### Data Acquisition

All patients underwent a whole-body  $^{18}\text{F}$ -FDG PET/CT using a Biograph 64 mCT scanner (Siemens). Patients fasted for at least 8 h to ensure a serum glucose level less than 10 mmol/L. The time difference between injection and acquisition was  $71 \pm 9$  min (range, 52–88 min) after injection of  $229.4 \pm 22.2$  MBq ( $6.2 \pm 0.6$  mCi; range, 181.3–270.1 MBq [4.9–7.3 mCi]) of  $^{18}\text{F}$ -FDG. PET raw data were acquired for 1 min per bed position with the exception of the bed position covering the liver, which was acquired for 3 min. CT was

performed with a tube voltage of 120 kV and a tube current of 50 mAs.

### Image Reconstruction

Four investigations listed in Table 1 were performed to explore the impact of reconstruction settings on image features. The numbers of subsets for reconstructions without TOF and with TOF were 24 and 21, respectively. The image slice thickness was 5 mm. In total, 96 images were produced for each patient. The PET images were corrected for attenuation and scatter based on the CT scan.

### Tumor Segmentation

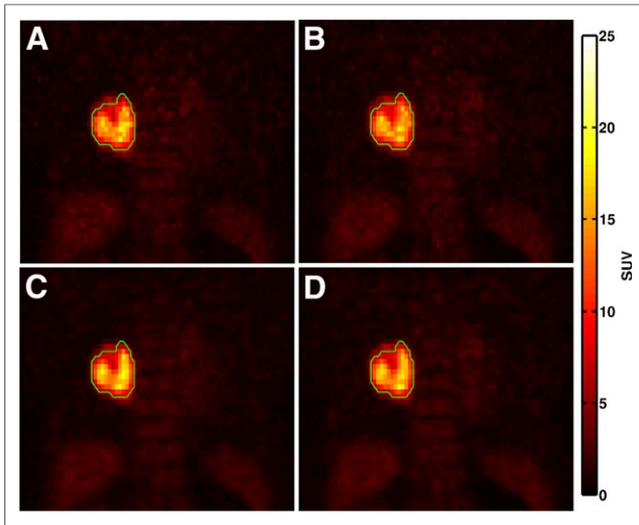
In this study, SUV was calculated as  $^{18}\text{F}$ -FDG uptake with decay correction normalized to injected dose and patient body weight. A tumor VOI was delineated using a fixed threshold set to 40% of the  $\text{SUV}_{\text{max}}$  in the lesion and followed by a morphologic closing operation to include necrotic regions since tumor heterogeneity is often associated with hypoxia and subsequent necrosis (27). In addition, a manual adjustment to exclude neighboring nodes or metastases was made for each VOI if necessary. As in the study by Orhac et al. (14), only lesions with a volume greater than  $5\text{ cm}^3$  were included in the subsequent analysis, because the impact of the partial-volume effect and textural measurements may not be relevant in small regions. Thus, the impact of the partial-volume effect can be neglected. Previous studies (22–26) demonstrated that OSEM + PSF + TOF produced better image quality than other methods in terms of SNR, contrast, and lesion detectability. Therefore, in this study, the tumor VOIs were delineated on the image reconstructed by OSEM + PSF + TOF with default settings and then applied to the other methods. This yielded 20 VOIs in non-small cell lung cancer, 2 VOIs in nasopharyngeal cancer, and 2 VOIs in lymphoma with a volume of  $36.1 \pm 42.3\text{ cm}^3$  (range, 5.3–153.1  $\text{cm}^3$ ).  $\text{SUV}_{\text{mean}}$ ,  $\text{SUV}_{\text{peak}}$ , and  $\text{SUV}_{\text{max}}$  of the lesions reconstructed by OSEM + PSF + TOF with default settings were  $6.4 \pm 2.7$  (range, 3.1–12.0),  $8.5 \pm 4.0$  (range, 4.0–18.8), and  $10.5 \pm 4.5$  (range, 5.3–21.2), respectively.

### Image Features

The TFs being used in  $^{18}\text{F}$ -FDG PET imaging can be categorized into 2 types: second-order and high-order features. A second-order feature is usually calculated from the gray level cooccurrence matrix (GLCM) (28). There are various matrices to calculate the high-order

**TABLE 1**  
List of Reconstruction Settings

Reconstruction algorithm	Variation over the default reconstruction settings	Impact of iteration number on image features (FWHM: 2.5 mm; grid size: 256 × 256)	Impact of FWHM on image features (iteration: 2; grid size: 256 × 256)	Impact of grid size on image features (iteration: 2; FWHM: 2.5 mm)
OSEM	Iteration: 2; FWHM: 2.5 mm; grid size: 256 × 256	Iteration: 1, 2, 3	FWHM: 2.5, 3.5, 4.5, 5.5 mm	Grid size: 256 × 256; 128 × 128
OSEM + PSF	Iteration: 2; FWHM: 2.5 mm; grid size: 256 × 256	Iteration: 1, 2, 3	FWHM: 2.5, 3.5, 4.5, 5.5 mm	Grid size: 256 × 256; 128 × 128
OSEM + TOF	Iteration: 2; FWHM: 2.5 mm; grid size: 256 × 256	Iteration: 1, 2, 3	FWHM: 2.5, 3.5, 4.5, 5.5 mm	Grid size: 256 × 256; 128 × 128
OSEM + PSF + TOF	Iteration: 2; FWHM: 2.5 mm; grid size: 256 × 256	Iteration: 1, 2, 3	FWHM: 2.5, 3.5, 4.5, 5.5 mm	Grid size: 256 × 256; 128 × 128



**FIGURE 1.** Coronal images of 71-y-old female patient diagnosed with non-small cell lung cancer (stage IV; body weight, 61.8 kg; height, 1.57 m; injected dose, 245.1 MBq) reconstructed by OSEM (A), OSEM + PSF (B), OSEM + TOF (C), and OSEM + PSF + TOF (D) with default reconstruction parameter settings (2 iterations and 24 subsets for OSEM and OSEM + PSF, 2 iterations and 21 subsets for OSEM + TOF and gaussian filter with FWHM of 2.5 mm). Lesion is located in right lung and contoured in green.

features including the gray-level run length matrix (GLRLM) (29), gray-level size zone matrix (GLSZM) (30), neighboring gray-level dependence matrix (NGLDM) (31), and neighbor gray-tone differ-

ence matrix (NGTDM) (32). To reduce image noise, resampling the original PET values is necessary. Orhac et al. (14) suggested that at least 32 gray levels should be used in the computation of TF. A bin size of 32, 64, and 128 was investigated in this study. In addition, 6 first-order statistics (FOS) based on histogram analysis were also included. The TF and FOS and their acronyms are summarized in Supplemental Table 1 (supplemental materials are available at <http://jnm.snmjournals.org>). In addition,  $SUV_{max}$ ,  $SUV_{mean}$ , and  $SUV_{peak}$  (a spheric VOI having a volume of 1 mL in a position that provides the maximal VOI average) were computed for comparison.

The 3-dimensional GLCM used in this study was extended from the original 2-dimensional GLCM by summing voxel triplet probabilities in a 2-dimensional image. Thirteen different directions and a spatial distance of 1 voxel displacement were selected (14). The average TF over the 13 directions was used in this study, and the same 13 directions were used to calculate GLRLM. The 26 nearest neighbors in 3 dimensions were used for NGLDM and NGTDM. In addition, only the intensity difference of zero between the voxel and its neighbors was considered for NGTDM.

### Data Analysis

To characterize the variation of image features over the different reconstruction parameters, we calculated the COV as follows:

$$COV = SD/mean \times 100(\%),$$

where SD and mean are the SD and the mean value of each image feature, respectively, over the different reconstruction settings. The mean value of the COV of all lesions for each TF and FOS was used to characterize feature variability. All features were categorized into 4 groups based on COV: very small ( $COV \leq 5\%$ ), small ( $5\% < COV \leq 10\%$ ), intermediate ( $10\% < COV \leq 20\%$ ), and large ( $COV > 20\%$ ) range of variation. In the last 3 studies, all features were categorized into 4 groups based on the

**TABLE 2**  
Change of Image Features over Default Reconstruction Settings

Feature	$COV \leq 5\%$	$5\% < COV \leq 10\%$	$10\% < COV \leq 20\%$	$COV > 20\%$
SUV	$SUV_{mean}$ , $SUV_{peak}$	$SUV_{max}$		
FOS	Entropy	COV, kurtosis, energy	Variance	Skewness
GLCM	Dissimilarity, energy, entropy, ID, SE, DE, IMC, IDN, IDMN, DM, SDN	Contrast, correlation, homogeneity, MP, SA, DV	Autocorrelation, SOS, SV	CS
GLRLM	GLNr, RP, LGRE, HGRE		SRE, LRE, RLN, SRLGE, SRHGE, LRLGE, LRHGE	
GLSZM	GLNz, LGZE	ZLN, HGZE, WVGLZ_S	SZE, LZE, SZLGE, SZHGE, LZLGE, LZHGE, WVGLZ_N	ZP
NGLDM	Entropy	SNE, NN, SM	LNE	
NGTDM		Coarseness, busyness, complexity, TS	Contrast	

ID = inverse difference; SE = sum entropy; DE = difference entropy; IMC = information measure of correlation; DM = diagonal moment; SDN = second diagonal moment; MP = maximum probability; SA = sum average; DV = difference variance; SOS = sum of squares; SV = sum variance; GLNr = grey-level nonuniformity for run; RP = run percentage; LGRE = low grey-level run emphasis; HGRE = high grey-level run emphasis; SRE = short-run emphasis; LRE = long-run emphasis; RLN = run-length nonuniformity; SRLGE = short-run low grey-level emphasis; SRHGE = short-run high grey-level emphasis; LRLGE = long-run low grey-level emphasis; LRHGE = long-run high grey-level emphasis; GLNz = grey-level nonuniformity for zone; LGZE = low grey-level zone emphasis; ZLN = zone-length nonuniformity; HGZE = high grey-level zone emphasis; WVGLZ\_S = weighted variance of gray-level size, S direction; SZE = short-zone emphasis; LZE = long-zone emphasis; SZLGE = short-zone low grey-level emphasis; SZHGE = short-zone high grey-level emphasis; LZLGE = long-zone low grey-level emphasis; LZHGE = long-zone high grey-level emphasis; WVGLZ\_N = weighted variance of gray-level size, N direction; SNE = small number emphasis; NN = number nonuniformity; SM = second moment; LNE = large number emphasis; TS = texture strength.

**TABLE 3**  
Impact of Iteration Number on Image Features

Feature	COV ≤ 5%	5% < COV ≤ 10%	10% < COV ≤ 20%	COV > 20%
SUV	SUV <sub>mean</sub> , SUV <sub>peak</sub>	SUV <sub>max</sub>		
FOS	COV, Entropy	Energy, kurtosis	Variance	Skewness
GLCM	Dissimilarity, energy, entropy, ID, SA, SE, DE, IMC, IDN, IDMN, DM, SDN	Autocorrelation, contrast, homogeneity, MP, SOS, SV, DV	Correlation	CS
GLRLM	GLNr, RLN, RP, LGRE, HGRE	SRE, LRE, SRLGE, SRHGE, LRLGE, LRHGE		
GLSZM	ZLN, LGZE	SZE, GLNz, LZE, HGZE, SZLGE, SZHGE, LZLGE, WVGLZ_S	LZHGE, WVGLZ_N	ZP
NGLDM	Entropy	SNE, NN, SM	LNE	
NGTDM		Contrast	Coarseness, busyness, complexity, TS	

ID = inverse difference; SA = sum average; SE = sum entropy; DE = difference entropy; IMC = information measure of correlation; DM = diagonal moment; SDN = second diagonal moment; MP = maximum probability; SOS = sum of squares; SV = sum variance; DV = difference variance; GLNr = grey-level nonuniformity for run; RLN = run-length nonuniformity; RP = run percentage; LGRE = low grey-level run emphasis; HGRE = high grey-level run emphasis; SRE = short-run emphasis; LRE = long-run emphasis; SRLGE = short-run low grey-level emphasis; SRHGE = short-run high grey-level emphasis; LRLGE = long-run low grey-level emphasis; LRHGE = long-run high grey-level emphasis; ZLN = zone-length nonuniformity; LGZE = low grey-level zone emphasis; SZE = short-zone emphasis; GLNz = grey-level nonuniformity for zone; LZE = long-zone emphasis; HGZE = high grey-level zone emphasis; SZLGE = short-zone low grey-level emphasis; SZHGE = short-zone high grey-level emphasis; LZLGE = long-zone low grey-level emphasis; WVGLZ\_S = weighted variance of gray-level size, S direction; LZHGE = long-zone high grey-level emphasis; WVGLZ\_N = weighted variance of gray-level size, N direction; SNE = small number emphasis; NN = number nonuniformity; SM = second moment; LNE = large number emphasis; TS = texture strength.

following rule: for each feature  $f$ , if it had a higher value of COV for one reconstruction algorithm than for the other reconstruction algorithms, it would be assigned to the group based on this higher value.

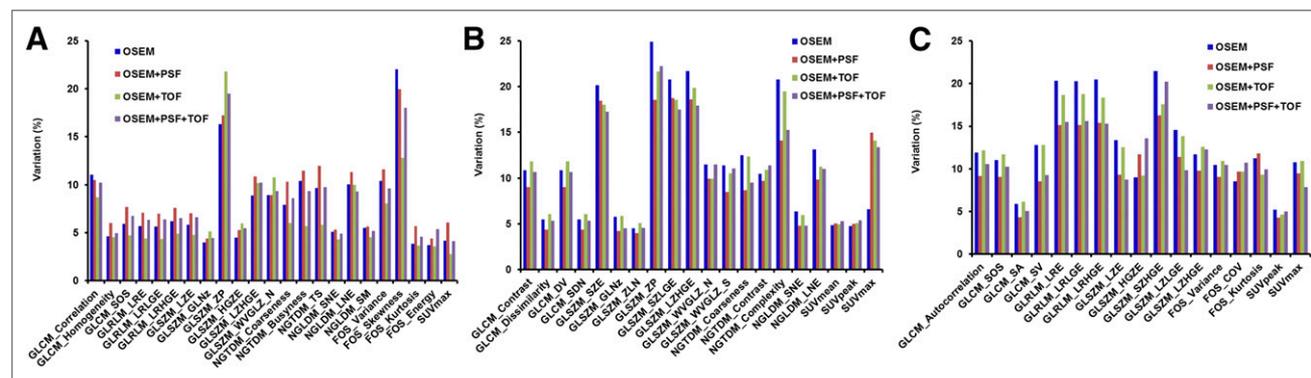
In addition, the quality of PET images reconstructed with default settings were analyzed using the SNR in the liver. A spheric mask with a diameter of 3 cm was placed in the liver, and the 4 default reconstructions were evaluated using the same mask. The SNR was defined as the ratio of the mean to the SD of the uptake within the mask.

## RESULTS

### Change of Image Features over Default Reconstruction Settings

PET images of a representative lung cancer patient reconstructed by the default reconstruction settings for OSEM, OSEM + TOF,

OSEM + PSF, and OSEM + TOF + PSF are shown in Figure 1. The delineated lesion is also shown in green in the figure. Consistent with the previous findings (22–26), OSEM + TOF + PSF produced visually the best image among the chosen 4 reconstruction algorithms with their default settings. In addition, quantitative assessment supported the fact that OSEM + TOF + PSF produced the best image with an SNR of 9.1 in the liver, whereas the SNR of OSEM, OSEM + PSF, and OSEM + TOF were 5.2, 5.5, and 8.5, respectively. The results with a bin size of 32, 64, and 128 were similar with only a few differences. For example, the COV of energy from GLCM changed from 5% for a bin size of 64 to a range of 5%–10% for a bin size of 32. In this study, 64 was selected as the reference bin size. SUV<sub>mean</sub> and SUV<sub>peak</sub> exhibited a small variation, with a COV less than 5%, and SUV<sub>max</sub> had



**FIGURE 2.** Image features with different range of variation for OSEM, OSEM + PSF, OSEM + TOF, and OSEM + PSF + TOF for each individual parameter: iteration number (A), FWHM (B), and grid size (C).

**TABLE 4**  
Impact of FWHM on Image Features

Feature	COV $\leq$ 5%	5% < COV $\leq$ 10%	10% < COV $\leq$ 20%	COV > 20%
SUV		SUV <sub>mean</sub> , SUV <sub>peak</sub>	SUV <sub>max</sub>	
FOS	Entropy	COV, kurtosis, energy		Variance, skewness
GLCM	Energy, entropy, ID, SE, DE, IMC, IDN, IDMN, DM	Dissimilarity, homogeneity, MP, SA, SDN	Autocorrelation, contrast, correlation, SOS, SV, DV	CS
GLRLM	GLNr, RP, LGRE, HGRE	RLN	SRE, LRE, SRLGE, SRHGE, LRLGE, LRHGE	
GLSZM	LGZE	GLNz, ZLN, HGZE	LZE, SZHGE, LZLGE, WVGLZ_N, WVGLZ_S	SZE, ZP, SZLGE, LZHGE
NGLDM	Entropy	SNE, NN, SM	LNE	
NGTDM		TS	Coarseness, contrast, busyness	Complexity

ID = inverse difference; SE = sum entropy; DE = difference entropy; IMC = information measure of correlation; DM = diagonal moment; MP = maximum probability; SA = sum average; SDN = second diagonal moment; SOS = sum of squares; SV = sum variance; DV = difference variance; GLNr = grey-level nonuniformity for run; RP = run percentage; LGRE = low grey-level run emphasis; HGRE = high grey-level run emphasis; RLN = run-length nonuniformity; SRE = short-run emphasis; LRE = long-run emphasis; SRLGE = short-run low grey-level emphasis; SRHGE = short-run high grey-level emphasis; LRLGE = long-run low grey-level emphasis; LRHGE = long-run high grey-level emphasis; LGZE = low grey-level zone emphasis; GLNz = grey-level nonuniformity for zone; ZLN = zone-length nonuniformity; HGZE = high grey-level zone emphasis; LZE = long-zone emphasis; SZHGE = short-zone high grey-level emphasis; LZLGE = long-zone low grey-level emphasis; WVGLZ\_N = weighted variance of gray-level size, N direction; WVGLZ\_S = weighted variance of gray-level size, S direction; SZE = short-zone emphasis; SZLGE = short-zone low grey-level emphasis; LZHGE = long-zone high grey-level emphasis; SNE = small number emphasis; NN = number nonuniformity; SM = second moment; LNE = large number emphasis; TS = texture strength.

a COV in the range of 5%–10%. Most of the features (58/61) had a COV less than 20%, indicating small differences across the 4 algorithms (Table 2).

#### Impact of Iteration Number on Image Features

Three features (skewness, cluster shade [CS], and zone percentage [ZP]) demonstrated a large variation (COV > 20%). SUV<sub>mean</sub>, SUV<sub>peak</sub>, and 22 features measuring heterogeneity showed a small variation (Table 3). The entropy from FOS, GLCM, and NGLDM were all insensitive to iteration number. In addition, there were 27 features in the group of small variations including contrast from GLCM and NGTDM. SUV<sub>max</sub> had a COV less than 10%. The remaining 9 features had a COV of between 10% and 20%. The results of each individual reconstruction algorithm were comparable (Supplemental Table 2), and the image features with different range of variation across the 4 reconstruction algorithms are shown in Figure 2A. There were only 8 features displaying differences between OSEM and OSEM + PSF. In addition, CS was the only parameter exhibiting a large variation in all reconstruction algorithms.

#### Impact of FWHM on Image Features

The number of features (TF and FOS) in the 4 groups (very small, small, intermediate, and large variation) were 16, 16, 21, and 8, respectively (Table 4). All of the TFs from GLRLM exhibited a COV less than 20%. GLRLM also yielded the TF with variation less than 20%, except for CS. Among the features derived from FOS, only variance and skewness had large variations (COV  $\geq$  20%), and 4 other features (entropy, COV, kurtosis, and energy) had small variations (COV  $\leq$  10%). The TF from GLSZM had a large range of COV, including 4 features (short-zone emphasis, ZP, short-zone low grey-level emphasis, and long-zone high grey-level emphasis) with COV larger than 20%. The results

of impact of the FWHM on TF and FOS for each individual reconstruction algorithm are shown in Supplemental Table 3. In addition, 20 image features including SUV<sub>mean</sub>, SUV<sub>peak</sub>, and SUV<sub>max</sub> exhibited different ranges of variation (Fig. 2B). Variance, skewness, CS, and ZP were in the large-variation group for all reconstruction algorithms. SUV<sub>mean</sub> and SUV<sub>peak</sub> exhibited small variation, and SUV<sub>max</sub> showed intermediate variation. Overall, the reconstruction algorithms all demonstrated similar performance.

#### Impact of Grid Size on Image Features

Two grid sizes (128  $\times$  128 and 256  $\times$  256) were used to evaluate the impact on TF and FOS with default iteration number and a FWHM of 2.5 mm (Table 5). More than half of the features (35/61) were sensitive to grid size, as demonstrated by the high COV ( $\geq$ 20%), including 1 of 6 features from FOS, 11 of 21 features from GLCM, 8 of 11 features from GLRLM, 6 of 13 features from GLSZM, 4 of 5 features from NGLDM, and 5 of 5 features from NGTDM. The entropy from FOS, difference entropy, inverse difference normalized (IDN), and inverse difference moment normalized (IDMN), low grey-level run emphasis, high grey-level run emphasis, and low grey-level zone emphasis were stable with respect to the grid size (COV  $\leq$  5%). Moreover, only sum average and sum entropy showed small variations (5% < COV  $\leq$  10%). Others (17/61 features) exhibited intermediate sensitivity to the grid size (10% < COV  $\leq$  20%). The grid size had a similar impact on TF and FOS for all the reconstruction algorithms (Supplemental Table 4), and the corresponding features, with different variations for these 4 reconstruction algorithms, are shown in Figure 2C. More than half of the 61 features (TF and FOS) demonstrated a large variation. SUV<sub>mean</sub> and SUV<sub>peak</sub> had a COV less than 10%. The COV of SUV<sub>max</sub> was between 10% and 20%.

**TABLE 5**  
Impact of Grid Size on Image Features

Feature	COV ≤ 5%	5% < COV ≤ 10%	10% < COV ≤ 20%	COV > 20%
SUV		SUV <sub>mean</sub> , SUV <sub>peak</sub>	SUV <sub>max</sub>	
FOS	Entropy		Kurtosis, variance, COV, energy	Skewness
GLCM	DE, IDN, IDMN	SA, SE	Autocorrelation, entropy, ID, SOS, SV	Contrast, correlation, CS, dissimilarity, energy, homogeneity, MP, DV, IMC, DM, SDN
GLRLM	LGRE, HGRE			LRE, SRE, GLNr, RLN, RP, SRLGE, SRHGE, LRLGE, LRHGE
GLSZM	LGZE		SZE, LZE, LZLGE, HGZE, SZLGE, LZHGGE	GLNz, ZLN, ZP, SZHGGE, WVGLZ_N, WVGLZ_S
NGLDM			SNE	LNE, NN, SM, entropy
NGTDM				Coarseness, contrast, busyness, complexity, TS

DE = difference entropy; SA = sum average; SE = sum entropy; ID = inverse difference; SOS = sum of squares; SV = sum variance; MP = maximum probability; DV = difference variance; IMC = information measure of correlation; DM = diagonal moment; SDN = second diagonal moment; LGRE = low grey-level run emphasis; HGRE = high grey-level run emphasis; LRE = long-run emphasis; SRE = short-run emphasis; GLNr = grey-level nonuniformity for run; RLN = run-length nonuniformity; RP = run percentage; SRLGE = short-run low grey-level emphasis; SRHGE = short-run high grey-level emphasis; LRLGE = long-run low grey-level emphasis; LRHGE = long-run high grey-level emphasis; LGZE = low grey-level zone emphasis; SZE = short-zone emphasis; LZE = long-zone emphasis; LZLGE = long-zone low grey-level emphasis; HGZE = high grey-level zone emphasis; SZLGE = short-zone low grey-level emphasis; LZHGGE = long-zone high grey-level emphasis; GLNz = grey-level nonuniformity for zone; ZLN = zone-length nonuniformity; SZHGGE = short-zone high grey-level emphasis; WVGLZ\_N = weighted variance of gray-level size, N direction; WVGLZ\_S = weighted variance of gray-level size, S direction; SNE = small number emphasis; LNE = large number emphasis; NN = number nonuniformity; SM = second moment; TS = texture strength.

## DISCUSSION

Our primary aim was to investigate the impact of reconstruction settings on image features derived from <sup>18</sup>F-FDG PET tumor images acquired with a PET/CT scanner incorporating PSF and TOF. Filtered backprojection was not included in this study because it is now rarely used clinically. Our results demonstrated that most of the features (58/61) had a COV less than 20% across the 4 algorithms with default settings, which were different from the results of variation across all possible reconstruction settings derived from FBP and OSEM in a previous study (11).

Although the results with a bin size of 32, 64, and 128 are similar, choosing the optimal bin size is still challenging. In a recent investigation of the prognostic value of the TF obtained from <sup>18</sup>F-FDG PET images in patients with oropharyngeal carcinoma (12), uniformity from GLCM was found to be a significant prognostic factor whereas its predictive value depended on the bin size. The optimal discretization size may depend on the noise and resolution of the PET images and should be carefully considered for different objectives.

Coarseness from NGTDM was considered to be correlated with human perception of image granularity and was deemed clinically useful. The coarseness from NGTDM can be used not only to predict response of chemoradiotherapy for non-small cell lung cancer (10) and esophageal cancer (18) but also to differentiate normal tissues from head and neck tumors and lymph nodes (8). However, it was found to be sensitive to reconstruction settings with a small variation (5% < COV ≤ 10%) over the 4 default reconstruction settings, intermediate variation (10% < COV ≤ 20%) for both impact of iteration number and FWHM, and large

variation (COV > 20%) to grid size. Therefore, it is desirable to standardize the reconstruction for multicenter clinical trials if such imaging biomarkers are to be used.

The smaller grid size meant that the bigger voxel size was more affected by the tissue fraction effect, and the image had a more uniform intensity distribution, which subsequently influenced TF and FOS. The grid size had the largest effect on image features, as demonstrated by 17 of the features with a COV between 10% and 20% and 35 of 61 features with a COV greater than 20% in this study (Table 5). Postreconstruction gaussian filters can improve the SNR of an image. However, there is a trade-off between noise and spatial resolution. A larger FWHM leads to poorer spatial resolution although the resultant image has less noise. The FWHM of the gaussian filter was found to be the second largest factor contributing to the variation of image features. Compared with grid size and FWHM, the iteration number had less impact on image features. In addition, TF and FOS were not sensitive to reconstruction algorithms with default settings and iteration number.

Volumetric tumor delineation is a prerequisite for computing TF and FOS. Although a large number of approaches have been proposed to segment tumors in PET images, accurate tumor segmentation is still a challenging task (33). A simple thresholding method (40% of SUV<sub>max</sub>) was used here, possibly leading to an imperfect tumor delineation. However, the inaccurate segmentation is unlikely to change the relative difference due to different reconstruction settings for each feature. Other limitations of the present study include a relatively small patient cohort and a single lesion site. Therefore, a larger patient cohort

study with different tumor types needs to be performed in the future. Moreover, it is absolutely necessary to investigate the clinical relevance and biologic basis of the PET texture features.

A new PET imaging biomarker should be insensitive to reconstruction settings or at least not worse than conventional indices such as SUV measures. In this work, as expected,  $SUV_{max}$ , measuring single-pixel uptake, was more sensitive to reconstruction settings than  $SUV_{mean}$  and  $SUV_{peak}$ . However,  $SUV_{max}$  is still more robust than some TF and FOS such as skewness from FOS, CS from GLCM, and ZP from GLSZM. The image features with a smaller variation than  $SUV_{max}$  are indicated for further investigation as to their clinical relevance. These features include the entropy from FOS, difference entropy, IDN and IDMN from GLCM, LGRE and HGRE from GLRLM, and LGZE from GLSZM.

## CONCLUSION

In this study, we analyzed the effect of reconstruction settings, including different algorithms, iteration number, FWHM of the postreconstruction gaussian filter, and grid size, on TFs and first-order features. Different TF and FOS had different sensitivities to reconstruction settings. Iteration number and the FWHM of the gaussian filter had a similar impact for all 4 reconstruction algorithms. Compared with iteration number and FWHM, grid size had a larger impact on image features. Careful selection of image features for research, clinical practice, and clinical trials is essential.

## DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. No potential conflict of interest relevant to this article was reported.

## REFERENCES

- Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012; 366:883–892.
- Van Allen EM, Wagle N, Stojanov P, et al. Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat Med*. 2014;20:682–688.
- Beyer T, Czernin J, Freudenberg LS. Variations in clinical PET/CT operations: results of an international survey of active PET/CT users. *J Nucl Med*. 2011;52: 303–310.
- Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake value. *J Nucl Med*. 2012;53: 1041–1047.
- El Naqa I, Grigsby P, Apte A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit*. 2009;42: 1162–1171.
- Brooks FJ, Grigsby PW. The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. *J Nucl Med*. 2014;55:37–42.
- Bundschuh RA, Dinges J, Neumann L, et al. Textural parameters of tumor heterogeneity in  $^{18}\text{F}$ -FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. *J Nucl Med*. 2014;55: 891–897.
- Cheng NM, Fang YH, Chang JT, et al. Textural features of pretreatment  $^{18}\text{F}$ -FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. *J Nucl Med*. 2013;54:1703–1709.
- Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in  $^{18}\text{F}$ -FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging*. 2013;40:133–140.
- Cook GJ, Yip C, Siddique M, et al. Are pretreatment  $^{18}\text{F}$ -FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med*. 2013;54:19–26.
- Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol*. 2010;49:1012–1016.
- Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour  $^{18}\text{F}$ -FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging*. 2013;40: 1662–1671.
- Lartzien C, Rogez M, Niaf E, Ricard F. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Health Inform*. 2014;18:946–955.
- Orlhac F, Soussan M, Maisonneuve JA, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in  $^{18}\text{F}$ -FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med*. 2014;55:414–422.
- Soussan M, Orhac F, Boubaya M, et al. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *PLoS One*. 2014;9:e94017.
- Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in  $^{18}\text{F}$ -FDG PET. *J Nucl Med*. 2012;53:693–700.
- Tixier F, Hatt M, Valla C, et al. Visual versus quantitative assessment of intratumour  $^{18}\text{F}$ -FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med*. 2014;55:1235–1241.
- Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline  $^{18}\text{F}$ -FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med*. 2011;52: 369–378.
- Xu R, Kido S, Suga K, et al. Texture analysis on  $^{18}\text{F}$ -FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions. *Ann Nucl Med*. 2014;28:926–935.
- Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging*. 2009;28:374–383.
- Yu H, Caldwell C, Mah K, et al. Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *Int J Radiat Oncol Biol Phys*. 2009;75:618–625.
- Schaefferkoetter J, Casey M, Townsend D, El Fakhri G. Clinical impact of time-of-flight and point response modeling in PET reconstructions: a lesion detection study. *Phys Med Biol*. 2013;58:1465–1478.
- Akamatsu G, Ishikawa K, Mitsumoto K, et al. Improvement in PET/CT image quality with a combination of point-spread function and time-of-flight in relation to reconstruction parameters. *J Nucl Med*. 2012;53:1716–1722.
- Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol*. 2011;56:2375–2389.
- Kadmas DJ, Casey ME, Conti M, Jakoby BW, Lois C, Townsend DW. Impact of time-of-flight on PET tumor detection. *J Nucl Med*. 2009;50:1315–1323.
- Surti S, Scheuermann J, El Fakhri G, et al. Impact of time-of-flight PET on whole-body oncologic studies: a human observer lesion detection and localization study. *J Nucl Med*. 2011;52:712–719.
- Lu X, Kang Y. Hypoxia and hypoxia-inducible factors: master regulators of metastasis. *Clin Cancer Res*. 2010;16:5928–5935.
- Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;SMC-3:610–621.
- Tang X. Texture information in run-length matrices. *IEEE Trans Image Process*. 1998;7:1602–1609.
- Thibault G, Fertil B, Navarro C, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. *Pattern Recognition Inf Process*. 2009:140–145.
- Sun C, Wee GW. Neighboring gray level dependence matrix for texture classification. *Comput Vision Graph*. 1983;23:341–352.
- Amadasum M, King R. Texture features corresponding to textural properties. *IEEE Trans Syst Man Cybern*. 1989;19:1264–1274.
- Zaidi H, El Naqa I. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. *Eur J Nucl Med Mol Imaging*. 2010;37:2165–2187.