



Impact of image reconstruction methods on quantitative accuracy and variability of FDG-PET volumetric and textural measures in solid tumors

Ali Ketabi^{1,2} · Pardis Ghafarian^{3,4} · Mohammad Amin Mosleh-Shirazi⁵ · Seyed Rabi Mahdavi⁶ · Arman Rahmim^{7,8} · Mohammad Reza Ay^{1,2} 

Received: 17 July 2018 / Revised: 20 August 2018 / Accepted: 11 September 2018
© European Society of Radiology 2018

Abstract

Objective This study aims to assess the impact of different image reconstruction methods on PET/CT quantitative volumetric and textural parameters and the inter-reconstruction variability of these measurements.

Methods A total of 25 oncology patients with 65 lesions (between 2017 and 2018) and a phantom with signal-to-background ratios (SBR) of 2 and 4 were included. All images were retrospectively reconstructed using OSEM, PSF only, TOF only, and TOFPSF with 3-, 5-, and 6.4-mm Gaussian filters. The metabolic tumor volume (MTV) and total lesion glycolysis (TLG) were measured. The relative percent error (Δ MTV and Δ TLG) with respect to true values, volume recovery coefficients, and Dice similarity coefficient, as well as inter-reconstruction variabilities were quantified and assessed. In clinical scans, textural features (coefficient of variation, skewness, and kurtosis) were determined.

Results Among reconstruction methods, mean Δ MTV differed by $-163.5 \pm 14.1\%$ to $6.3 \pm 6.2\%$ at SBR2 and $-42.7 \pm 36.7\%$ to 8.6 ± 3.1 at SBR4. Dice similarity coefficient significantly increased by increasing SBR from 2 to 4, ranging from 25.7 to 83.4% between reconstruction methods. Mean Δ TLG was -12.0 ± 1.7 for diameters > 17 mm and -17.8 ± 7.8 for diameters ≤ 17 mm at SBR4. It was -31.7 ± 4.3 for diameters > 17 mm and -14.2 ± 5.8 for diameters ≤ 17 mm at SBR2. Textural features were prone to variations by reconstruction methods ($p < 0.05$).

Conclusions Inter-reconstruction variability was significantly affected by the target size, SBR, and cut-off threshold value. In small tumors, inter-reconstruction variability was noteworthy, and quantitative parameters were strongly affected. TOFPSF reconstruction with small filter size produced greater improvements in performance and accuracy in quantitative PET/CT imaging.

Key Points

- *Quantitative volumetric PET evaluation is critical for the analysis of tumors.*
- *However, volumetric and textural evaluation is prone to important variations according to different image reconstruction settings.*
- *TOFPSF reconstruction with small filter size improves quantitative analysis.*

✉ Pardis Ghafarian
pardis.ghafarian@sbm.ac.ir

✉ Mohammad Reza Ay
mohammadreza_ay@sina.tums.ac.ir

¹ Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

² Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran

³ Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁴ PET/CT and Cyclotron Center, Masih Daneshvari Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁵ Ionizing and Nonionizing Radiation Protection Research Center and Department of Radio-Oncology, Shiraz University of Medical Sciences, Shiraz, Iran

⁶ Department of Medical Physics, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran

⁷ Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD, USA

⁸ Departments of Radiology and Physics & Astronomy, University of British Columbia, Vancouver, BC, Canada

Keywords PET-CT · Image reconstruction · Tumor burden · Radiation oncology

Abbreviations

COV	Coefficient of variation
CT	Computed tomography
FDG	Fluoro-deoxy-glucose
FWHM	Full width at half maximum
GE	General electric
HD	3D-OSEM algorithm referred to as HD
KBq	Kilo-becquerel
MBq	Mega-becquerel
MTV	Metabolic tumor volume
NEMA	National electrical manufacturers association
OSEM	Ordered subset expectation maximization
PET	Positron emission tomography
PSF	Point spread function
SBR	Signal-to-background ratios
SD	Standard deviation
SD _{ir}	Standard deviation of inter-reconstruction variation for each VOI
SD _{vi}	Standard deviation of voxel intensity distribution for each VOI
SUV	Standard uptake value
SUV _{max}	Maximum standard uptake value
SUV _{mean}	Mean standard uptake value
TLG	Total lesion glycolysis
TOF	Time of flight
TOFPSF	Time of flight and point spread function
VOI	Volume of interest
VRC	Volume recovery coefficients

Introduction

FDG-PET imaging holds a significant role and value for improved staging, image guidance to response assessment, and post-therapy follow-up in a range of cancers [1]. In addition, PET imaging, an increasingly integral part of radiation oncology, can be used with different orders of complexity in radiotherapy treatment planning [2].

Over the last few years, significant developments have been made to integrate novel quantitative imaging for treatment planning or post-treatment response evaluation. Quantitative analyses of PET images provide valuable information on the distribution of cancerous cells within a cancerous mass. PET/CT can be used to extract and sculpt 3D maps of tumor cells according to radiobiological relevant parameters that can be used to the concept of molecular imaging-based dose-painting [2, 3]. PET/CT-based dose-painting in which radiation is guided by the PET uptake is a paradigm of significant potential in radiation therapy prescription. In addition, there is significant interest in extraction and

characterization of PET-/CT-based volumetric and textural parameters, particularly metabolic tumor volume (MTV) and total lesion glycolysis (TLG), for assessment and prediction of treatment response as well as overall prognostication [4–17]. MTV calculation is based on the definition of tumor boundaries and does not indicate differences in the neoplastic cell densities within the tumor. TLG defined as MTV multiplied by the average SUV uptake (SUV_{mean}) within the MTV, which is an index of the neoplastic cell density; it thus includes both volumetric and activity information within the tumors.

One of the main contributors to the overall uncertainty in PET-based quantification is related to technical aspects of imaging (such as PET scan acquisition and image reconstruction methods and parameters) that impact the apparent size, shape, and uptake of tumors; there is neither consensus nor guidelines in volumetric assessment of tumors [18]. In recent years, with the introduction and development of new techniques such as time-of-flight (TOF) PET [19] and point spread function (PSF) modeling [20, 21] within the iterative reconstruction process, PET has experienced important changes [22, 23]. The impact of some image reconstruction methods on the standardized uptake value (SUV) has been evaluated in some studies [24–26].

Characterization of volumetric and textural features (spatial distributions of image intensity levels) in PET is an important consideration for PET-/CT-based dose-painting and response to therapy in radiation oncology. Changes in FDG uptake or tumor size and phenotypic heterogeneity are used as a measure that is very prone to variations by technical factors such as image reconstruction methods. We aimed to extend prior studies by evaluating the impact of different PET image reconstruction algorithms on quantitation of volumetric measures. In the current study, variability of conventional indices (MTV and TLG) measures to image reconstruction algorithms was investigated using ¹⁸F-FDG-PET/CT images of an image quality phantom. Additionally, tumor textural parameters such as coefficient of variation (COV), skewness, and kurtosis were quantified and assessed to determine the inter-reconstruction variability of these measurements in different image reconstruction methods.

Materials and methods

Data acquisition and image reconstruction

All PET acquisitions were performed on a Discovery-690 PET/CT scanner. This system, in default configuration, uses a fully 3D ordered subset expectation maximization (3D-

OSEM) algorithm (referred to as HD) [27]. Our data were retrospectively reconstructed with four different reconstruction algorithms: (i) HD = with no PSF or TOF, (ii) addition of PSF only, (iii) addition of TOF only, and (iv) TOFPSF = HD+TOF+PSF with post-smoothing Gaussian filters 3, 5, and 6.4 mm in full width at half maximum (FWHM). So, 12 different reconstruction methods were used as follows: HD3, HD5, HD6.4, PSF3, PSF5, PSF6.4, TOF3, TOF5, TOF6.4, TOFPSF3, TOFPSF5, and TOFPSF6.4. The coincidence time window was 4.1 ns, and the TOF time resolution was 555 ps. The image matrix was $256 \times 256 \times 47$ pixels with $2.73 \times 2.73 \times 3.27$ mm³ pixel size.

Phantom studies

Imaging protocol The National Electrical Manufacturers Association (NEMA) image quality phantom was used for all measurements. This phantom consists of six fillable inserts to simulate hot lesions with diameters of 10, 13, 17, 22, 28 and 37 mm suspended by rods that embedded in a body-shaped cavity. The background activity level of 5.3 kBq/ml was applied. Two different signal-to-background ratios of 2 (SBR2)

and 4 (SBR4) were chosen. The PET data were acquired in 3D list mode for 10 min per a bed position.

Data analysis SUV_{max} was defined as the cut-off threshold for contouring. MTV was defined for the voxels within the volume of interest (VOI) with $SUV \geq X\% \times SUV_{max}$ where X corresponds to the cut-off threshold value.

Volumetric accuracy was calculated using the percent error ΔMTV , i.e., the relative differences in MTV with respect to the true volume:

$$\Delta MTV = \frac{MTV_{recon.i} - True Volume_j}{True Volume_j} \times 100 \tag{1}$$

where $MTV_{recon.i}$ was the MTV corresponding to the *i*th reconstruction method and true volume_{*j*} is the actual size of *j*th inserts volume.

We assessed inter-reconstruction variability for each specific tumor size. First, six cut-off threshold values (X) were applied ranging from 40 to 90% at 10% increments and the percent error ΔMTV was determined for each reconstruction method in all target volumes. Next, the impact of different reconstruction methods on volumetric measures at 50% cut-off thresholding ($SUV_{max_50\%}$) was calculated.

Fig. 1 Percent error ΔMTV (%), mean ΔMTV , and SD_{ir} at SBR2; the smallest volume with 10-mm diameter (a), the medium volume with 17-mm diameter (b), the largest volume with 37-mm diameter (c). Positive ΔMTV indicates larger volumes for reconstructed images than for true volume. The ratio to SUV_{max} was applied at 10% intervals as the cut-off threshold for contouring

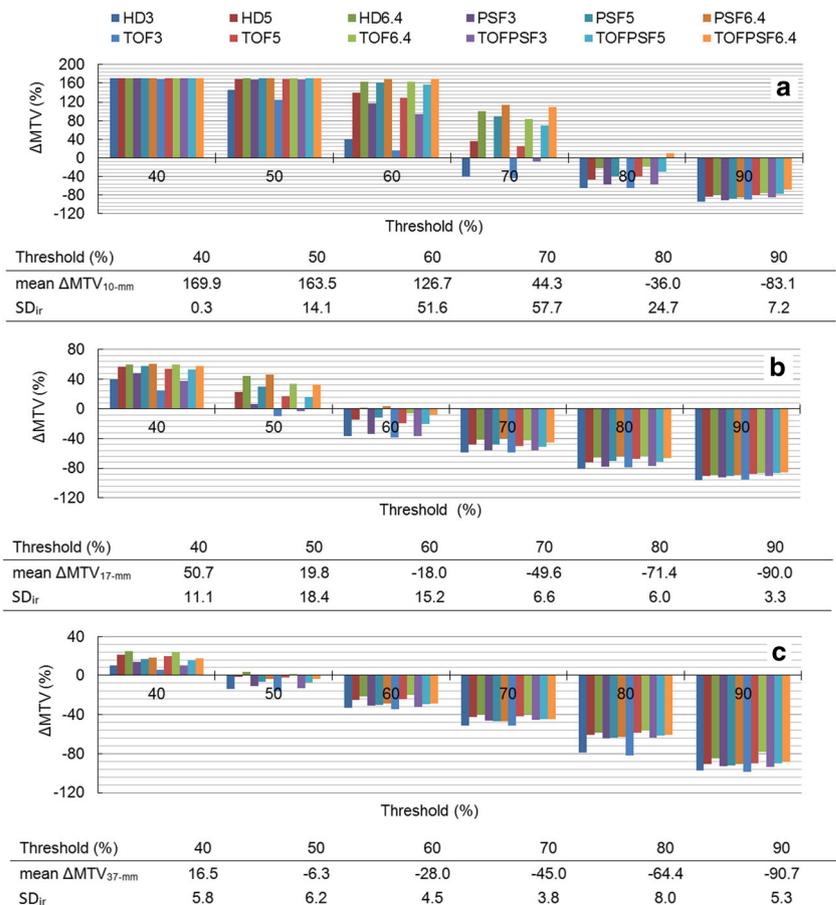
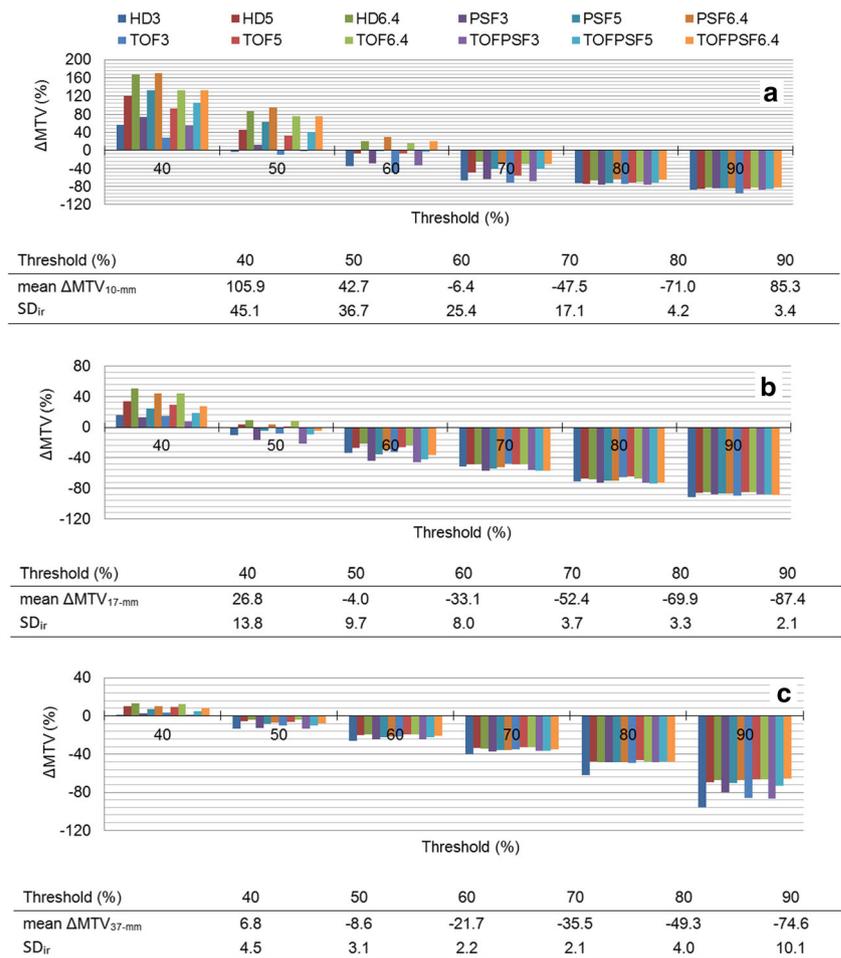


Fig. 2 Percent error Δ MTV (%), mean Δ MTV, and SD_{ir} at SBR4; the smallest volume with 10-mm diameter (a), the medium volume with 17-mm diameter (b), the largest volume with 37-mm diameter (c). Positive Δ MTV indicates larger volumes for reconstructed images than for true volume. The ratio to SUV_{max} was applied at 10% intervals as the cut-off threshold for contouring



Volumetric accuracy using the volume recovery coefficients (VRC) in measurements was also calculated. The VRC was defined as:

$$VRC = \frac{MTV_{recon.i}}{True\ Volume_j} \quad (2)$$

In addition, the Dice similarity coefficient [28] was used to determine the overlap between MTV from reconstructions and real target volumes, defined as follows:

Dice similarity coefficient

$$= \frac{2(MTV_{recon.i} \cap True\ Volume_j)}{MTV_{recon.i} + True\ Volume_j} \times 100 \quad (3)$$

Finally, to investigate the impact of reconstruction methods on a volumetric incorporating neoplastic cells density within the tumor, the quantitative value of TLG ($TLG = MTV \times SUV_{mean}$) was obtained. The relative differences in TLG with

respect to its true value, as the percent error Δ TLG, were calculated. Δ TLG was defined as follows:

$$\Delta TLG = \frac{TLG_{recon.i} - True\ TLG_j}{True\ TLG_j} \times 100 \quad (4)$$

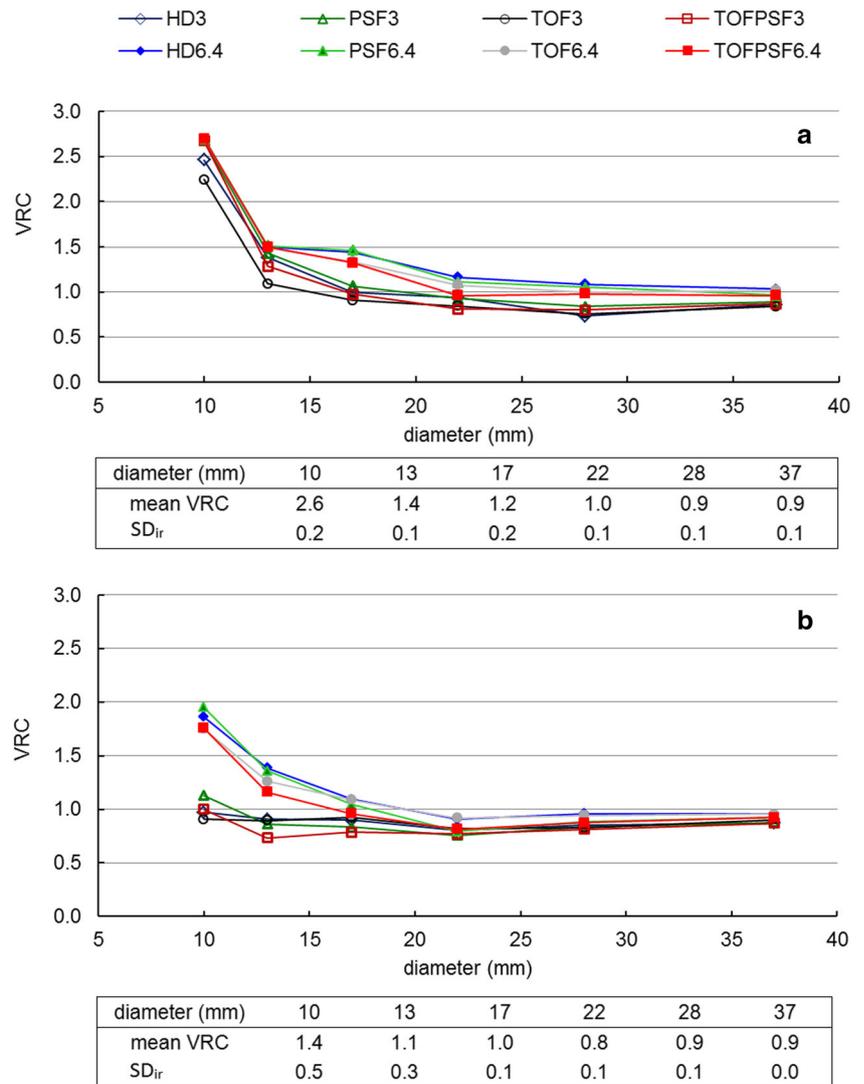
where $TLG_{recon.i}$ was the TLG corresponding to the i th reconstruction method and true TLG_j is the ideal value of TLG.

The mean Δ MTV, VRC, Dice similarity coefficient, and Δ TLG on reconstruction methods and the associated standard deviation of inter-reconstruction variation (SD_{ir}) for each VOI were calculated at the two SBRs (SBR2 and SBR4). SD_{ir} was chosen as a criterion to assess variability of volume estimates to reconstruction methods.

Patient data

The data for 25 patients (14 males and 11 females, between 2017 and 2018) with different types of solid tumors were retrospectively evaluated. The study was performed under a waiver of informed consent and approved by the Institutional Review Board and the authors used anonymous patient's

Fig. 3 Volume recovery coefficients (VRC), mean VRC, and SD_{ir} as function of insert diameter for different reconstruction methods using the Gaussian filters with FWHM of 3 mm (the smallest) and 6.4 mm (the largest) at SBR2 (a) and SBR4 (b). The 50% fraction (ratio to SUV_{max}) was applied as the cut-off threshold for MTV delineation



images. A total of 65 lesions including 26 primary tumors and 39 metastatic lesions were assessed. The mean age of patients was 43.0 ± 14.1 years. Malignancies included lymphoma, lung, colon, and unknown primary cancer. Patients were injected with 331.4 ± 71.0 MBq of ^{18}F -FDG. Patients fasted for at least 4 h before injection and scanned 60.8 ± 1.5 min post injection.

In addition, three different textural parameters including coefficient of variation (COV), skewness, and kurtosis were quantified and evaluated. The COV was measured as the ratio of the standard deviation of voxel intensity distribution (SD_{vi}) for each VOI and the mean of the activity concentration in the tumor volume (i.e., SD_{vi} /mean activity). Skewness and kurtosis were defined as measures of the asymmetry and peakedness of the activity distribution in the tumor volume, respectively. Reconstruction method PSF6.4 with three iterations and 18 subsets was applied in our routine clinical whole-body PET/CT imaging. For all abovementioned metrics in

patients, we evaluated the impact of other reconstruction methods in comparison with PSF6.4.

Statistical analysis was performed using SPSS, version 22.0 (IBM Corp., Armonk, New York, USA). Differences among the reconstruction methods under investigation were assessed using paired *t* test for the normal distribution and Wilcoxon's signed-rank test for those without normal distribution between the reconstruction methods with $p < 0.05$ as significance level.

Results

Comparison for different cut-off thresholds

Figures 1 and 2 depict the percent error ΔMTV (%), mean ΔMTV , and SD_{ir} for SBR2 and SBR4, respectively. The figures illustrate the inter-reconstruction variability of these measurements for 12 reconstruction methods and six different cut-

Table 1 Dice similarity coefficient between MTV and real target volume for different reconstruction methods at SBR2 and SBR4

Diameter	SBR2			SBR4		
	37 mm	17 mm	10 mm	37 mm	17 mm	10 mm
HD3	0.92	1.00	0.58	0.93	0.94	0.98
HD5	0.99	0.90	0.54	0.97	0.98	0.82
HD6.4	0.98	0.82	0.54	0.98	0.96	0.70
PSF3	0.94	0.97	0.54	0.94	0.91	0.94
PSF5	0.96	0.87	0.54	0.95	0.98	0.76
PSF6.4	0.98	0.81	0.54	0.96	0.98	0.68
TOF3	0.91	0.95	0.62	0.95	0.96	0.95
TOF5	0.99	0.92	0.54	0.97	1.00	0.86
TOF6.4	0.99	0.86	0.54	0.98	0.96	0.73
TOFPSF3	0.93	0.99	0.54	0.93	0.88	1.00
TOFPSF5	0.96	0.92	0.54	0.95	0.95	0.83
TOFPSF6.4	0.98	0.86	0.54	0.96	0.98	0.73
mean Dice similarity coefficient	0.96	0.91	0.55	0.96	0.96	0.83
SD _{ir}	0.03	0.06	0.02	0.02	0.03	0.12

The smallest (10-mm diameter), the medium (17-mm diameter), and the largest (37-mm diameter) volumes are presented. Mean Dice similarity coefficient and SD_{ir} among reconstruction methods are presented

off thresholds. For each reconstruction method, there was a turning-point that the value of Δ MTV changes from a positive to negative value by increasing cut-off threshold. The minimum Δ MTV and the maximum inter-reconstruction variabilities occurred in these turning points. Depending on the size of the target volume and the reconstruction methods, the turning points was obtained between the thresholds 50 to 80% at SBR2, and 40 to 60% at SBR4. The minimal Δ MTV resulted in the higher cut-off threshold by decreasing SBR and target volume size. Positive Δ MTV (according to Eq. 1) indicates larger volumes for reconstructed images than for true volume.

Comparison within 50% cut-off thresholding

By focusing on 50% cut-off thresholding in Figs. 1 and 2, a more detailed analysis was conducted to compare the reconstruction methods. The absolute value of Δ MTV from the lowest to the highest value was generally seen in TOFPSF, TOF, PSF, and HD reconstruction at both SBRs regardless of the filter size. For instance, Δ MTV in a range of target volumes in TOFPSF6.4, TOF6.4, PSF6.4, and HD6.4 were -40.4%, -43.7%, -46.7%, and -48.9% in SBR2; these were -8.1%, -15.3%, -15.9%, and -19.4% in SBR4.

Figure 3 illustrates the VRC, mean VRC, and SD_{ir} as a function of insert diameter for different reconstruction methods using the smallest and largest Gaussian filters at two SBRs. As can be seen in Fig. 3a, there is no significant difference between filter sizes in the smaller target volume at SBR2. The difference between filter sizes increases with increasing SBR (Fig. 3b).

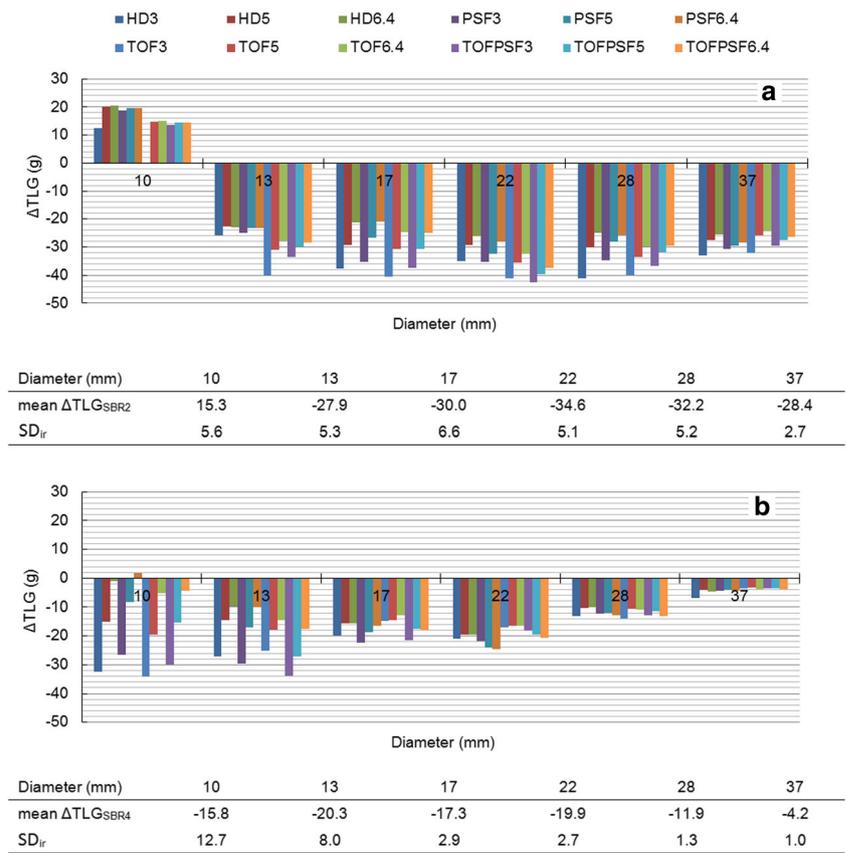
Table 1 quantifies the overlap between MTVs and real target volumes as the Dice similarity coefficient. Mean Dice similarity

coefficient and SD_{ir} among reconstruction methods are also presented. Minimum Dice value (0.83) and maximum SD_{ir} (0.12) was seen in the smallest insert at SBR4 that TOFPSF3 produced the best Dice value. The choice of a smaller filter in the small volumes is very important to achieve the higher Dice value. For each target volumes, mean Dice similarity coefficient increased by increasing SBR (ranging from 25.7 to 83.4%).

TLG for each specific insert size and all reconstruction methods were calculated from each MTV and its corresponding SUV_{mean}. Figure 4 compares percent error Δ TLG (%), mean Δ TLG, and SD_{ir} in different PET reconstruction methods and SBRs. The most TLG differences in small inserts were seen using the small filter size because it estimates a more realistic SUV_{mean} value.

The results on PSF6.4 and seven other reconstruction methods in clinical PET are compared in Fig. 5. Δ MTV, Dice similarity coefficient and Δ TLG are presented. The mean value and SD among tumor volumes for each reconstruction method are shown above the boxes. Inter-reconstruction differences between PSF6.4 and seven others were statistically significant. As illustrated in Fig. 5a, the MTVs for four algorithms were smaller at 3-mm filter size. PSF3 with -36.5 ± 12 mean Δ MTV had maximum MTV difference with PSF6.4. Dice similarity coefficient analysis also denoted a similar behavior for reconstruction methods. We subsequently investigated the inter-reconstruction variations in the tumor TLG estimation. The maximum TLG difference with PSF6.4 was observed in PSF3 with -14.5 ± 10.9 mean Δ TLG. Statistically significant inter-reconstruction differences were seen between PSF6.4 and seven others (p values < 0.0001 or p values < 0.01 as shown in the figures).

Fig. 4 Percent error Δ TLG (%), mean Δ TLG, and SD_{ir} in each specific tumor size and different reconstruction methods at SBR2 (a) and SBR4 (b). Positive Δ TLG indicates larger TLGs for reconstructed images than for ideal TLG. The 50% fraction (ratio to SUV_{max}) was applied as the cut-off threshold for MTV delineation



Box-plots of textural features for COV, skewness, and kurtosis are illustrated in Fig. 6. The mean value of each feature for each reconstruction method and the associated SD among patients are also shown. The *p* values for PSF6.4 method versus HD3, HD6.4, PSF3, TOF3, TOF6.4, TOFPSF3, and TOFPSF6.4 are summarized in Table 2. The volume of all lesions was smaller than the phantom insert with 17-mm diameter.

Discussion

We studied the impact of different PET image reconstruction methods on quantitative FDG-PET volumetric and textural parameters. Our results showed significant inter-reconstruction variability of these measurements. In smaller volumes, the appropriate cut-off threshold shifts to higher thresholds by decreasing the SBR. As such, the reconstruction method for these conditions must be chosen cautiously, and size-dependent thresholding is rational for accurate tumor volume delineation. Inter-reconstruction variability was significantly affected by the target volume size, SBRs, and the cut-off threshold value. Our results showed that for smaller target volumes, increasing inter-reconstruction variability was obtained.

Specifically, the effect of filter size was noteworthy. There was up to 74% variation in Δ MTV by changing from the smallest to the largest filter size. The smaller filter produced the best volumetric accuracy by decreasing both the target volume and SBR. Δ TLG analysis showed inter-reconstruction variability, SD_{ir} , to be significantly increased by increasing SBR and decreasing target volume. On the basis of our phantom and clinical findings, switching from no PSF to PSF modeling depicts greater dependence on tumor size. Also, consistent with prior literature [29, 30], we found that in the presence of PSF and TOF modeling, greater improvements in performance and volumetric accuracy were obtained.

Our work has a broad context. Quantitative cancer imaging using ^{18}F -FDG-PET is a promising tool in painting complex dose distributions [31, 32] and for prediction of response to treatment of patients [1]. However, accurate quantification is highly dependent on technical aspects of imaging which impact the apparent size and distribution of uptake in tumors. As such, deviations of volumetric parameters due to image reconstruction and segmentation methods need to be carefully assessed, which was the objective of the present work.

For all four reconstruction algorithms, the absolute value Δ MTV increased by decreasing the target volume diameter, as would be expected due to the partial volume

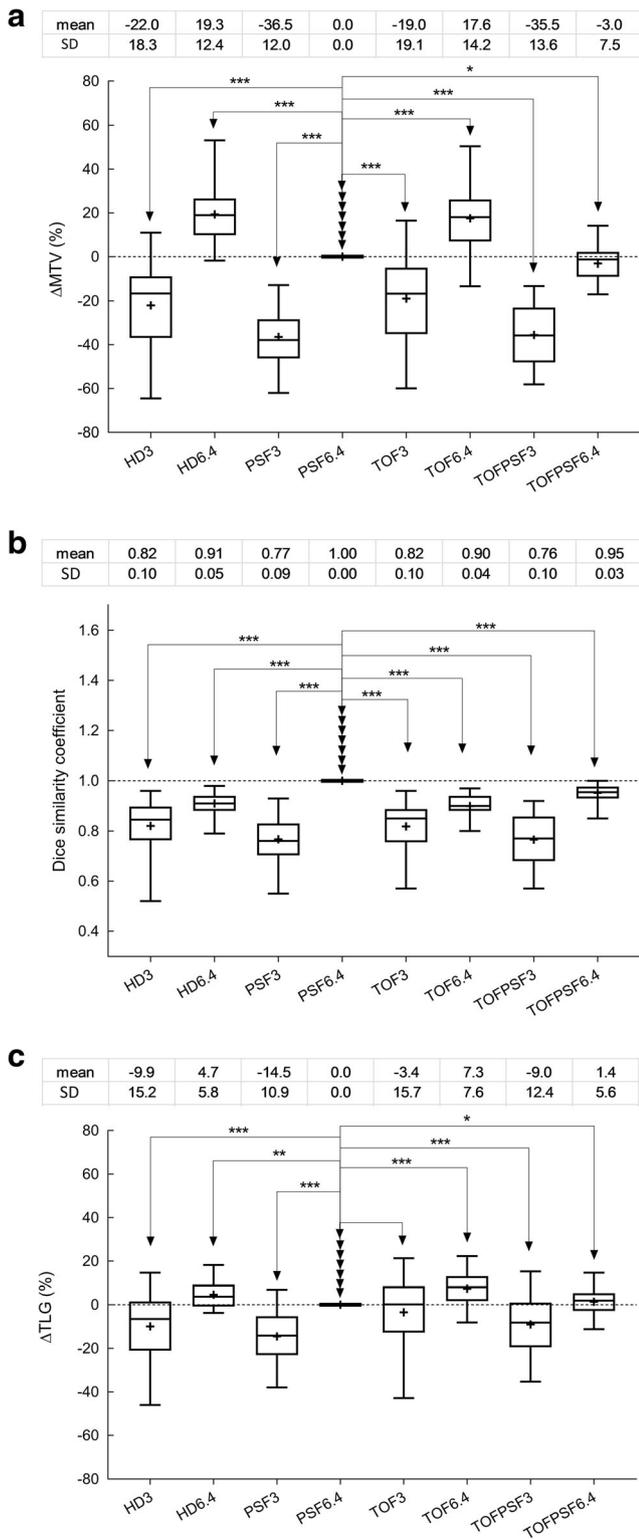


Fig. 5 Box-plots of Δ MTV (a), Dice similarity coefficient (b), and Δ TLG (c) for tumor volumes, comparing the PSF6.4 method with seven different reconstruction methods (** p values < 0.0001, * p values < 0.01). The mean value of each parameter for each reconstruction method and the associated SD among patient are presented

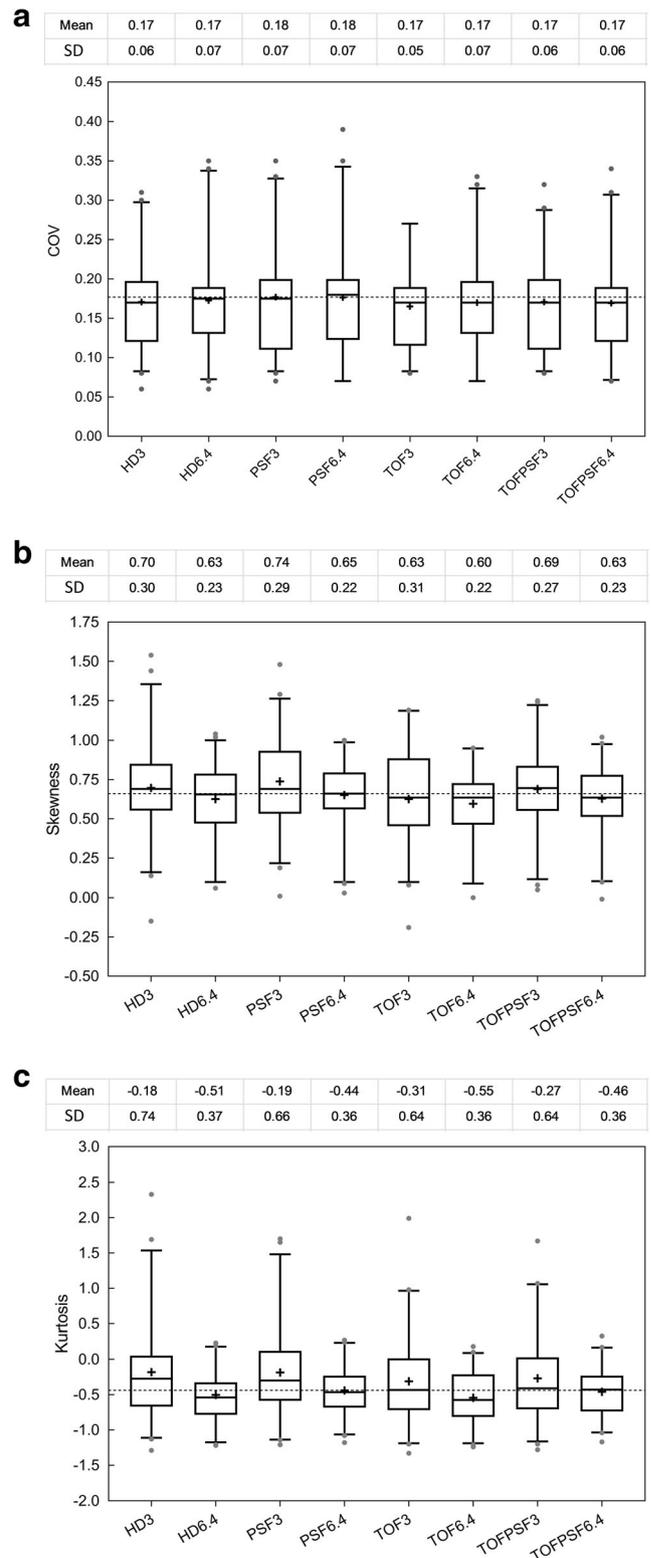


Fig. 6 Box and whisker (5–95 percentiles) plots of textural features are illustrated for COV (a), skewness (b) and kurtosis (c). The mean value of each feature for each reconstruction method and the associated SD among patients are presented

effect [33–35]. In such a condition, lesion delineation based on PET images results in an overestimation due to the limited spatial resolution. Hoetjes et al [36] showed that PSF reconstruction, as a partial volume correction method, could increase SUV by 5% up to 80% depending on tumor size. As illustrated in Figs. 1 and 2, the combination of PSF with HD and/or TOF tended to produce higher values of Δ MTV compared to non-PSF algorithms in a range of target volumes regardless of the filter size. This behavior occurs in both SBRs and is more pronounced in smaller target volumes. What is additionally interesting is that the interaction of PSF and TOF with lower smoothing in the quantification of smaller target volume, and one that needs more assessment and characterization. Indeed, TOFPSF in relation to the use of TOF only can also alter tumor detection [37, 38].

For higher SBR, the largest inter-reconstruction variability resulted in the smallest target volume in 40–50% thresholds. For lower SBR, MTV substantially overestimated in the smallest target volume so that the thresholds < 60% was not reliable for MTV delineation. However, inter-reconstruction variability was considerable for the higher thresholds (60–70%).

Large differences in VRC between reconstruction methods were seen in lesions with diameters ≤ 17 mm; VRC increased up to 2.7 and 2 at lower and higher SBRs in the smallest volume, respectively. This is attributed to the partial volume effect [33–35] which significantly impacts (lowers) SUV_{max} for lower volumes, thus overestimating MTVs as estimated using thresholding. Reconstruction methods with 3-mm filter size yielded the highest volume accuracy in the small volumes at higher SBR. The accuracy of larger volumes was not affected by SBRs for all reconstruction methods. For diameters > 17 mm, the VRC demonstrated differences of only 0.0 to 0.3 between SBRs. Differences in VRC ranging from 0.7 to 1.7 between SBRs was seen in the smallest target volume. The largest differences in VRC between SBRs were in TOFPSF3, PSF3, HD3, and TOF3 reconstruction methods (1.7, 1.5, 1.5, and 1.3, respectively).

Volumetric parameters (TLG and MTV) have proven to provide useful prognostic metrics both for staging and response assessment [6, 10–12, 17]. Our results, in concordance with a previous study [24] on patients with 85 lesions (all ≥ 2 cm), showed combination of PSF and/or TOF with HD-PET reconstruction had a variable effect on TLG

values. This becomes especially more important for small volumes at higher SBR whenever different smoothing filters were used. As illustrated in Fig. 4, although TLG is not significantly affected by reconstruction methods in large target volume, it was in small target volume that the effect on TLG was considerable. Inter-reconstruction comparison for small volume at lower SBR was not reliable because MTV segmentation as mentioned was not applicable. TLG slightly increased by increasing the size of the post-smoothing filter; this increase was more in non-PSF reconstruction methods than PSF methods.

Heterogeneity is a key feature of tumor mapping associated with dose-painting radiotherapy. Increased heterogeneity and poor prognosis are expected in higher COV, positive skewness, and higher kurtosis within a tumor [39]. Textural features were prone to variations by reconstruction methods. All TOF includes reconstruction methods that were statistically significantly different when comparing the different reconstruction methods for COV (see Table 2 with Fig. 6). The parameter skewness and kurtosis showed statistically significant differences between the PSF6.4 method and some reconstruction methods (summarized in Table 2). Bundschuh et al [40] reported that the COV had a higher area under the curve in receiver operating characteristic analysis than skewness and kurtosis.

The current work has some limitations, and our investigations into this area are still ongoing. We concentrated on few texture parameters (COV, skewness, and kurtosis). Further research is needed to assess the complex shape and texture or heterogeneity within a tumor. Although different target sizes, SBRs, and post-smoothing filters were used, the study only investigated fixed-threshold-based delineation.

The present work highlights the reconstruction method dependence of PET volumetric and textural parameters. Overall, we found that quantitative accuracy of small target volumes is more susceptible to change with image reconstruction methods. It is worth noticing that three of the six inserts with a small diameter (≤ 17 mm) have a large relative change (> 20%) in MTV and TLG with reconstruction methods. By contrast, only the change in filter size in large target volumes generates some difference in MTV and TLG estimates. Overall, it is our finding that image reconstruction techniques should be carefully considered and fully standardized for appropriate and consistent quantification efforts.

Table 2 Table of *p* values for the comparisons

Reconstruction method	HD3	HD6.4	PSF3	TOF3	TOF6.4	TOFPSF3	TOFPSF6.4
COV	0.18	0.15	0.75	0.04*	0.04*	0.03*	0.02*
Skewness	0.39	0.08	0.03*	0.70	0.00*	0.34	0.01*
Kurtosis	0.05*	0.04*	0.01*	0.46	0.01*	0.39	0.13

The PSF6.4 method versus other reconstruction methods is summarized

*Statistically significant *p* value

Conclusion

Quantification of volumetric PET parameters (MTV and TLG) is highly dependent on reconstruction methods. Inter-reconstruction variability was significantly affected by the target volume size, SBRs, and the cut-off threshold value. In small tumor volumes, inter-reconstruction variability was significant, and quantitative parameters were strongly affected. This was minimized in large tumor volume. TOFPSF reconstruction with small filter size produces greater improvement in performance and accuracy of quantitative analysis. Statistically significant differences were found among inter-reconstruction methods for first-order texture measures. Overall, the impact of image reconstruction techniques should be carefully considered and fully standardized for accurate and robust tumor quantification, texture analysis, and personalized therapy.

Funding This study has received funding by the Tehran University of Medical Sciences, Tehran, Iran, under grant number 28212; and Masih Daneshvari Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Mohammad Reza Ay, PhD, Professor of Medical Physics.

Conflict of interest The authors declare that they have no conflict of interest.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- Retrospective
- Diagnostic or prognostic study/experimental
- Performed at one institution

References

1. Wahl RL, Jacene H, Kasamon Y, Lodge MA (2009) From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 50:122S–150S
2. Thorwarth D, Geets X, Paiusco M (2010) Physical radiotherapy treatment planning based on functional PET/CT data. *Radiother Oncol* 96:317–324
3. Bentzen SM (2005) Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol* 6:112–117
4. Koyasu S, Nakamoto Y, Kikuchi M et al (2014) Prognostic value of pretreatment 18F-FDG PET/CT parameters including visual evaluation in patients with head and neck squamous cell carcinoma. *AJR Am J Roentgenol* 202:851–858
5. Wray R, Sheikhabaei S, Marcus C et al (2016) Therapy response assessment and patient outcomes in head and neck squamous cell carcinoma: FDG PET Hopkins criteria versus residual neck node size and morphologic features. *AJR Am J Roentgenol* 207:641–647
6. Pak K, Cheon GJ, Nam HY et al (2014) Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: a systematic review and meta-analysis. *J Nucl Med* 55:884–890
7. Abgral R, Keromnes N, Robin P et al (2014) Prognostic value of volumetric parameters measured by 18F-FDG PET/CT in patients with head and neck squamous cell carcinoma. *Eur J Nucl Med Mol Imaging* 41:659–667
8. Paidpally V, Chirindel A, Chung CH et al (2014) FDG volumetric parameters and survival outcomes after definitive chemoradiotherapy in patients with recurrent head and neck squamous cell carcinoma. *AJR Am J Roentgenol* 203:W139–W145
9. Chan SC, Chang JT, Lin CY et al (2011) Clinical utility of 18F-FDG PET parameters in patients with advanced nasopharyngeal carcinoma: predictive role for different survival endpoints and impact on prognostic stratification. *Nucl Med Commun* 32:989–996
10. Ryu IS, Kim JS, Roh JL et al (2014) Prognostic significance of preoperative metabolic tumour volume and total lesion glycolysis measured by 18F-FDG PET/CT in squamous cell carcinoma of the oral cavity. *Eur J Nucl Med Mol Imaging* 41:452–461
11. Abd El-Hafez YG, Moustafa HM, Khalil HF, Liao CT, Yen TC (2013) Total lesion glycolysis: a possible new prognostic parameter in oral cavity squamous cell carcinoma. *Oral Oncol* 49:261–268
12. Park GC, Kim JS, Roh JL, Choi SH, Nam SY, Kim SY (2012) Prognostic value of metabolic tumor volume measured by 18F-FDG PET/CT in advanced-stage squamous cell carcinoma of the larynx and hypopharynx. *Ann Oncol* 24:208–214
13. Lee SJ, Choi JY, Lee HJ et al (2012) Prognostic value of volume-based 18F-fluorodeoxyglucose PET/CT parameters in patients with clinically node-negative oral tongue squamous cell carcinoma. *Korean J Radiol* 13:752–759
14. Dibble EH, Alvarez AC, Truong MT, Mercier G, Cook EF, Subramaniam RM (2012) 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J Nucl Med* 53:709–715
15. Sheikhabaei S, Wray R, Young B et al (2016) 18F-FDG-PET/CT therapy assessment of locally advanced pancreatic adenocarcinoma: impact on management and utilization of quantitative parameters for patient survival prediction. *Nucl Med Commun* 37:231–238
16. Cherry S, Sorenson J, Phelps M (2012) *Physics in nuclear medicine*, 4th edn. Elsevier Saunders, Philadelphia, PA
17. Chung MK, Jeong HS, Park SG et al (2009) Metabolic tumor volume of [18F]-fluorodeoxyglucose positron emission tomography/computed tomography predicts short-term outcome to radiotherapy with or without chemotherapy in pharyngeal cancer. *Clin Cancer Res* 15:5861–5868
18. Lee JA (2010) Segmentation of positron emission tomography images: some recommendations for target delineation in radiation oncology. *Radiother Oncol* 96:302–307
19. Karp JS, Surti S, Daube-Witherspoon ME, Muehllehner G (2008) Benefit of time-of-flight in PET: experimental and clinical results. *J Nucl Med* 49:462–470
20. Rapisarda E, Bettinardi V, Thielemans K, Gilardi MC (2010) Image-based point spread function implementation in a fully 3D OSEM reconstruction algorithm for PET. *Phys Med Biol* 55:4131
21. Rahmim A, Qi J, Sossi V (2013) Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. *Med Phys* 40:064301

22. Cheebsumon P, Yaqub M, van Velden FH, Hoekstra OS, Lammertsma AA, Boellaard R (2011) Impact of [18F] FDG PET imaging parameters on automatic tumour delineation: need for improved tumour delineation methodology. *Eur J Nucl Med Mol Imaging* 38:2136–2144
23. Alessio AM, Rahmim A, Orton CG (2013) Resolution modeling enhances PET imaging. *Med Phys* 40:120601
24. Sheikhabaehi S, Marcus C, Wray R, Rahmim A, Lodge MA, Subramaniam RM (2016) Impact of point spread function reconstruction on quantitative 18F-FDG-PET/CT imaging parameters and inter-reader reproducibility in solid tumors. *Nucl Med Commun* 37:288–296
25. Ghafarian P, Ketabi A, Doroudinia A, Karam MB, Ay MR (2016) Effect of TOF and PSF in detection of lymph node metastases in head and neck of PET/CT images. *European Journal of Nuclear Medicine and Molecular Imaging* 43: S507–S507
26. Ketabi A, Ghafarian P, Mosleh-Shirazi M, Mahdavi S, Ay MR (2018) The influence of using different reconstruction algorithms on sensitivity of quantitative 18F-FDG-PET volumetric measures to background activity variation. *Iran J Nucl Med* 26:87–97
27. Bettinardi V, Presotto L, Rapisarda E, Picchio M, Gianoli L, Gilardi MC (2011) Physical performance of the new hybrid PET/CT Discovery-690. *Med Phys* 38:5394–5411
28. Zou KH, Warfield SK, Bharatha A et al (2004) Statistical validation of image segmentation quality based on a spatial overlap index 1: scientific reports. *Acad Radiol* 11:178–189
29. Akamatsu G, Ishikawa K, Mitsumoto K et al (2012) 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* 53:1716–1722
30. Schaefferkoetter J, Casey M, Townsend D, El Fakhri G (2013) Clinical impact of time-of-flight and point response modeling in PET reconstructions: a lesion detection study. *Phys Med Biol* 58: 1465
31. De Neve W (2015) Comparison of Adaptive dose painting by numbers with standard Radiotherapy for head and neck cancer. (C-ART-2). University Hospital, Ghent. Available via <https://clinicaltrials.gov/ct2/show/NCT01341535>. Accessed 25 Apr 2011
32. Guerrero Urbano T (2016) 18F-FDG-PET guided dose-painting with intensity modulated radiotherapy in oropharyngeal tumours (FiGaRO). Guy's & St Thomas' NHS Foundation Trust, United Kingdom. Available via <https://clinicaltrials.gov/ct2/show/NCT02953197>. Accessed 2 Nov 2016
33. Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF (2012) A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol* 57:R119
34. Rousset O, Rahmim A, Alavi A, Zaidi H (2007) Partial volume correction strategies in PET. *PET Clin* 2:235–249
35. Soret M, Bacharach SL, Buvat I (2007) Partial-volume effect in PET tumor imaging. *J Nucl Med* 48:932–945
36. Hoetjes NJ, van Velden FH, Hoekstra OS et al (2010) Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging* 37:1679–1687
37. Rahmim A, Tang J (2013) Noise propagation in resolution modeled PET imaging and its impact on detectability. *Phys Med Biol* 58: 6945
38. Kadmas DJ, Casey ME, Conti M, Jakoby BW, Lois C, Townsend DW (2009) Impact of time-of-flight on PET tumor detection. *J Nucl Med* 50:1315
39. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA (2010) Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging* 10:137
40. Bundschuh RA, Dinges J, Neumann L et al (2014) Textural parameters of tumor heterogeneity in 18F-FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. *J Nucl Med* 55:891–897