Radiological Physics and Technology (2019) 12:343–350
https://doi.org/10.1007/s12194-019-00525-7

Influence of image noise and object size on segmentation accuracy of FDG-PET imaging: a phantom experiment

Yoshiyuki Takahashi[1] · Tatsuya Tsuchitani[1] · Noriko Kotoura[1] · Kazuhiro Kitajima[2]

Received: 20 February 2019 / Revised: 20 July 2019 / Accepted: 20 July 2019 / Published online: 25 July 2019

Abstract
We aimed to evaluate the influence of noise and object size on segmentation accuracy of fluorodeoxyglucose positron emission tomography (FDG-PET) imaging. The scanned data of spherical phantoms were used. For the gradient method, 40% maximum standardized uptake value (SUV_{max}) method, and SUV of 2.5 threshold method, we evaluated the correlation between segmentation accuracy and background variability and that between segmentation accuracy and sphere diameters. For the gradient method, background variability did not affect segmentation accuracy, but sphere diameters had a small effect. As for the 40% SUV_{max} threshold method, both sphere diameters and background variability affected the segmentation accuracy. In the SUV of 2.5 threshold method, segmentation accuracy was affected by sphere diameters but not by background variability. With regard to segmentation accuracy of FDG-PET imaging, the gradient method may be more accurate and reliable compared to threshold methods when applied to images with varying noise or object size.

Keywords FDG-PET · Segmentation · Gradient method · SUV_{max}

1 Introduction

Fluorodeoxyglucose positron emission tomography combined with computed tomography (FDG-PET/CT) is widely used for the diagnosis, staging, treatment optimization, restaging, therapy monitoring, and prognostication of various types of malignant tumors [1]. Standardized uptake value (SUV) is the most widely used quantification index for the diagnosis and therapy assessment associated with metabolic activity [2, 3], although that depends heavily on physiological, physical, and procedural factors [4–7]. Furthermore, the maximum SUV (SUV_{max}) does not reflect three-dimensional (3D) information [8]. To overcome these limitations, metabolic tumor volume (MTV) and total lesion glycolysis are used as potential exploratory parameters for characterizing and providing information regarding the 3D structure and intra-tumoral biological variations of tumors [9–11]. Several methods for segmenting metabolic tumor volumes have been evaluated, including methods based on a fixed threshold of volume, volume gradient, and source-to-background ratio and manual methods [12–14]. Some studies have reported challenges with regard to the segmentation of PET images, issues with resolution, and noise [13, 14], although few studies have evaluated these factors in detail. Moreover, segmentation accuracy is usually evaluated by computing the absolute percent difference in the total volume between two segmentations. Being intuitive and simple, the percent volume difference is commonly used in clinical settings, but this metric alone does not convey sufficient information to determine the similarity between two segmentations [14]. As a commonly used index for segmentation evaluation, the dice similarity coefficient (DSC) was introduced [14]. The objective of the present study was to assess the influence of noise and object size on the segmentation accuracy of PET imaging with two different indices using fixed-threshold and gradient methods.

✉ Yoshiyuki Takahashi
yo-takahashi@hyo-med.ac.jp

[1] Department of Radiological Technology, Hyogo College of Medicine College Hospital, 1-1 Mukogawa-cho, Nishinomiya, Hyogo 663-8501, Japan

[2] Department of Radiology, Division of Nuclear Medicine and PET Center, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya, Hyogo 663-8501, Japan

2 Materials and methods

2.1 PET/CT

PET/CT images were acquired using a GEMINI TF 64 (PHILIPS Healthcare, Eindhoven, Netherlands). PET acquisition was performed in 3D mode, and the obtained images were reconstructed using the ordered subset expectation maximization (OSEM) reconstruction algorithm (33 subsets, 3 iterations, and 4 mm per slice), with attenuation correction based on 64-multi-detector CT (120 kVp, 30 mAs, and slice width 5 mm). The image matrix was 144 × 144. The pixel size was 4 mm. The slice thickness acquired on PET was 4 mm. For scatter correction, the time-of-flight extended single scatter simulation algorithm [15] was used.

2.2 Scanning of spherical phantoms

The National Electrical Manufacturers Association and International Electro-Technical Commission body phantom images filled with ^{18}F were acquired by the PET/CT camera. Six spherical phantoms, with diameters of 10, 13, 17, 22, 28, and 37 mm, were obtained in the phantom, while background activity was 2.65 MBq/mL with a source-to-background ratio of 4:1. Phantom images were acquired in a single-bed position over 30 min to cover the middle area of the phantom in list mode. Ten images at acquisition times of 10, 20, 30, 40, 50, 60, and 90 s and 3, 5, and 10 min were obtained from the list mode data, and each image was reconstructed with the relaxation parameters (λ) Normal ($\lambda = 1.0$), Smooth ($\lambda = 0.7$), Smooth A ($\lambda = 0.6$), and Smooth B ($\lambda = 0.5$). The relaxation parameters, incorporated in the OSEM algorithm, were the smoothing parameters. Finally, 44 types of phantom images were created.

2.3 Image analysis

The sphere volume was measured using the MIM Maestro software, ver. 6.8.7 (MIM Software Inc., Cleveland, OH, USA). The spheres of the 44 phantom images were measured using the gradient segmentation and fixed-threshold methods previously described [13]. Briefly, the gradient method relies on an operator-defined starting point near the center of the lesion. As the operator drags out from the center of the lesion, six axes extend outward, providing a visual feedback for the starting point of the gradient segmentation. Spatial gradients are interactively calculated along each axis, and the length of an axis is restricted when a large spatial gradient is detected along that axis. The maximal activity gradient along each axis is used to detect the tumor edge. The six axes are utilized in forming an ellipsoid, which is then used as the initial bounding region for gradient detection. The observers in the study were instructed to select the image slice with the largest tumor. Subsequently, the observers were instructed to locate a point near the center of the lesion in the selected slice and drag the cursor from that point to the six axes approximating the boundaries of the lesion. After releasing the mouse button, the edges of the structure were automatically calculated and outlined. For irregularly shaped structures that were not well defined on the six axes, the observers were instructed to use the gradient method one or more times to add to the initial contour by dragging the cursor from a point near the center of the omitted region. Operators added regions until they were visually satisfied that the entire structure was included in the contour. The fixed-threshold method includes all voxels that are greater than a definite percentage of the maximum voxel within an operator-defined sphere or a set value (in this study, 40% of the maximum voxel and SUV of 2.5 [16]). Cross-sectional circles are displayed in all three projections (axial, sagittal, and coronal), ensuring a 3D coverage of the lesion. In the present study, the edges of the lesion were automatically calculated and outlined using all the segmentation methods.

Segmentation accuracy was evaluated using two different formulas. One was for %Error, which was calculated as follows:

$$\%\text{Error} = |(\text{Vol}_{\text{PET}} - \text{Vol}_{\text{Phantom}})|/\text{Vol}_{\text{Phantom}} \times 100,$$

where Vol_{PET} represents the PET sphere volume using the gradient and fixed-threshold (40% SUV_{max} and SUV of 2.5) methods, and $\text{Vol}_{\text{Phantom}}$ represents the true sphere volume. The other formula was used for dice similarity coefficient (DSC), computed as follows:

$$\text{DSC} = 2|\text{Vol}_{\text{CT}} \cap \text{Vol}_{\text{PET}}|/(|\text{Vol}_{\text{CT}}| + |\text{Vol}_{\text{PET}}|),$$

where Vol_{CT} represents a spherical phantom lesion measured manually from the CT image using spherical volume of interest (VOI). Therefore, $|\text{Vol}_{\text{CT}} \cap \text{Vol}_{\text{PET}}|$ indicates the overlap of CT and PET.

Background variability (BGV) was used as the index of noise. We placed 12 circular regions of interest (ROIs) with diameters of 37 mm on the central slice and at ±1 and ±2 cm away from the central slice (total of 60 ROIs) in each PET image. For each ROI, the $\Delta\text{SUV}_{\text{mean}}$ was calculated as follows:

$$\Delta\text{SUV}_{\text{mean}} = (\text{SUV}_{\text{mean}}/\text{SUV}_{\text{TOT}} - 1),$$

where SUV_{mean} is the mean SUV of each ROI, and SUV_{TOT} is the mean SUV of all ROIs. Then, the standard deviation of $\Delta\text{SUV}_{\text{mean}}$ was defined as BGV.

2.4 Statistical analysis

To evaluate the effect of noise on segmentation, we analyzed the correlation between BGV and %Error and between BGV and DSC with the three segmentation methods using

Pearson's coefficient. Data obtained from a sphere with a diameter of 37 mm were used in this evaluation for minimizing the partial volume effect.

The relationship between object size and segmentation accuracy was analyzed using the correlation between sphere diameter and mean %Error and between sphere diameter and mean DSC for each of the three segmentation methods with a Jonckheere–Terpstra trend test. Mann–Whitney *U* test was used to compare differences among the mean %Error or DSC of each pair of the six sphere diameter sizes. Data obtained from the acquisition times of 30, 10, and 5 min were used in this evaluation to minimize the effect of noise.

A p value < 0.05 was considered statistically significant. The JMP Pro 13 (SAS Institute Inc.) and EZR (Saitama Medical Center, Jichi Medical University, Saitama, Japan) software packages were used for all the analyses.

3 Results

Figure 1 shows the relationship between BGV and %Error of the three segmentation methods. The correlation coefficient (r^2) values for the BGV and the %Error values obtained with the gradient, 40% SUV_{max}, and SUV of 2.5 threshold methods were 0.05 ($p = 0.154$), 0.56 ($p < 0.0001$), and 0.001 ($p = 0.84$), respectively (Fig. 1).

Figure 2 shows the relationship between BGV and DSC for the three segmentation methods. The correlation coefficient (r^2) values for the BGV and DSC values measured by the gradient, 40% SUV_{max}, and SUV of 2.5 threshold methods were 0.008 ($p = 0.06$), 0.003 ($p = 0.74$), and 0.06 ($p = 0.11$), respectively (Figs. 2, 3, 4, 5).

As shown in Table 1, a statistically significant linear trend was observed for the mean %Error with the gradient and

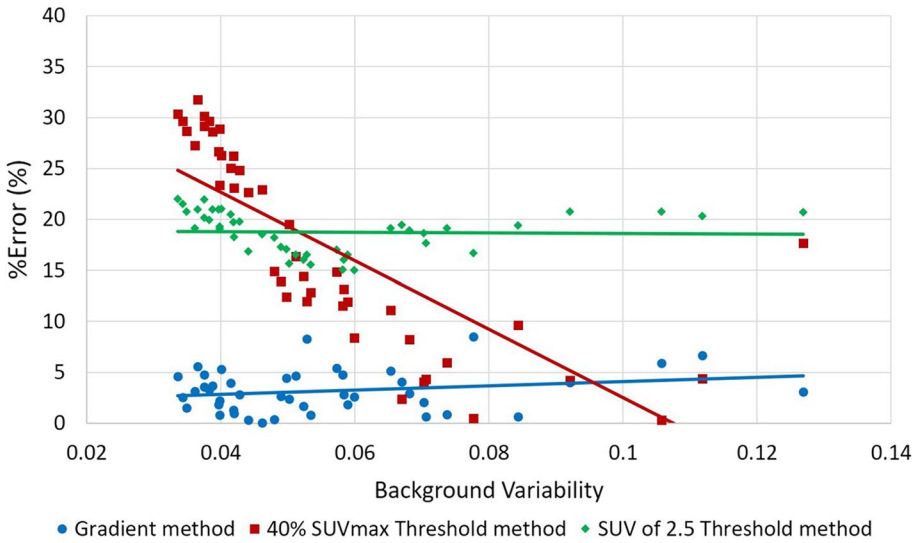

Fig. 1 Correlation of background variability and %Error with the gradient method, 40% maximum standardized uptake value threshold method, and SUV of 2.5 threshold method. The r^2 values with these methods were 0.05 ($p = 0.154$), 0.56 ($p < 0.0001$), and 0.001 ($p = 0.84$), respectively

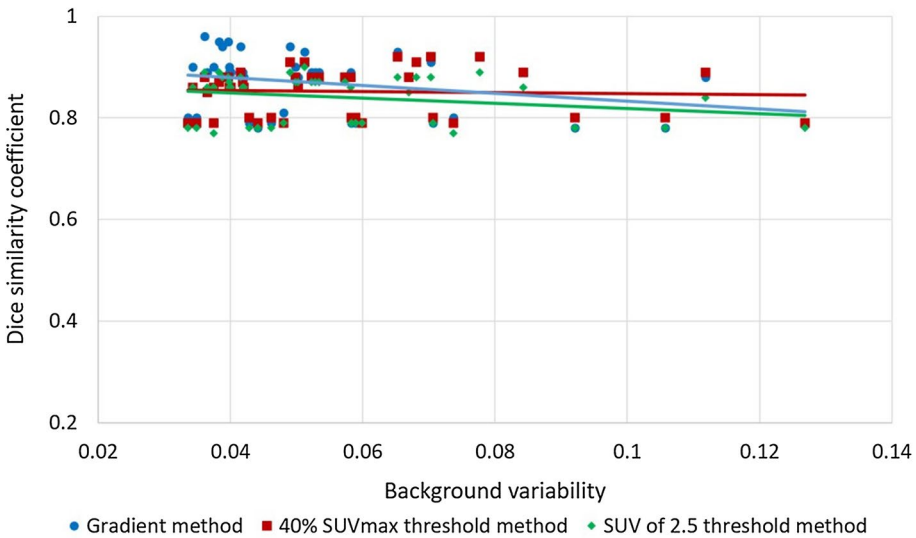

Fig. 2 Correlation of background variability and dice similarity coefficient with the gradient method, 40% maximum standardized uptake value threshold method, and SUV of 2.5 threshold method. The r^2 values with these methods were 0.008 ($p = 0.06$), 0.003 ($p = 0.74$), and 0.06 ($p = 0.11$), respectively

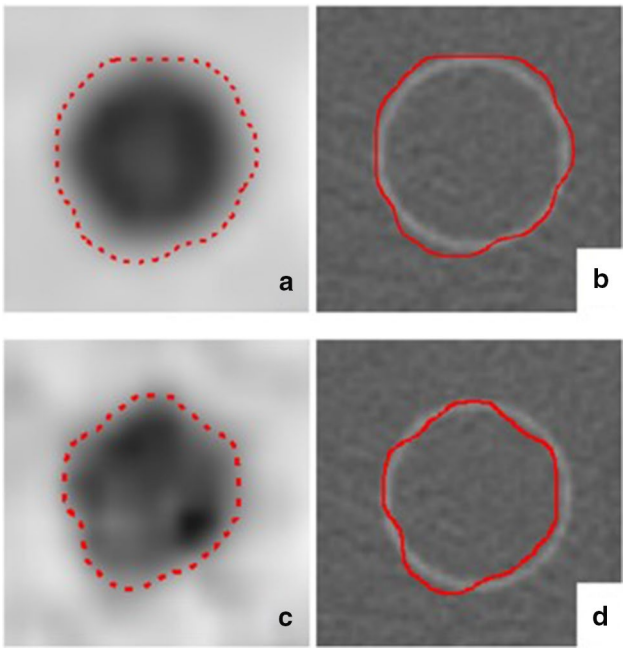

Fig. 3 Sphere images with similar dice similarity coefficient (DSC) and different %Error values. **a** Transverse PET imaging (acquisition time 10 min, reconstructed with relaxation parameters [λ] of Smooth [$\lambda = 0.7$], and background variability of 0.0349) and **b** CT imaging of segmented lesion using 40% maximum standardized uptake value (SUV_{max}) threshold method. The DSC value was 0.79 and %Error was 28.7%. **c** Transverse PET imaging (acquisition time 20 s, reconstructed with relaxation parameters [λ] of Smooth [$\lambda = 0.7$], and background variability of 0.0738) and **d** CT imaging of segmented lesion using 40% SUV_{max} threshold method. The DSC value was 0.79 and %Error was 6.0%

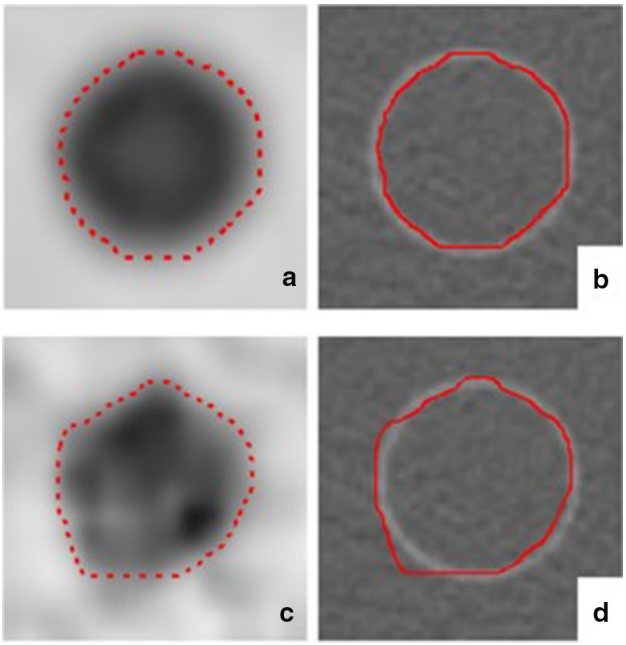

Fig. 4 Sphere images with similar dice similarity coefficient (DSC) and %Error values, despite the varying levels of noise. **a** Transverse PET imaging (acquisition time 30 min, reconstructed with relaxation parameters [λ] of Normal [$\lambda = 1.0$], and background variability of 0.0398) and **b** CT imaging of segmented lesion with gradient method. The DSC was 0.9 and %Error was 0.78%. **c** Transverse PET imaging (acquisition time 20 s, reconstructed with relaxation parameters [λ] of Normal [$\lambda = 1.0$], and background variability of 0.0843) and **d** CT imaging of segmented lesion using gradient method. The DSC was 0.89 and %Error was 0.63%

SUV of 2.5 threshold methods. Based on regression from a sphere diameter of 37 mm to 28, 22, 17, 13, and then 10 mm (gradient method only), the mean %Error of the gradient and SUV of 2.5 threshold methods were significantly increased ($p < 0.01$ and $p < 0.01$, respectively) (Fig. 6). As for the 40% SUV_{max} threshold method, no statistically significant linear trend was observed for mean %Error, while mean %Error for a sphere diameter of 13 mm was significantly greater than that for the sphere diameters of 17 ($p < 0.01$), 22 ($p < 0.01$), 28 ($p < 0.01$), and 37 ($p < 0.01$) mm (Table 1, Fig. 6). The 10-mm sphere could not be segmented from the background lesion with SUV of 2.5 threshold and 40% SUV_{max} threshold methods.

Similarly, there was a statistically significant linear trend for the mean DSC of the three segmentation methods (Table 1, Fig. 7). Based on the regression from a sphere diameter of 37 mm to 28, 22, 17, 13, and then 10 mm (gradient method only), the mean DSC values for the gradient, and 40% SUV_{max}, and SUV of 2.5 threshold methods were significantly increased ($p < 0.01$, $p < 0.01$, $p < 0.01$, respectively).

In addition, Fig. 8 shows the relationship between relaxation parameters and %Error of each sphere with the gradient method (Fig. 8a) and SUV of 2.5 threshold method (Fig. 8b) for the data obtained in an acquisition time of 30 min. With the change of relaxation parameters, %Error of the 10-mm and 13-mm spheres fluctuated the most while using the gradient method (Fig. 8a) and SUV of 2.5 threshold method (Fig. 8b), respectively.

4 Discussion

In the present study, we evaluated the effects of noise and object size on the segmentation accuracy of the gradient, 40% SUV_{max}, and SUV of 2.5 threshold methods on PET imaging using a sphere phantom.

Regarding the SUV of 2.5 threshold method, there were no significant correlations between BGV and %Error or DSC. Although the 3D shape of the VOI might be altered by noise, the effect was not enough to affect the measurement results. However, because of the SUV decline caused by partial volume effect with the decrease of sphere diameter,

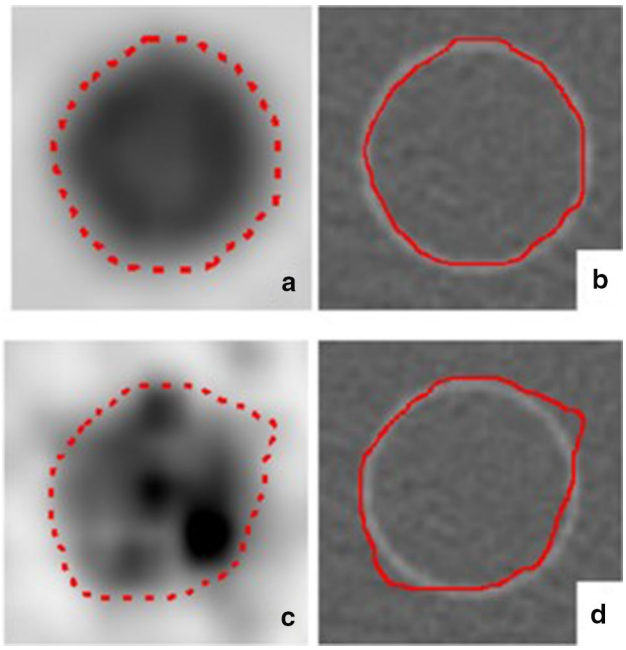

Fig. 5 Sphere images with similar %Error and different dice similarity coefficient (DSC) values. **a** Transverse PET imaging (acquisition time 10 min, reconstructed with relaxation parameters [λ] of Normal [$\lambda = 1.0$], and background variability of 0.0361) and **b** CT imaging of segmented lesion using gradient method. The DSC was 0.96 and %Error was 3.1%. **c** Transverse PET imaging (acquisition time 10 s, reconstructed with relaxation parameters [λ] of Normal [$\lambda = 1.0$], and background variability of 0.1269) and **d** CT imaging of segmented lesion using gradient method. The DSC was 0.79 and %Error was 3.1%

%Error was significantly increased and DSC significantly decreased.

As for the 40% SUV_{max} threshold method, there was a significant correlation between BGV and %Error. As noise increases, there is a gain in SUV_{max}, and a lesion segmented with the 40% SUV_{max} threshold method is determined based on SUV_{max}. Therefore, it is considered that the accuracy of the 40% SUV_{max} threshold method is largely affected by noise. There was no significant correlation between BGV and DSC. Even when the two measured %Error values are different, it is possible that the DSC values may be similar (Fig. 3). Thus, our result showing a relationship between BGV and DSC with the 40% SUV_{max} threshold method might not be relevant. As with the SUV of 2.5 threshold method, partial volume effect affected the segmentation accuracy of the 40% SUV_{max} threshold method.

In the gradient method, there were no significant correlations between BGV and %Error or DSC. Therefore, we concluded that segmentation accuracy with the gradient method is not easily affected by noise (Fig. 4), which is different from the 40% SUV_{max} threshold method. As the sphere diameter decreased, %Error significantly increased and DSC significantly decreased. The partial volume effect had an influence on the segmentation accuracy of the gradient method. Nevertheless, the effect with the gradient method was much lower than that with the other two methods, because the segmentation accuracy for the sphere diameter of 13 mm with the gradient method was much higher, and the 10-mm sphere could be segmented from the background lesion with the gradient method only. The %Error and DSC in the 22- and 17-mm spheres were worse for the gradient method compared with the 40% SUV_{max} threshold method. However, as described in the previous paragraph, the segmentation accuracy for the 40% SUV_{max} threshold method was affected by noise. If the noise on image changes, the %Error and DSC in the 22- and 17-mm spheres for the 40% SUV_{max} threshold method may fluctuate. Therefore, for the segmentation accuracy in the 22- and 17-mm spheres, the 40% SUV_{max} threshold method does not necessarily exceed the gradient method.

Table 1 Relation between %Error or dice similarity coefficient (DSC) of three segmentation methods and sphere diameter

Segmentation method	Sphere diameter (mm)						
	37 ($n=12$)	28 ($n=12$)	22 ($n=12$)	17 ($n=12$)	13 ($n=12$)	10 ($n=12$)	p
Gradient method							
%Error (mean ± SD) (%)	3.0 ± 1.6	5.3 ± 2.6	33.0 ± 9.9	36.3 ± 4.5	13.3 ± 3.5	139.8 ± 16.6	< 0.01
DSC (mean ± SD)	0.89 ± 0.06	0.90 ± 0.04	0.77 ± 0.06	0.75 ± 0.06	0.82 ± 0.06	0.40 ± 0.03	< 0.01
40% SUV_{max} threshold method							
%Error (mean ± SD) (%)	28.9 ± 1.5[a]	32.9 ± 1.1[a]	27.7 ± 3.4[a]	21.2 ± 7.1[a]	271.3 ± 57.8[a]		0.07
DSC (mean ± SD)	0.85 ± 0.04	0.85 ± 0.03	0.85 ± 0.04	0.82 ± 0.05	0.44 ± 0.05		< 0.01
SUV of 2.5 threshold method							
%Error (mean ± SD) (%)	20.6 ± 1.0	30.2 ± 1.4	49.1 ± 3.1	58.4 ± 5.8	72.5 ± 12.7		< 0.01
DSC (mean ± SD)	0.85 ± 0.04	0.81 ± 0.02	0.68 ± 0.03	0.60 ± 0.08	0.39 ± 0.15		< 0.01

[a]%Error in sphere diameter of 13 mm was significantly greater than that in sphere diameters of 17 mm ($p < 0.01$), 22 mm ($p < 0.01$), 28 mm ($p < 0.01$), and 37 mm ($p < 0.01$)

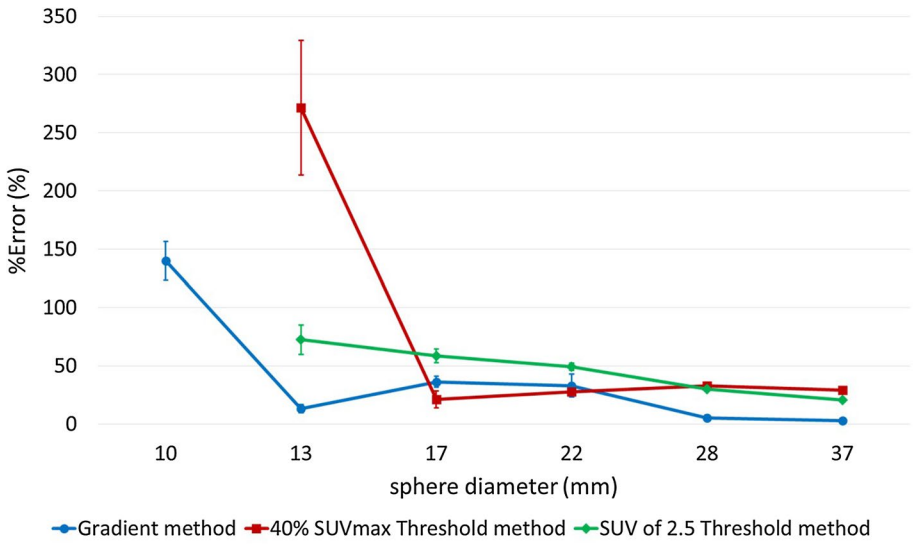

Fig. 6 Correlation of sphere diameter and %Error with the gradient method, 40% maximum standardized uptake value (SUV_{max}) threshold method, and SUV of 2.5 threshold method. Sizes of the error bars indicate standard deviation. The 10-mm sphere could not be segmented from the background lesion with the SUV of 2.5 threshold and 40% SUV_{max} threshold methods

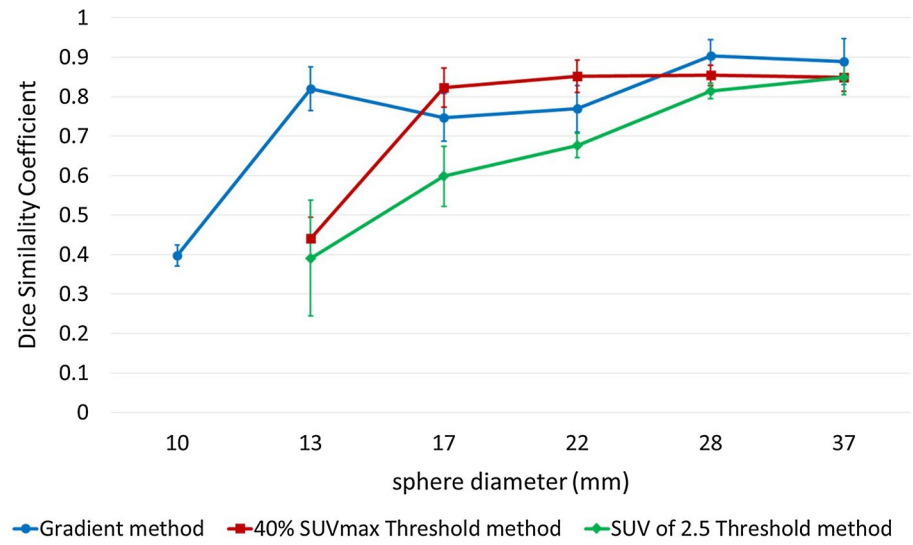

Fig. 7 Correlation of sphere diameter and dice similarity coefficient with the gradient method, 40% maximum standardized uptake value (SUV_{max}) threshold method, and SUV of 2.5 threshold method. Sizes of the error bars indicate standard deviation. The 10-mm sphere could not be segmented from the background lesion with the SUV of 2.5 threshold and 40% SUV_{max} threshold methods

In a previous study, Geets et al. [12] suggested that the gradient method is more accurate than the threshold methods for segmentation with phantoms and patient data. However, they only evaluated the relationship between object size and segmentation accuracy, and the effect of noise was not examined. Foster et al. [14] noted that for segmentation of PET images, resolution-related issues and noise are challenging.

To assess the accuracy of the segmentation method, we used %Error and DSC. In earlier studies, %Error alone was used for the evaluation of the segmentation method [13]; however, that method does not convey enough information to determine the similarity between two segmentations. For example, it is possible to use the segmentation method to measure the same volume as the actual volume (i.e., segmentation leaks into a non-object territory with the same volume, although the true volume can be obtained) (Fig. 5). DSC is used to determine the spatial overlap as the percentage of a segmented lesion from the surrogate. In our study, we could not suggest the obvious utility of DSC for segmentation evaluation. However, in some images, the size did not match between %Error and DSC (Figs. 3, 5). Therefore, the precise evaluation of a segmentation algorithm should be based on DSC and not solely on the absolute volume based on statistical evaluations [14].

In clinical situations, PET images can be blurred by respiratory movement [17]. In the present study, the gradient method and SUV of 2.5 threshold method were not affected by noise. By acquiring PET images with breath-holding for a short period of time, it might be possible to measure MTV more accurately, especially in lung lesions, although the SUV in such cases may be unreliable because of a large volume of noise.

In this study, we did not evaluate segmentation accuracy using complex objects or clinical data. It is possible that

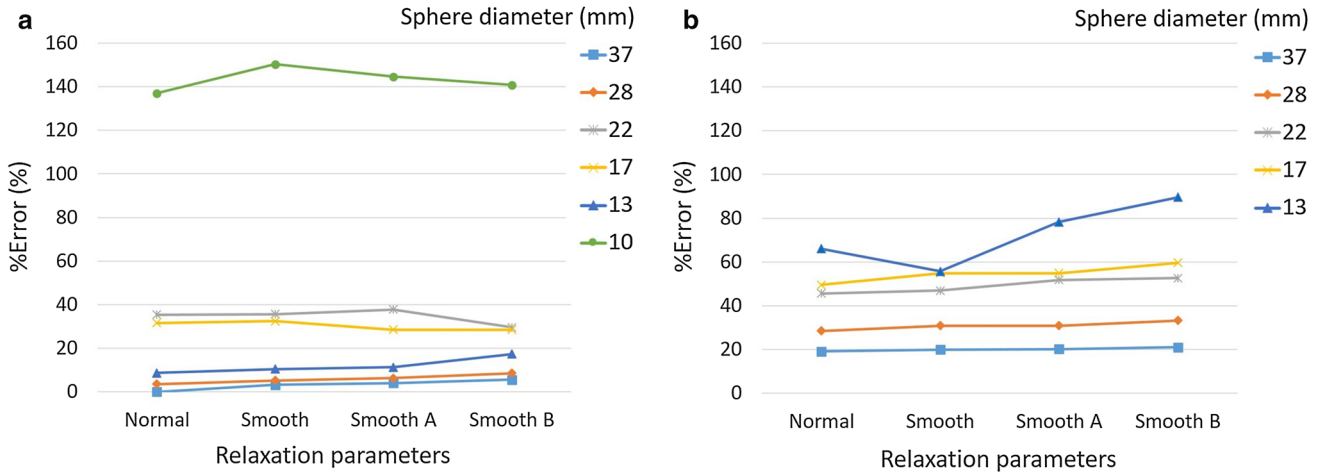

Fig. 8 Relationship between relaxation parameters and %Error of each sphere with the gradient method (**a**) and standardized uptake value (SUV) of 2.5 threshold method (**b**) on the data obtained in an acquisition time of 30 min. The 10-mm sphere could not be segmented from the background lesion with the SUV of 2.5 threshold method

the gradient method and SUV of 2.5 threshold method are affected by noise when examining complex or small objects. Furthermore, the smoothing parameter might have affected the segmentation accuracy, especially for small spheres, because of the partial volume effect (Fig. 8). Moreover, the voxel size in the PET image affects the segmentation accuracy. To evaluate resolution-related issues, the effect of voxel size or the smoothing parameter on segmentation accuracy should be investigated in detail.

5 Conclusion

For segmentation accuracy of FDG-PET imaging, the SUV of 2.5 threshold method was found to be affected by PVE but not by noise. As for the 40% SUV_{max} threshold method, both PVE and noise affected segmentation accuracy; thus, attention is required when segmenting images with different acquisition times or reconstruction parameters when using that method. The gradient method was not affected by noise and PVE had a small effect. Based on our findings, we concluded that the gradient method is more accurate and reliable than either of the threshold methods when applied to images with varying levels of noise or resolution.

Compliance with ethical standards

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

Human and animal rights This article does not contain any studies performed with human participants and animals.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. von Schulthess GK, Steinert HC, Hany TF. Integrated PET/CT: current applications and future directions. Radiology. 2006;238:405–22.
2. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [^{18}F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. Eur J Cancer. 1999;35:1773–82.
3. Basu S, Zaidi H, Holm S, Alavi A. Quantitative techniques in PET-CT imaging. Curr Med Imaging Rev. 2011;7:216–33.
4. Shankar LK, Hoffman JM, Bacharach S, Graham MM, Karp J, Lammertsma AA, et al. Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. J Nucl Med. 2006;47:1059–66.
5. Sugawara Y, Zasadny K, Neuhoff A, Wahl R. Re-evaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. Radiology. 1999;213:521–5.
6. Beaulieu S, Kinahan P, Tseng J, Dunnwald LK, Schubert EK, Pham P, et al. SUV varies with time after injection in (18)F-FDG PET of breast cancer: characterization and method to adjust for time differences. J Nucl Med. 2003;44:1044–50.
7. Srinivas SM, Dhurairaj T, Basu S, Bural G, Surti S, Alavi A. A recovery coefficient method for partial volume correction of PET images. Ann Nucl Med. 2009;23:341–8.
8. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. J Nucl Med. 2007;48:932–45.
9. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50(Suppl 1):122S–50S.
10. Kitajima K, Doi H, Kuribayashi K, Hashimoto M, Tsuchitani T, Tanooka M, et al. Prognostic value of pretreatment volume-based

quantitative $_{18}$F-FDG PET/CT parameters in patients with malignant pleural mesothelioma. Eur J Radiol. 2017;86:176–83.
11. Kitajima K, Miyoshi Y, Yamano T, Odawara S, Higuchi T, Yamakado K. Prognostic value of FDG-PET and DWI in breast cancer. Ann Nucl Med. 2018;32:44–53.
12. Geets X, Lee JA, Bol A, Lonneux M, Grégoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur J Nucl Med Mol Imaging. 2007;34:1427–38.
13. Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, et al. What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol Phys. 2012;82:1164–71.
14. Foster B, Bagci U, Mansoor A, Xu Z, Mollura DJ. A review on segmentation of positron emission tomography images. Comput Biol Med. 2014;50:76–96.
15. Surti S, Kuhn A, Werner ME, et al. Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. J Nucl Med. 2007;48:471–80.
16. Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. Int J Radiat Oncol Biol Phys. 2007;67:720–6.
17. Nehmeh SA, Erdi YE, Ling CC, Rosenzweig KE, Schoder H, Larson SM, et al. Effect of respiratory gating on quantifying PET images of lung cancer. J Nucl Med. 2002;43:876–81.