Size-dependent thresholding as an optimal method for tumor volume delineation on positron emission tomography—computed tomography: A Phantom study

Arun Gupta, Punit Sharma, Chetan D Patel, Sagar Maharjan, Anil Pandey, Rakesh Kumar, Arun Malhotra

Departments of Nuclear Medicine and Health Physics, Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India

ABSTRACT

Background: Use of a fixed threshold value for tumor volume delineation in positron emission tomography (PET) images will ignore the effect of size of the lesion and source to background ratio (SBR). The purpose of this Phantom study was to evaluate the effect of the size of the lesion and SBR on the threshold to be used for PET tumor volume delineation. Materials and Methods: Phantom used in the study comprised a sphere–cylinder assembly containing six spheres of different inner diameters (1.10, 1.35, 1.44, 1.50, 1.83 and 1.93 cm) with inner volumes of 0.70, 1.30, 1.50, 1.77, 3.22 and 3.82 cm³, respectively. The scans were acquired with SBR of 6:01, 7:01, 8:01 and 10:01. These SBRs were calculated from 42 patients with lymphoma to simulate clinical images. PET tumor volume was calculated using RT_Image software at different threshold values (40, 45, 50, 55, 60, 65, 70 and 75% of SUVmax) for each sphere at different SBRs. The threshold intensity value at which the calculated volume was nearly equal to actual volume of spheres was considered as the standardized threshold intensity (STI) value. Results: STI values depended on the diameter of the sphere and not on the SBR. It is found that 40% threshold is suitable for calculating the volume of any lesion with diameter greater than 1.83 cm, 60% for diameter greater than 1.35 cm but less than 1.83 cm, and 75% for diameter less than 1.35 cm. Conclusion: Size-dependent thresholding is an accurate and reproducible method of tumor volume delineation on PET–computed tomography (CT).

Keywords: Positron emission tomography—computed tomography, phantom thresholding, tumor volume

INTRODUCTION

Positron emission tomography (PET) staging is rapidly becoming a standard part of the evaluation of a majority of patients with cancer.[1] PET is used not only as a staging tool but also as a planning tool for external beam radiotherapy (EBRT) and radionuclide therapy. Accurate EBRT/radionuclide therapy, however, requires knowledge of tumor volumes. Also, reproducible measurements of tumor volume are helpful in evaluating the response to therapy and the need for changing treatment plan. So, automated tumor delineation in PET images is highly desirable for improved quantification, objective patient monitoring, and refinement of computed tomography (CT) based treatment planning in radiotherapy.

Anatomic volume estimation is usually based on CT and is considered accurate. Volume measurement with CT is usually straightforward as the margins of the tumor can be well demarcated. However, contouring PET images is difficult given the modest spatial resolution, the relatively high noise level in PET images, the fuzzy tumor edges, and observer subjectivity, as one can vary threshold levels.[2] The first method applied and still widely used to delineate lesion is the visual interpretation of the PET scan and the definition of contours as judged by the experienced nuclear medicine physician.[3,4] Methods based on standardized uptake value (SUV) have been suggested for volume delineation.[5,6] For non small cell lung cancer (NSCLC), for example, an SUV value of ≥2.5 is considered abnormal and highly suspicious for tumor. The use of this method for target delineation, though simplistic, has been thought by some to be problematic as some well-differentiated tumors have an SUV of <2.5 and some benign conditions have an SUV >2.5.[7,8] Small lesions and edges of moving targets have a partial volume effect

Address for correspondence:
Dr. Chetan D Patel, B-54, N.D.S.E Part-I, New Delhi – 110 029, India. E-mail: chetanruptipatel@hotmail.com
and lower SUV, thus possibly underestimating the tumor volume. Moreover, for other tumors such as lymphoma, head and neck, and esophageal cancer, the appropriate SUV value cutoff is not as clear. Constant threshold method, using a fixed percentage of the maximum voxel value in the tumor (SUV$_{\text{max}}$), is the most commonly used method, but there is lack of consensus as to the threshold value to be used and a 5% change in threshold contour level can translate into a 200% increase in volume.$^{39}$ Values of 15–60% have been used with most of the studies using a value of 40%.$^{10-15}$ Use of a single threshold value for all the tumors may not be accurate, as this ignores the effect of size of the lesion (partial volume effect) and source to background ratio (SBR).

The purpose of this Phantom study was to evaluate the effect of the size of the lesion and source to background ratio on the threshold to be used for PET tumor volume delineation.

**MATERIALS AND METHODS**

**Phantom specification**

Phantom that was used in the study was locally fabricated, and made of tissue equivalent material. It was designed to have hot spheres in warm background model to simulate the actual condition in PET study of the patients. It consisted of two parts: a hollow cylinder and sphere assembly. The diameter and height of the body of Phantom was 18.5 and 20 cm, respectively. The sphere assembly contained six spheres of different inner diameters (1.10, 1.35, 1.44, 1.50, 1.83 and 1.93 cm). The inner volumes of the spheres were 0.70, 1.30, 1.50, 1.77, 3.22 and 3.82 cm$^3$, respectively. These spheres were placed on the groove in the disk and sandwiched between the two disks to restrict the movement when the cylinder was filled with water. The largest sphere was placed at the center of the disk and the remaining five spheres were placed at the periphery. The sphere assembly was placed at the middle position of the cylinder.

**Source to background ratio**

To simulate the tumor to background ratio (TBR) in patients, SBRs were first calculated for 154 known lesions in 42 patients with lymphoma. The patient imaging data were loaded one by one in the RT_Imag software. The RT_Imag package runs on the Interactive Data language, Virtual Machine 7 (IDL, Research System Inc., Boulder, CO, USA). The program reads the F-18 fluorodeoxyglucose (FDG) PET images in DICOM format and also converts the intensity values automatically to SUVs. Elliptical regions of interest (ROIs) were drawn over the lesions and background area on the axial images [Figure 1]. SUV$_{\text{max}}$ of lesion and background was noted as shown by ROI statistics [Figure 1] respectively. The ratio of SUV$_{\text{max}}$ of lesion to background for all 154 lesions sites was calculated in the same way. The range of TBR was 2:1 to 36:1 (mean 10:1, median 9:1, SD).

**Preparation of Phantom**

Phantom studies were done with 6:1, 7:1, 8:1 and 10:1 SBRs. First, the cylinder was filled with water containing 0.909 µCi F-18 FDG/cm$^3$. The activity in the spheres for different SBRs was calculated and the spheres were then filled with water containing calculated concentration of activity [Table 1].

**Image acquisition**

The scans were acquired with a dedicated PET-CT scanner (Biograph 2, Siemens, Erlangen, Germany). It has Lutetium Oxyorthosilicate (LSO; Lu2SiO5:Ce) detectors, with an attenuation coefficient of 0.89/cm, photo fraction of 30%, and a decay constant of 40 nsec. The energy resolution at 511 KeV (% full width at half maximum (FWHM)) is 10 with a spatial resolution of 6 mm. In the PET/CT system, CT acquisition was performed on spiral dual slice CT. The CT parameters were 130 kVp, 30 mA, scan time 5 sec, rotation time 0.8 sec, slice width 5 mm, effective milliamperes 50 mAs and kernel T 20s standard. After CT, 3D PET acquisition was done for 3 minutes per bed position for two beds. The matrix size was 128 × 128 and no zoom was employed. CT-based attenuation correction of the emission images was employed. PET images were reconstructed by iterative method ordered subset expectation maximization (OSEM; two iterations and eight subsets). After completion of PET acquisition, the reconstructed attenuation corrected PET images, CT images and fused images of matching pairs of PET and CT images were available for review in axial, coronal and sagittal planes, as well as in maximum intensity projections, three-dimensional cine mode.

**Image processing**

PET imaging data of Phantom study were also analyzed with RT_
Image software. The SUV$_{\text{max}}$ values of spheres were determined by drawing elliptical ROIs over the spheres in multiple axial slices. At first the images of the Phantom were processed at a threshold of 40% of SUV$_{\text{max}}$. Volume of a particular sphere was calculated by drawing elliptical ROIs on all slices containing hot images of that sphere at that threshold value (40%), [Figures 2 and 3]. This process was repeated at different threshold values (45, 50, 55, 60, 65, 70 and 75% of SUV$_{\text{max}}$). The same method was applied for the remaining spheres (different diameters) of that Phantom and for Phantom studies at different SBRs. The threshold intensity value at which the calculated volume was nearly equal to actual volume of spheres was considered as the standardized threshold intensity (STI) value.

**RESULTS**

Table 2 shows SUVs and the volumes of the spheres at different thresholds compared to the actual volume. At the SBR of 6:1, the SUV$_{\text{max}}$ of the spheres ranged from 2.6 to 8.86. For the two largest spheres with diameters 1.93 and 1.83 cm, 40% threshold resulted in approximately the same volume as its actual volume with less than 3% variation and a threshold above 40% underestimated the actual volume. For the spheres with diameters 1.50, 1.44 and 1.35 cm, 40, 45, 50 and 55% threshold values overestimated the volume of spheres. Thresholding at 60% resulted in the calculated volumes that matched the actual volumes of spheres, with less than 7% variation. Above this threshold value, again the calculated volumes underestimated the actual volume. For the smallest sphere with diameter 1.10 cm, 40–70% thresholding values overestimated the sphere volume. Thresholding at 75% gives nearly equal volume as the actual volume, with a variation of nearly 6%. Similar findings were noted at the SBR of 7:1, 8:1 and 10:1. To summarize the observation, for the two large spheres, 40% threshold gave the best result; for the smallest sphere, 75% threshold best approximated the actual volume, while for the rest of the spheres, a threshold of 60% was the best.

| SBR | Size of sphere | Volume of sphere | SUV$_{\text{max}}$ | Volume at different thresholds |
|-----|----------------|------------------|-------------------|-------------------------------|
|     |                |                  |                   | 40%  | 45%  | 50%  | 55%  | 60%  | 65%  | 70%  | 75%  |
| 6:01| 1.93           | 3.82             | 8.86              | 3.9  | 3.42 | 2.95 | 2.75 | 2.57 | 1.8  | 1.43 | 1.04 |
|     | 1.83           | 3.22             | 8.4              | 3.23 | 3.04 | 2.75 | 2.57 | 2.28 | 1.9  | 1.24 | 0.95 |
|     | 1.5           | 1.77             | 6.5             | 3.14 | 2.66 | 2.19 | 2.09 | 1.66 | 1.14 | 0.85 | 0.57 |
|     | 1.44           | 1.5             | 5.25             | 3.23 | 2.76 | 2.19 | 1.71 | 1.42 | 1.24 | 1.14 | 0.66 |
|     | 1.35           | 1.3             | 5.12             | 3.14 | 2.19 | 2.09 | 1.52 | 1.24 | 1.14 | 0.95 | 0.57 |
|     | 1.1           | 0.7             | 2.6             | 3.61 | 3.52 | 3.23 | 1.9  | 1.33 | 1.14 | 0.95 | 0.66 |
| 7:01| 1.93           | 3.82             | 6.16             | 3.9  | 3.71 | 3.61 | 2.66 | 2.57 | 2.18 | 1.62 | 1.14 | 0.95 |
|     | 1.83           | 3.22             | 6.06             | 3.42 | 2.95 | 2.66 | 2.57 | 2.18 | 1.62 | 1.14 | 0.95 |
|     | 1.5           | 1.77             | 4.22             | 2.94 | 2.47 | 2.18 | 2   | 1.61 | 1.33 | 1.14 | 0.85 |
|     | 1.44           | 1.5             | 3.84             | 3.23 | 2.28 | 2   | 1.62 | 1.42 | 1.24 | 0.95 | 0.76 |
|     | 1.35           | 1.3             | 3.95             | 2.76 | 2.38 | 1.71 | 1.52 | 1.24 | 0.95 | 0.76 | 0.57 |
|     | 1.1           | 0.7             | 2.24             | 3.23 | 2.28 | 1.71 | 1.43 | 1.14 | 0.95 | 0.76 | 0.67 |
| 8:01| 1.93           | 3.82             | 5.01             | 3.9  | 3.33 | 3.14 | 2.57 | 2.38 | 2.09 | 1.81 | 1.14 |
|     | 1.83           | 3.22             | 4.67             | 3.33 | 3.04 | 2.76 | 2.38 | 1.81 | 1.52 | 1.33 | 1.14 |
|     | 1.5           | 1.77             | 3.5             | 3.14 | 2.28 | 2   | 1.9  | 1.71 | 1.14 | 0.86 | 0.57 |
|     | 1.44           | 1.5             | 3.46             | 2.66 | 2.09 | 1.9  | 1.62 | 1.43 | 1.25 | 0.94 | 0.77 |
|     | 1.35           | 1.3             | 3.08             | 2.47 | 2.09 | 1.9  | 1.52 | 1.2  | 1.05 | 0.85 | 0.56 |
|     | 1.1           | 0.7             | 1.47             | 4.75 | 2.95 | 2.66 | 1.62 | 1.43 | 1.24 | 0.95 | 0.64 |

SBR: source to background ratio

Figure 2: Image showing elliptical ROIs drawn over magnified images of hot spheres (D=1.93 and 1.83 cm) of the Phantom on axial slices

Figure 3: Image showing volume of sphere calculated by STI method. ROIs were drawn over all axial slices containing hot image of the sphere (D=1.93 cm) of Phantom. ROI statistics shows the volume of sphere calculated at each slice and total volume at 40% of SUV$_{\text{max}}$
Based on these Phantom studies, the STI values were determined [Table 3]. It was observed that the STI values depended on the diameter of the sphere and not on the SBR. From the analysis, it is seen that 40% threshold is suitable for calculating the volume of any lesion greater than 1.83 cm and 60% is suitable for a diameter greater than 1.35 cm and less than 1.83 cm. Similarly, 75% threshold values must be applied to calculate the volume of lesion with diameter less than 1.35 cm.

**DISCUSSION**

The use of FDG-PET for the delineation of the tumor volume is becoming increasingly popular in oncology for the estimation of metabolic tumor burden and for radiotherapy planning.[14] Although many methods have been proposed for tumor volume delineation in F-18 FDG PET images, there is no consensus regarding which method should be preferred.[15–17] Comparing the performance of these methods from the data published in the literature is almost impossible given the variety of situations in which evaluation studies have been conducted. The easiest method to delineate the outline of the FDG avid region is the visual interpretation of an experienced nuclear medicine physician.

However, application of FDG-PET data for target-volume delineation is not straightforward as identification of tumor boundaries on PET suffers from a relative low spatial resolution and a “blurry” appearance of lesions, leading to high level of subjectivity and intra and interobserver variability.[18] Authors have investigated the use of fixed threshold methods for segmenting PET images for better reproducibility. Schinagl et al. compared in 78 patients with Stages II–IV squamous cell carcinoma of the head and neck, gross tumor volume (GTV) delineated on CT with five PET-based GTVs obtained by visual interpretation (GTVVIS), applying an isocontour of a standardized uptake value of 2.5 (GTV2.5), using a fixed threshold of 40% (GTV40%) and 50% (GTV50%) of the maximum signal intensity, and applying an adaptive threshold based on the signal-to-background ratio (GTVsbr).[19] Volume and shape of the PET-GTV were influenced heavily by the choice of segmentation tool. Mean GTV50% and GTVsbr were very similar; GTV50% was less than GTV40%, which was less than GTVVIS which in turn was less than GTV2.5. Biehl et al. found that appropriate threshold for PET-GTV delineation in patients with nonsmall cell lung cancer is highly dependent on tumor size.[19] The optimal thresholds (mean±SD) were 15 ± 6% for tumors measuring greater than 5 cm, 24±9% for tumors measuring 3–5 cm, and 42±2% for tumors measuring less than 3 cm. Erdi et al. studied the effect of size of the lesion and variable background activity on the volume estimated from PET images of elliptical Jaszcak Phantom.[20] For sphere volumes larger than 4 ml, the best threshold value ranged from 36 to 44%, with the exact value depending on the SBRs. For smaller volumes (<4 ml), a decrease in the SBR increases the optimum threshold levels. Applying the optimum threshold values derived from the Phantom experiments in the studies performed on 10 patients with 17 primary or metastatic lung lesions, a good correlation between the volume from CT and the volume from PET study was noted (r=0.999, P= 0.02). Jentzen et al. used an iterative thresholding method (ITM) to estimate the PET volumes without anatomical prior knowledge and its application to clinical images.[21] ITM was based on threshold-volume curves at varying source-to-background (S/B) ratios acquired from a body Phantom. The measured S/B ratios of the lesions were estimated from PET images, and their volumes were iteratively calculated using the calibrated S/B-threshold-volume curves. The resulting PET volumes were then compared with the known sphere inner volume and CT volumes of tumors. ITM sufficiently estimated the clinical volumes in the range of 0.8–7.5 ml. Volumes larger than 7.5 ml showed greater deviations that are still acceptable. These findings are associated with the limitation of the ITM. The ITM is especially useful for lesions that are only visible on PET.[22]

Our study demonstrates that SBR does not have significant effect on the estimation of volumes from PET images in the range of SBR encountered in patients with lymphoma. The only determining factor for the threshold for PET volume estimation was the size of the sphere. For spheres ≥1.83 cm in diameter (corresponding to the volume of 3.22 ml), a fixed threshold of 40% resulted in approximately the same volume as its actual volume. This finding is similar to that reported by Erdi et al. However, we did not find an effect of variable SBR on the appropriate threshold values for spheres of smaller size.

The present study has certain limitations. Firstly, the SBR was calculated from patients with lymphoma which usually have high FDG uptake. Hence, the SBR in our study were higher, thus providing a limited range. Secondly, the Phantom used in the study had less number of spheres without much variation in the diameter (range between 1.1 and 1.93 cm). A further study incorporating a varied range of SBR and diameter of spheres may be more useful to simulate the routinely encountered clinical spectrum.

**CONCLUSION**

Size-dependent thresholding is an accurate and reproducible method of tumor volume delineation on PET–CT. Application of this method in varied clinical settings will further reinforce its usefulness in PET–CT imaging.
REFERENCES

1. Fletcher JW, Djulbegovic B, Soares HP, Siegel BA, Lowe VJ, Lyman GH, et al. Recommendations on the use of F-18 FDG PET in oncology. J Nucl Med 2008;49:480-508.

2. Pöttsch C, Hofheinz F, van den Hoff J. Vergleich der Inter-observer-variabilität bei manueller und automatischer volumenbestimmung in der PET. Nuklearmedizin 2006;45:A42.

3. Nestle U, Walter K, Schmidt S, Licht N, Nieder C, Motaref B, et al. 18 F-Deoxyglucose positron emission tomography (18 FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. Int J Radiat Oncol Biol Phys 1999;44:593-7.

4. Kiffer JD, Berlangieri SU, Scott AM, Quong G, Feigen M, Schumer W, et al. The contribution of 18 F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer. Lung Cancer 1998;19:167-77.

5. Paulino AC, Johnstone PAS. FDG-PET in radiotherapy treatment planning: Pandora’s box? Int J Radiat Oncol Biol Phys 2004;59:4-5.

6. Wang D, Schultz CJ, Jursinic PA, Bialkowski M, Zhu XR, Brown WD, et al. Initial experience of FDG-PET/CT guided IMRT of head and neck carcinoma. Int J Radiat Oncol Biol Phys 2006;65:143-51.

7. Paulino AC, Johnstone PA. Does SUV stand for silly useless value? Int J Radiat Oncol Biol Phys 2004;60:1006.

8. MacManus M, Hicks R, Bayne M, Leong T, Peters I, Ball D, et al. In regard to Paulino and Johnstone: Use of PET and CT imaging data in radiation therapy. Int J Radiat Oncol Biol Phys 2004;60:1005-6.

9. Ford EC, Kinahan PE, Hanlon L, Alessio A, Rejendran J, Schwartz DL, et al. Tumor delineation using PET in head and neck cancers: Threshold contouring and lesion volumes. Med Phys 2006;33:4280-8.

10. Biehl KJ, Kong FM, Dehdashti F, Jin JY, Mucic S, El Naqa I, et al. F-18 FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: Is a single standardized uptake value threshold approach appropriate? J Nucl Med 2006;47:1808-12.

11. Bradley J, Thorstad WL, Mucic S, Miller TR, Dehdashti F, Siegel BA, et al. Impact of FDG-PET on radiation therapy volume delineation in non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2004;59:78-86.

12. Miller T, Grigsby P. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated with radiation therapy. Int J Radiat Oncol Biol Phys 2002;53:353-9.

13. Erdi Y, Rosenzweig K, Erdi A, Macapinlac HA, Hu YC, Bhaban LF, et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). Radiother Oncol 2002;62:51-60.

14. Mah K, Caldwell CB, Ung YC, Danjoux CE, Balogh JM, Ganguli SN, et al. The impact of F-18 FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. Int J Radiat Oncol Biol Phys 2002;52:339-50.

15. Brianzoni E, Rossi G, Aneidei S, Berbellini A, Capoceri F, Cidda C, et al. Radiotherapy planning: PET/CT scanner performances in the definition of gross tumour volume and clinical target volume. Eur J Nucl Med Mol Imaging 2005;32:1392-9.

16. Grégoire V, Haustermans K, Geets X, Roels S, Lonneux M. PET-based treatment planning in radiotherapy: A new standard? J Nucl Med 2007;48:68S-77.

17. Riegel AC, Berson AM, Destian S, Nq T, Tena LB, Mitnick RJ, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. Int J Radiat Oncol Biol Phys 2006;65:726-32.

18. Schinagl DA, Vogel WV, Hoffmann AL, van Dalen JA, Oyen WJ, Kaanders JH. Comparison of five segmentation tools for F-18-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer. Int J Radiat Oncol Biol Phys 2007;69:1282-9.

19. Erdi YE, Mawlawi O, Larson SM, Imbriaco M, Yeung H, Finn R, et al. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer 1997;80:2505-9.

20. Jentzen W, Freudenburg L, Eising EG, Sonnenschein W, Knust J, Bockisch A, et al. Segmentation of PET volumes by iterative image thresholding. J Nucl Med 2007;48:108-14.

21. Berkowitz A, Basu S, Sinavas S, Sankaran S, Schuster A, Alavi A, et al. Determination of whole-body metabolic burden as a quantitative measure of disease activity in lymphoma: A novel approach with fluorodeoxyglucose-PET. Nuc Med Commun 2008;29:521-6.

How to cite this article: Gupta A, Sharma P, Patel CD, Maharjan S, Pandey A, Kumar R, Mahotra A. Size-dependent thresholding as an optimal method for tumor volume delineation on positron emission tomography-computed tomography: A Phantom study. Indian J Nucl Med 2011;26:22-6.

Source of Support: Nil. Conflict of Interest: None declared.