Phantom Study on Three-Dimensional Target Volume Delineation by PET/CT-Based Auto-Contouring

Zhang, Tejiao
Division of Radiological Science, Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University

Tachiya, Yosuke
Department of Medical Technology, Kyushu University Hospital

Sakaguchi, Yuichi
Division of Radiological Science, Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University

Mitsumoto, Katsuhiko
Division of Radiological Science, Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University

https://doi.org/10.15017/18984
Original Article

Phantom Study on Three-Dimensional Target Volume Delineation by PET/CT-Based Auto-Contouring

Tiejiao ZHANG1, Yosuke TACHIBANA2, Yuichi SAKAGUCHI3, Katsuhiko MITSUMOTO1, Tatsuya MITSUMOTO1, Nobuyoshi OHYA2 and Masayuki SASAKI1
1)Division of Radiological Science, Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University
2)Department of Medical Technology, Kyushu University Hospital

Abstract Objective: The aim of this study was to determine an appropriate threshold value for delineation of the target volume in PET/CT and to investigate whether we could delineate a target volume by phantom studies.

Methods: A phantom consisted of six spheres (φ 10–37 mm) filled with 18F solution. Data acquisition was performed PET/CT in non-motion and motion status with high 18F solution and in non-motion status with low 18F solution. In non-motion phantom experiments, we determined two types of threshold value, an absolute SUV (TSUV) and a percentage of the maximum SUV (T%). Delineation using threshold values was applied for all spheres and for selected large spheres (a diameter of 22 mm or larger). In motion phantom experiments, data acquisition was performed in a static mode (sPET) and a gated mode (gPET). CT scanning was performed with helical CT (HCT) and 4-dimentional CT (4DCT).

Results: The appropriate threshold values were aT%=27% and aTSUV=2.4 for all spheres, and aT%=30% and sTSUV=4.3 for selected spheres. For all spheres in sPET/HCT in motion, the delineated volumes were 84%–129% by the aT% and 34%–127% by the aTSUV. In gPET/4DCT in motion, the delineated volumes were 94%–103% by the aT% and 51%–131% by the aTSUV. For low radioactivity spheres, the delineated volumes were all underestimated.

Conclusion: A threshold value of T%=27% was proposed for auto-contouring of lung tumors. Our results also suggested that the respiratory gated data acquisition should be performed in both PET and CT for target volume delineation.

Key words: PET/CT, radiotherapy planning, threshold, auto-contouring, gross tumor volume, phantom study, FDG

Introduction

The metabolic image obtained by [18F] fluoro-2-deoxy-glucose (FDG)–positron emission tomography (PET) is now widely used for tumor diagnosis. The diagnostic ability of FDG-PET is considered to be superior to that of anatomic images in the differential diagnosis, staging, response evaluation, and predicting prognosis of many malignant tumors. FDG-PET has been reported to be especially useful for the clinical management of non-small cell lung cancer (NSCLC). Regarding tumor detection, FDG-PET shows higher sensitivity in detecting lymph node metastasis and distant metastasis than does conventional imaging in patients with NSCLC. Regarding exclusion of tumor involvement, FDG-PET can differentiate between malignant tissue and non-malignant tissue.

Recently, developments in radiation treatment
(RT) have improved the prognosis and quality of life (QOL) of patients with lung cancer\(^{12}\). Highly precise radiotherapies such as three-dimensional conformal radiation therapy (3D-CRT) and intensity-modulated radiation therapy (IMRT) require strict RT planning to focus the radiation dose on the tumor and to avoid unnecessary irradiation of normal tissue. The excellent image contrast of PET/CT led to the idea of using it in RT planning for patients with NSCLC\(^ {11,13}\). Using PET/CT in RT planning provides higher sensitivity of PET/CT results, focusing higher radiation doses on treatment volumes. Furthermore, the high specificity of PET/CT is considered to reduce unnecessary irradiation, including that of normal tissues. FDG-PET/CT imaging has been reported to significantly reduce inter-observer variation of the gross tumor volume (GTV) by the threshold setting\(^ {14}\). To date, however, no standardized method has been validated for the delineation of GTV by PET/CT using a threshold. The respiration-associated motion in PET/CT imaging is also a big problem when applying the PET/CT images for RT planning\(^ {15–17}\).

The aim of our study was to determine an appropriate threshold value for delineation of the target volume in PET/CT by phantom studies. We also investigated whether we could delineate a target volume in motion with this technology.

**Materials and methods**

**Phantom and Motion Table**

In this study, the NEMA (National Electrical Manufacturers Association) IEC (International Electrotechnical Commission) Body Phantom Set consisted of an ellipse cavity and 6 spheres (Model ECT/IEC–BODY/P). The internal diameters of the spheres were 10, 13, 17, 22, 28 and 37 mm, and the actual volumes of the spheres were 0.5, 1.2, 2.6, 5.6, 11.5, and 26.5 cm\(^3\), respectively. The wall thickness of the spheres was 1 mm. All spheres were filled with \(^{18}\)F solution (36.7 kBq/ml), and the background was filled with air. The radioactivity was equal to the standardized uptake value (SUV) of 15. The \(^{18}\)F solution of 6.61 kBq/ml (SUV = 2.7) was also used to examine low radioactivity subjects. The radioactivity of spheres with a diameter of 17 mm or smaller was underestimated due to the partial volume effect\(^ {18}\).

Thus this study was performed on two objective groups: one group consisted of all spheres (all), and the other consisted of spheres with a diameter of 22 mm or larger (selected).

To simulate respiratory motion, the NEMA phantom was placed on a motor–driven moving table of our own making. The moving table oscillated with a displacement of 2 cm in the cranio-caudal direction and a frequency of 15 cycles/1 minute. The motion tracking was recorded by a real-time position management (RPM) respiratory gating system (Varian Medical Systems, Palo Alto, CA). The RPM system includes a video camera that measures respiration by tracking the displacement of two infrared reflective markers rigidly mounted on a plastic block.

**PET/CT Protocol**

PET/CT scans were acquired in both a non-moving status (non-motion) and a moving status (motion) on a Discovery ST Elite scanner (GE Healthcare, Milwaukee, WI). Emission data were acquired in the 3D mode with 128 × 128 matrices (5.47 × 5.47 × 3.27 mm). The acquisition times per bed position were 3 minutes in static mode (sPET) and 8 minutes in gated mode (gPET). PET images were reconstructed using a 3D ordered subsets–expectation maximization (3D–OSEM) algorithm (VUE Point Plus, 2 iterations, 28 subsets, and a post–filter of 6 mm full width at half maximum (FWHM)). The gPET was divided into 8 respiratory phases using RPM data.

The 16–slice CT scanning was performed at 120 kV and auto mA (30–154 mA). The trans–axial field of view was 500 mm for helical CT (HCT) with a slice thickness of 5 mm and for 4–Dimen-
sional CT (4DCT) with a slice thickness of 3.75 mm. CT images were acquired using a matrix of 512 × 512 pixels. Gantry rotation was 0.5 sec/rotation for conventional HCT and 0.5 sec/rotation in the cine mode for 4DCT. The 4DCT was divided into 8 respiratory phases using RPM data.

The PET/CT images in non-motion were reconstructed with a combination of sPET and HCT data and were used as a standard. The sPET in motion was reconstructed with HCT (sPET/HCT), and the gated PET (gPET) with 4DCT (gPET/4DCT).

**Determination of the threshold value**

PET/CT images were analyzed using a workstation for RT planning (Eclipse; Varian Medical Systems, Palo Alto, CA). Reconstructed PET/CT images were converted to digital imaging and communication in medicine (DICOM) format and transferred to the Eclipse workstation. The threshold value to delineate the target volume was determined using sPET in non-motion (SUV = 15) as follows. In each sphere, a series of slices in the axial images was analyzed. First, the highest SUV in a sphere was determined as SUVmax. Next, the sphere volume in PET was measured at some different threshold values of the absolute SUV from 15.0 to 1.0 at 0.1 intervals. The threshold value whose delineated sphere volume was closest to the actual volume without any underestimation was determined as the appropriate threshold value. We examined two types of threshold value, an absolute SUV (T_{SUV}) and a percentage of the maximum SUV (T_{%}). The thresholds were determined by using all spheres (aT_{SUV}, aT_{%}) and by using spheres with a diameter of 22 mm or larger (sT_{SUV}, sT_{%}).

**Auto-contouring by the threshold value on different PET images**

To extract the sphere volume automatically, we applied the determined thresholds in non-motion for PET images in motion with different acquisition modes sPET/HCT and gPET/4DCT, and also for those in non-motion with low SUVs. The contoured sphere volumes by the threshold were compared with each actual volume and the relative volume (%V) expressed as \( V_c/V_a \times 100 \) (Vc: contoured volume, Va: actual volume).

**Results**

**Determination of the threshold value in non-motion**

In non-motion, the appropriate threshold values that can provide the sphere volumes nearest to actual volumes without any underestimation are shown in Table 1. The threshold value of spheres differed depending on sphere size; the smaller the sphere size, the smaller the threshold value. Based on the tests we performed, the appropriate threshold values were determined as aT_{%} = 27% and aT_{SUV} = 2.4 for all spheres, and sT_{%} = 30% and sT_{SUV} = 4.3 for selected spheres (Table 2).

The obtained %V of each sphere in non-motion, using these threshold values, is shown in Table 3 and Fig. 1. The %Vs obtained by aT_{%}, sT_{%}, and sT_{SUV} were less than 109% ; however, those by aT_{SUV} were large (ranging from 103% to 147%).

**Table 1** Threshold value determined for each sphere

<table>
<thead>
<tr>
<th>Sphere diameter</th>
<th>10mm</th>
<th>13mm</th>
<th>17mm</th>
<th>22mm</th>
<th>28mm</th>
<th>37mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_{%}</td>
<td>27.0</td>
<td>27.6</td>
<td>27.4</td>
<td>30.1</td>
<td>30.0</td>
<td>33.2</td>
</tr>
<tr>
<td>T_{SUV}</td>
<td>2.4</td>
<td>3.4</td>
<td>3.8</td>
<td>4.3</td>
<td>4.5</td>
<td>5.0</td>
</tr>
</tbody>
</table>

**Table 2** The approximate threshold values

<table>
<thead>
<tr>
<th>For all spheres (φ 10–37mm)</th>
<th>For selected spheres (φ 22–37mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_{%}</td>
<td>aT_{%}= 27%</td>
</tr>
<tr>
<td>T_{SUV}</td>
<td>aT_{SUV}= 2.4</td>
</tr>
<tr>
<td>sT_{%}</td>
<td>sT_{%}= 30%</td>
</tr>
<tr>
<td>sT_{SUV}</td>
<td>sT_{SUV}= 4.3</td>
</tr>
</tbody>
</table>

**Table 3** The auto-contoured sphere volume (%V) in non-motion

<table>
<thead>
<tr>
<th>Diameter (mm)</th>
<th>aT_{%} (27%)</th>
<th>sT_{%} (30%)</th>
<th>aT_{SUV} (2.4)</th>
<th>sT_{SUV} (4.3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>101</td>
<td>103</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>13</td>
<td>104</td>
<td>147</td>
<td>147</td>
<td>147</td>
</tr>
<tr>
<td>17</td>
<td>101</td>
<td>141</td>
<td>141</td>
<td>141</td>
</tr>
<tr>
<td>22</td>
<td>108</td>
<td>141</td>
<td>141</td>
<td>101</td>
</tr>
<tr>
<td>28</td>
<td>107</td>
<td>135</td>
<td>135</td>
<td>102</td>
</tr>
<tr>
<td>37</td>
<td>109</td>
<td>128</td>
<td>128</td>
<td>106</td>
</tr>
</tbody>
</table>

%V = V_c / V_a × 100; Vc : contoured volume, Va : actual volume
Auto-contouring in motion

The %V of auto-contoured sphere volumes in motion are shown in Table 4. For all spheres in sPET/HCT, the %V ranged from 84% to 129% by the autocontouring (Fig. 1) and from 34% to 127% by the automatic threshold (aT%). A sphere with a diameter of 10 mm was not visualized in sPET/HCT, and thus its volume could not be contoured. For selected spheres, both sT% and aTSV underestimated the %V. For all spheres in gPET/4DCT, the %Vs were 94–103% by the aT% (Fig. 1) and 51–131% by the automatic threshold. For selected spheres, all the auto-contoured sphere volumes were underestimated by either threshold (sT% or aTSV).

Auto-contouring of low SUV spheres in non-motion

The contoured sphere volumes with low SUVs in non-motion are shown in Table 5. The %V of most spheres obtained by using aT% was underestimated, and only the sphere of 37-mm diameter was accurately contoured. The sT% also underestimated the %V. The aTSV contoured extremely small volumes, ranging from 0% to

---

**Table 4** The auto-contoured sphere volume (%V) in motion

<table>
<thead>
<tr>
<th>Diameter (mm)</th>
<th>sPET/HCT</th>
<th>gPET/4DCT*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>aT% (27%)</td>
<td>sT% (30%)</td>
</tr>
<tr>
<td>10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>106</td>
<td>34</td>
</tr>
<tr>
<td>17</td>
<td>129</td>
<td>120</td>
</tr>
<tr>
<td>22</td>
<td>99</td>
<td>87</td>
</tr>
<tr>
<td>28</td>
<td>84</td>
<td>75</td>
</tr>
<tr>
<td>37</td>
<td>94</td>
<td>87</td>
</tr>
</tbody>
</table>

%V = Vc/Va × 100; Vc: contoured volume, Va: actual volume

*The mean %V of 8 phases
10%, because the $a_{TSU}$ was close to the radioactivity of the spheres (SU = 2.4). The $s_{TSU}$ did not contour the sphere volume at all, because the $s_{TSU} = 4.3$ was higher than the sphere SUV.

### Discussion

In this study, we investigated the threshold value for auto-contouring a 3D target volume using hot sphere phantoms. A threshold value of $a_{TS} = 27\%$ was considered to be appropriate in non-motion. However, this value could not be applied to a low SUV sphere or to a moving phantom. We also showed that gated acquisition is important in motion to use PET/CT for RT planning.

In phantom studies, a threshold determined by a percentage of the maximum SUV ($T_m$) has been reported to be useful. However, the values ranged widely, from 15% to 42%. Erdi et al. proposed a value of $T_m = 42\%$, because their study demonstrated the most accurate result for spherical phantoms with sizes larger than 4 cm$^3$. Recently, Okubo et al. reported that the PET delineations with a 35% threshold fit well to the actual spheres in an axial plane for both static and moving phantoms. The subjects of their study were selected spheres (22–37 mm), and thus their threshold of 35% should be compared with that of our $s_{TS}$ (30%). The $s_{TS}$ in our study was a little lower than their threshold. Although the reason for this difference has not been clarified, the presence of $^{18}$F solution in the background in phantom studies may have induced the difference.

Some authors have investigated the absolute SUV as a threshold. Because a SUV of 2.5 is generally considered to help distinguish between malignant and benign tissues, the use of a SUV of 2.5 as a threshold for GTV delineation has been investigated. In most papers, a threshold of SUV of 2.5 tended to overestimate tumor volume. In our study, the sphere volumes were overestimated with the threshold of $a_{TSU} = 2.4$. At the same time, higher SUVs, such as $s_{TSU} = 4.3$, underestimated the volume of spheres with a diameter less than 22 mm. These facts indicate that a lower value may be appropriate. In general, $T_{SU}$ is considered to underestimate volumes of small lesions, because the SUVs of small lesions are underestimated due to the partial volume effect. Furthermore, it is not considered to be used clinically, because the SUVs of some NSCLCs, for example, a bronchiolar alveolar carcinoma, are usually less than 2.5.

The radioactivities of small lesions are usually underestimated in PET images, due to poor spatial resolution. Full recovery can be obtained in lesions with a diameter of 3 times the FWHM or larger. The recovery coefficient of 90% or higher was observed with hot spheres with 22-mm diameter or larger in non-motion. This result was consistent with those provided in a previous report. In our study, an appropriate threshold value was $a_{TS} = 27\%$ for all spheres (φ 10–37 mm). Although the best match was observed using $s_{TS} = 30\%$ for spheres of 22–37 mm, the overestimation by $a_{TS} = 27\%$ was minimal and was considered to be acceptable.

In respiratory motion, PET acquisition in the static mode underestimated the radioactivity and overestimated the dimensions. We observed that the mean %counts were $-43.8 \pm -51.4\%$ by static acquisition and $-15.2 \pm -30.1\%$ by gated acquisition. A recent analysis of gated PET showed a significant influence of the breathing cycle on the measured SUVmax, with variations in SUV measurements of up to 24%.
dimensions of motion spheres by static acquisition were significantly higher than those of non-motion; however, those by gated acquisition were close to the standard values. The maximum FDG activity of small spheres decreased significantly, depending on the movement distance. Respiration motion greatly influences the threshold value in small tumors, so the elimination of motion effect is important for the accurate delineation of tumor volume. The data acquisition of PET/CT by either respiration gating or with breath holding is necessary. On the other hand, the internal target volume (ITV) must include the respiratory motion distance when the RT is performed without either respiratory gating or the breath-hold technique. The threshold value to delineate the blurring distance in the static acquisition must be examined.

Recently, several investigators reported that the use of a single-threshold model for delineating the GTV with PET/CT is not sufficient, because of the effects of the background radioactivity, target size, motion, and image reconstruction parameters. In response to these reports, some complex threshold models have been suggested. Nestl et al. proposed a \( I_{\text{mean}} \times 0.15 + I_{\text{background}} \) threshold model and reported that it gave good delineation for a tumor with inhomogeneity. In this model, \( I_{\text{mean}} \) was calculated as the mean intensity of all pixels surrounded by the 70% \( I_{\text{max}} \) isocontour within the tumor, and \( I_{\text{background}} \) as the mean intensity including the highest FDG uptake of the adjacent normal tissue. Black et al. observed a strong linear relationship between the threshold SUV and the mean target SUV, and proposed the following linear regressive function: threshold SUV = 0.307 × (mean target SUV) + 0.588. The background concentration and target volume indirectly affect the threshold SUV by way of their influence on the mean target SUV. Okubo et al. reported that three threshold levels for FDG activity were adopted according to tumor size for tumors of < 2 cm (2.5 SUV), 2 to 5 cm (\( T_{0.5} = 35\% \)), and > 5 cm (\( T_{0.5} = 20\% \)), and the multiple thresholds were applicable for the primary target delineation in PET/CT simulation. Clinical investigation by Biehl et al. revealed that the optimal threshold was 42% ± 2% for tumors less than 3 cm, 24% ± 9% for tumors measuring 3-5 cm, and 15% ± 6% for tumors measuring greater than 5 cm. Further examination with different radioactivity levels and different lesion sizes may require a different threshold model.

Our study had several limitations. The contrast and spatial resolution of PET images vary by PET device and reconstruction method. The threshold value would be different when such conditions are different. The background of the phantom in this study was filled with air because our study simulated a lung tumor. In cases of tumors in the trunk, we must examine the phantom filled with radioactive background. In this study, the distribution of radioactivity in hot spheres is homogeneous, and the margin of hot spheres is sharply demarcated. In clinical cases, the neoplastic tissue consists of histologically heterogeneous components and a pathologically irregular margin. Thus a clinical valuation in comparison with the histopathological examination is necessary before clinical application of this method.

In conclusion, a threshold value of \( a T_{\text{50}} = 27\% \) was proposed for auto-contouring of lung tumors. At the same time, our results suggested that the respiratory-gated data acquisition should be performed in both PET and CT for 3D target volume delineation.

Acknowledgement

We gratefully acknowledge Mr. Hiroyoshi Kawakami (GE Healthcare Japan), Mr. Naoyuki Tamamura (Nihon MediPhysics), and Mr. Daiki Magome (Kyushu University) for their technical assistance. This study was supported in part by a Grant-in-Aid for Scientific Research (C) from
the Japan Society for the Promotion of Science KAKENHI (22611012).

References


(Received for publication November 3, 2010)
ファントムを用いた PET/CT による腫瘍標的体積の自動輪郭抽出に関する検討

目的：我々はファントムを用いて PET/CT による腫瘍標的体積の自動輪郭抽出における至適閾値の決定とその有用性について検討した。方法：PET/CT 装置は Discovery STE を、ファントムは IEC Body Phantom を用いた。ファントムの 6 個の球（径 10〜37mm）には 18F 溶液を、周囲には空気を封入した。ファントムは静止状態に加えて呼吸性移動を模した移動状態（振幅：2 cm、周期：4 秒）にて撮像し、さらに低放射能濃度の静止状態にても撮像した。PET データ収集は static モード（sPET）3 分間と呼吸同期（gPET）8 分間で行い、CT はヘリカル CT（HCT）とシェードモード（4DCT）で撮像した。まず静止状態において各球の理論的体積を過小評価せずに抽出できる値を閾値とした。次にその閾値を用いて各条件にて撮像した sPET-HCT、gPET-4DCT から抽出された体積を理論値と比較した。なお閾値は、球の最大 SUV に対する割合（T%）および対象 SUV（TSUV）の 2 種類を求め、対象は 6 個全ての場合（all）と部分容積効果の影響のない大きい球のみ（径 37〜22mm）（selection）の 2 種類で検討した。結果：閾値は、全ての球では aT% は 27%，aTSUV は 2.4。大きい球のみでは sT% は 30%，sTSUV は 4.3 となった。全ての球を対象としたとき、sPET-HCT は aT% では 84〜129% を、aTSUV では 34〜127% を抽出したが、gPET-4DCT ではそれぞれ 94〜103%，51〜131% を抽出した。一方、大きい球を対象としたとき、sPET-HCT では 76〜87%，63〜90% を、gPET-4DCT では 89〜98%，84〜99% を抽出した。低放射能濃度の場合はすべての球を過小評価した。結論：PET/CT による腫瘍標的体積の自動輪郭抽出における至適閾値はファントムを用いた検討では 27% であった。また、呼吸性移動のある腫瘍の輪郭抽出を PET/CT で行う場合は、PET、CT ともに呼吸同期収集が必要と考えられた。