Defining PET Standardized Uptake Value Threshold for Tumor Delineation with Metastatic Lymph Nodes in Head and Neck Cancer†

Osamu Suzuki1,*, Kinji Nishiyama1, Masahiro Morimoto1, Hajime Hashiguchi2, Kanji Tanaka2, Aya Nakajima1, Satoaki Nakamura1 and Kunitoshi Yoshino3

1Department of Radiation Oncology, Osaka Medical Center for Cancer and Cardiovascular Disease, 2Ohmichikai Medical Corporation Morinomiya Clinic and 3Department of Head and Neck Surgery, Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka, Japan

*For reprints and all correspondence: Osamu Suzuki, Department of Radiation Oncology, Osaka Medical Center for Cancer and Cardiovascular Disease, 1-3-3 Nakamichi, Higashinari-ku, Osaka City, Osaka Prefecture 537-8511, Japan. E-mail: osamu-s@umin.ac.jp

†Presented as poster presentation at 50th American Society for Therapeutic Radiology and Oncology Annual Meeting, 21–25 November 2008, Boston, MA, USA.

Received November 29, 2011; accepted March 11, 2012

Objective: Hot spots of F-18 fluorodeoxyglucose positron emission tomograms are variable in size according to window settings of standardized uptake values. The purpose of this study was to determine the standardized uptake value threshold that represents the target volume.

Methods: Sixty-three patients who underwent fluorodeoxyglucose positron emission tomographic computed tomography and were diagnosed as having head and neck cancer with cervical lymphadenopathy were studied. The horizontal and vertical diameters of metastatic lymph nodes (LN-CT) were measured at the center of computed tomographic images. Of the corresponding nodes, the maximal standardized uptake value (SUVmax) and standardized uptake value profiles along the central horizontal and vertical axes were calculated on positron emission tomographic images (LN-PET). On the standardized uptake value profiles, the standardized uptake value levels (SUVeq) where the size of LN-PET was equivalent to the diameters of LN-CT were obtained. The regression formula between SUVeq and SUVmax was obtained. The regression formula of SUVeq was validated in subsequent 30 positron emission tomographic computed tomography studies.

Results: The mean horizontal and vertical diameters of LN-CT were 14.9 and 16.4 mm, respectively. SUVmax ranged from 1.88 to 9.07, and SUVeq was between 1.16 and 6.42. The regression formula between SUVeq and SUVmax was as follows: SUVeq = 1.21 + 0.34 x SUVmax (coefficient of correlation: R = 0.69). The validation study resulted in a good correlation between the volume of lymph nodes on computed tomography and positron emission tomographic computed tomography (R² = 0.93).

Conclusions: The formula with a relatively high coefficient of correlation is considered to indicate that SUVeq is not constant, but is a complex of an absolute standardized uptake value and is proportional to SUVmax.

Key words: PET–CT – delineation – SUV – radiation planning
INTRODUCTION

Recently, three-dimensional (3D) conformal radiotherapy and intensity-modulated radiation therapy have achieved dose distributions of improved conformity to the target volume. Computed tomography (CT) scans and magnetic resonance imaging (MRI) serving images of tumors and the surrounding related structures have been widely used in radiation therapy planning (RTP) (1).

In addition, the functional imaging of F-18 fluorodeoxyglucose positron emission tomography (FDG-PET) has been introduced in RTP. FDG-PET has a higher sensitivity, specificity, and accuracy for the detection of tumor involvement than CT or MRI for many cancer sites and results in a more accurate staging of malignancies, both for the primary tumor and for lymph nodes (LNs) and distant metastasis. In RTP, it is essential to delineate the tumor boundary. Because PET serves functional images, it has an advantage of clarifying the presence of malignant deposits that cannot be depicted with CT or MRI. However, it has a disadvantage that the border of lesions varies on window settings and there is no consensus on the optimal window setting to depict the hot areas equivalent to tumors has been achieved yet. Several kinds of approach to contour gross tumor volume (GTV) on PET images have been reported (2–17). Visual interpretation is most commonly used (2–7). However, this method is highly operator-dependent. Objective methods to depict FDG hot spots equivalent to GTV were cut-off standardized uptake value (SUV threshold). The SUV thresholds reported were a constant proportion to the maximal SUV (SUVmax) such as 40 or 50%, or an absolute SUV of 2.5, 3.0 and so on (6,8–12,15). Contrast-dependent adaptive thresholding methods have also been proposed (6,13,14). And gradient-based delineation was reported as a robust alternative to the threshold-based delineation (15).

There is still a disagreement on the optimal SUV threshold (16,17). The purpose of this study was to determine the SUV threshold to the contour target volume by finding a regression formula using an interpolation method with PET–CT data of LNs of head and neck cancer.

PATIENTS AND METHODS

PATIENTS

Sixty-three patients with head and neck cancer accompanied by cervical LN swelling (52 males and 11 females, 29–87 years) who underwent PET–CT scan from January 2007 until December 2007 were retrospectively enrolled in this study. LNs that exceeded 1 cm in the shortest diameter on CT images or whose SUVmax was 2.0 or more on PET images were diagnosed as metastatic. No pathologic proof was obtained in this study. The metastatic LN was regarded as GTV on CT images in this study. If there were multiple metastatic LNs, the one with the highest SUVmax was the study object. The cervical LNs whose SUVmax exceeded 10.0 were excluded. The primary sites of the subjects were the oropharynx in 15 cases, the hypopharynx in 12, the nasopharynx in 2, the tongue in 10, the larynx in 10, the salivary gland in 5, the oral cavity in 3 and other locations in 6. All images of PET and CT were obtained on integral PET–CT systems in the same patient position.

PET–CT ACQUISITION

In accordance with the PET–CT protocol, patients fasted for 6 h before examination. Complete anamnesis was obtained in all patients, and the serum glucose level was measured before $^{18}$F-FDG administration in order to verify that it was acceptable (cut-off 180 mg/dl). Patients were then injected with a solution of $^{18}$F-FDG at a dose of 3.7 MBq/kg and relaxed for an hour before imaging. PET scans were performed ~60 min after the intravenous administration of FDG at the above-mentioned dose.

First, a scout image was acquired to select the area of interest; a spiral CT examination was then performed, followed by the acquisition of 3D-PET emission data, with each bed position being 16.2 cm long and with an imaging time of 3 min per bed position.

The PET–CT systems used in this study were GEMINI GXL 16 (Philips, Hamburg, Germany) and Biograph LSO Duo (Siemens, Forchheim, Germany). On the Biograph LSO Duo, CT scan acquisitions were performed on a spiral dual-slice CT, with a slice thickness of 3.5 mm and a pitch of 1.15 mm. Images were acquired using a matrix of $512 \times 512$ pixels and a pixel size of 0.98 mm, and were reconstructed for the treatment planning system workstation with a thickness of 3.4 mm. PET data were acquired using a matrix of $128 \times 128$ pixels, with a slice thickness of 5.3 mm. The final imaging resolution for clinical practice was ~6.5 mm. CT-based attenuation correction of the emission images was employed. The PET images were reconstructed by the iterative method ordered subset expectation maximization (three iterations, eight subsets) with a filter of 7 mm. On the GEMINI GXL 16 integrated PET–CT scanner, CT scan acquisitions were performed on a spiral 16-slice CT, with a slice thickness of 4 mm and a pitch of 1. Images were acquired using a matrix of $512 \times 512$ pixels and a pixel size of 1.17 mm and were reconstructed for the treatment planning system workstation with a thickness of 3.4 mm. PET data were acquired using a matrix of $128 \times 128$ pixels, with a slice thickness of 4 mm. The PET images were reconstructed by the 3D-LOR Row-Action Maximization-likelihood (two iterations, relaxation parameter 0.025). The SUVs measured with the two PET–CT systems are regularly validated with cross-calibration using the same dose calibrator (CRC-15PET, Capintec, NJ, USA).

FINDING THE SUV THRESHOLD FROM THE REGRESSION FORMULA

The horizontal (right to left) and vertical (anterior to posterior) diameters of the LNs on axial CT images (LN-CT) were measured at the central level (Fig. 1). The SUVmax and SUV profiles along the horizontal and vertical axes of the corresponding LN on PET images (LN-PET) were obtained.
The SUV threshold (SUVeq) where the diameter of LN-PET was equal to that of LN-CT was calculated by the interpolation method (Fig. 2). One hundred and twenty-six (63/2) data of SUVmax and SUVeq were obtained and used in the study. The correlation between the SUVeq and SUVmax was standardized by obtaining the regression formula between SUVeq and SUVmax.

**VALIDATION OF THE SUVEQ FORMULA**

For validation of the regression formula obtained in this study, subsequent 30 PET–CT studies of head and neck cancer patients were evaluated. The volume and diameter on the right–left axis of cervical LN suspected of metastasis were measured on CT and PET, respectively, on radiation treatment planning system. On PET images, the volume and diameter depicted using SUV threshold calculated from the SUVeq formula were measured. The difference in diameter and volume between CT and PET was evaluated, respectively.

**RESULTS**

The LNs examined were all well demarcated and their diameters could be measured on CT images. The mean diameter of LN-CTs were 14.9 mm (range: 5.6–39.1 mm) and 16.4 mm (range: 7.0–46.9 mm), in horizontal and vertical directions, respectively. The mean serum glucose level measured before 18F-FDG administration was 96 ± 20 mg/dl, and only one data exceeded 150 mg/dl (164 mg/dl). The mean SUVmax was 5.16 (range: 1.88–9.07), and the SUVeq measured with the interpolation method was 2.93 on average (range: 1.16–6.42; Table 1).

The linear regression analysis demonstrated a significant correlation between SUVeq and SUVmax ($P < 0.001$). Figure 3 shows the linear regression analysis between SUVmax and SUVeq. The formula of the regression line was as follows: $SUV_{eq} = 1.21 + 0.34 \times SUV_{max}$ (coefficient of correlation: $R = 0.69$) (Fig. 3).

Figure 3 demonstrates that the scattering is narrow below 5.0 of SUVmax and dispersed over 5.0. Figure 4 indicates that the proportion of SUVeq to SUVmax became lower as SUVmax being larger. And that ranged 9% to 89% to SUVmax. The results indicated that the SUVeq was not a constant value, but was positively correlated with SUVmax.
From the validation phase, differences in diameters of cervical LNs between CT and PET were almost within 2 mm (Fig. 5). The volumes depicted on PET images using the threshold of SUVeq were indicated a good correlation with the volumes measured on CT (Fig. 6).

**DISCUSSION**

In this study, we used the LNs of head and neck cancer to determine the SUV threshold to the contour target volume on the PET images. Some investigators studied the relationship between the GTV on CT images and the SUV threshold to delineate the GTV on the PET image of the primary lesion of head and neck cancers (17–19). Van Baardwijk et al. (20) intended non-small cell lung cancers for determining the optimal SUV threshold to delineate the pathological GTVs. In the primary lesion of head and neck cancers, there could be superficial extensions of tumor that could not be detected by diagnostic imaging. In the lung cancer, there is respiratory tumor motion that leads ambiguous delineation of the tumor on PET images (21). Cervical lymphadenopathy is well contoured on CT images and considered to be less influenced with respiratory motion compared with the lung tumor. Thus, we selected cervical LNs as the object. However, they were not histologically but

| Table 1. Computed tomographic (CT) and positron emission tomographic (PET) data of studied lymph nodes |
|--------------------------------------------------|------------|------------|
| CT (mm)                                           | Min       | Max        | Ave        |
| Horizontal                                       | 5.6       | 39.1       | 14.9       |
| Vertical                                         | 7.0       | 46.9       | 16.4       |
| PET                                               |           |            |            |
| SUVmax                                           | 1.88      | 9.07       | 5.16       |
| SUVeq                                            | 1.16      | 6.42       | 2.93       |

Figure 3. The relationship between SUVeq and SUVmax. A linear regression analysis with the datasets of SUVeq and SUVmax. The formula of the regression line was SUVeq = 1.21 + 0.34 × SUVmax (coefficient of correlation: R = 0.69).

Figure 4. The proportion of SUVeq to SUVmax. The proportion of SUVeq to SUVmax ranged from 9 to 89%. That became lower as SUVmax became larger.

Figure 5. The differences in diameters of cervical LNs between CT and PET were indicated. The differences were almost within 2 mm.

Figure 6. The relationship between the LN volume on CT and the PET volume. Plots represent the correlation between the LN volume on CT and the SUVeq-derived volume. The black dots represent individual patients, and the gray lines represent the region where the SUVeq-derived volume is within ±50% of the LN volume on CT.
clinically diagnosed metastatic LNs. That might influence the results in some degree. Murakami et al. (22) investigated SUVmax and the size of cervical nodes on preoperative PET–CT to find the characteristics of pathologically positive LNs in head and neck cancer patients. They showed that a receiver-operating characteristic analysis of the SUVmax of nodal sizes suggested that the size-based cut-off value was 1.9 for nodes <10 mm in diameter, 2.5 for those 10–15 mm and 3.0 for >15 mm. These cut-off values yielded 79% sensitivity and 99% specificity for cervical LN metastasis. The data of LN size and SUVmax in the present study met the cut-off value described earlier. So, the author expects that the subjects of the present study are almost metastatic LNs in spite of lacking histological proof.

There were some approaches to contour GTV on PET images based on visual interpretation and SUV thresholding. Visual interpretation was most commonly used. However, this method was highly operator-dependent and led to poor reproducibility of the GTV contours (9–11). On the other hand, the GTV contours delineated by using an SUV threshold were quantitative and reproducible. Several kinds of objective methods to depict FDG hot spots equivalent to GTV

An absolute SUV of 2.5 was often reported as an SUV threshold to delineate GTV. Schinagl et al. compared GTVs on CT and PET images in head and neck cancer. The SUV-based volumes obtained by applying an iso-contour of SUV of 2.5 around the tumor were largely oversized in 35 of the 77 cases. They concluded that a threshold SUV of 2.5 was not useful for an automated target definition of head and neck cancer (27). Nestle et al. (6) also found that the threshold SUV of 2.5 was not suitable when attempting to delineate lesions surrounded by tissue with significant background activity. Burri et al. (19) compared GTVs derived by several kinds of PET delineations with the pathologic tumor volume in 18 patients with head and neck cancer. The PET-based tumor volume defined with a threshold SUV of 2.5 was likely to grossly overestimate the tumor volume by 150% than the pathologic tumor volume. These results indicate that the threshold SUV of 2.5 is not applicable for contouring tumor volume of the head and neck cancer.

Burri et al. also studied the PET-based tumor volume delineated with the threshold of a constant proportion of 40% to SUVmax in head and neck cancer. From their comparison with pathologic tumor volume, volumes with a threshold of 40% of SUVmax were also likely to grossly overestimate tumor volumes as with a threshold SUV of 2.5, but it was considered to offer the best compromise between the accuracy and reducing the risk of understimating the tumor extent among thresholding methods such as an SUV of 2.5, and a source-to-background method (19). Nestle et al. (6) also demonstrated that the 40% SUVmax concept was not generally suitable for target volume delineation of non-small cell lung cancer.

Thus, there has not been a definite agreement on using an absolute SUV or constant proportions to SUVmax of tumor as SUV threshold.

Contrast-dependent adaptive thresholding methods were also proposed (6,13,14,23–26,28). In the thresholding method with a signal-to-background ratio (SBR), scanner-specific variables necessary for the calculation of SBR were derived by phantom experiments (13). Daisne et al. (28) compared the PET volume of pharyngolaryngeal cancer to the pathological examination with an automatic segmentation algorithm based on the SBR method. Although GTVs delineated on PET images were more accurate than those on CT or MR images, the former were still larger than the latter. They stated that there were overestimation of extralaryngeal extension and underestimation of superficial spread of tumor. Schinagl et al. (27,29) compared five segmentation approaches of visual delineation, iso-contour of SUV 2.5, 40 or 50% to the SUVmax of lesions and SBR method for contouring FDG-PET-based target volumes of primary tumors and metastatic LNs in the head and neck cancer. Their studies revealed that PET frequently detected the extension of tumor outside CT-based GTVs, regardless of the applied segmentation methods. They concluded that the proper approach of FDG-PET-based target contouring was not identified.

Van Baardwijk et al. (20) reported a good correlation (Pearson’s correlation coefficient, 0.90) between diameters of GTV on PET images depicted with the SBR-based auto-contour and macroscopic tumor diameters on surgical specimens in 23 operable lung cancers. However, the GTV diameters derived from the auto-contour underestimated the diameter of surgical specimens in general (20). According to their phantom measurements, the SUV threshold to SUVmax varied from 31 to 47%. In cases where SBR was larger than 5, they reported that a 34% threshold to SUVmax could be used. Geets et al. (15) utilized a gradient-based method in the phantom study and validated with laryngeal tumors and stated that the contour based on SBR was not sufficient. Both the gradient-based method and the SBR method did not encompass macroscopic laryngeal specimens.

In this study, we obtained the threshold SUVeq that indicated the same size as the LN-CT using the interpolation method. In this interpolation method, the gradient of the SUV profile tends to be steeper if the SUVmax is high. That leads to a large increase and decrease in SUVeq, which was thought to be harboring some errors in the calculated SUVeq. So, we excluded LNs with a high SUVmax that exceeded 10. A regression formula was derived from the data of SUVmax and SUVeq (Fig. 3).

That was as follows: $SUV_{eq} = 1.21 + 0.34 \times SUV_{max}$.

This formula indicated that the SUVeq was not an absolute SUV or a constant proportion to the SUVmax. SUVeq ranged from 1.16 to 5.48. That indicated that SUVeq was not a constant figure. Figure 4 indicates that the proportion of SUVeq to SUVmax became lower as the SUVmax being larger. And that ranged from 9 to 89% to SUVmax. The range, being larger than the results of Van Baardwijk et al., indicated that SUVeq was not a constant proportion to SUVmax. The proportion of our formula (0.34) was the same as their 34% to SUVmax in the cases of SBR ≥ 5. It
was considered that in cases where SUVmax was higher, our formula derivates SUVeq being close to 34% of SUVmax with the proportion of the constant value of 1.21 to SUVmax being smaller.

From our formula, SUVeq was the sum of proportional of SUVmax and a constant figure. Black et al. (30) conducted a phantom study and also obtained a regression formula that indicated the threshold SUV being a function of mean SUV. The threshold SUV = 0.307 × SUVmean + 0.588 (SUVmean: mean target SUV). In this way, the formula by Black et al. also indicated that the threshold SUV was not an absolute SUV or a constant proportion to the SUVmean but a combination of the two parameters as our formula. Because it is difficult to obtain a mean SUV of tumor in vivo, we used the data of SUVmax and obtained a similar formula to that of Black et al. In a recent study reported by Murphy et al. (17), they intended to compare the pathological volume of oral cavity cancer to metabolic tumor volume (MTV) delineated with SUV threshold on PET images. The correlation between MTV and pathologic tumor volume was relatively poor using the SUV threshold of such as an absolute SUV, the proportion to SUVmax or SBR. They reported the regression formula that indicated that SUVmax and tumor grade were involved in calculating SUV threshold. In our validation study, the difference between CT and PET was almost within 2 mm in diameter and was almost within ±50% in volume. These results were better than the results of Murphy et al. and indicated that our formula was a validated one to depict the gross tumor burden in PET–CT.

A shortcoming of PET usage in RTP is the low resolution of PET image. In our validation study, geometrical miss between LN on CT and the volume depicted on PET image. In our validation study, the difference between CT and PET was almost within 2 mm in diameter and was almost within ±50% in volume. These results were better than the results of Murphy et al. and indicated that our formula was a validated one to depict the gross tumor burden in PET–CT.

In conclusion, the regression formula obtained in this study demonstrated that the SUVeq to contour GTV suitably is not constant, but is complex of an absolute SUV and is proportional to SUVmax. The formula validated a high coefficient of correlation was considered to be suitable for use in RTP.

Conflict of interest statement
None declared.

References


