

# A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: derivation from phantom measurements and validation in patient data

Andrea Schaefer · Stephanie Kremp · Dirk Hellwig ·  
Christian Rube · Carl-Martin Kirsch · Ursula Nestle

Received: 30 March 2008 / Accepted: 12 June 2008 / Published online: 26 July 2008  
© Springer-Verlag 2008

## Abstract

**Purpose** An easily applicable algorithm for the FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer was developed by phantom measurements and validated in patient data.

**Methods** PET scans were performed (ECAT-ART tomograph) on two cylindrical phantoms (phan1, phan2) containing glass spheres of different volumes (7.4–258 ml) which were filled with identical FDG concentrations. Gradually increasing the activity of the fillable background, signal-to-background ratios from 33:1 to 2.5:1 were realised. The mean standardised uptake value (SUV) of the region-of-interest (ROI) surrounded by a 70% isocontour ( $mSUV_{70}$ ) was used to represent the FDG accumulation of each sphere (or tumour). Image contrast was defined as:  $C = (mSUV_{70} - BG) / BG$  where BG is the mean background-SUV. For the spheres of phan1, the threshold SUVs (TS) best matching the known sphere volumes were determined. A regression function representing the relationship between  $TS / (mSUV_{70} - BG)$  and  $C$  was calculated and used for delineation of the spheres in phan2 and the gross tumour volumes (GTVs) of eight primary lung tumours. These GTVs were compared to those defined using CT.

**Results** The relationship between  $TS / (mSUV_{70} - BG)$  and  $C$  is best described by an inverse regression function which can be converted to the linear relationship  $TS = a \times mSUV_{70} + b \times BG$ . Using this algorithm, the volumes delineated in phan2 differed by only  $-0.4$  to  $+0.7$  mm in radius from the true ones, whilst the PET-GTVs differed by only  $-0.7$  to  $+1.2$  mm compared with the values determined by CT.

**Conclusion** By the contrast-oriented algorithm presented in this study, a PET-based delineation of GTVs for primary tumours of lung cancer patients is feasible.

**Keywords** Volume delineation ·  $^{18}F$ -FDG · PET · Lung cancer · GTV definition · Radiotherapy · Physics · PET oncology therapy PET/CT

## Introduction

In recent years, FDG-PET has become established as the most accurate diagnostic procedure for staging non-small-cell lung cancer (NSCLC) [1, 2]. As FDG-PET images the functional characteristics of tissue, this method has the potential to accurately demonstrate the three-dimensional extent of malignant spread [3]. FDG-PET is, therefore, attractive to be used as a tool for gross tumour volume (GTV) definition in radiotherapy planning. Historically, the GTV of NSCLC tumors has been defined by the International Commission on Radiation Units and Measurements (ICRU) based only on morphological imaging [4]. Therefore it is compromised by the diagnostic limits of these methods [5, 6].

As the functional target volume found with FDG-PET can often be assumed to be clinically more relevant than that historically achieved with CT, many institutions have

A. Schaefer (✉) · D. Hellwig · C.-M. Kirsch · U. Nestle  
Department of Nuclear Medicine,  
Saarland University Medical Center,  
66421 Homburg, Germany  
e-mail: andrea.schaefer@uks.eu

S. Kremp · C. Rube  
Department of Radiotherapy, Saarland University Medical Center,  
66421 Homburg, Germany

begun utilising FDG-PET for treatment planning. However, for technical reasons, PET data suffer from a comparatively low geometric resolution which poses technical problems when integrating the method into the radiotherapy planning process. Determination of a volume from FDG-PET is, therefore, a critical step. Various basic approaches were reported in the literature to accurately contour FDG-based GTVs [7–9]. Most commonly, visual differentiation by the experienced nuclear medicine physician is applied. Whilst profiting from clinical experience, this method also bears the risk of inter-observer variability [7]. Other authors proposed the use of absolute thresholds like a standardised uptake value (SUV) of 2.5 [8, 10]. However, due to biological and physical factors [11–16], an absolute threshold does not seem to be generally applicable for target volume definition [17]. Hence, automated or semi-automated methods of image segmentation have been attempted, e.g. the application of a fixed threshold value of maximal uptake [9]. However, depending on the tracer accumulation in the background of the lesion, this method may also fail.

In another basic approach developed by phantom measurements, some authors proposed to use signal-to-background ratios in algorithms for the segmentation of FDG-positive tissue [18, 19]. One investigation has shown that tumour volumes determined by such an algorithm in head-and-neck tumours (volumes <30 ml) do better correlate with the tumour volume determined pathologically than volumes derived from CT and MRI [18, 20–22]. As lung tumours do largely differ from head-and-neck tumours in terms of size, homogeneity and physiological background, an algorithm proposed for that region of the body does not necessarily apply to lung cancer targets. Another group developed an algorithm which uses the mean target SUV as the only parameter to iteratively calculate the threshold SUV for auto-contouring a volume [19]. Applying the algorithm to spheres of known volume and a series of NSCLC patients in comparison with a fixed image intensity threshold, smaller deviations in target/phantom delineation were observed when using the algorithm. Although this group reported the validation in phantoms, discrepant GTVs were determined by another group using this algorithm to evaluate patient data [23]. Therefore, until now, no method has been established for the universal use in FDG-based GTV contouring in lung cancer.

During recent years, our institution has focussed on FDG-PET in radiotherapy planning for lung cancer patients [6, 24–28]. In preparation of a clinical multi-centre trial on FDG-based radiotherapy planning for NSCLC involving centres with various PET, CT and planning systems, there was the need to address the question of how to define comparable target volumes using FDG-PET between the centres. A method for target volume delineation of lung

tumours was needed which is non-sensitive against varying physiological background conditions or tumour inhomogeneity and which can furthermore be simply used in various PET and/or planning systems. Therefore, the aim of the present investigation was to develop an algorithm for the GTV delineation of lung tumours, easily applicable, to be used in different systems. The method was developed by phantom measurements and validated by application to patient data.

## Materials and methods

### Phantoms

Two cylindrical phantoms formed from polymethylmethacrylate (PMMA) enclosing different glass spheres were used: one (phan1) with a liquid background volume of 21.48 l and up to eight fillable glass spheres with volumes ranging from 7.4 up to 258.0 ml (7.4, 13.7, 33.8, 64.0, 107.1, 171.3 and 258.0 ml), the other (phan2) with a liquid background volume of 7.24 l and four fillable glass spheres (volumes=7.4, 13.7, 64.0 and 171.3 ml). The phantoms are shown in Figs. 1 and 2, respectively. The sizes of all glass spheres used were within the range of tumour volumes observed in NSCLC patients. The geometry of phan1 was chosen to reflect the distributions of scatter and random events similar to those of patient studies. Phan2 was



**Fig. 1** Phan1 as used for evaluation of the contrast-oriented algorithm comprising ten fillable spheres of different diameters



**Fig. 2** Phan2 as used for validation of the contrast-oriented algorithm comprising four fillable spheres of different diameters

planned to be used for the purpose of quality control during a multi-centre trial. Therefore, its smaller geometry was chosen with regard to practicability.

#### Phantom measurements

For each series of measurements, 1,000 ml of a homogeneous solution of FDG was prepared and all spheres filled with this solution. Measurements were repeated starting with an initial activity concentration of FDG within the spheres of  $47.0 \pm 7.6$  kBq/ml at the time of the first acquisition. The background was filled with the lowest activity concentration of FDG used and then gradually increased, whereby the sphere-to-background ratio decreased from 33:1 to 2.5:1 for phan1. Hypothetical “uptake values” corresponding to the biological SUV based on the net phantom weight as “lean body weight” were also calculated for the spheres. These SUVs ranged from 25.3 down to 2.5 and were within the range observed in patients with lung tumours. Using phan2, sphere-to-background ratios of about 10:1, 5:1 and 3:1 were realised. For all measurements, the hypothetically isodistributed activity concentration within both phantoms varied between  $3.90 \pm 0.98$  kBq/ml, similar to that used in clinical practice.

To further investigate whether there is an influence of the overall activity concentration on the results, two additional series of measurements were performed using phan1: one realising a comparably low activity concentration (2 kBq/ml) and one a comparably high activity concentration (7.8 kBq/ml).

#### Patients

For the validation of the method developed by phantom measurements in patients, data sets of eight patients with histologically proven primary NSCLC were used. These were chosen out of a group of patients with both PET and CT scans acquired for radiotherapy planning [17]. The patients selected for the purpose of the actual comparison showed lesions which were clearly delineable in CT and, therefore, allowed CT-based GTV determination at a high level of certainty.

#### Acquisition of data

PET data were acquired on an ECAT-ART PET scanner (CTI/Siemens; slice thickness=3.375 mm, transaxial spatial resolution=6.2 mm) [29] comprising an emission scan followed by a “hot” transmission scan in singles mode using two  $^{137}\text{Cs}$  point sources [30]. A total field of view (FOV) of 28.2 cm was scanned in two overlapping bed positions (BP) (axial FOV of the scanner=16.2 cm). Acquisition parameters of the emission and transmission scan, respectively, were similar to those used in clinical routine (emission scan time=10 min/BP; “hot” transmission live time=192 s/BP, reconstructed image size=128×128 pixels of  $5.15 \times 5.15$  mm). The transmission images were reconstructed using filtered back projection (ramp filter, 0.25 Nyquist). Attenuation-corrected image reconstruction of the emission data was performed using the iterative OSEM method (two iterations, four subsets, 2-mm Gaussian filter) [31].

All patients underwent an additional PET acquisition of the thorax (two bed positions) in radiotherapy treatment position 160 min after the injection of 250 MBq FDG following the routine whole-body scan. The corresponding planning CT scan (Philips CT TWIN FLASH, 5 mm slice thickness,  $512 \times 512$  pixels of 0.98 mm, shallow breathing) was acquired on the same day. To ensure identical patient position during both examinations, the same positioning aid was used on both modalities and the patients were aligned and marked using external laser positioning systems as described before in more detail [17].

PET and CT data were transferred to a Hermes Workstation (Hermes Medical Solutions, Sweden) to perform automatic co-registration using a rigid normalised mutual-information algorithm. The co-registered patient data as well as the PET data of the phantom measurements were converted to DICOM data and then transferred to the radiotherapy planning system (Philips Pinnacle, version 6.2) of the department of radiotherapy [6, 32].

#### Image analysis: phan1

Evaluation of the images was performed on the PET workstation using the ECAT software (Siemens CTI, ECAT

version 7.2). According to our hypothesis, the optimum method to determine the TS for auto-contouring a volume in FDG-PET is influenced by two parameters: the background activity and the FDG accumulation of the lesion.

To account for the surrounding background FDG concentration, regions-of-interest (ROI) were drawn in every plane between the spheres at a distance of at least 5 mm from the target. The mean SUV of all these ROIs was used as background value (BG).

To account for the activity of the lesion itself, one has to consider the intensity as well as the inhomogeneity of the FDG accumulation. Therefore, firstly, the maximum SUVs of each sphere were determined by drawing ROIs around the spheres. Secondly, to obtain a lesion-specific parameter, which later also accounts for the biologic inhomogeneity of tumours, the following procedure was performed: ROIs were generated in each sphere by a threshold of 70% of its maximum SUV. For these ROIs, the mean SUV ( $mSUV_{70}$ ) was calculated. (ROIs are also presented in Fig. 4a, see the “Results” section).

The choice of the threshold of 70% was determined experimentally as a result of multiple tests in patient datasets. The rationale was to use voxels surrounded by a threshold which is sufficiently smaller than  $SUV_{max}$  in order to minimise the influence of non-representative global maximum values which are subject to large statistical variation. On the other side, the voxels used for the determination of this value representative for a lesion had to be located safely within the tumour, even in the case of a faint FDG accumulation in a rather high background to be dealt with.

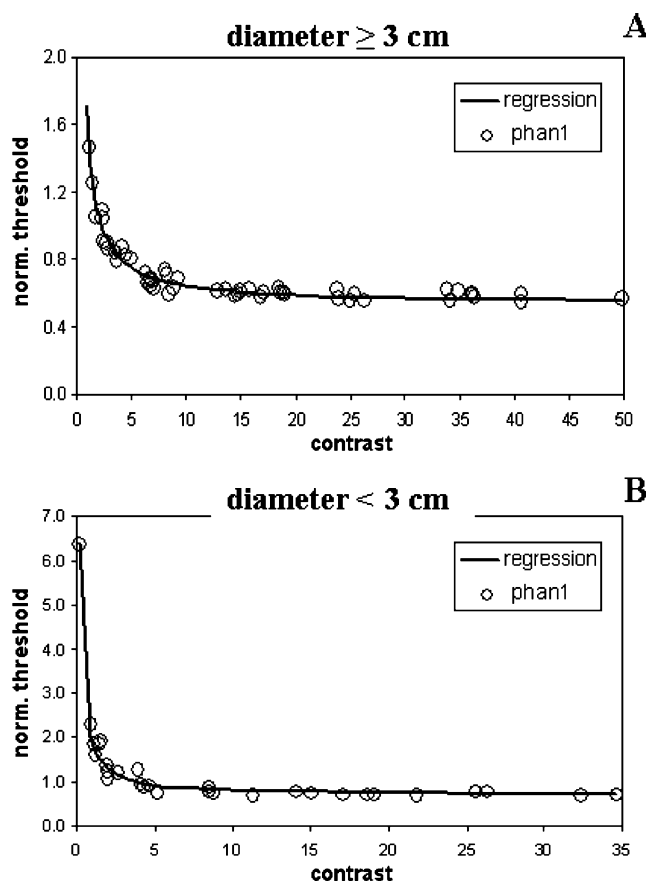
According to our hypothesis, the threshold for auto-contouring should depend on both  $mSUV_{70}$ , the mean SUV of the 70% isocontour, and background, BG. To test our hypothesis, the contrast ( $C$ ) was calculated for each sphere as the difference between the  $mSUV_{70}$  and the background relative to the background, that is:

$$C = (mSUV_{70} - BG)/BG. \quad (1)$$

The contrast was varied by gradually increasing the activity concentration of the background whereas the activity concentration of the spheres simultaneously decreases due to the physical half life of  $^{18}F$ . To find the threshold value that yielded a volume best matching the true volume, the spheres were auto-contoured in the attenuation-corrected slices varying the TS in steps of 0.1. Using the resulting optimum threshold values, a regression function was calculated representing the relationship between the normalised TS [ $TS/(mSUV_{70}-BG)$ ] and the image contrast of the spheres as defined above (Fig. 3; see the “Results” section).

Validation: phan2 and patients

The regression function derived for phan1 (Eqs. 2 and 3; see the “Results” section) was applied to define the



**Fig. 3** Normalised TS [ $TS/(mSUV_{70}-BG)$ ] as a function of the contrast  $C$ : results of measurements of phan1 for the spheres of diameter over 3 cm (a), for the spheres of diameter up to 3 cm (b). The inverse regression function curves are also shown

threshold that should fit the volumes of the spheres of phan2 as well as to define the PET-GTV of the NSCLC patients. In the patient studies, the background values were defined by the following procedure: the structure with the highest FDG uptake adjacent to a significant proportion of the lesion (e.g. lung, liver, mediastinum) was regarded as relevant background [17]. In every plane where the lesion was visually detectable, circular ROIs were placed into this relevant background at a distance of at least 5 mm from the target. The mean SUV of all these ROIs was taken as background value, BG, of the patient study. The  $mSUV_{70}$  was determined by the procedures described above (phan1).

The calculated volumes of the spheres in phan2 were compared with their true ones, whereas the PET-GTVs of the patients were compared with the GTVs defined by the radiooncologist in CT. These were generated according to current international recommendations using the soft-tissue window with respect to the lung window. For a better comparison, hypothetical spheric radii were calculated for the derived tumour volumes. This yielded a value comparable to clinical practice where the differences in radius would reflect the distance between the contours drawn in

the same image. It further eases the comparison of the order of magnitude of the differences in volumes measured to the resolution of the imaging system.

### Results

#### Phan1

The volumes determined by the best-fitting TS after stepwise variation matched the known spherical volumes within  $\pm 2\%$ . Figure 3a shows the normalised TS  $[TS/(mSUV_{70}-BG)]$  as a function of the contrast  $C$  for spheres with diameters equal or larger than 3 cm; in Fig. 3b, the values for spheres of diameters of up to 3 cm are presented. The evaluation showed that the relationship between  $(TS/(mSUV_{70}-BG))$  and  $C$  was best described by an inverse regression function providing the best regression parameters, that is:

$$TS/(mSUV_{70} - BG) = k/C + a. \tag{2}$$

Evaluating TS normalised by  $(mSUV_{70}-BG)$  instead of TS, this inverse equation could easily be converted by multiplication to a linear relationship between TS,  $mSUV_{70}$  and BG, that is:

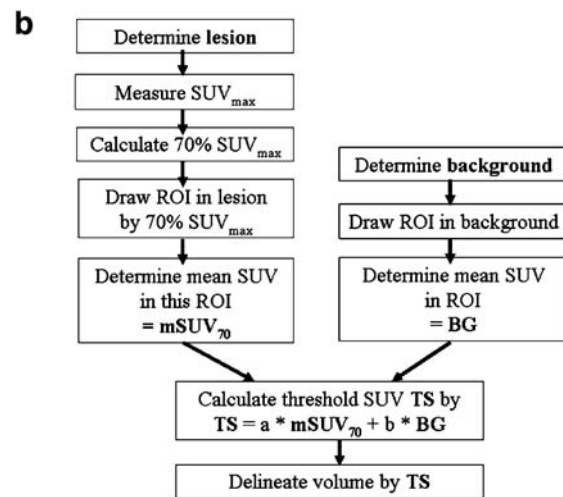
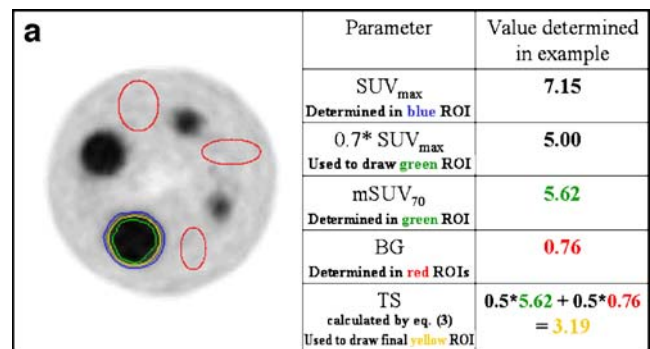
$$TS = a \times mSUV_{70} + b \times BG \tag{3}$$

with  $a=0.50$  and  $b=k-a=0.50$  ( $R=0.973$ ) for the ECAT system for spheres of diameter equal or larger than 3 cm (large spheres) and with  $a=0.67$  and  $b=0.60$  ( $R=0.993$ ) for spheres of diameter up to 3 cm (small spheres). The inverse regression function curves providing these parameters are also presented in Fig. 3. The distinction of using two sets of parameter was made based on the following procedure: for each sphere, the volume was calculated applying both sets of parameter. As the ratio of the calculated and the true volume should be preferably near one, it was necessary to apply one set of parameters for spheres of diameter equal or larger than 3 cm and the other for the smaller ones.

Figure 4a exemplifies the application of the method: a representative transaxial slice of phan1 is shown including some typical background ROIs (red) as well as the ROI drawn to find the  $SUV_{max}$  of the sphere (blue); the ROI generated by a 70% isocontour, which yields  $mSUV_{70}$  (green) and the resulting ROI, which was generated by the threshold for auto-contouring, TS, according to Eq. 3 (yellow). Results of the evaluation ( $SUV_{max}$ ,  $mSUV_{mean}$ , BG and TS) of this slice are also given in this figure. Additionally, the whole procedure is summarised in a flowchart and presented in Fig. 4b.

#### Influence of activity concentration

Data of two additional series of measurements were evaluated which realised comparatively low and high

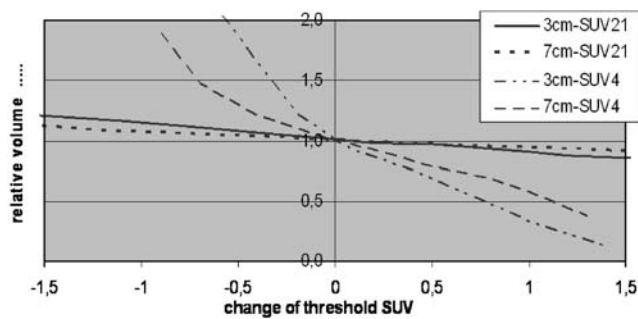


**Fig. 4** Example for the use of the contrast-oriented algorithm (a) and flowchart (b). **a** Representative transaxial slice of phan1 including typical ROIs and values determined. Red background ROIs; blue ROI generated by a threshold of  $SUV=2$  to measure  $SUV_{max}$ ; green ROI generated by a 70% ( $0.7 \times SUV_{max}$ ) isocontour, used to determine  $mSUV_{70}$ ; yellow final ROI for volume delineation, threshold TS calculated according to Eq. 3. **b** Flowchart

hypothetically isodistributed activity concentrations (2.0 versus 7.8 kBq/ml). SUVs of the spheres were 7 and 16, respectively. The thresholds for auto-contouring the spheres were calculated according to Eq. 3. We found the ratio of the calculated volumes of the spheres,  $V_{calc}$ , and the true ones,  $V_{true}$ , to be  $V_{calc}/V_{true} = 1.03 \pm 0.05$  for values of relatively high hypothetically isodistributed activity and  $V_{calc}/V_{true} = 1.05 \pm 0.02$  for the low ones, respectively. These results showed that, within the effective range of activity concentrations used, the relationship between TS,  $mSUV_{70}$  and BG as given in Eq. 3 as well as the values of the parameters  $a$  and  $b$  as stated above were found to remain valid.

#### Sensitivity of auto-contoured sphere volume to small changes in threshold SUV

The effect of small variations in TS on auto-contoured volumes is shown in Fig. 5 for the spheres of volumes of



**Fig. 5** Changes in relative auto-contoured sphere volume ( $V_{\text{calc}}/V_{\text{true}}$ ) for small changes in threshold SUV for spheres of diameters of 3 and 7 cm and theoretical SUVs of 4 and 21

171.3 ml (7 cm diameter) and 13.7 ml (3 cm diameter), respectively. We found that the relative change of volume ( $V_{\text{calc}}/V_{\text{true}}$ ) was increasing both with decreasing spherical volume and with decreasing SUV within the sphere. For low SUVs, the volume changed very rapidly with small variations of TS, the effect being more pronounced for spheres of small volumes. For example, considering a mean SUV of about 4 within the spheres and spherical volumes of 171.3 and 13.7 ml, respectively, the differences in comparison with the true volumes were found to be  $24 \pm 4\%$  and  $40 \pm 9\%$  respectively, for a change of TS of 0.5.

## Phan2

To validate the algorithm stated above (phan1, Eq. 3; Fig. 3) the volumes of the spheres of phan2 (Fig. 2) were calculated by applying TS determined according to this relationship and the parameters for large and small spheres, respectively. We found that the volumes of the two large spheres (171.3 and 64.0 ml) were slightly overestimated by less than 7% whereas the volumes of the two small spheres

(13.7 and 7.35 ml) were both underestimated by less than 8% and overestimated by less than 11%. These values corresponded to deviations in spherical radii of less than 1 mm for the large spheres. For the small spheres, the deviation in spherical radii varied between  $-0.4$  and  $+0.4$  mm. These results are summarised in Table 1.

## Application to patient data

The algorithm stated in Eq. 3 was applied to the image data of the eight patients with NSCLC. The calculated volumes are listed in Table 2 in comparison with the GTVs determined by the radiotherapist using CT. All patients but one (patient no. 2) showed high values of FDG uptake as is usually known for patients with inoperable NSCLC [33]. For a better comparison, the radii of equivalent spherical volumes are given and also the differences in spherical radii between PET and CT. The calculated radii showed differences of  $-0.7$  up to  $+1.2$  mm in lesion size compared with those determined by CT. The target ROI of one of the patients (patient no. 8) as found by auto-contouring following the contrast-oriented algorithm is presented in Fig. 6.

## Discussion

In the present investigation, we developed and evaluated an easily applicable contrast-dependent method for defining a GTV with FDG-PET, optimised for patients with locally advanced lung cancers. The phantom measurements showed that the TS for auto-contouring a volume in the FDG-PET image depends on the mean SUV within the volume ( $\text{mSUV}_{70}$ ) as well as on the background value

**Table 1** Results of volume delineation of the spheres of phan2 following the contrast-oriented algorithm (Eq. 3) in comparison with the true values

SUV <sub>max</sub>	True sphere volume (ml)	Measured sphere volume (ml)	Difference (%)	True sphere radius (mm)	Measured sphere radius (mm)	Difference (mm)
8.5	171.3	178.1	+3.9	34.5	34.9	+0.4
4.3		180.0	+5.1		35.0	+0.5
3.4		182.4	+6.5		35.2	+0.7
8.5	64.0	67.0	+4.7	24.8	25.2	+0.4
4.3		64.3	+0.4		24.9	+0.1
3.4		66.0	+3.2		25.1	+0.3
8.5	13.73	13.0	-5.6	14.9	14.6	-0.3
4.3		13.5	-1.7		14.8	-0.1
3.4		13.9	+1.2		14.9	+0.0
8.5	7.35	6.7	-8.8	12.1	11.7	-0.4
4.3		7.6	+3.4		12.2	+0.1
3.4		8.2	+11.6		12.5	+0.4

Differences in spheric volumes between measured and true values are given in % of the true volumes. Differences in spheric radius between calculated values and the true ones are given in mm

**Table 2** Results of FDG-based GTV delineation of the patients following the contrast-oriented algorithm (Eq. 3) in comparison with GTV definition using CT

Patient no.	Tumour localisation	SUV <sub>max</sub>	mSUV <sub>70</sub>	BG	GTV PET		GTV CT		Difference	
					Volume (ml)	Radius (cm)	Volume (ml)	Radius (cm)	Volume (%)	Radius (mm)
1	Peripherally, upper lobe	26.3	19.7	0.3	67.3	2.52	66.4	2.51	+1.3	+0.11
2	Peripherally, lower lobe	1.7	1.4	0.2	5.5	1.10	5.5	1.10	0.00	0.00
3	Peripherally, lower lobe	15.8	11.8	0.5	16.2	1.57	12.7	1.50	+27.8	+1.23
4	Peripherally, upper middle and lower lobe	16.8	11.3	0.3	20.8	1.70	20.7	1.70	+0.2	+0.01
5	Centrally	13.6	10.0	0.3	30.2	1.93	27.0	1.86	+11.7	+0.70
6	Peripherally, lower lobe	15.2	9.7	0.3	45.4	2.21	49.9	2.28	-9.0	-0.71
7	Peripherally, upper middle and lower lobe	21.7	16.3	0.3	72.5	2.59	69.0	2.54	+5.0	+0.42
8	Centrally	19.1	13.0	1.0	194.5	3.59	185.2	3.53	+5.1	0.60

Differences in volume and radius between PET and CT are also given. Radius values are calculated for spheres of equivalent volume

(BG). Considering the contrast as given in Eq. 1 and choosing an inverse regression allows transforming the calculated function into the mathematically simpler algorithm as presented in Eq. 3. This algorithm is easily applied and—adaptable to any system used—may become implemented into the software systems which are currently in use in clinical practice. The application of the method described in this study for the design of FDG-PET-based GTVs in lung cancer has the potential to reduce much of the inter- and intra-observer variation otherwise inherent in the use of FDG-PET-based radiotherapy planning today [7].

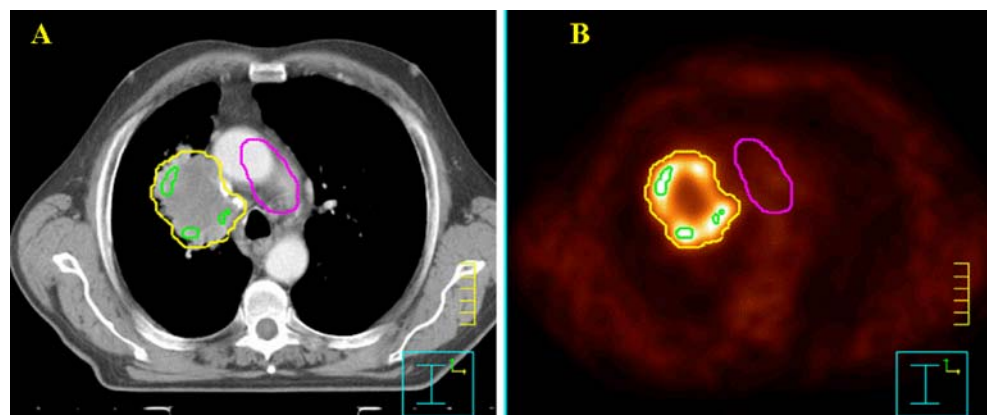
#### Validation in phantom

To validate the algorithm presented, we applied the method to another phantom (phan2) with different geometry which was chosen with regard to the phantom's use for quality control in a multi-centre trial. As phan2 is significantly smaller than phan1, those measurements show reduced scatter and randoms. Therefore, there was a need to validate carefully if the results determined for phan2 agree with

those of phan1. Using phan2, we observed a slight over-estimation of the volumes of the large spheres (diameter  $\geq 3$  cm) by up to 7%. This represents a calculated over-estimation in a radius of less than +0.7 mm which is much below the resolution as well as smaller than the pixel size in PET imaging used to date. For the small spheres (diameter  $< 3$  cm), we observed both under- and over-estimation of the volumes up to -6% and +12%, respectively. However, this difference in volume again represents a calculated difference in radius of less than -0.3 and +0.4 mm, respectively, which is also below the resolution of the system. Overall, the results for phan2 are in agreement with those observed for phan1. Applying the contrast-dependent method, adequate estimates were observed for objects having at least a diameter of more than twice the full-width-at-half-maximum (FWHM) of the PET scanner used. As expected, the accuracy of the results is limited by the spatial resolution of the scanner with relatively large voxels.

To calculate the threshold for auto-contouring, we decided to use two sets of parameters for the regression function: one set for spheres of diameter equal or larger

**Fig. 6** Transverse CT slice (a) and corresponding PET slice (b) of patient 8 including examples of target ROIs as found by auto-contouring the PET images following the contrast-oriented algorithm: green ROI generated by a 70% isocontour which is used to determine mSUV<sub>70</sub>; pink BG; yellow ROI generated by TS according to Eq. 3=GTV PET



than 3 cm and another for the spheres of diameter smaller than 3 cm. This distinction can be explained by the limited spatial resolution of the PET scanner which is more pronounced for the smaller spheres [29, 34]. Moreover, as lesion detectability not only depends on PET resolution but also on both the acquisition parameters and the reconstruction methodology used [35], this distinction may also reflect the influence of those parameters.

#### Validation in patient data

Unlike in real tumours, the FDG distribution in the spheres of the phantoms and in the background was homogenous. To validate the method in clinical practice, we applied the algorithm to lung tumours which were clearly delineable in CT images and, therefore, allowed GTV determination using CT at a high level of certainty. The volumes of primary tumours in eight patients with NSCLC were analysed and compared with the GTVs delineated by the radiotherapist in the CT data.

The volumes of lung tumours as measured by  $^{18}\text{F}$ -FDG-PET represent the accumulation averaged over several breathing cycles whilst the CT data rather represent a “snapshot” of the density at a short point in time within the breathing cycle. As predicted by Caldwell, volumes measured by FDG-PET may include not only the volume visible on CT but also the position of the tumour during the breathing cycle [36]. In concordance with this expectation, we observed over-estimation of the volumes by PET compared to CT up to 28%. This concordance is further supported by the fact that the tumour in which the largest over-estimation by FDG-PET occurred was located peripherally in the lower lobe of the right lung, in a region where extensive breathing movements are present. In no case was the volume defined by the contrast-oriented algorithm significantly smaller than the GTV delineated in CT. This is an important fact, since in radiotherapy planning, it is mandatory to safely encompass all malignant structures. Insufficient target volumes, being, e.g. too small to cover the whole tumour, may be a reason for local under-dosing and, therefore, tumour recurrence.

All patients but one (patient no. 2) showed relatively high values of FDG uptake. In the clinical situation, a  $\text{SUV}_{\text{max}}$  of 2.5 is often assumed for the differentiation between benign and malignant tumours. In fact, after histological examination, only very few primary lung cancers (e.g. patient no. 2) show a  $\text{SUV}_{\text{max}}$  below 2.5, whilst most tumours show a much higher  $\text{SUV}_{\text{max}}$ . For radiotherapy patients with inoperable NSCLC, a mean  $\text{SUV}_{\text{max}}$  of 17 is reported in the literature [33], whereas for all patients with NSCLC (including the operable), a mean  $\text{SUV}_{\text{max}}$  of 11 has been published [37].

#### Comparison with other methods

Various basic approaches for contouring FDG-based GTVs were reported in the literature. In an earlier publication, we compared the GTVs resulting from the methods most commonly used (visual contouring, absolute thresholding ( $\text{SUV}=2.5$ ), relative thresholding (40%  $\text{SUV}_{\text{max}}$ )) with those from the contrast-oriented algorithm described in this study in 25 primary lung tumours [17]. That investigation showed essential clinically relevant differences between the volumes contoured by the different methods. Apart from the visual delineation of the FDG-based GTVs by one observer, the volumes delineated by the contrast-oriented algorithm appeared to be most useful for radiotherapy treatment planning of lung cancer patients. In the context of this paper on clinical application, the question of the choice of the background region in patients was discussed extensively. In order to avoid redundancy, these details will not be discussed again in this report.

In another clinically oriented publication, we applied our algorithm to define the volume of 32 FDG-PET-positive lymph nodes using the parameters evaluated for spheres of diameter  $<3$  cm [38]. We compared both the delineability of the nodes and the calculated GTVs with those resulting from the above-mentioned methods used in the contouring of primary tumours. In contrast to the data for primary tumours, different methods of contouring did not lead to widely differing GTVs for lymph nodes, but to significant differences in the technical delineability. Whilst nearly all volumes visually detected could be defined by the contrast-oriented method ( $>95\%$ ), only 78% and 63% of the nodes could be delineated by an absolute SUV or a threshold of 40% of the  $\text{SUV}_{\text{max}}$ , respectively. We consider this result to be due to the ability of the method to cope with the low contrast between nodes and mediastinum.

Several authors have addressed the problem of GTV delineation in FDG-PET by automatic and semi-automatic methods which were also derived from phantom measurements. Different clinical settings were addressed by their publications. Most authors used spheres of comparably small volumes matched to the values of the expected patient tumour volumes [9, 18, 20–22, 39–42]. Daisne and colleagues who evaluated a signal-to-background algorithm, focussed on head-and-neck tumours and, therefore, used spheres of volumes from 2 to 17.15 ml corresponding to spheric diameters of less than 3.2 cm. Erdi and colleagues developed their method of image segmentation by using a percentage of the maximum FDG accumulation focussing on lung metastases caused by different neoplasms. They used spheres of volumes smaller than 5.5 ml (corresponding spheric diameters  $\leq 2.2$  cm) [9].

Our study addresses the question of GTV definition in patients with locally advanced NSCLC. Therefore, the



present analysis deals with spheres of relatively large volume sizes between 7.4 and 258.0 ml corresponding to spheric diameters between 2.5 and 8.0 cm. One other publication [19] reported on spheres of similar volumes (between 12.2 and 291 ml). The authors also focussed on patients with NSCLC, however, they evaluated an iterative method using the mean SUV of the lesion irrespective of the background

The contrast-oriented algorithm for contour delineation described in this study uses the parameter  $mSUV_{70}$  for the characterisation of the activity accumulation of the questionable lesion, whilst most other methods use the maximal SUV or maximal activity within the ROI [18]. The  $mSUV_{70}$  is defined as the mean SUV within a ROI surrounded by a threshold of 70% of the maximum activity in the sphere (or the lesion, respectively). In our eyes, using the mean SUV of a  $X\%$  isocontour instead of the maximal SUV with  $X$  sufficiently smaller than 100% substantially minimises the influence of statistically non-representative maximum values which have to be expected in clinical use due to the inhomogeneity of the FDG accumulation in tumours. The value of 70% of the maximum was chosen to make sure that the calculated  $mSUV_{70}$  is higher than the background value even for small faintly accumulating lesions within the mediastinum. Nevertheless, other values of  $X$  are also feasible, e.g. 60–80%, each of them resulting in different values of the parameters  $a$  and  $b$  used in Eq. 3. The value applied in clinical routine has to be adapted for the system used.

Daisne and coworkers who also proposed to use a signal-to-background algorithm in head-and-neck tumours tried to minimise the influence of statistically not representative maximum values by determining the threshold for auto-contouring as a percentage of the maximum activity which is defined as the average activity of the 9-voxel surrounding the hottest voxel [18, 20, 21]. However, the often inhomogenous uptake of FDG is not explicitly considered by this investigation. This is justified by the clinical application in the head-and-neck region because the volumes of the documented lesions are rather small (<20 ml).

Black and coworkers focussed on patients with NSCLC; however, they proposed an algorithm which uses the mean target SUV,  $SUV_{mean}$ , as the single parameter to determine the threshold-SUV for target volume delineation [19]. The authors stated that the mean target SUV calculation already includes the effects of background activity. Nevertheless, as the mean target SUV and the threshold-SUV are influenced by each other, an iterative method is needed to calculate the threshold-SUV which should converge to the expected value generally after not more than five iterations. This is in contrast to our aim of finding a uniformly applicable method for defining a GTV with FDG-PET in lung cancer

which can easily be used in clinical practice making iterative processes unsuitable. For comparison, we applied the algorithm proposed by Black et al. to our phantom data, resulting in the following parameters:  $SUV_{threshold} = 0.61 \times SUV_{mean} + 0.54$  ( $r = 0.97$ ). The application of this constant-threshold method in our patient data, however, consistently yielded smaller volumes than those delineated from CT (range = -4.8% to -38%). Similar results have been reported by Bayne and colleagues who also found discrepant GTVs using the algorithm of Black and colleagues to evaluate patient data [23]. The reason for the underestimation of tumour volumes may lie in the inhomogenous uptake of FDG in lung tumours or explicitly in the neglect of the background activity

### Limitations

There are some limitations to this study. First, there is no gold standard for the “true” volumes of real tumours measured by CT, which have to be compared to ungated PET data. Therefore, only estimates can be made about the influence of breathing excursions. The results of Caldwell et al. may give an idea of the physical order of magnitude of the size variations [36]. However, breathing excursions may vary depending on the size, localisation and surrounding of the tumour, as well as on the breathing mechanics in the individual patient. In the end, the results of clinical trials will have to show which contouring method can be used safely in patient care.

There are various approaches to the problem of breathing movements including gating in image acquisition and radiotherapy application [43, 44] or software correction techniques and hardware, like 4-D-CT [45]. In the future, also PET-MRI may provide new solutions. Most of these techniques, however, are still under investigation for their applicability in clinical routine [46] and their benefits to the patients still have to be evaluated. The setting addressed by the method presented in this study is the standard ungated irradiation of NSCLC patients as commonly used in clinical routine today.

Secondly, in cases of tumours with highly inhomogenous FDG uptake, e.g. central necrosis, the volumes delineated by our method may be too small. As the central necrosis within a tumour is part of the GTV, the volume resulting from auto-contouring has then to be visually corrected by the experienced nuclear medicine physician together with the radiotherapist. Moreover, as stated above, in a recent, more clinically oriented publication [17], we have compared the method described in this study with other methods currently used in clinical application. In this study, from the point of view of the radiotherapist, the GTVs in inhomogenous tumours generated by use of the contrast-oriented algorithm were more satisfactory than

the GTVs generated by the other methods. However, to date, no method of automatic contouring is able to fully cope with this problem.

Furthermore, no automatic or semi-automatic method can fully replace the experienced eye of the nuclear medicine physician who in a first diagnostic step will always have to decide which lesion has to be considered as a tumour in contrast to inflammatory lesions or normal tissue.

First measurements with our systems (ECAT-ART PET scanner/Philips Pinnacle planning system) have shown that parameters  $a$  and  $b$  will have to be adjusted for every combination of PET including the reconstruction method and parameters used, CT and planning system [24, 25]. Therefore, it has to be emphasised that although the method is uniformly applicable, the values of parameters  $a$  and  $b$  for Eq. 3 reported in this study are system-dependent. These values have to be separately adjusted for each system by phantom measurements as described in this study. In preparation of a multi-centre trial, we are currently investigating the validity and practicability of the algorithm described in this study by comparing the results of phantom measurements using different PET and radiotherapy planning systems in other centres. The results of these measurements will be published as soon as possible.

## Conclusion

By the contrast-oriented algorithm presented in this study, a PET-based delineation of GTVs for primary tumours of lung cancer patients is feasible. The TS for defining the target depends on  $mSUV_{70}$  as a measure of the activity of the lesion and on BG as a measure of the background activity. The algorithm is easily applicable and adaptable to any system used.

**Acknowledgement** We gratefully acknowledge the valuable help of Dipl. Ing. P. Donsch who constructed the phantoms. Furthermore, we thank Andrew Page for his help in the wording of the manuscript.

**Conflict of interest statement** None of the authors has any conflict between his or his family's financial interest and the content of this manuscript.

## References

- Baum RP, Hellwig D, Mezzetti M. Position of nuclear medicine modalities in the diagnostic workup of cancer patients: lung cancer. *Q J Nucl Med Mol Imaging* 2004;48:119–42.
- Mayor S. NICE issues guidance for diagnosis and treatment of lung cancer. *BMJ* 2005;330:439.
- Gambhir SS, Czernin J, Schwimmer J, Silverman DH, Coleman RE, Phelps ME. A tabulated summary of the FDG PET literature. *J Nucl Med* 2001;42:1S–93S.
- ICRU. Prescribing, recording and reporting photon beam therapy. Supplement to report 50. Bethesda, MD: ICRU; 1999.
- Grosu AL, Piert M, Weber WA, Jeremic B, Picchio M, Schratzenstaller U, et al. Positron emission tomography for radiation treatment planning. *Strahlenther Onkol* 2005;181:483–99.
- Nestle U, Kremp S, Grosu A. Practical integration of [(18)F]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): the technical basis, ICRU-target volumes, problems, perspectives. *Radiother Oncol* 2006;81:209–25.
- Pöttsch C, Hofheinz F, van den Hoff J. Vergleich der interobserver-variabilität bei manueller und automatischer volumenbestimmung in der PET. *Nuklearmedizin* 2006;45:A42.
- Knight SB, Delbeke D, Stewart JR, Sandler MP. Evaluation of pulmonary lesions with FDG-PET. Comparison of findings in patients with and without a history of prior malignancy. *Chest* 1996;109:982–8.
- 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.
- Hellwig D, Graeter TP, Gröschel A, Sybrecht GW, Schäfers HJ, Kirsch CM. FDG-PET for mediastinal staging of lung cancer: which SUV threshold makes sense? *J Nucl Med* 2007;48:1761–6.
- Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med* 2004;45:1519–27.
- Jaskowiak CJ, Bianco JA, Perlman SB, Fine JP. Influence of reconstruction iterations on 18F-FDG PET/CT standardized uptake values. *J Nucl Med* 2005;46:424–8.
- Lindholm P, Minn H, Leskinen-Kallio S, Bergman J, Ruotsalainen U, Joensuu H. Influence of the blood glucose concentration on FDG uptake in cancer—a PET study. *J Nucl Med* 1993;34:1–6.
- Menda Y, Bushnell DL, Madsen MT, McLaughlin K, Kahn D, Kernstine KH. Evaluation of various corrections to the standardized uptake value for diagnosis of pulmonary malignancy. *Nucl Med Commun* 2001;22:1077–81.
- Paquet N, Albert A, Foidart J, Hustinx R. Within-patient variability of 18F-FDG: standardized uptake values in normal tissues. *J Nucl Med* 2004;45:784–8.
- Schoder H, Erdi YE, Chao K, Gonen M, Larson SM, Yeung HW. Clinical implications of different image reconstruction parameters for interpretation of whole-body PET studies in cancer patients. *J Nucl Med* 2004;45:559–66.
- Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342–8.
- Daisne JF, Sibomana M, Bol A, Doumont T, Lonnew M, Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol* 2003;69:247–50.
- Black QC, Grills IS, Kestin LL, Wong CY, Wong JW, Martinez AA, et al. Defining a radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys* 2004;60:1272–82.
- Daisne JF, Duprez T, Weynand B, Lonnew M, Hamoir M, Reyckler H, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology* 2004;233:93–100.
- Geets X, Daisne JF, Gregoire V, Hamoir M, Lonnew M. Role of 11-C-methionine positron emission tomography for the delineation

- tion of the tumor volume in pharyngo-laryngeal squamous cell carcinoma: comparison with FDG-PET and CT. *Radiother Oncol* 2004;71:267–73.
22. Geets X, Daisne JF, Tomsej M, Duprez T, Lonneux M, Gregoire V. Impact of the type of imaging modality on target volumes delineation and dose distribution in pharyngo-laryngeal squamous cell carcinoma: comparison between pre- and per-treatment studies. *Radiother Oncol* 2006;78:291–7.
  23. Bayne M, MacManus M, Hicks R, Leong T, Peters L, Ball D. Can a mathematical formula help define a radiation target volume using positron emission tomography? In regard to Black et al. (*Int J Radiat Oncol Biol Phys* 2004;60:1272–1282). *Int J Radiat Oncol Biol Phys* 2005;62:299–300. author reply 300.
  24. Kremp S, Nestle U, Sebastian-Welsch C, Schaefer A, Kirsch CM, Rube C. Möglicher Nutzen der FDG-PET zur Abbildung der Tumorbewegung in der 3-D Bestrahlungsplanung von Bronchialkarzinomen mittels CT-PET fusionierter Bildgebung. *Strahlenther Onkol* 2006;182:51.
  25. Kremp S, Schaefer-Schuler A, Nestle U, Sebastian-Welsch C, Rube C, Kirsch C-M. Comparison of CT and CT-PET-fusion based 3D treatment plans in the percutaneous radiotherapy of lung cancer. *Radiother Oncol* 2004;73:S447–8.
  26. Nestle U, Hellwig D, Schmidt S, Licht N, Walter K, Ukena D, et al. 2-Deoxy-2-[18F]fluoro-D-glucose positron emission tomography in target volume definition for radiotherapy of patients with non-small-cell lung cancer. *Mol Imaging Biol* 2002;4:257–63.
  27. Nestle U, Walter K, Schmidt S, Licht N, Nieder C, Motaref B, et al. 18F-deoxyglucose positron emission tomography (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.
  28. Nestle U, Hellwig D, Fleckenstein J, Walter K, Ukena D, Rube C, et al. Comparison of early pulmonary changes in 18FDG-PET and CT after combined radiochemotherapy for advanced non-small-cell lung cancer: a study in 15 patients. *Front Radiat Ther Oncol* 2002;37:26–33.
  29. Bailey DL, Young H, Bloomfield PM, Meikle SR, Glass D, Myers MJ, et al. ECAT ART—a continuously rotating PET camera: performance characteristics, initial clinical studies, and installation considerations in a nuclear medicine department. *Eur J Nucl Med* 1997;24:6–15.
  30. Watson C, Schaefer A, Luk WK, Kirsch CM. Clinical evaluation of single-photon attenuation correction for 3D whole body PET. *IEEE Trans Nucl Sci* 1999;46:1024–31.
  31. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imag* 1994;13:601–9.
  32. Kremp S, Hellwig D, Schaefer-Schuler A, Sebastian-Welsch C, Ukena D, Rube C. Integration von CT/PET fusionierter Bildgebung in die 3-D-Bestrahlungsplanung des Bronchialkarzinoms. *Strahlenther Onkol* 2003;179:6. abstract.
  33. Borst GR, Belderbos JS, Boellaard R, Comans EF, Jaeger KD, Lammertsma AA, et al. Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. *Eur J Cancer* 2005;41:1533–41.
  34. Townsend DW, Wensveen M, Byars LG, Geissbuhler A, Tochon-Danguy HJ, Christin A, et al. A rotating PET scanner using BGO block detectors: design, performance and applications. *J Nucl Med* 1993;34:1367–76.
  35. Kadrmaz DJ, Christian PE. Comparative evaluation of lesion detectability for 6 PET imaging platforms using a highly reproducible whole-body phantom with (22)Na lesions and localization ROC analysis. *J Nucl Med* 2002;43:1545–54.
  36. Caldwell CB, Mah K, Skinner M, Danjoux CE. Can PET provide the 3D extent of tumor motion for individualized internal target volumes? A phantom study of the limitations of CT and the promise of PET. *Int J Radiat Oncol Biol Phys* 2003;55:1381–93.
  37. Vesselle H, Pugsley JM, Vallieres E, Wood DE. The impact of fluorodeoxyglucose F 18 positron-emission tomography on the surgical staging of non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2002;124:511–9.
  38. Nestle U, Schaefer-Schuler A, Kremp S, Groeschel A, Hellwig D, Rube C, et al. Target volume definition for (18)F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* 2007;34:453–62.
  39. Ciernik IF, Dizendorf E, Baumert BG, Reiner B, Burger C, Davis JB, et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853–63.
  40. Ciernik IF, Huser M, Burger C, Davis JB, Szekely G. Automated functional image-guided radiation treatment planning for rectal cancer. *Int J Radiat Oncol Biol Phys* 2005;62:893–900.
  41. Davis JB, Reiner B, Huser M, Burger C, Szekely G, Ciernik IF. Assessment of 18F PET signals for automatic target volume definition in radiotherapy treatment planning. *Radiother Oncol* 2006;80:43–50.
  42. Erdi YE, Rosenzweig K, Erdi AK, Macapinlac HA, Hu YC, Braban LE, et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). *Radiother Oncol* 2002;62:51–60.
  43. Nehmeh SA, Erdi YE, Ling CC, Rosenzweig KE, Squire OD, Braban LE, et al. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Med Phys* 2002;29:366–71.
  44. Jin JY, Ajlouni M, Chen Q, Yin FF, Movsas B. A technique of using gated-CT images to determine internal target volume (ITV) for fractionated stereotactic lung radiotherapy. *Radiother Oncol* 2006;78:177–84.
  45. Nehmeh SA, Erdi YE, Pan T, Pevsner A, Rosenzweig KE, Yorke E, et al. Four-dimensional (4D) PET/CT imaging of the thorax. *Med Phys* 2004;31:3179–86.
  46. Grgic A, Nestle U, Schaefer-Schuler A, Kremp S, Kirsch C-M, Hellwig D. FDG-PET-based radiotherapy planning in lung cancer: optimum breathing protocol and patient positioning—an intra-individual comparison. *Int J Radiat Oncol Biol Phys*. 2008; doi:10.1016/j.ijrobp.2008.03.063.