REVIEW ARTICLE

PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques

Habib Zaidi · Issam El Naqa

Abstract Historically, anatomical CT and MR images were used to delineate the gross tumour volumes (GTVs) for radiotherapy treatment planning. The capabilities offered by modern radiation therapy units and the widespread availability of combined PET/CT scanners stimulated the development of biological PET imaging-guided radiation therapy treatment planning with the aim to produce highly conformal radiation dose distribution to the tumour. One of the most difficult issues facing PET-based treatment planning is the accurate delineation of target regions from typical blurred and noisy functional images. The major problems encountered are image segmentation and imperfect system response function. Image segmentation is defined as the process of classifying the voxels of an image into a set of distinct classes. The difficulty in PET image segmentation is compounded by the low spatial resolution and high noise characteristics of PET images. Despite the difficulties and known limitations, several image segmentation approaches have been proposed and used in the clinical setting including thresholding, edge detection, region growing, clustering, stochastic models, deformable models, classifiers and several other approaches. A detailed description of the various approaches proposed in the literature is reviewed. Moreover, we also briefly discuss some important considerations and limitations of the widely used techniques to guide practitioners in the field of radiation oncology. The strategies followed for validation and comparative assessment of various PET segmentation approaches are described. Future opportunities and the current challenges facing the adoption of PET-guided delineation of target volumes and its role in basic and clinical research are also addressed.

Keywords PET · Segmentation · Radiation therapy · Treatment planning · Validation

Introduction

The foundation of molecular imaging-guided radiation therapy (MIGRT) lies in the use of advanced multimodality imaging technology for improved delineation of tumour target volumes, thus relating the absorbed dose information to image-based patient representations. This was made possible by the introduction and clinical acceptance of dual-modality single photon emission computed tomography (SPECT)/computed tomography (CT) and positron eEmission tomography (PET)/CT systems in the clinical setting in the late 1990s which has revolutionized the practice of diagnostic imaging [1, 2]. Combined PET/CT systems have been operational for almost a decade since their commercial introduction. The complementarity between the intrinsically aligned anatomical (CT) and functional or metabolic (PET) information provided in a "one-stop shop" and the opportunity to reduce scanning time through the use of CT images for attenuation correction of the PET data contributed to the success of this technology and its wide adoption by the medical imaging community.

H. Zaidi (✉)
Division of Nuclear Medicine, Geneva University Hospital,
1211 Geneva 4, Switzerland
e-mail: habib.zaidi@hcuge.ch

H. Zaidi
Geneva Neuroscience Center, Geneva University,
1205 Geneva, Switzerland

I. El Naqa
Department of Radiation Oncology,
Washington University School of Medicine,
St. Louis, MO 63110, USA

On the other hand, radiation therapy has gone through a series of revolutions in the last century [3] advocating today the opportunity to generate highly conformal radiation dose distributions through the use of advanced techniques such as intensity-modulated radiation therapy (IMRT) using tomotherapy [4] and volumetric-modulated arc therapy [5] and many other techniques including CyberKnife, heavy ion therapy [6] and many other radiation therapy devices for dose conformity. The improved dose conformity and steep dose gradients have necessitated enhanced patient localization and beam targeting techniques for radiotherapy treatments.

The most critical component of radiation therapy treatment planning is the delineation of the gross tumour volume (GTV) [7]. Historically, anatomical cross-sectional images (mainly CT and in some instances MR) are used to delineate the treatment volumes [8, 9], and radiation treatment beams are planned to completely cover the treatment volume with the aim to deliver a uniform dose distribution to it. CT can provide useful anatomical information and the electronic density required for dosimetry calculations for radiation therapy treatment planning. However, it has poor soft tissue contrast, which might be insufficient for target and critical organ delineation. MRI provides different signatures of the human body and presents better soft tissue contrast compared to CT. Many investigators reported significant differences in terms of target volume delineation when using MRI compared to CT [10, 11]. It has been argued that coregistered CT/MRI data can be used with confidence for accurate delineation of target volumes and critical organs (using MRI) and dose computation (using CT). However, MRI suffers from spatial distortions, intensity inhomogeneity and does not provide electronic tissue density information required for dose calculation in radiation therapy. Moreover, the criteria followed to delineate the GTV are not standardized leading to high inter- and intra-observer variability [12].

The above-discussed limitations together with the enhanced capabilities offered by advanced treatment units to deliver nonuniform dose patterns by design stimulated the emergence of the biological target volume (BTV) concept [13] made possible by the commercial availability of multimodality molecular imaging platforms [14]. It has therefore been argued that molecular PET/CT imaging may be of additional value and even allow a more correct delineation of the BTV [15–20]. The most interesting aspect related to the use of PET for BTV delineation is mainly the depiction of functional or metabolic rather than morphological information provided by CT or MRI. However, the application of PET for this purpose is not trivial and has many drawbacks. Among the concerns associated with PET-guided radiation therapy treatment planning is the complexity of the procedures that need to be

developed for accurate delineation of target regions from typical blurred and noisy functional images suffering from many instrumentation- and physics-related factors [16, 21]. The most important challenges are image segmentation and imperfect system response function. Image segmentation is defined as the process of classifying the voxels of an image into a set of distinct classes. Medical image segmentation has been identified as the key problem of medical image analysis and remains a popular and challenging area of research [22–25]. The difficulty in PET image segmentation is compounded by the low spatial resolution and high noise characteristics of PET images. Despite the difficulties and known limitations, several image segmentation approaches have been proposed and used in the clinical setting including thresholding, region growing, classifiers, clustering, edge detection, Markov random field models, deformable models and many other approaches [26–29]. The number of scientific contributions related to this subject has been increasing steadily, which motivated the writing of this survey as a snapshot of the dynamically changing field of PET image segmentation. A detailed description of the various approaches proposed in the literature is given. We also briefly discuss some important considerations and limitations of widely used techniques. The strategies followed for validation and comparative assessment of various PET segmentation approaches are outlined. Future opportunities and the challenges facing the adoption of PET-guided delineation of target volumes and its role in basic and clinical research will also be addressed.

Molecular PET/CT imaging-guided radiation therapy treatment planning

Early attempts to use nuclear medicine imaging and particularly PET for radiation therapy treatment planning date back to the late 1990s [30–37]. An important contribution came from Ling et al. [13] who established the concept of "biological imaging" and moved forward the role of PET in radiation therapy thus allowing it to enter the clinical arena. Since that time, the technical aspects of PET/CT-guided radiation therapy have been described more thoroughly in the scientific literature [15, 21, 38–45]. The success of these initial studies prompted significant interest from the major medical imaging equipment manufacturers who now all have introduced commercial PET/CT scanners equipped with the required accessories (flat couch insert, positioning system, respiratory gating, ...) and software tools (e.g. virtual simulation, visualization and segmentation tools, support of DICOM RT object definition, ...) for clinical use. The typical workflow for PET/CT-guided radiation therapy treatment planning usually involves (with

few exceptions) strong collaboration between two clinical departments (nuclear medicine and radiation oncology) [21, 46]. With the growing availability of large bore dual-modality PET/CT scanners of the last generation equipped with fixed radiation therapy positioning laser systems in the scanner room, a "one-stop shop" providing diagnostic PET/CT and radiation therapy planning CT scan in only one session has become possible. It should however be noted that the 80-cm bore required for stereotactic body radiotherapy planning and to accommodate various other positioning aids is available only from one scanner manufacturer at the present time.

The main motivation stimulating the use of PET/CT in radiation therapy is the efficacy of ^{18}F-fluorodeoxyglucose (FDG) PET imaging in a wide variety of malignant tumours with sensitivities, specificities and accuracy often in the high 90th percentile range [47]. In that sense, it might provide superior visualization compared to CT simulation, which in some cases might miss some areas that light up on the PET study including the detection of distant metastases, or shed light on the actual lesion volume, which might in reality be smaller on the PET study than on the CT alone. Moreover, discrepancies between anatomical (CT/MRI) and metabolic (PET) findings are very often reported in the literature where the addition of PET significantly impacted patient management and changed the treatment plans in 25–50% of the cases [16, 20, 27, 38, 39, 48–53]. Last and not least, inter- and intra-observer variability was considerably reduced when PET information was available for target volume delineation [54–58]. Both state-of-the-art FDG PET and novel PET probes' applications in the process of radiation therapy treatment planning can be found elsewhere [21, 44, 45, 59] and are beyond the scope of this review, which is solely focused on FDG PET.

Current indications for ^{18}F-FDG PET/CT-guided radiation therapy treatment planning fall under two classes: established and experimental. Well-established indications include head and neck cancer, lung cancer, gynaecological and oesophageal cancer whereas experimental indications comprise colorectal cancer, breast cancer, lymphoma and malignant melanoma and many other malignancies.

A technical discussion of issues related to margin set-up and link between GTV, clinical treatment volume (CTV), planning treatment volume (PTV) and the BTV is beyond the scope of this review. Interested readers could consult the International Commission on Radiation Units and Measurements (ICRU) report defining the relationship between the above-mentioned quantities for use with conventional imaging modalities [7] and updates by other groups following incorporation of PET information [13, 52]. We emphasize that the discussed PET image segmentation algorithms tend to capture the BTV while CTV margins are related to micro-disease extensions. Moreover, due to the

varying biodistribution and contrast of the various probes (see "Challenges and future directions"), different segmentation methods might have to be used in the era of multitracer multimodality imaging [60].

Among the many physical degrading factors impacting image quality and quantitative accuracy, partial volume effect is recognized as one of the most important factors impacting tumour imaging [61]. Owing to the limited spatial resolution of clinical PET systems, the resulting images will be blurred by the system response and as a consequence smaller lesions will appear larger. Although the total number of counts is preserved, they are distributed over a larger volume. Partial volume is usually addressed in the context of "small" objects, i.e. with dimensions smaller than around 2–3 times the full-width at half-maximum of the scanner's point spread function [62]. Several techniques and corrective measures have been developed to address this problem in the context of molecular PET imaging which operate either directly on reconstructed images or are incorporated within the image reconstruction framework with many of them relying on an adjunct coregistered anatomical image (CT or MRI). It should be noted that their availability and implementation in clinical packages for use in the clinical setting are still lacking. Respiratory motion is another challenging issue for lesions located in the thorax and upper abdomen that has been addressed in many different ways [63, 64]. The most successful approaches attempt to overcome the limitations of the traditional approach, which allows obtaining individually reconstructed noisy images through respiratory gating by incorporating motion estimation and correction within the image reconstruction process to obtain images of enhanced quality [65–70].

Survey of PET image segmentation techniques

Over the past few years, several methods have been proposed for target volume definition in radiotherapy treatment planning based on incorporating PET physiological information. In particular, FDG PET, a glucose analogue, is currently used in many cancer centres around the world to improve BTV definition, which is traditionally identified on CT simulation images in radiotherapy clinical routine. Accurate volume definition is particularly important in radiotherapy because it constitutes the target of the radiation beam; under-dosing of tumour may lead to recurrence while over-dosing of surrounding normal tissues might lead to severe and possibly lethal side effects to the patient such as brain or lung injury [71]. There are several ways to categorize PET segmentation approaches such as the cancer site, the injected radiotracer or the image processing technique. It is noted that there could be

differences and overlaps between the sites, the tracers or the techniques. Nevertheless, in this survey, we find that it would be pedagogically convenient to categorize PET segmentation based on the techniques used and refer to differences in sites or tracer-specific variations as appropriate. There is a plethora of segmentation methods that could be applied to nuclear medicine imaging, particularly in cardiovascular imaging; the interested reader is referred to [24]. According to our literature survey of existing methods, we identified four broad categories of PET segmentation methodologies: (1) image thresholding methods, (2) variational approaches, (3) learning methods and (4) stochastic modelling-based techniques. The main characteristics and limitations of these categories are summarized in Table 1.

Thresholding methods

This is by far the most widely used PET segmentation approach in clinical practice for BTV delineation for radiotherapy planning. The only competing approach with thresholding is possibly visual interpretation of PET scans and identification of lesion boundary by consensus reading of an experienced nuclear medicine physician and radiation oncologist [33]. However, visual inspection is susceptible to the window level settings and suffers from inter-observer variability (Fig. 1). Therefore, several segmentation methods based on thresholding have emerged to reduce this subjectivity. In the thresholding approach, a threshold value (T) is selected to make a hard decision of separating the lesion foreground from the noise background within the

whole image or a selected region of interest (ROI) within the image $I(\mathbf{x})$, thus:

$$Lesion = T[I(\mathbf{x})] = \begin{cases} 1, I(\mathbf{x}) \geq T \\ 0, I(\mathbf{x}) < T \end{cases} \tag{1}$$

where I is the intensity image and $\mathbf{x}$ is the spatial coordinate (x, y, z). Various methods have been proposed in the literature to determine the "optimal" threshold value [72–75]. However, almost uniformly across all PET thresholding applications, the PET image voxels are converted into standardized uptake values (SUVs) prior to threshold selection. SUV is a decay-corrected measurement of activity per unit mass of tissue adjusted for administered activity per unit of body weight [76]:

$$SUV = \frac{Mean\ ROI\ concentration\ (MBq/g)}{Injected\ dose\ (MBq)/Body\ weight(g)} \times \frac{1}{decay\ factor} \tag{2}$$

Several technical and patient-related factors can affect the reliability of SUV [77]. These include but are not limited to (1) time interval between injection and PET scanning, (2) degree of infiltration of administered tracer at the site of injection, (3) the often neglected residual activity in the syringe, (4) correction for radioactive decay of the injected tracer and (5) partial volume effects due to limited spatial resolution of the PET scanner [78, 79]. The interested reader is referred to Table 1 in both above-referenced reviews.

Table 1 Summary of main pros/cons of the various categories of PET image segmentation techniques

Category	Characteristics	Limitations
Manual techniques	Visual interpretation and manual delineation of contours. Very simple to use. Tools to transfer RT objects to treatment planning systems available from most vendors	Time consuming. Susceptibility to window level settings. Suffer from intra- and inter-observer variability. Consensus reading by nuclear medicine physician and radiation oncologist hardly practical in busy clinical departments
Thresholding techniques	Most frequently used due to their simple implementation and high efficiency	Hard decision making. Too sensitive to PVE, tumour heterogeneity and motion artefacts. Some methods focus on volume, others focus on intensity differences. Combination of both seems to provide best results [95]
Variational approaches	Subpixel accuracy, boundary continuity and relatively efficient. They are mathematically well developed and allow for incorporation of priors such as shape	Sensitive to image noise. As a PDE, stability and convergence could be subject to numerical fluctuations, especially if the parameters are not properly selected
Learning methods	Utilize pattern recognition power. Two main types: supervised (classification) and unsupervised (clustering)	Computational complexity especially in supervised methods, which require time-consuming training. Feature selection besides commonly used intensity is a flexibility but can also be a challenge
Stochastic models	Exploit statistical differences between tumour uptake and surrounding tissues. Most natural to deal with the noisy nature of PET	Effect of initialization and convergence to local optimal solutions are concerns, especially when compromises are made to improve efficiency

Fig. 1 Illustration of the operator dependence of manual delineation of the contours on the window level settings (SUV thresholding) used for display of PET images. The images show a representative slice of a NSCLC patient. **a** Anatomical CT image. **b** Fused PET/CT image and example PET images displayed using two different percentages of the maximum threshold (SUV_{max}) (**c, d**). Note the differences in the maximum diameters estimated on the horizontal axis of the transaxial images: 5.36 cm on the CT (**a**) vs 3.47 and 5.48 cm on the two PET images displayed using different window level settings (**c, d**)
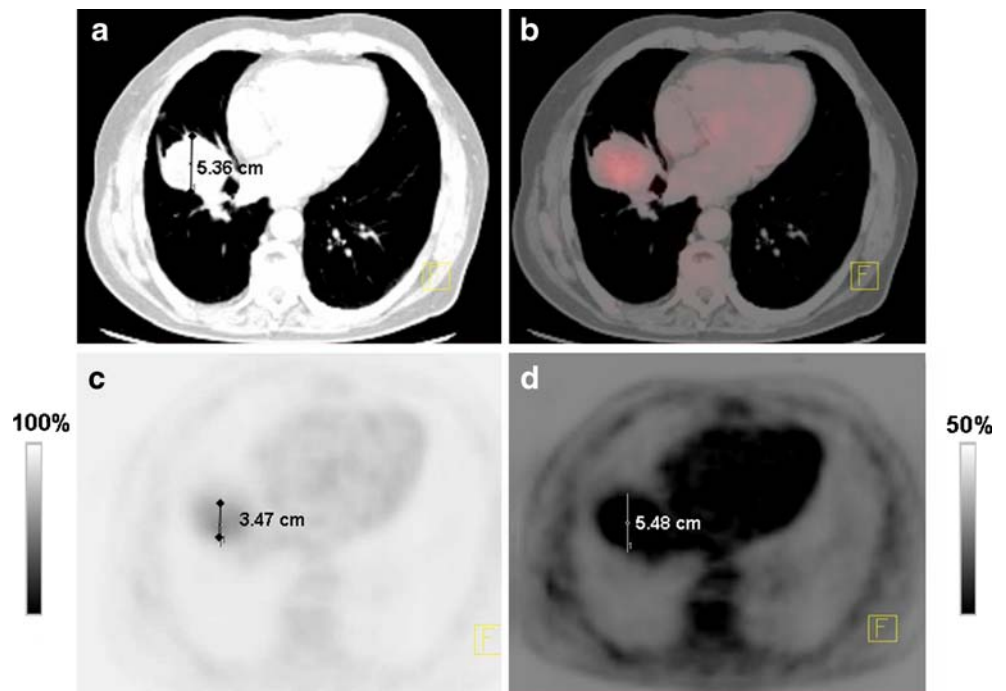
Among the earliest thresholding methods used is thresholding by percentage of 40–50% of the maximum SUV based on phantom studies of stationary spheres with FDG [80]. Subsequently, a value of 40% was adopted by several groups for BTV delineation of non-small cell lung cancer (NSCLC) [48], cervical cancer [81] and head and neck cancer (HNC) [40], for instance. However, later studies have shown that the optimal fraction selection for accurate quantification is dependent on the tumour volume [82, 83]. Furthermore, Biehl et al. have shown a logarithmic relationship between the tumour volume and the optimal selected threshold using manually delineated CT simulation data as gold standard in a cohort of NSCLC patients (Fig. 2):

$$\%Threshold = 59.1 - 18.5 \times \log 10(\text{tumour volume}) \quad (3)$$

This inverse relationship between threshold and volume confirms the observations seen in phantom studies [84, 85]. An alternative method based on the absolute 2.5 SUV isocontour was proposed by Paulino et al. [86]. Proponents of this threshold argue that this is a value that distinguishes between benign and malignant tumours. However, tumour inhomogeneity and motion artefacts may hinder the application of this approach. The choice of an empirical SUV threshold-based cut-off to differentiate malignant lesions is controversial and an isocontour of a 2.5 SUV has been shown to fail to provide successful delineation in a large number of cases (45% for FDG PET [87] and 78% for FET PET [88]), particularly when using tracers presenting with low uptake in the target. Among the problems

associated with the use of fixed thresholding techniques is that they do not take the background into consideration. In addition, the threshold required for accurate edge location is a function of the source size. Inhomogeneity of tumours is a further confounding factor for the optimal percentage thresholding [89]. In order to avoid the controversies associated with fixed thresholds, some authors suggested adapting the threshold value to the signal to background ratio (SBR) [90, 91]. This technique requires calibration data that are applicable for quantification of tumour volumes in clinical oncology. This is usually carried out using a body phantom consisting of an elliptical water-filled cavity
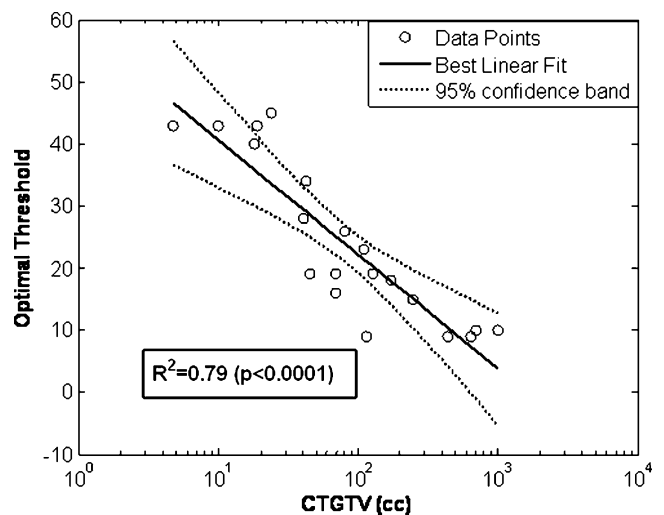

Fig. 2 Optimal percentage maximum SUV threshold selection in NSCLC. Reprinted with permission from [82]

with spherical inserts suspended by plastic rods of various inner diameters filled with activity concentrations in the range of SBR observed in clinical conditions to derive the parameters required for implementation of the adaptive thresholding algorithm based on SBR estimates. The measured SBR and the resultant threshold are then fitted using an inverse function (*Threshold* = $a + b \times 1/SBR$) to yield the best regression parameters for the equation which allows estimation of the optimal threshold independent of any a priori knowledge of the lesion characteristics and tumour location.

Black et al. proposed using the following function to define the optimal threshold [90]:

$$Threshold = (0.307 \times mean-target-SUV + 0.588) \quad (4)$$

According to this function, the threshold value is dependent on the mean target SUV and is not related to the background concentration or the target volume. This is in contrast to the multivariate analysis results by Brambilla et al. [92] and the discussion above. A different function that depends on each slice-specific image contrast defined as the difference between maximum voxel intensity within the target (IT_{max}) and effective mean voxel intensity of the background region (IB_{mean}) was also reported by Drever et al. [93]:

$$Threshold = (Contrast\ level \times (IT_{\max} - IB_{mean}) + IB_{mean}) \quad (5)$$

Nestle et al. [94] used the following formulation:

$$Threshold = \beta I_{mean} + I_{bkg} \quad (6)$$

where β=0.15, I_{mean} is computed as the mean intensity of all pixels surrounded by the 70% I_{max} isocontour within the tumour and I_{bkg} is the mean intensity in a neighbouring organ sufficiently far from the target and having the highest tracer uptake. The above formulation was slightly revised more recently and optimized using specially designed phantoms containing spheres of different diameters [95]:

$$Threshold = \left(a \times SUV_{mean(70\%)} + b \times BKG\right)/SUV_{\max} \quad (7)$$

where a and b are sphere size and PET scanner-dependent parameters reported in this paper as $a=b=0.50$ for large spheres (diameter ≥3 cm) and with $a=0.67$ and $b=0.60$ for small spheres (diameter <3 cm) for the ECAT ART system (Siemens Medical Solutions, Knoxville, TN, USA). An iterative thresholding scheme to the SBR approach was introduced by Jentzen et al. [96], in which the authors suggested the following initialization:

$$\%Threshold = 7.8\%/V(ml) + 61.7\% \times B/S + 31.6\% \quad (8)$$

An alternative iterative technique based on a mathematical fit derived from Monte Carlo simulation studies to estimate the thresholds uses the following formulation [97]:

$$\%Threshold = a_0 + \exp^{[a+(b/V)+c\ \log V]} \quad (9)$$

where the parameters resulting from the least-squares fit describing the correlation between the lesion volume (V) and the corresponding optimum threshold are a_0=5, a=3.568, b=0.197 and c=−0.1069.

Despite their limitations, thresholding approaches have an intuitive basis that the other methods lack. Specific review of thresholding applications for target definition in HNC could be found in [87, 98] and for NSCLC in [94].

Variational approaches

These methods attempt to exploit information provided by intensity variation (gradient differences between the foreground lesion and the background) for the segmentation task. Several methods fall under this category [99–101]. This includes simple edge or ridge detectors such as the Sobel operator and the Watershed transform (WT) evaluated in [102]. However, their results did not show improvements compared with thresholding methods. This is in contrast with the results obtained by Geets et al. [103], in which the WT was applied in conjunction with cluster analysis on pre-processed images using a bilateral denoising filter and a Landweber's deblurring filter with a Gaussian kernel. This pre-processing reduced the known sensitivity of the WT to noise amplitude in PET images.

More recently, there has been more focus on applying deformable active contour models to PET segmentation [104–106]. These models are geometric representations for curves (in 2D) or surfaces (in 3D) and are defined explicitly or implicitly in the imaging domain. The active models move (deform) under the influence of so-called internal force-like equations, which represent the curve or the surface bending characteristics, or so-called external forces, which are computed from the image data, typically as directional gradients [107, 108]. The idea is that contours are characterized by sharp variations in the image intensity. Hence, the objective is to match deformed and reference contours by means of energy minimization (or dynamic force equilibrium). Mathematically, if the deformable contour/surface is represented by $C(s) = \{x(s), y(s),$

$z(s)\}$, $s \in [0, 1]$, then its movement is governed by the following function:

$$J(C(t)) = \int_0^1 \alpha(s) \left| \frac{\partial C(s;t)}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 C(s;t)}{\partial s^2} \right|^2 ds + \gamma \int_0^1 P(C(s;t)) ds$$

(10)

where the first term corresponds to internal energy and controls the tension and rigidity of the contour.[1] The second term corresponds to the external energy, where P represents the potential energy given as $g(|\nabla I|)$, and g is selected to be a monotonically decreasing function of the image intensity (I) gradient. Using calculus of variation techniques, the solution to Eq. 10 is obtained by solving the associated Euler-Lagrange partial differential equation (PDE). This type of parametric active contours is known in the literature as "snakes" [109–111]. In order to improve the capture range of the snake-based algorithms, Xu and Prince proposed the gradient vector flow (GVF) snake [112]. This is achieved by replacing the external forces component (∇P) by GVF, which is computed as a diffusion of the gradient vectors in the image. Hsu et al. applied this concept to segment liver PET images. However, in their approach they estimated the external forces by solving a Poisson PDE instead and the algorithm was initialized by a Canny edge detection method [105]. To further overcome the limitations encountered in the snake-based parametric deformable models such as initialization requirement and topological adaptation (splitting or merging of model parts as encountered in necrotic tumours), a family of geometric deformable models known as the level set methods were developed [113]. These techniques are based on using the geometric concept of evolving level sets by solving iteratively the following evolutionary PDE:

$$\frac{\partial \phi}{\partial t} = V(\kappa) |\nabla \phi|$$

(11)

where ϕ is an implicit function (e.g. a signed distance) that represents the evolving level set, where at the contour $\phi(C) = 0$. V is a velocity function proportional to the curvature and inversely proportional to the image gradient. The algorithm starts by some initial contour, then the curve evolves under the influence of the internal (contour curvature) and external forces (image gradient) until it reaches the boundary of the object where balance between the internal and the external forces is achieved as governed by the solution of Eq. 11. Efficient solutions are developed

by using the marching method [107]. This method was applied to segmentation of PET images in NSCLC [104] and HNC [106]. El Naqa et al. [104] applied an iterative deblurring algorithm while Li et al. [106] used RG as pre-processing steps to improve the active contour robustness. Sample results illustrating application of the level set approach are shown in Fig. 3.

Learning methods

These are techniques developed in pattern recognition to estimate dependencies from data [114]. In the case of lesion segmentation in PET, the learning task becomes to discriminate uptakes in the lesion voxels from surrounding normal tissue voxels based on a set of extracted features from these images. There are two common types of statistical learning: supervised and unsupervised [114, 115]. Supervised learning is used to estimate an unknown (*input*, *output*) mapping from known labelled samples called the training set (e.g. classification of lesions given a certain number of example images). In unsupervised learning, only input samples are given to the learning system (e.g. clustering or dimensionality reduction).

Commonly used classifiers include k-nearest neighbour (KNN), artificial neural network (ANN) and support vector machine (SVM). These methods have seen limited application in PET; however, they are widely applied in segmentation of anatomical MR images [116–118] and X-ray images [119]. This is partly due to the high heterogeneity of PET images that complicates the identification of proper stable patterns for training purposes from a finite number of observations [89]. On the other hand, clustering methods have shown more promise for PET segmentation of tumours [89, 120–123] in addition to many other applications in nuclear medicine [124, 125]. There are several clustering methods in the literature [126, 127]. Among the frequently used ones are the k-means algorithm, the fuzzy C-means (FCM) algorithm and the expectation maximization (EM) algorithm. The k-means is the most widely used due to its simplicity. The algorithm is first initialized to K-initial cluster centres, which are selected either manually or randomly. Then, the cluster memberships are updated iteratively by applying a hard decision distance metric (e.g. Euclidean distance). This is achieved by minimizing the following objective function:

$$J(\mathbf{x}, \mathbf{c}) = \sum_{i=1}^N \sum_{k=1}^K \| \mathbf{x}_i - \mathbf{c}_k \|^2$$

(12)

where N is the number of voxels, K is the number of tissue classes, $\mathbf{x}_i$ is a feature vector at the ith location and c is the kth class cluster centre. However, such a technique is quite

---

[1] The first-order derivative suppresses stretching and makes the contour behave like an elastic string. The second-order derivative suppresses bending and makes the model behave like a rigid rod.

Fig. 3 **a** Fused PET/CT with a necrotic lesion area magnified before and after deblurring. **b** Auto-segmentation using level set method. *Left* initial contour, *middle* evolved contour after 40 iterations and *right* the estimated contour (*red*). The algorithm converged in 100 iterations
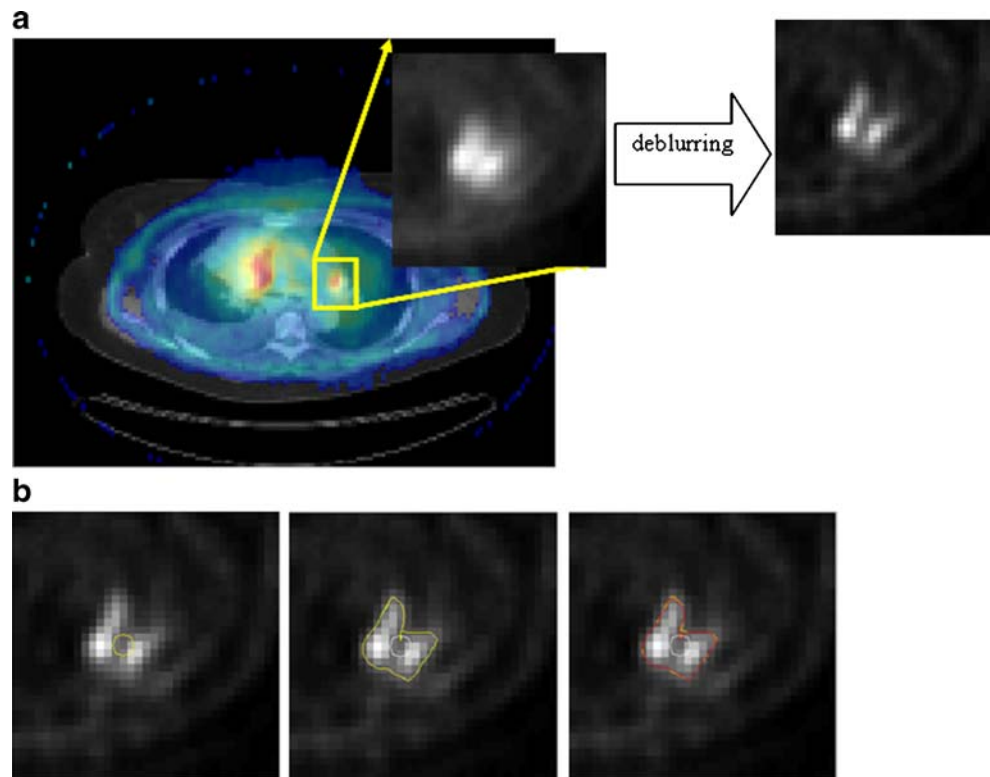
sensitive to initial cluster selection and is not robust to noise and spatial inhomogeneities. Several modifications have been proposed in the literature to alleviate these problems. One approach is to replace the hard decision approach in the k-means by a soft decision approach by applying the fuzzy set theory [128]. The algorithm proceeds in a similar fashion to k-means, but a voxel would belong to more than one class, and cluster association is built in the algorithm by using fuzzy membership functions [129, 130]. The membership function represents the amount of stochastic overlap "fuzziness" between the tumour region and surrounding regions. In the FCM case, the fuzzy member function at any iteration n is given by:

$$u_{ik}^{(n)} = \frac{\left\| \mathbf{x}_i - \mathbf{c}_k^{(n)} \right\|^{-2}}{\sum_{k=1}^{K} \left\| \mathbf{x}_i - \mathbf{c}_k^{(n)} \right\|^{-2}} \tag{13}$$

and the update for cluster centres is:

$$c_k^{(n+1)} = \frac{\sum_{i=1}^{N} \left(u_{ik}^{(n)} \right)^b \mathbf{x}_i}{\sum_{i=1}^{N} \left(u_{ik}^{(n)} \right)^b} \tag{14}$$

where $\mathbf{x}_i$ again is the feature vector at the ith location, $\mathbf{c}_k^{(n)}$ is the kth centroid at the nth iteration and b is an exponent >1. A variation of this method was applied in [124], in which the algorithm starts with an oversized number of clusters to avoid misidentification of conflicting regions. Then, this is

followed by a merging process to reach the desired or natural number of clusters according to a priori anatomical knowledge. Similar clustering techniques were also used for unsupervised lesion segmentation operating on dynamic PET images based on time-activity curve (TAC) shape differences between malignant and healthy tissues [131]. For lesion segmentation, the TAC slope values were k-means clustered into two clusters.

To overcome the above-described limitations of the FCM algorithm, a new fuzzy segmentation technique based on the standard FCM algorithm and adapted to typical oncological PET data was proposed by Belhassen and Zaidi [123]. PET images are first smoothed using a nonlinear anisotropic diffusion filter [132] and are then added as a second input to the proposed FCM algorithm to conduct the objective function optimization with knowledge about spatial constraint, thus incorporating spatial information (FCM-S). In addition, a methodology was developed to integrate the à trous wavelet transform [133] into the standard FCM algorithm (FCM-SW) to allow handling of heterogeneous lesions' uptake. This is achieved by adding a regularization term to the FCM objective function using the transformation result of the PET image by the à trous wavelet transform with the aim to incorporate information about lesion heterogeneity.

Stochastic modelling

Methods belonging to this category aim to exploit the statistical differences in intensity distribution between

tumour PET uptake and its tissue surroundings. Aristophanous et al. proposed a Gaussian mixture model (GMM) for segmenting NSCLC PET images. In this approach, it was assumed that the image intensities are independent and identically distributed with a Gaussian probability density function that could be divided into three regions: background, the uncertain and the target regions [134]. Hence, the likelihood function is written as:

$$L(\pi, \mu, \sigma) = \prod_{i=1}^{N} f(x_i / \pi, \mu, \sigma) = \prod_{i=1}^{N} \sum_{k=1}^{K} \frac{\pi_k}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}}$$

(15)

where N is the number of voxels, K is the number of classes, π are the mixing parameters and μ, σ are the Gaussian parameters. The maximum likelihood estimates of the unknown parameters are obtained using the EM algorithm and the probability of voxel x_i belonging to class k is given by:

$$p_{ik} = \frac{\pi_k f_k(x_i / \mu_k, \sigma_k)}{\sum\limits_{m=1}^{K} \pi_m f_m(x_i / \mu_m, \sigma_m)}$$

(16)

It should be noted that EM is a general approach for maximum likelihood estimation. The most widely used approach for image segmentation uses a two-step EM algorithm where the E-step is dedicated to the computation of the probabilities and the M-step estimates the cluster parameters assuming that the intensity distribution of each class may not be Gaussian and assigns belonging probabilities according to non-Gaussian distributions, e.g. [135, 136].

On the other hand, the spatial independence assumption could be relaxed by using hidden Markov models [133, 137]:

$$p_{ik} = \frac{\exp(-H(g(x_i), k)}{\sum\limits_{m=1}^{K} \exp(-H(g(x_i), m)}$$

(17)

where H is a Gibbs potential function and g is a neighbourhood partition (clique). Hatt et al. utilized a fuzzy membership of the voxels [138] into the Markov chain model in their implementation [137] referred to as FHMC, whereas Montgomery et al. used a multiscale decomposition based on the wavelet transform to improve performance [133]. More recently, Hatt et al. modified their FHMC algorithm into fuzzy locally adaptive Bayesian (FLAB) to improve segmentation of small objects [89]. This Bayesian approach uses adaptive estimation of priors using a family of distributions from the Pearson's system.

Contrasting of techniques

There is a large variability in terms of computational complexity and amount of user interaction required by the various image segmentation techniques. Despite their limitations, visual delineation performed by experts is still the most widely used technique [58]. However, manual techniques are labour intensive and suffer from intra-observer variability whereas thresholding techniques are simple to put into practice although scanner-specific calibration might be required for implementation of the adaptive thresholding method. The high computational burden associated with supervised methods which require time-consuming training is also worth emphasizing. In a clinical setting, the balance between algorithmic complexity and the validity of results obtained is an important criterion when selecting a PET image segmentation technique. Even though the techniques which make a large number of assumptions can frequently be undemanding, it is not necessarily true that complex algorithms will always perform better. The extra complexity must be used judiciously and justified for the particular application at hand. In many cases, extra complexity can just as easily result in unreliability as in improved results.

Although thresholding methods are the most frequently used techniques, possibly this is driven by their simplicity to implement and their efficiency for clinical practice. However, application of the proper threshold technique could be a challenging task due to limited resolution of PET images, blurriness due to partial volume effect or motion artefacts and limited statistical noise fluctuations due to limited photon count. In addition, a threshold technique does not take into account variations in tumour heterogeneity or possible overlap in intensity between the tumour and surrounding normal tissues resulting in under- or over-estimation of the tumour extent depending on the selected threshold value. The comparison study conducted by Nestle et al. indicated the presence of up to 41% variability between the visualization definition of BTV and four different thresholding methods that were attributed to maximum SUV, tumour homogeneity and lesion size [94]. The results of this study and similar ones have motivated the investigation of other segmentation techniques for defining BTVs in PET images for radiotherapy treatment planning purposes. However, such methods need to be carefully applied, otherwise inferior performances could be obtained as revealed in [102], where the authors have noticed that the direct application of the Sobel edge detector or the WT have failed to correctly identify the correct size of experimental volumes compared to thresholding segmentation.

To overcome these difficulties, several authors have explored active deformable models, learning methods and stochastic approaches, which seemed to provide better

results when compared with threshold segmentation methods. For instance, active deformable models allow the incorporation of smoothness and other prior information into the segmentation algorithm, which would make it more robust to boundary gaps in addition to enjoying inherent subpixel accuracy due to its continuum nature. However, variational methods, especially gradient-based methods, suffer from sensitivity to image noise [139]. Therefore, proper pre-processing techniques such as denoising or deblurring could be required. Another approach is to replace gradient-based methods by region-based approaches such as the Mumford-Shah model to improve robustness to noise [140]. Moreover, most deformable models may require the selection of proper parameters to achieve satisfactory results in clinical practice.

Learning methods based on classification require training of the method on data with known labels (ground truth known). However, this could be a challenging task due to variability of PET tracer uptakes and biodistribution depending on the biomarker concentration in the blood (e.g. glucose concentration for FDG), tracer dose, interval between tracer injection and scanning and scanning time. In addition, the PET images need to be labelled properly to identify the ground truth for training purposes (e.g. the structures contoured by a panel of experienced radiologists). Therefore, application of these methods would require extraction of robust features to these effects in addition to a ground truth. The issue of ground truth is partially resolved using unsupervised methods such as clustering techniques; however, these methods may still need incorporation of prior knowledge to estimate the right number of clusters and their connectiveness. Issues related to spatial correlation are mitigated through incorporation of additional neighbourhood constraints.

Stochastic models allow the incorporation of voxel's intensity uncertainty directly into the model; however, these models require the use of a proper noise model. In many cases, the Gaussian assumption is used because it simplifies the computational burden and speeds up convergence. Spatial constraints could be incorporated by using the Markov property; however, these models may require additional constraints to achieve the smoothness observed in the deformable models.

The introduction of integrated PET/CT scanners has provided new opportunities to develop new methods for defining BTV in radiotherapy treatment planning by combining coregistered physiological information from PET images with anatomical information from CT images [1]. Bradley et al. have shown that a significant reduction in observer variability could be achieved by combining PET and CT for lung cancer target definition [42]. This was also shown in several HNC studies [29, 141]. The feasibility of this approach has been also demonstrated for different cancer sites [142]. Hence, it could be conjectured that the combination of PET with CT will provide an improvement over CT alone for targeting in future radiation therapy treatment planning. Consequently, a framework based on the generalized level set approach and a *soft*-AND model for concurrent segmentation of multimodality images (e.g. PET/CT) has recently been developed. Measurement of the overlap index (Dice similarity metric), discussed below, showed significant improvement in target definition by using combined imaging information [143].

Validation and comparison of PET image segmentation techniques

Validation of accuracy (fidelity to the truth) and precision (reproducibility) are very crucial steps for any clinical use of a computer algorithm [144]. However, a rather challenging, even a problematic issue for validation of segmentation algorithms, is the identification of a gold standard (i.e. the benchmark) [25, 144]. Segmentation methods yield binary classification results (a voxel belongs to the object or does not). There are basically four different strategies allowing the assessment of the accuracy of PET image segmentation techniques. Table 2 summarizes these strategies and provides a concise summary of their advantages, drawbacks and limitations. These include manual segmentation by experts in the field, the use of simulated or experimental phantom studies where the ground truth (tumour volume) is known a priori, the comparison with correlated anatomical GTVs defined on CT or MRI and the comparison of tumour volumes delineated on clinical PET data with actual tumour volumes measured on the macroscopic specimen derived from histology, in case a PET scan was undertaken before surgery. It should be emphasized that such correlative analysis relies on a high degree of registration accuracy between multimodality images which is still challenging to perform in a clinical setting [145].

Contouring by experts has often been labelled as the "truth". However, such an approach suffers from intra- and inter-observer variability [58, 146, 147] besides being time-consuming and exhaustive to carry out. In the second approach, one would rely on constructing phantoms, either physically or digitally. Evaluation and validation of quantitative analysis procedures using simulated or experimental phantom studies is very popular in medical imaging physics research. Physical phantoms are typically constructed for a specific standardization task (e.g. NEMA phantoms). Such physical phantoms have been widely used for testing and evaluating PET segmentation algorithms with static [80, 148] or dynamic constructions [149, 150]. For testing PET image segmentation methods, a realistic anthropomorphic thorax phantom (Radiology Support

Table 2 Outlined strategies for validation and comparison of PET image segmentation techniques and their main advantages/disadvantages

Category	Key advantages	Limitations/drawbacks	Representative references
Manual contouring by experts	Current gold standard	Time-consuming, subjective, large intra- and inter-observer variability	[53–58, 212, 216]
Simulation studies	Controlled study with known conditions and segmentation ground truth, flexibility in adjusting experimental parameters	Simulation conditions may not be representative of complex real-life PET/CT imaging conditions and varying patient anatomy and physiology	[83, 89, 123]
Experimental phantom studies	Better representation of scanner imaging and physical conditions (e.g. PVE) than could be simulated numerically with known ground truth	Human anatomy is too complex and cannot be represented by simple materials or geometrical shapes, difficulty in designing dynamic phantoms that mimic respiratory motion	[51, 80, 83, 85, 90, 91, 95–97, 133, 148, 217–221]
Clinical studies (without macroscopic specimen)	Real-life clinical images of patient disease are used	Ground truth unknown	[38, 48, 82, 87, 88, 98, 142, 182, 183, 207, 222–229]
Clinical studies (with macroscopic specimen)	Real-life clinical images of scanner capability and patient disease are produced with an independent ground truth	Need to carefully account for specimen shrinkage artefacts in vitro in order to match the in vivo conditions	[50, 57, 103, 123, 170–178, 186]

Devices Inc., Long Beach, CA, USA) that is depicted in Fig. 4 could be used. This phantom was thoroughly used by Montgomery et al. [133] for assessment and characterization of their novel segmentation algorithm. This phantom consists of lungs, cardiac insert and a liver compartment that can be filled independently. In addition, there is a plastic vial placed in the pelvis to simulate the bladder and thus take into account out of field of view activity. Techniques similar to those suggested by Turkington et al. [151] were used to fabricate a set of tumours of different size that can be used to produce wall-less small radioactive lesions using moulted wax. This design has many advantages compared to using commercial plastic or glass spheres that have an inactive wall as they do not reflect the real situation in the patient and lead to quantification errors in the presence of background activity [152]. The lesions can be mixed with ^{18}F activity and embedded in the various compartments/organs provided (e.g. left and right lungs to mimic lung cancer studies) within the physical anthropomorphic thorax phantom. This arrangement provides properties similar to the human thorax suitable for mimicking whole-body oncological FDG PET studies.

Fig. 4 Photograph of the anthropomorphic thorax phantom (**a**), rendered CT image of the phantom (**b**) and wall-less small radioactive lesions of different shape and size fabricated using moulted wax beeswax tumours (**c**) that can be inserted into the lungs of the phantom to simulate typical lesions in lung cancer patients
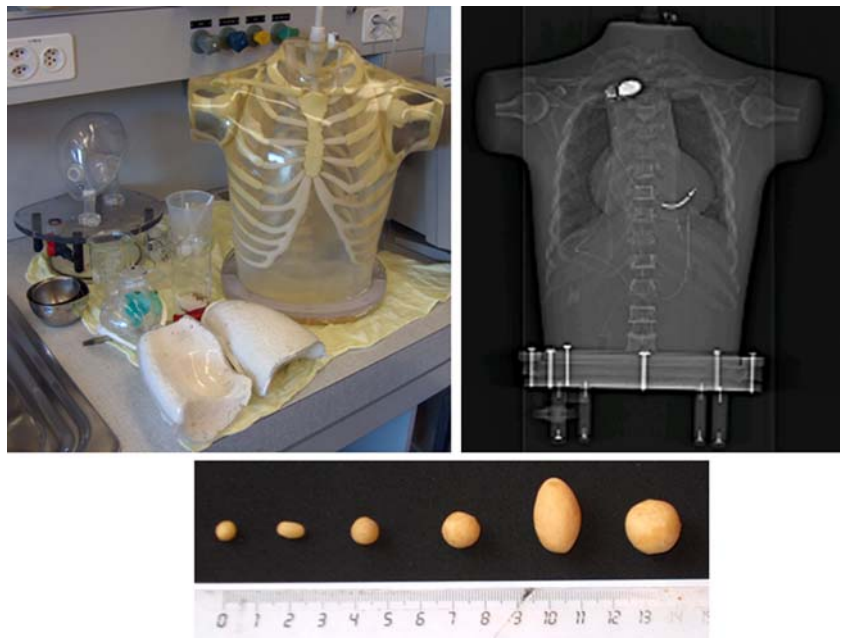
Simple and more sophisticated computational anthropomorphic anatomical and physiological models are also widely used in the literature [153, 154]. Among the most popular models for functional imaging is the 4D nonuniform rational B-splines (NURBS)-based cardiac-torso (NCAT)—now known as the 4D extended cardiac-torso (XCAT)—phantom [155]. This computational model was developed to realistically model human anatomy based on the "Visible Human" data. Dynamic cardiac and respiratory models are incorporated into the phantom to mimic real heart and breathing motion. The organ shapes are formed using NURBS [156] and now include more realistic modelling of normal and abnormal variations in anatomy and in the respiratory motion [157]. Typical strategies used in simulation-based generation of realistic oncological PET data to assess the performance of image segmentation techniques showing a simulated tumour in the right lung at end of exhalation with a hypothetical average SUV=6.5, a typical average SUV in NSCLC, are shown in Fig. 5. Accurate simulation techniques combining Monte Carlo modelling techniques and clinical data are now becoming available and being exploited for comprehensive assessment of segmentation algorithms [158–160]. Modelling combined PET/CT systems using dedicated PET [161–163] and X-ray CT [164–166] simulators as performed by Ay and Zaidi [167] might offer some additional advantages, particularly when respiratory motion is addressed. Recent developments describing more realistic dynamic 4D anatomical models incorporating accurate modelling of respiratory and cardiac motions might help to achieve this goal [153, 154].

Notwithstanding the advantages offered by phantom studies, some solutions to the problem of PET segmentation are less suitable for routine clinical applications than they are in phantom simulations. The accuracy reached in phantom studies is unlikely to be reached in clinical investigations. The true clinical feasibility of the methods described in the previous section has yet to be fully investigated. The results reported in the literature about the accuracy of PET segmentation very much depend on the type of phantoms, lesion size, noise, type of scanner, etc.

The application of PET image segmentation in the clinical setting requires a rigorous assessment in order to
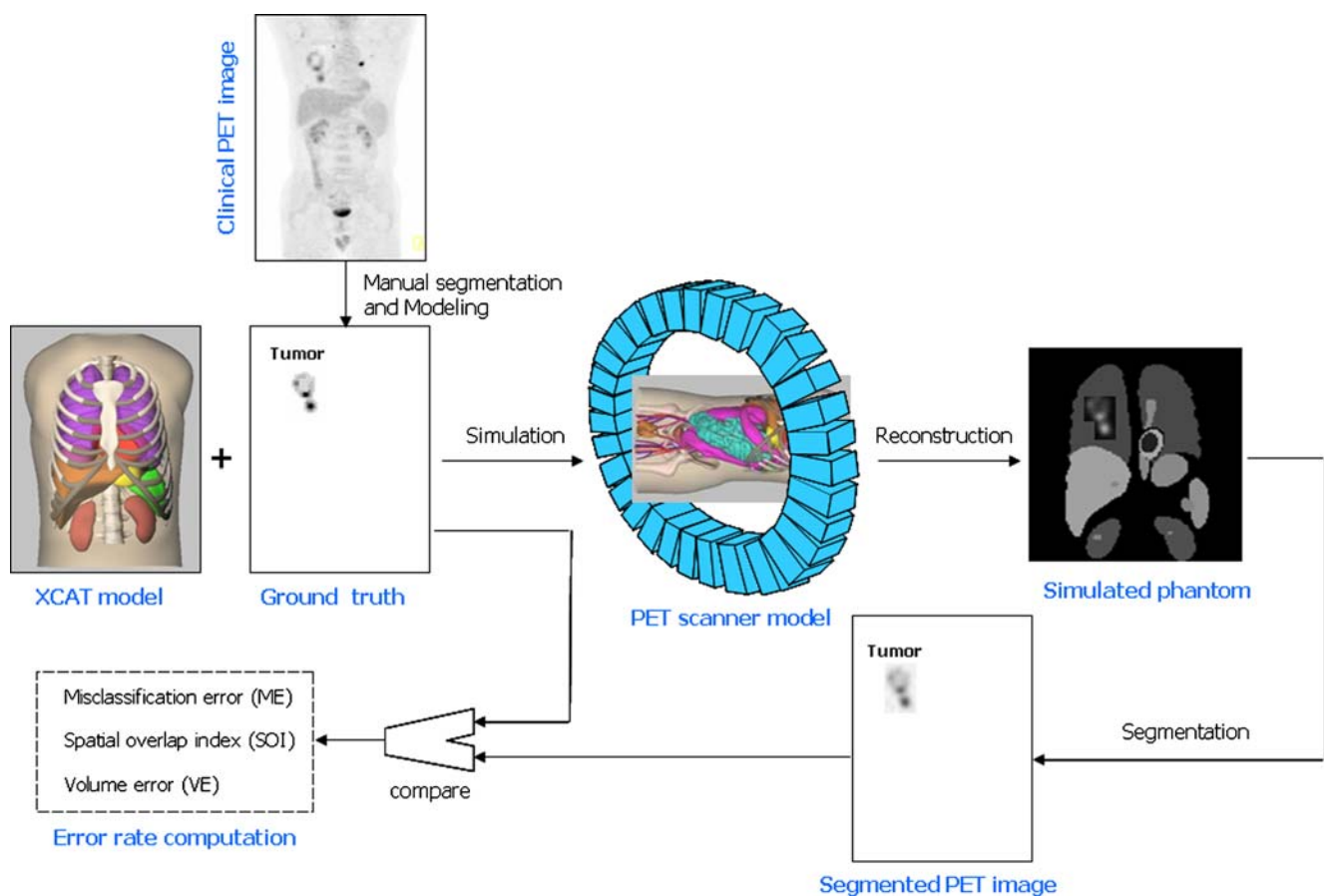

Fig. 5 Principle of simulation-based validation of PET image segmentation techniques showing the 4D NCAT anatomical model representing typical tracer distribution in the body to which a lesion of known characteristics was added. The 3D rendered simulated FDG PET image with a lesion in the right lung at end of exhalation with an SUV=6.5 is also shown. For lung cancer studies, motion could be simulated using a respiratory cycle of 5 s and diaphragm extent of 2 cm

define the most appropriate algorithm for a specific combination of a probe and targeted application and give an estimate of expected accuracy. There is no single figure of merit that summarizes algorithm performance, since performance ultimately depends on the diagnostic task being performed. Well-established figures of merit known to have a large influence on many types of task performance are generally used to assess the performance of image segmentation procedures [168, 169].

It has been argued that when using clinical data where the ground truth is not known *a priori*, it is generally unacceptable to use an imaging modality as gold standard against which results from another imaging modality are compared. However, comparisons of GTVs delineated on structural imaging (CT or MRI) with BTVs obtained using various strategies for functional PET-based image segmentation techniques might provide some useful hints. One such example is shown in Fig. 6 which illustrates a patient study presented with a glioblastoma exemplifying differences in target volume definition obtained by using various segmentation methods as compared to GTV defined manually on MRI [88]. PET image segmentation techniques included manual delineation of contours, a 2.5 SUV cut-off, a fixed threshold of 40 and 50% of the maximum signal intensity, SBR-based adaptive thresholding (BTV$_{SBR}$), gradient find (BTV$_{GF}$) and RG (BTV$_{RG}$). For this particular patient, all PET-based techniques define almost similar contours on this plane, but overall the estimated lesion volume is different.

Yet another attractive approach would be to use registered histological specimens [170]. Only very few studies reported on the use of macroscopic surgical specimens for validation of PET-based BTV delineation techniques [50, 57, 103, 171–177]. The only studies where the 3D volumetric macroscopic specimen (in contrast to lesion volume or size defined on the one to three major axes) was available clearly demonstrated the superiority of PET compared to other structural modalities for pharyngolaryngeal squamous cell carcinoma [171]. More studies are still required before claiming superiority of PET for other tumour types and localizations. However, one needs to be cautious about ex vivo tumour shrinkage in such cases where timing between imaging and surgery play a significant role. In the study by Daisne et al., a special procedure was developed to allow a three-dimensional coregistration of the macroscopic specimen with the imaging modalities [171]. Fresh surgical specimens were placed in a polystyrene cast containing three longitudinally placed wood rods that were equally spaced in the transverse plane of the specimen and the cast was filled with a 16% gelatin solution and kept at −20°C for 48 h and thereafter at −80°C for at least 72 h. The authors suggest that such fixation and freezing procedures resulted in no retraction compared to other methods as evident by their animal data. Radiology-pathology correlation is even more challenging in the lung. A limited number of NSCLC patients (5) were investigated in a feasibility study that highlighted that ignoring deformations of the lung might result in underestimation of the microscopic spread [50]. Dahele et al. developed an original framework allowing the 3D reconstruction of resected NSCLC specimens [178]. The procedure incorporates a number of distinctive features, including agar embedding of the specimen and whole-mount digital histopathology.

Fig. 6 Gadolinium-enhanced T1-weighted MRI (**a**), corresponding ^{18}F-FET PET (**b**) and fused PET/MR (**c**) transaxial slices of a clinical study with a glioblastoma showing differences in target volume definition. Indicated are **d** the GTV delineated on MRI (GTV$_{MRI}$) and **e** enhanced details of PET-based BTVs obtained by manual delineation of contours (BTV$_{man}$; *magenta*), an isocontour of a SUV of 2.5 (BTV$_{2.5}$; *purple*), a fixed threshold of 40% (BTV$_{40\%}$; *green*) and 50% (BTV$_{50\%}$; *cyan*) of the maximum signal intensity, SBR-based adaptive thresholding (BTV$_{SBR}$; *yellow*), gradient find (BTV$_{GF}$; *blue*) and RG (BTV$_{RG}$; *red*) segmentation algorithms. Note that GTV$_{MRI}$ overestimates the tumour extension relative to BTV$_{man}$. Reprinted with permission from [88]
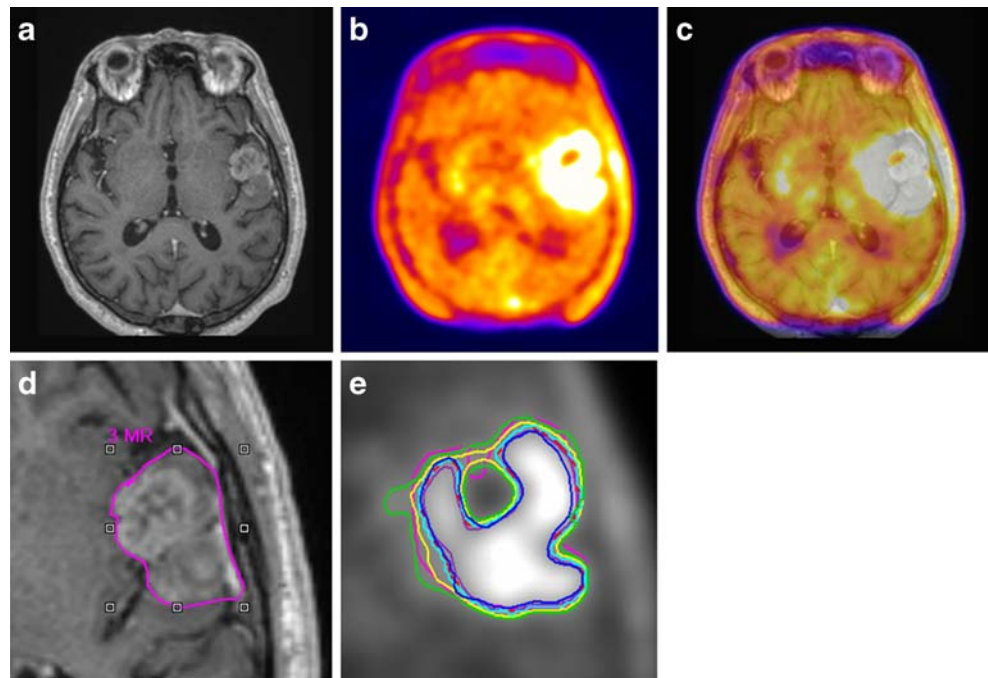
The inherent difficulty of the procedure related to shrinkage of the specimen outside after surgical excision restricted the use of this methodology in clinical research. As a result, only few studies were published so far covering pharyngolaryngeal squamous cell carcinoma [171], NSCLC [50] and oropharyngeal and oral cavity cancer [175]. Such deformations need to be accounted for properly before use in comparison studies [178]. Moreover, some preclinical studies reported discrepancies between PET images and the underlying microscopic reality derived from autoradiography owing to the finite resolution of PET scanners [179].

A limited number of studies reported the comparative evaluation of PET image segmentation techniques using clinical data [57, 86–88, 94, 102, 148, 171, 174, 176, 177, 180–184]. The main difference between the segmentation methods is the underlying empirical derivation or mathematical model used to differentiate between the tumour and background. Comparison studies underlined the importance of the PET segmentation technique without definitively suggesting a particular procedure for clinical use [87, 88]; however, it has clearly been shown that PET segmentation using the adaptive thresholding technique outperformed structural imaging procedures (CT and MRI) for GTV delineation where macroscopic surgical specimen derived from histology served as gold standard for assessment [171]. Thus, it can be argued that at present PET image segmentation is still an open research area that requires further research and development efforts where users are applying the most reliable technique available to them, although ultimately it may become clear which method is best for a particular application.

To provide better insight into the performance of the different categories of PET image segmentation techniques described in the preceding sections, we show qualitative and quantitative results using clinical data. It should be noted that none of the methods proposed so far is well enough developed or understood to allow it to be used clinically without a careful knowledgeable check. A recent study compared nine segmentation techniques [148]. Representative segmentation results of an FDG PET/CT image of a patient presenting with histologically proven NSCLC are shown in Fig. 7. The GTV/BTVs defined on the irregularly shaped non-homogeneous lesion using nine segmentation techniques are depicted on both the CT and FDG PET transaxial slices. These include manual delineation performed by an experienced radiation oncologist on both the CT (CT_{man}) and PET (PET_{man}) images, four semi-automated methods comprising the SBR-based adaptive thresholding technique [91], RG [28], Black et al. technique [90], Nestle et al. technique [94] and three fully automated methods: standard FCM [120], the FCM-S which incorporates nonlinear anisotropic diffusion filtering thus allowing the integration of spatial contextual information and the FCM-

SW algorithm which also considers inhomogeneity of tracer uptake through the use of the à trou wavelet transform [123]. Table 3 summarizes the results obtained using the various PET image segmentation techniques for the clinical study shown in Fig. 7 (maximum diameter determined by histology is 7 cm). The overlap fraction used to assess geographic mismatch between the BTVs delineated using the different techniques with respect to the BTV_{SBR} is also shown. The overlap volume of BTV_X (X refers to one of the above-referenced segmentation techniques) and BTV_{SBR} is expressed as the overlap volume of BTV_X and BTV_{SBR} relative to BTV_{SBR}.

For extraction of quantitative metrics of segmentation accuracy, besides the comparison of spatially independent volume changes that is commonly practiced, several spatially dependent statistical metrics could be used. The following metrics recommended for general segmentation analysis [185] could be further adopted to better assess the accuracy of the different PET segmentation methods with respect to a predefined benchmark.

The two most common techniques currently used to assess the performance of image segmentation techniques are linear regression and Bland-Altman analysis. The first consists in performing paired correlation between two segmentation algorithms by computing the slope, intercept and the correlation coefficient for the regression plot using lesion volume [103, 175] or maximum lesion diameter along one of the three axes as a figure of merit [57, 148, 186]. A large correlation coefficient (which actually reflects the spread of the data points around the regression line) is always associated with strong agreement. On the other hand, Bland-Altman analysis attempts to quantify the level of agreement between the estimates resulting from the use of two algorithms using a statistical method consisting in depicting the difference between the estimates against the average of the estimates [187]. The two estimation methods are assumed to "agree" if the differences lie within the 95% confidence interval (limits of agreement) defined as mean ± 1.96×SD In such case, one technique could, in theory, replace another with confidence for estimation of the same parameter. Despite their popularity, such approaches have many shortcomings and caution should be observed for interpretation of the results.

Receiver-operating characteristics (ROC) curve This is a plot of the sensitivity (true classification fraction) versus specificity (false-positive fraction) for a continuum of threshold values. The overall accuracy is summarized by the area under the ROC curve (AUC). An extension of this metric called localization ROC (LROC) is used to resolve the spatial localization issue [188].

The ROC technique could be used as a metric to evaluate the performance of a "continuous classifier" and
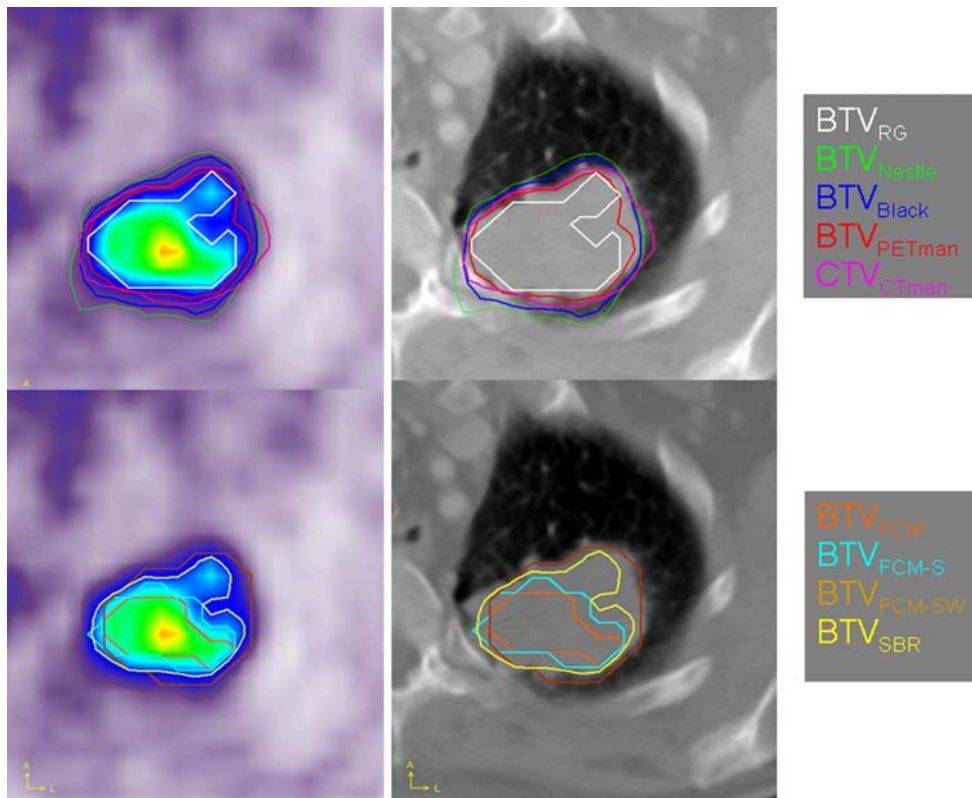
Fig. 7 Representative segmentation results of FDG PET/CT image of a patient presenting with histologically proven NSCLC. The GTVs defined on the non-homogeneous lesion with irregular shape (maximum diameter determined by histology is 7 cm) using nine segmentation techniques are depicted on both the CT (*left*) and FDG PET (*right*) transaxial slices. These include manual delineation performed by an experienced radiation oncologist on both the CT (CT_{man}) and PET (PET_{man}) images, four semi-automated methods comprising the SBR-based adaptive thresholding technique [91], RG [28], Black et al. technique [90], Nestle et al. technique [94] and three fully automated methods: standard FCM [120], the FCM-S which incorporates nonlinear anisotropic diffusion filtering to incorporate information about the spatial context and the FCM-SW algorithm which also considers inhomogeneity of tracer uptake through the use of the à trou wavelet transform [123]

Table 3 Summary of comparative PET image segmentation results of the clinical study shown in Fig. 7 (maximum diameter determined by histology is 7 cm) showing the maximum diameter and volume of the lesion estimated by each segmentation technique. The overlap fraction used to assess geographic mismatch between the GTV/BTVs delineated using the different techniques with respect to the BTV delineated using the SBR technique is also shown. Note the large differences between the estimated volumes despite the small differences between the maximum diameters measured by the various segmentation techniques

Segmentation method	Max. diameter (cm)	Volume (cc)	Overlap fraction w.r.t. SBR
CT_{man}	6.5	98	0.60
PET_{man}	7.5	80	0.70
RG	6	61	0.95
Nestle et al.	7	138	0.47
Black et al.	6.5	117	0.55
SBR	6.5	65	-
FCM	4.8	32	0.49
FCM-S	5.5	45	0.66
FCM-SW	6.5	76	0.87

is a popular visual method for assessing the overall classification accuracy. Numerous ROC techniques were developed including non-parametric, semi-parametric and parametric transformation models for estimating and comparing ROC curves derived from continuous data as described in [185, 189]. A significant advance in ROC analysis in the absence of a gold standard (i.e. without knowing the ground truth) was reported by Henkelman et al. who demonstrated that ROC parameters could be assessed by using additional tests on the same patients [190]. The statistical basis of this approach was verified and its relevance further confirmed in subsequent studies [191]. More recently, a more sophisticated technique using the maximum likelihood approach allowing the assessment and comparison of various estimation methods without the use of a gold standard was suggested [192, 193]. The method estimates the relative accuracy and consistency of the assessed algorithms without a priori knowledge of the gold standard approach. Simulated studies seem to suggest that the technique is capable of estimating accurately the parameters of a regression line without a gold standard

(i.e. without the x-axis). The technique seems appropriate for evaluation of image segmentation techniques, particularly that the approach was successful in estimating the volume of a physical phantom using two imaging modalities (SPECT and CT) [194].

Spatial overlap index Typically, the Dice similarity coefficient (DSC) is used, which is defined in terms of pixel ratio of the overlapping regions, where at any given threshold DSC values would range from 0, indicating no spatial overlap between two sets of binary segmentation results, to 1 indicating complete overlap [169]:

$$DSC = \frac{2(A \cap B)}{(A + B)} \qquad (18)$$

The multiclass type I (*TI*) and type II (*TII*) errors were also suggested to quantify the error based on mis-segmented voxels [168, 195]. The percentages of the two error types are given by:

$$TI = 100 \times \frac{Number\ of\ voxels\ of\ class\ k\ not\ classified\ as\ k}{Total\ number\ of\ voxels\ of\ class\ k} \qquad (19)$$

$$TII = 100 \times \frac{Number\ of\ voxels\ of\ other\ classes\ than\ k\ classified\ as\ k}{Total\ number\ of\ voxels\ of\ other\ classes\ than\ k} \qquad (20)$$

Mutual information This is a measure of uncertainty (entropy), which is routinely used with registration algorithms [196]; however, the metric is applicable to segmentation analysis as well [197]. Entropy measures, and particularly mutual information, have found widespread use in nuclear medicine applications. There are many variants of mutual information and joint entropy measures; however, they do not seem to be widely used for this purpose. In a recent assessment of voxel similarity measures, Holden et al. have demonstrated clear preference for those based on entropy measures including mutual information [198]. A normalized version of this metric is recommended, which is given by:

$$NMI = \frac{H(A) + H(B)}{H(A, B)} \qquad (21)$$

where $H(\cdot)$ is the Shannon entropy.

Challenges and future directions

An overview of current state-of-the art developments in PET image segmentation techniques is provided in this review. Approximate methods suitable for clinical routine applications and more sophisticated approaches for research applications, where there is greater emphasis on accurate quantitative analysis, have been proposed. It is gratifying to see in overview the progress that PET image segmentation has made in the last decade following the successful introduction of PET/CT in the radiation therapy treatment planning process, from operator-dependent manual delineation of structures, through simple thresholding, the use of classifiers and fuzzy clustering, and more recently active contour models, dynamic PET data analysis and atlas-guided approaches incorporating prior information [24]. Recent developments have been enormous particularly in the last 5 years, the main opportunities striving towards improving the accuracy, precision and computational speed through efficient implementation in conjunction with decreasing the amount of operator interaction. The application of PET image segmentation is well established in research environments and is still limited in clinical settings to academic institutions with advanced physics and technical support. As the above-mentioned challenges are met, and experience is gained, implementation of validated techniques in commercial software packages will be useful to attract the interest of the clinical community and increase the popularity of these tools. It is expected that with the availability of computing power, more complex and ambitious computer intensive segmentation algorithms will become clinically feasible.

Despite the remarkable progress that image segmentation has made during the last few years, performance validation in a clinical setting remains the most challenging issue [144]. Notwithstanding the outstanding growth achieved as can be witnessed by the enormous number of publications in the field, many challenging issues still remain to be solved through research. There is no shortage of challenges and opportunities for PET image segmentation today. Tumour heterogeneity and stability of tracer uptake is one of the challenges facing automated delineation of BTVs [199]. Very few studies have addressed the issue of tumour heterogeneity which cannot be appropriately handled using simple thresholding techniques [89, 123]. It should be noted that most published studies reporting on the use of image segmentation in the context of PET-guided radiation therapy focused on FDG as tracer whereas very few investigations paid attention to other probes. Evolving areas for application of other promising radiotracers for this purpose [200–203] include but are not limited to markers of tumour proliferation (e.g. ^{18}F-fluorothymidine, FLT), amino acid metabolism (^{11}C-methionine, ^{11}C-tyrosine and ^{18}F-fluoroethyltyrosine, FET), cell membranes and fatty acid metabolism (^{11}C-acetate or ^{11}C-choline and ^{18}F-choline), somatostatin receptors (^{68}Ga-DOTA-TOC) and hypoxia (^{18}F-FMISO, ^{18}F-FAZA, ^{64}Cu-ATSM and

[18F]-EF5) which have already shown their potential in patient management or dose painting in radiation therapy [204, 205]. In addition, novel tracers have been developed that specifically bind to certain intra- or extracellular compounds of various tumours, such as [18F]-DOPA (metabolism of amine precursor uptake and decarboxylase, APUD) which is now widely adopted as a diagnostic tool for the imaging of neuroendocrine tumours [206]. Published reports so far involved the use of tracers such as [11C]-methionine [53, 180, 207], [18F]-choline [208–210], [18F]-FLT [211] and [18F]-FET [88, 212]. In addition, while the role of hypoxia markers was demonstrated in dose painting techniques [205], its relevance is not well established in target volume definition. Experiments performed on animal models have shown the limitations of PET-guided dose painting techniques which should be considered with caution given the finite spatial resolution of current clinical PET scanners [179]. It would be interesting to explore optimal strategies aiming at tailoring image segmentation techniques to match the specificity of each particular cancer imaging probe and tumour type [60]. Despite the worthwhile research carried out and the remarkable achievements, the field is still in its infancy and plenty of research opportunities still exist particularly in connection with the design of experimental clinical protocols and tracers or combinations of tracers to be used for each application, which remain open research questions.

PET image segmentation methods developed so far were rather general and as such were not optimized to address a specific problem or type of tumour and/or localization. The availability of integrated PET/CT imaging systems provides new opportunities as well as new challenges in radiotherapy treatment planning and delivery; they have also presented many technical challenges on how to integrate different modalities information properly affecting their visualization and delineation. One of the challenges faced by scanner manufacturers is to provide commercial platforms that can easily handle multimodality images from any DICOM-compatible imaging modality including hybrid PET/CT and incorporate this information with minimum effort into the radiation therapy treatment planning software. There is still a lack of commercial tools available in treatment planning systems allowing importation and realignment of functional PET images and display of SUV units for user-defined automatic thresholding. Some investigators have argued for a potential role of PET/MRI to replace PET/CT as a platform of choice [213]. Besides the soft tissue discrimination, blood flow measures and indirect biochemical content could be incorporated [214]. Currently MRI is still in its infancy for radiotherapy application. However, its role is more in relation to better soft tissue discrimination compared to CT. In addition, functional MRI (fMRI) and its ability to compete with PET in radiation oncology as a delineation tool has not been well demonstrated yet.

PET will undoubtedly continue to play an important role in the management of patients with cancer [215]. In this regard, it should be emphasized that many different computational approaches and algorithmic design paths have been and continue to be pursued in both academic and corporate settings, which offer different compromises in terms of performance and versatility but in most cases improve the clinical workflow efficiency. It is still uncertain which algorithms will be the most successful to deserve incorporation into future software supplied by either medical imaging or radiation therapy treatment planning vendors, but it is certain that technical advances will continue and will enable more accurate quantification of molecular targets using PET. In the clinical setting, it has become standard practice to use simplified techniques compared to the often complex methods developed for research where there is greater emphasis on quantitative accuracy. As a general guideline, fixed thresholding techniques should not be used as adaptive thresholding techniques which take into account the background could be easily implemented in the clinic following careful implementation of the experimental calibration procedure required to derive imaging protocol- and scanner-specific parameters. For those departments with limited physics support, manual delineation techniques are to be preferred to fixed thresholding approaches. Fully automated PET image segmentation techniques remain the best option and will likely find their way to the clinic in the near future.

References

1. Townsend DW. Multimodality imaging of structure and function. Phys Med Biol 2008;53:R1–39.
2. Hasegawa B, Zaidi H. Dual-modality imaging: more than the sum of its components. In: Zaidi H, editor. Quantitative analysis in nuclear medicine imaging. New York: Springer; 2006. p. 35–81.
3. Bernier J, Hall EJ, Giaccia A. Radiation oncology: a century of achievements. Nat Rev Cancer 2004;4:737–47.
4. Fenwick JD, Tomé WA, Soisson ET, Mehta MP, Rock Mackie T. Tomotherapy and other innovative IMRT delivery systems. Semin Radiat Oncol 2006;16:199–208.
5. Ling C, Zhang P, Archambault Y, Bocanek J, Tang G, Losasso T. Commissioning and quality assurance of RapidArc radiotherapy delivery system. Int J Radiat Oncol Biol Phys 2008;72:575–81.
6. Jäkel O, Karger CP, Debus J. The future of heavy ion radiotherapy. Med Phys 2008;35:5653–63.
7. ICRU. Prescribing, recording and reporting photon beam therapy. ICRU Report 62. Washington: International Commission on Radiation Units and Measurements; 1999.

8. Austin-Seymour M, Chen GT, Rosenman J, Michalski J, Lindsley K, Goitein M. Tumor and target delineation: current research and future challenges. Int J Radiat Oncol Biol Phys 1995;33:1041–52.

9. Evans PM. Anatomical imaging for radiotherapy. Phys Med Biol 2008;53:R151–91.

10. Papiez L, Langer M. On probabilistically defined margins in radiation therapy. Phys Med Biol 2006;51:3921–39.

11. Khoo VS, Adams EJ, Saran F, Bedford JL, Perks JR, Warrington AP, et al. A comparison of clinical target volumes determined by CT and MRI for the radiotherapy planning of base of skull meningiomas. Int J Radiat Oncol Biol Phys 2000;46:1309–17.

12. Chaney E, Ibbott G, Hendee WR. Methods for image segmentation should be standardized and calibrated. Med Phys 2005;32:3507–10.

13. Ling C, Humm J, Larson S, Amols H, Fuks Z, Leibel S, et al. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. Int J Radiat Oncol Biol Phys 2000;47:551–60.

14. Zaidi H, Alavi A. Current trends in PET and combined (PET/CT and PET/MR) systems design. PET Clin 2007;2:109–23.

15. Chapman JD, Bradley JD, Eary JF, Haubner R, Larson SM, Michalski JM, et al. Molecular (functional) imaging for radiotherapy applications: an RTOG symposium. Int J Radiat Oncol Biol Phys 2003;55:294–301.

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

17. 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.

18. Lecchi M, Fossati P, Elisei F, Orecchia R, Lucignani G. Current concepts on imaging in radiotherapy. Eur J Nucl Med Mol Imaging 2008;35:821–37.

19. Mah D, Chen CC. Image guidance in radiation oncology treatment planning: the role of imaging technologies on the planning process. Semin Nucl Med 2008;38:114–8.

20. Messa C, Di Muzio N, Picchio M, Gilardi MC, Bettinardi V, Fazio F. PET/CT and radiotherapy. Q J Nucl Med Mol Imaging 2006;50:4–14.

21. Zaidi H, Vees H, Wissmeyer M. Molecular PET/CT imaging-guided radiation therapy treatment planning. Acad Radiol 2009;16:1108–33.

22. Olabarriaga SD, Smeulders AW. Interaction in the segmentation of medical images: a survey. Med Image Anal 2001;5:127–42.

23. Udupa JK, Saha PK. Fuzzy connectedness and image segmentation. Proc IEEE 2003;91:1649–69.

24. Boudraa A, Zaidi H. Image segmentation techniques in nuclear medicine imaging. In: Zaidi H, editor. Quantitative analysis of nuclear medicine images. New York: Springer; 2006. p. 308–57.

25. Zaidi H. Medical image segmentation: quo vadis. Comput Methods Programs Biomed 2006;84:63–7.

26. van Baardwijk A, Baumert BG, Bosmans G, van Kroonenburgh M, Stroobants S, Gregoire V, et al. The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning. Cancer Treat Rev 2006;32:245–60.

27. Greco C, Rosenzweig K, Cascini GL, Tamburrini O. Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC). Lung Cancer 2007;57:125–34.

28. Graves EE, Quon A, Loo Jr BW. RT_Image: an open-source tool for investigating PET in radiation oncology. Technol Cancer Res Treat 2007;6:111–21.

29. Ahn PH, Garg MK. Positron emission tomography/computed tomography for target delineation in head and neck cancers. Semin Nucl Med 2008;38:141–8.

30. Rahn AN, Baum RP, Adamietz IA, Adams S, Sengupta S, Mose S, et al. Value of 18F fluorodeoxyglucose positron emission tomography in radiotherapy planning of head-neck tumors. Strahlenther Onkol 1998;174:358–64. German.

31. Munley MT, Marks LB, Scarfone C, Sibley GS, Patz Jr EF, Turkington TG, et al. Multimodality nuclear medicine imaging in three-dimensional radiation treatment planning for lung cancer: challenges and prospects. Lung Cancer 1999;23:105–14.

32. Gross MW, Weber WA, Feldmann HJ, Bartenstein P, Schwaiger M, Molls M. The value of F-18-fluorodeoxyglucose PET for the 3-D radiation treatment planning of malignant gliomas. Int J Radiat Oncol Biol Phys 1998;41:989–95.

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

34. Scarfone C, Jaszczak RJ, Gilland DR, Greer KL, Munley MT, Marks LB, et al. Quantitative pulmonary single photon emission computed tomography for radiotherapy applications. Med Phys 1999;26:1579–88.

35. 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.

36. Vanuytsel LJ, Vansteenkiste JF, Stroobants SG, De Leyn PR, De Wever W, Verbeken EK, et al. The impact of 18F-fluoro-2-deoxy–glucose positron emission tomography (FDG-PET) lymph node staging on the radiation treatment volumes in patients with non-small cell lung cancer. Radiother Oncol 2000;55:317–24.

37. Levivier M, Wikler D, Goldman S, David P, Metens T, Massager N, et al. Integration of the metabolic data of positron emission tomography in the dosimetry planning of radiosurgery with the gamma knife: early experience with brain tumors. Technical note. J Neurosurg 2000;93 Suppl 3:233–8.

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

39. Paulino AC, Thorstad WL, Fox T. Role of fusion in radiotherapy treatment planning. Semin Nucl Med 2003;33:238–43.

40. Scarfone C, Lavely WC, Cmelak AJ, Delbeke D, Martin WH, Billheimer D, et al. Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. J Nucl Med 2004;45:543–52.

41. Yap JT, Carney JP, Hall NC, Townsend DW. Image-guided cancer therapy using PET/CT. Cancer J 2004;10:221–33.

42. Bradley JD, Perez CA, Dehdashti F, Siegel BA. Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. J Nucl Med 2004;45 Suppl 1:96S–101.

43. Brunetti J, Caggiano A, Rosenbluth B, Vialotti C. Technical aspects of positron emission tomography/computed tomography fusion planning. Semin Nucl Med 2008;38:129–36.

44. Pan T, Mawlawi O. PET/CT in radiation oncology. Med Phys 2008;35:4955–66.

45. Nestle U, Weber W, Hentschel M, Grosu A-L. Biological imaging in radiation therapy: role of positron emission tomography. Phys Med Biol 2009;54:R1–25.

46. Macapinlac HA. Clinical applications of positron emission tomography/computed tomography treatment planning. Semin Nucl Med 2008;38:137–40.

47. Czernin J, Allen-Auerbach M, Schelbert HR. Improvements in cancer staging with PET/CT: literature-based evidence as of September 2006. J Nucl Med 2007;48:78S–88.

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

49. Xing L, Siebers J, Keall P. Computational challenges for image-guided radiation therapy: framework and current research. Semin Radiat Oncol 2007;17:245–57.

50. Stroom J, Blaauwgeers H, van Baardwijk A, Boersma L, Lebesque J, Theuws J, et al. Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumors. Int J Radiat Oncol Biol Phys 2007;69:267–75.

51. 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.

52. Nestle U, Kremp S, Grosu AL. Practical integration of [18F]-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.

53. Grosu AL, Weber WA, Astner ST, Adam M, Krause BJ, Schwaiger M, et al. 11C-methionine PET improves the target volume delineation of meningiomas treated with stereotactic fractionated radiotherapy. Int J Radiat Oncol Biol Phys 2006;66:339–44.

54. Kalff V, Hicks RJ, MacManus MP, Binns DS, McKenzie AF, Ware RE, et al. Clinical impact of (18)F fluorodeoxyglucose positron emission tomography in patients with non-small-cell lung cancer: a prospective study. J Clin Oncol 2001;19:111–8.

55. Caldwell CB, Mah K, Ung YC, Danjoux CE, Balogh JM, Ganguli SN, et al. Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of 18FDG-hybrid PET fusion. Int J Radiat Oncol Biol Phys 2001;51:923–31.

56. Fox JL, Rengan R, O'Meara W, Yorke E, Erdi Y, Nehmeh S, et al. Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non-small-cell lung cancer? Int J Radiat Oncol Biol Phys 2005;62:70–5.

57. van Baardwijk A, Bosmans G, Boersma L, Buijsen J, Wanders S, Hochstenbag M, et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. Int J Radiat Oncol Biol Phys 2007;68:771–8.

58. Steenbakkers RJHM, Duppen JC, Fitton I, Deurloo KEI, Zijp LJ, Comans EFI, et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. Int J Radiat Oncol Biol Phys 2006;64:435–48.

59. Sovik A, Malinen E, Olsen DR. Strategies for biologic image-guided dose escalation: a review. Int J Radiat Oncol Biol Phys 2009;73:650–8.

60. Basu S. Selecting the optimal image segmentation strategy in the era of multitracer multimodality imaging: a critical step for image-guided radiation therapy. Eur J Nucl Med Mol Imaging 2009;36:180–1.

61. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. J Nucl Med 2007;48:932–45.

62. Rousset O, Rahmim A, Alavi A, Zaidi H. Partial volume correction strategies in PET. PET Clin 2007;2:235–49.

63. Rahmim A, Rousset O, Zaidi H. Strategies for motion tracking and correction in PET. PET Clin 2007;2:251–66.

64. Nehmeh SA, Erdi YE. Respiratory motion in positron emission tomography/computed tomography: a review. Semin Nucl Med 2008;38:167–76.

65. Li T, Thorndyke B, Schreibmann E, Yang Y, Xing L. Model-based image reconstruction for four-dimensional PET. Med Phys 2006;33:1288–98.

66. Qiao F, Pan T, Clark J, John W, Mawlawi O. Joint model of motion and anatomy for PET image reconstruction. Med Phys 2007;34:4626–39.

67. Lamare F, Ledesma Carbayo MJ, Cresson T, Kontaxakis G, Santos A, Cheze Le Rest C, et al. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. Phys Med Biol 2007;52:5187–204.

68. Rahmim A, Dinelle K, Cheng J-C, Shilov MA, Segars WP, Lidstone SC, et al. Accurate event-driven motion compensation in high-resolution PET incorporating scattered and random events. IEEE Trans Med Imaging 2008;27:1018–33.

69. Büther F, Dawood M, Stegger L, Wübbeling F, Schäfers M, Schober O, et al. List mode-driven cardiac and respiratory gating in PET. J Nucl Med 2009;50:674–81.

70. Rahmim A, Tang J, Zaidi H. Four-dimensional (4D) image reconstruction strategies in dynamic PET: beyond conventional independent frame reconstruction. Med Phys 2009;36:3654–70.

71. Perez CA. Principles and practice of radiation oncology. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2004.

72. Otsu N. A thresholding selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–6.

73. Reddi SS, Rudin SF, Keshavan HR. An optimal multiple threshold scheme for image segmentation. IEEE Trans Syst Man Cybern 1984;14:661–5.

74. Kittler J, Illingworth J. Minimum error thresholding. Pattern Recognit 1986;19:41–7.

75. Pal NR, Pal SK. A review on image segmentation techniques. Pattern Recognit 1993;26:1277–94.

76. Huang S-C. Anatomy of SUV. Nucl Med Biol 2000;27:643–6.

77. Keyes JW Jr. SUV: standard uptake value or silly useless value? J Nucl Med 1995;36:1836–9.

78. Basu S, Zaidi H, Houseni M, Udupa J, Acton P, Torigian D, et al. Novel quantitative techniques for assessing regional and global function and structure based on modern imaging modalities: implications for normal variation, aging and diseased states. Semin Nucl Med 2007;37:223–39.

79. Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med 2009;50:11S–20.

80. 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.

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

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

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

84. Zaidi H. Organ volume estimation using SPECT. IEEE Trans Nucl Sci 1996;43:2174–82.

85. Yaremko B, Riauka T, Robinson D, Murray B, Alexander A, McEwan A, et al. Thresholding in PET images of static and moving targets. Phys Med Biol 2005;50:5969–82.

86. Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume

in intensity-modulated radiotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2005;61:1385–92.

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

88. Vees H, Senthamizhchelvan S, Miralbell R, Weber D, Ratib O, Zaidi H. Assessment of various strategies for 18F-FET PET-guided delineation of target volumes in high-grade glioma patients. Eur J Nucl Med Mol Imaging 2009;36:182–93.

89. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging 2009;28:881–93.

90. 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.

91. Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, Grégoire 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.

92. Brambilla M, Matheoud R, Secco C, Loi G, Krengli M, Inglese E. Threshold segmentation for PET target volume delineation in radiation treatment planning: the role of target-to-background ratio and target size. Med Phys 2008;35:1207–13.

93. Drever L, Robinson DM, McEwan A, Roa W. A local contrast based approach to threshold segmentation for PET target volume delineation. Med Phys 2006;33:1583–94.

94. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rübe 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.

95. Schaefer A, Kremp S, Hellwig D, Rübe C, Kirsch C-M, Nestle U. 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. Eur J Nucl Med Mol Imaging 2008;35:1989–99.

96. Jentzen W, Freudenberg L, Eising EG, Heinze M, Brandau W, Bockisch A. Segmentation of PET volumes by iterative image thresholding. J Nucl Med 2007;48:108–14.

97. Nehmeh SA, El-Zeftawy H, Greco C, Schwartz J, Erdi YE, Kirov A, et al. An iterative technique to segment PET lesions using a Monte Carlo based mathematical model. Med Phys 2009;36:4803–9.

98. Greco C, Nehmeh SA, Schöder H, Gönen M, Raphael B, Stambuk HE, et al. Evaluation of different methods of 18F-FDG-PET target volume delineation in the radiotherapy of head and neck cancer. Am J Clin Oncol 2008;31:439–45.

99. Marr D, Hildreth E. Theory of edge detection. Proc R Soc Lond B Biol Sci 1980;207:187–217.

100. Huertas A, Medioni G. Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks. IEEE Trans Pattern Anal Mach Intell 1986;8:651–64.

101. Canny JF. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 1986;8:679–98.

102. Drever LA, Roa W, McEwan A, Robinson D. Comparison of three image segmentation techniques for target volume delineation in positron emission tomography. J Appl Clin Med Phys 2007;8:93–109.

103. Geets X, Lee J, Bol A, Lonneux M, Grégoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur J Nucl Med Mol Imaging 2007;34:1427–38.

104. El Naqa I, Bradley J, Deasy J, Biehl K, Laforest R, Low D. Improved analysis of PET images for radiation therapy. 14th

International Conference on the Use of Computers in Radiation Therapy. Seoul, Korea; 2004. pp 361–63.

105. Hsu C-Y, Liu C-Y, Chen C-M. Automatic segmentation of liver PET images. Comput Med Imaging Graph 2008;32:601–10.

106. Li H, Thorstad WL, Biehl KJ, Laforest R, Su Y, Shoghi KI, et al. A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours. Med Phys 2008;35:3711–21. Erratum. pp 5958.

107. Sethian JA. Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and material science. 2nd ed. Cambridge: Cambridge University Press; 1999.

108. Xu C, Pham DL, Prince JL. Image segmentation using deformable models. In: Sonka M, Fitzpatrick JM, editors. Handbook of medical imaging: medical image processing and analysis. Bellingham: SPIE Press; 2002. pp. 129–74.

109. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. Int J Comput Vis 1988;1:321–31.

110. Kass M, Witkin A, Terzopoulos. Snakes: active contour models. First International Conference on Computer Vision. London; 1987. pp. 259–68.

111. Liang J, McInerney T, Terzopoulos D. United snakes. Med Image Anal 2006;10:215–33.

112. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Trans Image Process 1998;7:359–69.

113. Osher S, Sethian JA. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. J Comput Phys 1988;79:12–49.

114. Duda RO, Hart PE, Stork DG. Pattern classification. 2nd ed. New York: Wiley; 2001.

115. Jain AK, Duin RPW, Mao J. Statistical pattern recognition: a review. IEEE Trans Pattern Anal Mach Intell 2000;22:4–37.

116. Clarke LP, Velthuizen RP, Phuphanich S, Schellenberg JD, Arrington JA, Silbiger M. MRI: stability of three supervised segmentation techniques. Magn Reson Imaging 1993;11:95–106.

117. Vaidyanathan M, Clarke LP, Velthuizen RP, Phuphanich S, Bensaid AM, Hall LO, et al. Comparison of supervised MRI segmentation methods for tumor volume determination during therapy. Magn Reson Imaging 1995;13:719–28.

118. Suri JS, Singh S, Reden L. Computer vision and pattern recognition techniques for 2-D and 3-D MR cerebral cortical segmentation (part I): a state-of-the-art review. Pattern Anal Appl 2002;5:46–76.

119. El Naqa I, Yang Y. Techniques in the detection of micro-calcification (MC) clusters in digital mammograms. In: Leondes T, editor. Medical imaging systems: technology and applications. Singapore: World Scientific Publishing Co. Pte. Ltd.; 2005. pp. 15–36.

120. Boudraa AE, Champier J, Cinotti L, Bordet JC, Lavenne F, Mallet JJ. Delineation and quantitation of brain lesions by fuzzy clustering in positron emission tomography. Comput Med Imaging Graph 1996;20:31–41.

121. Zhu W, Jiang T. Automation segmentation of PET image for brain tumors. IEEE Nucl Sci Symp Conf Rec 2003;4:2627–29.

122. Kim J, Wen L, Eberl S, Fulton R, Feng DD. Use of anatomical priors in the segmentation of PET lung tumor images. IEEE Nucl Sci Symp Conf Rec 2007;4:4242–45.

123. Belhassen S and Zaidi H. A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET. Med Phys 2010;37:1309–1324.

124. Zaidi H, Diaz-Gomez M, Boudraa AO, Slosman DO. Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. Phys Med Biol 2002;47:1143–60.

125. Acton PD, Pilowsky LS, Kung HF, Ell PJ. Automatic segmentation of dynamic neuroreceptor single-photon emission tomography images using fuzzy clustering. Eur J Nucl Med 1999;26:581–90.

126. Bezdek JC, Hall LO, Clark MC, Goldgof DB, Clarke LP. Medical image analysis with fuzzy models. Stat Methods Med Res 1997;6:191–214.

127. Jain AK, Murty MN, Flynn PJ. Data clustering: a review. ACM Comput Surv 1999;31:264–323.

128. De Luca A, Termini S. A definition of non-probabilistic entropy in the setting of fuzzy sets theory. Inform Control 1972;20:301–12.

129. Hall LO, Bensaid AM, Clarke LP, Velthuizen RP, Silbiger MS, Bezdek JC. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. IEEE Trans Neural Netw 1992;3:672–82.

130. Pham DL, Prince JL. An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities. Pattern Recognit Lett 1999;20:57–68.

131. Janssen MH, Aerts HJ, Ollers MC, Bosmans G, Lee JA, Buijsen J, et al. Tumor delineation based on time-activity curve differences assessed with dynamic fluorodeoxyglucose positron emission tomography-computed tomography in rectal cancer patients. Int J Radiat Oncol Biol Phys 2009;73:456–65.

132. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell 1990;12:629–39.

133. Montgomery D, Amira A, Zaidi H. Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. Med Phys 2007;34:722–36.

134. Aristophanous M, Penney BC, Martel MK, Pelizzari CA. A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. Med Phys 2007;34:4223–35.

135. Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based tissue classification of MR images of the brain. IEEE Trans Med Imaging 1999;18:897–908.

136. Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26:839–51.

137. Hatt M, Lamare F, Boussion N, Turzo A, Collet C, Salzenstein F, et al. Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET. Phys Med Biol 2007;52:3467–91.

138. Salzenstein F, Pieczynski W. Parameter estimation in hidden fuzzy Markovian fields and image segmentation. Graph Models Image Process 1997;59:205–20.

139. Long DT, King MA, Sheehan J. Comparative evaluation of image segmentation methods for volume quantitation in SPECT. Med Phys 1992;19:483–9.

140. Chan TF, Vese LA. Active contours without edges. IEEE Trans Image Process 2001;10:266–77.

141. Guido A, Fuccio L, Rombi B, Castellucci P, Cecconi A, Bunkheila F, et al. Combined (18)F-FDG-PET/CT imaging in radiotherapy target delineation for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2009;73:759–63.

142. 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.

143. El Naqa I, Yang D, Apte A, Khullar D, Mutic S, Zheng J, et al. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Med Phys 2007;34:4738–49.

144. Jannin P, Krupinski E, Warfield S. Validation in medical image processing. IEEE Trans Med Imaging 2006;25:1405–9.

145. Slomka P, Baum R. Multimodality image registration with software: state-of-the-art. Eur J Nucl Med Mol Imaging 2009;36:S44–55.

146. Fiorino C, Reni M, Bolognesi A, Cattaneo GM, Calandrino R. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. Radiother Oncol 1998;47:285–92.

147. Giraud P, Elles S, Helfre S, De Rycke Y, Servois V, Carette MF, et al. Conformal radiotherapy for lung cancer: different delineation of the gross tumor volume (GTV) by radiologists and radiation oncologists. Radiother Oncol 2002;62:27–36.

148. Belhassen S, Llina Fuentes CS, Dekker A, De Ruysscher D, Ratib O, Zaidi H. Comparative methods for 18F-FDG PET-based delineation of target volumes in non-small-cell lung cancer [abstract]. J Nucl Med 2009;50:27P.

149. Boucher L, Rodrigue S, Lecomte R, Bénard F. Respiratory gating for 3-dimensional PET of the thorax: feasibility and initial results. J Nucl Med 2004;45:214–9.

150. El Naqa I, Low DA, Bradley JD, Vicic M, Deasy JO. Deblurring of breathing motion artifacts in thoracic PET images by deconvolution methods. Med Phys 2006;33:3587–600.

151. Turkington TG, Degrado TR, Sampson WH. Small spheres for lesion detection phantoms. IEEE Nucl Sci Symp Conf Rec 2001;4:2234–37.

152. Bazañez-Borgert M, Bundschuh RA, Herz M, Martínez MJ, Schwaiger M, Ziegler SI. Radioactive spheres without inactive wall for lesion simulation in PET. Z Med Phys 2008;18:37–42.

153. Zaidi H, Xu XG. Computational anthropomorphic models of the human anatomy: the path to realistic Monte Carlo modeling in radiological sciences. Annu Rev Biomed Eng 2007;9:471–500.

154. Zaidi H, Tsui BMW. Review of computational anthropomorphic anatomical and physiological models. Proc IEEE 2009;97:1938–53.

155. Segars WP. Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom [PhD Thesis]: University of North Carolina, Chapel Hill, NC, USA; 2001.

156. Piegl L, Tiller W. The NURBS book. New York: Springer; 1997.

157. Segars WP, Tsui BMW. MCAT to XCAT: the evolution of 4D computerized phantoms for imaging research. Proc IEEE 2009;97:1954–68.

158. Aristophanous M, Penney BC, Pelizzari CA. The development and testing of a digital PET phantom for the evaluation of tumor volume segmentation techniques. Med Phys 2008;35:3331–42.

159. Tomei S, Reilhac A, Visvikis D, Odet C, Giammarile F, Mognetti T, et al. Development of a database of realistic simulated whole body 18F-FDG images for lymphoma. Proc IEEE Nuclear Science Symposium and Medical Imaging Conference. Dresden, Germany: IEEE; 2008. pp. 4958–63.

160. Le Maitre A, Segars WP, Marache S, Reilhac A, Hatt M, Tomei S, et al. Incorporating patient specific variability in the simulation of realistic whole body 18F-FDG distributions for oncology applications. Proc IEEE 2009;97:2026–38.

161. Zaidi H, Herrmann Scheurer A, Morel C. An object-oriented Monte Carlo simulator for 3D positron tomographs. Comput Methods Programs Biomed 1999;58:133–45.

162. Jan S, Santin G, Strul D, Staelens S, Assie K, Autret D, et al. GATE: a simulation toolkit for PET and SPECT. Phys Med Biol 2004;49:4543–61.

163. Harrison RL, Vannoy SD, Haynor DR, Gillispie SB, Kaplan MS, Lewellen TK. Preliminary experience with the photon history generator module for a public-domain simulation system for emission tomography. Records of IEEE Nuclear Science Symposium and Medical Imaging Conference; 1993. pp. 1154–58.

164. Ay M, Zaidi H. Development and validation of MCNP4C-based Monte Carlo simulator for fan- and cone-beam x-ray CT. Phys Med Biol 2005;50:4863–85.

165. Kyriakou Y, Riedel T, Kalender WA. Combining deterministic and Monte Carlo calculations for fast estimation of scatter intensities in CT. Phys Med Biol 2006;51:4567–86.

166. Malusek A, Sandborg M, Carlsson GA. CTmod-A toolkit for Monte Carlo simulation of projections including scatter in computed tomography. Comput Methods Programs Biomed 2008;90:167–78.

167. Ay M, Zaidi H. Assessment of errors caused by X-ray scatter and use of contrast medium when using CT-based attenuation correction in PET. Eur J Nucl Med Mol Imaging 2006;33:1301–13.

168. Zhang YJ. A survey on evaluation methods for image segmentation. Pattern Recognit Lett 1996;29:1335–46.

169. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, et al. Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 2004;11:178–89.

170. Edwards PJ, Nijmeh AD, McGurk M, Odell E, Fenlon MR, Marsden PK, et al. Validation of PET imaging by alignment to histology slices. Int Conf Med Image Comput Comput Assist Interv 2005;8:968–75.

171. Daisne JF, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler 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.

172. Mamede M, Abreu ELP, Oliva MR, Nosé V, Mamon H, Gerbaudo VH. FDG-PET/CT tumor segmentation-derived indices of metabolic activity to assess response to neoadjuvant therapy and progression-free survival in esophageal cancer: correlation with histopathology results. Am J Clin Oncol 2007;30:377–88.

173. Burri RJ, Rangaswamy B, Kostakoglu L, Hoch B, Genden EM, Som PM, et al. Correlation of positron emission tomography standard uptake value and pathologic specimen size in cancer of the head and neck. Int J Radiat Oncol Biol Phys 2008;71:682–8.

174. Venel Y, Garhi H, de Muret A, Baulieu J-L, Barillot I, Prunier-Aesch C. Comparaison de six méthodes de segmentation du volume tumoral sur la 18F-FDG TEP-TDM avec le volume de référence anatomopathologique dans les cancers bronchopulmonaires non à petites cellules. Médecine Nucléaire 2008;32:339–53.

175. Seitz O, Chambron-Pinho N, Middendorp M, Sader R, Mack M, Vogl TJ, et al. 18F-Fluorodeoxyglucose-PET/CT to evaluate tumor, nodal disease, and gross tumor volume of oropharyngeal and oral cavity cancer: comparison with MR imaging and validation with surgical specimen. Neuroradiology 2009;51:677–86.

176. Yu J, Li X, Xing L, Mu D, Fu Z, Sun X, et al. Comparison of tumor volumes as determined by pathologic examination and FDG-PET/CT images of non-small-cell lung cancer: a pilot study. Int J Radiat Oncol Biol Phys 2009;75:1468–74.

177. Yu HM, Liu YF, Hou M, Liu J, Li XN, Yu JM. Evaluation of gross tumor size using CT, (18)F-FDG PET, integrated (18)F-FDG PET/CT and pathological analysis in non-small cell lung cancer. Eur J Radiol 2009;75:1468–74.

178. Dahele M, Hwang D, Peressotti C, Sun L, Kusano M, Okhai S, et al. Developing a methodology for three-dimensional correlation of PET-CT images and whole-mount histopathology in non-small-cell lung cancer. Curr Oncol 2008;15:62–9.

179. Christian N, Lee JA, Bol A, De Bast M, Jordan B, Grégoire V. The limitation of PET imaging for biological adaptive-IMRT assessed in animal models. Radiother Oncol 2009;91:101–16.

180. Geets X, Daisne JF, Gregoire V, Hamoir M, Lonneux M. Role of 11-C-methionine positron emission tomography for the delineation of the tumor volume in pharyngo-laryngeal squamous cell carcinoma: comparison with FDG-PET and CT. Radiother Oncol 2004;71:267–73.

181. Topkan E, Yavuz AA, Aydin M, Onal C, Yapar F, Yavuz MN. Comparison of CT and PET-CT based planning of radiation therapy in locally advanced pancreatic carcinoma. J Exp Clin Cancer Res 2008;27:41.

182. Ford EC, Lavely WC, Frassica DA, Myers LT, Asrari F, Wahl RL, et al. Comparison of FDG-PET/CT and CT for delineation of lumpectomy cavity for partial breast irradiation. Int J Radiat Oncol Biol Phys 2008;71:595–602.

183. Visser EP, Philippens MEP, Kienhorst L, Kaanders JHAM, Corstens FHM, de Geus-Oei L-F, et al. Comparison of tumor volumes derived from glucose metabolic rate maps and SUV maps in dynamic 18F-FDG PET. J Nucl Med 2008;49:892–8.

184. Grgic A, Nestle U, Schaefer-Schuler A, Kremp S, Kirsch CM, Hellwig D. FDG-PET-based radiotherapy planning in lung cancer: optimum breathing protocol and patient positioning—an intraindividual comparison. Int J Radiat Oncol Biol Phys 2009;73:103–11.

185. Zou KH, Wells WM, Kikinis R, Warfield SK. Three validation metrics for automated probabilistic image segmentation of brain tumours. Stat Med 2004;23:1259–82.

186. Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruysscher D, Oellers M, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. Int J Radiat Oncol Biol Phys 2010: in press.

187. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

188. Swensson RG. Unified measurement of observer performance in detecting and localizing target objects on images. Med Phys 1996;23:1709–25.

189. Zou KH, Warfield SK, Fielding JR, Tempany CM, William MW, Kaus MR, et al. Statistical validation based on parametric receiver operating characteristic analysis of continuous classification data. Acad Radiol 2003;10:1359–68.

190. Henkelman RM, Kay I, Bronskill MJ. Receiver operator characteristic (ROC) analysis without truth. Med Decis Making 1990;10:24–9.

191. Beiden SV, Campbell G, Meier KL, Wagner RF. On the problem of ROC analysis without truth: the EM algorithm and the information matrix. Proc SPIE 2000;3981:126–34.

192. Hoppin JW, Kupinski MA, Kastis GA, Clarkson E, Barrett HH. Objective comparison of quantitative imaging modalities without the use of a gold standard. IEEE Trans Med Imaging 2002;21:441–9.

193. Kupinski MA, Hoppin JW, Clarkson E, Barrett HH, Kastis GA. Estimation in medical imaging without a gold standard. Acad Radiol 2002;9:290–7.

194. Hoppin JW, Kupinski MA, Wilson DW, Peterson T, Gershman B, Kastis G, et al. Evaluating estimation techniques in medical imaging without a gold standard: experimental validation. Proc SPIE 2003;5034:230–7.

195. Zaidi H, Ruest T, Schoenahl F, Montandon M-L. Comparative evaluation of statistical brain MR image segmentation algorithms and their impact on partial volume effect correction in PET. Neuroimage 2006;32:1591–607.

196. Maes F, Vandermeulen D, Suetens P. Medical image registration using mutual information. Proc IEEE 2003;91:1699–722.

197. Viola P. Alignment by maximization of mutual information. [PhD Thesis]. Massachusetts Institute of Technology; Cambridge, 1995.

198. Holden M, Hill DL, Denton ER, Jarosz JM, Cox TC, Rohlfing T, et al. Voxel similarity measures for 3-D serial MR brain image registration. IEEE Trans Med Imaging 2000;19:94–102.

199. Aerts HJ, Bosmans G, van Baardwijk AA, Dekker AL, Oellers MC, Lambin P, et al. Stability of (18)F-deoxyglucose uptake locations within tumor during radiotherapy for NSCLC: a prospective study. Int J Radiat Oncol Biol Phys 2008;71:1402–7.

200. Kumar R, Dhanpathi H, Basu S, Rubello D, Fanti S, Alavi A. Oncologic PET tracers beyond [(18)F]FDG and the novel quantitative approaches in PET imaging. Q J Nucl Med Mol Imaging 2008;52:50–65.

201. Lewis JS, Welch MJ, Tang L. Workshop on the production, application and clinical translation of "non-standard" PET

nuclides: a meeting report. Q J Nucl Med Mol Imaging 2008;52: 101–6.

202. Bading JR, Shields AF. Imaging of cell proliferation: status and prospects. J Nucl Med 2008;49 Suppl 2:64S–80.

203. Dunphy MPS, Lewis JS. Radiopharmaceuticals in preclinical and clinical development for monitoring of therapy with PET. J Nucl Med 2009;50:106S–21.

204. Koch CJ, Evans SM. Non-invasive PET and SPECT imaging of tissue hypoxia using isotopically labeled 2-nitroimidazoles. Adv Exp Med Biol 2003;510:285–92.

205. Grosu AL, Souvatzoglou M, Röper B, Dobritz M, Wiedenmann N, Jacob V, et al. Hypoxia imaging with FAZA-PET and theoretical considerations with regard to dose painting for individualization of radiotherapy in patients with head and neck cancer. Int J Radiat Oncol Biol Phys 2007;69:541–51.

206. Jager PL, Chirakal R, Marriott CJ, Brouwers AH, Koopmans KP, Gulenchyn KY. 6-L-18F-fluorodihydroxyphenylalanine PET in neuroendocrine tumors: basic aspects and emerging clinical applications. J Nucl Med 2008;49:573–86.

207. Tang BN, Van Simaeys G, Devriendt D, Sadeghi N, Dewitte O, Massager N, et al. Three-dimensional Gaussian model to define brain metastasis limits on (11)C-methionine PET. Radiother Oncol 2008;89:270–7.

208. Ciernik IF, Brown DW, Schmid D, Hany T, Egli P, Davis JB. 3D-segmentation of the 18F-choline PET signal for target volume definition in radiation therapy of the prostate. Technol Cancer Res Treat 2007;6:23–30.

209. Wang H, Vees H, Miralbell R, Wissmeyer M, Steiner C, Ratib O, et al. (18)F-fluorocholine PET-guided target volume delineation techniques for partial prostate re-irradiation in local recurrent prostate cancer. Radiother Oncol 2009;93:220–5.

210. Weber D, Wang H, Cozzi L, Dipasquale G, Khan H, Ratib O, et al. RapidArc, intensity modulated photon and proton techniques for recurrent prostate cancer in previously irradiated patients: a treatment planning comparison study. Radiat Oncol 2009;4:34.

211. Patel DA, Chang ST, Goodman KA, Quon A, Thorndyke B, Gambhir SS, et al. Impact of integrated PET/CT on variability of target volume delineation in rectal cancer. Technol Cancer Res Treat 2007;6:31–6.

212. Weber DC, Zilli T, Buchegger F, Casanova N, Haller G, Rouzaud M, et al. [(18)F]Fluoroethyltyrosine-positron emission tomography-guided radiotherapy for high-grade glioma. Radiat Oncol 2008;3:44.

213. Zaidi H, Mawlawi O. Simultaneous PET/MR will replace PET/CT as the molecular multimodality imaging platform of choice. Med Phys 2007;34:1525–8.

214. Pichler BJ, Wehrl HF, Kolb A, Judenhofer MS. Positron emission tomography/magnetic resonance imaging: the next generation of multimodality imaging? Semin Nucl Med 2008;38: 199–208.

215. Hillner BE, Siegel BA, Liu D, Shields AF, Gareen IF, Hanna L, et al. Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the National Oncologic PET Registry. J Clin Oncol 2008;26:2155–61.

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

217. Davis JB, Reiner B, Huser M, Burger C, Szekely G, Ciernik IF. Assessment of (18)F PET signals for automatic target volume definition in radiotherapy treatment planning. Radiother Oncol 2006;80:43–50.

218. Drever L, Roa W, McEwan A, Robinson D. Iterative threshold segmentation for PET target volume delineation. Med Phys 2007;34:1253–65.

219. Vauclin S, Doyeux K, Hapdey S, Edet-Sanson A, Vera P, Gardin I. Development of a generic thresholding algorithm for the delineation of 18FDG-PET-positive tissue: application to the comparison of three thresholding models. Phys Med Biol 2009;54:6901–16.

220. Day E, Betler J, Parda D, Reitz B, Kirichenko A, Mohammadi S, et al. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. Med Phys 2009;36:4349–58.

221. van Dalen JA, Hoffmann AL, Dicken V, Vogel WV, Wiering B, Ruers TJ, et al. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. Nucl Med Commun 2007;28:485–93.

222. Erdi YE, Rosenzweig K, Erdi AK, Macapinlac HA, Hu Y-C, 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.

223. Vrieze O, Haustermans K, De Wever W, Lerut T, Van Cutsem E, Ectors N, et al. Is there a role for FGD-PET in radiotherapy planning in esophageal carcinoma? Radiother Oncol 2004;73:269–75.

224. van Loon J, Offermann C, Bosmans G, Wanders R, Dekker A, Borger J, et al. 18FDG-PET based radiation planning of mediastinal lymph nodes in limited disease small cell lung cancer changes radiotherapy fields: a planning study. Radiother Oncol 2008;87:49–54.

225. Breen SL, Publicover J, De Silva S, Pond G, Brock K, O'Sullivan B, et al. Intraobserver and interobserver variability in GTV delineation on FDG-PET-CT images of head and neck cancers. Int J Radiat Oncol Biol Phys 2007;68:763–70.

226. Schinagl DA, Hoffmann AL, Vogel WV, van Dalen JA, Verstappen SM, Oyen WJ, et al. Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer? Radiother Oncol 2009;91:95–100.

227. Murakami R, Uozumi H, Hirai T, Nishimura R, Katsuragawa S, Shiraishi S, et al. Impact of FDG-PET/CT fused imaging on tumor volume assessment of head-and-neck squamous cell carcinoma: intermethod and interobserver variations. Acta Radiol 2008;49:693–9.

228. El-Bassiouni M, Ciernik IF, Davis JB, El-Attar I, Reiner B, Burger C, et al. [18FDG] PET-CT-based intensity-modulated radiotherapy treatment planning of head and neck cancer. Int J Radiat Oncol Biol Phys 2007;69:286–93.

229. Deantonio L, Beldi D, Gambaro G, Loi G, Brambilla M, Inglese E, et al. FDG-PET/CT imaging for staging and radiotherapy treatment planning of head and neck carcinoma. Radiat Oncol 2008;3:29.