Heliyon 6 (2020) e05267

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Intensity threshold based solid tumour segmentation method for Positron Emission Tomography (PET) images: A review

Mahbubunnabi Tamal *

Department of Biomedical Engineering, Imam Abdulrahman Bin Faisal University, PO Box 1982, Dammam, 31441, Saudi Arabia

ARTICLE INFO

Keywords:
Biomedical engineering
Medical imaging
Oncology
Image processing
Computer-aided engineering
Intensity threshold
Segmentation
PET
Standardized uptake value (SUV)

ABSTRACT

Accurate, robust and reproducible delineation of tumour in Positron Emission Tomography (PET) is essential for diagnosis, treatment planning and response assessment. Since standardized uptake value (SUV) – a normalized semiquantitative parameter used in PET is represented by the intensity of the PET images and related to the radiotracer uptake, a SUV based threshold method is a natural choice to delineate the tumour. However, determination of an optimum threshold value is a challenging task due to low spatial resolution, and signal-to-noise ratio (SNR) along with finite image sampling constraint. The aim of the review is to summarize different fixed and adaptive threshold-based PET image segmentation approaches under a common mathematical framework Advantages and disadvantages of different threshold based methods are also highlighted from the perspectives of diagnosis, treatment planning and response assessment. Several fixed threshold values (30%–70% of the maximum SUV of the tumour (SUV_{maxT})) have been investigated. It has been reported that the fixed threshold-based method is very much dependent on the SNR, tumour to background ratio (TBR) and the size of the tumour. Adaptive threshold-based method, an alternative to fixed threshold, can minimize these dependencies by accounting for tumour to background ratio (TBR) and tumour size. However, the parameters for the adaptive methods need to be calibrated for each PET camera system (e.g., scanner geometry, image acquisition protocol, reconstruction algorithm etc.) and it is not straight forward to implement the same procedure to other PET systems to obtain similar results. It has been reported that the performance of the adaptive methods is also not optimum for smaller volumes with lower TBR and SNR. Statistical analysis carried out on the NEMA thorax phantom images also indicates that regions segmented by the fixed threshold method are significantly different for all cases. On the other hand, the adaptive method provides significantly different segmented regions only for low TBR with different SNR. From this viewpoint, a robust threshold based segmentation method that will be less sensitive to SUV_{maxT}, SNR, TBR and volume needs to be developed. It was really challenging to compare the performance of different threshold-based methods because the performance of each method was tested on dissimilar data set with different data acquisition and reconstruction protocols along with different TBR, SNR and volumes. To avoid such difficulties, it will be desirable to have a common database of clinical PET images acquired with different image acquisition protocols and different PET cameras to compare the performance of automatic segmentation methods. It is also suggested to report the changes in SNR and TBR while reporting the response using threshold based methods.

1. Introduction

Positron Emission Tomography (PET), a modality of nuclear medicine imaging, shows a very high sensitivity and quantitative capability (picomolar concentrations of radiotracer can be imaged) over other imaging techniques making it suitable for direct in vivo biochemical studies and three-dimensional (3D) whole body imaging (Reader and Zweit, 2001). Availability of wide range of physiologically relevant image contrast also makes PET a flexible imaging modality. Since its introduction in medicine, PET has extensively been used in tumour diagnosis and staging (Czernin et al., 2007; Valk et al., 2004). Recently PET is being used in treatment planning (Ford et al., 2009) and monitoring early response to therapy with development of novel radio tracers (Kenny et al., 2007).

Two PET images feature are conventionally considered to be the most relevant in clinical settings - Standardized uptake value (SUV) and

* Corresponding author.
E-mail addresses: mtamal@iau.edu.sa, mtamal@yahoo.com.

https://doi.org/10.1016/j.heliyon.2020.e05267
Received 24 September 2019; Received in revised form 14 May 2020; Accepted 12 October 2020

metabolic tumour volume (MTV). SUV – a normalized semiquantitative parameter that is derived using the intensity of the PET images. Tumour and surrounding background organs or tissues radiotracer uptake is conventionally represented by SUV. MTV is defined by the volume with similar SUV (Chung et al., 2009). The objectives of PET images segmentation vary based on whether the SUV or MTV is subsequently going to be used for diagnosis, treatment planning or response assessment.

The purpose of delineation of PET lesion varies based on the ultimate utilization of the segmentation ROI. In radiotherapy planning, the objective of PET image is delineation is to have a significant overlap between tumour's anatomical volume and MTV. On the other hand, for diagnosis and response to therapy assessment in oncological PET, SUV along with MTV and full quantitative (kinetic) parameters are also examined. In clinical practice, SUV is used as a image based biomarker to measure and quantify treatment response (Weber, 2005). In conventional response assessment, generally maximum SUV (SUV_{max}) representing a single-voxel value is used (Huang et al., 2011). However, response assessment based on SUV value extracted from a single voxel may lead to uncertainty caused by the presence of high level of PET image noise (Krak et al., 2005). Peak (Wahl et al., 2009) or mean SUV (Nahmias and Wahl, 2008) (SUV_{peak} or SUV_{mean}) are proposed as an alternative to SUV_{max} that are more robust to noise. SUV_{peak} is normally defined as the average of SUVs centered around a high radiotracer uptake part within a ROI. Whereas SUV_{mean} is the average SUV over all the voxels located within the ROI. Since both SUV_{peak} and SUV_{mean} represent average SUV values over few voxels, they are less sensitive to image noise. Because of this reason, they are suitable to be used in quantification of treatment response with lower uncertainty (Krak et al., 2005). However, several biological and technical factors along with region of interest (ROI) segmentation and peak definition impact the reproducibility of these parameters (Adams et al., 2010; Cheebsumon et al., 2011b; Vanderhoek et al., 2012).

Though manual segmentation of tumour is considered as the gold standard, it is sensitive to inter and intra observer variability (Maroy et al., 2008), particularly when only PET is used for delineation (Breen et al., 2007). Some of the challenges can be circumvented by fusing of CT anatomical information with PET (McKay et al., 2018; Townsend, 2008). However, the manual delineation problem still remains challenging as underlying physiological edge does not always necessarily relate to anatomical boundary (Greco et al., 2007; Messa et al., 2006).

To reduce the variability arising from manual segmentation, a number of automatic techniques were proposed. They can be categorized broadly into four groups (Zaidi and El Naqa, 2010) - a) thresholding (Erdi et al., 1997), b) variational approaches (El Naqa et al., 2007), c) learning methods (Belhassen and Zaidi, 2010) and d) stochastic modelling-based (Aristophanous et al., 2007). Comprehensive reviews of all these techniques have already been presented by other authors (Foster et al., 2014; Zaidi and El Naqa, 2010).

Among these methods, SUV based threshold is a natural choice for segmenting tumour from background as it is related to the intensity of the image (Erdi et al., 1997) and is widely used due to its simplicity and efficiency in implementation in clinical practice. Several papers were published either to review (Foster et al., 2014; Lee, 2010; Zaidi and El Naqa, 2010) or to compare (Cheebsumon et al., 2011a; Tylski et al., 2010) different threshold methods. The purpose of this article is it to review threshold-based methods that are generally used in clinical settings using a common mathematical framework. This review article also highlights methods, working principle, advantages and limitations of threshold-based PET image segmentation from two different perspectives –response to treatment and radiotherapy planning. A few recommendations are discussed in this paper.

2. Threshold based segmentation

Several studies suggested that threshold based segmentation method is robust to noise and resolution compared to other segmentation methods (e.g., clustering (Belhassen and Zaidi, 2010), image gradient based (Geets et al., 2007) etc.). However, due to low signal-to-noise ratio (SNR), finite image sampling constraint and poor spatial resolution of PET camera, the SUVs within the tumour and its surrounding area are not easily discernible (Figure 1). Because of this reason determination of an optimum threshold value is a difficult task (Schaefer et al., 2008).

Threshold based PET image segmentation methods proposed in the literature can be broadly categorized in two classes based on how the optimum threshold value is determined – 1) fixed threshold method and 2) adaptive threshold method. Figure 2 shows the further subgroups.

3. Fixed threshold (T_{Fixed})

In a fixed threshold based PET image segmentation method, a fixed percentage (T_{Fixed}) of the SUV_{max} of the tumour (SUV_{maxT}) is used to determine the SUV_{Fixed} value and can be given as

$$SUV_{Fixed} = T_{Fixed} \times SUV_{maxT} \tag{1}$$

Voxels having SUV greater than or equal to SUV_{Fixed} are included in the segmented ROI. Generally, the value of T_{Fixed} is defined empirically. A fixed empirical threshold value can provide a good separation between the tumour and background and yields a good the segmented volume closely matches the true volume (van Dalen et al., 2007). Theoretically, T_{Fixed} value 50% can provide a good estimated volume for noise free image. However, the SUV_{mean} extracted from this segmented volume is smaller than the true SUV_{mean} due to poor resolution of PET camera and partial volume effects (PVE) specifically for tumour with diameter two or three times smaller than the resolution of the imaging system (Kessler et al., 1984; Soret et al., 2007). The problem becomes even challenging in clinical cases where the true volume is unknown and presence of high noise in PET images. In such cases, unreliable results may be obtained if threshold is determined by using only one maximum pixel (Figure 3).

Erdi et al. (1997) proposed that in case where PVE are negligible, e.g., for volumes ($\geq$4 mL), by considering only tumour to background ratio (TBR) 36–44% fixed threshold values can be used as reasonable approximations. For tumours with volume bigger than 4 ml and TBR more than 5, 40% fixed threshold is the most suitable approximation. For each individual TBR, a functional form can be fitted using the following equation

$$T_{Fixed} = A \times \exp^{-(C \times V)} \tag{2}$$

V represents volume in mL. A and C values are determined experimentally by fitting Eq. (2) to the measured data and their values vary with TBRs. For bigger volume the method works well as can be seen from Figure 2. Previous investigation hypothesized that T can be a function of V and TBR

$$T_{Fixed} = f(V, TBR) \tag{3}$$

Another study suggested that the T_{Fixed} value is independent of TBR and can be determined directly from the size of the tumour as (van Dalen et al., 2007)

$$T_{Fixed} = \begin{cases} 40\% \, for > 15 \text{ mm diametr tumour} \\ 50\% \, fo \, r < 12 \text{ mm diametr tumour} \end{cases} \tag{4}$$

A range of other T_{Fixed} values ranging between 30 – 90 % have been studied and evaluated on patient and phantom. Some of them are summarized in Table 1.

The detail investigation of the fixed threshold reveals that there is no conclusive consensus on the value of that the optimum fixed threshold value, T_{Fixed} is available and T_{Fixed} is not only dependent on the volume and TBR but also scanner specification, reconstruction and data acquisition protocols, noise level etc. Since SNR and Full Width Half Maximum (FWHM) as a measure of image resolution include all these parameters, T_{Fixed} can be expressed as

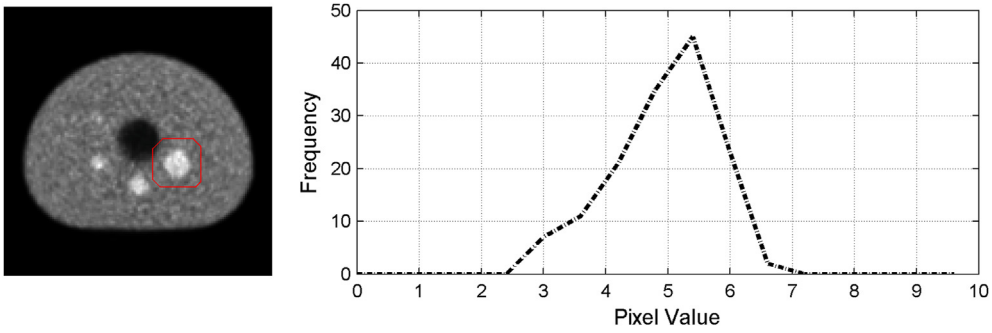

Figure 1. Left - an example PET image of a NEMA thorax phantom. Right - a histogram of the ROI marked by red colour on the left image. Since the histogram is not bimodal it is challaenging to determine a threshold level form the histogram.

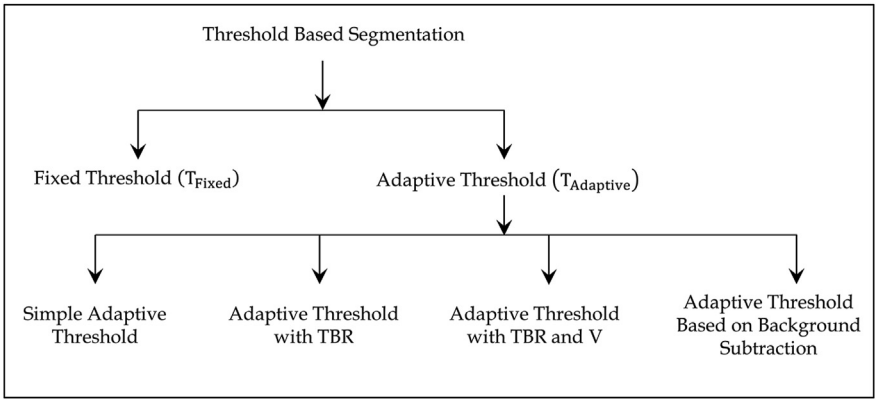

Figure 2. Overview of intensity threshold-based segmentation methods for PET images.

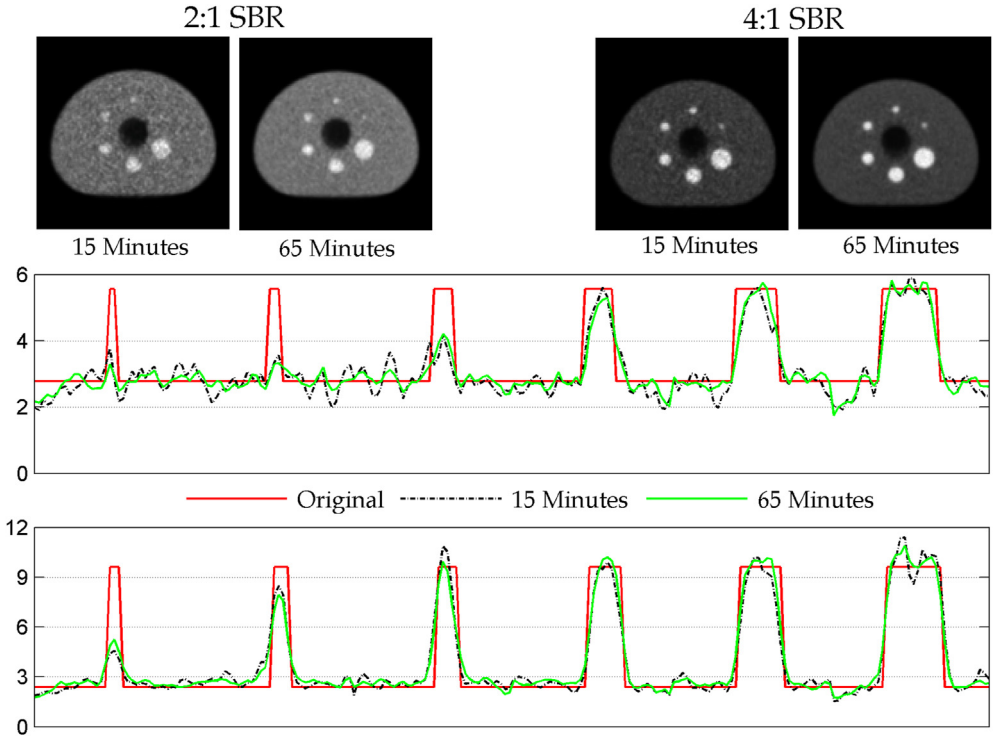

Figure 3. Top row: PET NEMA thorax phantom with 2:1 and 4:1 tumour to background (TBR) ratio or contrast with 15- and 65-minutes acquisition duration to represent different signal to noise ratio (SNR). Middle row: horizontal profiles through all the six spheres for 2:1. Bottom row: 4:1.

Table 1. Fixed threshold method.

	Segmentation method	Object type Or anatomical location	Radiotracer and Data acquisition duration	Object Size	TBR or Contrast	Validation procedure	Reconstruction algorithm parameters	Reference
1	T_{Fixed} 42%	An elliptical Jaszczak phantom	^{18}F 15 min	0.4–5.5 mL	7.4, 5.5, 3.1, and 2.8	Validated on phantom study	Hann filter with decay, randoms, scatter, and attenuation correction.	(Erdi et al., 1997)
		10 patients 17 primary or metastatic lung lesions	^{18}F-FDG 45 min	0.88–76.94 mL	2.3 to 9.2		Using standard clinical parameters	
2	T_{Fixed} 34%	Moving phantom	Over 20 min	^{22}Na 3 sphere with inner diameters of 1.3, 2.9, and 6.6 cm.	N/A	Compared to actual sphere volume	Using an iterative, ordered subsets, expectation maximization algorithm.	(Caldwell et al., 2003)
3	T_{Fixed} 30%, 40%, 50%, 60%, 70% and 90%	Non–Small Cell Lung Cancer (NSCLC) patient	^{18}F-FDG 3 min per bed position	N/A	N/A	Compared with the PET relapse volume	Fourier rebinning and attenuation-weighted ordered-subset expectation maximization	(Calais et al., 2015)
4	T_{Fixed} 40%, 50%, 60% and 70%	Recurrent High Grade Glioma Patient	^{18}F-FET 40 min	N/A	N/A	Compared with MRI	Siemens Biograph 6 PET/CT scanner. Algorithm details not available	(Debus et al., 2018)
5	T_{Fixed} 50% and 70%	Phantom and Simulation	^{18}F-FDG 3-5, 7–10, 30–40 min	10, 15, 20, and 30-mm diameter	2, 4, and 8	Validated on phantom study	Ordered Subset Expectation Maximization (OSEM) with 2 iterations and 16 subsets, with or without a 6-mm Gaussian filter	(Boellaard et al., 2004)
6	T_{Fixed} 50%	40 patients Intact squamous cell carcinoma	^{18}F-FDG 5 min per bed position.	2–456 mL	N/A	Compared to CTGTV	Using OSEM iterative reconstruction with two iterations and 28 subsets.	(Paulino et al., 2005)
7	T_{Fixed} 50%	101 patient Non–small-cell lung cancer (NSCLC)	^{18}F-FDG 40-min	N/A	N/A	Compared to CTGTV	Iterative Algorithm (OSEM 180).	(Moureau-Zabotto et al., 2005)
8	T_{Fixed} 40%, SUV2.5	19 patients Non–small-cell lung cancer (NSCLC)	^{18}F-FDG Over 45 min	12–922 mL	N/A	Compared to CTGTV	N/A	(Hong et al., 2007)
9	Manual, T_{Fixed} 40%, 50%, TBR	78 patients Squamous cell carcinoma of the head and neck area	^{18}F-FDG 16 min for the two bed positions	17.4–27.9 mL	N/A	Compared to GTVvis	2D ordered subset Expectation maximization iterative algorithm (four iterations, 16 subsets).	(Schinagl et al., 2007)
10	Manual, SUV2.5 T_{Fixed} 40%, TBR	25 patients Non–small-cell lung cancer (NSCLC)	^{18}F-FDG	9.3–666.2 mL	N/A	Validated on phantom study	Iterative reconstruction into 128 × 128 pixels of 5.1 mm	(Nestle et al., 2005)
11	T_{Fixed} 43%, SUV2.5	The Jaszczak deluxe flangeless ECT Phantom	^{18}F-FDG 3 min per bed	6 cylinders of diameter (3 at 25 mm and 3 At 8, 12, and 16 mm) and a length of 38.1 mm	2.5:1	Validated on phantom study	Ordered subset expectation Maximization (four iterations and eight subsets)	(Day et al., 2009)
		18 patients rectal or anal cancer	^{18}F-FDG 2–4 min per bed position	10.5–129.5 mL	N/A			
12	TBR T_{Fixed} 31%–47%.	A spherical lucite phantom	^{18}F-FDG Two 10 min transmission scans	Six spheres with volumes ranging from 0.5 to 32 mL	8.7, 6.5, 4.7, 3.0, 2.1, and 1.5	Validated on phantom study and macroscopic examination.	3D Filtered Back Projection (FBP) algorithm, a Fourier rebinning (FORE) algorithm and a FORE rebinning algorithm With an OSEM reconstruction	(van Baardwijk et al., 2007)
		33 consecutive patients' NSCLC 23 Tumors	^{18}F-FDG 5-min bed positions (total of seven	N/A	N/A		Reconstructed iteratively With a reconstruction increment of 5 mm	

(*continued on next page*)

Table 1 (*continued*)

Segmentation method	Object type Or anatomical location	Radiotracer and Data acquisition duration	Object Size	TBR or Contrast	Validation procedure	Reconstruction algorithm parameters	Reference
		bed positions with Overlap)					
13 T_{Fixed} 42% and 50%, FCM, FLAB	IEC phantom	^{18}F (1, 2 and 5 min)	10, 13, 17, 22, 28, and 37 mm diameters	4:1 and 8:1	Validated on phantom and simulated scans studies.	OPL-EM(seven iterations, one subset), RAMLA 3D, TF ML-EM, FORE-OSEM, and OSEM (four iterations and eight subsets).	(Hatt et al., 2011b)
	10 tumors from simulated PET database	^{18}F-FDG Routine clinical image a protocols were used	N/A	N/A			
14 Manual, TBR, T_{Fixed} 40% and 50%, SUV2.5	The NEMA IEC/2001 image quality body phantom	^{18}F 15 min	0.5, 1.2, 2.6, 5.6, 11.5, and 26.5 mL	10.0, 8.6, 6.9, 5.0, 3.3, and 2.1.	Validated on preliminary phantom study.	2D (NAW-OSEM) iterative normalized attenuation-weighted ordered subsets expectation maximization. With four iterations and eight subsets	(Vees et al., 2009)
	18 patients High-grade gliomas	^{18}F-FET 30 min	20.30–40.95 mL	N/A			
15 Manual, tbgd20%, tbgd40%, SUV2.5, T_{Fixed} 40%	16 patients with esophageal squamous cell carcinoma (ESCC)	^{18}F-FDG 6 or 7 bed positions at 2 min per Bed position.	(5.90 ± 2.38) cm	N/A	Compared to region of gross tumor in vivo	Ordered subset Expectation maximization (OSEM) algorithm	(Yu et al., 2009)
16 Gradient Based, TBR, T_{Fixed} 40%,	10 patients NSCLC.	^{18}F-FDG 35 min	N/A	8.7, 6.5, 4.7, 3.0, 2.1, and 1.5	Validated on phantom and HNSCC Patient materials	3D LOR-OSEM (line of response ordered-subset expectation Maximization)	(Wanet et al., 2011)

$$T_{Fixed} = f(V, TBR, SNR, FWHM) \quad (5)$$

Closed investigation of the PET data also reveals that for smaller lesions, volumes are overestimated if lower T_{Fixed} values (<40%) are used and the overestimation amount is very much dependent not only on TBR but also on SNR (Figure 3).

4. Adaptive threshold ($T_{Adaptive}$)

To overcome the dependency of fixed threshold-based method on different parameters (V, TBR, SNR and FWHM), several adaptive threshold based PET image segmentation methods were proposed. The main idea of adaptive threshold method is that, the threshold value is independently determined for each individual setting of parameters and hence, it is adaptable ($T_{Adaptive}$). Adaptive threshold can be sub-categorized into four groups – a) simple adaptive threshold, b) adaptive threshold with TBR, c) adaptive threshold with TBR and V and d) adaptive threshold based on background subtraction.

4.1. Simple adaptive threshold

In this case, $T_{Adaptive}$ is derived using SUV_{maxT} and $SUV_{Adaptive}$ similar to Eq. (1)

$$T_{Adaptive} = 100 \times \frac{SUV_{Adaptive}}{SUV_{maxT}} \quad (6)$$

But in this case, $SUV_{Adaptive}$ is dependent on the mean SUV of the tumour (SUV_{meanT}) and background (SUV_{meanBG}) and given as

$$SUV_{Adaptive} = f(SUV_{meanT}, SUV_{meanBG}) = (\alpha \times SUV_{meanT}) + (\beta \times SUV_{meanBG}) \quad (7)$$

α and β values are empirically determined by fitting Eq. (7) to the phantom data containing spheres with different sizes and TBRs. $T_{Adaptive}$ values determined in such a way not only depend on the scanner, data acquisition and image reconstruction protocols but also on volume and TBR. α and β required to be determined for each scanner, reconstruction algorithm, volume and TBR. α and β values are also dependent on SNR via acquisition durations. An example of regression to determine α and β for two acquisition durations (15 and 65 min) is shown in Figure 4 (Schaefer et al., 2008) and the values are given in Table 2.

Several proposed values for α and β are shown in Table 3.

Once α and β values are chosen using phantom data with known volume and TBR, to determine $T_{Adaptive}$ for patient data using Eq. (7), the values of SUV_{meanT} and $SUV_{meanBG}$are required to be estimated. SUV_{meanT} is derived from the average of the SUVs of those voxels located within the tumour and within the isocontour of 70% of SUV_{maxT}. SUV_{meanBG} is determined by placing a spherical ROI on the background and at a certain distance from the lesion to avoid partial volume effect (Nestle et al., 2005, 2007; Schaefer et al., 2008). If the background is not uniform, the tissue with the highest radiotracer uptake is considered as background (Nestle et al., 2005). The alternative way to calculate the average background value is to fit a Gaussian with the histogram of the ROI with the assumption that 80% voxels of the initial ROI belongs to the background (Tylski et al., 2010).

The level of threshold ($SUV_{Adaptive}$) can also be calculated employing an iterative way by fixing the second term of Eq. (7) to 0.588. In this procedure, it is assumed that either the threshold is not separately effected by the background activity and the sphere volume or there is only indirect effect via their influences on SUV_{meanT} (Black et al., 2004)

$$SUV_{Adaptive} = (\alpha \times SUV_{meanT}) + 0.588 \quad (8)$$

where $\alpha = 0.307$. The iterative process of the algorithm generally starts with $T = 0.307$ to calculate SUV_{Fixed} to get the initial estimated volume to determine SUV_{meanT}as in Eq. (1). This initial SUV_{meanT} is then used as to

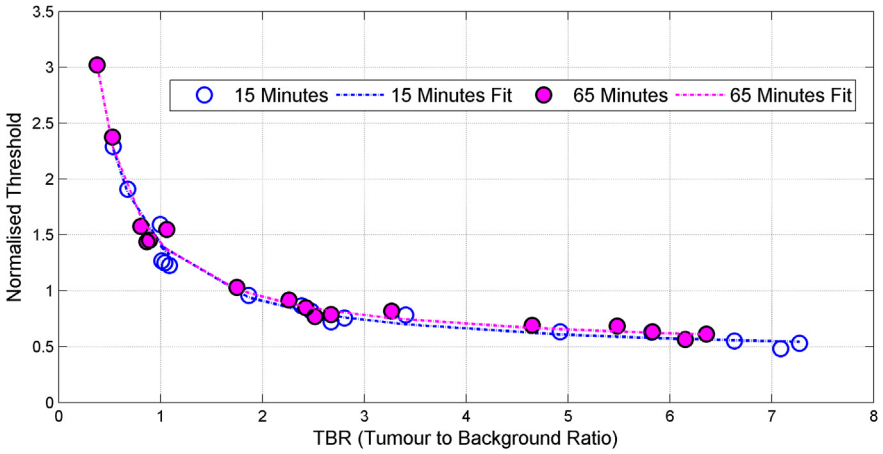

Figure 4. An example of regression to determine α and β for two acquisition durations (15 and 65 min) to represent two different SNR.

Table 2. α and β for two acquisition durations.

Duration	α	β
15 Minutes	0.40	0.59
65 Minutes	0.44	0.52

start the iterative process of determining the new $T_{Adaptive}$ using Eq. (8) This procedure is repeated until less than one voxel difference two subsequent segmented regions provided by two subsequent iterations are detected.

This is to be noted that none of the above mentioned threshold methods explicitly include TBR in their respective regression formulas. With the decrease in background uptake or increase in tumour uptake, TBR increases as a result $T_{Adaptive}$ also increases. In other words, $T_{Adaptive}$ can be considered to be proportional to the radiotracer uptake of both tumour and background. Several different other methods were proposed to explicitly account for TBR (Daisne et al., 2003; Jentzen et al., 2007) and the definition of TBR is very much dependent on the estimations of tumour and background ROIs.

4.2. Adaptive threshold with TBR

In this method, $T_{Adaptive}$ is calculated directly by explicitly accounting for TBR (Daisne et al., 2003)

$$T_{Adaptive} = \alpha + \left(\beta \times \frac{1}{TBR}\right) \tag{9}$$

The paper did not mention the values for α and β. In another work, the values of α and β were reported to be 44.1 and 70.4 (Schinagl et al., 2007). Threshold is inversely related to TBR. Similar findings were also reported by other researchers (Erdi et al., 1997).

4.3. Adaptive threshold with TBR and V

The volume and TBR are not explicitly accounted for in their respective functional forms for all the above mentioned methods. An analytical expression that accounts for variations in volume to generate different TBR curves can be written as (Jentzen et al., 2007)

$$T_{Adaptive} = \frac{0.078}{V} + \left(0.617 \times \frac{1}{TBR}\right) + 0.316 \tag{10}$$

Eq. (10) is derived by fitting it to phantom containing spheres of known volumes. However, for patient data, where the actual volume is unknown, the threshold value is first defined by approximating the TBR. A corresponding suitable TBR curve defined by Eq. (10) is then chosen. The final threshold is determined via iteratively updating both the threshold and volume.

The tumour cross sectional area (A)can also be used instead of the volume (Matheoud et al., 2011).

For $A \leq 133\ mm^2$

$$T_{Adaptive} = \alpha + \left[\beta \times \left(1 - \frac{1}{TBR}\right)\right] + (\gamma \times A) \tag{11}$$

where $\alpha = 55.94$, $\beta = 56.96$, $\gamma = -0.25$.

For cross sectional area larger than $133\ mm^2$ $(A > 133\ mm^2)$, $T_{Adaptive}$ can be considered not to be directly related to either volume or area and can be expressed as

$$T_{Adaptive} = \alpha + \left[\beta \times \left(1 - \frac{1}{TBR}\right)\right] \tag{12}$$

where α and β are reported to be 33.13 and 60.27 respectively.

In another study, threshold is considered to be related to the volume exponentially (Nehmeh et al., 2009). In this case, threshold is proposed to be derived using Monte Carlo simulation along with the background subtracted phantom data

$$T_{Adaptive} = \alpha + \exp^{\left[\beta + \frac{\gamma}{V} + (\delta \times V)\right]} \tag{13}$$

The parameters resulting from the fits are reported to be $\alpha = 5$, $\beta = 3.568$, $\gamma = 0.197$ and $\delta = -0.1069$. Computed Tomography (CT) data is used to define V.

Table 3. α and β values.

Reference	α	β	Scanner	Reconstruction algorithm
(Nestle et al., 2005)	0.15	1.0	Siemens ECAT ART PET	Iterative reconstruction
(Tylski et al., 2010)	0.25	1.0	Siemens Biograph PET/CT	OSEM (6 iterations, 8 subsets)
(Nestle et al., 2007)	0.7	0.5	Siemens ECAT ART PET	Iterative reconstruction
(Schaefer et al., 2008)	0.50 (lesion diameter≥3cm) 0.67 (lesion diameter<3cm)	0.50 0.60	Siemens ECAT ART PET	OSEM (2 iterations, 4 subsets, 2 mm Gaussian filter)

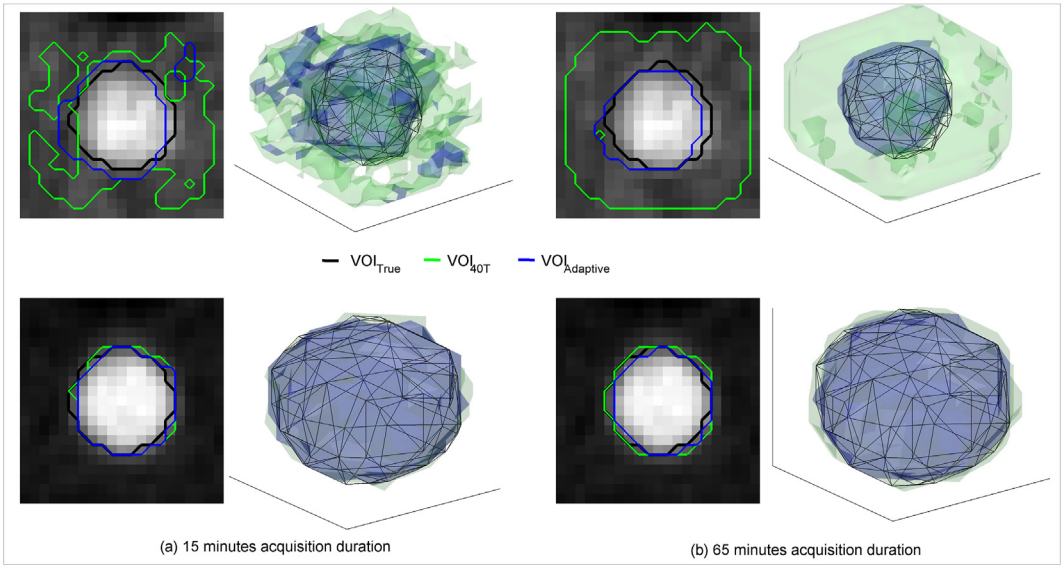

Figure 5. Representative segmentations of a sphere of a NEMA PET phantom with two different TBR (2:1 – top row and 4:1 – bottom row). Fixed threshold method is sensitive to TBR compared to the adaptive threshold methods.

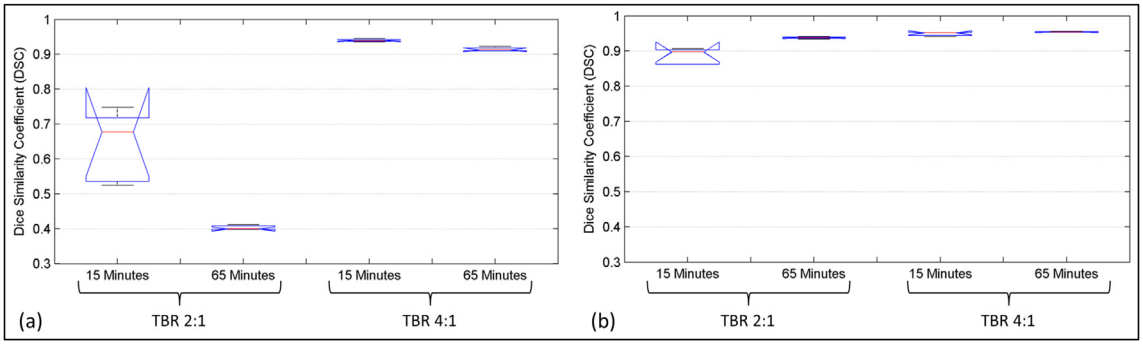

Figure 6. Box plot of segmentation comparison for different acquisition durations: (a) 40% fixed threshold and (b) adaptive threshold. Segmented regions for fixed threshold is significantly different for all cases. Adaptive method provides significantly different regions only for low TBR.

4.4. Background subtraction based adaptive threshold

The main idea behind background subtraction based adaptive threshold method is to determine the optimum threshold by subtracting background. In this method, threshold is determined based on the percentage level of contrast that eliminates the requirement of volume estimation (Drever et al., 2006)

$$SUV_{Adaptive} = \left[\frac{\%TBR}{100} \times (SUV_{maxT} - SUV_{meanBG})\right] + SUV_{meanBG} \quad (14)$$

where the %TBR is determined by the difference between maximum intensity within the tumour (SUV_{maxT}) and average intensity of the background region (SUV_{meanBG}) at each slice location of the phantom.

Another method similar to the method in Eq. (14) has been proposed (Davis et al., 2006)

$$T_{Adaptive} = [T_{rel} \times (SUV_{maxT} - SUV_{meanBG})] + SUV_{meanBG} \quad (15)$$

In this method, a relative threshold, T_{rel} is used instead of %TBR. T_{rel} needs to be first determined for estimating $T_{Adaptive}$. T_{rel} is defined as a fixed ratio of the normalized background subtracted SUV_{maxT} and depends on source diameter. It is generally estimated by setting the background uptake to 0% and maximum activity level to 100% and after subtraction of the background activity. For larger object with greater than 12.5 mm diameter, SUV_{maxT} of single voxel was used. For other cases, average of multiple voxels were used. Eq. (15) has a similar form of Eq. (9) with the only difference that $SUV_{Adaptive}$ needs to be calculated first and can be expressed as

Table 4. *p* value of one way ANOVA test.

	Segmentation Group 1	Segmentation Group 2	*p* value (40%)	*p* value (Adaptive)
1	15 Minutes, TBR 2:1	65 Minutes, TBR 2:1	0.000	0.000
2	15 Minutes, TBR 2:1	15 Minutes, TBR 4:1	0.000	0.000
3	15 Minutes, TBR 2:1	65 Minutes, TBR 4:1	0.000	0.000
4	65 Minutes, TBR 2:1	15 Minutes, TBR 4:1	0.000	0.468
5	65 Minutes, TBR 2:1	65 Minutes, TBR 4:1	0.000	0.144
6	15 Minutes, TBR 4:1	65 Minutes, TBR 4:1	0.854	0.856

$$SUV_{Adaptive} = SUV_{maxT}\left[\alpha + \left(\beta \times \frac{1}{TBR}\right)\right] \quad (16)$$

where $\alpha = T_{rel}$ and $\beta = 1 - T_{rel}$.

One of the challenging task in performance assessment of all these threshold based methods is that they were not evaluated on the same set of patient PET images., On top of it, α and β values are different for each individual model of scanner and image reconstruction protocol. Only a subset of the proposed methods, mainly the adaptive threshold based method that account for the background activity have been evaluated using the same set of data (Tylski et al., 2010).

Representative segmentations of 40% fixed threshold and an adaptive threshold (Schaefer et al., 2008) method is shown in Figure 5 for two different acquisition durations (15 and 65 min) for two different TBR.

An example boxplot of segmentation comparison between 40% fixed threshold and adaptive threshold based segmentation method (Schaefer et al., 2008) for two acquisition durations (15 and 65 min) and contrast levels (2:1 and 4:1 TBR) is shown in Figure 6 using dice similarity coefficient (DSC) for 37 mm sphere for 5 realization. The corresponding p values of one way ANOVA test are shown in Table 4.

It is clearly evident form Figure 6 and Table 4 that 40% threshold method provides different segmentation regions depending on the noise levels for TBR 2:1. On the other hand, for TBR 4:1 the segmented regions for two noise levels are not significant ($p = 0.854$). Pairwise comparison of segmented regions provided by 40% threshold indicates that approximately 83% regions are significantly different (5 out of 6) based on noise and TBR. The adaptive threshold provides much improved results specially with the difference between segmented regions being smaller. However, for low TBR, the segmented regions are still significantly different ($p < 0.05$) with 50% segmented regions (3 out of 6) being significantly different.

5. Discussion

The challenges of delineation of PET images for treatment response assessment and radiation oncology are different. In radiation oncology, PET images are required to be delineated only once. A more accurate delineation of biological target volume (Gregoire et al., 2007; McKay et al., 2018) and dose distribution (Schwartz et al., 2005) are made possible due to availability of combined PET/CT scanner. On the other hand, PET images are acquired to be delineation multiple times for treatment assessment – 1) at baseline (sometimes more than once for reproducibility) and 2) after the treatment. Based on the solid tumour volume changes on CT images, World Health Organization (WHO) and Response Evaluation Criteria in Solid Tumours (RECIST) criteria recommend four treatment outcomes (Gregoire et al., 2007; Huang et al., 2011), for solid tumour - i) completer response (CR), ii) partial response (PR), iii) stable disease (SD) and iv) progressive disease (PD). PET based response criteria is yet to be established (Huang et al., 2011). In practice, changes in semiquantative measures of SUV_{max}, SUV_{mean} or SUV_{peak} before and after treatment are used for response assessment. However, metabolic volume may increase or shrink without having any significant changes in the SUV values. On the other hand, changes in CT volumes or SUV values do not describe the changes in heterogeneous radiotracer uptake of the tumour due to treatment (Therasse et al., 2000). Determination of a gold standard for evaluating PET segmentation method using anatomical volume delineated using CT and/or MRI, therefore, may not be suitable in case of assessing response to treatment. Tumour delineation is also effected by heterogeneous uptake of radiotracer (Hatt et al., 2011a). Conversely, heterogeneity metric itself is effected by tumour delineation method (El Naqa et al., 2009).

Different PET image segmentation methods were proposed over the years ranging from fuzzy logic based methods (Hatt et al., 2009) to more recent deep learning based methods (Chen et al., 2019; Leung et al., 2020). All these methods have their own limitations. One of the major limitations in PET imaging is the variability of radiotracer uptake. Since PET camera acquires in vivo metabolic activity, reproducibility is one of the most challenging tasks in PET imaging unlike other imaging modalities. On the other hand, due to limited number of PET centers around the globe, the available images are not big in numbers for training and if the training set does not contain any specific PET image, there is a high probability of the segmentation method to fail. The image acquisition and reconstruction protocols are also different for different radiotracers and scanner models that makes to have a good training data set more challenging. Also, since deep learning based method works as a black box, it is difficult to identify exactly the cause of such failure. Because of all these reasons, though there are several limitations of threshold based methods, fixed and adaptive threshold are still being used because of its simplicity in implementation and predictability (Le Thiec et al., 2020; Pfaehler et al., 2020; Raynor et al., 2019).

For any specific radiotracer, the dependency of T_{Fixed} on the size of the tumour can be explained by PVE. Since due to spill out effect (provided that the background is cold), the SUV_{maxT} for smaller tumour is always lower compared to the big ones, the value of T_{Fixed} also becomes smaller. Because of this reason, fixed threshold method generally provides overestimated segmented volumes for smaller objects. Similarly, same fixed threshold will provide overestimated volumes for a tracer with lower uptake. For example ^{18}F-FLT or FET uptake is generally significantly lower compared to ^{18}F–FDG and also can vary between patients even between tumours of the same patient (Debus et al., 2018; Kenny et al., 2007). A big range of fixed thresholds (30–70%) is proposed to segment volume with high ^{18}F-FDG uptake on initial PET/CT scans because of potential radiotherapy dose escalation of target volumes (Calais et al., 2015; van Baardwijk et al., 2007). This big margin indicates that when the tumour uptake is high (high SNR) with a cold background (high TBR), a reasonable fixed threshold can provide a reasonable estimation of the true volume.

On the contrary, fixed threshold fails if SNR and TBR are low. In such case, adaptive threshold provides better alternatives by allowing to adapt the threshold value ($T_{Adaptive}$) to the data. However, inclusion of too many parameters in the functional form to determine $T_{Adaptive}$ can make it susceptible to minute changes in the data. A recent study highlights that manual selection of optimal gradient is the most suitable and robust technique in comparison to threshold based segmentation method (Werner-Wasik et al., 2012). Gradient based method is also least effected by variations of physical properties of the PET cameras and different image acquisition and reconstruction protocols. However, gradient based methods are more susceptible to local changes in intensity and a noisy voxel can be detected as a false edge. To overcome the limitation of gradient based method either noise need to be removed using filter before segmentation (Geets et al., 2007) or ROI can be segmented using global gradient information (Tamal, 2019a) or by combining filtering with global gradient information (Tamal, 2019b). Alternatively, multimodal gradient based segmentation can also be considered that can overcome the limitations of either modality (Sbei et al., 2020). With the recent increasing interest in radiomics analysis where textural features are investigated (Aerts et al., 2014; d'Amico et al., 2020), gradient based segmentation method may provide more robust information regarding the texture than intensity threshold method.

The challenges of validation of different automatic segmentation methods are crucial and generally the performance is evaluated in terms of robustness, accuracy and reproducibility (Jannin et al., 2002, 2006). For the purpose of performance assessment, either physical phantom or simulation study where the true volume is known is considered the most suitable option. In real patient studies, identification of gold standard is challenging since the ground truth is never known (Zaidi and El Naqa, 2010). Among several options surgical specimen, information from other imaging modalities (e.g., CT or MRI) are occasionally considered (Daisne et al., 2004). These approaches have their own limitations (Lee, 2010; Stroom et al., 2007). On the other hand, several volume overlap measures, e.g., Dice's similarity index, Hausdorff's contour-to-contour distances, barycentre distance etc. are used. The accuracy of different

segmentation methods can be validated using different volume overlap and statistical measures with different models of phantom (both physiological and mathematical) representing different simple and complicated clinical cases (Lee, 2010; Zaidi and El Naqa, 2010).

6. Conclusion

A review of threshold methods (both fixed and adaptive) is provided using a common mathematical framework. Fixed threshold-based methods are sensitive to SNR, TBR and size. Since threshold based methods are dependent on noise sensitiveSUV_{maxT} value, further investigation is required to evaluate their performance at different noise levels. This is because, due to response to the treatment, a tumour volume not only can reduce but its uptake can also decrease resulting in reduced TBR and SNR. Because of all these reasons, a fixed threshold method may not be suitable for diagnosis, treatment planning and response assessment. It may also not provide similar accuracy for tracer with lower uptake. On the other hand, adaptive threshold based methods are less sensitive to changes in TBR and size. However, the regression functions of the adaptive methods sometimes may contain PET system (e.g. scanner, data acquisition protocol, image acquisition and reconstruction protocols etc.) dependent multiple parameters and therefore, it is not straight forward to implement the same procedure to other clinical PET systems to achieve similar results. The performance of adaptive method is also not optimum for smaller volumes with lower TBR and SNR. From this viewpoint, a robust threshold based technique that is not sensitive to SUV_{maxT}, SNR, TBR and volume needs to be developed. It was really challenging to compare the performance of different threshold-based methods because the performance of each method was tested on dissimilar data set with different data acquisition and reconstruction protocols along with different TBR, SNR and volumes. To avoid such difficulties, it will be desirable to have a common database of clinical PET images acquired with different image acquisition protocols and different PET cameras to compare the performance of automatic segmentation methods. It is also suggested to report the changes in SNR and TBR while reporting the response using threshold based methods.

Declarations

Author contribution statement

Mahbubunnabi Tamal: All authors listed have significantly contributed to the development and the writing of this article.

Funding statement

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Saudi Arabia for funding this research work.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Adams, M.C., Turkington, T.G., Wilson, J.M., Wong, T.Z., 2010. A systematic review of the factors affecting accuracy of SUV measurements. AJR Am. J. Roentgenol. 195, 310–320.

Aerts, H.J., Velazquez, E.R., Leijenaar, R.T., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M.M., Leemans, C.R., Dekker, A., Quackenbush, J., Gillies, R.J., Lambin, P., 2014. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 5, 4006.

Aristophanous, M., Penney, B.C., Martel, M.K., Pelizzari, C.A., 2007. A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. Med. Phys. 34, 4223–4235.

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

Black, Q.C., Grills, I.S., Kestin, L.L., Wong, C.Y., Wong, J.W., Martinez, A.A., Yan, D., 2004. Defining a radiotherapy target with positron emission tomography. Int. J. Radiat. Oncol. Biol. Phys. 60, 1272–1282.

Boellaard, R., Krak, N.C., Hoekstra, O.S., Lammertsma, A.A., 2004. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 45, 1519–1527.

Breen, S.L., Publicover, J., De Silva, S., Pond, G., Brock, K., O'Sullivan, B., Cummings, B., Dawson, L., Keller, A., Kim, J., Ringash, J., Yu, E., Hendler, A., Waldron, J., 2007. Intraobserver and interobserver variability in GTV delineation on FDG-PET-CT images of head and neck cancers. Int. J. Radiat. Oncol. Biol. Phys. 68, 763–770.

Calais, J., Thureau, S., Dubray, B., Modzelewski, R., Thiberville, L., Gardin, I., Vera, P., 2015. Areas of high 18F-FDG uptake on preradiotherapy PET/CT identify preferential sites of local relapse after chemoradiotherapy for non-small cell lung cancer. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 56, 196–203.

Caldwell, C.B., Mah, K., Skinner, M., Danjoux, C.E., 2003. 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. 55, 1381–1393.

Cheebsumon, P., van Velden, F.H., Yaqub, M., Frings, V., de Langen, A.J., Hoekstra, O.S., Lammertsma, A.A., Boellaard, R., 2011. Effects of image characteristics on performance of tumor delineation methods: a test-retest assessment. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 52, 1550–1558.

Cheebsumon, P., Yaqub, M., van Velden, F.H., Hoekstra, O.S., Lammertsma, A.A., Boellaard, R., 2011. Impact of [(1)(8)F]FDG PET imaging parameters on automatic tumour delineation: need for improved tumour delineation methodology. Eur. J. Nucl. Med. Mol. Imag. 38, 2136–2144.

Chen, L., Shen, C., Zhou, Z., Maquilan, G., Albuquerque, K., Folkert, M.R., Wang, J., 2019. Automatic PET cervical tumor segmentation by combining deep learning and anatomic prior. Phys. Med. Biol. 64, 085019.

Chung, M.K., Jeong, H.S., Park, S.G., Jang, J.Y., Son, Y.I., Choi, J.Y., Hyun, S.H., Park, K., Ahn, M.J., Ahn, Y.C., Kim, H.J., Ko, Y.H., Baek, C.H., 2009. Metabolic tumor volume of [18F]-fluorodeoxyglucose positron emission tomography/computed tomography predicts short-term outcome to radiotherapy with or without chemotherapy in pharyngeal cancer. Clin. Canc. Res. : Off. J. Am. Ass. Cancer Res. 15, 5861–5868.

Czernin, J., Allen-Auerbach, M., Schelbert, H.R., 2007. Improvements in cancer staging with PET/CT: literature-based evidence as of September 2006. J. Nucl. Med.: Off. Pub. Soc. Nuclear Med. 48 (Suppl 1), 78S–88S.

d'Amico, A., Borys, D., Gorczewska, I., 2020. Radiomics and artificial Intelligence for PET imaging analysis. Nucl. Med. Rev. Cent. E Eur. 23, 36–39. Central & Eastern Europe.

Daisne, J.-F., Sibomana, M., Bol, A., Doumont, T., Lonneux, M., GrËgoire, V., 2003. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. RADION Radiother. Oncol. 69, 247–250.

Daisne, J.F., Duprez, T., Weynand, B., Lonneux, M., Hamoir, M., Reychler, H., Gregoire, V., 2004. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 233, 93–100.

Davis, J.B., Reiner, B., Huser, M., Burger, C., Szekely, G., Ciernik, I.F., 2006. Assessment of 18F PET signals for automatic target volume definition in radiotherapy treatment planning. Radiother. Oncol. : J. Eur. Soc. Ther. Radiol. Oncol. 80, 43–50.

Day, E., Betler, J., Parda, D., Reitz, B., Kirichenko, A., Mohammadi, S., Miften, M., 2009. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. Med. Phys. 36, 4349–4358.

Debus, C., Waltenberger, M., Floca, R., Afshar-Oromieh, A., Bougatf, N., Adeberg, S., Heiland, S., Bendszus, M., Wick, W., Rieken, S., Haberkorn, U., Debus, J., Knoll, M., Abdollahi, A., 2018. Impact of (18)F-FET PET on target volume definition and tumor progression of recurrent high grade glioma treated with carbon-ion radiotherapy. Sci. Rep. 8, 7201.

Drever, L., Robinson, D.M., McEwan, A., Roa, W., 2006. A local contrast based approach to threshold segmentation for PET target volume delineation. Med. Phys. 33, 1583–1594.

El Naqa, I., Grigsby, P., Apte, A., Kidd, E., Donnelly, E., Khullar, D., Chaudhari, S., Yang, D., Schmitt, M., Laforest, R., Thorstad, W., Deasy, J.O., 2009. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern Recogn. 42, 1162–1171.

El Naqa, I., Yang, D., Apte, A., Khullar, D., Mutic, S., Zheng, J., Bradley, J.D., Grigsby, P., Deasy, J.O., 2007. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Med. Phys. 34, 4738–4749.

Erdi, Y.E., Mawlawi, O., Larson, S.M., Imbriaco, M., Yeung, H., Finn, R., Humm, J.L., 1997. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer 80, 2505–2509.

Ford, E.C., Herman, J., Yorke, E., Wahl, R.L., 2009. 18F-FDG PET/CT for image-guided and intensity-modulated radiotherapy. J. Nucl. Med.: Off. Pub. Soc. Nuclear Med. 50, 1655–1665.

Foster, B., Bagci, U., Mansoor, A., Xu, Z., Mollura, D.J., 2014. A review on segmentation of positron emission tomography images. Comput. Biol. Med. 50, 76–96.

Geets, X., Lee, J.A., Bol, A., Lonneux, M., Gregoire, V., 2007. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur. J. Nucl. Med. Mol. Imag. 34, 1427–1438.

Greco, C., Rosenzweig, K., Cascini, G.L., Tamburrini, O., 2007. Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC). Lung Canc. 57, 125–134.
Gregoire, V., Haustermans, K., Geets, X., Roels, S., Lonneux, M., 2007. PET-based treatment planning in radiotherapy: a new standard? J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 48 (Suppl 1), 68S–77S.
Hatt, M., Cheze-le Rest, C., van Baardwijk, A., Lambin, P., Pradier, O., Visvikis, D., 2011. Impact of tumor size and tracer uptake heterogeneity in (18)F-FDG PET and CT non-small cell lung cancer tumor delineation. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 52, 1690–1697.
Hatt, M., Cheze Le Rest, C., Albarghach, N., Pradier, O., Visvikis, D., 2011. PET functional volume delineation: a robustness and repeatability study. Eur. J. Nucl. Med. Mol. Imag. 38, 663–672.
Hatt, M., Cheze le Rest, C., Turzo, A., Roux, C., Visvikis, D., 2009. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans. Med. Imag. 28, 881–893.
Hong, R., Halama, J., Bova, D., Sethi, A., Emami, B., 2007. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. Int. J. Radiat. Oncol. Biol. Phys. 67, 720–726.
Huang, W., Zhou, T., Ma, L., Sun, H., Gong, H., Wang, J., Yu, J., Li, B., 2011. Standard uptake value and metabolic tumor volume of (1)(8)F-FDG PET/CT predict short-term outcome early in the course of chemoradiotherapy in advanced non-small cell lung cancer. Eur. J. Nucl. Med. Mol. Imag. 38, 1628–1635.
Jannin, P., Fitzpatrick, J.M., Hawkes, D.J., Pennec, X., Shahidi, R., Vannier, M.W., 2002. Validation of medical image processing in image-guided therapy. IEEE Trans. Med. Imag. 21, 1445–1449.
Jannin, P., Krupinski, E., Warfield, S., 2006. Validation in medical image processing. IEEE Trans. Med. Imag. 25, 1405–1409.
Jentzen, W., Freudenberg, L., Eising, E.G., Heinze, M., Brandau, W., Bockisch, A., 2007. Segmentation of PET volumes by iterative image thresholding. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 48, 108–114.
Kenny, L., Coombes, R.C., Vigushin, D.M., Al-Nahhas, A., Shousha, S., Aboagye, E.O., 2007. Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3'-deoxy-3'-[18F]fluorothymidine positron emission tomography. Eur. J. Nucl. Med. Mol. Imag. 34, 1339–1347.
Kessler, R.M., Ellis Jr., J.R., Eden, M., 1984. Analysis of emission tomographic scan data: limitations imposed by resolution and background. J. Comput. Assist. Tomogr. 8, 514–522.
Krak, N.C., Boellaard, R., Hoekstra, O.S., Twisk, J.W., Hoekstra, C.J., Lammertsma, A.A., 2005. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. Eur. J. Nucl. Med. Mol. Imag. 32, 294–301.
Le Thiec, M., Testard, A., Ferrer, L., Guillerminet, C., Morel, O., Maucherat, B., Rusu, D., Girault, S., Lacombe, M., Hamidou, H., Meyer, V.G., Rio, E., Hiret, S., Kraeber-Bodere, F., Campion, L., Rousseau, C., 2020. Prognostic impact of pretherapeutic FDG-PET in localized anal cancer. Cancers 12.
Lee, J.A., 2010. Segmentation of positron emission tomography images: some recommendations for target delineation in radiation oncology. Radiother. Oncol.: J. Eur. Soc. Ther. Radiol. Oncol. 96, 302–307.
Leung, K.H., Marashdeh, W., Wray, R., Ashrafinia, S., Pomper, M.G., Rahmim, A., Jha, A.K., 2020. A physics-guided modular deep-learning based automated framework for tumor segmentation in PET. Phys. Med. Biol.
Maroy, R., Boisgard, R., Comtat, C., Frouin, V., Cathier, P., Duchesnay, E., Dolle, F., Nielsen, P.E., Trebossen, R., Tavitian, B., 2008. Segmentation of rodent whole-body dynamic PET images: an unsupervised method based on voxel dynamics. IEEE Trans. Med. Imag. 27, 342–354.
Matheoud, R., Della Monica, P., Loi, G., Vigna, L., Krengli, M., Inglese, E., Brambilla, M., 2011. Influence of reconstruction settings on the performance of adaptive thresholding algorithms for FDG-PET image segmentation in radiotherapy planning. J. Appl. Clin. Med. Phys. 12, 3363.
McKay, M.J., Taubman, K.L., Foroudi, F., Lee, S.T., Scott, A.M., 2018. Molecular imaging using PET/CT for radiation therapy planning for adult cancers: current status and expanding applications. Int. J. Radiat. Oncol. Biol. Phys. 102, 783–791.
Messa, C., Di Muzio, N., Picchio, M., Gilardi, M.C., Bettinardi, V., Fazio, F., 2006. PET/CT and radiotherapy. In: The Quarterly Journal of Nuclear Medicine and Molecular Imaging, 50. official publication of the Italian Association of Nuclear Medicine, pp. 4–14.
Moureau-Zabotto, L., Touboul, E., Lerouge, D., Deniaud-Alexandre, E., Grahek, D., Foulquier, J.N., Petegnief, Y., Gres, B., El Balaa, H., Kerrou, K., Montravers, F., Keraudy, K., Tiret, E., Gendre, J.P., Grange, J.D., Houry, S., Talbot, J.N., 2005. Impact of CT and 18F-deoxyglucose positron emission tomography image fusion for conformal radiotherapy in esophageal carcinoma. Int. J. Radiat. Oncol. Biol. Phys. 63, 340–345.
Nahmias, C., Wahl, L.M., 2008. Reproducibility of standardized uptake value measurements determined by 18F-FDG PET in malignant tumors. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 49, 1804–1808.
Nehmeh, S.A., El-Zeftawy, H., Greco, C., Schwartz, J., Erdi, Y.E., Kirov, A., Schmidtlein, C.R., Gyau, A.B., Larson, S.M., Humm, J.L., 2009. An iterative technique to segment PET lesions using a Monte Carlo based mathematical model. Med. Phys. 36, 4803–4809.
Nestle, U., Kremp, S., Schaefer-Schuler, A., Sebastian-Welsch, C., Hellwig, D., Rube, C., Kirsch, C.M., 2005. 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. : Off. Pub. Soc. Nuclear Med. 46, 1342–1348.
Nestle, U., Schaefer-Schuler, A., Kremp, S., Groeschel, A., Hellwig, D., Rube, C., Kirsch, C.M., 2007. Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. Eur. J. Nucl. Med. Mol. Imag. 34, 453–462.
Paulino, A.C., Koshy, M., Howell, R., Schuster, D., Davis, L.W., 2005. 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. 61, 1385–1392.
Pfaehler, E., Burggraaff, C., Kramer, G., Zijlstra, J., Hoekstra, O.S., Jalving, M., Noordzij, W., Brouwers, A.H., Stevenson, M.G., de Jong, J., Boellaard, R., 2020. PET segmentation of bulky tumors: strategies and workflows to improve inter-observer variability. PloS One 15, e0230901.
Raynor, W.Y., Jonnakuti, V.S., Zirakchian Zadeh, M., Werner, T.J., Cheng, G., Zhuang, H., Hoilund-Carlsen, P.F., Alavi, A., Baker, J.F., 2019. Comparison of methods of quantifying global synovial metabolic activity with FDG-PET/CT in rheumatoid arthritis. Int. J. Rheumatic Dis. 22, 2191–2198.
Reader, A.J., Zweit, J., 2001. Developments in whole-body molecular imaging of live subjects. Trends Pharmacol. Sci. 22, 604–607.
Sbei, A., ElBedoui, K., Barhoumi, W., Maktouf, C., 2020. Gradient-based generation of intermediate images for heterogeneous tumor segmentation within hybrid PET/MRI scans. Comput. Biol. Med. 119, 1–15.
Schaefer, A., Kremp, S., Hellwig, D., Rube, C., Kirsch, C.M., Nestle, U., 2008. 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. Imag. 35, 1989–1999.
Schinagl, D.A., Vogel, W.V., Hoffmann, A.L., van Dalen, J.A., Oyen, W.J., Kaanders, J.H., 2007. 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. 69, 1282–1289.
Schwartz, D.L., Ford, E.C., Rajendran, J., Yueh, B., Coltrera, M.D., Virgin, J., Anzai, Y., Haynor, D., Lewellen, B., Mattes, D., Kinahan, P., Meyer, J., Phillips, M., Leblanc, M., Krohn, K., Eary, J., Laramore, G.E., 2005. FDG-PET/CT-guided intensity modulated head and neck radiotherapy: a pilot investigation. Head Neck 27, 478–487.
Soret, M., Bacharach, S.L., Buvat, I., 2007. Partial-volume effect in PET tumor imaging. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 48, 932–945.
Stroom, J., Blaauwgeers, H., van Baardwijk, A., Boersma, L., Lebesque, J., Theuws, J., van Suylen, R.J., Klomp, H., Liesker, K., van Pel, R., Siedschlag, C., Gilhuijs, K., 2007. Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumors. Int. J. Radiat. Oncol. Biol. Phys. 69, 267–275.
Tamal, M., 2019. A fully automatic global gradient measure based 3D region growing solid tumour segmentation method (3D-GGM-RG) for low contrast and low count positron emission tomography. J. Med. Imag. Health Info. 9, 2022–2030.
Tamal, M., 2019. A hybrid region growing tumour segmentation method for low contrast and high noise Nuclear Medicine (NM) images by combining a novel non-linear diffusion filter and global gradient measure (HNDF-GGM-RG). Heliyon 5, e02993.
Therasse, P., Arbuck, S.G., Eisenhauer, E.A., Wanders, J., Kaplan, R.S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A.T., Christian, M.C., Gwyther, S.G., 2000. New guidelines to evaluate the response to treatment in solid tumors. European organization for Research and treatment of cancer, national cancer institute of the United States, national cancer institute of Canada. J. Natl. Cancer Inst. 92, 205–216.
Townsend, D.W., 2008. Multimodality imaging of structure and function. Phys. Med. Biol. 53, R1–R39.
Tylski, P., Stute, S., Grotus, N., Doyeux, K., Hapdey, S., Gardin, I., Vanderlinden, B., Buvat, I., 2010. Comparative assessment of methods for estimating tumor volume and standardized uptake value in (18)F-FDG PET. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 51, 268–276.
Valk, P.E., Bailey, D.L., Townsend, D.W., Maisey, M.N., 2004. Positron Emission Tomography: Basic Science and Clinical Practice. Springer.
van Baardwijk, A., Bosmans, G., Boersma, L., Buijsen, J., Wanders, S., Hochstenbag, M., van Suylen, R.J., Dekker, A., Dehing-Oberije, C., Houben, R., Bentzen, S.M., van Kroonenburgh, M., Lambin, P., De Ruysscher, D., 2007. 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. 68, 771–778.
van Dalen, J.A., Hoffmann, A.L., Dicken, V., Vogel, W.V., Wiering, B., Ruers, T.J., Karssemeijer, N., Oyen, W.J., 2007. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. Nucl. Med. Commun. 28, 485–493.
Vanderhoek, M., Perlman, S.B., Jeraj, R., 2012. Impact of the definition of peak standardized uptake value on quantification of treatment response. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 53, 4–11.
Vees, H., Senthamizhchelvan, S., Miralbell, R., Weber, D.C., Ratib, O., Zaidi, H., 2009. Assessment of various strategies for 18F-FET PET-guided delineation of target volumes in high-grade glioma patients. Eur. J. Nucl. Med. Mol. Imag. 36, 182–193.
Wahl, R.L., Jacene, H., Kasamon, Y., Lodge, M.A., 2009. From RECIST to PERCIST: evolving Considerations for PET response criteria in solid tumors. J. Nucl. Med. : Off. Pub. Soc. Nuclear Med. 50 (Suppl 1), 122S–150S.
Wanet, M., Lee, J.A., Weynand, B., De Bast, M., Poncelet, A., Lacroix, V., Coche, E., Gregoire, V., Geets, X., 2011. Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: a comparison with threshold-based approaches, CT

and surgical specimens. Radiother. Oncol. : J. Eur. Soc. Ther. Radiol. Oncol. 98, 117–125.
Weber, W.A., 2005. PET for response assessment in oncology: radiotherapy and chemotherapy. Br. J. Radiol. 28 (Suppl), 42–49.
Werner-Wasik, M., Nelson, A.D., Choi, W., Arai, Y., Faulhaber, P.F., Kang, P., Almeida, F.D., Xiao, Y., Ohri, N., Brockway, K.D., Piper, J.W., Nelson, A.S., 2012. What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int. J. Radiat. Oncol. Biol. Phys. 82, 1164–1171.
Yu, W., Fu, X.L., Zhang, Y.J., Xiang, J.Q., Shen, L., Jiang, G.L., Chang, J.Y., 2009. GTV spatial conformity between different delineation methods by 18FDG PET/CT and pathology in esophageal cancer. Radiother. Oncol. : J. Eur. Soc. Ther. Radiol. Oncol. 93, 441–446.
Zaidi, H., El Naqa, I., 2010. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. Eur. J. Nucl. Med. Mol. Imag. 37, 2165–2187.