

Semiautomatic methods for segmentation of the proliferative tumour volume on sequential FLT PET/CT images in head and neck carcinomas and their relation to clinical outcome

Anne I. J. Arens · Esther G. C. Troost · Bianca A. W. Hoeben · Willem Grootjans · John A. Lee · Vincent Grégoire · Mathieu Hatt · Dimitris Visvikis · Johan Bussink · Wim J. G. Oyen · Johannes H. A. M. Kaanders · Eric P. Visser

Received: 25 June 2013 / Accepted: 28 November 2013 / Published online: 18 December 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract

Purpose Radiotherapy of head and neck cancer induces changes in tumour cell proliferation during treatment, which can be depicted by the PET tracer ^{18}F -fluorothymidine (FLT). In this study, three advanced semiautomatic PET segmentation methods for delineation of the proliferative tumour volume (PV) before and during (chemo)radiotherapy were compared and related to clinical outcome.

Methods The study group comprised 46 patients with 48 squamous cell carcinomas of the head and neck, treated with accelerated (chemo)radiotherapy, who underwent FLT PET/CT prior to treatment and in the 2nd and 4th week of therapy. Primary gross tumour volumes were visually delineated on

CT images (GTV_{CT}). PVs were visually determined on all PET scans (PV_{VIS}). The following semiautomatic segmentation methods were applied to sequential PET scans: background-subtracted relative-threshold level (PV_{RTL}), a gradient-based method using the watershed transform algorithm and hierarchical clustering analysis ($\text{PV}_{\text{W\&C}}$), and a fuzzy locally adaptive Bayesian algorithm (PV_{FLAB}).

Results Pretreatment PV_{VIS} correlated best with PV_{FLAB} and GTV_{CT} . Correlations with PV_{RTL} and $\text{PV}_{\text{W\&C}}$ were weaker although statistically significant. During treatment, the PV_{VIS} , $\text{PV}_{\text{W\&C}}$ and PV_{FLAB} significant decreased over time with the steepest decline over time for PV_{FLAB} . Among these advanced segmentation methods, PV_{FLAB} was the most robust in segmenting volumes in the third scan (67 % of tumours as compared to 40 % for $\text{PV}_{\text{W\&C}}$ and 27 % for PV_{RTL}). A decrease in PV_{FLAB} above the median between the pretreatment scan and the scan obtained in the 4th week was associated with better disease-free survival (4 years 90 % versus 53 %).

Conclusion In patients with head and neck cancer, FLAB proved to be the best performing method for segmentation of the PV on repeat FLT PET/CT scans during (chemo)radiotherapy. This may potentially facilitate radiation dose adaptation to changing PV.

Keywords ^{18}F -Fluorothymidine PET · Head and neck cancer · Delineation · Radiotherapy · Outcome

Anne I. J. Arens and Esther G. C. Troost contributed equally to this work.

A. I. J. Arens (✉) · W. Grootjans · W. J. G. Oyen · E. P. Visser
Department of Nuclear Medicine, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
e-mail: anne.arenas@radboudumc.nl

E. G. C. Troost · B. A. W. Hoeben · J. Bussink ·
J. H. A. M. Kaanders
Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, The Netherlands

E. G. C. Troost
MAASTRO clinic, GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

J. A. Lee · V. Grégoire
Department of Radiation Oncology, Université Catholique de Louvain, St-Luc University Hospital, Brussels, Belgium

M. Hatt · D. Visvikis
INSERM UMR1101, Laboratoire de Traitement de l'Information Médicale (LaTIM), Brest, France

Introduction

In patients with advanced stage squamous cell carcinoma of the head and neck (HNSCC), radiotherapy with or without concomitant chemotherapy is commonly the treatment of

choice. Although accurate delineation of the gross tumour volume (GTV) is pivotal for high precision radiation therapy, there is significant uncertainty in tumour delineation using anatomical imaging (i.e. CT and MRI) due to interobserver and intraobserver variability [1]. HNSCC displays a number of characteristics that may adversely affect treatment outcome including intrinsic radiation sensitivity, tumour cell hypoxia and tumour cell repopulation during radiotherapy [2, 3]. The last of these tumour resistance mechanisms may be counteracted by concurrently administering targeted therapy such as epidermal growth factor receptor (EGFR) inhibitors, or by accelerating radiation schedules [4, 5]. Alternatively, enhanced dose delivery to the proliferative tumour subvolume may compensate for prolonged overall treatment times in conventional once-daily radiation schedules. Ideally, one would like to monitor tumour proliferation during treatment to steer counteractive measures based on the repopulative response.

PET is a noninvasive imaging modality capable of depicting functional characteristics of the entire tumour volume before and during treatment. ^{18}F -Fluorothymidine (FLT), a thymidine analogue, is a PET tracer that reflects cellular proliferation [6]. In contrast to ^{18}F -fluorodeoxyglucose (FDG), FLT does not accumulate in inflammatory tissue [7], which is frequently found in/near primary tumours of the head and neck or is induced during the course of (chemo)radiotherapy. It has previously been demonstrated that FLT accumulates in HNSCC, and that uptake diminishes significantly during therapy, preceding morphological changes [3, 8–10]. Similarly, FLT uptake in HNSCC has been shown to result in reproducible standardized uptake values (SUVs) before and during treatment [11, 12], and is therefore suitable for early response monitoring [12]. Thus, FLT PET may help identify highly proliferative tumour (sub)volumes, and select patients eligible for individual treatment modifications, ultimately leading to improved treatment outcome.

In order to transfer tailored therapy approaches from a research setting to clinical practice, the proliferative tumour volume (PV) should be accurately and objectively delineated. Several semiautomatic segmentation methods for PET functional volume determination have been developed over the last few years. Most of them have been validated using synthetic/simulated datasets, simplistic phantom acquisitions and visual tumour delineation, and occasionally against histopathological specimens [13–17]. In addition, the majority of clinical evaluation studies using such algorithms have considered FDG tumour imaging.

The goal of the present study was to compare three semiautomatic PET segmentation methods for derivation of PV in primary HNSCC on sequential FLT PET images before and during (chemo)radiotherapy. The abilities of these methods to deliver usable and coherent PV segmentations were assessed and compared with that

of visual delineation as a reference. Furthermore, segmented PVs were related to clinical outcome.

Materials and methods

Patients

From July 2006 until September 2008, eligible patients aged 18 years or older with newly diagnosed UICC stage II – IV HNSCC, treated with (chemo)radiotherapy with curative intent, were included in the study. Exclusion criteria were surgery as primary tumour therapy, palliative treatment and pregnancy. The Institutional Review Board approved the study and written informed consent was obtained from all patients. To evaluate clinical outcome, patients were followed until death or otherwise for at least 3 years. Data were censored on 31 December 2011.

Treatment

Eighteen patients were treated with three-dimensional conformal radiotherapy delivering a dose of 68 Gy in 2-Gy fractions to the primary tumour and metastatic cervical lymph nodes, and 44 Gy to electively treated nodal levels. The other patients were treated with intensity-modulated radiation therapy using a simultaneous integrated boost technique, delivering a dose of 68 Gy in 2-Gy fractions to the primary tumour and metastatic cervical lymph nodes, and 50.3 Gy in 1.48-Gy fractions to electively treated nodal levels. In all patients, an accelerated fractionation schedule was used with an overall treatment time of 5.5 weeks, delivering two fractions daily during the last 1.5 weeks of treatment. In accordance with institutional guidelines, 14 patients with cT3 and cT4 tumours, without contraindications and age below 70 years, were concomitantly treated with intravenous cisplatin at 40 mg/m² once weekly.

FLT PET/CT imaging

FLT was obtained from the Department of Nuclear Medicine and PET Research, Free University Medical Center, Amsterdam, The Netherlands. The procedure for FLT synthesis and purification has been described elsewhere [3, 18]. FLT PET/CT of the head and neck region was performed using a PET/CT scanner (Biograph Duo; Siemens/CTI, Munich, Germany). All patients received a FLT PET/CT scan prior to treatment, and a second FLT PET/CT scan was performed in 44 patients followed by a third scan in 28 patients. The first 14 patients accrued to the study did not receive a third FLT PET/CT scan because this was added in an amendment to the study protocol authorized by the Institutional Review Board 7 months after initiation of the study. All scans were acquired with the patient in the supine position, immobilized with an

individual head support and a rigid customized mask to increase positioning accuracy and to prevent movement during image acquisition. Emission images of the head and neck area were recorded 60 min after intravenous injection of approximately 250 MBq of FLT. Two bed positions were recorded with a 7-min acquisition time per bed position. PET images were reconstructed using the ordered-subsets expectation maximization iterative algorithm with parameters optimized for the head and neck region (four iterations, 16 subsets, and a 5-mm FWHM, three-dimensional gaussian filter) and a matrix size of $128 \times 128 \times 94$. Correction for radioactive decay to the start of the scan was applied to the PET images. Simultaneously, an intravenous contrast-enhanced (Optiray 300; Tyco Healthcare/Mallinckrodt Inc., Santa Monica, CA) CT scan for anatomic correlation and PET attenuation correction was acquired (120 mAs, 130 kV, rotation time 1.0 s, slice collimation 2.5 mm, and feed 4.3 mm).

Image analysis

PET and CT images were imported into the radiotherapy planning system Pinnacle³ (version 8.0d; Philips Radiation Oncology Systems, Andover, MA). With this software, consecutive CT and PET scans were resampled and registered to the first CT scan using cross-correlation and rigid transformation. Tumour SUV_{peak} was defined as the mean SUV of the hottest voxel in the tumour and its eight surrounding voxels in one transaxial slice. Two experienced radiation oncologists delineated the GTV of the primary tumour on the CT scans in consensus using all available clinical and anatomical information. The PVs on consecutive FLT PET images were visually delineated (PV_{VIS}) by a trained radiation oncologist. In previous studies, the value of SUV_{peak} and two commonly used segmentation methods, the 50 % isocontour of SUV_{peak} and signal-to-background ratio (SBR) [13] methods, have already been addressed in patients in this cohort and were therefore not reassessed here [3, 9]. The following three semiautomatic segmentation methods for determining the PV were applied.

The first was the iterative background-subtracted relative-threshold level (RTL) method using Pinnacle³ scripts (PV_{RTL}). This model-based method provides threshold levels independent of the SBR and is based on the convolution by the scanner's point spread function (PSF) of a sphere that takes into account differences in sphere size and scanner resolution [14]. Thresholds for RTL were iteratively derived using the adopted PV_{VIS} and a manually defined background ROI in the neck musculature (about 10 cm^3) at a sufficient distance from the vertebrae, the primary tumour, and lymph node metastases. This iterative algorithm was applied until the threshold level changes were less than 0.1 %. In the second method, the PV derived from the watershed transform and

hierarchical clustering (W&C; $PV_{W\&C}$) segmentation was created using the Matlab-based program IMRE viewer (Université Catholique de Louvain, Belgium) [15]. This is a gradient-based method which allows better estimation of the gradient intensity by first denoising with an edge-preserving filter and partial volume effect correction with a constrained iterative deconvolution algorithm before gradient-based delineation is applied. A PSF with a FWHM of 8 mm was chosen to deconvolve the images. Visual inspection of the segmented volumes allowed more accurate positioning of the volume of interest in which the algorithm performed the segmentation. In the final method, the PV derived from the fuzzy locally adaptive Bayesian (FLAB) algorithm (PV_{FLAB}) was obtained using specifically designed software (ImageD; LaTIM INSERM laboratory, Brest, France) [19]. This approach is based on computing the probability that each voxel belongs to a given class (i.e. tumour or background) by taking into account the intensity with respect to the statistical distributions of the voxels in the various regions of the image, as well as spatial correlation with neighbouring voxels. In cases where the automatic FLAB algorithm initialization procedure led to unsatisfactory results on visual inspection, the user exploited the available option of manually adjusting the initialization parameters (mean and/or standard deviation of tumour and/or background uptake) and restarting the automatic FLAB volume estimation process.

Images were not segmented when the radiation oncologist and nuclear medicine physician were unable to identify any FLT accumulation within known lesions. Before accepting the results, both medical specialists visually inspected the segmented PV obtained by the semiautomatic algorithms. Segmentation was deemed successful if it adhered to the anatomical boundaries of the primary tumour as defined in the first CT scan, and did not delineate a large area of evidently normal tissue.

Statistical analysis

Differences between groups were tested using linear mixed model analysis (random intercept patient effect, fixed method effects). Using untransformed PV as dependent variable led to relatively large skewness (6.38) of the residuals. Therefore, the volumes were logarithmically transformed, resulting in a considerably smaller skewness of 0.96, enabling parametric analysis. Correlations were determined using the Pearson test.

Median follow-up was determined using the inverse Kaplan-Meier method. The changes in GTVs based on CT and PVs obtained by FLT PET during treatment were correlated with long-term clinical outcome. Four-year locoregional control (LRC), disease-free survival (DFS) and overall survival (OS) (\pm standard error of the mean) were analysed using Kaplan-Meier estimates and compared using the log-rank test.

LRC was defined as no recurrence of primary tumour or cervical lymphadenopathy at the date of last follow-up. DFS was defined as freedom from locoregional or distant recurrence at the date of last follow-up. OS was defined as the time to the date of death from any cause, or patients were censored at the date of last follow-up. Survival parameters were related to the date of the first PET/CT scan. Continuous variables were dichotomized at the median. Two-sided p values <0.05 were considered to indicate significance.

Statistical analyses were performed using SAS version 9.2 (SAS institute Inc., Cary, NC) and GraphPad Prism version 5.04 (GraphPad Software, La Jolla, CA).

Results

Patient characteristics

The study accrued 52 patients with in total 54 primary HNSC C. Six patients were excluded from the present analysis, two due to technical difficulties and four because they were treated in a clinical phase III trial incorporating modification of the tumour oxygenation status [20] which deviated from the standard institutional treatment guideline. Therefore, the final population eligible for analysis consisted of 46 patients undergoing (chemo)radiotherapy with 48 HNSCCs in total. Their average age was 60 years at enrolment in the study, ranging from 39 to 75 years. Median follow-up was 52 months (range 8 – 65 months). The pretreatment FLT PET/CT scan was performed at a median of 4 days (range 0 – 14) before the start of treatment in all patients, the second scan in the 2nd week (after a median of 7 fractions, range 4 – 12 fractions) in 44 patients with 46 tumours, and the third scan in the 4th week of (chemo)radiotherapy (after a median of 18 fractions, range 15 – 19 fractions) in 28 patients with 30 tumours. Patient and tumour characteristics, as well as events are presented in Table 1.

Comparison of CT-delineated GTVs and PET-segmented PVs on pretreatment scans

Figure 1 shows the CT-delineated GTVs and the PET-segmented PVs on the pretreatment scans. The mean PVs ranged from 9.4 to 14.4 cm³. All segmentation methods resulted in statistically significantly different PVs ($p = 0.003$). PV_{VIS} was the largest (mean volume \pm standard deviation, 14.4 \pm 12.7 cm³) and correlated best with PV_{FLAB} and GTV_{CT} (Pearson correlation coefficient 0.85, $p < 0.0001$; and 0.88, $p < 0.00001$, respectively). Correlations of PV_{VIS} with PV_{RTL} and PV_{W&C} were weaker, although still statistically significant (Pearson correlation coefficient 0.77, $p < 0.0001$, for PV_{RTL}; 0.77, $p < 0.0001$, for PV_{W&C}).

Table 1 Patient and tumour characteristics, as well as events

Characteristic	No. (%) of patients
Gender	
Female	9 (20)
Male	37 (80)
T stage	
1	1 (2)
2	25 (52)
3	15 (31)
4	7 (15)
N stage	
0	21 (44)
1	6 (12)
2	21 (44)
3	–
UICC stage	
II	12 (25)
III	12 (25)
IVA	24 (50)
Histological grade	
1	4 (8)
2	24 (50)
3	11 (23)
Unknown	9 (19)
Primary tumour site	
Oral cavity	1 (2)
Oropharynx	25 (52)
Larynx	14 (29)
Hypopharynx	8 (17)
Treatment	
Radiotherapy	32 (70)
Concurrent chemoradiotherapy	14 (30)
Events	
Locoregional recurrence	8 (17)
Distant metastases	5 (11)
Tumour-related death	8 (17)
Death due to other cause	11 (24)

Delineation of PVs on repeat FLT PET/CT scans

Previous studies have shown that FLT SUV_{peak} decreases significantly during treatment [3, 9]. CT volumes did not show a significant decrease until the 4th week of treatment (mean decrease 1.6 mL/7 % between the first and second scan, $p = 0.1$; mean decrease 4.8 mL/41 % between the first and third scan, $p < 0.0001$). Regarding PVs, PV_{VIS} showed a significant decrease between one FLT PET/CT scan and the next ($p < 0.0001$; Fig. 1). Changes in PV_{RTL} were small and not statistically significant (overall difference between scans, $p = 0.33$). PV_{W&C} and PV_{FLAB} both showed a significant

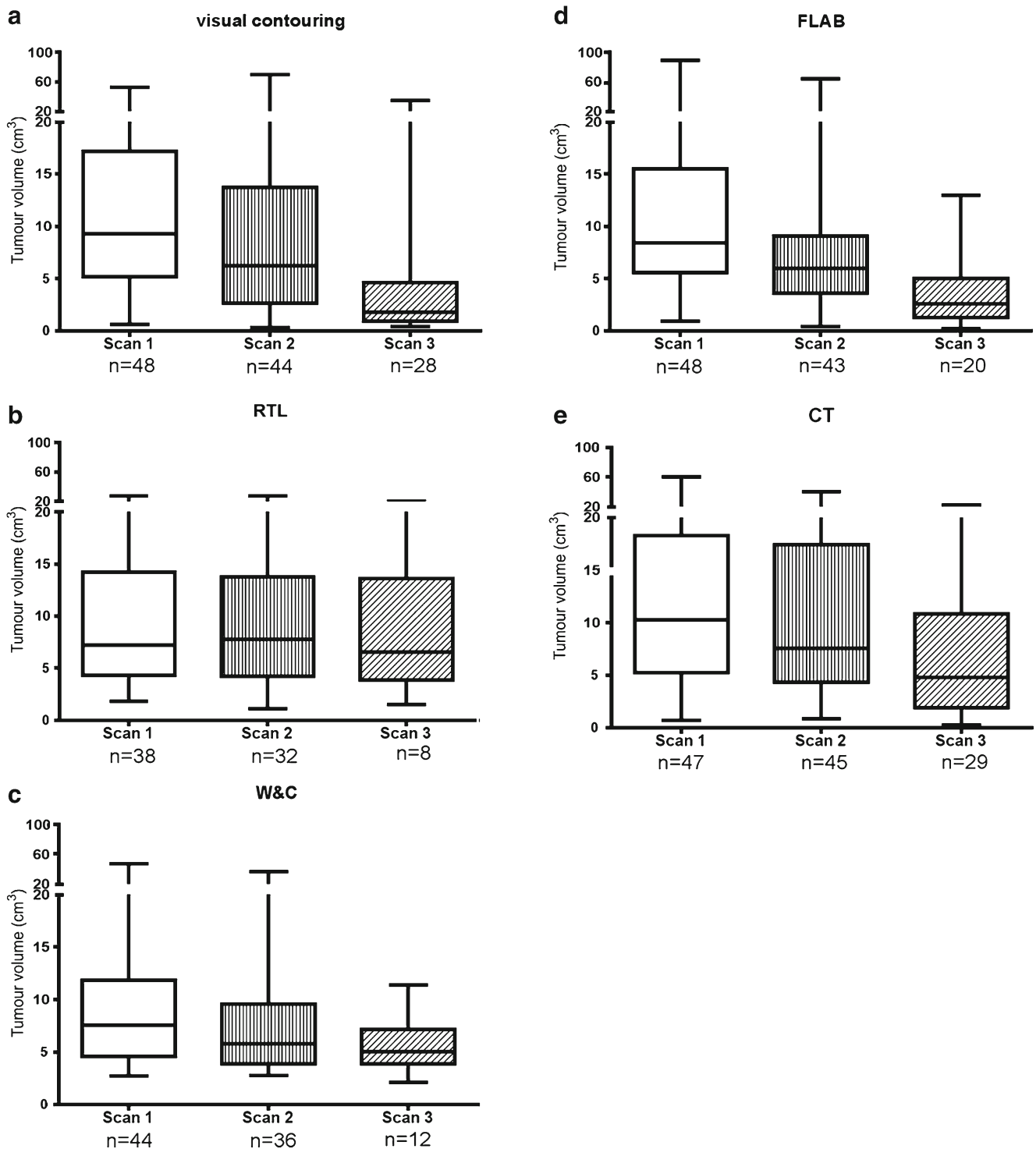


Fig. 1 Plots of successfully delineated PVs from segmentation methods on consecutive FLT PET scans: **a** visual delineation, **b** RTL, **c** W&C and **d** FLAB (**e** CT volumes per scan). The bottom and top of each box

represent the 25 % and 75 % percentiles. The horizontal line within the boxes is the median. The lower and higher extremes of the whiskers represent the minimum and maximum values

reduction, with a steeper decrease in PV_{FLAB} than in $PV_{W\&C}$ ($p < 0.0001$ for decreases in PV_{FLAB} between all scans, and in $PV_{W\&C}$ between the first and second scan; and $p = 0.16$ for decrease in $PV_{W\&C}$ between the second and third scan).

As illustrated in Figs. 2 and 3, the RTL, W&C and FLAB methods showed a considerable decrease in the percentage of PV delineations deemed successful on the second and third scan. FLAB demonstrated a higher success rate than RTL and

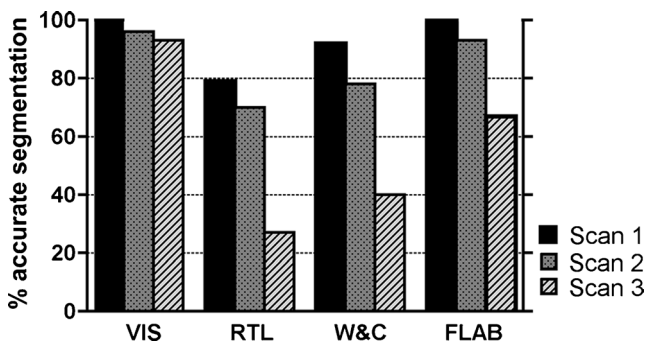


Fig. 2 Rates of accurate tumour segmentation using the three segmentation methods RTL, W&C and FLAB as a percentage of all PET/CT scans performed in comparison with visual delineation on PET images

W&C (e.g. in the third scan, 67 % of tumours as compared to 40 % for $PV_{W\&C}$ and 27 % for PV_{RTL}).

Influence of the primary tumour site and the treatment regimen on PV

Since the delineation of oropharyngeal tumours on FLT PET scans may be hampered by proliferative activity in nearby lymphoid tissues [21], the influence of the tumour site on the performance of the segmentation methods was assessed. On the pretreatment scans, there was no statistically significant difference in the proportion of PV delineations deemed successful between the various segmentation methods when comparing oropharyngeal and oral cavity tumours with HNSCC at other sites ($p=0.36$). However, the percentage of segmentations deemed successful during treatment was in general lower for tumours in the oropharynx and oral cavity than in tumours

at other sites (Table 2). The change in PV during (chemo)radiotherapy was not related to the primary tumour site (data not shown).

There were no significant differences in PV changes over time between the patients receiving radiotherapy and those receiving chemoradiotherapy. The number of successful segmentations on the second scans was comparable between the treatment groups. No conclusions could be drawn regarding the percentage of accurate segmentations on the third scans in relation to treatment due to the small number of patients receiving chemoradiotherapy (Table 3).

Correlation between semiautomatically segmented PV changes in sequential FLT PET scans and clinical outcome

Changes in the CT-based GTVs during (chemo)radiotherapy were not statistically significantly correlated with treatment outcome. Due to the low number of successful segmentations during the course of (chemo)radiotherapy, PV_{RTL} was not related to treatment outcome. Changes in $PV_{W\&C}$ between pretreatment and the subsequent two scans were not significantly predictive of outcome. For PV_{FLAB} , a median decrease between the pretreatment scan and the scan acquired in the 4th week of therapy of more than 7.39 cm^3 was predictive of a better 4-year DFS (90 ± 9.5 versus 53 ± 17.6 %, $p=0.040$; Fig. 4). Significant correlations were not found between PV_{FLAB} obtained at 2 weeks and outcome parameters, nor between PV_{FLAB} obtained at 4 weeks and LCR or OS at 4 years (results not shown).

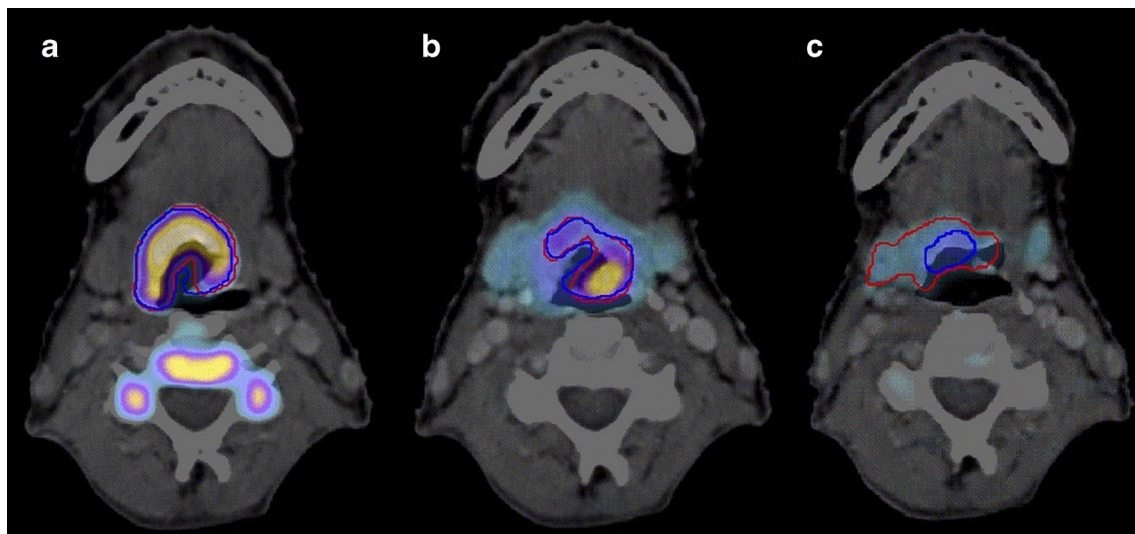


Fig. 3 PV_{RTL} (red) and PV_{FLAB} (blue) on the scans before treatment (a), and after 7 fractions (b) and 17 fractions (c) of radiotherapy in a patient with a cT4N0M0 oropharyngeal tumour who was treated with radiotherapy only. When FLT uptake is evidently higher in the tumour than in surrounding tissues, both segmentation algorithms give approximately

the same segmentation (a, b). As the tumour FLT signal approaches background levels (c) and the proliferating tumour volume decreases significantly, the RTL algorithm produces a delineation that clearly expands beyond the initial tumour boundaries (as delineated on the primary CT scan), and was thus termed “unsuccessful”

Table 2 Percentages of PV segmentations by RTL, W&C and FLAB deemed successful on consecutive scans in patients with oropharyngeal and oral cavity tumours versus HNSCC at other sites

Scan	Segmentation method	Site of HNSCC			
		Oropharynx and oral cavity		Other	
		Successful segmentation (%)	No. of patients	Successful segmentation (%)	No. of patients
1	RTL	73	26	86	22
	W&C	88		91	
	FLAB	100		100	
2	RTL	65	26	75	20
	W&C	77		80	
	FLAB	88		100	
3	RTL	33	15	20	15
	W&C	20		60	
	FLAB	67		80	

Discussion

In this study in patients with advanced HNSCC, of three semiautomated PV segmentation methods, the highest percentage of delineations deemed successful was obtained using the FLAB algorithm. The reduction in PV_{FLAB} during therapy was also correlated with clinical outcome.

Common automated segmentation methods using fixed thresholds such as the 50 % isocontour of the SUV_{peak}, are easy to use but subject to criticism because the threshold is fixed arbitrarily. These methods are sensitive to contrast

Table 3 Comparison of percentages of PV segmentations by RTL, W&C and FLAB deemed successful on the consecutive scans in patients treated with radiotherapy versus chemoradiotherapy

Scan	Segmentation method	Treatment			
		Radiotherapy		Chemoradiotherapy	
		Successful segmentation (%)	No. of patients	Successful segmentation (%)	No. of patients
1	RTL	82	34	71	14
	W&C	91		93	
	FLAB	100		100	
2	RTL	68	34	67	12
	W&C	74		92	
	FLAB	91		100	
3	RTL	33	24	0	6
	W&C	38		50	
	FLAB	71		50	

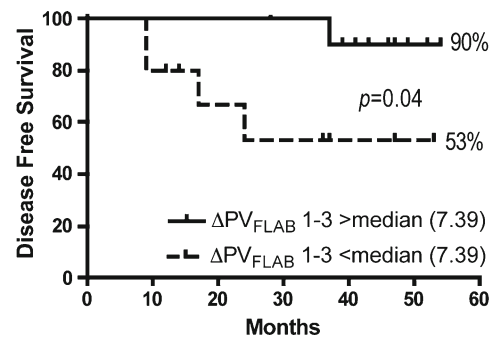


Fig. 4 Patients with a PV_{FLAB} decrease above the median between the baseline scan and the scan in the 4th week of (chemo)radiotherapy demonstrated a better DFS than patients with a PV_{FLAB} decrease below the median

variations as well as to noise [19, 22]. In a recent study [9], we showed that PV_{50%} leads to grossly overestimated functional volumes on repeat imaging, and therefore this method was dismissed from in-depth analysis. The same holds true for the adaptive threshold segmentation method SBR developed by Daisne et al. [13]. The decrease in signal-to-background contrast during treatment hampers this segmentation method. The RTL method, being independent of the tumour-to-background ratio because of background subtraction, and being based on tumour radius and image resolution only [14, 23], was the least robust advanced semiautomatic segmentation method in this study, especially for the third scan. On visual inspection of the results, the segmented PV_{RTL} encompassed normal vital structures, such as vertebrae, vessels, submandibular gland, tongue and soft palate, and thus the resulting segmentations were considered unsuccessful. An example of this can be seen in Fig. 3. These unsuccessful segmentations were probably due to a reduced performance of RTL when PVs became smaller than the FWHM of the PSF of the PET/CT scanner during (chemo)radiotherapy, and to a decrease in contrast between tumour and surrounding tissue on repeated PET/CT scans [14].

Also with the gradient-based W&C method, the number of successful segmentations diminished considerably, although the overall success rate was higher than with the RTL method. “Semiautomatic” corrections imply that the volume of interest for the algorithm can be changed and small changes in the binning step can be made. As with the previously discussed methods, the dissociation between low FLT uptake in PET-positive volumes and the background on pretreatment scans hampered the accuracy of this method. The uptake registered as background uptake might even be increased with the W&C method because of preprocessing steps. Notably, the gradient-based methodologies have been validated for accuracy against phantom images [15, 24] as well as histopathology in head and neck tumours [15] and lung tumours [25, 26]. They have also been applied to clinical datasets to derive image-based parameters of clinical prognostic value [27, 28]. However, the

W&C method was validated using FDG datasets only, and in this study using repeat FLT PET scans, performance of this method decreased during therapy.

The FLAB segmentation method resulted in the highest percentage of segmentations deemed successful and the delineated tumour volumes significantly decreased over time. However, it should be emphasized that this high success rate resulted from manual adjustment of the input parameters (mean and/or standard deviation of tumour and/or background uptake); this user-dependent operation was performed in 6 % of the first, 28 % of the second and 55 % of the third scans. The possibility to manually adjust the input parameters during segmentation thus seems to be an advantage over the other semiautomatic segmentation methods, but also induces operator dependency. FLAB is less sensitive to image noise than the other segmentation approaches tested here. Its robustness, reproducibility and repeatability have been assessed for (sequential) FDG and pretreatment FLT PET imaging of primary tumours at different sites [29–33]. Furthermore, a number of FDG image-derived parameters obtained with the FLAB algorithm have been proven to be of prognostic and predictive value [33, 34]. To date, no FLAB volume measurements on sequential FLT PET images have been performed.

A limitation of this study was that the results of the segmentations could not be validated against histopathological resection specimens, as this treatment regimen was chosen for organ preservation purposes. This also precluded immunohistochemical verification of proliferation decreases during therapy as found with the consecutive PET/CT scans. However, a previous study has validated the use of FLT against immunohistochemical proliferation markers for primary HNSCC [35]. Since we could not validate the segmentation methods against a ground-truth, we compared the semiautomated segmentations with manually delineated PVs. PV_{VIS} is highly interoperator and intraoperator dependent and reliable application requires a learning curve and a standardized approach such as, for example, PET image window level settings. However, when standards are applied this method can be used for qualitative assessment of the capacity of automated algorithms to delineate a PV; for example, if an intratumoral volume with an FLT uptake level above that of the background and of surrounding tissues is discerned qualitatively. An “ideal” automated segmentation algorithm would be capable of segmenting a PV within the physical boundaries of the tumour. If the tested method proves unable to achieve this, it may be assessed as “unsuccessful”.

In an earlier analysis of the current patient cohort, we observed that a decrease in SUV_{peak} of 45 % or more or a decrease in PV_{VIS} of 41 % (median) or more in the second treatment week is associated with a better 3-year DFS rate [9].

Among the semiautomated algorithms evaluated here, only a decrease in PV_{FLAB} was predictive of 4-year DFS, but not until the fourth treatment week. The combined results from the studies in this cohort indicate that, although a decrease in PV_{FLAB} is useful for prognostication purposes, early alterations in treatment might best be founded on SUV-based observations. However, SUV by itself does not convey information about the PV. The FLAB segmentation method can supply robust volumes before and during therapy. Defining the SUV_{peak} and the PV on FLT PET prior to treatment may help identify patients who will benefit from customized antiproliferative treatment, such as accelerated radiotherapy [5], or from the addition of antiproliferative systemic treatment to radiotherapy, e.g. EGFR inhibition [4]. In patients not eligible for these two treatment approaches, enhanced dose delivery to the FLAB-defined proliferative tumour subvolume during treatment, as this changes during therapy, may compensate for prolonged overall treatment times in conventional once-daily radiation schedules [3]. FLAB-based FLT PET segmentation may be a guidance tool with respect to target volume reduction in patients who show an excellent proliferative response during therapy and may receive over-treatment when the original target volume is adhered to. These assumptions need to be explored in further investigation to warrant eventual clinical application of PV_{FLAB} .

The present study indicated that FLAB is a promising candidate for radiation target volume adaptation based on sequential FLT PET scanning. However, there is a need for improved automatic initialization of the FLAB method before it can be applied with reduced user dependency in common clinical practice. Similarly, results obtained with the gradient-based method could be improved by adjusting preprocessing steps for FLT data and/or allow manual adjustment.

Conclusion

In this study, the FLAB algorithm performed best in terms of segmentation of the proliferative FLT PET tumour volume and correlation with patient outcome, when compared to the other semiautomatic segmentation methods RTL and W&C. With further operator-independent automation of the method, it may be a good candidate for FLT PET-guided individualized treatment choices, radiotherapy delivery strategies and response assessment.

Acknowledgments We thank Jorn van Dalen, Ph.D., Ton Feuth, Ph.D., Aswin Hoffmann, Ir, Karin Huisman, B.Sc, and Aniek Maessen, B.Sc, as well as the technologists of the Department of Nuclear Medicine and Radiotherapy for their contribution to this work.

Disclosures None.

References

- Hermans R, Feron M, Bellon E, Dupont P, Van den Bogaert W, Baert AL. Laryngeal tumor volume measurements determined with CT: a study on intra- and interobserver variability. *Int J Radiat Oncol Biol Phys.* 1998;40:553–7.
- Kim JJ, Tannock IF. Repopulation of cancer cells during therapy: an important cause of treatment failure. *Nat Rev Cancer.* 2005;5:16–25.
- Troost EG, Bussink J, Hoffmann AL, Boerman OC, Oyen WJ, Kaanders JH. 18F-FLT PET/CT for early response monitoring and dose escalation in oropharyngeal tumors. *J Nucl Med.* 2010;51:866–74.
- Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med.* 2006;354:567–78.
- Overgaard J, Hansen HS, Specht L, Overgaard M, Grau C, Andersen E, et al. Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHAN CA 6 and 7 randomised controlled trial. *Lancet.* 2003;362:933–40.
- Shields AF, Grierson JR, Dohmen BM, Machulla HJ, Stayanoff JC, Lawhorn-Crews JM, et al. Imaging proliferation in vivo with [¹⁸F]FLT and positron emission tomography. *Nat Med.* 1998;4:1334–6.
- van Waarde A, Cobben DC, Suurmeijer AJ, Maas B, Vaalburg W, de Vries EF, et al. Selectivity of 18F-FLT and 18F-FDG for differentiating tumor from inflammation in a rodent model. *J Nucl Med.* 2004;45:695–700.
- Menda Y, Boles Ponto LL, Dornfeld KJ, Tewson TJ, Watkins GL, Schultz MK, et al. Kinetic analysis of 3'-deoxy-3'-(18F)-fluorothymidine ((18F)-FLT) in head and neck cancer patients before and early after initiation of chemoradiation therapy. *J Nucl Med.* 2009;50:1028–35.
- Hoeben BA, Troost EG, Span PN, Herpen CM, Bussink J, Oyen W, et al. 18F-FLT PET during radiotherapy or chemoradiotherapy in head and neck squamous cell carcinoma is an early predictor of outcome. *J Nucl Med.* 2013;54:532–40.
- Everitt S, Hicks RJ, Ball D, Kron T, Schneider-Kolsky M, Walter T, et al. Imaging cellular proliferation during chemo-radiotherapy: a pilot study of serial 18F-FLT positron emission tomography/computed tomography imaging for non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys.* 2009;75:1098–104.
- de Langen AJ, Klabbbers B, Lubberink M, Boellaard R, Spreuwenberg MD, Slotman BJ, et al. Reproducibility of quantitative 18F-3'-deoxy-3'-fluorothymidine measurements using positron emission tomography. *Eur J Nucl Med Mol Imaging.* 2009;36:389–95.
- Boles Ponto LL, Menda Y, Dornfeld K, Tewson TJ, Watkins GL, Sunderland J, et al. Stability of 3'-deoxy-3'-[18F]fluorothymidine standardized uptake values in head and neck cancer over time. *Cancer Biother Radiopharm.* 2010;25:361–3. doi:10.1089/cbr.2009.0709.
- Daisne JF, Sibomana M, Bol A, Doumont T, Lonnew M, Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol.* 2003;69:247–50.
- 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.
- Geets X, Lee JA, Bol A, Lonnew M, Gregoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging.* 2007;34:1427–38.
- Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruyscher 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;77:301–8.
- Zaidi H, Abdoli M, Fuentes CL, El Naqa IM. Comparative methods for PET image segmentation in pharyngolaryngeal squamous cell carcinoma. *Eur J Nucl Med Mol Imaging.* 2012;39:881–91.
- Machulla HJ, Blocher A, Kuntzsch M, Piert M, Wei R, Grierson JR. Simplified labeling approach for synthesizing 3'-deoxy-3'-18F-fluorothymidine. *J Radioanal Nucl Chem.* 2000;243:843–6.
- 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.
- Janssens GO, Rademakers SE, Terhaard CH, Doornaert PA, Bijl HP, van den Ende P, et al. Accelerated radiotherapy with carbogen and nicotinamide for laryngeal cancer: results of a phase III randomized trial. *J Clin Oncol.* 2012;30:1777–83.
- Troost EG, Vogel WV, Merks MA, Slootweg PJ, Marres HA, Peeters WJ, et al. 18F-FLT PET does not discriminate between reactive and metastatic lymph nodes in primary head and neck cancer patients. *J Nucl Med.* 2007;48:726–35.
- Krak NC, Boellaard R, Hoekstra OS, Twisk JW, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging.* 2005;32:294–301.
- Vriens D, Visser EP, de Geus-Oei LF, Oyen WJ. Methodological considerations in quantification of oncological FDG PET studies. *Eur J Nucl Med Mol Imaging.* 2010;37:1408–25.
- Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, et al. 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.* 2012;82:1164–71.
- Wanet M, Lee JA, Weynand B, De Bast M, Poncelet A, Lacroix V, et al. 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.* 2011;98:117–25.
- Cheebsumon P, Boellaard R, de Ruyscher D, van Elmpst W, van Baardwijk A, Yaqub M, et al. Assessment of tumour size in PET/CT lung cancer studies: PET- and CT-based methods compared to pathology. *EJNMMI Res.* 2012;2:56.
- Liao S, Penney BC, Wroblewski K, Zhang H, Simon CA, Kampalath R, et al. Prognostic value of metabolic tumor burden on 18F-FDG PET in nonsurgical patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging.* 2012;39:27–38.
- Dibble EH, Alvarez AC, Truong MT, Mercier G, Cook EF, Subramaniam RM. 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J Nucl Med.* 2012;53:709–15.
- Hatt M, Cheze Le Rest C, Albarghach N, Pradier O, Visvikis D. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging.* 2011;38:663–72.
- Hatt M, Cheze-Le Rest C, Aboagye EO, Kenny LM, Rosso L, Turkheimer FE, et al. Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. *J Nucl Med.* 2010;51:1368–76.
- Heijmen L, de Geus-Oei LF, de Wilt JH, Visvikis D, Hatt M, Visser EP, et al. Reproducibility of functional volume and activity concentration in 18F-FDG PET/CT of liver metastases in colorectal cancer. *Eur J Nucl Med Mol Imaging.* 2012;39:1858–67.
- Hatt M, Groheux D, Martineau A, Espie M, Hindie E, Giacchetti S, et al. Comparison between 18F-FDG PET image-derived indices for early prediction of response to neoadjuvant chemotherapy in breast cancer. *J Nucl Med.* 2013;54:341–9.

33. Hatt M, van Stiphout R, le Pogam A, Lammering G, Visvikis D, Lambin P. Early prediction of pathological response in locally advanced rectal cancer based on sequential 18F-FDG PET. *Acta Oncol.* 2013;52:619–26.
34. Hatt M, Visvikis D, Albarghach NM, Tixier F, Pradier O, Cheze-le RC. Prognostic value of 18F-FDG PET image-based parameters in oesophageal cancer and impact of tumour delineation methodology. *Eur J Nucl Med Mol Imaging.* 2011;38:1191–202.
35. Troost EG, Bussink J, Slootweg PJ, Peeters WJ, Merks MA, van der Kogel AJ, et al. Histopathologic validation of 3'-deoxy-3'-18F-fluorothymidine PET in squamous cell carcinoma of the oral cavity. *J Nucl Med.* 2010;51:713–9.