Radiotherapy and Oncology 123 (2017) 63–70

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Target definition in head and neck

Modality-specific target definition for laryngeal and hypopharyngeal cancer on FDG-PET, CT and MRI

Hans Ligtenberg [a,*,1], Elise Anne Jager [a,1], Joana Caldas-Magalhaes [a], Tim Schakel [a], Frank A. Pameijer [b], Nicolien Kasperts [a], Stefan M. Willems [c], Chris H.J. Terhaard [a], Cornelis P.J. Raaijmakers [a,2], Marielle E.P. Philippens [a,2]

[a] Department of Radiotherapy; [b] Department of Radiology; and [c] Department of Pathology, University Medical Center Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 19 August 2016
Received in revised form 21 December 2016
Accepted 5 February 2017
Available online 1 March 2017

Keywords:
GTV
CTV
Pathology
Head and neck cancer
Imaging

ABSTRACT

Background and purpose: The goal of this study was to improve target definition by deriving modality-specific margins for clinical target volumes (CTV) for laryngeal and hypopharyngeal cancer on CT, MRI and 18-FDG-PET.

Material and methods: Twenty-five patients with T3/T4 laryngeal/hypopharyngeal cancer underwent CT, MRI and 18-FDG-PET scans before laryngectomy. HE-sections were obtained from the surgical specimen and tumor was delineated ($tumor_{HE}$). The GTVs on CT and MRI were delineated in consensus. PET-based GTVs were automatically segmented. The three-dimensionally reconstructed specimen was registered to the various images. Modality-specific CTV margins were derived and added to the GTVs to achieve adequate tumor coverage. The resulting CTVs were compared with each other, to $tumor_{HE}$, and to CTV_{CT10} constructed on CT with the clinical margin of 10 mm.

Results: CTV margins of 4.3 mm (CT), 6.1 mm (MRI) and 5.2 mm (PET) were needed to achieve adequate tumor coverage. The median volumes of the resulting modality-specific CTVs were 44 ml (CT), 48 ml (MRI) and 39 ml (PET), while the CTV_{10mm} was 80 ml.

Conclusion: For laryngohypopharyngeal tumors, 45–52% target volume reduction compared with CTV_{10mm} is achievable when modality-specific CTV margins are used. PET-based CTVs were significantly smaller compared to CT- and MRI-based CTVs.

In the past two decades, radiotherapy has experienced impressive technological innovations with increased precision in dose delivered to the tumor [1–3]. Accurate definition of the target volumes is essential to exploit these advances. In radiotherapy practice, disease extent is derived from imaging data, since pathological data are usually absent. Incorrect target definition can lead to unnecessary damage to healthy tissues or undertreatment near the tumor boundary, particularly for head-and-neck cancer, where the target is surrounded by critical anatomical structures [1,2]. Despite the significance, accurate and consistent definition of the target volume remains a major challenge in radiotherapy [3,4].

The gross tumor volume (GTV) is delineated on an image as a first estimate for the tumor. This volume is expanded with a margin to include microscopic tumor spread that is invisible on the images to create the clinical target volume (CTV). In the ICRU83 report, the CTV is defined as "a volume of tissue that contains a demonstrable GTV and/or subclinical malignant disease…" [3]. While the GTV represents the tumor visible in imaging, the CTV is more closely related to the histopathological tumor extent. No clear GTV definition exist for histo(pathology).

In spite of this, comparison of GTV delineation with (histo)-pathology is used to determine the accuracy of imaging modalities for target definition in literature [5]. To overcome the discrepancy between GTV in imaging and tumor volume in (histo)pathology, research groups made different choices to come to a surrogate GTV in macroscopic pathology [5] or microscopic histology [6] for head-and-neck tumors.

Another issue is that tumor visibility and, consequently, GTV delineation depends on the imaging modality and the data

* Corresponding author at: Department of Radiotherapy, Q00.3.11, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands.
E-mail address: h.ligtenberg@umcutrecht.nl (H. Ligtenberg).
[1] Joint first authors.
[2] Joint last authors.

http://dx.doi.org/10.1016/j.radonc.2017.02.005
0167-8140/

acquisition parameters. Further, interpretation of the images, differences in experience level and delineation guidelines result in inter-observer variation [7–9]. Assessment of the necessary CTV margin is hampered by these limiting factors.

In addition to variations between imaging modalities and between observers, microscopic extent can be highly variable depending on tumor site and type. As a consequence, "delineation of the CTV is currently based on clinical experience", as the ICRU 83 report states [3], which is reflected by a large variation in CTV margins between institutes [10,11] and between observers [7,10,12], ranging from 5 mm to 2 cm [6]. In our institute, in general a 10-mm CTV margin for head-and-neck cancer is added to the GTV determined on computed tomography (CT), as suggested by Caudell [11].

CT has traditionally been used for tumor contouring. However, with the development of magnetic resonance imaging (MRI) and positron emission tomography (PET), research for target definition has advanced accordingly to exploit the additional information available from these imaging modalities. MRI is known to have a range of excellent soft tissue contrasts, which could enhance tumor visibility [13]. Fluorodeoxyglucose (FDG)-PET provides information on the uptake of the radiopharmaceutical glucose analog, reflecting metabolic differences between tumor and surrounding tissues. The difference in metabolic properties can result in a high contrast image, although with low spatial resolution. These modality-specific qualities lead to modality-specific variations in GTV delineation [5,9,13], which might influence the CTV margin needed for adequate tumor coverage.

Here, we have compared the gross tumor volumes defined on MRI, CT and FDG-PET datasets with the corresponding tumor delineations on hematoxylin-and-eosin (HE) stained histopathology, which is considered to be the gold standard, in order to investigate the impact of different modalities on target definition of primary tumor. The main goal of this study was to derive modality-specific (MRI, CT and FDG-PET) clinical target definitions for laryngeal and hypopharyngeal tumors. The margins for these CTVs were optimized to ensure coverage of microscopic tumor extent defined by the histopathological "true tumor" volumes.

Materials and methods

Patient selection

Thirty-six patients, treated with a total laryngectomy (TLE) for primary T3/T4 laryngeal or hypopharyngeal cancer were included in this study according to the inclusion criteria as further described in this section. In total nine patients were excluded. The first six patients were used for optimization of the pathology-imaging registration procedure. Patient 12 was excluded because of a biopsy between preoperative imaging and surgery. The tumor of patient 21 was too large for our standard whole-mount procedure. The exclusion of patient 30 was due to a tumor which fragmented during surgery.

Resultantly, 27 patients (median age: 62 years, range: 49–79 years, two female and 25 male) with primary T3 ($N = 4$) or T4 ($N = 23$) laryngeal (supra-&glottic:2, supraglottic:7, transglottic:4, glottic:2) or hypopharyngeal (12) squamous cell carcinoma were included in the study. The patients underwent TLE as primary treatment between March 2009 and August 2014.

Furthermore, if the time interval between imaging with a particular modality and surgery was $\geq$20 days, images from that modality were excluded from further analysis. For two patients (11 and 32) all imaging was performed $\geq$20 days prior to surgery. Therefore, in total 25 patients were suitable for analysis with 25, 23 and 22 for CT, MRI and PET respectively.

Criteria of exclusion were contraindications for MRI or for CT contrast administration as defined in the protocols of the Radiology department, and insulin-dependent diabetes mellitus. Patients are numbered according to their study numbers. The study was approved by our ethics review board.

Preoperative image acquisition

Before surgery, all patients underwent CT, 1.5T MRI and 18-FDG-PET scans while immobilized in a head-and-shoulder radiotherapy mask with a head-and-neck support. The CT-scan was obtained after intravenous contrast agent administration. Two small flexible surface coils (Flex S coils, Philips Medical Systems) [14] were used for the MRI scan. Details on the imaging parameters are shown in Suppl. Tables 1–3.

Surgical specimen processing

The pathology procedure was described in detail previously [15]. Briefly, the fresh larynx specimen was fixated in 10%-formaldehyde directly after surgery. After fixation, the specimen was scanned on the CT for registration purposes, as described in the next section. Subsequently, the specimen was embedded in an agarose block and transversely sliced in approximately 3-mm-thick slices, which were then photographed and digitized. The thick slices were reconstructed in 3D into a digitized specimen using the contour of the agarose block to steer the registration as described by Caldas-Magalhaes et al. [15]. From each slice, a 4-μm section was obtained and stained with HE. Histopathological analysis was performed by a dedicated head-and-neck pathologist who delineated all tumor tissue on the HE-sections using a microscope, generating a three-dimensional structure referred to as tumor_{HE}. Subsequently, the HE-sections were digitized and registered to the 3D-reconstructed specimen. The structure tumor_{HE} was used as the gold standard to validate tumor definition from the various imaging modalities. The HE-sections were registered to the corresponding thick slice using cartilage landmarks to perform point-based rigid registration with scaling. Shrinkage of the HE-sections, which amounted to 12%, was taken into account. Shrinkage of the thick slices was limited and therefore not taken into account [15].

Image registration

Registration between the 3D-reconstructed histopathology block and the three imaging data sets was needed to compare the GTV delineations with the pathological tumor contour. The registration was performed in three main steps (Suppl. Fig. 1).

The first step was an automated rigid registration of the CT scans to the specimen using the CT of the specimen ($\text{specimen}_{\text{CT}}$) as an intermediate step [15]. Firstly, the contour of the pathology specimen and the contour of $\text{specimen}_{\text{CT}}$ where rigidly registered. Secondly, the $\text{specimen}_{\text{CT}}$ was registered to the CT of the patient. To guide this registration step a region of interest (ROI) was drawn around the thyroid and cricoid cartilage. The MRI and PET were rigidly registered to the CT based on mutual information. Additionally, the PET image was translated by overlaying the center of gravity of the PET-based GTV with the pathology based GTV, which corrects for larynx movement but might give a positive bias to the final results.

Secondly, the registration was optimized by manual rigid registration based on tumor and anatomical structures visible both on the specimen and on the MRI and/or CT images. The high soft-tissue contrast on MRI and the high contrast between cartilage and soft-tissue on CT was exploited here. The single

transformation derived from this registration step was applied to all imaging.

Lastly, due to deformations of the specimen further non-rigid transference of the tumor contour from pathology to the MRI was performed if necessary. The procedure was based on anatomical structures visible on both histopathology and MRI.

GTV delineation

Three dedicated head-and-neck specialists (two radiation oncologists and one radiologist) manually delineated the GTV in consensus on CT as well as on MRI datasets. An endoscopic report was available during the delineation sessions. For MRI, the GTV was delineated on the T_1-weighted gadolinium(Gd)-enhanced transverse MRI while applying dedicated delineation guidelines that have been previously validated [16]. All tumor delineation on CT and MRI was performed in Volumetool [17].

The GTV on PET was automatically segmented using a Gaussian mixture model [18], which was adapted for laryngeal and hypopharyngeal tumors [19]. The Gaussian mixture model determines a patient-specific threshold based on the signal intensity histogram of a volume of interest by fitting Gaussian distributions to the histogram using expectation maximization. These distributions, representing regions with different intensity levels in the image, are classified to tumor or background.

The sensitivity and positive predictive value (PPV) were derived for overlap analysis between the tumor_{HE} volume and the modality-specific GTV. The sensitivity is defined as the percentage coverage of the tumor_{HE} volume by the GTV, and the PPV is defined as the percentage of the GTV that covers the tumor_{HE} volume.

Modality-specific derivation of the CTV margin

From the GTV and tumor_{HE} contours, a common contour was derived for each patient per modality. The common contour is the delineation enclosing the overlapping volume of the GTV and tumor_{HE}. The three-dimensional distances from each point of the tumor_{HE} to the closest point of the common contour were determined (Fig. 1) for each patient per modality, yielding a distribution of distances per patient per modality. These measures are referred to as "tumor-intersection distances" and quantify the amount of tumor that is not included in the GTV.

We aimed to derive margins for each imaging modality to ensure coverage of at least 95% of the surface of tumor_{HE} that is not already included in the GTV identified on images for 95% of patients. Thus, the 95th percentile of the tumor-intersection distances was calculated (p95-distances) for each patient per modality. This 95th percentile cut-off point was chosen to limit the influence of residual registration errors and deformations.

The modality specific margin was then defined as the second highest p-95-distance for each modality to cover all tumor tissue in all but one patient, who is considered to be an outlier due to registration or deformation errors.

To accurately compare the various imaging modalities on a high resolution, all imaging was resampled to a resolution of $0.43 \times 0.43 \times 0.67$ mm.

CTV analysis

Each CTV was constructed by ally expanding the GTV using the derived modality-specific margin. Subsequently, the CTV for CT and MRI was corrected for anatomical boundaries such as air, pharyngeal constrictor muscles and vertebrae (Fig. 2, Fig. 3). As PET imaging does not give anatomical information, correction for air and vertebrae of PET CTVs was based on CT data.

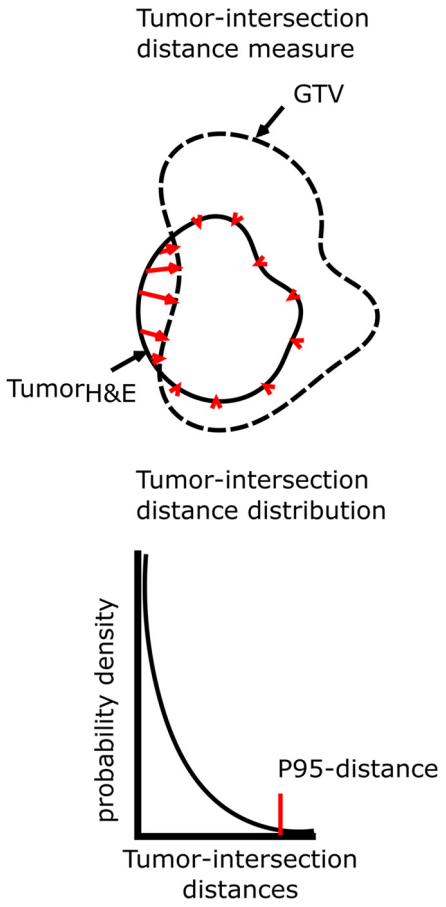

Fig. 1. P95-distance derivation scheme. For all patients, GTV and tumor_{HE} were geographically compared and the intersection of both contours was determined. The distance from the tumor_{HE} to the contour of the intersection of the tumor_{HE} with the GTV was measured (red arrows). For each patient, the 95th percentile of these distances is calculated (red bar in 'Tumor-intersection distance measure' figure). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These various CTVs were compared to the CTVs derived from our current clinical practice where the CT-based GTVs are expanded with a 10 mm margin (CTV_{CT10}).

Statistical analysis

Statistical analysis was performed in GraphPad Prism Version 6.07. The non-parametric Wilcoxon-signed-ranked test was used because there was no justification to assume a particular distribution. Differences were assumed to be significant for p-values < 0.05.

Results

Large variation in tumor_{HE} volume was observed (Table 1, Suppl. Table 5). Nearly all GTVs overestimated the tumor volume. However, the tumor_{HE} was not completely covered by the GTV contoured on each patient, which is reflected by a sensitivity less than one for all modalities and patients. The sensitivity was highest for CT compared to MRI and PET. However, the CT images also had the lowest positive predictive value (PPV) (Suppl. Table 4). Therefore, from these parameters, the optimal image modality for clinical practice could not be derived.

The median value of the p95-distances, quantifying the margin needed for adequate tumor coverage, was smaller for CT than for

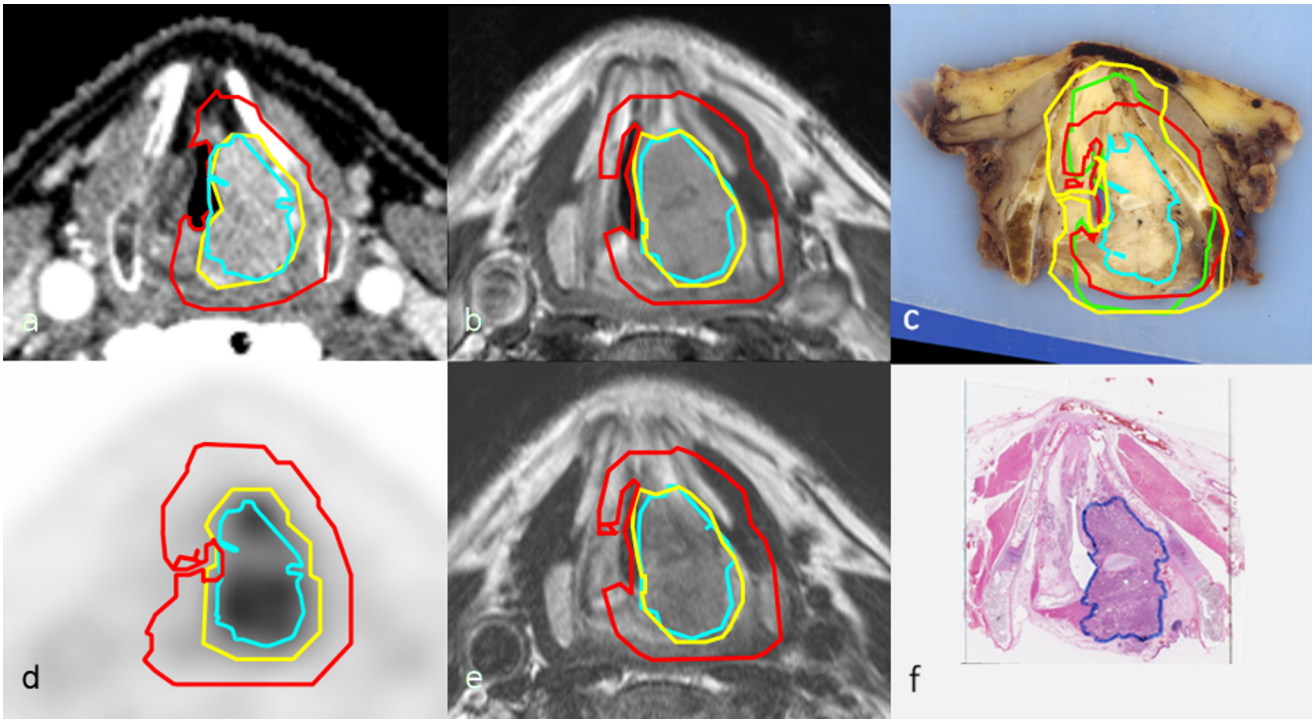

Fig. 2. Laryngeal tumor delineated on (a) contrast-enhanced CT, (b, e) MRI (T_1-weighted Gd-enhanced, T_2 weighted, respectively) and (d) FDG-PET with corresponding (c) macro- and (f) histo-pathology of patient 34 with a T4 stage glottic laryngeal carcinoma. For a, b, d and e the contours of $tumor_{HE}$ (blue), GTV (yellow) and CTV (red) are shown. For c, the contours of $tumor_{HE}$ (blue) and CTV_{CT} (green), CTV_{MRI} (red) and CTV_{PET} (yellow) are given. In f, the manually delineated pathology contour (dark blue) is depicted. The $tumor_{HE}$ contour shown in a–e is the non-rigidly transferred delineation of the pathology delineation shown in f. The difference of the pathology based tumor delineation between the imaging modalities is caused by slice thickness differences between MRI and CT/PET resulting in sampling differences.

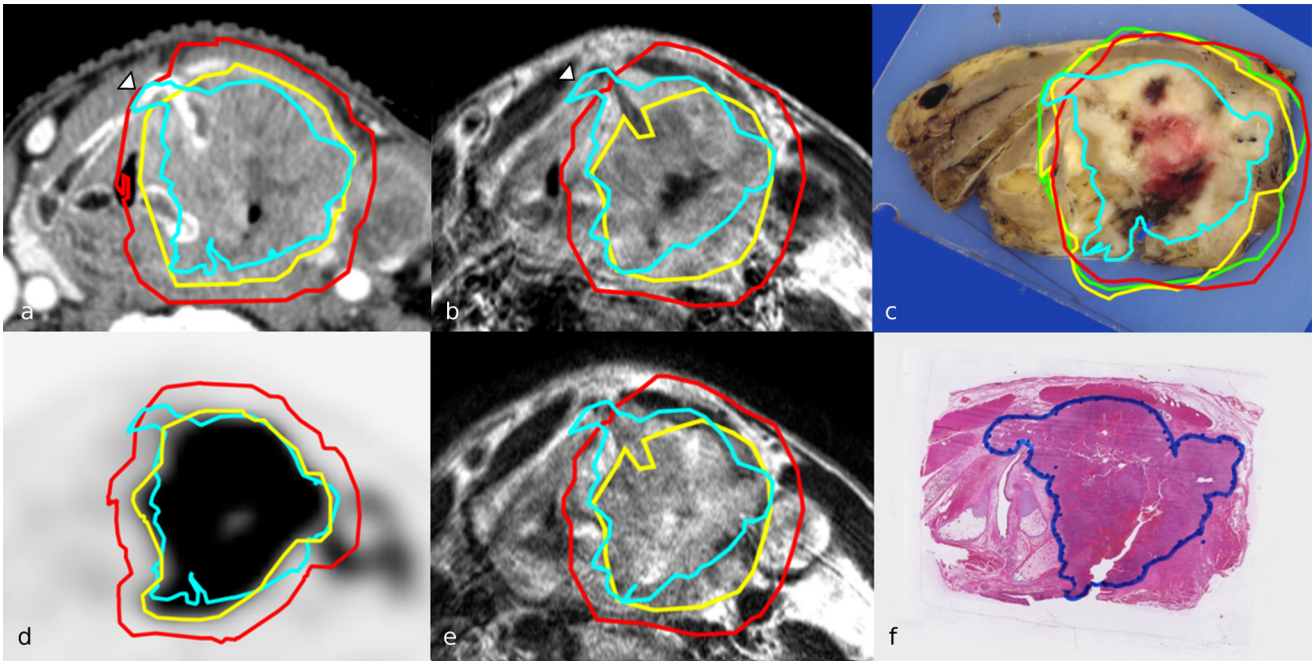

Fig. 3. Comparison of GTV delineation in MRI, CT and PET with tumor on histopathology illustrated by patient 24 with a T4-hypopharyngeal tumor with thyroid cartilage invasion: (a) contrast-enhanced CT, (b, e) MRI (T1w Gd-enhanced, T2w) and (e) FDG-PET with corresponding (c) macro- and (f) histo-pathology. For a, b, d and e the contours of $tumor_{HE}$ (blue), GTV (yellow) and CTV (red) are shown. For c, the contours of $tumor_{HE}$ (blue) and CTV_{CT} (green), CTV_{MRI} (red) and CTV_{PET} (yellow) are given. In f the manually delineated pathology contour (dark blue) is depicted. The $tumor_{HE}$ contour shown in a-e is the non-rigidly transferred delineation of the pathology delineation shown in f. (a) For CT, the tumor extension in the thyroid cartilage was missed by the GTV although just included in the CTV (indicated by arrow). (b) On MRI, a part of the thyroid cartilage invaded by tumor tissue and some other small parts of the tumor volume were not included in the GTV. The cartilage was not fully enclosed by the CTV margin (indicated by arrow point). (c) On PET, small parts of the tumor volume were missed in the GTV as well as a part of the thyroid cartilage that was invaded. The whole tumor was covered after adding a CTV margin to the GTV. The difference in depiction of the pathology based tumor delineation between the various imaging modalities is caused by differences in sampling due to differences in slice thickness between MRI and CT/PET. The analyses are performed on resampled images.

Table 1
Tumor$_{H\&E}$ volume, GTV, distances and CTV for all modalities.

Parameters	Modality	Median	Range	*p*-Values		
				Tumor$_{H\&E}$	MRI	PET
Tumor$_{H\&E}$ volume [ml]	Pathology	8.7	3.4–68.6	–	–	–
GTV [ml]	CT	17.5	5.9–88.7	**<0.0001**	**<0.0001**	**<0.0001**
	MRI	15.5	4.9–66.3	**<0.0001**	–	0.44
	PET	14.5	6.1–82.7	**<0.0001**	0.44	–
P95-distance [mm]	CT	1.6	0.4–5.6	–	**0.015**	0.12
	MRI	1.7	0.4–6.3	–	–	0.74
	PET	2.3	0.7–6.2	–	0.74	–
Maximum distance [mm]	CT	4.6	1.7–10.7	–	0.50	0.72
	MRI	4.3	1.9–12.3	–	–	0.48
	PET	5.3	2.0–12.1	–	0.48	–
CTV [ml]	CT	43.9 (55%)	18.9–152.0	–	0.052	**0.0022**
	MRI	47.7 (60%)	24.4–152.3	–	–	**0.0005**
	PET	38.7 (48%)	19.8–159.0	–	**0.0005**	–
	CT10 mm	79.9 (100%)	39.2–226.3	–	–	–
Tumor volume not included in CTV [ml]	CT	0.01	0.00–0.39	–	–	–
	MRI	0.01	0.00–0.34	–	–	–
	PET	0.03	0.00–0.53	–	–	–
	CT10 mm	>0.01	0.00–0.24	–	–	–

Tumor$_{H\&E}$ volume: histopathological tumor volume based on hematoxylin-eosin staining; p95-distance: 95th percentile of the distances measured form the tumor$_{HE}$ to the tumor$_{HE}$-GTV intersection.
Numbers in bold are significantly different (*p*-values < 0.05; Wilcoxon-signed-ranked test).

PET and MRI (Table 1). The modality-specific CTV margins are 4.3 mm (CT), 5.2 mm (PET) and 6.1 mm (MRI) (Fig. 4a).

The CTVs derived using modality-specific margins resulted in similar tumor coverage (Figs. 2 and 3, Suppl. Table 5). These CTVs were compared with the tumor$_{HE}$ volume and the CTV$_{CT10}$. Modality-specific CTVs were all smaller than the clinically used CTV$_{CT10}$ ($p < 0.0001$) with a reduction of the median volume of 45% (CT), 40% (MRI) and 52% (PET) (Table 1). PET-based CTVs were smaller than CT- and MRI-based CTVs. No significant differences were found between CT- and MRI-based CTVs no differences were

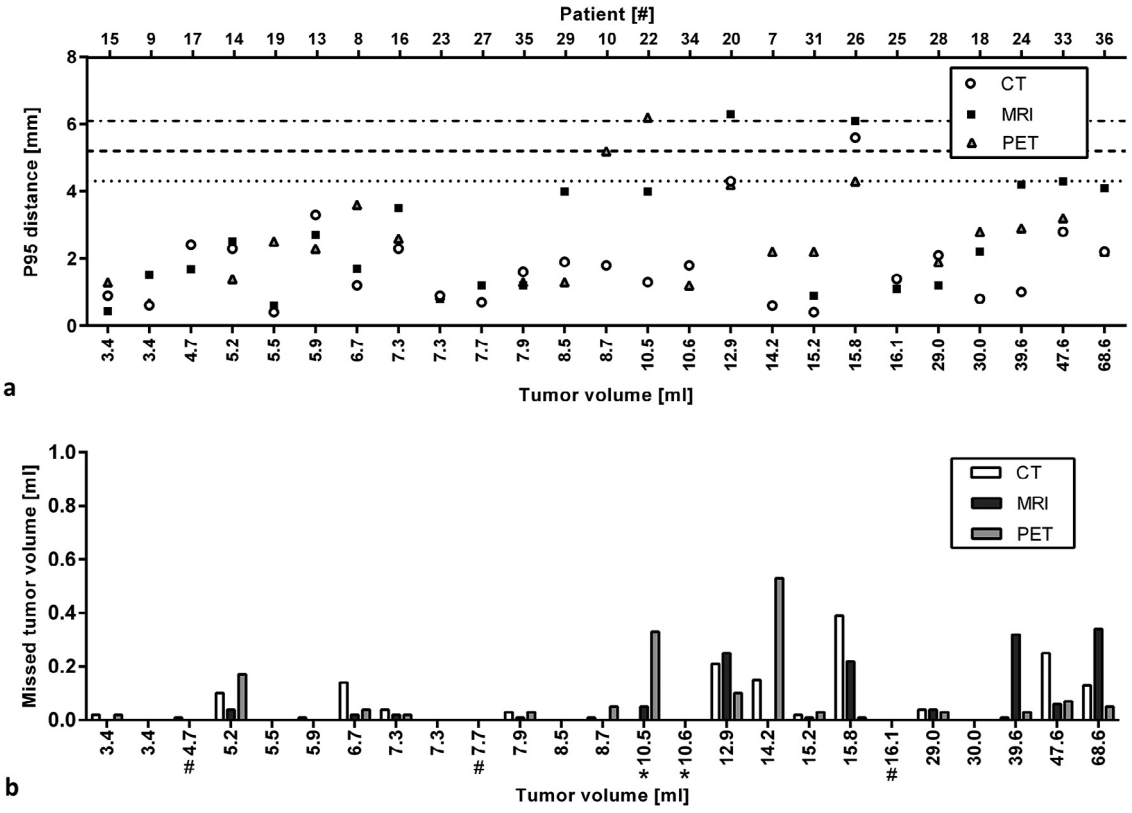

Fig. 4. Determination of the margin required to cover the tumor for the GTV delineation of the different imaging modalities and the resulting missed volume after application of this margin in all patients. (a) p95 distances for CT (circle), MRI (square) and PET (triangle) plotted against the tumor$_{HE}$ volume. Target margins are represented for CT (dotted line), MRI (dashed-dotted line) and PET (dashed line). (b) Missed tumor$_{HE}$ volume after adding the target margin ranked according to the tumor$_{HE}$ volume for CT (white), MRI (black) and PET (gray). # MRI excluded due to time interval ≥20 days. *PET excluded due to time interval ≥20 days.

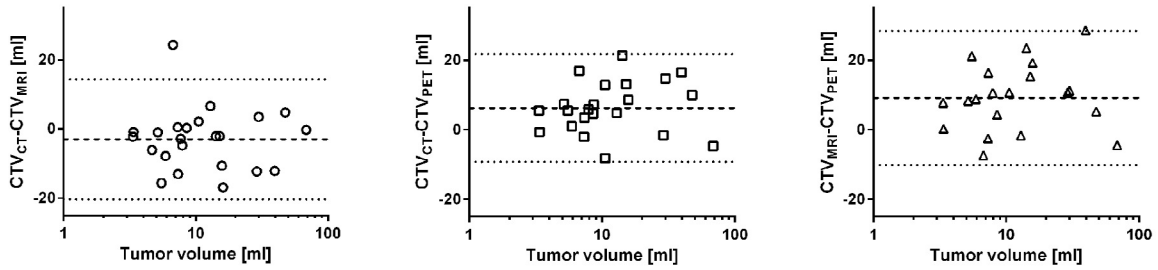

Fig. 5. Difference of target volumes between CT and MRI (a), CT and PET (b) and MRI and PET (c) versus tumor$_{HE}$ volume. Each point represents a patient. Mean difference between target volume and tumor$_{HE}$ volume (thick dashed line) and 95% confidence interval (thin dotted lines) are depicted.

found (Table 1). No correlation was found between the tumor$_{HE}$ volume and the volume differences of the CTVs for the different modalities (Fig. 5).

As a result of our CTV definition, tumor tissue was located outside the CTVs for several patients. Less tumor tissue was missed for CTVs of smaller tumors (<10 ml) than of larger tumors (Fig. 4b, Table 1). The volume located outside the CTV is a small fraction of the tumor volume with median values of 0.1% for all modalities and maximum fractions of 2.5% (CT), 1.9% (MRI) and 3.7% (PET). Cartilage invasion was not entirely covered in 3 (CT), 2 (MRI) and 4 (PET) patients (Fig. 3).

Discussion

In this study, clinical target definition for laryngeal and hypopharyngeal squamous cell carcinoma was improved by introducing the first evidence-based margins for CTV construction dedicated for each imaging modality. Expansion of the GTV by these margins resulted in a reduction of the average volume of the CTV by 45–52% compared with CTVs constructed using a margin of 10 mm on the GTV identified on the CT image, which is our current clinical practice. This reduction improves normal tissue complication probability (NTCP) [20,21]. The CTV margins were designed to achieve adequate tumor coverage while minimizing the inclusion of healthy tissue and amounted to 4.3 mm (CT), 5.2 mm (PET) and 6.1 mm (MRI). Since all modality-specific CTVs had equal tumor coverage and PET-based CTVs were smallest, i.e. the lowest overestimation, PET-based CTVs were considered to be most accurate. Nevertheless, the differences between the various imaging modalities were small.

Remarkably, PET-based CTVs were most accurate for target definition. This might be attributed to the high contrast between tumor and non-tumorous tissue and an accurate automatic segmentation using the Gaussian mixture model resulting in more adequate exclusion of non-tumorous tissue in GTV$_{PET}$ compared to CT and MRI. Consequently, the image analysis resulted in a small GTV and a moderate CTV margin. The low resolution of PET imaging appeared to be sufficient for target definition. PET-based target definition has been used for head-and-neck tumors, resulting in smaller target volumes compared with CT [5,15,22].

CT-based CTVs were larger than PET-based CTVs, although the CTV margin was smallest for CT. Low contrast between tumor and surrounding soft tissue on contrast-enhanced CT makes differentiation between tumorous and non-tumorous tissue difficult. This challenge may have persuaded the observers to enlarge the GTV in order to be reassured that all tumorous tissue was included [23].

Target definition on MRI was comparable with CT and was less accurate than PET, despite the advantage offered by superior soft tissue contrast and anatomical detail. The CTV margin for MRI was larger compared to the other imaging modalities. Apparently, the signal intensity of the tumor and the surrounding tissue is heterogeneous, and peritumoral inflammation may mimic neoplastic invasion [24], leading to difficulties in interpretation of the large variety of signal intensities [16]. New MRI contrasts, such as diffusion weighted imaging (DWI), may improve GTV delineation and target definition [25].

The tumor tissue missed outside the GTV was mostly due to partial volume effects. Moreover, in all imaging modalities, cartilage invasion was difficult to distinguish and therefore quite frequently missed in the GTV delineation. This is inconvenient as cartilage invasion is used for TNM-stage classification and influences the choice of treatment. The preferred treatment of patients with observed invasion through the outer cortex of the thyroid is surgery as radiation treatment leads to higher recurrence rates [26].

The main purpose of our study is the translation of these histological findings into a validated CTV concept, which has been shown to decrease target volumes and thus is expected to reduce complications to organs at risk. In literature, no consensus on how to determine a CTV margin exists, but some valuable remarks are made in ICRU report 83. It states that the CTV contains "subclinical malignant disease with a certain probability of occurrence considered relevant for therapy. There is no general consensus on what probability is considered relevant for therapy, but typically a probability of occult disease higher than from 5% to 10% is assumed to require treatment." In other words, the CTV is not considered to cover all tumor tissue for all patients. Here, we aimed to derive target margins that would ensure nearly full coverage of the histopathologically-identified tumor for at least 95% of the patients. These margins resulted in a maximum missed volume of 3.7% (0.5 ml) with a median of 0.1%. A margin that ensured 100% tumor coverage for all patients would overestimate the volume relevant for therapy for the majority of patients.

The general concept is that the CTV includes microscopic tumor islands, which are separated from the gross tumor bulk. In our study, however, isolated tumor was not observed on approximately 800 HE-sections. Future research with cytokeratin staining of histopathological sections, which has a high sensitivity for detection of squamous cell carcinoma, will be performed to confirm the absence of isolated tumor islands.

In this study, isotropic CTV margins were used and afterward corrected for clear anatomical boundaries such as air cavity, vertebrae and pharyngeal constrictor muscles. Further reduction of the CTV might have been possible when anisotropic margins were used. However, due to the limited number of patients involved in this study, derivation of guidelines for anisotropic margins was not possible.

A limitation of our study is that tumors eligible for radiotherapy are typically lower staged (T1b-T3) than tumors in the present study (T4-tumors). Higher stage tumors are often larger than lower stage tumors, which could have an impact on the CTV margin necessary to ensure tumor coverage. In our study, less tumor$_{HE}$ volume

was missed on the modality-specific CTVs for smaller tumors (<10 ml) compared to larger tumors [Fig. 4b]. This result indicates that smaller tumors have less microscopic tumor extent or that observers delineate smaller tumors more accurately. Based on this finding, we expect that the CTV margins derived in the present study are also sufficient for lower staged tumors.

Although image registration was optimized, registration mismatch could not fully be prevented. We manually improved the automatic rigid registration process which was described in a previous publication [15]. Despite this additional non-rigid registration, some deformations remained and resulted in a suboptimal registration between histology and the imaging modalities. To account for this, we accepted limited tumor volume outside the CTV, which was 2% at most.

It is important to note that for CT and MRI the CTV margin is based on manual GTV delineations, which also include observer delineations errors, i.e. visible gross tumor that is not delineated. In the determination of the CTV margin, both human errors and misses due to lack of visibility are jointly analyzed, although delineation errors are by definition not included in the CTV margin. To minimize human errors, we used consensus delineation by three observers. Therefore, we strongly advised to review the GTV delineation by a second observer and include imaging of various modalities to reduce delineation errors. We observed that in case of delineation errors, the delineations on one of the other did include the missed tumor part.

Another issue is that inter- and intra-observer variation can be large for manual GTV delineations on CT and MRI [9,23], whereas automatic GTV segmentation on PET does not suffer from these variations. This implies more standardization and easier validation than manual GTV delineations. Conversely, automatic segmentation can vary considerably with acquisition and reconstruction method. Using other PET segmentation methods, for example 50% SUVmax or manual delineation, might result in suboptimal GTV delineations, increasing the CTV margin needed. Furthermore, PET cannot be used as a stand-alone modality because of the lack of anatomical information which is needed to reduce the CTV by excluding anatomical structures like uninvolved muscle and cartilage. It should be noted that all observers involved in this study were experienced in delineating tumors in head-and-neck.

Generally, a fixed margin for CTV is used independent of imaging modality. In our study, this would have led to large differences in CTV following the differences in GTV. These differences would be partly compensated with the introduction of our new CTV approach using modality-specific margins which also substantially reduce the CTV. The final CT- and MRI-based CTVs were comparable, but were slightly larger than PET-based CTVs. Whether these differences are clinically significant should be analyzed in a broader perspective. For example, costs and availability of the various imaging modalities should be taken into account. MRI and CT scans are less expensive than PET scans, while MRI and PET are generally less easily accessible than CT. When PET is available, we suggest to use automatic segmentation for primary target definition to decrease inter- and intra-observer variability. However, CT or MRI based target definitions are good alternatives when proper, i.e. modality specific, margins are used. As MRI is currently not widely used in radiotherapy, also experience and closely following guidelines is important [16].

A combination of the various imaging modalities might further improve target definition, because the modalities provide complementary information. CT can provide anatomical information, MRI can add soft-tissue contrast and PET shows the metabolic activity. However, further research is needed to decide which combination is optimal for target definition.

In conclusion, although in general GTV overestimated the tumor volume in all modalities, CTV margins were needed to achieve adequate tumor coverage. In our study, modality-specific CTV margins of 4.3 mm (CT), 6.1 mm (MRI) and 5.2 mm (PET) were determined. This resulted in CTVs with 45–52% volume reduction compared with a clinical CTV based on GTV delineation on CT with a margin of 10 mm. PET-based CTVs were smallest, and CT and MRI-based CTVs were not significantly different from each other.

Conflict of interest statement

All authors declare having no conflict of interest related to the content of this manuscript.

Acknowledgments

This work was supported by two research grants from the Dutch Cancer Society (UU 2011-5152 and UU 2011-5216).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.radonc.2017.02.005.

References

[1] Rasch C, Steenbakkers R, van Herk M. Target definition in prostate, head, and neck. Semin Radiat Oncol 2005;15:136–45.
[2] Grégoire V, Coche E, Cosnard G, Hamoir M, Reychler H. Selection and delineation of lymph node target volumes in head and neck conformal radiotherapy. Proposal for standardizing terminology and procedure based on the surgical experience. Radiother Oncol 2000;56:135–50.
[3] Report 83: prescribing, recording, and reporting photon-beam Intensity-Modulated Radiation Therapy (IMRT). J ICRU 2010;10:41–53.
[4] Burnet NG, Thomas SJ, Burton KE, Jefferies SJ. Defining the tumour and target volumes for radiotherapy. Cancer Imaging 2004;4:153–61.
[5] Daisne J-F, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 2004;233:93–100.
[6] Campbell S, Poon I, Markel D, Vena D, Higgins K, Enepekides D, et al. Evaluation of microscopic disease in oral tongue cancer using whole-mount histopathologic techniques: implications for the management of head-and-neck cancers. Int J Radiat Oncol Biol Phys 2012;82:574–81.
[7] Purdy JA. Current ICRU definitions of volumes: limitations and future directions. Semin Radiat Oncol 2004;14:27–40.
[8] Daisne J-F, Sibomana M, Bol A, Cosnard G, Lonneux M, Grégoire V. Evaluation of a multimodality image (CT, MRI and PET) coregistration procedure on phantom and head and neck cancer patients: accuracy, reproducibility and consistency. Radiother Oncol 2003;69:237–45.
[9] Jager EA, Kasperts N, Caldas-Magalhaes J, Philippens MEP, Pameijer FA, Terhaard CHJ, et al. GTV delineation in supraglottic laryngeal carcinoma: interobserver agreement of CT versus CT-MR delineation. Radiat Oncol 2015;10:1–9.
[10] Moeckli R, Jeanneret Sozzi W, Mirimanoff R-O, Ozsahin M, Zouhair A, Valley JF, et al. Physical considerations on discrepancies in target volume delineation. Z Med Phys 2009;19:224–35.
[11] Caudell JJ, Meredith RF, Spencer SA, Keene KS, Dobelbower MC, Bonner JA. Margin on gross tumor volume and risk of local recurrence in head-and-neck cancer. Int J Radiat Oncol Biol Phys 2010;76:164–8.
[12] Hong TS, Tome WA, Harari PM. Heterogeneity in head and neck IMRT target design and clinical practice. Radiother Oncol 2012;103:92–8.
[13] Bird D, Scarsbrook AF, Sykes J, Ramasamy S, Subesinghe M, Carey B, et al. Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma. BMC Cancer 2015;15:844.
[14] Verduijn GM, Bartels LW, Raaijmakers CPJ, Terhaard CHJ, Pameijer FA, van den Berg CAT. Magnetic resonance imaging protocol optimization for delineation of gross tumor volume in hypopharyngeal and laryngeal tumors. Int J Radiat Oncol Biol Phys 2009;74:630–6.
[15] Caldas-Magalhaes J, Kasperts N, Kooij N, Van Den Berg CAT, Terhaard CHJ, Raaijmakers CPJ, et al. Validation of imaging with pathology in laryngeal cancer: accuracy of the registration methodology. Int J Radiat Oncol Biol Phys 2012;82:289–98.
[16] Jager EA, Ligtenberg H, Caldas-Magalhaes J, Schakel T, Philippens ME, Pameijer FA, et al. Validated guidelines for tumor delineation on magnetic resonance imaging for laryngeal and hypopharyngeal cancer. Acta Oncol 2016;55:1305–12.

[17] Bol GH, Kotte ANTJ, van der Heide UA, Lagendijk JJW. Simultaneous multi-modality ROI delineation in clinical practice. Comput Methods Programs Biomed 2009;96:133–40.
[18] Aristophanous M, Penney BC, Martel MK, Pelizzari CA. A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. Med Phys 2007;34:4223–35.
[19] Caldas Magalhaes J, Raaijmakers CP, Aristophanous M, Lee JA, Kasperts N, Jager EA, et al. FDG-PET semi automatic segmentation methods for GTV delineation in laryngeal and hypopharyngeal cancer. Int J Radiat Oncol 2014;90:S536.
[20] Vugts CAJM, Terhaard CHJ, Philippens MEP, Pameijer FA, Kasperts N, Raaijmakers CPJ. Consequences of tumor planning target volume reduction in treatment of T2–T4 laryngeal cancer. Radiat Oncol 2014;9:195.
[21] Chen AM, Yu Y, Daly ME, Farwell DG, Benedict SH, Purdy JA. Long-term experience with reduced planning target volume margins and intensity-modulated radiotherapy with daily image-guidance for head and neck cancer. Head Neck 2014;36:1766–72.
[22] Leclerc M, Lartigau E, Lacornerie T, Daisne J-F, Kramar A, Grégoire V. Primary tumor delineation based on 18FDG PET for locally advanced head and neck cancer treated by chemo-radiotherapy. Radiother Oncol 2015;116:87–93.
[23] Caldas-Magalhaes J, Kooij N, Ligtenberg H, Jager EA, Schakel T, Kasperts N, et al. The accuracy of target delineation in laryngeal and hypopharyngeal cancer. Acta Oncol (Madr) 2015;54:1181–7.
[24] Becker M, Zbaren P, Casselman JW, Kohler R, Dulguerov P, Becker CD. Neoplastic invasion of laryngeal cartilage: reassessment of criteria for diagnosis at MR imaging. Radiology 2008;249:551–9.
[25] Schakel T, Hoogduin JM, Terhaard CHJ, Philippens MEP. Diffusion weighted MRI in head-and-neck cancer: geometrical accuracy. Radiother Oncol 2013;109:394–7.
[26] Castelijns JA, Becker M, Hermans R. Impact of cartilage invasion on treatment and prognosis of laryngeal cancer. Eur Radiol 1996;6:156–69.