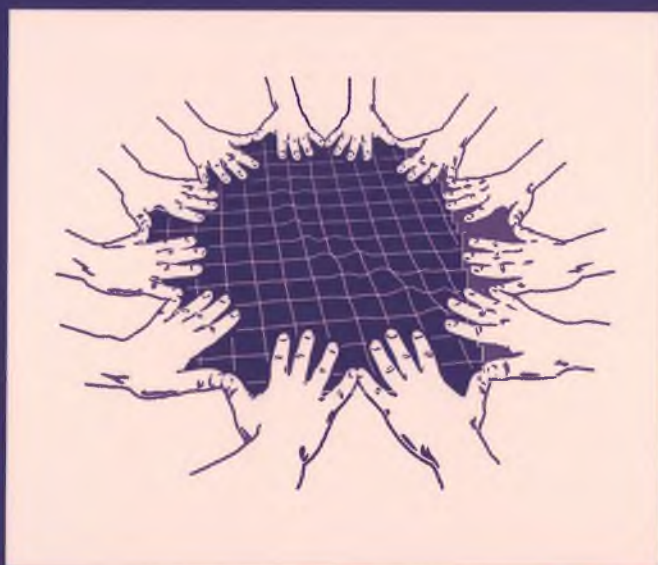


Radiotherapy for head and neck cancer

The role of FDG-PET



Dominic Schinagl

Radiotherapy for head and neck cancer

The role of FDG-PET

Dominic Schinagl

ISBN

978-90-9027849-0

Cover Artwork

© Moderat, reprinted after kind permission Monkeytown Records

Lay-out

Promotie In Zicht, Arnhem

Print

Ipskamp Drukkers, Enschede

Radiotherapy for head and neck cancer

The role of FDG-PET

Proefschrift

ter verkrijging van de graad van doctor aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 6 november 2013
om 12.30 uur precies

door

Dominic Ambrose Xavier Schinagl

geboren op 24 maart 1972

te Heerlen

Promotoren

Prof. dr. J.H.A.M. Kaanders

Prof. dr. W.J.G. Oyen

Manuscriptcommissie

Prof. dr. P. Wesseling

Prof. dr. B. van Ginneken

Prof. dr. D. de Ruyscher (*UZ Leuven, België*)

Contents

1	General Introduction and Outline	7
2	From anatomical to biological target volumes: the role of PET in radiation treatment planning. <i>Cancer Imaging 2006;6:S107-116</i>	15
3	Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. <i>Q J Nucl Med Mol Imaging 2008;52:74-83</i>	33
4	Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose positron emission tomography based target volume definition in head and neck cancer. <i>Int J Radiat Oncol Biol Phys 2007;69:1282-1289</i>	51
5	Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head and neck cancer? <i>Radiother Oncol 2009;91:95-100</i>	69
6	Can FDG-PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. <i>Eur J Nucl Med Mol Imaging 2011;38:1449-1458</i>	85
7	Pathology-based validation of FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer. <i>Eur J Nucl Med Mol Imaging: accepted for publication</i>	105
8	General discussion and future perspectives adapted from Innovations in radiotherapy planning of head and neck cancers: role of PET. <i>J Nucl Med 2010; 51: 66-76</i> Clinical evidence on PET-CT for radiation therapy planning in head and neck tumours. <i>Radiother Oncol 2010;96:328-334</i>	123
9	Summary	147
10	Summary in Dutch (Nederlandse Samenvatting)	155
	Dankwoord	161
	List of Publications	163
	Curriculum Vitae	167
A	Appendix: Color Figures	168

1



General Introduction and Outline

Squamous cell carcinomas of the head and neck have an annual incidence of 850000 patients worldwide and 2400 patients in the Netherlands [1]. It is the fifth most common malignancy in men, and the eighth in women. In western counties, patients typically have a history of alcohol and nicotine abuse, although there is increasing evidence concerning the role of the human papilloma virus in carcinogenesis in relatively young non-smoking oropharyngeal cancer patients [2-5]. In parts of Africa and in parts of Asia, infection with the Epstein-Barr virus and betel quid chewing play a significant role in the development of nasopharyngeal and oral cancer, respectively [4,6]. Patients can often present with an advanced stage of the disease: large primary tumors and frequently uni- or bilateral cervical lymph node metastases. If there are no signs of distant metastases (primarily lung- and/or bone metastases), intensive local or loco-regional treatment is recommended. Treatment can consist of surgical resection of the primary tumor and cervical lymph node levels (sometimes followed by postoperative (chemo)radiotherapy), definitive (chemo) radiotherapy of the primary tumor and (potentially) affected regional lymph nodes. Chemotherapy can be used as an induction treatment potentially leading to downstaging of the tumor, or it can be used to enhanced radiotherapy efficacy [7-13]. In most cases radiotherapy is the prime modality, as organ preservation strategies are increasingly preferred [12-14].

Radiotherapy has seen tremendous technological progress in recent years. Dose delivery with high geometrical precision is increasingly possible due to the introduction of stereotactic radiotherapy, radiosurgery, intensity modulated radiotherapy (IMRT) and three-dimensional planning of brachytherapy [15]. The current standard of target volume definition is based on information gathered by physical examination, and anatomical imaging such as computed tomography (CT) and magnetic resonance imaging (MRI). CT with intravenous iodine contrast agent is the imaging modality of choice for tumors of the larynx and hypopharynx. It enables localization and delineation of the primary tumor, (metastatic) lymph nodes, and surrounding tissues that are possibly invaded by the primary tumor, such as the thyroid cartilage. MRI with intravenous administration of gadolinium is the preferred modality for tumors of the oral cavity and oropharynx. MRI delivers supreme soft tissue contrast, whereby it is possible to, for instance, differentiate between malignant tumor infiltration and benign deep tongue musculature. In recent years new methods of tumor visualization have been introduced in oncology. Imaging techniques such as positron emission tomography (PET), single-photon emission computed tomography (SPECT), magnetic resonance spectroscopy (MRS) and diffusion weighted MRI (DW-MRI) are able to visualize biological characteristics of tumors, providing information on metabolism, physiology and molecular biology of tumor tissue [15]. PET is a highly sensitive imaging modality with a relatively good temporal, but limited spatial resolution. The administered radionuclide emits positrons that annihilate with electrons producing two photons simultaneously traveling in opposite directions. These photons

are subsequently detected on a ring of crystals arranged around the volume of interest. Coincidental events result in a “sinogram” that is finally reconstructed for image display. PET enables *in vivo* imaging of biologically active molecules, such as peptides, hormones, metabolites and pharmaceuticals. A variety of tumor characteristics, such as, glucose metabolism, hypoxia, proliferation and receptor expression can be non-invasively visualized with PET using specific radiopharmaceuticals. PET may support the localization and delineation of the primary tumor and possibly of the regional metastatic lymph nodes. The PET signal may also hold relevant information on overall tracer uptake within the lesion and on its intratumoral heterogeneity. ^{18}F -fluorodeoxyglucose (FDG) is the most widely used PET tracer in oncology, applied for tumor detection, staging, and treatment response monitoring [16]. FDG-PET can provide important complementary information for treatment planning in head and neck cancer, facilitating normal tissue sparing and dose escalation to resistant tumor subvolumes [17-19]. A well known limitation however, is the false positive PET signal caused by FDG accumulation by inflammatory cells; in head and neck cancer this is potentially frequent as the primary squamous cell carcinomas are often ulcerative lesions accompanied by reactive cervical lymph nodes.

Outline of the thesis

This thesis deals with the role of FDG-PET in radiation treatment planning of head and neck carcinomas. The current standard of target volume definition is based on information gathered by physical examination, and anatomical imaging modalities such as CT and MRI. Molecular imaging techniques for tumor visualization such as PET provide insight into tumor characteristics and can be complementary to the anatomical data.

In **chapter 2** an introductory overview regarding the role of PET in radiation treatment planning is given. The focus is on the potential impact of adding PET information to the current standard of target volume definition. The following three issues are discussed: First, can PET identify the primary tumor? Second, can biological tumor characteristics be visualized? Third, can intratumoral biological heterogeneity be identified?

The integration of PET information into CT-based radiation therapy planning requires a reliable image registration procedure. The aim in **chapter 3** is to select the optimal method for software fusion of dedicated PET and CT, and to validate the procedure for external beam radiation in the head and neck area.

The application of FDG-PET data for target volume delineation is not straightforward, as identification of tumor boundaries on PET suffers from a relative low spatial resolution and a “blurry” appearance of lesions. Furthermore, FDG-PET usually is interpreted qualitatively in diagnostic nuclear medicine, whereas in radiation oncology a more quantitative

approach is required for tumor contouring. In **chapter 4**, we compare different methods for tumor delineation with FDG-PET relative to CT-based delineation for radiation treatment planning in patients with head and neck cancer.

Radiation oncologists often encounter the dilemma of marginally enlarged lymph nodes when delineating the radiation target volume of a head and neck cancer patient. A number of CT criteria are used to assess the presence or absence of metastatic tumor within lymph nodes, these are often a compromise between sensitivity and specificity. FDG-PET has not provided a clear advantage in the nodal staging of head and neck cancer when compared to CT and MRI. However, this does not disqualify FDG-PET as a potential useful tool for delineation of the lymph node target volume in radiotherapy planning. Therefore, **chapter 5** describes the comparison of several FDG-PET segmentation methods to the current CT-based practice of lymph node assessment in order to evaluate the role of FDG-PET in radiotherapy target volume definition of the neck in head and neck cancer patients eligible for primary radiotherapy.

Various treatment strategies have been developed to improve outcome in head and neck cancer. However, it remains difficult to select patients for these intensified treatments. The FDG uptake of a head and neck squamous cell carcinoma may be an independent prognostic factor. In **chapter 6**, we assess the prognostic value of the determination of primary tumor volume from CT and FDG-PET scans, and various ways of quantifying FDG uptake in patients with head and neck cancer treated with (chemo)radiotherapy and provide an overview of the available literature.

The data on the accuracy of radiotherapy target volume delineation using FDG-PET are limited. In **chapter 7**, we correlate the volume of metastatic lymph nodes in head and neck cancer patients assessed by CT and various FDG-PET-based segmentation methods with the volume as determined by pathological examination.

Chapter 8 provides a general discussion including future perspectives. A summary of the thesis is given in **chapter 9**.

References

- [1] Brokstein B. *Head and Neck Cancer*. Northwestern University Medical school: Kluwer Academic Publishers;2003
- [2] Cognetti DM, Weber RS, Lai SY. Head and neck cancer: an evolving treatment paradigm. *Cancer* 2008;113:1911-1932.
- [3] Hashibe M, Brennan P, Benhamou S, *et al*. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck cancer Epidemiology Consortium. *J Natl Cancer Inst* 2007;99:777-789.
- [4] Pelucchi C, Gallus S, Garavello W, *et al*. Alcohol and tobacco use, and cancer risk for upper aerodigestive tract and liver. *Eur J Cancer Prev* 2008;17:340-344.
- [5] Psyrri A, DiMaio D. Human papillomavirus in cervical and head-and-neck cancer. *Nat Clin Pract Oncol* 2008;5:24-31.
- [6] Chou J, Lin YC, Kim J, *et al*. Nasopharyngeal carcinoma – review of the molecular mechanisms of tumorigenesis. *Head Neck* 2008;30:946-963.
- [7] Adelstein DJ, Leblanc M. Does induction chemotherapy have a role in the management of locoregionally advanced squamous cell head and neck cancer? *J Clin Oncol* 2006;24:2624-2628.
- [8] Bernier J, Dornge C, Ozsahin M, *et al*. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Eng J Med* 2004;350:1945-1952.
- [9] Bernier J, Vermorken JB, Koch WM. Adjuvant therapy in patients with resected poor-risk head and neck cancer. *J Clin Oncol* 2006;24:2629-2635.
- [10] Brizel DM, Esclamado R. Concurrent chemoradiotherapy for locally advanced, nonmetastatic, squamous carcinoma of the head and neck: consensus, controversy, and conundrum. *J Clin Oncol* 2006;24:2612-2617.
- [11] Cooper JS, Pajak TF, Forastiere AA, *et al*. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous cell carcinoma of the head and neck. *N Eng J Med* 2004;350:1937-1944.
- [12] Forastiere AA, Zhang Q, Weber RS, *et al*. Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. *J Clin Oncol* 2013;31:845-852.
- [13] Mendenhall WM, Arndt RJ, Palta JR. Intensity-modulated radiotherapy in the standard management of head and neck cancer: promises and pitfalls. *J Clin Oncol* 2006;24:2618-2623.
- [14] Parsons JT, Mendenhall WM, Stringer SP, *et al*. Squamous cell carcinoma of the oropharynx: surgery, radiation therapy, or both. *Cancer* 2002;94:2967-2980.
- [15] Schinagel DAX, Kaanders JH, Oyen WJ. From anatomical to biological target volumes: the role of PET in radiation treatment planning. *Cancer Imaging* 2006;6:S107-S116.
- [16] Juweid ME, Cheson BD. Positron-emission tomography and assessment of cancer therapy. *N Engl J Med* 2006;354:496-507.
- [17] Madani I, Duthoy W, Derie C, *et al*. Positron emission tomography guided, focal-dose escalation using intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;68:126-135.
- [18] Schwartz DL, Ford EC, Rajendran J, *et al*. FDG-PET/CT guided intensity modulated head and neck radiotherapy for head and neck cancer. *Head Neck* 2005;27:478-487.
- [19] Vanderstraeten B, Duthoy W, De GW, *et al*. [18F]fluoro-deoxy-glucose positron emission tomography ([18F] FDG-PET) voxel intensity-based intensity-modulated radiation therapy for head and neck cancer. *Radioth Oncol* 2006;79:249-258.

2 |

From anatomical to biological target volumes: the role of PET in radiation treatment planning

Dominic A.X. Schinagl
Johannes H.A.M. Kaanders
Wim J.G. Oyen

Abstract

Progress in radiation oncology requires a re-evaluation of the methods of target volume delineation beyond anatomical localization. New molecular imaging techniques for tumour visualisation such as positron emission tomography (PET) provide insight in tumour characteristics and can be complementary to the anatomical data of computed tomography or magnetic resonance imaging. In this review, three issues will be discussed: Firstly, can PET identify a tumor more accurately? Secondly, can biological tumour characteristics be visualised? Finally, can intratumoural heterogeneity of these characteristics be identified?

Introduction

In the past decade, there has been substantial technological progress in radiation oncology. Dose delivery with high geometric precision is possible due to the introduction of stereotactic radiotherapy, radiosurgery, intensity modulated radiotherapy (IMRT) and three-dimensional planning of brachytherapy. These developments require a re-evaluation of the standard methods for target volume delineation. The current standard of target volume definition is based on information gathered by physical examination, computed tomography (CT) and magnetic resonance imaging (MRI). In recent years new methods for tumor visualisation have been introduced in oncology. Imaging techniques such as positron emission tomography (PET), single-photon emission computed tomography (SPECT) and magnetic resonance spectroscopy (MRS) are able to visualise biological characteristics of tumours, providing information on metabolism, physiology and molecular biology of tumour tissue. These so called “functional” or “molecular” imaging modalities complement the anatomical data supplied by CT and MRI and include several potential advances.

First, the primary tumour can be identified more accurately. If carefully validated this could resize and reshape the gross tumour volume (GTV). This consequently could increase cure rates by reducing the chance of geographically missing part of the tumour during the treatment. When imaging modalities become more accurate, the inter- and intraobserver variation in tumour delineation decreases, which implies an enormous increase in the standard of care. More accurate tumour identification could also lead to an increased normal tissue sparing. Second, tumour characteristics relevant for radiation sensitivity can be visualized. Functional imaging may identify the degree of radiosensitivity of tumours, leading to an individualization of the radiation treatment. For example, the addition of a hypoxic cell sensitizer like nimorazole could be advantageous for radiation treatment when the tumour demonstrates a certain level of hypoxia [1]. Third, intratumoural biological heterogeneity can be identified. Ling *et al.* introduced the concept of “biological target volume” [2]. This biological target volume represents a subvolume of the tumour with specific characteristics on functional or molecular imaging techniques. Subvolumes that are relatively resistant to radiation receive an extra dose delivered with high precision on a small volume, which is called ‘dose painting’ or ‘dose sculpting’. This could increase cure rates without increasing the chances of late radiation-induced toxicity.

In this review we will focus on PET as this molecular imaging technique is becoming widely available at an astonishing rate. Significant clinical work has already been done to investigate its possible role in radiation treatment planning. The following three issues will be discussed:

- (1) Can PET identify the primary tumour more accurately?
- (2) Can biological tumour characteristics be visualised?
- (3) Can intratumoural biological heterogeneity be identified?

Can PET identify the primary tumour more accurately?

Coregistration

CT is the reference imaging modality for radiation treatment planning as it provides electron density information of the various tissues which is needed for the dose calculation algorithms. However, CT-images lack contrast between soft-tissue structures and tumour extension. E.g. the assessment of oral cavity and oropharyngeal tumours is severely hampered by scatter artefacts of dental fillings. Compared to CT, MRI has shown to be more accurate in evaluating soft-tissue and can be more sensitive for bone invasion of head and neck tumours, but it also has its limitations like geometric distortions at field of view edges, and artefacts at interfaces of bone and air. The observation that malignant tumour cells are characterized by increased glycolysis resulted in the development of whole body imaging using PET and the fluorine-18 labelled glucose analogue fluorodeoxyglucose (FDG). The high sensitivity of FDG-PET is related to the upregulation of glucose transporters on the cell membrane as well as increased hexokinase activity in a wide variety of malignancies [3]. Following phosphorylation by hexokinase, FDG is trapped in cells and leads to an uptake into tissue in proportion to the overall glucose metabolism. FDG uptake, however, is not cancer-specific, as increased glucose metabolism is also seen in inflammatory processes, muscle activity and brain activity. To incorporate PET data into CT-based tumour delineation there are three options of image fusion: visual fusion, where the physician compares two separate imaging modalities viewed next to each other, software fusion, where both modalities acquired on separate machines are overlaid in an integrated set of images (co-registration), and hardware fusion in which both data sets are acquired on one single machine, e.g. a hybrid PET/CT scanner. The co-registration of CT (and/or MRI) and PET images has successfully been demonstrated by many groups [4-6]. If performed properly, co-registration matches the performance of a hybrid PET/CT [6].

Validation

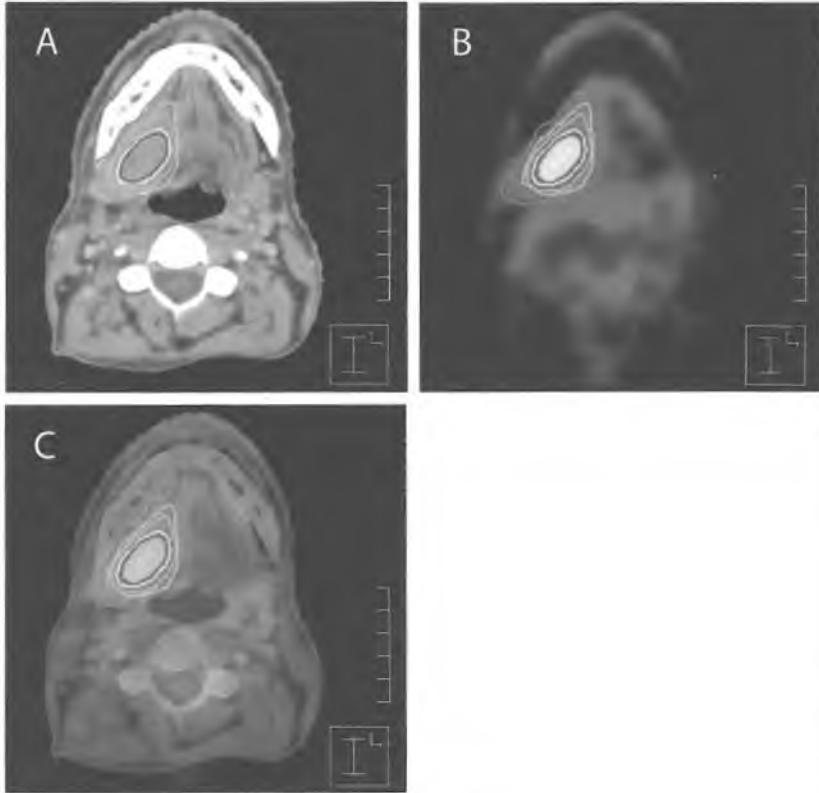
In 2005 Ng *et al.* reported on 124 patients with cancer of the oral cavity, all eligible for surgery [7]. Non co-registered FDG-PET and CT or MRI were obtained to determine their performance in detecting the primary tumour with histopathology as the gold standard. This largest study to date confirmed earlier data showing that 122 of 124 tumours were correctly detected by FDG-PET versus 108 on CT or MRI. FDG-PET missed a small superficial tumour and misinterpreted a floor of mouth tumour for a tongue tumour. However, no attempt was made to accurately delineate the tumour. The best evidence to date that FDG-PET can identify the primary tumour in head and neck cancer more accurately than conventional imaging is provided by Daisne *et al.* [8]. They compared the role of co-registered CT, MRI and FDG-PET in delineating the primary tumour in nine patients with laryngeal cancer who were scheduled for laryngectomy. Compared to the reference surgical specimen, all modalities overestimated the extension of the tumour. The average GTV on histological examination

was 12.6 cm^3 , whereas averages for the various imaging modalities were 16.3 cm^3 (PET), 20.8 cm^3 (CT) and 23.8 cm^3 (MRI). PET was closest to depict the true tumour volume, but all three imaging modalities (including PET) failed to identify a small fraction (approximately 10%) of the macroscopic tumour, mainly superficial mucosal extension. Application of radiopharmaceuticals other than FDG which depict different characteristics of tumour cells may also play a role in accurate tumour definition, e.g. the amino acid based radiopharmaceuticals O-2- ^{18}F fluoroethyl-L-tyrosine (FET) and methyl ^{11}C -methionine (MET). In brain tumour imaging FET and MET both have an advantage over FDG, as they do not accumulate in normal brain cells. FET-PET was superior in delineating human glioma's compared with MRI, which was demonstrated by the use of stereotactic biopsies in 28 glioma patients [9]. When comparing FET-PET to FDG-PET and CT in 18 patients with head and neck cancer, both FET-PET and FDG-PET were superior to CT in the detection of tumor [10]. FET-PET showed no uptake in physiologic tissue and inflammatory tissue, resulting in a higher specificity for tumour detection than FDG-PET. A drawback of FET-PET was a lower sensitivity, caused by a relatively low tumour uptake compared to FDG-PET. MET-PET images in patients with brain tumours were investigated by correlating MET uptake with histological examination of stereotactic biopsies [11]. Solid parts of brain tumours as well as brain tissue with infiltrating tumour cells were detected with high sensitivity (87%) and specificity (89%). Kracht *et al.* emphasized its value as a delineation tool for treatment planning in neurosurgery and radiotherapy [11]. Grosu *et al.* investigated the role of MET-PET in target volume definition of meningioma patients [12]. In skull-base tumours they found MET-PET to be superior in defining tumour infiltration in the surrounding structures compared to CT/MRI. Furthermore, they demonstrated that MET-PET significantly decreased the interobserver variability in tumour delineation as compared to CT/MRI-based delineation [13].

Reduction of interobserver variability

Interobserver variability of radiation target volume definition is a widely recognized problem. In laryngeal cancer for example, target definition using only CT leads to significant inter- and intraobserver variations in delineation of the GTV [14]. When imaging modalities become more accurate, the interobserver variability will decrease. A reduction in the interobserver variability is seen by incorporating FDG-PET data in the treatment planning [13,15-18]. Ciernik *et al.* reported a reduction in the mean volume difference of 26.6 cm^3 to 9.1 cm^3 when FDG-PET was incorporated into the CT-based GTV definition of 39 patients with various solid tumours [17]. This was associated with a reduction of the standard deviation from 38.6 cm^3 to 14.4 cm^3 . For 30 patients with NSCLC Caldwell *et al.* found a large variation in GTVs between the observers. The mean ratio of largest to smallest GTV was 2.31 when CT based and 1.56 when PET-CT based, indicating a clear improvement [16]. When analysing target definition of 22 patients with NSCLC by 11 observers, Steenbakkers *et al.* found that the amount of disagreement was reduced from 45% (CT based) to 18% (PET-CT based) [18].

Fig. 1 Planning CT scan (A), corresponding FDG-PET scan (B) and fusion image (C) show differences in target volume definition. Volume GTV_{CT} (red) = 47.5 cm^3 , GTV_{VIS} (green) = 43.8 cm^3 , $GTV_{40\%}$ (yellow) = 20.1 cm^3 , $GTV_{2.5}$ (orange) = 32.6 cm^3 , GTV_{SBR} (blue) = 15.7 cm^3 . Note that GTV_{SBR} is significantly smaller than GTV_{CT} and GTV_{VIS} .



Ashamalla *et al.* reported that for 19 patients with NSCLC the interobserver GTV variability decreased from a mean volume difference of 28.3 cm^3 (CT based) to 9.1 cm^3 (PET-CT based), with a respective decrease in standard deviation from 20.99 to 6.47 [15]. Co-registration of CT and FDG-PET reduces this variability even further, which was demonstrated in NSCLC by comparing co-registered to non-registered images [19].

Thresholding

Studies comparing GTV definition using FDG-PET (GTV_{PET}) and T1 weighted contrast enhanced MRI (GTV_{MRI}) in high-grade astrocytomas consistently showed that GTV_{PET} was smaller than GTV_{MRI} [20,21]. Similar studies comparing GTV_{PET} to delineation using CT or

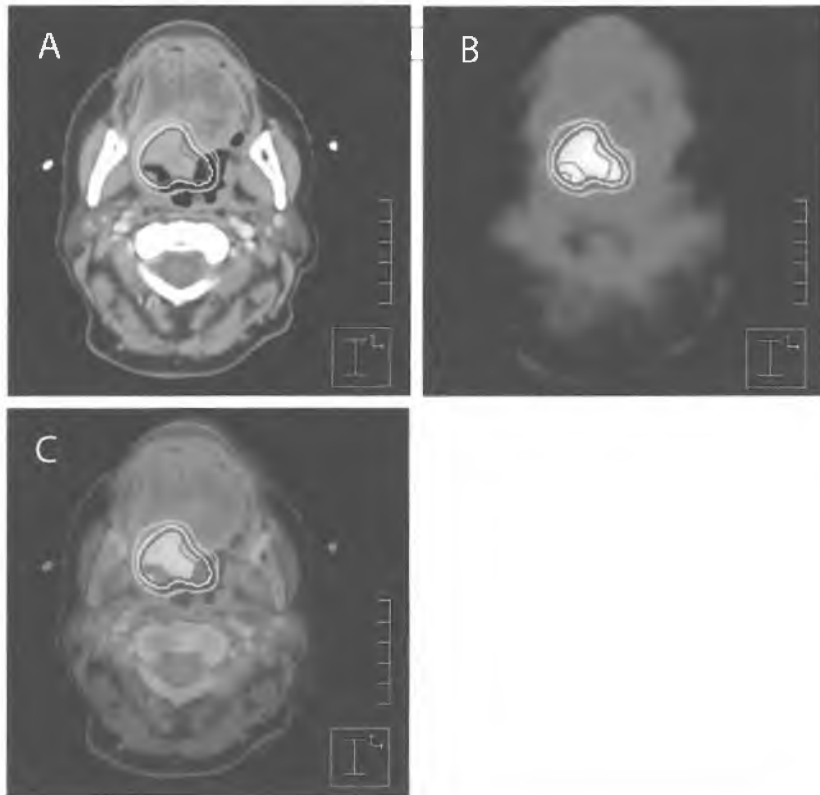
MRI ($GTV_{CT/MRI}$) have been performed in head and neck cancer [17,22-26]. Results show either no difference between GTVs or that GTV_{PET} was significantly smaller than $GTV_{CT/MRI}$. A reason for this might be that the optimal way to delineate a tumour on PET is not clearly defined, resulting in variations in the methods for defining GTV_{PET} . In clinical nuclear medicine, PET studies are usually interpreted qualitatively, whilst in radiation oncology a more quantitative approach is necessary as edge detection is required for tumour contouring [27].

PET images can be interpreted by visual assessment only (GTV_{VIS}), or by choosing thresholds, i.e. segmenting a lesion on the basis of a given level of radioactivity. This threshold could be any fixed cut-off of the standardized uptake value (SUV), and some investigators choose a cut-off of 2.5 ($GTV_{2.5}$) [28,29]. Thresholds are more commonly defined as a fixed percentage of the maximum tumour activity, for example 40% ($GTV_{40\%}$). Erdi *et al.* performed two-dimensional analyses of phantom experiments and stated that a threshold between 36% and 44% of the maximum activity would lead to an adequate segmentation [30]. Ciernik *et al.* recommended a fixed threshold of 50% of the maximum activity, also based on phantom data [17]. A more sophisticated method developed at the University St Luc in Brussels uses an adaptive threshold based on the signal-to-background ratio (SBR) (GTV_{SBR}) [31,32]. This method aims to incorporate specific PET imaging properties by deriving a mathematical function from phantom-measurements of objects of various sizes under various signal-to-background conditions. Nestle *et al.* recently compared these four PET-delineation methods in 25 NSCLC patients. They observed substantial differences in the resulting GTVs and demonstrated that the choice of a tool for target volume definition based on PET images is far from trivial [28]. Examples of two cases where different delineation tools were applied are given in figures 1 and 2.

Site specific issues

If a lung tumour causes atelectasis, it becomes extremely difficult to define the actual tumour on CT. In this situation most radiation oncologists would, when in doubt, prefer to define a larger target volume and possibly encompass part of the atelectasis rather than risk missing part of the tumour with a smaller target volume. FDG-PET may assist in more accurate definition of the border between tumour and atelectasis. This potential application of FDG-PET needs to be validated by resection specimen analysis. If this is proven to be reliable, radiation oncologists will confidently delineate a smaller target than they have done in the past. This can potentially result in a reduction of pulmonary toxicity by reducing the mean lung dose in patients currently eligible for high dose radiotherapy [33]. It also means that patients who previously were not eligible for this treatment because of pre-existent impaired pulmonary function, could then become eligible as result of the reduction in expected radiation damage to the lung.

Fig. 2 Planning CT (A), corresponding FDG-PET (B) and fusion image (C) show differences in target volume definition. It also illustrates PET activity in the air cavity



Several issues need to be addressed when using FDG-PET in radiation treatment planning of NSCLC. Although most primary NSCLC are visualized with FDG-PET, bronchoalveolar carcinoma shows limited or no increased FDG-uptake [34,35]. Furthermore, a post-obstruction pneumonia may also cause increased FDG accumulation, while malignant mediastinal involvement can be missed in those patients with relatively high FDG uptake in the heart. Issues associated with organ motion have already been addressed for CT scanning by gating image acquisition to the respiratory cycle. Similarly, it is possible to gate the linear accelerator. Gating the PET acquisition is also possible but experience is still relatively limited [36,37]. In head and neck cancer, FDG-PET may miss small lesions with insufficient tumour cells to delineate tumour, e.g. in superficial tumour infiltration of the mucosal lining [8] or micrometastatic disease in the regional lymph nodes. False positive findings can occur in inflammatory lymph nodes, tonsils, salivary glands and in areas of

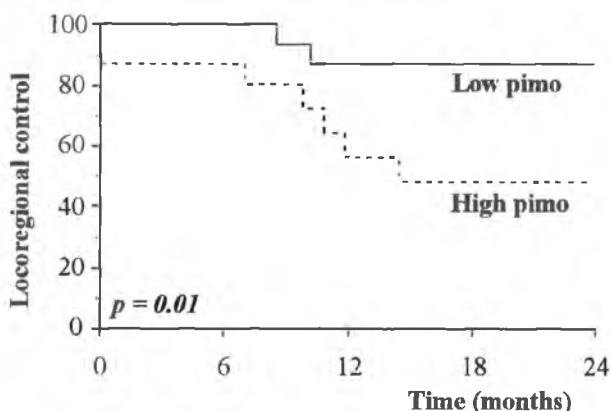
muscle activity (soft palate, base of tongue) [38]. When using segmented PET to delineate head and neck cancer, the GTV may become artificially enlarged by an inherent enclosure of part of an air cavity (fig 2). This needs to be corrected by using the CT as the anatomical template and subsequently 'trimming' part of the GTV_{PET} which is clearly in an air cavity. Thus, integration of FDG-PET into target volume definition is not trivial, but feasible. The GTV_{PET} generally is smaller than the GTV_{CT} or GTV_{MRI} for the primary tumour. However, accuracy for primary tumour delineation has only been fully validated for nine laryngeal cancers. Further studies need to elucidate the best methodology for FDG-PET based tumour delineation.

Can biological tumour characteristics be visualized?

Relevant factors for treatment outcome

Tumour hypoxia is a strong contributor to radiation resistance [39]. Kaanders *et al.* correlated the hypoxic fraction measured by staining tumor biopsies with pimonidazole, an exogenous hypoxia marker, with loco-regional tumour control after radiotherapy of advanced head and neck cancer [40]. Tumour control inversely correlated with the hypoxic fraction (fig3).

Fig. 3 Influence of tumour oxygenation on outcome after radiotherapy for head and neck cancer. Tumour hypoxia was measured by administering pimonidazole followed by immunohistochemical staining. High pimonidazole binding (low oxygen tension) was associated with significantly worse outcome. Reprinted with permission from Kaanders *et al.* "Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer" *Cancer Research* 2002;62:7066-7074.



Overgaard *et al.* demonstrated in a large randomised placebo controlled trial that administration of the hypoxic cell radiosensitizer nimorazole to radiotherapy could improve loco-regional control and disease free survival in patients with head and neck cancer [1]. Various treatment modifications are available to counteract hypoxia induced radioresistance: irradiating during hyperoxic gas breathing under normobaric or hyperbaric conditions, adding a hypoxic cell sensitizer (nimorazole) or a hypoxic cytotoxin (tirapazamine), or increasing the radiation dose. To various extent these modifications lead to increased toxicity. As not every patient benefits from these treatment intensifications, careful selection of patients is necessary, and PET could be of value as a predictive tool. Another tumour characteristic associated with radioresistance is tumour cell proliferation, especially in squamous cell carcinomas. This can be counteracted by shortening the overall treatment time, which has been shown to be effective in several randomised clinical trials [41,42]. As this also is a treatment intensification, PET could possibly help in selecting patients by identifying highly proliferating tumours, and hereby sparing the increased toxicity to these patients who are not likely to benefit from this approach [43]. Epidermal growth factor receptor (EGFR) inhibition is another strategy to counteract tumour cell proliferation. EGFR plays a key role in cellular proliferation of head and neck cancer, and Bentzen *et al.* recently discovered that the amount of EGFR expression in tumour biopsies could reliably be used to select the dose-fractionation scheme that had the greatest chance of benefiting the patient [44]. Bonner *et al.* reported that adding an EGFR-inhibitor (cetuximab) to the radiation treatment in a randomised clinical trial resulted in increased tumour control with only limited increase of toxicity [45].

Which radiopharmaceuticals are available to image these aspects?

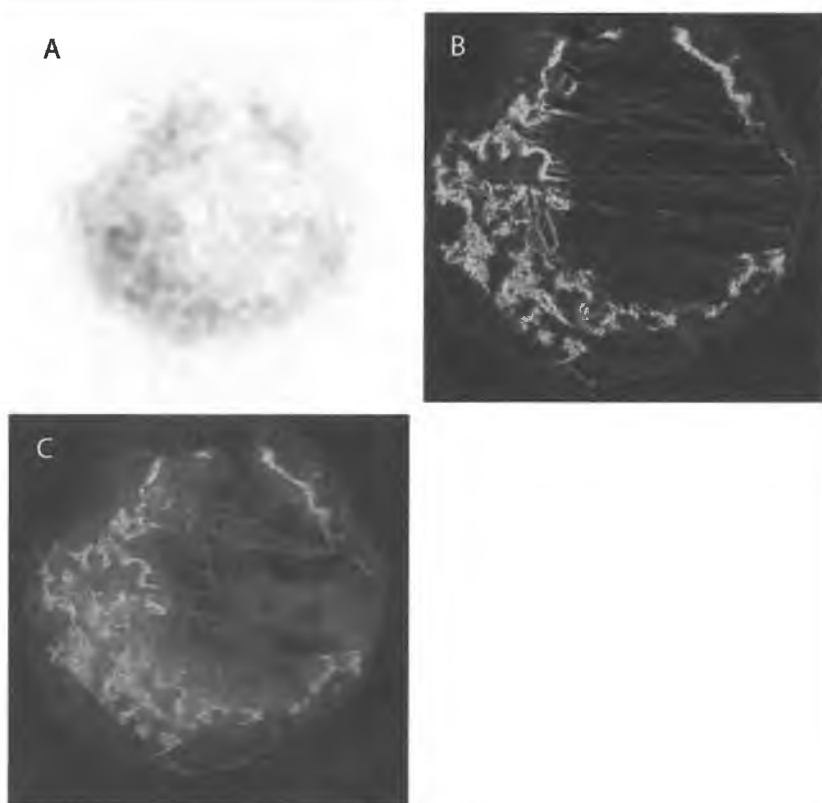
For imaging of tumour hypoxia both imidazole- and nonimidazole-containing agents have been developed. Imidazole containing radiopharmaceuticals are ^{18}F -fluoromisonidazole (FMISO) and ^{123}I -iodoazomycin arabinoside (IAZA). Non-imidazole tracers are $^{99\text{m}}\text{Tc}$ 4,9 diaza-3,3,10,10-tetramethyldodecan-2,11-dione-dioxime ($^{99\text{m}}\text{Tc}$ HL91) and ^{64}Cu -diacetyl-bis(N-4-methylthiosemicarbazone) (^{64}Cu -ATSM). Cell proliferation can be identified by labelling DNA precursors like thymidine or deoxyuridine, which are incorporated in DNA replication during the S phase of cycling cells. Clinical studies show a correlation between ^{18}F -3-deoxy-3-fluorothymidine (FLT) uptake and the Ki-67 labelling index. The latter is an accepted immunohistochemical marker to measure proliferation [46,47]. Further clinical validation of FLT against histopathological standards is in progress. At present, new radiopharmaceutical are developed in the preclinical phase to quantify tumour EGFR expression with PET, e.g. Ga-68-EGF and Zr-89-cetuximab [48,49].

Validation

FMISO was shown to bind selectively to hypoxic cells at radiobiologically relevant oxygen levels [50]. Piert *et al.* demonstrated in pigs that FMISO retention occurred at relevant pO_2

levels after restricting blood supply to part of the liver [51,52]. pO_2 levels were measured using invasive oxygen-sensitive Eppendorf histograph needle electrodes. In patients, hypoxia can also be detected by the exogenous marker pimonidazole. Bioreduction and irreversible binding of this marker occurs at pO_2 levels of below 10 mm Hg and can be visualised by immunohistochemistry in tumour sections [53]. To validate the usefulness of FMISO for detection of hypoxia in human tumours, Troost *et al.* compared FMISO autoradiography with pimonidazole immunohistochemistry in xenografted human squamous cell carcinomas and glioblastomas under various conditions of oxygen supply [54]. They found a significant correlation between the pimonidazole derived hypoxic fractions and the mean FMISO signal intensity, but also noticed that the correlation varied between the tumour lines stressing the need for validation in clinical studies with different tumour

Fig. 4 FMISO autoradiography (A), pimonidazole immunohistochemistry (B) and fusion image (C) of a xenografted human head and neck cancer. Courtesy of E.Troost, Dept. Radiation Oncology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.



entities (fig 4). Furthermore they observed that in the registration of changes of oxygenation status, FMISO was less accurate as the reference pimonidazole.

Temporal changes-treatment induced changes

Little is known about the temporal stability of hypoxia imaging, and there are at least three aspects to consider. Firstly, hypoxia can be transient due to structural and functional abnormalities of the tumour microvessels [55]. These abnormalities cause disturbances in the blood supply leading to temporal shutdown of vessels [55-57]. So, areas identified as 'normoxic' could be 'hypoxic' at a different time point. These changes occur at the microregional level and little is known about the sensitivity of PET-scanning for such changes. Secondly, the lifetime of chronically hypoxic cells is short, varying from a few hours to several days [58,59]. By the time radiation treatment commences, the hypoxic cells that were imaged will have already died and be replaced. Although the lifetime of individual cells is short, it is unlikely that the overall hypoxic pattern of a tumor will change significantly in a period of a few days. However, if intervals between PET-imaging and treatment exceed one or two weeks, this may become a relevant problem. Thirdly, irradiation itself can cause rapid changes in oxygenation and perfusion [60]. So even if the information obtained by functional imaging correlates with a relevant tumour characteristic, and even if that characteristic has an impact on clinical decision making on treatment selection, one has to be aware that the temporal stability of the imaged data may be limited. Currently, studies of temporal stability of hypoxia maps are in progress.

Can intratumoural biological heterogeneity be identified?

Concept of dose painting

The concept of dose painting was proposed by Ling *et al.* [2]. The idea is to visualize tumour subvolumes with a potential resistance to irradiation and to paint some additional dose onto that volume. Chao *et al.* applied this in a pilot study demonstrating an IMRT plan where a subvolume of the oropharyngeal tumour, identified by increased ^{64}Cu -ATSM retention, received an extra dose of 10 Gy [61]. Despite the dose escalation in this dose painting exercise the parotid glands could still adequately spared with the high precision IMRT technique.

Spatial resolution

At the level of tumour microenvironment it is conceivable that new cells formed in the proliferating cell compartment push older cells away from the blood vessels resulting in a gradual depletion of oxygen and nutrients. Tumour cells in the hypoxic compartment would be pushed further down the oxygen gradient and eventually die of oxygen deficiency and starvation [62]. The dynamics of hypoxic tumour cells in xenografted

human head and neck cancer was analysed by consecutive injection of two different hypoxia markers [59]. Over time pimonidazole positive cells were being pushed away from the vasculature and cell debris with pimonidazole adducts appeared in the necrotic regions. Meanwhile, new hypoxic cells appeared at the 'hypoxic front' identified by a second marker, CCI-103F, administered at a later point in time. This 'pattern of hypoxia' is measured in micrometres and therefore can not be detected by *invivo* imaging techniques such as PET due to limited spatial resolution. So PET images hypoxia at a more global level. Furthermore the radiation dose delivery also has a certain resolution. Given these limitations, dose painting will most likely only be feasible with subvolumes greater than 0.5 cm^3 . PET imaging would then have to identify subvolumes within the tumour with a larger than average content of hypoxic cells.

Conclusion

Integrating biological or molecular information of tumours into radiation oncology might help in deciding not only *where* but also *how* radiation therapy should be delivered [63]. The addition of functional imaging to the standard anatomically based target volume definition has already shown significant advantages, especially when images are co-registered. The reduction of interobserver variability is obvious and can increase the standard of care for all patients, not only by reducing geographically missing parts of the cancer, but also by reducing the irradiated volume of normal tissues and organs at risk. As more studies become available validating the various functional imaging properties, more treatment related decisions (dose painting, shortening overall treatment time, adding a sensitizer) will have to be tested in clinical trials. The goals are challenging and clear: Firstly, to enlarge the therapeutic window by increasing the tumour control probability and decreasing the normal tissue complication probability. Secondly, to develop predictive assays that can serve as selection tools for patients that are likely to profit from intensified treatments. We believe that molecular imaging can play an important role in achieving these goals.

References

- [1] Overgaard J, Eriksen JG, Nordsmark M, *et al.* Palsma osteopontin, hypoxia, and response to the hypoxia sensitizer nimorazole in radiotherapy of head and neck cancer; results from the DAHANCA 5 randomised double-blind placebo-controlled trial. *Lancet Oncol* 2005;6:757-764.
- [2] Ling CC, Humm J, Larson S, *et al.* Towards multidimensional Radiotherapy (MD-CRT): biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys* 2000;47:551-560.
- [3] Gambhir SS, Czernin J, Schwimmer J, *et al.* A tabulated summary of the FDG-PET literature. *J Nucl Med* 2001;42(Suppl 5):15-93S.
- [4] Vogel WV, Oyen WJ, Barentsz JO, *et al.* PET/CT: panacea, redundancy, or something in between? *J Nucl Med* 2004;45(Suppl 1):15S-24S.
- [5] Vogel WV, van Dalen JA, Schinagl DA, *et al.* Correction of an image size difference between positron emission tomography (PET) and computed tomography (CT) improves image fusion of dedicated PET and CT. *Nucl Med Commun* 2006;27:515-519.
- [6] Vogel WV, Schinagl DA, van Dalen JA, *et al.* Validated image fusion of dedicated PET and CT for external beam Radiation therapy in the head and neck area. *Q J Nucl Med Mol Imaging* 2008;52:74-83.
- [7] Ng SH, Yen TC, Liao CT, *et al.* 18F-FDG-PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of 124 patients with histologic correlation. *J Nucl Med* 2005;46:1136-1143.
- [8] Daisne JF, Duprez T, Weynd B, *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.
- [9] Pauleit D, Floeth F, Hamcher K *et al.* O-2-18F-fluoroethyl-L-tyrosine PET combined with MRI improves the diagnostic assessment of cerebral gliomas. *Brain* 2005;128(Pt 3):678-687.
- [10] Pauleit D, Zimmermann A, Stoffels G, *et al.* 18F-FET-PET compared with 18F-FDG-PET and CT in patients with head and neck cancer. *J Nucl Med* 2006;47:256-261.
- [11] Kracht LW, Miletic H, Busch S, *et al.* Delineation of brain tumor extent with 11C-L-methionine positron emission tomography: local comparison with stereotactic histopathology. *Clin Cancer Res* 2004;10:7163-7170.
- [12] Grosu AL, Lachner R, Wiedenmann N, *et al.* Validation of a method for automatic image fusion (BrainLAB System) of CT data and 11C-methionine-PET data for stereotactic radiotherapy using a LINAC: First clinical experience. *Int J Radiat Oncol Biol Phys* 2003;56:1450-1463.
- [13] Grosu AL, Weber WA, Astner ST, *et al.* 11C-methionine PET improves the target volume delineation of meningiomas treated with stereotactic fractionated radiotherapy. *Int J Radiat Oncol Biol Phys* 2006;66: 339-344.
- [14] Hermans R, Feron M, Bellon E, *et al.* Laryngeal tumor volume measurements determined with CT: a study on intra- and interobserver variability. *Int J Radiat Oncol Biol Phys* 1998;40:553-557.
- [15] Ashamalla H, Rafia S, Parikh K, *et al.* The contribution of integrated PET/CT to the evolving definition of treatment volumes in radiation treatment planning in lung cancer. *Int J Radiat Oncol Biol Phys* 2005;63:1016-1023.
- [16] Caldwell CB, Mah K, Ung YC, *et al.* Observer variation in contouring Gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of 18FDG-hybrid PET fusion. *Int J Radiat Oncol Biol Phys* 2001;51:923-931.
- [17] Ciernik IF, Dizendorf E, Baumert BG, *et al.* Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
- [18] Steenbakkers RJ, Duppen JC, Fitton I, *et al.* Reduction of observer variation using matched CT-PET for lung cancer delineation: a three dimensional analysis. *Int J Radiat Oncol Biol Phys* 2006;64:435-448.
- [19] Fox JL, Rengan R, O'Meara W, *et al.* Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non-small-cell lung cancer? *Int J Radiat Oncol Biol Phys* 2005;62:70-75.
- [20] Gross MW, Weber WA, Feldmann HJ, *et al.* The value of 18F-fluorodeoxyglucose PET for the 3D radiation treatment planning of malignant gliomas. *Int J Radiat Oncol Biol Phys* 1998;41:989-995.
- [21] Tralins KS, Douglas JG, Stelzer KJ, *et al.* Volumetric analysis of 18F-FDG-PET in glioblastoma multiforme: prognostic information and possible role in definition of target volumes in radiation dose escalation. *J Nucl Med* 2002;43:1667-1673.
- [22] Heron DE, Andrade RS, Flickinger J, *et al.* Hybrid PET-CT simulation for Radiation treatment planning in head and neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.

- [23] Nishioka T, Shiga T, Shirato, *et al.* Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057.
- [24] Paulino AC, Koshy M, Howell R, *et al.* Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
- [25] Scarfone C, Lavelly WC, Cmelak AJ, *et al.* Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. *J Nucl Med* 2004;45:543-552.
- [26] Schwartz DL, Ford E, Rajendran J, *et al.* FDG-PET/CT imaging for preradiotherapy staging of head and neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2005;61:129-136.
- [27] Bradley JD, Perez CA, Dehdashti F, *et al.* Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. *J Nucl Med* 2004;45(Suppl 1): 96S-101S.
- [28] Nestle U, Kremp S, Schaefer-Schuler A, *et al.* Comparison of different methods for delineation of 18F-FDG-PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1348.
- [29] Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys* 2004;59:4-5.
- [30] Erdi YE, Mawlawi O, Larson SM, *et al.* Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. *Cancer* 1997;80(Suppl 12):250S-2509.
- [31] Black QC, Grills IS, Kestin LL, *et al.* Defining a Radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys* 2004;60:1271-1282.
- [32] Daisne JF, Sibomana M, Bol A *et al.* Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiation Oncol* 2003;69:247-250.
- [33] van der WA, Nijsten S, Hochstetler M, *et al.* Increased therapeutic ratio by 18FDG-PET CT planning in patients with clinical CT stage N2-N3M0 non-small-cell lung cancer: a modelling study. *Int J Radiat Oncol Biol Phys* 2005;61:649-655.
- [34] Higashi K, Ueda Y, Seki H, *et al.* Fluorine-18-FDG-PET imaging is negative in bronchioalveolar lung carcinoma. *J Nucl Med* 1998;39:1016-1020.
- [35] Pieterman RM, van Putten JW, Meuzelaar JJ *et al.* Preoperative staging of non-small-cell lung cancer with positron-emission tomography. *N Engl J Med* 2000;343:254-261.
- [36] Nehmeh SA, Erdi YE, Ling CC, *et al.* Effect of respiratory gating on quantifying PET images of lung cancer. *J Nucl Med* 2002;43:876-881.
- [37] Nehmeh SA, Erdi YE, Pan T, *et al.* Four-dimensional (4D) PET/CT imaging of the thorax. *Med Phys* 2004;31:3179-3186.
- [38] Grosu AL, Pierl M, Weber WA, *et al.* Positron emission tomography for radiation treatment planning. *Strahlenther Onkol* 2005;181:483-499.
- [39] Gray LH, Conger AD, Ebert M, *et al.* The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br J Radiol* 1953;26:638-648.
- [40] Kaanders JH, Wijffels KI, Marres HA, *et al.* Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer. *Cancer Res* 2002;62:7066-7074.
- [41] Fu KK, Pajak TF, Trotti A, *et al.* Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: First report of RTOG 9003. *Int J Radiat Oncol Biol Phys* 2000;48:7-16.
- [42] Overgaard J, Hansen HS, Specht L, *et al.* Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHANCA 6 and 7 randomised controlled trial. *Lancet* 2003; 362:933-940.
- [43] Buck AK, Halter G, Schirrmeister H, *et al.* Imaging proliferation in lung tumors with PET: 18F-FLT versus 18F-FDG. *J Nucl Med* 2003;44:1426-1431.
- [44] Bentzen SM, Atasoy BM, Daley FM, *et al.* Epidermal growth factor receptor expression in pretreatment biopsies from head and neck squamous cell carcinoma as a predictive factor for a benefit from accelerated radiation therapy in a randomised controlled trial. *J Clin Oncol* 2005;23:5560-5567.
- [45] Bonner JA, Harari PM, Giralt J, *et al.* Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2006;354:567-578.

- [46] Buck AK, Schirrmester H, Hetzel M, *et al.* 3-deoxy-3-18F-fluorothymidine-positron emission tomography for non-invasive assessment of proliferation in pulmonary nodules. *Cancer Res* 2002;62:3331-3334.
- [47] Vesselle H, Grierson J, Muzi M, *et al.* In vivo validation of 3'-deoxy-3'-18F-fluorothymidine (FLT) as a proliferation imaging tracer in humans: correlation of FLT uptake by positron emission tomography with Ki-67 immunohistochemistry and flow cytometry in human lung tumors. *Clin Cancer Res* 2002;8:3315-3323.
- [48] Perk LR, Visser GW, Vosjan MJ, *et al.* (89)Zr as a PET surrogate radioisotope for scouting biodistribution of the therapeutic radiometals (90)Y and (177)Lu in tumorbearing nude mice after coupling to the internalizing antibody cetuximab. *J Nucl Med* 2005;46:1898-1906.
- [49] Velikyan I, Sundberg AL, Lindhe O, *et al.* Preparation and evaluation of (68)Ga-DOTA-hEGF for visualization of EGFR expression in malignant tumors. *J Nucl Med* 2005;46:1881-1888.
- [50] Rasey JS, Koh WJ, Grierson JR, *et al.* Radiolabelled fluoromisonidazole as an imaging agent for tumor hypoxia. *Int J Radiat Oncol Biol Phys* 1989;17:985-991.
- [51] Piert M, Machulla H, Becker G, *et al.* Introducing 18F-fluoromisonidazole positron emission tomography for the localization and quantification of pig liver hypoxia. *Eur J Nucl Med* 1999;26:95-109.
- [52] Piert M, Machulla HJ, Becker G, *et al.* Dependency of the 18F-fluoromisonidazole uptake on oxygen delivery and tissue oxygenation in the porcine liver. *Nucl Med Biol* 2000;27:693-700.
- [53] Raleigh JA, Calkins-Adams DP, Rinker LH, *et al.* Hypoxia and vascular endothelial growth factor expression in human squamous cell carcinomas using pimonidazole as a hypoxia marker. *Cancer Res* 1998;58:3765-3768.
- [54] Troost EG, Laverman P, Kaanders JH, *et al.* Imaging hypoxia after oxygenation modification: comparing 18F-FMISO autoradiography with pimonidazole immunohistochemistry in human xenograft tumor. *Radiother Oncol* 2006;80:157-164.
- [55] Vaupel P, Thews O, Hoeckel M. Treatment resistance of solid tumors: role of hypoxia and anaemia. *Med Oncol* 2001;18:243-259.
- [56] Dewhirst MW, Ong ET, Braun RD, *et al.* Quantification of longitudinal tissue pO₂ gradients in window chamber tumours: impact on tumour hypoxia. *Br J Cancer* 1999;79:1717-1722.
- [57] Vaupel P, Kallinowski F, Okunieff P. Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumours: a review. *Cancer Res* 1989;49:6449-6465.
- [58] Durand RE, Raleigh JA. Identification of nonproliferating but viable hypoxic tumor cells in vivo. *Cancer Res* 1998;58:3547-3550.
- [59] Ljungkvist AS, Bussink J, Kaanders JH, *et al.* Hypoxic cell turnover in different solid tumor Lines. *Int J Radiat Oncol Biol Phys* 2005;62:1157-1168.
- [60] Bussink J, Kaanders JH, Rijken PF, *et al.* Changes in blood perfusion and hypoxia after irradiation of a human squamous cell carcinoma xenograft tumor line. *Radiat Res* 2000;153:398-404.
- [61] Chao KS, Bosch WR, Mutic S, *et al.* A novel approach to overcome hypoxic tumor resistance: Cu-ATSM-guided intensity modulated radiation therapy. *Int J Radiat Oncol Biol Phys* 2001;49:1171-1182.
- [62] Thomlinson RH, Gray LH. The histological structure of some human lung cancers and the possible implications for radiotherapy. *Br J Cancer* 1955;9:539-549.
- [63] Bentzen SM. Theragnostic imaging for Radiation oncology: dose painting by numbers. *Lancet Oncol* 2005;6:112-117.

3 |

Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area

Wouter V. Vogel

Dominic A.X. Schinagl

Jorn A. van Dalen

Johannes H.A.M. Kaanders

Wim J.G. Oyen

Abstract

Background and purpose: Integration of PET information into CT-based IMRT allows adaptation of the target volume to functional parameters, but only when the image registration procedure is reliable. The aim of this study was to select the optimal method for software fusion of dedicated PET and CT, and to validate the procedure for IMRT in the head-neck area.

Materials and methods: 15 patients with HNSCC underwent separate CT and FDG-PET, both in a custom-moulded rigid mask fitted with 4 multimodality fiducial markers. Five image registration methods were applied. PET emission and CT were registered manually (ME), and using the landmarks (LM). PET transmission and CT were registered manually (MT), using a mutual information-based method (MI) and using an iterative closest point method (ICP). The error in image registration using each of the methods was determined by evaluation of the markers.

Results: LM showed an average registration error of 1.4 mm at the location of the markers, and 0.3 mm in the planning area. However, this method proved to be laborious. Apart from LM the best method was ICP, with registration errors of 3.0 and 2.0 mm, respectively. For ME the respective errors were 4.7 and 3.5 mm, for MT 3.6 and 2.7 mm, and for MI 5.3 and 4.1 mm.

Conclusions: Image fusion of dedicated PET and CT of the head-neck area can be performed reliably with no need for laborious markers, using the operator-independent ICP method. The achieved accuracy permits implementation of dedicated PET images in external beam radiation therapy.

Introduction

External beam radiation therapy of head and neck squamous cell carcinoma (HNSCC) demands accurate dose delivery. Nowadays this can best be achieved with intensity modulated external beam radiation therapy (IMRT), with a good effective spatial accuracy in dose delivery and with high achievable dose gradients [1]. As a next step, integration of functional and anatomical information by fusion of PET and CT images allows subsequent consideration of biological tumor characteristics in the determination of the IMRT target volume [2-4].

For IMRT, accurate detection and localization of tumor sites is essential. Primary tumors in the head and neck area are often small at the time of discovery, and lymph node metastases tend to be small in size and multiple in number [5]. Computer Tomography (CT) imaging provides the anatomical reference and electron density information that is mandatory for 3-dimensional planning in IMRT, and can visualize structural information conveniently. Major drawbacks of CT are a low sensitivity for small lymph node metastases and low specificity in marginally enlarged lymph nodes or atypical lesions [6]. The inability of CT imaging to differentiate tissue characteristics contributes to this problem, despite the application of intravenous contrast.

Additional functional and molecular information can be provided with Positron Emission Tomography (PET). The radiopharmaceutical ^{18}F -fluorodeoxyglucose (FDG) quantitatively visualizes glucose metabolism, thus providing a tool for discrimination of normal and malignant tissues. The clinical value of FDG-PET in staging of malignancy in the head and neck area has been demonstrated by some [6-9], although others have reported less impressive results [10]. Radiotherapy planning may benefit from improved tumor detection, and from quantitative evaluation of intra-tumoral variations in metabolic activity such as glucose metabolism. Other evaluable biological parameters include hypoxia and proliferation, as visualized with ^{18}F -Misonidazole (FMISO) and ^{18}F -Fluorodeoxy-L-thymidine (FLT), respectively.

Using IMRT, a high dose can be delivered accurately to known tumor sites [11,12]. Perhaps equally important, a significant dose reduction to adjacent non-tumor sites can be achieved, resulting in fewer complications and side-effects such as central nervous system damage, mucositis, and loss of parotid gland function [13-15]. Furthermore, the 3-dimensional approach of IMRT facilitates intra-tumoral variations in dose delivery, tailored to (regional) specific functional and molecular tumor characteristics. Planning and application of such highly optimized IMRT procedures depends on precise tumor imaging. Any error in the anatomical registration of CT and PET images may lead to erroneous localization or interpretation of lesions, which may subsequently result in suboptimal radiation treatment.

Currently, many perform image fusion of CT and PET for planning of external beam radiation therapy using their own preferred approach to image registration. Some use an integrated PET/CT scanner, others apply software fusion of dedicated CT and PET. Different methods of image registration are available for the latter approach, each with their specific characteristics, and with variable accuracy of image registration. The relative accuracy of the available methods is currently not known. We have hypothesized that for successful application of image fusion in external beam radiation therapy planning – especially for IMRT – the error in image registration should not exceed existing limits in spatial accuracy of image quality and dose delivery. In this paper we evaluate different approaches to high accuracy software PET/CT image fusion, and validate the selected optimal technique for application in IMRT procedures in the head and neck area.

Materials and Methods

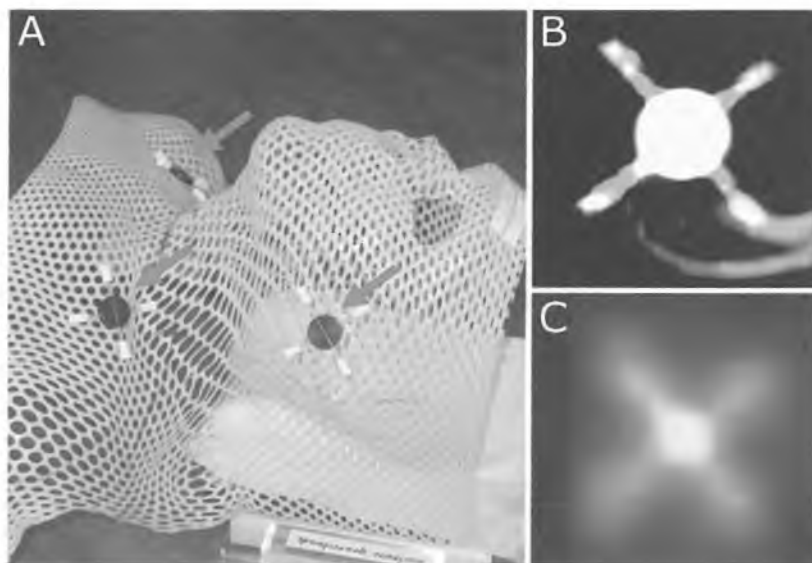
A total of 15 patients (mean age 59 years, range 49-74 years) referred for external beam radiation therapy for newly diagnosed HNSCC were included. Twelve patients had carcinoma of the larynx, 3 had carcinoma of the tongue. PET and CT scans were acquired within a maximum interval of one week.

Imaging

CT scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters were 130 kV, 120 mAs, slice distance and slice thickness 3 mm. Images were acquired from the top of the lungs to the base of the skull, with intravenous contrast.

FDG-PET was acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded. However, glucose levels had to be appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection). A 3D emission scan of the head and neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 minutes after intravenous injection of 250 MBq FDG (Mallinckrodt Medial, Petten, The Netherlands). The acquisition time per bed position was 5 minutes for emission and 3 minutes for the Germanium-based transmission, resulting in a total scanning time of 16 minutes for the two bed positions. Emission and transmission scans were reconstructed using a 2D ordered subset expectation maximization (OSEM) iterative algorithm, with parameters optimized for low photon attenuation in the head and neck area as described elsewhere [16]. As during previous studies, a structural difference in real image size between the CT and PET devices of 2.0% in the transversal direction and 0.8% in the axial direction was observed [17], this was corrected by applying a scaling factor.

Fig. 1 Multimodality markers. Example of the multi-modality markers used for landmark registration and image fusion accuracy evaluation (A) Markers (red arrows) as placed on the mask (B) Marker as seen on CT filled with iodine-containing contrast. (C) Marker as seen on PET filled with FDG solution. The center of the markers could be determined with an inter-operator variability well below 1 mm on both CT and PET images.



Patient positioning

During all imaging procedures, patients were placed in radiotherapy position within a custom-moulded rigid mask covering the head, neck and shoulders. Maximum reproducibility in positioning was assured by the use of additional support systems; e.g. a flat scanning bed, a customized head support cushion, an intra-oral mould for positioning of the tongue when needed, a standard cushion supporting the knees, and a laser positioning system.

All patients were scanned with 4 multi-modality crosshair fiducial markers firmly attached to the fixation mask in a rectangular configuration around the center of the IMRT planning area, halfway in the anterior-posterior plane. The crosshair markers consisted of two 5 cm glass capillaries positioned in a 90° angle. The capillaries were filled with either 300 mg/ml iodine-containing contrast for CT or 1 MBq/ml FDG for PET. Figure 1 shows an example of a mask with fiducial markers in place.

Image registration

The image registration procedure was performed on a PC with in-house developed image viewing and registration software based on the visualization toolkit VTK [18] and the insight segmentation and registration toolkit ITK [19]. The software allows rigid-body image registration, i.e. based on 3 translation and 3 rotation parameters. Anatomical image registration of PET images to CT images was performed using five different methods:

1. Method ME: Manual registration of PET emission images to CT. Anatomically correct registration in the field of interest was performed by an experienced investigator, through manual adjustment of the 6 free parameters. The operator used an interface where 3 variable sections of orthogonal planes (transverse, coronal and sagittal) through the 2 images were displayed simultaneously. Hence, by performing registration in only one section at the time, the 3D registration problem was reduced to a series of 2D problems. The method is therefore fast and easy, but may suffer from operator-dependency.
2. Method LM: Landmark-based registration of PET emission images to CT. Registration was optimized by manual identification of the centers of the multimodality markers, followed by a minimization of $\sum_{i=1}^N ((r_i - T * r'_i)^2) / N$, where r_i are the coordinates of CT landmark i , r'_i are the coordinates of PET landmark i , T is the transformation, and $N=4$ for the number of markers. This minimization provides (by taking the square-root of this minimization result) the fiducial registration error (FRE) [20]. This method is subject to operator-dependency due to the manual localization of the markers.
3. Method MT: Manual registration of PET transmission images to CT. Parallel to method ME, anatomically correct registration in the field of interest was pursued through manual adjustment of the 6 free parameters. At the end of the procedure the transmission images were substituted with the emission images. This method is also fast and easy, but may again suffer from operator-dependency.
4. Method MI: Mutual information based registration of PET transmission images to CT. This method optimizes a functional measuring the similarity of all geometrically corresponding voxel pairs for some feature. The mutual information metric implementation follows the method as specified by Voila and Wells [21,22]. In this implementation, probability densities are estimated from the image data using the Parzen-Window scheme [23]. This method is available in ITK [19]. Its main advantages are that it can work directly with image data as no pre-processing or segmentation is needed. Furthermore, it has an efficient implementation based on stochastic approximation. The parameters of this method [19] have been tuned to our application: the Parzen window width was set to 2, the number of samples to 50, the learning rate to 0.0005, the translation scale to 100 and the number of iterations to 20,000. This provided robust results. At the end of the registration procedure the

transmission images were substituted with the emission images. This method is not operator dependant, but does require some computation time.

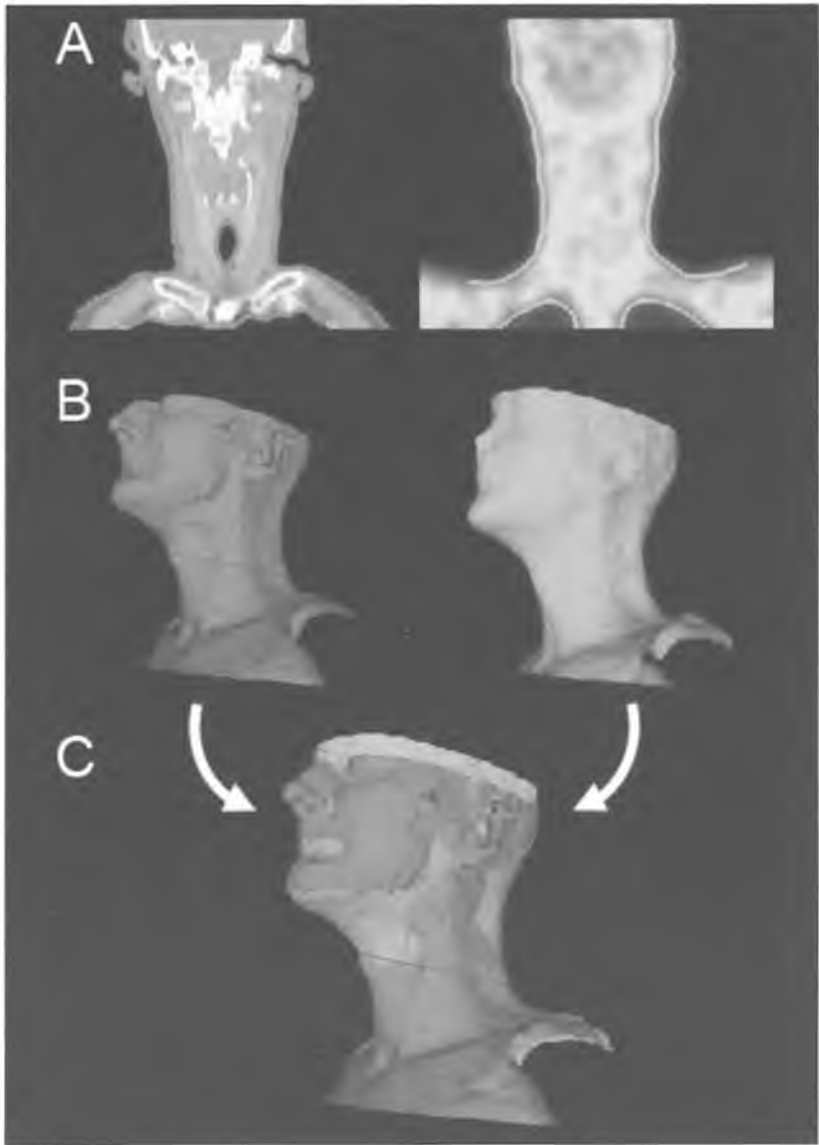
5. Method ICP: Iterative closest point registration, using surface models of the body contours acquired from PET transmission and CT. A thresholding technique was used on both image sets to create triangulated iso-surfaces of the body, using the VTK method `vtkContourFilter` [18]. Normals were computed for these surfaces for each of the polygonal facets, using the VTK method `vtkPolyDataNormals` [18]. Using these normals, a faceted shading of the surface was obtained that was used as an input for the ICP algorithm. The ICP algorithm [19] has three stages and iterates. Firstly, random points on the first model (of the transmission PET) are associated to random points on the other model (of the CT) by the nearest neighbor criterion. Secondly, transformation parameters are estimated using a mean square cost function. Next, the transformation is applied. Then, the iteration procedure starts by re-associating points, etc. The procedure continues until convergence has been reached. In this study 5000 random points were used for each model and the convergence criterion was to stop when the change between two successive iterations fell below a threshold of 0.001 mm, or when the number of iterations exceeded 3000. At the end of the procedure the transmission PET images were substituted with the emission images. The ICP method is illustrated in figure 2. This method is not operator dependant, but does require some computation time.

For the automatic registration methods MI and ICP the algorithm input was restricted to a volume of interest by defining a 3-dimensional box, containing the head and neck area, but excluding bodyparts that tend to keep freedom of movement within the mask (i.e. the shoulders). With the exception of the landmark-based method LM, the fiducial markers were not considered during the image registration process. In the PET transmission images the markers were not visible, and all representations of the markers were removed from the emission images prior to image registration.

Assessment of image registration

The fiducial markers were used for assessment of the accuracy in image registration. The location of the center of each marker was determined on both CT and PET emission images. The inter-operator variation of the manual localization of the center of the markers was evaluated by analysis of 12 markers in a separate session by two operators. The difference in position of corresponding markers on CT and PET was determined, representing the error in image registration at the location of the markers. Subsequently, for each patient the location of a hypothetical marker in the center of the IMRT planning area was determined, by calculating the geometrical center of all surrounding markers. The difference in position of this hypothetical marker on CT and PET was determined to estimate the image registration error in the region of the IMRT planning area.

Fig. 2 Iterative closest point registration (ICP) of PET transmission images to CT (A) On both CT and the PET transmission scan, a surface model of the body contour is generated using a thresholding technique (B) Both surface models are represented as a 3-dimensional structure. (C) Automatic image registration is performed by iteratively minimizing the distance between the surface models.



Statistical analysis

Differences in the image registration errors for the applied methods were evaluated using a repeated measures ANOVA with Bonferoni's correction for multiple comparisons, using GraphPad InStat version 3 (GraphPad Software, San Diego CA, USA). The level of significance was set at 0.05.

Results

The inter-operator variability in determination of the center of the markers was well below 1 mm in both types of scans. For CT images the average difference between the two operators was 0.54 mm (S.D. 0.28 mm). For PET images the average difference was 0.42 mm (S.D. 0.19 mm).

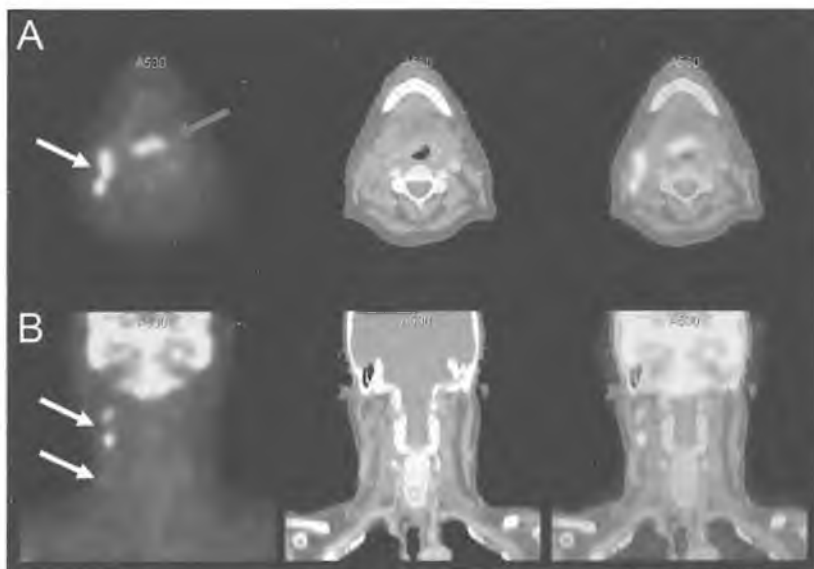
Application of LM achieved the best results ($P < 0.001$ as compared to all other methods). LM showed an average registration error of 1.4 mm (range 0.3 – 3.8, SD 0.8) at the location of the markers, and a calculated average registration error in the center of the planning area of 0.3 mm (range 0.0 – 0.6, SD 0.2).

The second-best method proved to be ICP, with an average error in image registration of 3.0 mm (range 0.5 – 8.8, SD 0.9) at the location of the markers and 2.0 mm (range 0.6 – 4.3, SD 1.1) at the center of the planning area. The average error in the center was above the limit of 3 mm in three cases (20%), the largest error in the planning area being 4.3 mm.

For ME the respective registration errors were 4.7 mm (range 1.2 – 13.6, SD 2.2) and 3.5 mm (range 0.7 – 6.3, SD 1.6), for MT 3.6 mm (range 0.5 – 10.6, SD 1.5) and 2.7 mm (range 1.1 – 6.1, SD 1.6), for MI 5.3 mm (range 1.0 – 18.3, SD 2.0) and 4.1 mm (range 1.9 – 6.8, SD 1.4). As compared to ICP, ME and MI were significantly less accurate in image registration ($P < 0.05$). MT was not significantly different as compared to method ICP. An example of the results after image registration is shown in figure 3. The quantitative results are depicted in figure 4.

During the PET image acquisition, small positioning problems occurred in 4 patients, due to inexperience of the PET operating personnel with the fixation mask. Examples are a minor backwards tilt of the head within the mask, or an incorrectly attached part of the mask near a shoulder. This resulted in minor visually discernible positioning differences between PET and CT scans, in all cases below 1 cm and mostly well outside the IMRT planning area. All 3 patients who showed a registration error above 3 mm using method ICP were subject to such errors.

Fig. 3 Example of image fusion using method ICP Transverse (A) and coronal (B) sections of FDG-PET (left), CT (middle) and fused PET-CT (right) in a patient with carcinoma at the base of the tongue (blue arrow) and multiple pathological lymph nodes on the right side of the neck (white arrows). The extent of the malignancy was difficult to appreciate on the CT, while fused PET/CT images allowed clear delineation of the primary tumor, as well as identification of pathological lymph nodes.



Discussion

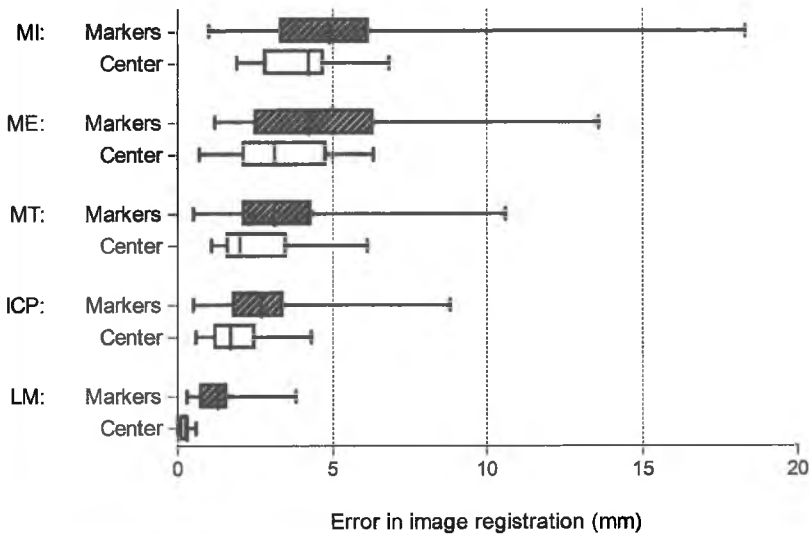
Image fusion may be performed with software-based image registration of dedicated PET and CT. However, image registration of dedicated PET and CT has a risk for introducing errors [24]. Although high accuracy in image registration has been shown using phantoms [25], a time interval between scans and repeated positioning may result in introduction of differences when imaging patients. Furthermore, the image registration procedure itself is hampered by limited visualization of normal anatomical structures on FDG-PET. The risk for potentially significant errors in PET/CT image registration emphasizes the need for a structured and validated approach.

At least three fiducial markers are needed to define position and orientation of a rigid object in a 3-dimensional space. In these studies 4 markers were used, in a rectangular configuration in the coronal plane around the planning area to equally represent the whole head/neck area. The anterior and posterior sides of the mask were not marked with

separate markers, but these areas were considered not very prone to additional local positioning errors because of rigid attachment to the scanning table and adequate fixation of the nose.

Fiducial markers can also be used for evaluation of the local image registration error. This local error is influenced by non-rigid positioning differences within the mask. The extent of such positioning differences is unclear. According to van Lin et al. the error in repeated positioning in a customized mask is in the range of 3–4 mm [26], but this includes other factors such as visual and manual correlation with laser guides. In this paper, the relatively good results of image registration suggest that on average patient positioning was adequate, and well within the range found by van Lin et al. Nevertheless, some of the patients were indeed subject to slightly suboptimal positioning, due to relative inexperience of PET personnel in positioning of the customized mask. This suggests that the accuracy of the image registration may improve even further, as experience grows.

Fig. 4 Accuracy of image registration. The errors in image registration, as performed by the different applied techniques (MI = mutual information, ME = manual emission images, MT = Manual transmission images, ICP = Iterative closest point, LM = landmark). The graph shows that the lowest error in the center of the IMRT field can be achieved using the Landmark-based method (LM). From the methods that do not depend on fiducial markers, the ICP method proves the most accurate.



In this study, the errors in image registration were measured by markers placed well outside the of the target volume for radiotherapy, which can be considered a 'worst case scenario'. The much more relevant error in the center of the planning area was estimated by the error in a mathematically derived hypothetical marker, defined by the geometrical center of all available surrounding markers. A drawback of this method is that it represents a theoretical value. However, there is no reliable direct method to evaluate the errors in image registration within the neck of a patient, due to the lack of anatomical landmarks that are well demarcated both on CT and PET transmission images, the unfeasibility of marker placement in such areas, and unavoidable physiologic processes such as swallowing. Wong et al. have described anatomical landmarks that may be identified in the head and neck area on both PET and CT [27], but we considered the visibility of such landmarks very poor, and also considered the suggested accuracy of localization of 1.3 – 8.2 mm insufficient. Furthermore, the artificial markers were chosen as these do not depend on tracer biodistribution and uptake, and as manual localization will be less operator dependent. Therefore, the calculated error in the center of the planning area, derived from fiducial markers at the edges, was considered the best available indicator of the real error in image registration.

Fitzpatrick et al. have previously published a method for quantitative evaluation of registration errors using recognisable landmarks [28,29]. The so-called "Target Registration Error" (TRE) expresses the displacement between any two corresponding points in the images, in relation to the fiducial registration error (FRE) of available landmarks. This TRE is considered to be the most accurate mathematical representation of the real registration error at a specific point. Given the configuration and number of the markers in our study, the calculated average position of the hypothetical marker in the center of the planning area approaches the local TRE as defined by Fitzpatrick. The error in areas further away from the planning area (e.g. at the anterior and posterior borders of the head and neck area) will be somewhat underestimated, but this is considered not very relevant for treatment planning for areas in the center field of view.

The spatial resolution of PET imaging is in the range of 4 – 7 mm (full width half maximum) for most currently available scanners, although the image quality may be better in the head and neck area where photon attenuation is low [16]. The geometrical uncertainty in IMRT dose delivery is in the range of 2 – 3 mm [1]. Therefore, an upper limit for the error in image registration of 3 mm can be considered acceptable for application in IMRT planning.

The best results in image registration accuracy were achieved by the landmark-based method. These values were considered a 'gold standard' in rigid-body image registration accuracy as they reflect the mathematically achievable accuracy, including all unavoidable

positioning differences, mask deformations, patient motion due to e.g. swallowing, and sampling errors in determining the position of the markers. The average registration error at the location of the landmarks is better than the values observed when using anatomical landmarks by Wong et al., probably because the manual localization procedure of artificial markers is more accurate and less operator independent [27]. Furthermore, our results with LM are very similar to those reported with integrated PET/CT scanners, which provide an excellent tool for image fusion for IMRT planning in the head and neck area [30,31].

In practice, LM proved to be relatively laborious and cumbersome, as for each procedure the markers had to be filled manually, placed on the masks and subsequently removed again. This may be overcome by using commercially available multimodality markers, containing a solid long lived PET source. However, an adequate method with no need for markers may still be preferable.

The ICP method proved to be the second best method, as compared to the landmark-based registration method. With ICP, the average error in image registration was 2.0 mm in the center of the planning area, which was within the limits of accuracy of IMRT [1]. ICP proved to be significantly better than the MT and mutual information based methods. A further advantage of ICP over the manual methods is the operator independency. An advantage over the mutual information method is the independency of initial registration differences between PET and CT. The results of the ICP method in this study were better than previously published results of software PET/CT fusion of the brain in a single patient using a mutual information algorithm by Lavelly et al. [25]. Obviously, a direct comparison with the head-neck area is not possible. Also, Lavelly et al. did show more precise registration in a phantom, but it is not very realistic to extrapolate these results to clinical procedures.

Using ICP, three patients still showed an error of more than 3 mm in image registration in the center of the planning area. In these cases, less experienced personnel positioned the patient in the fixation mask. Thus, it is likely that a smaller registration error of the ICP method (and possibly the other methods) than reported in our series is feasible, when experience in patient positioning increases. This also illustrates that patient positioning could be the limiting factor in image registration, rather than the ICP method itself. This supports the conclusion that PET-CT image registration using the ICP method can be applied in the IMRT procedure.

The method ME performed significantly less accurate than the ICP method. Furthermore, as opposed to the ICP method, both manual registration methods (ME and MT) are operator-dependent, which is a clear disadvantage. Therefore, the manual methods were considered suboptimal for procedures in the head and neck area as compared to ICP.

The performance of the MI based algorithm was somewhat disappointing, as the method performed significantly less accurate than the ICP method, with an average error larger than 3 mm. We have observed that the algorithm consistently converges to a reasonable registration, but tends to deviate slightly in the final registration result as an intelligent evaluation of edges and symmetry seems to be lacking, as opposed to the ICP method. Furthermore, the results may vary with the extent and direction of the registration mismatch that the MI algorithm is confronted with initially. The results of an MI-based algorithm depend on several parameters, such as the number of iterations. We fine-tuned our parameters, but were unable to improve the results to the level of the ICP results. Therefore, we consider the Parzen Window implementation of the MI method suboptimal for procedures in the head and neck area as compared to ICP.

The ICP method in this study was based on registration of PET transmission images to CT. This implies that PET without transmission images cannot be registered using the ICP method. This is not a relevant problem, as PET-based quantitative tissue characterization for IMRT already requires correction for photon attenuation. A further important advantage of using transmission images is the independence of the choice of the radiopharmaceutical (i.e. FLT, F-Misonidazole). This allows image registration also when delineation of anatomical landmarks in the emission images is relatively poor and tumor uptake is low. Theoretically, image registration based on transmission images may fail in case of patient motion between the acquisition of emission and transmission, but the use of a fixation mask will prevent such problems. The fixation mask itself is visible on CT images and may slightly influence PET transmission images, despite its limited thickness and density. Visualization of the mask on PET transmission images would be homogeneous and symmetrical, and is not considered a problem for image registration. Especially the ICP method is unaffected by symmetrical factors, such as the threshold level on PET images. The evaluation of image registration with fiducial markers is not influenced by the presence of the mask.

The accuracy and robustness as achieved by the observer independent ICP method has not been published before. Only a few studies have evaluated the accuracy of rigid image registration in the head and neck area. As stated above, one study has shown the feasibility of image fusion in 30 patients using manually selected anatomical landmarks [27], although the method is very operator dependent and the average registration error of 3.8 mm (range 1,3 – 8,2 mm) is considered insufficient for high-accuracy procedures such as IMRT. Another study by Klabbers et al. has also demonstrated the feasibility of mutual information based registration with transmission images [32], but the image registration was only evaluated for the object as a whole using a full-circle method, which provides no information about local and systematic registration errors. The claimed average registration error of 4 mm is exclusive additional errors due to rotation, and neglects local

errors due to positioning differences. A study by Nishioka et al. used image registration based on the brain only, and used visual evaluation only [33].

We have evaluated 5 different approaches to PET-CT image registration. More techniques are available, as published in an extensive listing by Maintz et al. [20]. In this study, a method has been selected from each category of techniques, unless inappropriate (i.e. not applicable for inter-modality registration). Other techniques are not yet established, such as non-rigid (elastic) registration.

The high accuracy of image registration presented in this study is only valid for rigid-body image registration. This limits the applicability of this approach to IMRT planning in the head and neck area, where measures can be taken to ensure reproducible positioning. It is likely that image registration of non-rigid body parts where immobilization and positioning tools are less effective – such as the chest and abdomen – will be less accurate, or will require different approaches to image registration, tailored to the specific situation.

For hybrid PET/CT scanning, phantom studies have demonstrated an image registration accuracy in the range of 0-2 mm, and for patients in the range of 1-3 mm in the head and neck area [30]. Furthermore, patient repositioning between the acquisition of PET and CT images is no longer required. However, many hospitals do not have access to an integrated PET/CT scanner system. As the clinical demand for functional imaging in radiation therapy planning is high, also in hospitals without a PET/CT scanner, validation of software image fusion remains of relevance.

Conclusions

High-accuracy rigid-body image registration of dedicated PET and CT of the head and neck area can be performed adequately and reliably, with or without multi-modality markers. Without using markers, the iterative closest point registration method of CT and PET-transmission images proved to be the most accurate operator-independent image registration method. The achieved accuracy – with an average error lower than 3 mm in image registration – permits implementation of dedicated PET images in IMRT planning and therapy. Our results validate the use of software image fusion of PET and CT for IMRT planning in the head and neck area, thus permitting application of this technique whenever dedicated PET and CT are available.

References

- [1] Low DA, Mutic S, Dempsey J, *et al.* Quantitative dosimetric verification of an IMRT planning and delivery system. *Radiother Oncol* 1998;49:305-316
- [2] Ling CC, Humm J, Larson S, *et al.* Towards multidimensional Radiotherapy (MD-CRT): biological imaging conformality. *Int J Radiat Oncol Biol Phys* 2000;47:551-560
- [3] Scarfone C, Lavelly WC, Cmelak AJ, *et al.* Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. *J Nucl Med* 2004;45:543-552
- [4] Solberg TD, Agazaryan N, Goss BW, *et al.* A feasibility study of 18F-fluorodeoxyglucose positron emission tomography targeting and simultaneous integrated boost for intensity-modulated radiosurgery and Radiotherapy. *J Neurosurg* 2004;101 suppl3:381-389
- [5] Stuckensen T, Kovacs AF, Adams S, *et al.* Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. *J Craniomaxillofac Surg* 2000;28:319-324
- [6] Bruschini P, Giorgetti A, Bruschini L, *et al.* Positron emission tomography (PET) in the staging of head and neck cancer: comparison between PET and CT. *Acta Otorhinolaryngol Ital* 2003;23:446-453
- [7] Braams JW, Pruim J, Freling NJ, *et al.* Detection of lymph node metastases of squamous cell cancer of the head and neck with FDG-PET and MRI. *J Nucl Med* 1995;36:211-216
- [8] Laubenbacher C, Saumweber D, Wagner-Manslau C, *et al.* Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous cell carcinomas. *J Nucl Med* 1995;36:1747-1757
- [9] Teknos TN, Rosenthal EL, Lee D, *et al.* Positron emission tomography in the evaluation of stage III and IV head and neck cancer. *Head Neck* 2001;23:1056-1060
- [10] Stoeckli SJ, Steinert H, Pfaltz M, *et al.* Is there a role for positron emission tomography with 18F-fluorodeoxyglucose in the initial staging of nodal negative oral and oropharyngeal squamous cell carcinoma. *Head Neck* 2002;24:345-349
- [11] DeNeve W, Duthoy W. Intensity-modulated Radiation therapy for head and neck cancer. *Expert Rev Anticancer Ther* 2004;4:425-434
- [12] Zhou J, Fei D, Wu Q. Potential of intensity-modulated Radiotherapy to escalate doses to head and neck cancers: what is the maximal dose? *Int J Radiat Oncol Biol Phys* 2003;57:673-682
- [13] Lin A, Kim HM, Terrel Je. Quality of life after parotid-sparing IMRT for head and neck cancer: a prospective longitudinal study. *Int J Radiat Oncol Biol Phys* 2003;57:61-70
- [14] Parliament MB, Scrimger RA, Anderson SG. Preservation of oral health-related quality of life and salivary flow rates after inverse-planned intensity modulated Radiotherapy (IMRT) for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2004;58:663-673
- [15] Cozzi L, Fogliata A, Bolsi A, *et al.* Three-dimensional conformal vs. Intensity-modulated Radiotherapy in head and neck cancer patients: comparative analysis of dosimetric and technical parameters. *Int J Radiat Oncol Biol Phys* 2004;58:617-624
- [16] Vogel WV, Wensing BM, van Dalen JA, *et al.* Optimized PET reconstruction of the head and neck area: improved diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2005;32:1276-1282
- [17] Vogel WV, van Dalen JA, Schinagl DA, *et al.* Correction of an image size difference between PET and CT improves image fusion of dedicated PET and CT. *Nucl Med Commun* 2006;27:515-519
- [18] Schroeder W. The visualization toolkit: an object-oriented approach to 3D graphics, 3rd edn New York: Kitware Inc 2003
- [19] Ibanez L, Schroeder W, Ng L, *et al.* The ITK Software Guide: The Insight Segmentation and Registration Toolkit, version 1.4 New York: Kitware Inc 2003
- [20] Maintz JB, Viergever MA. A survey of medical image registration. *Med Image Anal* 1998;2:1-36
- [21] Viola P, Wells WM. Alignment by maximization of mutual information. *Proc 5th Int Conf on Computer Vision (Boston)* (Cambridge MA: MIT Press) 1995;16-23
- [22] Viola P, Wells WM. Alignment by maximization of mutual information. *Int J of Computer Vision* 1997;137-154
- [23] Duda and Hart P. Pattern Classification and Scene Analysis (New York: Wiley)
- [24] Vogel WV, Oyen WL, Barentsz JO, *et al.* PET/CT: panacea, redundancy, or something in between? *J Nucl Med* 2004;45 suppl1:155-245

- [25] Lavelly WC, Scarfone C, Cevikalp H, *et al.* Phantom validation of coregistration of PET and CT for image-guided Radiotherapy. *Med Phys* 2004;31:1083-1092
- [26] Van Lin EN, Huizenga H, Kaanders JH, *et al.* Set-up improvement in head and neck Radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support. *Radioth Oncol* 2003;68:137-148
- [27] Wong WL, Hussain K, Chevetton E, *et al.* Validation and clinical application of computer-combined computed tomography and positron emission tomography tomography with 2-[18F]fluoro-2-deoxy-D-glucose head and neck images. *Am J Surg* 1996;172:628-632
- [28] Fitzpatrick JM, West JB, Maurer CR. Predicting error in rigid-body point-based registration. *IEEE trans Med Imaging* 1998;17:694-702
- [29] Fitzpatrick JM, Hill D, Maurer CR. Image registration. *Handbook of Medical Imaging* (Bellingham, Washington: SPIE Press);447-513
- [30] Ciernik IF, Dizendorf E, Baumert BG, *et al.* Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863
- [31] Heron DE, Andrade RS, Flickinger J, *et al.* Hybrid PET-CT simulation for radiation treatment planning in head and neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424
- [32] Klabbers BM, de Munck JC, Slotman BJ *et al.* Matching PET and CT scans of the head and neck area: development of method and validation. *Med Phys* 2002;29:2230-2238
- [33] Nishioka T, Shiga T, Shirato H, *et al.* Image fusion between 18FDG-Pet and MRI/CT for Radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057

4 |

Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose positron emission tomography based target volume definition in head and neck cancer

Dominic A.X. Schinagl
Wouter V. Vogel
Aswin L. Hoffmann
Jorn A. Van Dalen
Wim J.G. Oyen
Johannes H.A.M. Kaanders

Abstract

Background and purpose: Target volume delineation for radiation treatment in the head and neck area is traditionally based on physical examination, CT, and MRI. Additional molecular imaging with FDG-PET may improve definition of the gross tumor volume (GTV). In this study, five methods for tumor delineation on FDG-PET are compared with CT-based delineation.

Materials and methods: Seventy-eight patients with a stage II-IV squamous cell carcinoma of the head and neck area underwent co-registered CT and FDG-PET. The primary tumor was delineated on CT (GTV_{CT}), and five PET-based GTVs were obtained: visual interpretation (GTV_{VIS}), applying an isocontour of a standardized uptake value of 2.5 (GTV_{SUV}), using a fixed threshold of 40% and 50% of the maximum signal intensity ($GTV_{40\%}$ and $GTV_{50\%}$), and applying an adaptive threshold based on the signal to background ratio (GTV_{SBR}). Absolute GTV volumes were compared, and overlap analyses were performed.

Results: The GTV_{SUV} method failed to provide successful delineation in 45% of cases. For the other PET delineation methods, the volume and shape of the GTV were heavily influenced by the choice of the segmentation tool. On average, all threshold-based PET-GTVs were smaller than on CT. Nevertheless, PET frequently detected significant tumor extension outside the GTV_{CT} (15%-34% of PET-volume).

Conclusions: The choice of a segmentation tool for target volume definition of head and neck cancer based on FDG-PET images is not trivial, as it influences both volume and shape of the resulting GTV. With adequate delineation, PET may add significantly to CT and physical examination-based GTV definition.

Introduction

Progress in radiation oncology enables delivery of radiation treatment with increasing geometric precision. This requires a re-evaluation of target volume delineation, which is traditionally based on physical examination, computed tomography (CT) and magnetic resonance imaging (MRI). In recent years, new methods have been introduced for visualization of tumor tissue. In addition to anatomical data as supplied by CT and MRI, 'functional' and 'molecular' imaging techniques, such as positron emission tomography (PET), single-photon emission computed tomography (SPECT) and magnetic resonance spectroscopy (MRS), allow visualization of biological characteristics, with several potential advances. The primary tumor may be identified more accurately, with consequences for the size and shape of the gross tumor volume (GTV). Tumor characteristics relevant for radiation sensitivity can be visualized (e.g. hypoxia), which may assist in the selection of patients for customized treatments [1]. Also, intra-tumoral heterogeneity of these characteristics may be identified, providing an opportunity for 'dose painting' [2]. Finally, when imaging modalities become more accurate, the inter- and intra-observer variations in tumor delineation will decrease, resulting in improved standard of care.

Metabolic information, as provided by imaging ^{18}F -fluoro-deoxy-glucose (FDG) with PET, has been incorporated into target volume delineation by many groups [3]. Tumor localizations can be identified and localized with high sensitivity, due to the high contrast resolution of PET. However, application of FDG-PET data for target volume delineation is not straightforward, as identification of tumor boundaries on PET suffers from a relatively low spatial resolution and a 'blurry' appearance of lesions. Furthermore, FDG-PET is usually interpreted qualitatively in diagnostic nuclear medicine, whilst in radiation oncology a more quantitative approach is required for tumor contouring [4]. Currently, various methods for FDG-PET based target volume definition are in use. Visual interpretation is the most commonly used method [5-12]. This method, however, is susceptible to the window-level settings of the images and is highly operator dependent. Therefore, other – more objective – methods have been explored. Examples are isocontouring based on either a standardized uptake value (SUV) of 2.5 around the tumor [10, 13, 14], a fixed threshold of the maximum signal intensity [15-21], or on a threshold which is adaptive to the signal to background ratio [22]. The utility of these methods for tumor delineation in the head and neck area is currently unknown.

The choice of method for tumor delineation on FDG-PET may influence GTV determination, with consequences for the outcome of radiation therapy. The aim of this study was to compare different methods for tumor delineation with FDG-PET, relative to CT-based delineation, for radiation therapy planning in head and neck cancer patients.

Materials and methods

Patients

Seventy-eight patients (59 males and 19 females, median age 61 years, range 43-86 years) with stage II-IV squamous cell carcinoma of the head and neck area, eligible for primary curative radiotherapy, were prospectively enrolled from June 2003 until July 2006. The tumor characteristics are summarized in table 1. The study was approved by the Ethics Committee of the Radboud University Nijmegen Medical Centre and all patients provided informed consent.

Table 1 Tumor characteristics

Tumor site	Oral cavity	6
	Oropharynx	31
	Hypopharynx	9
	Larynx	32
T stage	T1	1
	T2	16
	T3	39
	T4	22
Total		78

Image Acquisition

Prior to treatment, a CT scan and an FDG-PET scan were acquired in radiation treatment position, the patient being immobilized by a custom-made rigid mask covering head, neck and shoulders. Maximum reproducibility in positioning was assured by the use of additional support systems: a flat scanning bed, a customized head support cushion, an intra-oral mould when indicated, a standard cushion supporting the knees, and a laser positioning system. The median interval between CT and FDG-PET was 3 days (range 0-10 days). CT scan was always performed prior to PET-scan. CT scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters were 130 kV, 120 mAs, slice distance and slice thickness 3 mm, scanning from above the frontal sinuses to below the clavicles, with intravenous contrast.

FDG-PET was acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded. However, glucose levels had to be appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection). A 3D

emission scan of the head and neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 minutes (median 64 minutes, S.D. \pm 11.4) after intravenous injection of 250 MBq FDG (Mallinckrodt Medial, Petten, The Netherlands). The acquisition time per bed position was 5 minutes for emission and 3 minutes for the Germanium-based transmission scan, resulting in a total scanning time of 16 minutes for the two bed positions. Emission and transmission scans were reconstructed using a 2D ordered subset expectation maximization (OSEM) iterative algorithm, with parameters (4 iterations, 16 subsets) optimized for low photon attenuation in the head and neck area as described elsewhere [23].

The image registration procedure has been described in detail previously [24]. In brief, three-dimensional surface models were automatically derived from both the CT and the PET transmission images. These models were anatomically co-registered using an operator-independent iterative closest point algorithm, with an average registration error of 2.0 mm at the centre of the planning area. Afterwards, the PET transmission images were replaced with the PET emission images. In addition, a second PET data set was generated in which the original values were replaced with calculated SUV values. SUV was defined as the voxel value of detected activity (in Bq/ml) multiplied by the weight of the patient (in kg) divided by the activity at the beginning of the scan (in Bq) multiplied by 1000 [25]. The CT and the two PET data sets were transferred via DICOM to a Pinnacle³ treatment planning system (Philips Medical Systems, Andover, MA, USA) for target volume definition and subsequent volume analysis.

Target Volume Definition

The primary tumor was delineated on CT and FDG-PET images by two experienced radiation oncologists in consensus. The role of FDG-PET in detection of metastatic lymph nodes will be the subject of a separate analysis.

On CT images, manual delineation of the GTV_{CT} was performed according to current clinical protocols, using information gathered from physical examination, available diagnostic workup imaging modalities (CT and/or MRI, examination under general anesthesia) and the CT in treatment position. When the radiation oncologists were drawing the GTV_{CT} contours, the FDG-PET images were blinded.

Five PET-based GTVs were obtained using different delineation approaches. Visual delineation (GTV_{vis}) was performed by contouring FDG activity clearly above normal background activity. Localizations with increased FDG uptake were classified malignant in consensus with an experienced nuclear medicine physician. The other (threshold based) GTVs were obtained using in-house developed software scripts for the Pinnacle³ treatment planning system. SUV-based delineation was obtained by applying an isocontour of

$SUV=2.5$ (GTV_{SUV}) around the tumor. Two thresholds were based on fixed percentages of the maximum signal intensity in the primary tumor, of 40% ($GTV_{40\%}$) and 50% ($GTV_{50\%}$) respectively. Finally, an adaptive threshold delineation (GTV_{SBR}) based on the signal-to-background ratio (SBR) was performed, as developed at Université St. Luc in Brussels, Belgium [22]. The maximum signal intensity was defined as the mean activity of the hottest voxel and its eight surrounding voxels in a transversal slice. The mean background activity was obtained in a manually defined region of interest of approximately 10 cm^3 in the left neck musculature, far away from the primary tumor and any involved lymph nodes. Prior to delineation, scanner-specific variables that were needed for calculation of the GTV_{SBR} were derived by a phantom experiment as described by Daisne *et al.* [22]. In brief, hot spheres with different sizes in a 6.5L Jaszczak phantom were imaged at different image contrast ratios. Optimal delineation thresholds were determined by minimizing the square difference between true and measured sphere volumes. These results were used to find the parameters a and b in the algorithm: $\text{Threshold}=a+bx1/SBR$ (in our setting $a=44.1$, and $b=70.4$). These data were consistent in replicate experiments.

For all FDG-PET delineation methods, air cavities were excluded from the GTV. This was achieved using automatic contouring of air cavities on CT (Hounsfield units 0-900) and subsequent subtraction from the GTV using an in-house developed Pinnacle³ script. Results obtained by automated delineation algorithms were checked visually before acceptance. A delineation was considered unsuccessful if the resulting GTV included significant volumes of tissue that were clearly normal at visual interpretation.

Volume analysis

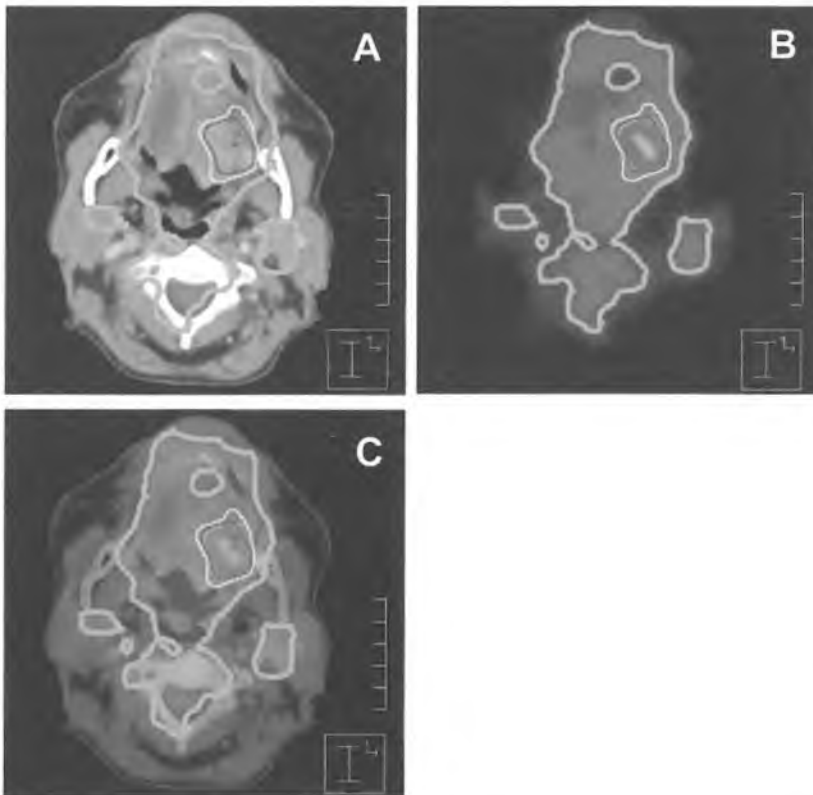
Absolute volumes of different GTVs were compared using paired t-test. In addition, for all PET GTVs, in-house developed scripts were used to calculate: 1. the overlap volume of GTV_{CT} and GTV_{PET} , where overlap was expressed as the overlap volume of GTV_{CT} and GTV_{PET} relative to GTV_{CT} (overlap fraction_{CT} (OF_{CT})), and as the overlap volume of GTV_{CT} and GTV_{PET} relative to the five PET-based GTVs (overlap fraction_{PET} (OF_{PET})); 2. the volume enclosed by GTV_{CT} but not by GTV_{PET} relative to GTV_{CT} , which is $1 - OF_{CT}$; and 3. the volume enclosed by GTV_{PET} but not by GTV_{CT} relative to GTV_{PET} , which is $1 - OF_{PET}$. Correlations between overlap fractions were assessed by linear regression analysis.

Results

Seventy-eight patients were included in this study. Of these, 77 data sets were available for analysis; one patient was excluded as the primary tumor, a T2N2cM0 oropharyngeal carcinoma, was not visualized by FDG-PET.

The GTV_{VIS} could be generated for all 77 patients. The GTV_{SBR} segmentation tool resulted in an unsuccessful volume definition in two patients. This was observed in four patients for both the $GTV_{40\%}$ and the $GTV_{50\%}$, two of whom also had an unsatisfactory GTV_{SBR} . The GTV_{SUV} determination was not successful in 35 patients, including the four patients mentioned above. As a consequence, this latter method was not further evaluated. Unsuccessful delineation was not correlated with specific tumor subsites or T-stages. All unsatisfactory volumes were largely oversized, being at least 300 cm^3 . An example of an inadequate GTV_{SUV} is depicted in figure 1.

Fig. 1 CT scan (A), corresponding FDG-PET scan (B) and fused image (C) of a patient with a T4N2M0 tongue carcinoma, showing differences in target volume definition. Indicated are GTV_{CT} (red), GTV_{VIS} (light green), GTV_{SUV} (orange), $GTV_{40\%}$ (yellow), $GTV_{50\%}$ (blue), and GTV_{SBR} (dark green). GTV_{SUV} is unsuccessful in this case, due to inclusion of large areas of normal background tissue. Note that all other PET-based delineations indicate greater tumor extension towards lateral side and less towards medial side compared to CT delineation. Also note that on this transversal slice $GTV_{40\%}$ and GTV_{SBR} are indistinguishable.



The mean absolute volumes obtained with the various delineation procedures are shown in table 2 and figure 2. GTV_{CT} and GTV_{VIS} yielded comparable volumes but the three threshold-based PET-GTVs (40%, 50% and SBR) all were smaller than the CT-based GTV ($p \leq 0.0001$ for all comparisons). Furthermore, $GTV_{50\%} < GTV_{40\%} < GTV_{VIS}$ ($p \leq 0.0003$), indicating that these methods resulted in significantly differently sized GTVs. The mean volumes of $GTV_{50\%}$ and GTV_{SBR} were very similar.

Table 2 GTV volume and overlap fractions for various segmentation tools

	mean absolute volume (95% CI) in cm^3	mean OF_{CT} (95% CI)	mean OF_{PET} (95% CI)
GTV_{CT}	22.7 (17.4-27.9)	-	-
GTV_{VIS}	21.5 (16.5-26.6)	0.61 (0.56-0.66)	0.66 (0.63-0.7)
$GTV_{40\%}$	16.4 (13.2-19.6)	0.55 (0.51-0.58)*	0.72 (0.67-0.77) ^o
$GTV_{50\%}$	10.5 (8.2-12.7)	0.39 (0.36-0.43)**	0.80 (0.76-0.85) ^{oo}
GTV_{SBR}	11.2 (8.2-12.9)	0.43 (0.35-0.51)***	0.85 (0.74-0.97) ^{ooo}

CI: Confidence Interval

OF_{CT} : Overlap Fraction_{CT}

OF_{PET} : Overlap Fraction_{PET}

p-values (relative to GTV_{VIS}).

* $p = 0.0007$, ** $p < 0.0001$, *** $p < 0.0001$, ^o $p = 0.01$, ^{oo} $p < 0.0001$, ^{ooo} $p = 0.001$

The results of the overlap analyses are also shown in table 2. The mean OF_{CT} varied from 0.39 to 0.61, depending on the segmentation tool used. The mean OF_{PET} varied from 0.66 to 0.85. A clear trend was observed with OF_{CT} decreasing and OF_{PET} increasing, from GTV_{VIS} to $GTV_{40\%}$ to $GTV_{50\%}$ to GTV_{SBR} . This indicates that in this order the PET volume not only decreased but was also increasingly incorporated within the CT volume. All overlap fractions were significantly different from each other, except overlap fractions of $GTV_{50\%}$ versus GTV_{SBR} . The mean GTV fraction delineated by CT but not by PET, $1 - OF_{CT}$ varied from 0.39 to 0.61. The mean GTV fraction delineated by PET but not by CT, $1 - OF_{PET}$ varied from 0.15 to 0.34. The latter is further detailed in table 3, categorized for tumor subsite. The segmentation tools performed very differently with no clear difference by tumor site or stage and all methods resulted in a large percentage of patients having more than 20% of the GTV_{PET} outside the GTV_{CT} -domain. $GTV_{50\%}$ and GTV_{SBR} resulted in less PET-volume outside the CT-volume than GTV_{VIS} and $GTV_{40\%}$. An example of tumor tissue delineated by FDG-PET but not on CT is shown in figure 3.

The absolute volumes, OF_{CT} , and OF_{PET} of the methods $GTV_{50\%}$ and GTV_{SBR} were similar. The overlap fraction of $GTV_{50\%}$ relative to GTV_{CT} , versus the overlap fraction of GTV_{SBR}

Fig. 2 Mean absolute volumes of the various GTV methods. Error bars indicate standard deviation of the mean

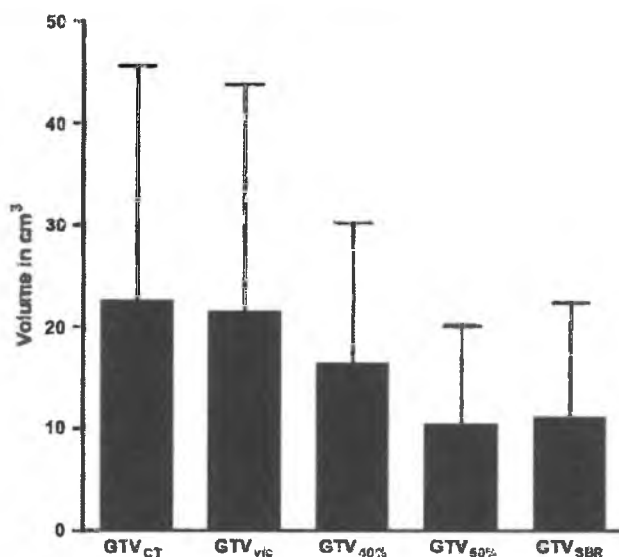


Table 3 GTV_{PET} located outside the GTV_{CT}-volume

Cases with $\geq 10\%$ of GTV_{PET} located outside the GTV_{CT}-volume

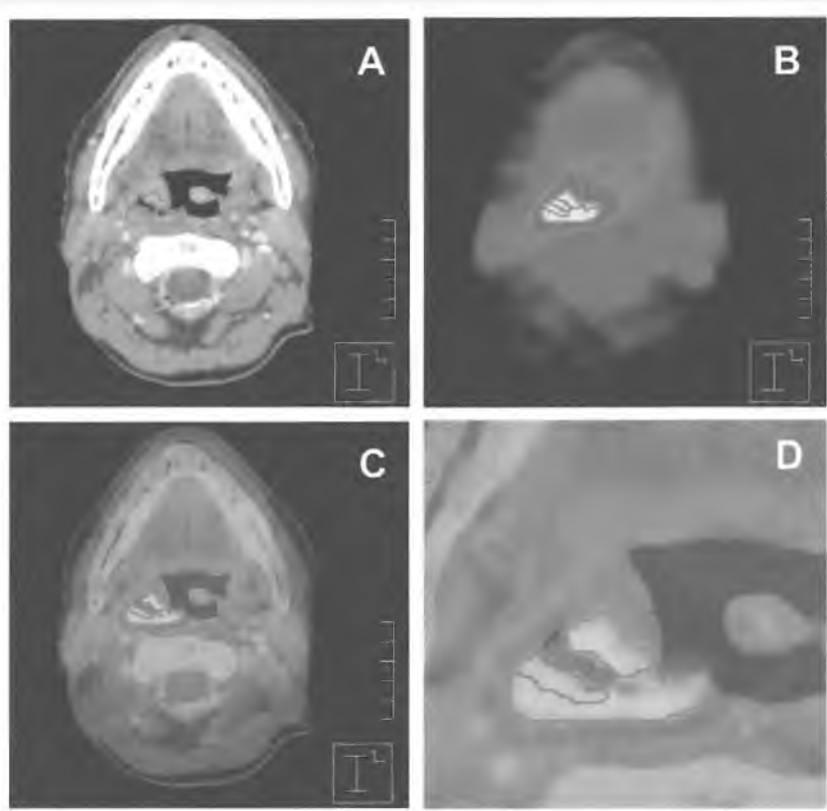
	GTV _{VIS}	GTV _{40%}	GTV _{50%}	GTV _{SBR}
Oral cavity/oropharynx	30 (83%)	23 (68%)	17 (50%)	15 (43%)
Larynx/hypopharynx	35 (85%)	29 (74%)	20 (51%)	22 (55%)
Total	65 (84%)	52 (71%)	37 (51%)	37 (49%)

Cases with $\geq 20\%$ of GTV_{PET} located outside the GTV_{CT}-volume

	GTV _{VIS}	GTV _{40%}	GTV _{50%}	GTV _{SBR}
Oral cavity/oropharynx	19 (53%)	19 (56%)	8 (24%)	9 (26%)
Larynx/hypopharynx	30 (73%)	24 (62%)	13 (33%)	14 (35%)
Total	49 (64%)	43 (59%)	21 (29%)	23 (31%)

relative to GTV_{CT} showed a strong correlation between the two methods (Pearson correlation $r=0.85$, $p<0.0001$). Nevertheless, the overlap of GTV_{50%} and GTV_{SBR} relative to each other showed a similarity less than 90% in 26 cases. This indicates that, although the

Fig. 3 CT scan (A), corresponding FDG-PET scan (B), fused image (C) and enhanced detail of fused image (D) of a patient with a T3N2cM0 oropharyngeal carcinoma, showing tumor tissue delineated by FDG-PET, but not by CT. Indicated are GTV_{CT} (red) and GTV_{SBR} (dark green).



GTV_{50%} and GTV_{SBR} segmentation tools yield similar average GTV volumes and overlap fractions, on an individual patient basis, there is a geographical mismatch in a substantial number of cases.

Discussion

In this study we compared five segmentation tools for FDG-PET based target volume definition in a large cohort of head and neck cancer patients. There were three important observations. First, the GTV_{SUV} method using a fixed threshold of 2.5 failed to provide

successful delineation in a large number of cases. Second, the volume and shape of the GTV on PET largely depends on the segmentation tool used. Third, PET frequently detected extension of tumor tissue outside the GTV_{CT} , regardless of the applied segmentation method.

Segmentation using a SUV-threshold of 2.5 resulted in an unsatisfactory large GTV in nearly half of the patients. In the remaining patients it resulted in a volume that was larger than the corresponding CT-volume, whilst the other automated segmentation tools produced volumes smaller than GTV_{CT} . The SUV of 2.5 was chosen as an arbitrary cut-off value as it represents the level that some reports use to consider a lesion as malignant, e.g. when staging non-small cell lung cancer [14,15], although the use of any SUV to differentiate benign from malignant is highly questionable [26]. Furthermore, extrapolation to tumor sites other than lung cancer is even more debatable. Lung tumors are frequently much larger than head and neck tumors, and lung tumors are often surrounded by large areas of very low FDG activity (i.e. normal lung tissue) whereas in head and neck cancer there is always a significant amount of background activity, because of physiological FDG-uptake in surrounding muscular tissue. For segmentation of primary lung cancer, Nestle *et al.* have found the value of 2.5 for SUV satisfactory [10], but they also reported that when attempting to delineate a lesion surrounded by tissue with significant background activity that the GTV_{SUV} tool was not suitable. Based on our results, we conclude that this method is not useful for automated target definition of head and neck cancer.

The other PET segmentation tools were successful in most, or all (GTV_{VIS}), cases. We found no explanation why a $GTV_{40\%}$ and $GTV_{50\%}$ could not be generated in four cases. Miller *et al.* reported that thresholding at 40% only works when the SBR is larger than 10 [20]. These four patients all had a SBR below 10 (range 3.6 to 6.4). However, 31 other patients who did have a satisfactory $GTV_{40\%}$ and $GTV_{50\%}$ also had a SBR below 10. We can not conclude that it was a low SBR that caused the unsatisfactory GTVs.

The mean absolute GTV volume derived from visual interpretation of PET-images was comparable to GTV_{CT} . All threshold-based PET segmentation tools, besides the rejected SUV-based segmentation, resulted in volumes smaller than GTV_{CT} . This was also described by Daisne *et al.* [27], who compared the role of co-registered CT, MRI and FDG-PET in delineating the primary tumor in laryngeal cancer patients prior to laryngectomy, with contouring based on the GTV_{SBR} tool. Compared to the reference surgical specimen all modalities overestimated the tumor extension, but GTV_{SBR} came closest at depicting the true tumor volume. The differences were significant, with an average GTV on histological examination of 12.6 cm^3 , whereas averages for PET, CT and MRI were 16.3 cm^3 , 20.8 cm^3 and 23.8 cm^3 , respectively. We also observed this effect for CT and GTV_{SBR} in our patients with cancer of the larynx, oral cavity, oropharynx and hypopharynx.

Most other studies comparing GTV definition using FDG-PET against CT or MRI in head and neck cancer have used visual interpretation of PET images [5,9,28]. All studies demonstrated significant differences between PET and CT volumes in a large proportion of the patients. Compared to the GTV_{CT} , the average GTV_{PET} was smaller in one study [5], of similar size in one study [9] and in another study larger in 40% of the cases, but smaller in the remaining 60% [28]. These variable results may, at least partly, reflect the subjective nature and operator dependency of the visual interpretation method. Paulino *et al.* used a fixed threshold of 50% of the maximum signal intensity [21]. The resulting GTV_{PET} was almost a factor two smaller relative to the GTV_{CT} which is in good agreement with our results.

Despite the fact that on average threshold-based GTV_{PET} volumes were smaller than GTV_{CT} , in many cases a significant part of the GTV_{PET} was located outside the volume defined as GTV_{CT} . This suggests that GTV definition using PET may include tumor extension that are not unequivocally depicted on CT. Note that our GTV_{CT} already included tumor extensions found at clinical examination and examination under anesthesia. As FDG may also accumulate in inflammatory tissue, this is a potential caveat resulting in larger PET-based GTV not due to cancer. The relevance of peritumoral inflammation needs further evaluation, as this could not be discriminated in the current study.

FDG-PET may also underestimate tumor volume in specific situations. Tumor parts that are small, or that are located in a background with a relatively intense signal, could be missed. Both Daisne *et al.* and Ng *et al.* observed that, like CT and MRI, FDG-PET failed to identify superficial mucosal tumor extension [27,29]. Nevertheless, using PET, Ng *et al.* missed only one small tumor location in a series of 124.

Given the different characteristics of GTV delineation with CT and PET, PET-based target volume definition can be used as complementary information to the conventional methods, but at present should not replace the CT-based volume until the smaller PET-based GTVs have been proven to be oncologically safe. This proof should ideally be acquired through histological validation studies, for then one can reliably decide how to use the additional information. GTV may be reduced using PET when dubious densities as seen on CT prove to be FDG-negative. This will reduce overtreatment and possibly reduce side-effects. GTV may be expanded based on PET, when additional FDG-positive locations are not explained by clinically evident benign inflammation. This will reduce the risk of geographical misses. We envisage that the future role of PET in target volume definition will be as a complementary tool, adding to other imaging and clinical information.

The PET delineation methods resulted in GTV volumes with significant differences. Only the $GTV_{50\%}$ and GTV_{SBR} methods produced comparable volumes and overlap fractions. These segmentation tools, however, were not equivalent as geographical similarity

between $GTV_{50\%}$ and GTV_{SBR} proved to be less than 90% in 26 of 73 cases. This might be explained by the thresholds that were generated by the GTV_{SBR} algorithm, which ranged from 45.8% to 63.7%. This illustrates that for an individual patient $GTV_{50\%}$ and GTV_{SBR} were not interchangeable. The differences between the other PET-based methods were larger both with regard to volume and overlap. This underlines the need for validated GTV definition with PET.

All segmentation tools have inherent limitations. The main weaknesses of GTV_{VIS} are that the resulting GTV is strongly influenced by the window-level setting of the data set and that it is a pure subjective approach, leading to substantial intra- and inter-observer variability [30]. Using a fixed threshold (i.e. $GTV_{40\%}$ or $GTV_{50\%}$) as advocated by many research groups [15-21] is debatable. This seems to perform reasonably well in phantom-based experiments using symmetrical volumes with homogeneous activity and a sharp demarcation from the background activity. However, tumors may display a heterogeneous distribution of radioactivity. They may also be more complex in shape, which may limit the performance of the segmentation tool in the clinical setting [31]. Furthermore, these methods imply that there is no dependency on the background activity. When trying to apply a fixed threshold on Daisne's laryngectomy data, thresholds ranging from 36% to 73% were necessary to fit the true tumor volume on histology [30, 32]. The GTV_{SBR} tool needs to be calibrated to the specific institutional image acquisition and reconstruction settings and it is not ideal for low SBR images [33]. However, it is the only tool that has been the subject of a histological validation study, and its threshold is adapted to the signal to background ratio of an individual patient. A novel iterative method for lesion delineation and volumetric quantification has recently been presented, whereby the background is subtracted from the signal, which makes it independent of the signal-to-background ratio, and seems to be a more robust segmentation tool [34]. It will be of great value to investigate its performance in head and neck cancer patients. A method that is not dependent on observer variations and SBR ratio, and that has been validated properly, is preferable. These criteria apply best to the GTV_{SBR} segmentation tool.

Conclusions

This study shows that FDG-PET may have important consequences for GTV definition, but that the choice of a segmentation tool for target volume definition of head and neck cancer based on PET images is not trivial. The absolute PET volume is dependent on the segmentation method used. Delineation using an SUV value of 2.5 is insufficient, and the other evaluated methods show inconsistencies. The SBR method seems preferable, since it uses a threshold adapted to the signal to background ratio of an individual patient and it does not depend on observer variability.

In general, PET volumes were smaller than CT volumes, but PET also identified possible tumor areas that were not contoured by the conventional CT-based method. This could potentially improve the accuracy of GTV definition. Additional histological validation studies are necessary before routine usage of FDG-PET data to optimize CT-derived target volumes.

References

1. Rischin D, Hicks RJ, Fisher R, *et al.* Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a substudy of Trans-Tasman Radiation Oncology Group Study 98.02. *J Clin Oncol* 2006;24:2098-2104.
2. Ling CC, Humm J, Larson S, *et al.* Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys* 2000;47:551-560.
3. van Baardwijk A, Baumert BG, Bosmans G, *et al.* The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning. *Cancer Treat Rev* 2006;32:245-260.
4. Bradley JD, Perez CA, Dehdashti F, *et al.* Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. *J Nucl Med* 2004;45 Suppl 1:96S-101S.
5. Heron DE, Andrade RS, Flickinger J, *et al.* Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.
6. Hebert ME, Lowe VJ, Hoffman JM, *et al.* Positron emission tomography in the pretreatment evaluation and follow-up of non-small cell lung cancer patients treated with radiotherapy: preliminary findings. *Am J Clin Oncol* 1996;19:416-421.
7. Kiffer JD, Berlangieri SU, Scott AM, *et al.* The contribution of 18F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer. *Lung Cancer* 1998;19:167-177.
8. Nestle U, Walter K, Schmidt S, *et al.* 18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. *Int J Radiat Oncol Biol Phys* 1999;44:593-597.
9. Nishioka T, Shiga T, Shirato H, *et al.* Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;15:1051-1057.
10. Nestle U, Kremp S, Schaeffer-Schuler A, *et al.* Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1348.
11. Messa C, Ceresoli GL, Rizzo G, *et al.* Feasibility of [18F]FDG-PET and coregistered CT on clinical target volume definition of advanced non-small cell lung cancer. *Q J Nucl Med Mol Imaging* 2005;49:259-266.
12. Riegel AC, Berson AM, Destian S, *et al.* Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys* 2006;65:726-732.
13. Hong R, Halama J, Bova D, *et al.* 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* 2007;67:720-726.
14. Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys* 2004;59:4-5.
15. Bradley J, Thorstad WL, Mutic S, *et al.* Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;59:78-86.
16. Brianconi E, Rossi G, Ancidei S, *et al.* Radiotherapy planning: PET/CT scanner performances in the definition of gross tumour volume and clinical target volume. *Eur J Nucl Med Mol Imaging* 2005;32:1392-1399.
17. Ciernik IF, Dizendorf E, Baumert BG, *et al.* Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
18. Giraud P, Grahek D, Montravers F, *et al.* CT and (18)F-deoxyglucose (FDG) image fusion for optimization of conformal radiotherapy of lung cancers. *Int J Radiat Oncol Biol Phys* 2001;49:1249-1257.
19. Mah K, Caldwell CB, Ung YC, *et al.* The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int J Radiat Oncol Biol Phys* 2002;52:339-350.
20. Miller TR, Grigsby PW. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. *Int J Radiat Oncol Biol Phys* 2002;53:353-359.
21. Paulino AC, Koshy M, Howell R, *et al.* Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
22. Daisne JF, Sibomana M, Bol A, *et al.* Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol* 2003;69:247-250.

23. Vogel WV, Wensing BM, van Dalen JA, *et al.* Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2005;32:1276-1282.
24. Vogel WV, Schinagel DA, van Dalen JA, *et al.* Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. *Q J Nucl Med Mol Imaging* 2008;52:74-83.
25. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med* 2004;45:1431-1434.
26. Westterterp M, Pruim J, Oyen W, *et al.* Quantification of FDG PET studies using standardized uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. *Eur J Nucl Med Mol Imaging* 2007;34:392-404.
27. Daisne JF, Duprez T, Weynand B, *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.
28. Scarfone C, Lavelly WC, Cmelak AJ, *et al.* Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. *J Nucl Med* 2004;45:543-552.
29. Ng SH, Yen TC, Liao CT, *et al.* 18F-FDG PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of 124 patients with histologic correlation. *J Nucl Med* 2005;46:1136-1143.
30. Grégoire V, Bol A, Geets X, *et al.* Is PET-based treatment planning the new standard in modern radiotherapy? The head and neck paradigm. *Semin Radiat Oncol* 2006;16:232-238.
31. Hoffmann AL, van Dalen J, Lee J, *et al.* Assessment of ^{18}F PET signals for automatic target volume definition in radiotherapy treatment planning: regarding Davis *et al.* *Radiother Oncol* 2007;83:102-103.
32. Grégoire V, Daisne JF, Geets X. Comparison of CT- and FDG-PET-defined GT: in regard to Paulino *et al.* *Int J Radiat Oncol Biol Phys* 2005;63:308-309.
33. Grégoire V, Haustermans K, Geets X, *et al.* PET-based treatment planning in radiotherapy: a new standard? *J Nucl Med* 2007;48 Suppl 1:68S-77S.
34. van Dalen JA, Hoffmann AL, Dicken V, *et al.* A novel iterative method for lesion delineation and volumetric quantification with FDG-PET. *Nucl Med Comm* 2007;28:485-493

5

Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head and neck cancer?

Dominic A.X. Schinagl
Aswin L. Hoffmann
Wouter V. Vogel
Jorn A. Van Dalen
Suzan M.M. Verstappen
Wim J.G. Oyen
Johannes H.A.M. Kaanders

Abstract

Background and purpose: The role of FDG-PET in radiotherapy target volume definition of the neck was evaluated by comparing eight methods of FDG-PET segmentation to the current CT-based practice of lymph node assessment in head-and-neck cancer patients.

Materials and methods: Seventy-eight head-and-neck cancer patients underwent co-registered CT- and FDG-PET scans. Lymph nodes were classified as "enlarged" if the shortest axial diameter on CT was ≥ 10 mm, and as "marginally enlarged" if 7-10mm. Subsequently, lymph nodes were assessed on FDG-PET applying eight segmentation methods: visual interpretation (PET_{VIS}), applying fixed thresholds at a standardized uptake value (SUV) of 2.5 and at 40% and 50% of the maximum signal intensity of the primary tumor (PET_{SUV} , $PET_{40\%}$, $PET_{50\%}$) and applying a variable threshold based on the signal-to-background ratio (PET_{SBR}). Finally, $PET_{40\%N}$, $PET_{50\%N}$ and PET_{SBRN} were acquired using the signal of the lymph node as the threshold reference.

Results: Of 108 nodes classified as "enlarged" on CT, 75% were also identified by PET_{VIS} , 59% by $PET_{40\%}$, 43% by $PET_{50\%}$ and 43% by PET_{SBR} . Of 100 nodes classified as "marginally enlarged", only a minority was visualized by FDG-PET. The respective numbers were 26%, 10%, 7% and 8% for PET_{VIS} , $PET_{40\%}$, $PET_{50\%}$ and PET_{SBR} . $PET_{40\%N}$, $PET_{50\%N}$ and PET_{SBRN} respectively identified 66%, 82% and 96% of the PET_{VIS} -positive nodes.

Conclusions: Many lymph nodes that are enlarged and considered metastatic by standard CT-based criteria appear to be negative on FDG-PET scan. Alternately, a small proportion of marginally enlarged nodes are positive on FDG-PET scan. However, the results are largely dependent on the PET segmentation tool used and until proper validation FDG-PET is not recommended for target volume definition of metastatic lymph nodes in routine practice.

Introduction

Lymph node involvement in squamous cell carcinoma (SCC) of the head-and-neck is an indicator of poor prognosis, reducing the cure rate by almost 50% [1]. Standard diagnostic **workup for assessing cervical lymph node status is performed by computed tomography (CT) or magnetic resonance imaging (MRI).** The sensitivity (50%–80%) and specificity (70%–90%) of CT and MRI are comparable [2,3]. For marginally enlarged lymph nodes, examination by ultrasound imaging with fine-needle aspiration cytology (FNAC) is superior to CT and MRI if performed by an experienced radiologist (sensitivity and specificity up to 76% and 100%, respectively) [4,5].

Radiation oncologists often encounter the dilemma of marginally enlarged lymph nodes when delineating the radiation target volume of a head-and-neck cancer patient. A number of CT-criteria are used to assess the presence or absence of metastatic tumor within lymph nodes. These include signs of central necrosis, conglomeration of the nodes and minimal axial nodal diameter. The presence of central nodal necrosis is the most reliable criterion determining the presence of metastatic disease [6]. The size criterion of shortest axial diameter is a compromise between sensitivity and specificity. Many radiologists accept the CT-criteria proposed by van den Brekel *et al.*: minimal axial diameter of 11mm for jugulodigastric nodes and 10mm for all other cervical nodes. These criteria yielded a sensitivity of 89% and a specificity of 73% [7]. However, a subsequent ultrasound study suggested that a minimal axial diameter of 7mm for jugulodigastric and 6mm for other cervical nodes was the optimal compromise for sensitivity/specificity [8]. The radiation oncologist has to decide whether or not to include these marginally enlarged lymph nodes in the high dose target volume.

Metabolic information, as provided by imaging ^{18}F -fluorodeoxyglucose (FDG) with positron emission tomography (PET), has not provided a clear advantage in the nodal staging of head-and-neck cancer when compared to CT and MRI [9,10]. However, this does not disqualify FDG-PET as a potential useful tool for delineation of the lymph node target volume in radiotherapy planning.

FDG-PET has been incorporated into target volume delineation by many groups [11]. Various methods for FDG-PET based target volume definition are in use. Visual interpretation is the most commonly applied method [12-15]. This method, however, is susceptible to subjective window-level settings of the images and is thus highly operator dependent. Therefore, more objective methods have been explored. Examples are isocontouring based on either a standardized uptake value (SUV) [13,16-18], a fixed threshold of the maximum signal intensity [19-23], or on a threshold adaptive to the signal-to-background ratio (SBR) [24].

In a recent study, we demonstrated that FDG-PET may have important consequences for GTV definition of the primary tumor in head-and-neck cancer, but that the choice of the PET segmentation tool is not trivial [25].

In this study, eight methods of FDG-PET segmentation are compared to the current CT-based practice of lymph node assessment in order to evaluate the role of FDG-PET in radiotherapy target volume definition of the neck, in head-and-neck cancer patients eligible for primary radiotherapy.

Materials and methods

Patients

Seventy-eight patients (59 males and 19 females, median age 61 years, range 43-86 years) with stage II-IV SCC of the head-and-neck area, eligible for primary curative radiotherapy, were prospectively enrolled from June 2003 until July 2006. The tumor characteristics are summarized in Table 1. Staging was performed according to standard clinical protocol. FDG-PET was performed only for research purposes and did not influence staging. The study was approved by the Ethics Committee of the Radboud University Nijmegen Medical Centre and all patients provided informed consent.

Table 1 Tumor characteristics.
Staging was performed according to standard clinical protocol.

Tumor site	Oral cavity	6
	Oropharynx	31
	Hypopharynx	9
	Larynx	32
T stage	T1	1
	T2	16
	T3	39
	T4	22
N stage	N0	21
	N1	10
	N2a	0
	N2b	17
	N2c	29
	N3	1
Total		78

Image Acquisition

Prior to treatment, a CT scan and an FDG-PET scan were acquired in radiation treatment position, the patient being immobilized by a custom-made rigid mask covering head, neck and shoulders. Maximum reproducibility in positioning was assured by the use of additional support systems: a flat scanning bed, a customized head support cushion, an intra-oral mould when indicated, a standard cushion supporting the knees, and a laser positioning system. The median interval between CT and FDG-PET was 3 days (range 0-10 days).

CT scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters were 130 kV, 120 mAs, slice thickness 3mm, scanning from above the frontal sinuses to below the clavicles, with intravenous contrast.

FDG-PET scans were acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded. However, glucose levels had to be appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection). A 3D emission scan of the head-and-neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 minutes (median 64 minutes, S.D. \pm 11.4) after intravenous injection of 250 MBq FDG (Covidien, Petten, The Netherlands). The acquisition time per bed position was 5 minutes for emission- and 3 minutes for the Germanium-based transmission scans, resulting in a total acquisition time of 16 minutes for the two bed positions. Emission and transmission scans were reconstructed using an iterative 2D ordered subset expectation maximization (OSEM) algorithm, with parameters (4 iterations, 16 subsets) optimized for low photon attenuation in the head-and-neck area as described elsewhere [26].

The image registration procedure has been described in detail previously [27]. In brief, three-dimensional surface models of both imaging modalities were derived and an operator-independent iterative closest point algorithm coregistered these models with an average registration error of 2.0mm. In addition, a second PET data set was generated in which the original voxel values were replaced with calculated SUV values. SUV was defined as the voxel value of detected activity (in Bq/ml) multiplied by the weight of the patient (in kg) divided by the activity at the beginning of the scan (in Bq) [28].

The CT and the two PET data sets were transferred via DICOM to the Pinnacle³ treatment planning system (Philips Medical Systems, Andover, MA, USA) in order to apply the PET-segmentation tools for metastatic lymph node identification.

Lymph Node Identification

The lymph nodes were identified and delineated on CT images by two experienced radiation oncologists in consensus (Node_{CT}).

On CT images, manual delineation of the Node_{CT} was performed according to our current clinical protocol, using all available information gathered from physical examination, CT and/or MRI, ultrasound with FNAC and the contrast enhanced CT in treatment position. CT criteria for pathologic nodes were: a lymph node ≥ 7 mm in the shortest axis or the aspect of central necrosis or a conglomeration of at least three nodes. Lymph nodes measuring 7 to 10 mm in the shortest axis were classified as "marginally enlarged". Nodes ≥ 10 mm in the shortest axial diameter were classified as "enlarged". The FDG-PET images were blinded during Node_{CT}-definition. CT images were not blinded during assessment of nodal disease by PET. All patients underwent ultrasonic evaluation of cervical lymph node regions as part of the routine diagnostic work-up. Fine needle aspiration was performed by the radiologist when suspicious nodes, mostly with a shortest axial diameter ≥ 7 mm, were detected. Nodes with clear clinical and radiological signs of metastatic disease were not routinely aspirated. Of 78 patients, 55 underwent FNAC of a total of 85 lymph nodes. It must be stressed that ultrasound with FNAC was a routine diagnostic procedure with the purpose of staging only. FNAC was not performed with the intention to provide a substrate for validation of the PET findings at the level of individual nodes. Eight PET segmentation tools were used to assess FDG uptake in the nodes. Visual delineation (PET_{VIS}) was performed by identifying FDG activity clearly above normal background activity and classified "positive" in consensus with an experienced nuclear medicine physician. For the other (threshold based) segmentation tools in-house developed scripts for the Pinnacle³ treatment planning system were used. SUV-based identification of pathologic nodes was obtained by applying an isocontour of 2.5 (PET_{SUV}). Two thresholds were based on fixed percentages of the maximum signal intensity in the primary tumor of 40% (PET_{40%}) and 50% (PET_{50%}), respectively. These thresholds were based on data derived from phantom experiments and have previously been used by other investigators [19-23]. Finally, an adaptive threshold tool (PET_{SBR}) based on the signal-to-background ratio (SBR) of the primary tumor was used, as developed at Université St. Luc in Brussels, Belgium [24]. The calibration and implementation of the PET_{SBR} method was described in detail previously [25]. Results obtained by automated delineation algorithms were checked visually before acceptance. A delineation was considered unsuccessful if the resulting lymph node volumes included significant amounts of tissue that was clearly normal at visual interpretation. For all the nodes identified by the PET_{VIS} method additional segmentation procedures were performed now using the lymph node itself as reference and not the primary tumor. The scripts for the PET_{40%}, PET_{50%} and PET_{SBR} segmentation methods were again applied whereby now the maximum signal intensity of the specific lymph node was used instead of the primary tumor, yielding the parameters PET_{40%N}, PET_{50%N} and PET_{SBRN} respectively.

Statistics

The number of lymph nodes identified as pathologic by each method were compared using Wilcoxon Signed Rank and McNemar tests.

Results

Seventy-eight patients were included in this study. PET_{VIS} could be generated for all 78 patients. For the segmentation methods depending on the FDG activity of the primary tumor (PET_{40%}, PET_{50%}, PET_{SBR}) 77 patients were available for analysis as the primary tumor of one patient with a T2N2cM0 oropharyngeal carcinoma was not visualized by FDG-PET. The PET_{SBR} segmentation tool resulted in an unsuccessful segmentation in four patients. This was observed in nine patients for both PET_{40%} and PET_{50%}, four of whom also had an unsuccessful PET_{SBR}. SUV data was unavailable for one patient. In 50 of the remaining 77 patients the PET_{SUV} segmentation was unsuccessful, including the nine patients mentioned above. As a consequence, the PET_{SUV} method was not further evaluated. Unsuccessful segmentation was not correlated with specific tumor subsites or stages.

Table 2 shows the number of lymph nodes identified by CT and the various PET segmentation methods. For the 78 patients 208 nodes were classified as "enlarged" or "marginally enlarged"

Table 2 Number of lymph nodes identified by CT, PET_{VIS}, PET_{40%}, PET_{50%} and PET_{SBR}.

	Total number of nodes	Number of nodes identified per patient				p-value*
		Mean (SD)	Median [I _Q _{0.25} -0.75]	Minimum	Maximum	
CT ≥ 7 mm (N = 78)	208	2.67 (3.51)	2.0 [0–3.0]	0	22	
CT ≥ 10 mm (N = 78)	108	1.38 (1.74)	1.0 [0–2.0]	0	9	
PET_{VIS} (N = 78)	112	1.44 (1.62)	1.0 [0–2.0]	0	6	<0.001
PET_{40%} (N = 69)	68	0.99 (1.35)	1.0 [0–1.5]	0	6	<0.001
PET_{50%} (N = 69)	46	0.67 (1.05)	0 [0–1.0]	0	6	<0.001
PET_{SBR} (N = 74)	50	0.68 (1.08)	0 [0–1.0]	0	6	<0.001

*Wilcoxon Signed Rank test p-value relative to CT ≥ 7mm

according to the CT-criteria. The number of nodes identified by each of the PET methods was significantly smaller ($p \geq 0.001$). Of the nodes classified as “enlarged” (i.e. ≥ 10 mm) on CT, 81 (75%) were visualized by PET_{VIS}, 55 (59%) by PET_{40%}, 40 (43%) by PET_{50%} and 42 (43%) by PET_{SBR} (Table 3). Of the nodes classified as “marginally enlarged” (i.e. 7–10 mm) on CT only a minority was visualized by PET. The respective numbers were 26%, 10%, 7% and 8% for PET_{VIS}, PET_{40%}, PET_{50%} and PET_{SBR}. Both for the identification of “enlarged” and “marginally enlarged” nodes there were significant ($p \leq 0.004$) differences between the PET segmentation methods (Table 3). The performance of the various PET methods in relation to the shortest axial nodal diameter as measured by CT is shown in Figure 1.

Table 3 Number of lymph nodes identified by the various PET segmentation methods.

	CT ≥ 7 mm (N = 208)	p-value	“marginally enlarged” CT 7–10 mm (N = 100)	p-value	“enlarged” CT ≥ 10 mm (N = 108)	p-value	Number of patients
PET _{VIS}	107/208 (51%)		26/100 (26%)		81/108 (75%)		78
PET _{40%}	64/183 (35%)	<0.001*	9/90 (10%)	<0.001*	55/93 (59%)	0.004*	69
PET _{50%}	46/183 (25%)	<0.001*	6/90 (7%)	<0.001*	40/93 (43%)	<0.001*	69
PET _{SBR}	50/193 (26%)	<0.001*	8/96 (8%)	<0.001*	42/97 (43%)	<0.001*	74
PET _{40%N}	73/107 (68%)	<0.001†	14/26 (54%)	0.001†	59/81 (73%)	<0.001†	78
PET _{50%N}	89/107 (83%)	<0.001†	18/26 (69%)	0.016†	71/81 (88%)	0.016†	78
PET _{SBRN}	103/107 (96%)		25/26 (96%)		78/81 (96%)		78

* Difference relative to PET_{VIS} (McNemar test)

† Difference relative to PET_{SBRN} (McNemar test)

Different results were obtained when segmentation was performed using the maximum intensity of the individual nodes as reference instead of the primary tumor. Of the 112 nodes identified by PET_{VIS}, 74 (66%) were visualized by PET_{40%N}, 92 (82%) by PET_{50%N} and 108 (96%) by PET_{SBRN}. Note that there is an increasing order in the number of nodes visualized by PET_{40%N}, PET_{50%N} and PET_{SBRN} whereas this was the reverse when the primary tumor is used as the reference for segmentation of the nodes. The performance of PET_{40%N}, PET_{50%N} and PET_{SBRN} with respect to “enlarged” and “marginally enlarged” nodes is shown in table 3. An example of discordance between CT- and PET findings is shown in Figure 2, including a large node with a necrotic centre that is PET negative. Discrepancies between the various PET segmentation methods are shown in Table 4. The largest discrepancies were found between PET_{VIS} and the other methods and to a lesser extent between PET_{40%} and the other methods. There was very little discordance between PET_{50%} and PET_{SBR}.

Fig. 1 Percentage of nodes positive for the various FDG-PET segmentation methods in relation to the shortest axial nodal diameter.

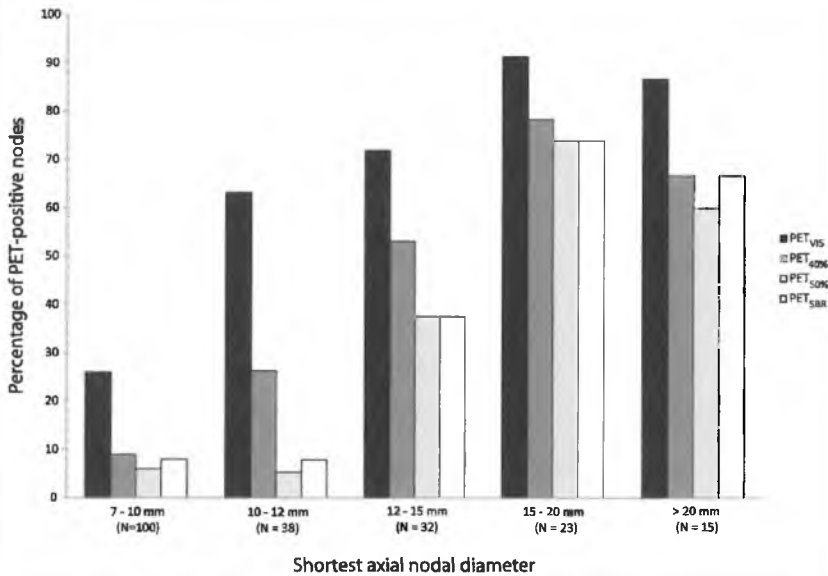


Table 4 Discrepancies between the various PET segmentation methods.

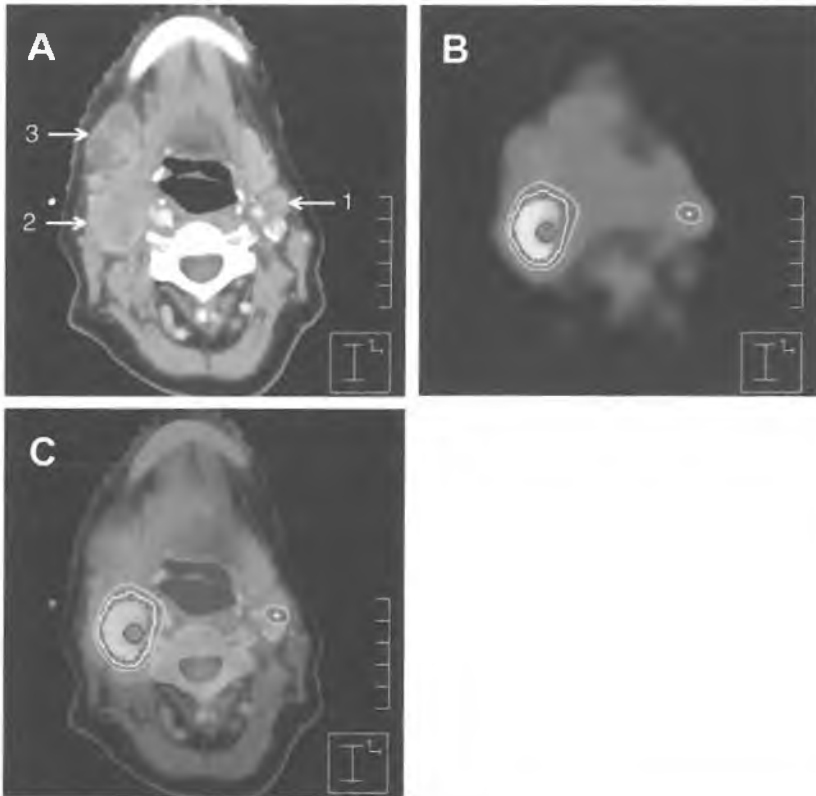
	Number of PET negative nodes out of 208 nodes ≥ 7 mm on CT-scan	Positive by other segmentation method			
		PET _{VIS}	PET _{40%}	PET _{50%}	PET _{SBR}
PET _{VIS}	101	—	4/101	3/101	3/101
PET _{40%}	119	32/119	—	0/119	0/119
PET _{50%}	137	49/137	18/137	—	1/137
PET _{SBR}	143	55/143	19/143	2/143	—

In five cases FDG-PET was positive whilst the nodes were classified as “benign” by our standard diagnostic workup procedure. One palpable node, 9mm in diameter on CT, FNAC-negative on two separate occasions, was positive for PET_{VIS} and PET_{40%}. Four nodes not detected by physical examination, one of 7mm in the shortest axis, FNAC-negative, but PET_{VIS}-positive and three of 6mm which were positive for PET_{VIS} and PET_{40%}.

Discussion

Generally accepted CT-criteria for assessment of lymph nodes are nodal size, aspect of central necrosis and conglomeration of (at least 3) nodes [29]. Radiation oncologists use these criteria to decide which nodes should be included in the radiotherapy target volume for head-and-neck cancer. Usually the cut-off level for size is set at 10mm shortest axial diameter [7,9,29,30]. However, approximately 25% of nodes ≥ 10 mm will not contain metastatic disease, and 25% of nodes < 10 mm will harbor metastases [7]. In this study we aimed to explore whether FDG-PET can assist in the decision making of radiotherapy target volume definition of cervical nodes in head-and-neck cancer. The sensitivity, specificity and accuracy of FDG-PET in identifying neck node metastases has been the subject of various histologically controlled studies, and vary for head-and-neck cancer from 33% to 90%, 76% to 98%, and 63% to 96%, respectively [9,29-32]. These variations could be the result of differences in selection of patients, histopathological evaluation or assessment of PET data [32]. False positive FDG-PET findings can be produced by inflammatory response in reactive lymph nodes. All these studies used visual interpretation of PET images which is highly susceptible to inter-operator variation. For that reason it would be preferable to use a semi-automatic threshold-based segmentation tool that is less operator dependent. It should be noted, however, that given the differences in performance of these segmentation tools compared to visual interpretation as demonstrated in the current study, this would yield different sensitivity and specificity rates. An important observation by Dammann *et al.* was that FDG-PET allowed correction of nearly all false positive CT and MRI findings [29]. They compared nodal staging by CT, MRI and FDG-PET in 64 patients with SCC of the head-and-neck and validated the results by histopathology. The current study suggests that the number of lymph nodes considered as metastatic according to the CT-criteria might be reduced with 25% when FDG-PET with visual interpretation is used. It could be reduced to less than half when semi-automatic segmentation tools are used. Dammann *et al.* also found that false negative findings with CT and MRI were less frequent and harder to correct with PET [29]. It has also been reported by other investigators that the sensitivity of metastatic lymph node detection is not significantly increased by FDG-PET [32,33]. This could be caused by the spatial resolution of PET, whereby small intranodal tumor deposits could go undetected, or by a low signal-to-background ratio through uptake in the surrounding muscular tissue. Nevertheless, the current results show that of the lymph nodes measuring 7-10mm on CT-scan, still 26% were found positive on PET, albeit by visual interpretation. The rate was significantly lower ($< 10\%$) with $PET_{40\%}$, $PET_{50\%}$ and PET_{SBR} (Figure 1). In this context it is again important to note that the results are largely dependent on the PET segmentation tool used.

Fig. 2 (A) CT scan, (B) corresponding FDG-PET scan, and (C) fused image of a patient with a T4 oropharyngeal carcinoma show three lymph nodes. Left (arrow1) a node that is 13mm in the shortest axial diameter that was identified by PET_{75} (light green) and $PET_{20\%}$ (yellow). Right (arrow2) a node of 23mm that was identified by PET_{75} , $PET_{20\%}$, $PET_{50\%}$ (blue) and PET_{SBR} (dark green). The pattern of FDG-uptake in this particular node suggests intranodal tumor heterogeneity. Right (arrow3) an overtly metastatic node of 21mm with central necrosis that was not identified by FDG-PET.



Segmentation using a SUV-threshold of 2.5 resulted in unsuccessful identification in 65% of the patients. The arbitrary cut-off value of 2.5 represents the level that in some reports has been used to consider a lesion as malignant, e.g. when staging non-small cell lung cancer [17,18], although the use of any SUV to differentiate benign from malignant is questionable [34]. Nestle *et al.* found the value of 2.5 satisfactory when segmenting lung cancer [13], but also reported that this method was not useful when a lesion was surrounded by significant background activity. Previously we concluded that this method is not useful for automated primary tumor delineation in head-and-neck cancer [25].

Based on the results of the current study, we conclude that this method is also not useful for automated metastatic lymph node identification. Recently however, Murakami *et al.* showed that size-based SUV cut-off values could be utilized in order to identify malignant lymph node in patients with head-and-neck cancer [35]. In a histologically controlled study of 23 patients they found that a SUV of 1.9 for nodes <10mm in diameter, 2.5 for nodes 10–15mm, and 3.0 for nodes >15mm yielded a sensitivity of 79% and a specificity of 99%.

The other PET segmentation tools were successful in most, or all (PET_{VIS}), cases. We found no explanation why a $PET_{40\%}$ and $PET_{50\%}$ could not be generated in nine of the 78 cases. Miller *et al.* reported that thresholding at 40% only works when the SBR is larger than 10 [22]. These nine patients all had an SBR below 10 (range 3.6 to 6.4). However, 26 other patients in whom successful $PET_{40\%}$ and $PET_{50\%}$ were generated also had an SBR below 10, which indicates that this is not necessarily a major obstacle.

When the maximum signal activity of lymph nodes was used as the reference for threshold settings instead of the activity of the primary tumor the results differ considerably. The PET_{SBRN} method was able to segment almost all PET_{VIS} -positive nodes. These numbers decreased (82% and 66%) as the threshold decreased ($PET_{50\%N}$ and $PET_{40\%N}$) making some lymph nodes undetectable for the tools as the threshold falls below the activity level of the surrounding benign tissue. A reverse effect was seen when the primary tumor was used as the reference. Given the discrepancies between the segmentation tools demonstrated in the current study and the dependency on the threshold reference, we believe that these segmentation tools should only be used in daily clinical practice after proper validation, preferably against histopathological assessment of tumor extensions. We feel therefore that, as yet, there are no strong arguments to recommend FDG-PET as a routine for radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer.

Combined PET/CT scanners are rapidly replacing standalone PET scanners and also find their way to radiotherapy departments. The suggested superiority of image fusion with these integrated scanners over software fusion of data acquired by standalone equipment is a matter of debate [36]. The average registration error of the standalone FDG-PET and CT images was 2.0mm in our study [27] and as there are no data indicating that PET/CT can outperform high quality software fusion of PET and CT [36] we do not believe that the results of the current study would differ when a PET/CT were used.

A potential future role for FDG-PET in radiation treatment planning may be in the dose level chosen for the nodes. FDG-PET could possibly identify a category of “intermediate risk” lymph nodes, i.e. nodes that likely contain only a small tumor burden. This may

include the “marginally enlarged” but PET-positive nodes of 7-10mm. An intermediate dose level may well be sufficient in such cases. This approach could limit toxicity, without jeopardizing the rates of cure. Possibly also nodes ≥ 10 mm but PET-negative may classify as “intermediate risk” although here it will be more difficult to set criteria as with increasing size the likelihood of metastatic disease rapidly increases. Also, PET can be false negative in larger nodes containing substantial amount of necrosis which might explain our observation that with nodal sizes >20 mm there is a decrease in the rate of PET positive nodes (Figure 1). However, at this point the intermediate risk concept remains theoretical until histological studies have established the value of metabolic information provided by FDG-PET in this setting. Only then, could one consider testing the intermediate dose level hypothesis in clinical studies.

Conclusions

A substantial number of lymph nodes with a shortest axial diameter ≥ 10 mm on CT-scan and considered metastatic by generally accepted radiological criteria appear to be negative on FDG-PET. Alternately, a proportion of nodes between 7 and 10mm are positive on FDG-PET. The results, however, are largely dependent on the PET segmentation tool used and until proper validation FDG-PET is not recommended for target volume definition of metastatic lymph nodes in routine practice. Which method is to be preferred, visual interpretation or one of the semi-automatic segmentation procedures, is the subject of an ongoing validation study.

References

- [1] Brockstein B, Haraf DJ, Rademaker AW, et al. Patterns of local failure, prognostic factors and survival in locoregionally advanced head-and-neck cancer treated with concomitant chemoradiotherapy: a 9-year, 337-patient, multi-institutional experience. *Ann Oncol* 2004;15:1179-1186.
- [2] Van den Brekel MW, Castelijns JA, Stel HV, et al. Modern imaging techniques and ultrasound-guided aspiration cytology for the assessment of neck node metastases: a prospective comparative study. *Eur Arch Otorhinolaryngol* 1993;250:11-17.
- [3] Wide JM, White DW, Woolgar JA, et al. Magnetic resonance imaging in the assessment of cervical nodal metastasis in oral squamous cell carcinoma. *Clin Radiol* 1999;54:90-94.
- [4] Takes RP, Righi P, Meeuwis CA, et al. The value of ultrasound with ultrasound-guided fine-needle aspiration biopsy compared to computed tomography in the detection of regional metastases in the clinically negative neck. *Int J Radiat Oncol Biol Phys* 1998;40:1027-1032.
- [5] Van den Brekel MW, Castelijns JA, Stel HV, et al. Occult metastatic neck disease: detection with US and US-guided fine-needle aspiration cytology. *Radiology* 1991;180:457-461.
- [6] Van den Brekel MW. Lymph node metastases: CT and MRI. *Eur J Radiol* 2000;33:230-238.
- [7] Van den Brekel MW, Stel HV, Castelijns JA, et al. Cervical lymph node metastasis: assessment of radiologic criteria. *Radiology* 1990;177:379-384.
- [8] Van den Brekel MW, Castelijns JA, Snow GB. The size of lymph nodes in the neck on sonograms as a radiologic criterion for metastasis: how reliable is it? *Am J Neuroradiol* 1998;19:695-700.
- [9] Adams S, Baum RP, Stuckensen T, Bitter K, Hor G. prospective comparison of 18F-FDG-PET with conventional imaging modalities (CT, MRI, US) in lymph node staging of head-and-neck cancer. *Eur J Nucl Med* 1998;25:1255-1260.
- [10] Stokkel MP, ten Broek FW, Hordijk GJ, Koole R, van Rijk PP. Preoperative evaluation of patients with primary head-and-neck cancer using dual-head 18fluorodeoxyglucose positron emission tomography. *Ann Surg* 2000;231:229-234.
- [11] Van Baardwijk AA, Baumert BG, Bosmans G, et al. The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning. *Cancer Treat Rev* 2006;32:245-260.
- [12] Heron DE, Andrade RS, Flickinger J, et al. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.
- [13] Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1248.
- [14] Nishioka T, Shiga T, Shirato H, et al. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057.
- [15] Riegel AC, Berson AM, Destian S, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys* 2006;65:726-732.
- [16] Hong R, Halama J, Bova D, Sethi A, Emami B. 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* 2007;67:720-726.
- [17] Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys* 2004;59:4-5.
- [18] Bradley J, Thorstad WL, Mutic S, et al. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;59:78-86.
- [19] Brianzoni E, Rossi G, Ancidei S, et al. Radiotherapy planning: PET/CT scanner performances in the definition of gross tumour volume and clinical target volume. *Eur J Nucl Med Mol Imaging* 2005;32:1392-1399.
- [20] Ciernik IF, Dizendorf E, Baumert BG, et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
- [21] Mah K, Caldwell CB, Ung YC, et al. The Impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int J Radiat Oncol Biol Phys* 2002;52:339-350.
- [22] Miller TR, Grigsby PW. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. *Int J Radiat Oncol Biol Phys* 2002;53:353-359.

- [23] Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
- [24] Daisne JF, Sibomana M, Bol A, et al. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol* 2003;69:247-250.
- [25] Schinagel DA, Vogel WV, Hoffmann AL, et al. Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:1282-1289.
- [26] Vogel WV, Wensing BM, van Dalen JA, et al. Optimised PET reconstruction of the head-and-neck area: improved diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2005;32:1276-1282.
- [27] Vogel WV, Schinagel DA, van Dalen JA, Kaanders JH, Oyen WJ. Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head-and-neck area. *Q J Nucl Med Mol Imaging* 2008;52:74-83.
- [28] Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med* 2004;45:1431-1434.
- [29] Dammann F, Horger M, Mueller-Berg M, et al. Rational diagnosis of squamous cell carcinoma of the head-and-neck region: comparative evaluation of CT, MRI, and 18FDG-PET. *Am J Roentgenol* 2005;184:1326-1331.
- [30] Ng SH, Yen TC, Chang JT, et al. Prospective study of [18F]fluoro-deoxyglucose positron emission tomography and magnetic resonance imaging in oral cavity squamous cell carcinoma with palpably negative neck. *J Clin Oncol* 2006;24:4371-4376.
- [31] Stuckensen T, Kovacs AF, Adams S, Baum RP. Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. *J Craniomaxillofac Surg* 2000;28:319-324.
- [32] Wensing BM, Vogel WV, Marres HA, et al. FDG-PET in the clinically negative neck in oral squamous cell carcinoma. *Laryngoscope* 2006;116:809-813.
- [33] Stoeckli SJ, Steinert H, Pfaltz M, Schmidt S. Is there a role for positron emission tomography with 18F-fluorodeoxyglucose in the initial staging of nodal negative oral and oropharyngeal squamous cell carcinoma. *Head Neck* 2002;24:345-349.
- [34] Westerterp M, Pruim J, Oyen W, et al. Quantification of FDG-PET studies using standardized uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. *Eur J Nucl Med Mol Imaging* 2007;34:392-404.
- [35] Murakami R, Uozumi H, Hirai T, et al. Impact of FDG-PET/CT imaging on nodal staging for head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2007;68:377-382.
- [36] Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, Corstens FH. PET/CT: panacea, redundancy, or something in between? *J Nucl Med* 2004;45:155-245.

6 |

Can FDG-PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study

Dominic A.X. Schinagl

Paul N. Span

Wim J.G. Oyen

Johannes H.A.M. Kaanders

Abstract

Background and purpose: In head-and-neck cancer (HNC) various treatment strategies have been developed to improve outcome, but selecting patients for these intensified treatments remains difficult. Therefore, identification of novel pre-treatment assays to predict outcome is of interest. In HNC there are indications that pre-treatment tumour F-18-fluoro-deoxy-glucose (FDG) uptake may be an independent prognostic factor. The aim of this study was to assess the prognostic value of FDG uptake and computed tomography (CT)- and FDG-PET-based primary tumour volume measurements in HNC patients treated with (chemo)radiotherapy.

Materials and methods: Seventy-seven stage II-IV HNC patients eligible for definitive (chemo)radiotherapy underwent co-registered pre-treatment CT and FDG-PET. Primary tumour volume was determined on CT (GTV_{CT}) and FDG-PET. Five PET-segmentation methods were applied: interpreting FDG-PET visually (PET_{VIS}), applying an isocontour at a standardized uptake value (SUV) of 2.5 ($PET_{2.5}$), using fixed thresholds of 40% and 50% ($PET_{40\%}$, $PET_{50\%}$) of the maximum intratumoural FDG-activity (SUV_{MAX}) and applying an adaptive threshold based on the signal-to-background (PET_{SBR}). Mean FDG-uptake for each PET-based volume was recorded (SUV_{mean}). Subsequently, to determine the metabolic volume, the integrated SUV (iSUV) was calculated, being the product of PET-based volume and SUV_{mean} . All these variables were analyzed as potential predictors for local control (LC), regional recurrence free survival (RRFS), distant metastasis free survival (DMFS), disease-free survival (DFS) and overall survival (OS).

Results: In oral cavity/oropharynx tumours PET_{VIS} was the only volume-based method able to predict LC. Both PET_{VIS} and GTV_{CT} were able to predict DMFS, DFS and OS in these subsites. iSUV's were associated with LC, DMFS, DFS and OS, while SUV_{mean} and SUV_{MAX} were not. In hypopharyngeal/laryngeal tumours none of the variables were associated with outcome.

Conclusions: Thus far, there is no role for pre-treatment FDG-PET as a predictor of (chemo) radiotherapy outcome in HNC in daily routine yet. However, this potential application needs further exploration, focusing both on FDG-PET based primary tumour volume, iSUV and SUV_{MAX} of the primary tumour.

Introduction

In head-and-neck cancer various treatment strategies have been developed to improve outcome. However, it remains difficult to select patients for these intensified treatments despite careful evaluation of the clinical factors such as tumour size/stage, lymph node involvement and anatomic subsite. Therefore, identification of novel pre-treatment factors that potentially predict treatment response and long-term outcome is of great interest [1].

The development of molecular imaging techniques, such as positron emission tomography (PET), allows the non-invasive study of pathophysiology of cancers.

In head-and-neck cancer there are indications that pre-treatment tumour F-18-fluoro-deoxy-glucose (FDG) uptake may be an independent prognostic factor [1]. Many research groups have studied the incorporation of FDG-PET into radiation treatment planning, and several ways of using PET data have been described. Visual interpretation is the most commonly used method [2-5]. This method, however, is susceptible to variations due to the window-level settings of the images and is highly operator dependent. Therefore, more objective methods have been explored. Examples are isocontouring based on either a standardized uptake value (SUV) of 2.5 around the tumour [3;6-8], a fixed threshold of the maximum signal intensity [9-13], or a threshold which is adaptive to the signal to background ratio [3,14]. We recently demonstrated that FDG-PET may have important consequences for GTV definition of the primary tumour in head-and-neck cancer, but that the choice of PET-segmentation tool is not trivial [15]. The aim of this study was to assess the prognostic value of CT- and FDG-PET-based primary tumour volume, and various ways of quantifying FDG uptake in head-and-neck cancer patients treated with (chemo) radiotherapy and to provide an overview of the available literature.

Materials and methods

Patients

Seventy-seven patients (58 males and 19 females, median age 61 years, range 43-86 years) with stage II-IV squamous cell carcinoma of the head-and-neck area, eligible for primary curative radiotherapy, were prospectively enrolled from June 2003 until July 2006. FDG-PET was performed only for research purposes, and did not influence treatment. The tumour characteristics are summarized in table 1. No information on human papillomavirus relatedness can be provided. The study was approved by the Ethics Committee of the Radboud University Nijmegen Medical Centre and all patients provided informed consent.

Table 1 Tumour characteristics of 77 patients.

Tumour site	Oral cavity	6
	Oropharynx	30
	Hypopharynx	9
	Larynx	32
T stage	T1	1
	T2	15
	T3	39
	T4	22
N stage	N0	21
	N1	10
	N2a	0
	N2b	17
	N2c	28
	N3	1
Histological grade	1	4
	2	37
	3	33
	Unknown	3
Total		77

Treatment

All patients were discussed in a multidisciplinary conference for tumour classification and treatment recommendations. Our protocol recommended treating primary tumour and metastatic lymph nodes to a dose of 68-70 Gray (Gy). This was combined with concomitant weekly intravenous cisplatin 40 mg/m² for large unresectable tumours. Elective lymph node regions were treated to 44 Gy.

Image Acquisition

Before treatment, a CT scan and FDG-PET scan were acquired in radiation treatment position with the patient immobilized by using a custom-made rigid mask covering the head, neck, and shoulders. Maximum reproducibility in positioning was insured by the use of additional support systems: a flat scanning bed, customized head-support cushion, intraoral mould when indicated, standard cushion supporting the knees, and laser positioning system as previously described [15]. Prior to treatment, a CT-scan and an FDG-PET-scan were acquired in radiation treatment position. CT-scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters were 130 kV, 120 mAs, slice

distance and slice thickness 3 mm, scanning the head-and-neck area, with intravenous contrast. FDG-PET-scans were acquired using a full-ring dedicated PET-scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded. However, glucose levels had to be appropriately regulated (glucose level at time of FDG injection < 10 mmol/L, no insulin administration before FDG injection). A 3D emission scan of the head-and-neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 min (median 64 min, S.D. ± 11.4) after intravenous injection of 250 MBq FDG (Covidien, Petten, The Netherlands). The acquisition time per bed position was 5 min for emission and 3 min for the Germanium-based transmission scan, resulting in a total scanning time of 16 min for the two bed positions. Image reconstruction has been described in detail previously [16].

Three-dimensional surface models were automatically derived from both the CT and PET images. These models were anatomically co registered by using an operator-independent iterative closest point algorithm, with an average registration error of 2.0 mm at the centre of the planning area as described previously [17]. SUV was defined as the voxel value of detected activity multiplied by the weight of the patient divided by the activity at the beginning of the scan.

The CT and the two PET data sets were transferred via DICOM to a Pinnacle3 treatment planning system (Philips Medical Systems, Andover, MA, USA) for target volume definition.

Target Volume Definition

The primary tumour was delineated on CT and FDG-PET images by two experienced radiation oncologists in consensus. The volume of the metastatic lymph nodes was not included. The role of FDG-PET in the delineation of metastatic lymph nodes has been analyzed previously [18].

On CT images, manual delineation of the GTV_{CT} was performed according to current clinical protocols, using information gathered from physical examination, available diagnostic workup imaging modalities (CT and/or MRI, examination under general anaesthesia) and the CT in treatment position. When the radiation oncologists were drawing the GTV_{CT} contours, the FDG-PET images were blinded.

Five PET-based volumes were obtained using different delineation approaches. Visual delineation (PET_{vis}) was performed by contouring FDG activity clearly above normal background activity. Localizations with increased FDG uptake were classified malignant in consensus with an experienced nuclear medicine physician. The other (threshold based) volumes were obtained using in-house developed software scripts for the Pinnacle³ treatment planning system. SUV-based delineation was obtained by applying an isocontour

of $SUV=2.5$ ($PET_{2.5}$) around the tumour. Two thresholds were based on fixed percentages 40% ($PET_{40\%}$) and 50% ($PET_{50\%}$) of the maximum signal intensity in the primary tumour (SUV_{MAX}). Finally, an adaptive threshold delineation (PET_{SBR}) based on the signal-to-background ratio (SBR) was performed, as developed at Université St. Luc in Brussels, Belgium [14]. The calibration and implementation of the PET_{SBR} method were described in detail previously [15]. Results obtained by automated delineation algorithms were checked visually before acceptance. A delineation was considered unsuccessful if the resulting volume included significant volumes of tissue that were clearly normal at visual interpretation.

The mean FDG uptake of each PET-based volume was recorded ($SUV_{meanVIS}$, $SUV_{mean2.5}$, $SUV_{mean40\%}$, $SUV_{mean50\%}$, $SUV_{meanSBR}$). This was multiplied by the corresponding volume resulting in the integrated SUV ($iSUV_{VIS}$, $iSUV_{2.5}$, $iSUV_{40\%}$, $iSUV_{50\%}$, $iSUV_{SBR}$).

Treatment outcome analysis

Follow-up visits included history, inspection of the upper aerodigestive tract and palpation of the neck. Local and regional recurrences were proven by histology and cytology, respectively. Distant metastases were identified by either pathologic or radiologic proof.

Statistics

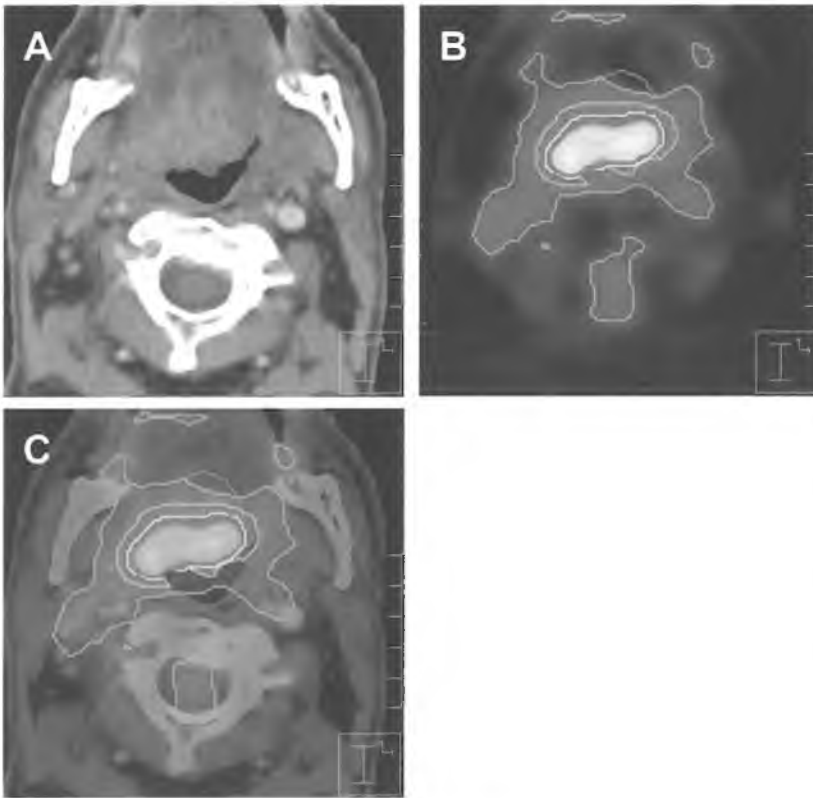
All statistical analyses were performed using SPSS version 16.0 (SPSS Inc. Chicago IL). Differences between two categories were established using t-tests or Mann-Whitney U testing, when appropriate. Normality of distributions were assessed using Kolmogorov-Smirnov tests. Variables were entered as continuous variables in Cox regression analyses to preclude the need of establishing a cut-off value for local control (LC), regional recurrence free survival (RRFS), distant metastasis free survival (DMFS), disease-free survival (DFS) and overall survival (OS). A $p < 0.05$ was *a priori* considered as statistically significant.

Results

Tumour volume measurements

For CT based primary tumour volume measurements 77 data sets were available. PET_{VIS} was generated for all 77 patients, the PET_{SBR} segmentation tool resulted in unsuccessful volume definition in two patients. A delineation was considered unsuccessful if the resulting GTV included significant volumes of tissue that were clearly normal at visual interpretation. This was observed in four patients for both $PET_{40\%}$ and $PET_{50\%}$, two of whom also had an unsatisfactory PET_{SBR} . The $PET_{2.5}$ segmentation tool was unsuccessful in 35 patients, including the four patients mentioned. As a consequence this latter method was not further evaluated. All unsuccessful volume definitions were largely oversized, being at least 300 cm^3 and clearly incorporated benign tissue. An unsuccessful delineation

Fig. 1 Planning CT (A), corresponding FDG-PET (B) and fusion image (C) of patient with T3N2bM0 oropharyngeal carcinoma show differences in target-volume definition. Indicated are gross tumour volume (GTV) delineated on CT (GTV_{CT}, red, absolute volume of 34.0cm³) and PET-based GTVs obtained by visual interpretation (PET_{VIS}, light green, volume 33.8cm³), applying an isocontour of a standardized uptake value (SUV) of 2.5 (PET_{2.5}, orange), using a fixed threshold of 40% (PET_{40%}, yellow, volume 14.0cm³) and 50% (PET_{50%}, blue, volume 13.4cm³) of the maximum signal intensity, applying an adaptive threshold based on the signal-to-background ratio (SBR, PET_{SBR}, dark green, volume 15.0cm³). GTV_{2.5} was unsuccessful in this case because of inclusion of large areas of normal background tissue. Note that on this transversal slice PET_{50%} and PET_{SBR} are indistinguishable



did not correlate with specific tumour subsites or T stages. An example of an inadequate PET_{2.5} is shown in figure 1. The mean absolute tumour volume for the various methods were 22.7 cm³, 21.5 cm³, 16.4 cm³, 10.5 cm³ and 11.2 cm³ for GTV_{CT}, PET_{VIS}, PET_{40%}, PET_{50%}, and PET_{SBR} respectively. GTV_{CT} and PET_{VIS} yielded similar mean absolute volumes, but the threshold based methods (PET_{40%}, PET_{50%}, and PET_{SBR}) were all smaller than GTV_{CT}

($p \leq 0.0001$ for all comparisons). Overlap and mismatch analyses performed in order to evaluate the location of the acquired volumes, showed that more than 20% of the FDG-PET based tumour volume was located outside the GTV_{CT}-domain in 64%, 59%, 29% and 31% of the cases for PET_{VIS}, PET_{40%}, PET_{50%} and PET_{SBR}, respectively.

Treatment and treatment outcome

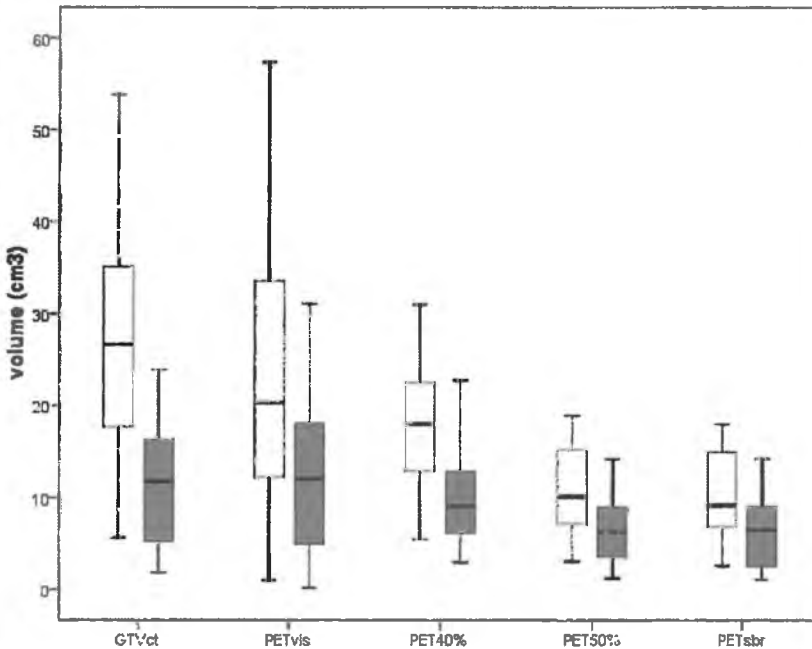
The median primary tumour radiation dose was 68 Gy (range 64-72 Gy). Three patients were not treated; one patient died just prior to radiotherapy, the other patient refused primary radiotherapy and the latter developed distant metastases prior to radiotherapy. After a median follow up of 46 months (range 2.5-76) LC, RRFs, DMFS, DFS and OS at 2 years were 84%, 95%, 86%, 73% and 77%, respectively. Follow up was at least 24 months or until patients' death. After primary treatment, five patients did not obtain a complete remission. These patients did not have significantly different CT- or PET- based tumour volumes than the patients who did obtain a complete remission. No recurrences were seen in the areas treated to an elective dose.

Prognostic value of CT and PET

Primary tumour volume (PET- or CT-based), SUV_{mean}, SUV_{MAX} and iSUV were not able to predict the chance of obtaining a complete remission. The CT- and PET-based tumour volumes of the patients who have achieved a complete remission (n=69) are shown in figure 2. There was a significant difference in the volumes of oral cavity and oropharyngeal tumours as compared to laryngeal and hypopharyngeal tumours ($p \leq 0.004$, Mann-Whitney). SUV_{MAX} for oral cavity/oropharyngeal tumours and laryngeal/hypopharyngeal tumours was 9.7 and 10.0, respectively. We analyzed LC, RRFs, DMFS, DFS and OS in patients who achieved a complete remission (n=69) after primary treatment using primary tumour volume (PET- or CT-based), SUV_{mean}, SUV_{MAX} and iSUV as continuous variables in Cox regression survival analyses.

In hypopharyngeal and laryngeal tumours, none of the CT- and PET parameters were associated with any of the outcome-related endpoints. SUV_{MAX} and SUV_{mean} also had no prognostic value in oral cavity and oropharyngeal tumours. The other results for oral cavity and oropharyngeal tumours are presented in table 2. In these head-and-neck subsites PET_{VIS} was able to predict LC, whereas the other volume-based methods were not. Both PET_{VIS} and GTV_{CT} were able to predict DMFS, DFS and OS. Furthermore, all iSUV methods were able to predict LC, DMFS, DFS, and OS, albeit sometimes with borderline significance (p-values between 0.051 and 0.055). Figure 3 shows individual data points of GTV_{CT} and PET_{VIS} in relation to LC and DFS of oral cavity/oropharyngeal tumours with a follow-up of at least 24 months. Albeit that mean values differ significantly, figure 3 also shows that there is a large overlap in the volume range between patients with and without recurrence or death, indicating that the discriminative power of GTV_{CT} and PET_{VIS} is limited.

Fig. 2 Box-and-Whisker plot depicting 5% and 95% confidence intervals (Whiskers), 25% and 75% confidence intervals (Box), and median of CT- and PET-based tumour volumes of oral cavity/oropharyngeal tumours (white) and hypopharyngeal/laryngeal tumours (black). There was a significant difference in the volumes of oral cavity and oropharyngeal tumours as compared to laryngeal and hypopharyngeal tumours ($p \leq 0.004$, Mann-Whitney)



Discussion

In this study we assessed the prognostic value of CT- and FDG-PET-based primary tumour volume measurements, mean FDG (SUV_{mean}) and maximum FDG uptake (SUV_{MAX}), and integrated SUV (iSUV) in a large cohort of head-and-neck cancer patients treated with (chemo)radiotherapy.

Interestingly, PET_{vis} was able to predict LC in oral cavity and oropharyngeal tumours, but GTV_{CT} was not, while the mean volumes of PET_{vis} and GTV_{CT} were similar. Other studies confirm the lack of prognostic potential of CT-based primary tumour volume in oral cavity and oropharyngeal tumours [33;34]. Our observation that PET_{vis} is associated with LC is novel. It remains questionable however if visual assessment can be a reliable prognostic

tool given the operator dependent nature of this method. Both GTV_{CT} and PET_{VIS} could predict DMFS, DFS and OS in these subsites. For CT-based primary tumour volume this was also observed by Chao *et al.* in 31 patients with oropharyngeal cancer treated with definitive (chemo)radiotherapy [35]. Apparently, in oropharynx tumours local radiotherapy response does not so much depend on the primary tumour volume, but possibly more on biological characteristics of the tumour [36]. On the other hand, these results do suggest that metastatic potential is associated with the primary tumour volume in this head-and-neck subsite. One other study of 59 stage III-IV head-and-neck cancer patients treated with definitive (chemo)radiotherapy found a correlation between PET-based primary tumour volume, using the PET_{2.5} method, and PFS [28]. After further analyses they also found that a volume $\geq 9.3 \text{ cm}^3$ was associated with a decreased OS.

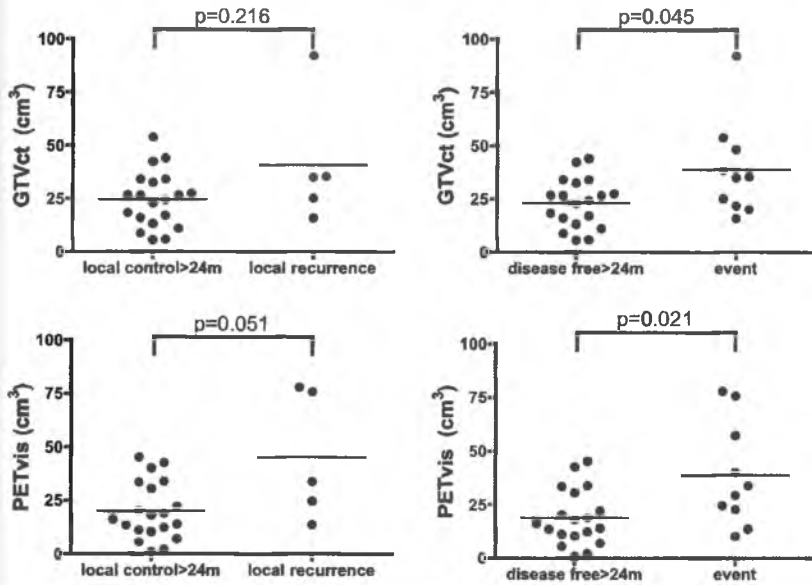
All the iSUV methods (the product of the PET-based primary tumour volume and the SUV_{mean} within that volume, reflecting the metabolic volumes) were able to predict LC, DMFS, DFS, OS, in oral cavity and oropharynx tumours, albeit sometimes with borderline significance. iSUV is a new variable fully representing the total metabolic activity within a predefined tumour volume. La *et al.* also correlated iSUV to treatment outcome, albeit based on cumulative volumes of both the primary tumour and the PET-avid lymph nodes [27]. However, they hypothesized that the effect was due to the volume and not the product of volume and SUV_{mean}. In contrast, our data indicate that of all PET-based volume measurements only PET_{VIS} has a predictive value, while this is the case for practically all iSUV methods. This suggests that the product of volume and SUV_{mean} provides more robust parameter which could possibly be a surrogate for both tumour aggressiveness and the total cancer cell mass.

Table 2 Primary tumour volume (PET- or CT-based) and PET-based integrated SUV as variables of treatment outcome prediction in patients with oral cavity and oropharynx tumours who achieved a complete remission (n=31) after definitive (chemo)radiotherapy.

Outcome	GTV _{CT}	PET _{VIS}	PET _{40%}	PET _{50%}	PET _{52%}	iSUV _{VIS}	iSUV _{40%}	iSUV _{50%}	iSUV _{52%}
LC	>0.1	0.031	>0.1	>0.1	>0.1	0.021	0.025	0.039	0.033
RRFS	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1
DMFS	0.003	0.046	0.080	0.064	>0.1	0.055	0.023	0.023	0.024
DFS	0.024	0.016	>0.1	>0.1	>0.1	0.033	0.041	0.054	0.051
OS	0.018	0.023	>0.1	>0.1	>0.1	0.026	0.038	0.052	0.040

LC = Local Control; RRFS = Regional Recurrence Free Survival; DMFS = Distant Metastasis Free Survival; DFS = Disease Free Survival; OS = Overall Survival. Variables assessed using Cox regression analyses. The numbers in indicate p-values.

Fig. 3 Panels showing GTV_{CT} and PET_{VIS} in relation to local control (a) and disease free survival (b) of oral cavity/oropharyngeal tumours with a follow-up of at least 24 months. Differences were analyzed using Mann-Whitney U tests.



In hypopharyngeal and laryngeal tumours we found no association between GTV_{CT} or PET_{VIS} and treatment outcome, whereas several studies have demonstrated the prognostic value of CT-determined tumour volume for outcome after definitive radiation therapy for these subsites as well as for nasopharyngeal cancer [37]. We do not have a solid explanation for this observation, except for the fact that we obtain high tumour control rates (LC at 2 years of 86%) compared to several other studies and consequently relatively few events which reduces the discriminative power of any pre-treatment test.

None of the three semi-quantitative methods of PET-based tumour volume calculation ($PET_{40\%}$, $PET_{50\%}$ and PET_{SBR}) showed associations with outcome in any of the head-and-neck subsites. It should be noted that all three semi-quantitative methods produced significantly smaller variability. This may also reduce discriminative power.

As the absolute volumes of FDG-PET based tumour sometimes partly located outside the GTV_{CT} -domain were small, it was not possible to trace the exact origin of a recurrence precisely to whether it was located outside the GTV_{CT} -domain, but within the FDG-PET based tumour volume.

Table 3 Summary of studies on treatment outcome prediction using SUV_{MAX} on pre-treatment FDG-PET of head-and-neck cancer patients treated with definitive (chemo)radiotherapy

Study	Patients	Tumoursite	Variable
Lee (21)	41	Nasopharynx (n=41)	SUV_{MAX} PT and/or MLN
Machtay (22)	60 A	Oral cavity/Oropharynx (n=44) Hypopharynx/Larynx (n=16)	SUV_{MAX} PT and/or MLN
Suzuki (30)	45	Nasopharynx (n=16) Oropharynx (n=20) Hypopharynx (n=3) Others (n=6)	SUV_{MAX} PT
Brun (19)	47	Nasopharynx (n=6) Oral cavity/Oropharynx (n=30) Hypopharynx/Larynx (n=10) Maxilla (n=1)	SUV_{MAX} PT
Schwartz (25)	54 B	Oral cavity/Oropharynx (n=34) Hypopharynx/Larynx (n=20)	SUV_{MAX} PT
Allal (1)	120 C	Oral cavity/Oropharynx (n=78) Hypopharynx/Larynx (n=39) Unknown (n=3)	SUV_{MAX} PT or MLN
Thorwarth(31)	12	Oral cavity/Oropharynx (n=6) Hypopharynx/Larynx (n=5) Unknown (n=1)	SUV_{MAX} PT or MLN
Roh (24)	79 D	Hypopharynx/Larynx (n=79)	SUV_{MAX} PT
Halfpenny (20)	58 E	Nasopharynx (n=1) Oral cavity/Oropharynx (n=55) Hypopharynx (n=1) Maxilla (n=1)	SUV_{MAX} PT
Minn (23)	37 F	Nasopharynx (n=5) Oral cavity/Oropharynx (n=16) Hypopharynx/Larynx (n=15) Parotid gland (n=2)	SUV_{MAX} PT
Chung (26)	82	Nasopharynx (n=63) Oropharynx (n=13) Hypopharynx (n=6)	SUV_{MAX} PT
Soto(29)	61	Nasopharynx (n=2) Oral cavity/Oropharynx (n=46) Hypopharynx/Larynx (n=9) Unknown (n=4)	SUV_{MAX} PT or MLN
Seol(28)	59	Oropharynx (n=13) Hypopharynx/Larynx (n=46)	SUV_{MAX} PT

Prediction	Treatment results
DFS worse when $SUV_{MAX} > 8.0$	3y DFS 74.3%
DFS and OS worse when $SUV_{MAX} \geq 9.0$	if $SUV_{MAX} \geq 9.0$: 2y DFS 37% if $SUV_{MAX} < 9.0$: 2y DFS 76%
DFS not correlated to SUV_{MAX}	if $SUV_{MAX} \geq 5.5$: 2y DFS 48% if $SUV_{MAX} < 5.5$: 2y DFS 76%
DFS and OS worse when $SUV_{MAX} > 9.0$	LC 78% ('during follow up time')
LC and DFS worse when $SUV_{MAX} \geq 9.0$	if $SUV_{MAX} \geq 9.0$: 2y LC 73% if $SUV_{MAX} < 9.0$: 2y LC 96% if $SUV_{MAX} \geq 9.0$: 2y DFS 69% if $SUV_{MAX} < 9.0$: 2y DFS 93%
LC and DFS worse when $SUV_{MAX} > 4.8$	4y LC 75% 4y DFS 59%
LC not correlated to SUV_{MAX}	LC 58% ('during follow up time')
LC and DFS worse when $SUV_{MAX} > 8.0$	3y LC 79% 3y DFS 50%
OS worse when $SUV_{MAX} > 10.0$	no LC or DFS data provided
OS worse when SUV_{MAX}	no LC or DFS data provided
DFS not correlated to SUV_{MAX}	DFS 78% ('after mean followup 35 months')
LRF not correlated to SUV_{MAX}	2y LRF 17%
PFS and OS not correlated to SUV_{MAX}	no LC or DFS data provided

Table 3 Continued.

Study	Patients	Tumoursite	Variable
La(27)	85	Nasopharynx (n=22) Oral cavity/Oropharynx (n=49) Hypopharynx/Larynx (n=12) Unknown (n=2)	SUV _{MAX} PT or MLN
Vernon(32)	42	Nasopharynx (n=3) Oral cavity/Oropharynx (n=27) Hypopharynx/Larynx (n=8) Unknown (n=4)	SUV _{MAX} PT or MLN
Current study	74	Oral cavity/Oropharynx (n=36) Hypopharynx/Larynx (n=38)	SUV _{MAX} PT

PT = Primary Tumour, MLN = Metastatic Lymph Node, LC = Local Control, LRF = Local-Regional Failure, DFS = Disease Free Survival, OS = Overall Survival, PFS = Progression Free Survival

A: 19 patients definitive (chemo)radiotherapy, 41 patients surgery and (chemo)radiotherapy

B: 27 patients definitive (chemo)radiotherapy, 17 patients surgery and radiotherapy, 8 patients surgery, 1 patient chemotherapy, 1 patient not reported.

In our cohort the SUV_{MAX} of the primary tumour was not able to predict radiation treatment outcome. Table 3 summarizes the results of a literature search for studies examining the role of pre-treatment FDG-PET SUV_{MAX} of head-and-neck cancer patients treated with definitive (chemo)radiotherapy in predicting outcome. Fifteen studies were identified of which eight showed that SUV_{MAX} could possibly play a role in predicting radiation treatment response [1;19-25] and seven showed that it does not [26-32]. These inconsistencies could be caused by the heterogeneity of treatment modalities, the heterogeneity of tumour sites, the use of several endpoints (i.e. LC, LRF, DFS, or OS), various SUV_{MAX} cut-off values, and using either the SUV_{MAX} of the primary tumour or the SUV_{MAX} of a metastatic lymph node. It is important to note that of the eight studies demonstrating association between SUV_{MAX} and outcome, six included substantial numbers who were treated with surgery. Together, of the 408 patients included in these six studies, 227 (55%) underwent primary surgery. In fact, the study by Brun *et al.* is the only one indicating that SUV_{MAX} is a prognostic factor in a population treated with definitive (chemo)radiotherapy alone, and using only the SUV_{MAX} of the primary tumour, finding that DFS and OS was worse when SUV_{MAX} > 9.0 [19]. Thus, based on this overview of the literature an unequivocal conclusion about the predictive role of pre-treatment FDG-PET SUV_{MAX} in head-and-neck cancer patients treated with definitive (chemo)radiotherapy cannot yet be made. Possibly larger cohorts of patients with homogeneous tumour and treatment characteristics stratified for the various subsites are able to establish a role for a SUV_{MAX} cut-off value in order to investigate future treatment individualization. Ideally these studies should use the same type of treatment and the same definition of treatment outcome.

Prediction	Treatment results
DFS and OS not correlated to SUV_{MAX}	2y DFS 70% 2y OS 78%
DFS and OS not correlated to SUV_{MAX}	2y DFS 71% 2y OS 83%
LC, DFS and OS not correlated to SUV_{MAX}	2y LC 84% 2y DFS 73%

C: 73 patients definitive (chemo)radiotherapy, 31 patients surgery and radiotherapy, 16 patients surgery

D: 37 patients definitive (chemo)radiotherapy, 34 patients surgery and radiotherapy, 6 patients surgery, 2 patients (chemo)radiotherapy and surgery.

E: 5 patients definitive radiotherapy, 40 patients surgery and radiotherapy, 13 patients surgery

F: 16 patients definitive radiotherapy, 2 patients surgery and radiotherapy, 19 patients radiotherapy and surgery.

Using pre-treatment primary tumour volume based on FDG-PET is appealing, and has not been extensively reported yet. In the current study, $PET_{V_{IS}}$ proved to be the only PET-based volume able to predict treatment outcome, and only in the oral cavity and oropharyngeal tumours. It should be noted that the discriminative potential of $PET_{V_{IS}}$ may be limited because of the large overlap between data points of patients with and without recurrence.

The volumes generated by semi-automated PET segmentation methods were not useful for outcome prediction.

Thorwarth *et al.* demonstrated that cumulative FDG-PET-based volumes of both the primary tumour and the PET-avid lymph nodes could not predict treatment outcome in a small series of head-and-neck cancer patients treated with definitive (chemo)radiotherapy [31]. Their PET-based volume was generated by encompassing all voxels showing a higher intensity than 40% of the maximum value. La *et al.* correlated DFS and OS of 85 head-and-neck cancer patients treated with definitive (chemo)radiotherapy to a FDG-PET-based cumulative volumes of both the primary tumour and the PET-avid lymph nodes [27]. Their PET-based volume was generated by encompassing all voxels showing a higher intensity than 50% of the maximum value. Recently, Chung *et al.* correlated the DFS of 82 pharyngeal cancer patients treated with definitive (chemo)radiotherapy to a FDG-PET-based cumulative volumes of both the primary tumour and the PET-avid lymph nodes [26]. Their PET-based volume, generated by encompassing all voxels showing an $SUV \geq 2.5$, a significant prognostic factor for DFS, whereas stage, histological grade, and SUV_{MAX} were not.

In our cohort the PET_{2.5} segmentation method resulted in an unsuccessful delineation in 35 patients, factors that might explain this finding have been addressed in a previous report [15].

The use of a molecular imaging modality such as FDG-PET in order to identify a robust variable on which treatment response and long-term outcome can be based remains attractive. Thus far, there is no role for pre-treatment FDG-PET as a predictor of outcome in head-and-neck cancer in daily routine, given the inconsistencies between studies and the low levels of evidence. However, this potential application of FDG-PET needs further exploration, focusing both on FDG-PET based primary tumour volume, iSUV and SUV_{MAX} of the primary tumour. Preferentially these questions should be incorporated in prospective phase III trials with strict criteria on treatment- and outcome parameters. Other research questions are worth considering such as adding the data of a repeat FDG-PET during treatment to the data acquired by a pre-treatment FDG-PET and the use of different PET tracers such as ¹⁸F-fluoromisonidazole and 3'-deoxy-3'-¹⁸F-fluorothymidine, to image hypoxia and tumour cell proliferation respectively, which are well known tumour characteristics relevant to radiation response [38].

Conclusions

Three major findings of this study are: First, in oral cavity and oropharyngeal tumours PET_{VIS} was the only volume-based method able to predict LC. Both PET_{VIS} and GTV_{CT} were associated with DMFS, DFS and OS in these subsites. Second, in oral cavity and oropharyngeal tumours the volume- and SUV-derived parameters iSUV_{VIS}, iSUV_{40%}, iSUV_{50%}, iSUV_{SBR} were consistently associated with LC, DMFS, DFS and OS, while SUV_{mean} and SUV_{MAX} were not. Third, in hypopharyngeal and laryngeal tumours none of the CT- and PET parameters were correlated with treatment outcome.

Given the inconsistencies between studies and low level of evidence thus far, there is no role for pre-treatment FDG-PET as a predictor of outcome in head-and-neck cancer in daily routine yet. Due to the heterogeneous nature of head and neck cancers, the difficulty of obtaining large number of patients, and the variation in results, one has to be careful interpreting the results from our and similar studies, as they are based on a relatively low number of events. However, this potential application of FDG-PET needs further exploration, focusing both on FDG-PET based primary tumour volume, iSUV and SUV_{MAX} of the primary tumour. Preferentially these questions should be incorporated in prospective phase III trials with strict criteria on treatment- and outcome parameters.

References

- 1 Allal AS, Slosman DO, Kebdani T, Allaoua M, Lehmann W, Dulguerov P. Prediction of outcome in head-and-neck cancer patients using the standardized uptake value of 2-[[18F]fluoro-2-deoxy-D-glucose. *Int J Radiat Oncol Biol Phys* 2004;59:1295-1300.
- 2 Heron DE, Andrade RS, Flickinger J, et al. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.
- 3 Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1348.
- 4 Nishioka T, Shiga T, Shirato H, et al. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning for oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057.
- 5 Riegel AC, Berson AM, Destian S, et al. Variability of gross tumor volume delineation in non-small-lung cancer. *Int J Radiat Oncol Biol Phys* 2006;65:726-732.
- 6 Bradley J, Thorstad WL, Mutic S, et al. Impact of FDG-PET on radiation therapy volume delineation in non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;59:78-86.
- 7 Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target volume delineation in CT-based radiation treatment planning. *Int J Radiat Oncol Biol Phys* 2007;67:720-726.
- 8 Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys* 2004;59:4-5.
- 9 Brianconi E, Rossi G, Ancidei S, et al. Radiotherapy planning: PET/CT scanner performances in the definition of gross tumor volume and clinical target volume. *Eur J Nucl Med Mol Imaging* 2005;32:1392-1399.
- 10 Ciernik IF, Dizendorf E, Baumert BG, et al. Radiation treatment planning with an integrated positron emission and computed tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
- 11 Mah K, Caldwell CB, Ung YC, et al. The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int J Radiat Oncol Biol Phys* 2002;52:339-350.
- 12 Miller TR, Grigsby PW. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation treatment. *Int J Radiat Oncol Biol Phys* 2002;53:353-359.
- 13 Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
- 14 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;69:247-250.
- 15 Schinagel DA, Vogel WV, Hoffmann AL, van Dalen JA, Oyen WJ, Kaanders JH. Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:1282-1289.
- 16 Vogel WV, Wensing BM, van Dalen JA, et al. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2005;32:1276-1282.
- 17 Vogel WV, Schinagel DA, van Dalen JA, Kaanders JH, Oyen WJ. Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. *Q J Nucl Med Mol Imaging* 2008;52:74-83.
- 18 Schinagel DA, Hoffmann AL, Vogel WV, et al. Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer? *Radiother Oncol* 2009;91:95-100.
- 19 Brun E, Kjellen E, Tenvall J, et al. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. *Head Neck* 2002;24:127-135.
- 20 Halfpenny W, Hain SF, Biassoni L, Maisey MN, Sherman JA, McGurk M. FDG-PET. A possible prognostic factor in head and neck cancer. *Br J Cancer* 2002;86:512-516.
- 21 Lee SW, Nam SY, Im KC, et al. Prediction of prognosis using standardized uptake value of 2-[[18F]fluoro-2-deoxy-d-glucose positron emission tomography for nasopharyngeal carcinomas. *Radiother Oncol* 2008;97:211-216.

- 22 Machtay M, Natwa M, Andrei J, et al. Prediction of prognosis using standardized uptake value as a prognostic factor for outcome in head and neck cancer. *Head Neck* 2009;31:195-201.
- 23 Minn H, Lapela M, Kleini PJ, et al. Prediction of survival with fluorine-18-fluoro-deoxyglucose and PET in head and neck cancer. *J Nucl Med* 1997;38:1907-1911.
- 24 Roh JL, Pae KH, Choi SH, et al. 2-[18F]-Fluoro-2-deoxy-D-glucose positron emission tomography as guidance for primary treatment in patients with advanced-stage resectable squamous cell carcinoma of the larynx and hypopharynx. *Eur J Surg Oncol* 2007;33:790-795.
- 25 Schwartz DL, Rajendran J, Yueh B, et al. FDG-PET prediction of head and neck squamous cell cancer outcomes. *Arch Otolaryngol Head Neck Surg* 2004;130:1361-1367.
- 26 Chung MK, Jeong HS, Park SG, et al. 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 Cancer Res* 2009;15:5861-5868.
- 27 La TH, Filion EJ, Turnbull BB, et al. Metabolic tumor volume predicts for recurrence and death in head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2009;74:1335-1341.
- 28 Seol YM, Kwon BR, Song MK, et al. Measurement of tumor volume by PET to evaluate prognosis in patients with head and neck cancer treated by chemo-radiation therapy. *Acta Oncol* 2010;49:201-208.
- 29 Soto DE, Kessler ML, Pierr M, Eisbruch A. Correlation between pretreatment FDG-PET biological target volume and anatomical location of failure after radiation therapy for head and neck cancers. *Radiother Oncol* 2008;89:13-18.
- 30 Suzuki K, Nishioka T, Homma A, et al. Value of fluorodeoxyglucose positron emission tomography before radiotherapy for head and neck cancer: does standardized value predict treatment outcome? *Jpn J Radiol* 2009;27:237-242.
- 31 Thorwarth D, Eschmann SM, Holzner F, Paulsen F, Alber M. Combined uptake of [18F]FDG and [18F]FMISO correlates with radiation treatment outcome in head and neck cancer patients. *Radiother Oncol* 2006;80:151-156.
- 32 Vernon MR, Maheshwari M, Schultz CJ, et al. Clinical outcomes of patients receiving integrated PET/CT-guided radiotherapy for head and neck carcinoma. *Int J Radiat Oncol Biol Phys* 2008;70:678-684.
- 33 Hermans R, Op de Beeck K, Van den Boogaert W, et al. The relation of CT-determined tumor parameters and local and regional outcome of tonsillar cancer after definitive radiation treatment. *Int J Radiat Oncol Biol Phys* 2001;50:37-45.
- 34 Nathu RM, Mancuso AA, Zhu TC, Mendenhall WM. The impact of primary tumor volume on local control for oropharyngeal squamous cell carcinoma treated with radiotherapy. *Head Neck* 2000;22:1-5.
- 35 Chao KS, Ozyigit G, Blanco AI, et al. Intensity-modulated radiation therapy for oropharyngeal carcinoma: impact of tumor volume. *Int J Radiat Oncol Biol Phys* 2004;59:43-50.
- 36 Hoogsteen IJ, Marres HA, Wijffels KI, et al. Colocalization of carbonic anhydrase 9 expression and cell proliferation in human head and neck squamous cell carcinoma. *Clin Cancer Res* 2005;11:97-106.
- 37 Hermans R. Head and neck cancer: how imaging predicts treatment outcome. *Cancer Imaging* 2006;6:S145-S153.
- 38 Troost EG, Schinagel DA, Bussink J, et al. Innovations in radiotherapy planning of head and neck cancers: role of PET. *J Nucl Med* 2010;51:66-76.

7 |

Pathology-based validation of FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer

Dominic A.X. Schinagl

Paul N. Span

Franciscus J. van den Hoogen

Matthijs A.W. Merkkx

Petrus J. Slootweg

Wim J.G. Oyen

Johannes H.A.M. Kaanders

Abstract

Background and purpose: FDG-PET is increasingly incorporated into radiation treatment planning of head and neck cancer (HNC). However, there are only limited data on the accuracy of radiotherapy target volume delineation by FDG-PET. The purpose of this study was to validate FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer against the pathological standard.

Materials and methods: Twelve HNC patients with 28 metastatic lymph nodes eligible for therapeutic neck dissection underwent preoperative FDG-PET/CT. The metastatic lymph nodes were delineated on CT (Node_{CT}) and 10 PET segmentation tools were used to assess FDG-PET based nodal volumes: interpreting FDG-PET visually (PET_{VIS}), applying an isocontour at a standardized uptake value (SUV) of 2.5 (PET_{SUV}), two segmentation tools with a fixed threshold of 40% and 50% ($\text{PET}_{40\%}$, $\text{PET}_{50\%}$), and two adaptive threshold based methods (PET_{SBR} , PET_{RTL}). The latter four tools were applied with the primary tumor as reference and also with the lymph node itself as reference. Nodal volumes were compared with the true volume as determined by pathological examination.

Results: Both Node_{CT} and PET_{VIS} showed good correlations with the pathological volume. PET segmentation tools using the metastatic node as reference all performed well but not better than PET_{VIS} . The tools using the primary tumor as reference correlated poorly with pathology. PET_{SUV} was unsatisfactory in 35% of the cases due to merging of the contours of adjacent nodes.

Conclusion: FDG-PET accurately estimates metastatic lymph node volume, but beyond the detection of lymph node metastases (staging) it has no added value over CT alone for the delineation of routine radiotherapy target volumes. If FDG-PET is used in radiotherapy planning, treatment adaptation or response assessment, we recommend an automated segmentation method for purposes of reproducibility and inter-institutional comparison.

Introduction

Progress in radiation oncology enables delivery of radiation treatment with increasing geometric precision. This requires a re-evaluation of target volume delineation, which is traditionally based on physical examination, and anatomical imaging using computed tomography (CT) and magnetic resonance imaging (MRI). Incorporating metabolic information, as provided by imaging F-18-fluorodeoxy-glucose (FDG) with positron emission tomography (PET) allows to study the pathophysiology of cancers non-invasively and has three potential advances: increased accuracy of tumour demarcation, visualization of tumour characteristics relevant for radiation sensitivity and identification of intratumoral biological heterogeneity [1].

Many research groups have studied the incorporation of FDG-PET into radiation treatment planning, and several ways of using PET information have been described [2]. Visual interpretation is the most commonly used method [3-6]. This method, however, is susceptible to the window-level settings of the images and is highly operator dependent. Therefore, more objective methods have been explored such as isocontouring based on either a standardized uptake value (SUV) [4,7-9], a fixed threshold of the maximum signal intensity [10-14], a threshold that is adaptive to the signal to background ratio (SBR) [4,15] or an iterative background-subtracted relative threshold level (RTL) [16]. It has been demonstrated that FDG-PET may have important consequences for GTV definition of the primary tumour in head and neck cancer, but that the choice of PET-segmentation tool is not trivial [17]. FDG-PET may also have a role in target volume definition of metastatic lymph nodes in head and neck cancer, but again the choice of segmentation tool may influence the results [18].

Studies that have validated these segmentation tools against histopathology are sparse. In head and neck cancer two studies have prospectively compared FDG-PET determined primary tumour volume to the histopathological assessment of laryngectomy specimens [19,20]. Daisne *et al.*, analyzed the tumours of nine patients using a threshold based PET-segmentation tool (PET_{SBR}). Caldas-Magalhaes *et al.*, analyzed the tumours of ten patients using the visual interpretation method (PET_{vis}). In both studies GTV assessed on PET best correlated with pathology-based GTV whereas both CT and MRI greatly overestimated the volume.

A PET segmentation method using a combination of a fuzzy measure and a locally adaptive Bayesian-based classification (3-FLAB) has recently been proposed [21]. This automatic approach combines statistical and fuzzy modelling to address specific issues associated with 3D-PET images, such as noise and partial volume effects. In a recent study its accuracy has been assessed using images of 18 lung cancer patients whereby the maximum tumour diameters were compared to the pathological standard [21].

A gradient-based segmentation method using the watershed transform and hierarchical cluster analysis, developed by Geets *et al.*, showed good results after tumour volume validation using seven of the nine laryngectomy specimens previously analysed by Daisne *et al.* [20,22].

Zaidi *et al.* tested several segmentation methods again using seven of the nine laryngectomy specimens previously analysed by Daisne *et al.* and found good performance for a spatial wavelet-based algorithm (FCM-SW), which incorporates spatial information during the segmentation process [20,23]. This recent publication provides a good overview of the current PET segmentation methods [23].

The aim of the current study was to correlate the volume of metastatic lymph nodes in head and neck cancer patients assessed by CT and various FDG-PET-based segmentation methods with the volume as determined by pathological examination.

Methods and materials

Patients

Twelve patients (all male, median age 60 years, range 50-75 years) with histologically proven squamous cell carcinoma metastases in one or more cervical lymph node eligible for therapeutic modified radical neck dissection were prospectively enrolled from July 2008 until February 2011. The primary tumour site was unknown (n=4), oropharynx (n=3), oral cavity (n=2), larynx (n=1), hypopharynx (n=1) and auricular skin (n=1). The study was approved by the Ethics Committee of the Radboud University Nijmegen Medical Centre and all patients provided written informed consent.

Preoperative image acquisition

The median interval between imaging and surgery was 6 days (range 1-23 days, in two patients the interval being more than two weeks). The patients underwent preoperative integrated FDG-PET and CT imaging on a hybrid PET/CT scanner (Biograph Duo; Siemens/CTI). All scans were performed with the patient supine and immobilized with an individual head support and a rigid customized mask covering the head and neck area to reduce movement artefacts during image acquisition. Patients with diabetes mellitus were not excluded, provided that glucose levels were appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection).

Emission images of the head and neck area were recorded 60 minutes after intravenous injection of 250MBq of FDG (Covidien, Petten, The Netherlands), with 7 minutes per bed position in 3-dimensional mode. PET images were reconstructed using the ordered-

subsets expectation maximization iterative algorithm with parameters optimized for the head and neck area (i.e., 4 iterations, 16 subsets, and a 5-mm 3-dimensional Gaussian filter [24]. In addition, CT images were acquired using 80 mAs, 130 kV and a 3-mm slice width.

After reconstruction, standardized uptake value (SUV) PET images were created with in-house developed software correcting for injected dose, decay of the tracer, and patient body weight. The PET and SUV-PET images were resliced using the CT format as a reference. All image data sets were imported into the Pinnacle³ treatment planning system (Philips Medical Systems, Andover, MA, USA) in order to perform volume analyses.

Lymph Node Analysis

Pathology proven and FDG-PET positive metastatic lymph nodes were delineated on CT images by two experienced radiation oncologists in consensus (Node_{CT}). Ten PET segmentation tools were used to assess FDG uptake in the nodes. Visual delineation (PET_{VIS}) was performed by identifying FDG activity clearly above normal background activity.

For the other (threshold based) segmentation tools in-house developed scripts for the Pinnacle³ treatment planning system were used. SUV-based segmentation of pathologic nodes was obtained by applying an isocontour of 2.5 (PET_{SUV}). Two thresholds were based on fixed percentages of the maximum signal intensity in the primary tumor of 40% (PET_{40%}) and 50% (PET_{50%}), respectively. An adaptive threshold tool (PET_{SBR}) based on the SBR of the primary tumor was used, as developed at Université St. Luc in Brussels, Belgium [15]. The calibration and implementation of the PET_{SBR} method was described in detail previously [17]. Finally, an in-house developed tool (PET_{RTL}) was used, which generates an iterative background-subtracted relative threshold level (RTL) as previously described in detail [16].

The PET_{40%}, PET_{50%}, PET_{SBR} and PET_{RTL} methods all used the PET signal of the primary tumour as reference. The scripts for these segmentation methods were again applied whereby now the maximum signal intensity of the specific metastatic lymph node itself was used, yielding the parameters PET_{40%N}, PET_{50%N}, PET_{SBRN}, and PET_{RTLN} respectively.

Results obtained by semi-automated delineation algorithms were checked visually before accepting the results. The PET-volume of a pathologic lymph node was considered 0 if the segmentation method failed to identify FDG-uptake (i.e. in case of a high segmentation-threshold). If the segmented PET-volume included significant amounts of tissue that was clearly normal at visual interpretation of the CT-scan (e.g. clearly uninvolved muscle or fatty tissue), the lymph node was considered "unidentified". If the segmentation procedure was influenced by one or more adjacent metastatic lymph nodes and the resulting volume

encompassed more than one metastatic node, this was considered “clustering”. If lymph nodes remained unidentified or if clustering occurred, the segmentation results were excluded from further analysis.

Pathology procedure

FDG-PET/CT images were available at the pathology department. Directly after surgery, the relevant lymph nodes were identified and excised from the neck dissection specimen by an experienced pathologist. Perinodal fibrous and fatty tissues were meticulously removed. The volume of the lymph node was then measured using water submersion.

Statistics

All statistical analyses were performed using SPSS version 20.0 (SPSS Inc. Chicago IL) and GraphPad Prism 4.0c (GraphPad Software, San Diego CA) for Mac. Linear regression analyses were performed of a segmentation method against the standard, i.e. volume as determined by pathological examination, and fits were reported as slope and 95% confidence intervals. Difference plots were generated by plotting the % difference from histopathological volume value. A $p < 0.05$ was *a priori* considered as statistically significant.

Results

A total of 28 metastatic lymph nodes were delineated on CT-scan and recovered in the 12 neck specimens. The lymph node characteristics are shown in table 1. The median pathological volume of the metastatic lymph nodes was 3.5 cm^3 (range $0.5 - 65 \text{ cm}^3$).

PET_{VIS} was generated for all 28 nodes. Segmentation tools requiring the maximum signal intensity of a primary tumour could not be generated in five patients, of whom four had an unknown primary tumour and one had a small tumour located in the oral cavity which was undetected by FDG-PET.

The volume of two metastatic lymph nodes was 0 for both PET_{SUV} and PET_{RTL}.

One metastatic lymph node remained unidentified with PET_{40%N}, PET_{50%N}, PET_{SBRN}, and PET_{RTL}. Clustering occurred once with PET_{40%N} and PET_{RTL}, twice with PET_{SBRN}, and ten times with PET_{SUV}.

Figure 1 shows the metastatic lymph node volumes as determined by submersion of the dissected node, CT measurement, and the FDG-PET-based segmentation methods. It demonstrates that the volumes generated by segmentation methods using the maximum PET signal intensity of the metastatic lymph node as reference (i.e. PET_{40%N}, PET_{50%N},

Table 1 Lymph node sizes.

Largest diameter on CT	
< 1 cm	1
1-2 cm	10
2-3 cm	11
3-4 cm	1
4-6 cm	5
Total	28
Pathological volume	
< 1 cm ³	1
1 – 2 cm ³	3
2 – 3 cm ³	2
3 – 4 cm ³	9
4 – 10 cm ³	7
>10 cm ³	6
Total	28

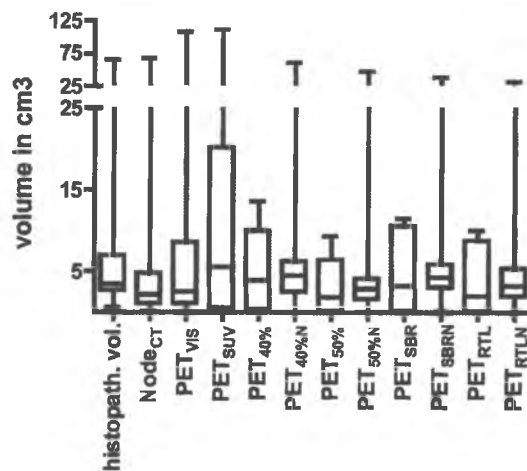
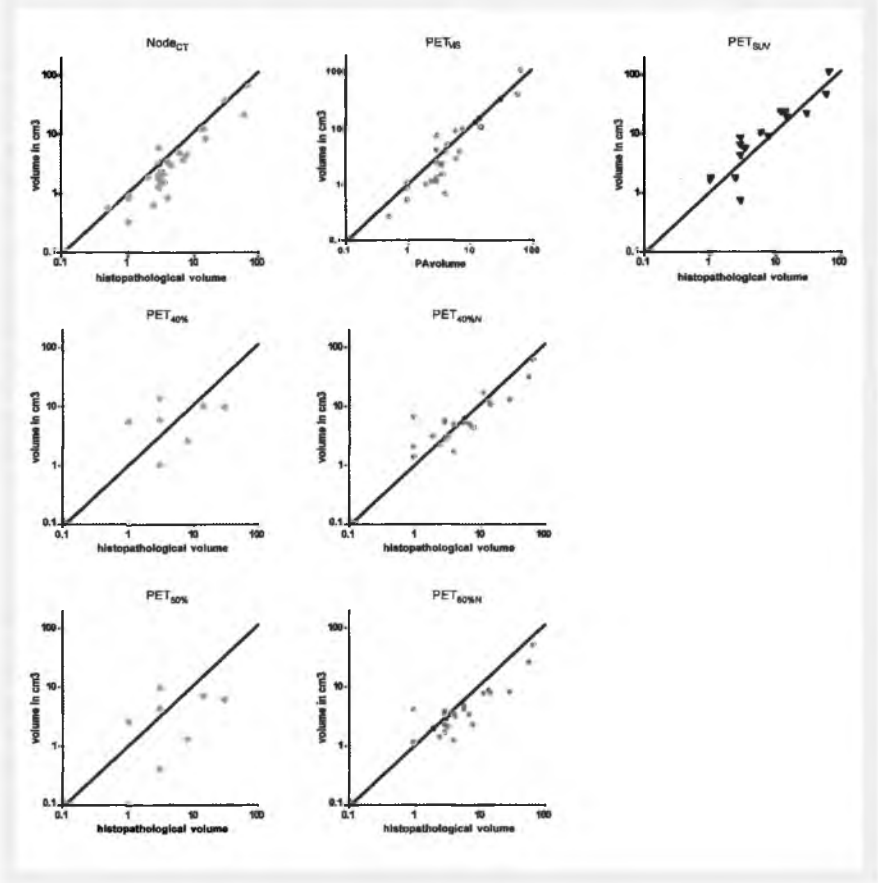
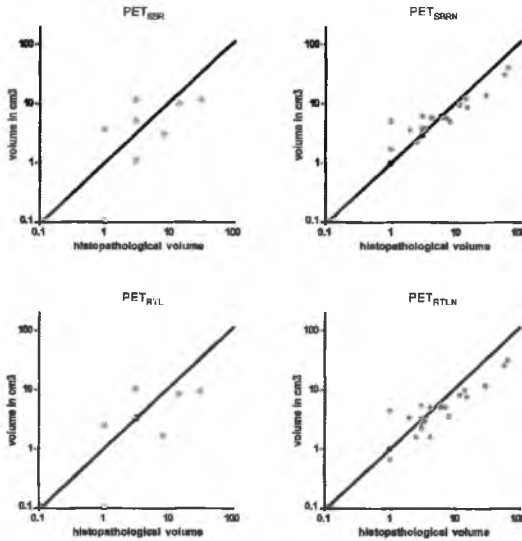
Fig. 1 Box and whisker plot showing 5% and 95% confidence intervals (whiskers), 25% and 75% confidence intervals (boxes), and median of the metastatic lymph node volumes as determined by pathological examination, CT measurement, and the various FDG-PET-based segmentation methods.

Fig. 2a-1 correlation of CT- and FDG-PET based (PET_{V15} , PET_{SUV} , $PET_{40\%}$, $PET_{40\%N}$, $PET_{50\%}$, and $PET_{50\%N}$) metastatic lymph node volumes in relation to the corresponding pathological volume. The line in the correlation plots represents identity ($y = x$).



PET_{SBRN} and PET_{RTL}) were in the same range as the pathological volumes. PET_{SUV} and the segmented volumes where FDG-uptake in the primary tumour served as the reference all demonstrated much larger variability. Figure 2 shows the correlation of CT- and FDG-PET based metastatic lymph node volumes in relation to the corresponding pathological volume and the difference plots. The resulting slopes and 95% confidence intervals after linear regression analyses of the various difference plots are shown in table 2. The data clearly indicate that the methods using the maximum PET signal of the primary tumour as reference (i.e. $PET_{40\%}$, $PET_{50\%}$, PET_{SBR} and PET_{RTL}) perform poorly with slopes significantly deviating from 1 and large confidence intervals. Both $Node_{CT}$ and PET_{V15} were closely

Fig. 2a-2 correlation of CT- and FDG-PET based (PET_{SER} , PET_{SERH} , PET_{RVL} , and PET_{RTLH}) metastatic lymph node volumes in relation to the corresponding pathological volume. The line in the correlation plots represents identity ($y = x$).



related to the pathological volume with slopes (confidence intervals) being 0.7708 (0.635 to 0.907) and 1.1440 (0.958 to 1.330) respectively.

The smallest FDG-PET positive metastatic lymph node was 0.5 cm^3 , its $Node_{CT}$ and PET_{VIS} were 0.6 cm^3 and 0.3 cm^3 , respectively. There were three nodes with a volume of 1.0 cm^3 . Their $Node_{CT}$ and PET_{VIS} were 0.4 cm^3 , 0.9 cm^3 , 1.0 cm^3 and 1.1 cm^3 , 0.8 cm^3 , 0.6 cm^3 , respectively.

Fig. 2b Difference plots of CT- and FDG-PET based metastatic lymph node volumes in relation to the corresponding pathological volume. The 95% confidence intervals are shown as dotted lines

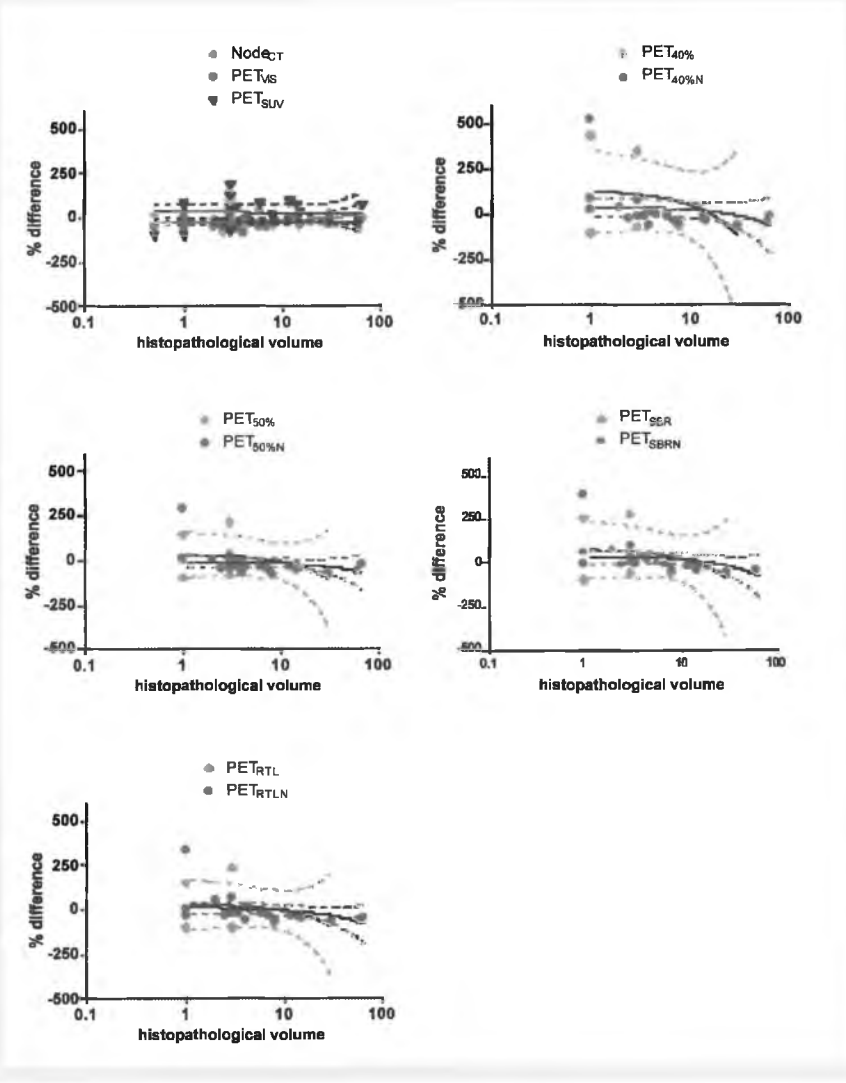


Table 2 Linear regression of CT and PET segmentation against pathological volume.

	Slope	confidence interval (range)	p-value
Node _{CT} (N=28)	0.7708	0.635 to 0.907 (0.272)	<0.0001
PET _{VIS} (N=28)	1.1440	0.958 to 1.330 (0.372)	<0.0001
PET _{SUV} (N=18)	1.2550	1.006 to 1.504 (0.498)	<0.0001
PET _{40%} (N=8)	0.4391	0.052 to 0.826 (0.774)	0.0315
PET _{40%N} (N=26)	0.7236	0.624 to 0.824 (0.120)	<0.0001
PET _{50%} (N=8)	0.2786	0.014 to 0.543 (0.529)	0.0415
PET _{50%N} (N=27)	0.5522	0.479 to 0.626 (0.147)	<0.0001
PET _{SBR} (N=8)	0.4716	0.163 to 0.780 (0.617)	0.0086
PET _{SBRN} (N=25)	0.5587	0.506 to 0.611 (0.105)	<0.0001
PET _{RTL} (N=8)	0.3786	0.114 to 0.643 (0.529)	0.0117
PET _{RTLN} (N=26)	0.4594	0.414 to 0.505 (0.091)	<0.0001

Discussion

In this study, we have compared volume of metastatic lymph nodes in head and neck cancer patients assessed by CT and various FDG-PET-based segmentation methods to the pathological standard.

As expected, segmentation methods using the maximum PET signal intensity of the metastatic lymph node as reference all performed better than the methods using the signal of the primary tumour as reference. In the latter situation the nodal volume is likely to be underestimated if the primary tumour has higher FDG uptake than the node, whereas the nodal volume can be overestimated if the primary tumour shows lower uptake. A segmentation method with reference directly related to the object of study is likely to perform best. Of the remaining segmentation methods there was not one method clearly outperforming the others. The non-automated PET_{VIS} and also Node_{CT} did not show poorer results, but actually to some extent performed better than the automated methods. Generally Node_{CT} slightly underestimated (slope 0.7708) and PET_{VIS} slightly overestimated (slope 1.1440) metastatic lymph node volume.

Tracer uptake heterogeneity can influence the delineation results of any threshold-based method due to their binary nature. However, variation in tracer uptake within the GTV was found not to be of influence on the precision of any of the delineation methods assessed here (data not shown).

Due to partial-volume effect (PVE), there may be underestimation of the actual FDG uptake, especially in small tumour masses [25]. PVE can cause intensity values in an image to differ from what they ideally should be due to two distinct phenomena [25]. The first is the image blurring due to finite spatial resolution of the imaging system, which is limited by the detector design and the reconstruction process, the second is image sampling due to the voxel grid [25]. The educational publication by Soret *et al.* provides a good overview of PVE. PVE should be kept in mind when analyzing relatively small metastatic lymph nodes, especially as correction for PVE may be associated with methodological issues [25]. No PVE correction was implemented in our study. However, in metastatic lymph nodes larger than 1 cm^3 the performance of all methods of segmentation did not appear to be related to the tumour mass, i.e. errors in volume estimates were independent over the range of node size studied.

Volume estimates based on CT and visual interpretation of PET-images were reasonably accurate and comparable. In a previous study of 78 head and neck cancer patients the primary tumour volumes based on CT and PET_{VIS} were also comparable, but there was often a significant geographical mismatch as PET frequently detected extension of tumour tissue outside the CT-based tumour volume [17]. For primary tumour delineation the addition of FDG-PET data might therefore be beneficial, especially in cases where the extension of the primary tumour is difficult to discriminate from the surrounding nonmalignant tissue or in areas where CT- and MR-images are disturbed by artefacts, e.g. dental fillings. However, as most metastatic lymph nodes are surrounded by fatty tissue, the CT-based delineation of a metastatic node is less error prone. The benefit of FDG-PET for delineation of metastatic lymph nodes might be limited to selected cases, for example nodal metastases with extensive extracapsular spread and invasion of muscular structures, skinny individuals, or situations where use of CT- or MR-contrast enhancing agents is contra-indicated.

If FDG-PET is used for radiotherapy target volume definition of lymph node metastases, the results of the current study do not support a clear preference for any automated segmentation method. PET_{VIS} yielded acceptable results even though it has its well known shortcomings, i.e. dependent of window-level settings and interoperator variability. The shortcomings of visual delineation, however, do limit that our observations in our single-centre study can be extrapolated to other settings. Therefore, objective PET segmentation methods are preferable. All currently analyzed semi-automated PET segmentation methods are threshold-based, using either fixed or adaptive approaches incorporating the background activity. Segmentation methods using adaptive thresholds such as PET_{SBRN} and PET_{RTLN} have a theoretical advantage over the methods using fixed thresholds ($\text{PET}_{40\%N}$ and $\text{PET}_{50\%N}$) since these provide a more tailored solution (i.e. adapt to the environment and background of the object of segmentation. This study shows that their performance is good, but they do not perform better than $\text{PET}_{40\%N}$ (table2).

All these approaches are system-dependent, they require manual delineation of background regions of interest, and their performance depends on parameters requiring optimization using phantom acquisitions for each scanner and reconstruction. An interesting novel approach to PET lesion segmentation uses a combination of a fuzzy measure and a locally adaptive Bayesian-based classification (3-FLAB). This automatic approach combines statistical and fuzzy modelling to address specific issues associated with 3D-PET images, such as noise and partial volume effects. In a recent study its accuracy has been assessed on both simulated and clinical images of complex shapes containing inhomogeneous activities and small regions [21]. It was demonstrated that accuracy and robustness of this algorithm are higher compared with adaptive threshold methods. Other novel approaches include the gradient-based segmentation method using the watershed transform and hierarchical cluster analysis as developed by Geets *et al.* and the spatial wavelet-based algorithm (FCM-SW), which incorporates spatial information during the segmentation process and was recently tested by Zaidi *et al.* [22,23]. The recent publication by Zaidi *et al.* provides a good overview of the current PET segmentation methods [23].

To our knowledge the current study is the first to prospectively validate study FDG-PET volume assessment of metastatic lymph nodes in head and neck cancer or any other tumour site. In head and neck cancer two groups have prospectively compared FDG-PET based tumour volume to the pathological standard and three other groups performed retrospective analyses [19,20,26-28]. The focus was, however, mostly on the primary tumour and only one of the retrospective studies included cervical lymph node metastases in the volume assessments but results were not reported separately from primary tumour volumes [27].

Pathologic tumour volume was overestimated by PET_{VIS} in one study [19], and underestimated by PET_{VIS} in another study [27]. Tumour volume was overestimated by segmentation tools using various fixed SUV thresholds in three studies [26-28]. Also, PET_{SBR} overestimated true tumour volume but was significantly better than CT and MRI [20]. One study found that CT and $PET_{40\%}$ were closest to the histopathological standard [27].

The pathology procedure used in the current study has potential shortcomings. First, perinodal fibrous and fatty tissue was carefully removed but in some cases small residuals remained. Second, the lymphatic tissue is not always completely replaced by tumour, especially in the smaller nodes. And, finally, no strict correction could be made for areas of intranodal necrosis.

A limitation of this study is that the total number of patients is relatively small.

On the basis of this study, it seems justified to conclude that FDG-PET will not improve the overall accuracy of radiotherapy target volume definition of metastatic lymph nodes.

This does not disqualify other potential applications of FDG-PET in radiation treatment of head and neck cancer patients, such as increased accuracy for identification of the tumour and its metastases, visualization of tumour characteristics relevant for radiation sensitivity and identification of intratumoral biological heterogeneity and also its potential role in treatment response prediction and evaluation [1,18,29,30]. For all these situations a standardized method of PET-signal interpretation is highly preferable. Such a method should be operator independent, easy to use, and calibration should be straightforward, in order to facilitate multicentre introduction. This would allow proper comparison of results of various research groups, and should be the basis of new multicentre study protocols.

In conclusion, FDG-PET accurately estimates metastatic lymph node volume, but we do not recommend FDG-PET for routine radiotherapy target volume delineation of lymph node metastases as it has no added value over CT alone for this indication unless in selected circumstances where demarcation from the surrounding tissues is difficult. If FDG-PET is used in radiotherapy planning, treatment adaptation or response assessment, we recommend an automated segmentation method for purposes of reproducibility and to facilitate inter-institutional comparisons.

References

- Schinagl DA, Kaanders JH, Oyen WJ. From anatomical to biological target volumes: the role of PET in radiation treatment planning. *Cancer Imaging* 2006;6:S107-S116.
- Troost EG, Schinagl DA, Bussink J, Boerman OC, van der Kogel AJ, Oyen WJ, et al. Innovations in radiotherapy planning of head and neck cancers: role of PET. *J Nucl Med* 2010;51:66-76.
- Heron DE, Andrade RS, Flickinger J, Johnson J, Agarwala SS, Wu A, et al. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.
- Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1348.
- Nishioka T, Shiga T, Shirato H, Tsukamoto E, Tsuchiya K, Kato T, et al. Image fusion between 18Fdg-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057.
- Riegel AC, Berson AM, Destian S, Ng T, Tena LB, Mitnick RJ, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys* 2006;65:726-732.
- Bradley J, Thorstad WL, Mutic S, Miller TR, Dehdashti F, Siegel BA, et al. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;59:78-86.
- Hong R, Halama J, Bova D, Sethi A, Emami B. 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* 2007;67:720-726.
- Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys* 2004;59:4-5.
- Brianzoni E, Rossi G, Ancidei S, Berbellini A, Capocchetti F, Cidda C, et al. Radiotherapy planning: PET/CT scanner performances in the definition of gross tumor volume and clinical target volume. *Eur J Nucl Med Mol Imaging* 2005;32:1392-1399.
- Ciernik IF, Dizendorf E, Baumert BG, Reiner B, Burger C, Davis JB, et al. Radiation treatment planning with an integrated positron emission and computed tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
- Mah K, Caldwell CB, Ung YC, Danjoux CE, Balogh JM, Ganguli SN, et al. The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int J Radiat Oncol Biol Phys* 2002;52:339-350.
- Miller TR, Grigsby PW. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. *Int J Radiat Oncol Biol Phys* 2002;53:353-359.
- Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
- Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Int J Radiat Oncol Biol Phys* 2003;69:247-250.
- Van Dalen JA, Hoffmann AL, Dicken V, Vogel WV, Wiering B, Ruers TJ, et al. A novel iterative method for lesion detection and volumetric quantification with FDG-PET. *Nucl Med Commun* 2007;28:485-493.
- Schinagl DA, Vogel WV, Hoffmann AL, van Dalen JA, Oyen WJ, Kaanders JH. Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:1282-1289.
- Schinagl DA, Hoffmann AL, Vogel WV, van Dalen JA, Verstappen SM, Oyen WJ, et al. Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer? *Radiother Oncol* 2009;91:95-100.
- Caldas-Magalhaes J, Kasperts N, Kooij N, van den Berg CA, Terhaard CH, Raaijmakers CP, et al. Validation of imaging with pathology in laryngeal cancer: accuracy of the registration methodology. *Int J Radiat Oncol Biol Phys* 2012;82:e289-e298.

20. Daisne JF, Duprez T, Weynand B, Lonnew M, Hamoir M, Reyckler H, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology* 2004;233:93-100
21. Hatt M, Cheze le RC, Descourt P, Dekker A, De RD, 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;77:301-308
22. 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;34:1427-1438
23. Zaidi H, Abdoli M, Fuentes CL, El Naqa I. Comparative methods for PET image segmentation in pharyngolaryngeal squamous cell carcinoma. *Eur J Nucl Med Mol Imaging* 2012;39:881-891
24. Vogel WV, Wensing BM, van Dalen JA, Krabbe PF, van den Hoogen FJ, Oyen WJ. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2005;32:1276-1282
25. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med* 2007;48:932-945
26. Baek CH, Chung MK, Son YI, Choi JY, Kim HJ, Yim YJ, et al. Tumor volume assessment by 18F-FDG PET/CT in patients with oral cavity cancer with dental artifacts on CT or MR images. *J Nucl Med* 2008;49:1422-1428
27. Burri RJ, Rangaswamy B, Kostakoglu L, Hoch B, Genden EM, Som PM, et al. Correlation of positron emission tomography standard uptake value and pathologic specimen size in cancer of the head and neck. *Int J Radiat Oncol Biol Phys* 2008;71:682-688
28. Seitz O, Chambron-Pinho N, Middendorp M, Sader R, Mack M, Vogl TJ, et al. 18F-fluorodeoxyglucose-PET/CT to evaluate tumor, nodal disease, and gross tumor volume of oropharyngeal and oral cavity cancer: comparison with MR imaging and validation with surgical specimen. *Neuroradiology* 2009;51:677-686
29. Arens AI, Troost EG, Schinagl D, Kaanders JH, Oyen WJ. FDG-PET/CT in radiation treatment planning of head and neck squamous cell carcinoma. *Q J Nucl Med Mol Imaging* 2011;55:521-528
30. Schinagl DA, Span PN, Oyen WJ, Kaanders JH. Can FDG PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. *Eur J Nucl Med Mol Imaging* 2011;38:1449-1458

8 |

General discussion and future perspectives

Adapted from

**Clinical evidence on PET-CT for radiation
therapy planning in head and neck tumors**

Esther G.C. Troost

Dominic A.X. Schinagl

Johan Bussink

Wim J.G. Oyen

Johannes H.A.M. Kaanders

Radiother Oncol 2010;96:328-334

**Innovations in radiotherapy planning of
head and neck cancers: role of PET**

Esther G.C. Troost

Dominic A.X. Schinagl

Otto C. Boerman

Albert J. van der Kogel

Johan Bussink

Wim J.G. Oyen

Johannes H.A.M. Kaanders

J Nucl Med 2010;51:66-76

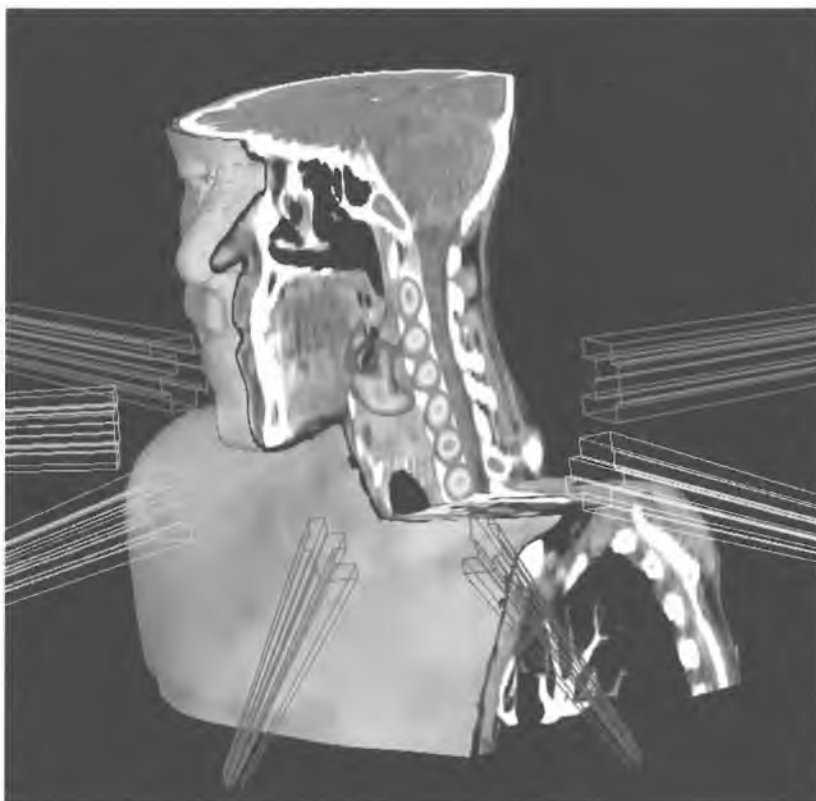
High-precision radiotherapy for head and neck tumors

The field of radiation oncology has changed dramatically after the wide introduction of computer-optimized intensity modulated radiation therapy (IMRT) in the beginning of the 21st century. IMRT is based on the use of numerous radiation beams with optimized nonuniform intensities resulting from inverse treatment planning. The algorithm for beam fluence calculations is guided by dose-volume objectives for the target volume and organs at risk delineated by the radiation oncologist. IMRT can thus achieve much better dose conformity than conventional radiotherapy techniques. With this technique, different dose prescriptions to multiple target sites can be delivered. It also facilitates boosting of high radiation doses to the primary tumor while reducing the dose to radiation-sensitive tissues adjacent to the tumor such as the salivary glands and swallowing structures [1-5]. Due to the highly conformal dose distribution and steep dose gradients used in IMRT, knowledge about the localization and boundaries of the primary tumor and of the cervical lymph node metastases is of increasing importance. For this purpose, biological imaging using positron emission tomography (PET) may augment traditional imaging methods such as computed tomography (CT) and magnetic resonance imaging (MRI) (Fig.1.)

FDG-PET

In 2008 a multidisciplinary expert panel developed recommendations for the use of FDG-PET in oncology practice [6]. Their recommendations on the use of FDG-PET for the detection and staging of head and neck tumors are briefly summarized. The expert panel concluded that FDG-PET should not be added to conventional anatomic imaging in the routine diagnostic work-up of primary head and neck tumors [6]. This conclusion was drawn because the available data were too uncertain as to whether FDG-PET can determine the anatomic extent of the primary tumor more accurately than CT or MRI. This recommendation remains unaltered in an evidence based guideline for the use of PET in head and neck cancer published in 2013, and the only routine indication for FDG-PET is for the detection of an unknown primary head and neck tumor as FDG-PET clearly outperforms CT or MRI [7]. Regarding the detection of cervical lymph node metastases, the expert panel concluded that FDG-PET had a higher specificity, sensitivity, positive predictive value and negative predicted value than CT and MRI. Therefore, its use in routine nodal staging was recommended [6]. This recommendation, however, did not incorporate a meta-analysis by Kyzas *et al.* that was published shortly thereafter [8]. This meta-analysis reviewed 35 studies using FDG-PET for the pretreatment evaluation of the lymph node status. The authors concluded that there was no solid evidence to support the routine application of FDG-PET, as the sensitivity and specificity improved by only 5%

Fig. 1 FLT-PET/CT scan for image-guided high-precision radiation treatment planning in oropharyngeal cancer



- 7% compared with conventional imaging modalities, but direct comparison revealed no statistically significant differences. In the subset of studies only enrolling patients without clinically apparent cervical lymph node metastases, the sensitivity was only 50% and not better than conventional imaging methods, specifically ultrasound with fine-needle aspiration cytology. From these contradictory recommendations, it is clear that this issue is unresolved and requires further study. In 2012, a review stated that FDG-PET does not improve accuracy for nodal staging of head and neck tumours to a clinically relevant degree when compared to CT or MRI [9]. In 2013, the evidence based guideline for the use of PET in head and neck cancer recommended FDG-PET for nodal staging only when conventional imaging was equivocal, or when treatment may be significantly modified.

For the detection of distant metastases, FDG-PET might be beneficial in patients with advanced-stage disease, in whom the odds of having distant metastases are greater [6]. In these patients, the FDG-PET findings may alter the treatment intention from curative to palliative and thus affect the total dose and fractionation scheme. Additionally, it may reduce treatment-related side effects in those patients, as the selected treatment volume is often confined to the primary tumor or the metastatic lymph nodes causing discomfort or pain.

As discussed, the value of FDG-PET for staging of the primary tumor and the cervical lymph nodes is controversial. However, in the meantime the incorporation of FDG-PET data for radiation treatment purposes is performed in an increasing number of patients.

Delineation of radiation therapy target volume: primary tumor

There are several potential advantages with the use of FDG-PET for target volume delineation: reduction of inter-observer variability in gross tumor volume (GTV) delineation, reduction of the size of the GTV, identification of tumor extensions that were missed by CT or MRI, and the possibility of identifying parts of the GTV potentially requiring an additional radiation dose. Drawbacks in the use of FDG-PET are: the limited spatial resolution, the lack of a standardized method of signal segmentation, and false-positive FDG-PET readings caused by inflammation.

A reduction of interobserver variability has been demonstrated for non-small cell lung cancer when FDG-PET was incorporated in GTV delineation [10-11]. In patients with head and neck cancer, this finding has been less consistent. Ciernik *et al.* investigated the value of FDG-PET in 39 patients with various solid tumors, of which 12 were head and neck cancer [12]. The investigators found both increase ($\geq 25\%$) and decrease ($\leq 25\%$) in half the patients when GTV delineation was based on CT alone compared with FDG-PET/CT. When GTV delineation was compared between two experienced radiation oncologists, the mean volume difference of 26.6 cm³ by CT alone was reduced to 9.1 cm³ with FDG-PET/CT [12]. Riegel *et al.* found conflicting results when two experienced radiation oncologists and two neuroradiologists delineated 16 patients with head and neck cancer [13]. On average, the GTV based on FDG-PET/CT were larger than the corresponding CT-based volumes. Furthermore, the authors observed a large discrepancy between the GTV delineation of the two radiation oncologists, with one delineating larger volumes on CT and the other on FDG-PET/CT [13]. An important difference between these two studies relates to the thresholding of the FDG-PET signal. Ciernik *et al.* chose a fixed threshold of 50% of the maximum signal intensity of the primary tumor, whereas Riegel *et al.* used a discretionary window-level setting [12-13].

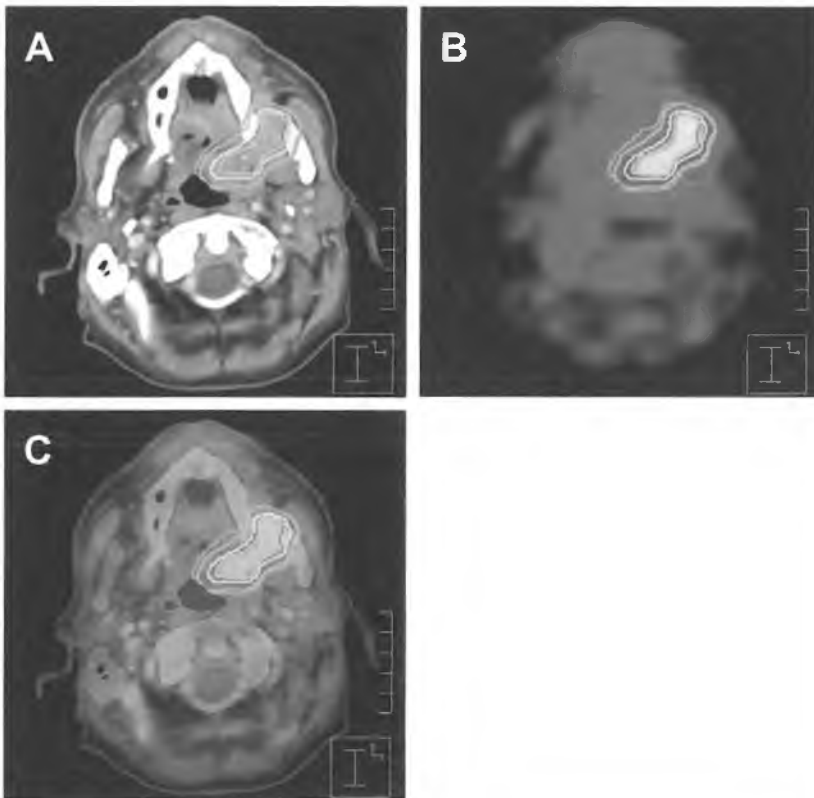
A reduction in the size of the GTV using FDG-PET has been demonstrated in a landmark study comparing the role of co-registered CT, MRI and FDG-PET in GTV delineation of laryngeal cancer in patients scheduled for laryngectomy [14]. FDG-PET was closest to depict true tumor volume when compared to the reference surgical specimen. All modalities overestimated the extension of the tumor, with an average of 29%, 65%, and 89% for FDG-PET, CT and MRI, respectively. However, all three imaging modalities, including FDG-PET, failed to identify a small fraction of the macroscopic tumor (approximately 10%), mainly consisting of superficial mucosal extensions.

Before PET-based GTVs can reliably and reproducibly be incorporated into high-precision radiotherapy planning, operator-independent segmentation tools have to be developed and validated. Visual interpretation of the PET signal is most commonly applied but is highly susceptible to the window-level settings of the images and interpretation differences [13,15,16]. This is why more objective methods such as isocontouring based on a chosen standardized uptake value (SUV) were explored, e.g. of 2.5, or thresholds acquired through phantom experiments such as a fixed threshold of the maximum tumor signal intensity (40% or 50%) or a variable threshold adaptive to the signal-to-background ratio [12,17-19].

In a study 78 head and neck cancer patients eligible for primary (chemo)radiotherapy, five commonly used methods of FDG-PET signal segmentation were compared [20]. It showed that the volume and the shape of the resulting GTV were influenced heavily by the choice of the segmentation tool (Fig. 2.) Visual interpretation of the PET signal yielded volumes close to those of the CT-based GTV delineation, whereas all automated segmentation methods resulted in significantly smaller GTVs based on clinical information and CT alone [20]. Furthermore, in a large percentage of patients (between 29% and 64%, depending on the segmentation tool used) more than 20% of the FDG-PET-based GTV was located outside the GTV based on clinical information and CT. This suggests that tumor could be identified by FDG-PET that was missed using the standard methods of GTV delineation. However, in the absence of histologic validation it is unknown in what percentage of cases this was caused by peritumoral inflammation, resulting in a false-positive reading of the FDG-PET signal.

In recent years, promising segmentation tools have been developed taking into account the underlying PET physics. Geets *et al.* have published a gradient-based segmentation tool based on watershed transform and hierarchical cluster analysis [21]. Van Dalen and co-workers developed an iterative background-subtracted relative-threshold level (RTL) method [22]. Finally, Hatt *et al.* have analyzed a segmentation tool that uses a combination of a fuzzy measure and a locally adaptive Bayesian-based classification (3-FLAB), which combines statistical and fuzzy modeling to address specific issues associated with 3D-PET images, such as noise and partial volume effects [23]. Although all groups validated their

Fig. 2 Planning CT scan (A), corresponding FDG-PET scan (B) and fusion image (C) of a patient with T4N2cM0 oropharyngeal cancer show differences in target volume delineation. Indicated are GTV delineated on CT (GTV_{CT} , red), and PET-based GTVs obtained by visual interpretation (GTV_{vis} , light green), using fixed threshold of 40% ($GTV_{40\%}$, yellow) and 50% ($GTV_{50\%}$, blue) of the maximum signal intensity, applying an adaptive threshold based on the signal-to-background ratio (GTV_{sbr} , dark green, largely covered by $GTV_{50\%}$, in blue) and applying an isocontour of a standardized uptake value (SUV) of 2.5 ($GTV_{2.5}$, orange). The respective volumes ranged from 15.1cm³ ($GTV_{50\%}$) to 59.7cm³ (GTV_{vis}).



new tools, broader experience implementing these in the research setting is compulsory. In the future, a standardized method of PET-signal interpretation is preferable. Such a method should be operator independent, easy to use, and calibration should be straightforward, in order to facilitate multicentre introduction. This would allow proper comparison of results in various research groups, and should be the basis of new multicentre study protocols.

Delineation of radiation therapy target volume: cervical lymph node metastases

In head and neck cancer, delineation studies incorporating FDG-PET have mostly concentrated on the primary tumor. CT-based delineation of metastatic lymph nodes usually is less prone to error due to better discrimination from the surrounding fatty tissue. This can be more difficult in cases with large, matted nodes. FDG-PET might be helpful in these situations, although one should be aware of the caveat of FDG-PET negative necrotic parts.

In a recent study, it was shown that the segmented cervical lymph node volumes again depended on the segmentation tool applied [24]. The potential value of FDG-PET may further be in the decision-making whether marginally enlarged lymph nodes should be included in the boost volume and to which dose levels these nodes should be treated.

It can be concluded that FDG-PET can provide important complementary information for radiotherapy planning in head and neck cancer. The GTV may be reduced which can facilitate the sparing of nearby normal structures and allow dose escalation to relatively small boost volumes. Furthermore, FDG-PET may identify areas of tumor spread not recognized by anatomical imaging, which can potentially improve the accuracy of GTV definition. However, to address the clinical value of these concepts, additional histologic validation studies and properly designed clinical studies for the evaluation of local tumor control and the radiation-induced toxicity are necessary before FDG-PET can be used safely in routine daily practice.

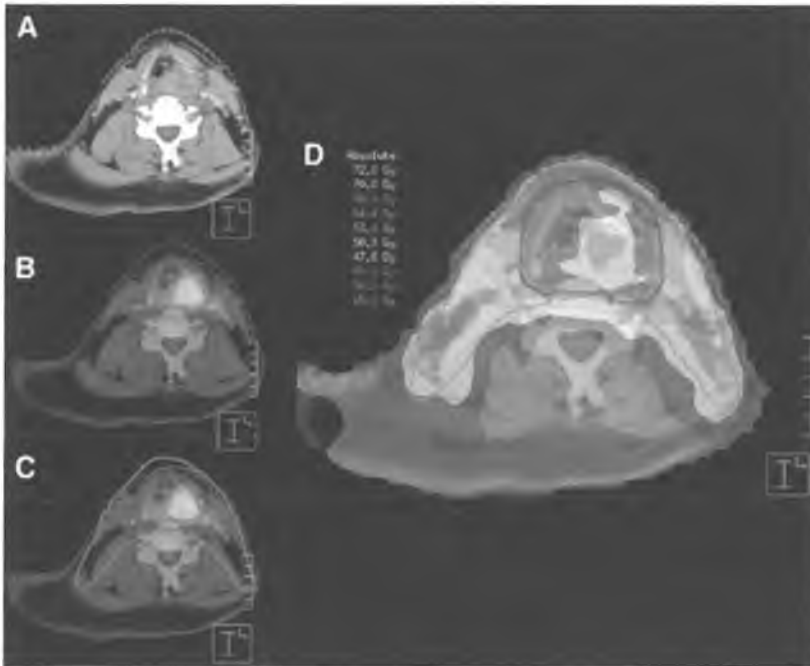
Dose escalation based on FDG-PET

The metabolic activity detected by FDG-PET may be indicative of tumor cell density or aggressiveness. Therefore, FDG-PET may be used to direct dose escalation to FDG-avid subvolumes of the tumor applying either uniform or voxel intensity-based dose escalation [5,25,26]. Schwartz *et al.* were the first to deliver a uniform escalated dose of 75 Gy in a theoretical planning study involving 20 patients with head and neck cancer [25]. Vanderstraeten *et al.* applied voxel-intensity-based dose escalation, whereby the FDG signal intensity in the PET voxel is proportionally related to the dose described to that voxel; that is, the higher the PET signal, the higher the prescribed dose [26]. The feasibility of dose escalation using a uniform dose distribution was demonstrated in a phase I clinical trial including 41 head and neck cancer patients, delivering doses up to 77.5 Gy, partly in 3 Gy fractions employing IMRT with simultaneous integrated boost [27]. Data on treatment outcome applying this approach are still pending. Figure 3 shows a FDG-PET-CT guided dose escalation plan of a head and neck cancer patient.

Initial clinical results after FDG-PET-CT guided IMRT planning

Recently, retrospective studies addressed the impact of integration of FDG-PET-CT data

Fig. 3 Planning CT scan (A), corresponding FDG-PET scan (B), fusion image (C) and calculated dose distribution (D) of a patient with T3N2M0 hypopharyngeal cancer. Red = GTV_{CT} (absolute volume of 39 cm³); light blue = GTV_{GR} (absolute volume of 131 cm³). Also illustrated are planning target volume to 50.3 Gy (pink), to 68.0 Gy (dark blue). Additional dose of 4.0 Gy is directed to GTV_{GR} using IMRT with integrated boost technique in accelerated scheme. A subvolume of GTV_{CT} thus receives a total dose of 72.0 Gy.



into IMRT planning on clinical treatment outcome. A case-control study compared 45 patients with stage IV-A pharyngeal carcinomas treated with FDG-PET-CT based IMRT with a historical matched cohort receiving three-dimensional conformal radiotherapy without FDG-PET [28]. The 2-year overall survival and event-free survival rates of patients treated with FDG-PET-CT-based IMRT were 91% and 80% and significantly better than for the control group. A similar study reported 2-year overall survival and disease-free survival rates of 83% and 71% respectively, for 42 patients with head and neck cancer of various stage and subsites [29]. Toxicity profiles in this second study were reported as favorable.

Although encouraging, the results of these studies must be interpreted cautiously because they suffer from a number of flaws including small and heterogeneous patient populations, short follow-up, and use of historical controls. Furthermore, it remains unclear

from both studies whether the suggested improvements in tumor control must be attributed to improved radiotherapy techniques, or the introduction of FDG-PET-CT or to other factors.

Adaptive radiation treatment planning based on repetitive FDG-PET

Thus far, only one proof of concept study on ten patients with pharyngo-laryngeal squamous cell carcinomas has addressed the impact of adaptive radiation treatment planning in the head and neck [30]. The patients were repetitively imaged using contrast-enhanced CT, MRI and dynamic FDG-PET before start of treatment and then once weekly during week 2-5. GTVs were delineated in CT and MRI, and segmented on PET using the gradient-based segmentation method [21]. Furthermore, the clinical target volume (CTV), planning target volume (PTV) and organs at risk (i.e. parotid glands, spinal cord, oral cavity) were defined and treatment plans calculated using the SIB IMRT approach. The GTVs delineated from functional imaging were at all times significantly smaller than those defined on anatomical imaging. During the course of treatment, the CTVs and PTVs progressively decreased; at 45 Gy the mean volumes had decreased by 51% and 48%, respectively. However, these findings did not translate into significantly reduced average doses to the organs at risk.

This adaptive treatment planning approach combined with highly conformal dose delivery poses the possibility of dose escalation impacting on tumor control. However, clinical trials need to first address the safety of this approach and assess the possible improvement in outcome.

Imaging biological tumor characteristics relevant to radiation treatment response

Three major tumor characteristics adversely affect treatment outcome and prognosis after radiation therapy: tumor cell hypoxia, repopulation during the course of treatment, and intrinsic radioresistance. These factors largely determine the outcome of radiotherapy in terms of local control and regional tumor control but ultimately also the risk of distant metastases and survival. PET enables noninvasive biologic profiling of the tumor before and during radiation treatment, with the potential to tailor therapy according to individual characteristics.

Hypoxia

Hypoxia is a feature of many solid tumors and in particular squamous cell carcinomas of the cervix and the head and neck [31,32]. Tumor cell hypoxia can result from two mechanisms: limited diffusion capacity of oxygen due to large distance from the supplying

blood vessel (chronic hypoxia), or impaired perfusion of the supplying vessel due to temporary vasoconstriction or endovascular obstruction (acute hypoxia) [33]. Treatment modifications are available, but at the cost of increased morbidity [34,35]. To individualize treatment and to select patients for these treatment modifications, assessment of the tumor oxygenation status is compulsory. In accessible tumors of the head and neck or uterine cervix, this assessment can be done by invasive polarographic electrode measurements or by immunohistochemical staining of markers in tumor biopsies [36-38]. The advantage of the polarographic electrodes is that the entire tumor can be mapped using multiple tracks. However, its clinical use is limited by the invasive nature of the procedure, the restriction to accessible tumors, and the inability to distinguish between normal, necrotic and tumor tissue. Immunohistochemical staining of tumor biopsy samples results in high-resolution images that can be analyzed for several endogenous markers of interest. Unfortunately, the tumor biopsy samples are often small and represent only a fraction of the entire tumor. Furthermore, exogenous markers require intravenous administration before biopsy samples can be taken. Finally, the acquisition of a tumor biopsy often requires the use of general anesthesia, and this procedure is not attractive for repetitive measurement. Noninvasive imaging using PET can provide a spatial map of the intratumoral distribution of hypoxia before and during treatment. This information can potentially be used not only as a selection instrument for treatment modification, but also for optimization of radiotherapy planning and delivery.

¹⁸F-fluoromisonidazole (FMISO) is a nitroimidazole PET tracer that is reduced and bound to cell constituents under hypoxic conditions. In the early 1990s, FMISO-PET was applied in several small clinical trials on different primary tumors [39-41]. Since then, FMISO-PET has been extensively used for the delineation of hypoxia in head and neck tumors [42-48]. Importantly, in head and neck cancer it was shown that the level of hypoxia depicted by FMISO-PET before treatment was correlated with locoregional failure [42,46,49]. Apart from its prognostic value, Rishin *et al.* published data supporting the predictive value of FMISO-PET [46]. They performed FMISO-PET scans in patients with advanced-stage head and neck carcinomas that were treated with chemoradiotherapy with or without tirapazamine (a hypoxic cytotoxin). Patients with hypoxic primary tumors treated with tirapazamine experienced significantly fewer local failures than patients treated with chemoradiotherapy alone. Furthermore, the absence of hypoxia on FMISO-PET was associated with a low risk of locoregional failure when treated with chemoradiotherapy alone [46]. FMISO-PET can thus serve as a predictive tool allowing treatment selection based on biologic tumor characteristics. Ultimately, reduction of side effects in patients not benefiting from treatment modification will be feasible.

Recently, Zips *et al.*, explored the use of FMISO-PET before and during treatment as a way to select patients at high risk of developing local recurrence. They performed FMISO-PET

scans in 25 stage III/IV head and neck cancer patients at four time points during chemoradiotherapy, at baseline, at end of week 1 (after 8-10Gy), at end of week 2 (after 18-20Gy) and during week 5 (after 50-60Gy) [50]. In this exploratory study they found that FMISO imaging during the initial phase of treatment carries strong prognostic value for identifying patients at risk from local recurrence. Baseline imaging was not found to be as strong in comparison, presumably as different treatment-related reoxygenation profiles exist within individual tumours.

Apart from tumor characterization, first attempts were made to delineate a biologic target volume and to escalate the dose to the primary tumor based on FMISO-PET [51-53]. Two theoretic planning studies proved the feasibility of dose escalation to the FMISO-PET-detected hypoxic subvolume using IMRT [51,52]. Rajendran *et al.* demonstrated that, using an IMRT technique, the dose to the FMISO-PET-detected hypoxic subvolume could be escalated by an additional 10Gy [52]. Lee *et al.* achieved a dose of 84 Gy in hypoxic areas without exceeding the normal-tissue tolerance [51]. Their attempt to further escalate the dose to 105 Gy in hypoxic regions was successful in only one of the two plans studied. In a third study, Thorwarth *et al.* compared IMRT planning with dose painting by numbers based on dynamic FMISO-PET data [53]. Thereby, spatially variant doses are delivered to the tumor according to dose-escalation factors determined on the basis of the dynamic FMISO-PET scan. With this approach, the tumor control probability was increased from 56% to 70% while the same level of toxicity was maintained [53]. However, one has to be cautious in interpreting the data because the number of patients included in this study was very small.

Clinical experience with hypoxic PET tracers other than FMISO is increasing. $^{60}\text{Cu}(\text{II})$ -diacetyl-bis(N4-methylthiosemicarbazone) (^{60}Cu -ATSM) was introduced into the clinic after successful preclinical studies demonstrating a strong correlation between tracer uptake and a low level for partial pressure of oxygen [54]. It was the first hypoxia-related PET tracer for which the potential use of a selective boost to the hypoxic subvolume was illustrated [55]. However, partly because of its limited specificity, especially if imaging is performed at early time points after administration, this compound did not find its way into larger-scale clinical studies.

^{18}F -fluoro-erythronitroimidazole (FETNIM), ^{18}F -fluoro-azomycin-arabinoside (FAZA), and ^{18}F -2-(2-nitro-imidazol-1-yl)-N-(3,3,3-trifluoropropyl)-acetamide (F-EF3) are members of a new generation of nitro-imidazoles. FETNIM showed a higher and more heterogeneously distributed tracer uptake in tumors than in adjacent neck muscle [56]. Furthermore, a high uptake of FETNIM before radiation therapy was associated with a trend toward poor overall survival [57]. FAZA has similar tracer characteristics to FETNIM and was proven feasible and of sufficient quality for clinical use in patients with head and neck cancer

[58,59]. Grosu *et al.* incorporated FAZA-PET into radiation treatment planning and detected hypoxic subvolumes of different sizes and distributions (representing on average 11% of the primary tumor volume and 8% of the metastatic lymph node volume) [60]. Dose escalation to 80.5Gy in FAZA-PET-detected hypoxic areas was shown to be feasible. Mortensen *et al.* performed FAZA-PET scans in 40 head and neck cancer patients receiving primary radiotherapy in order to identify hypoxia [62]. They detected a large inter-tumor variability in FAZA uptake, which was associated with poor outcome in patients with hypoxic tumors. F-EF3 was used in a phase I study of patients with head and neck cancer [61]. In that study, the use of this tracer was shown to be safe, but the number of advanced-stage tumors showing increased tracer uptake was disappointingly low.

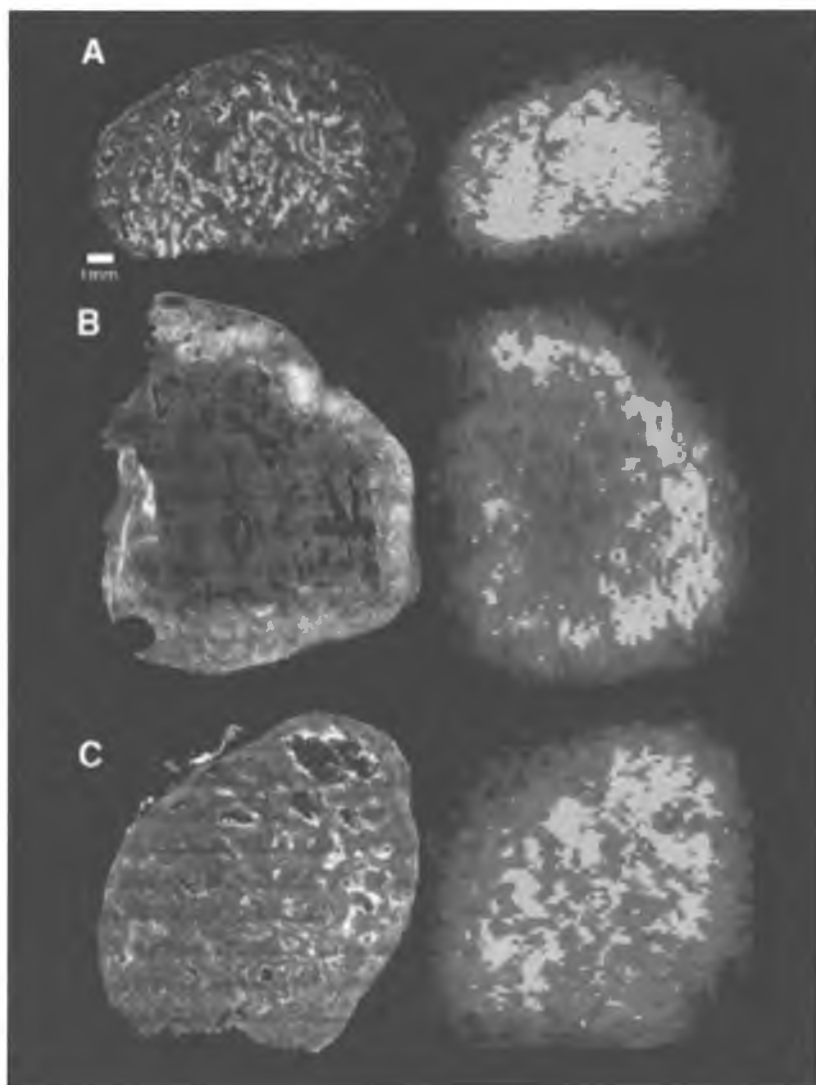
In summary, although numerous hypoxic or hypoxia-related PET tracers are available for clinical use, their prognostic and predictive value needs to be assessed in larger clinical studies before implementation for patient selection. Preferably, the PET tracer used must also visualize changes in the oxygenation status caused by treatment modifications counteracting hypoxia, such as carbogen breathing [63]. More important, the concept of dose painting to hypoxic subvolumes either by uniform doses or by dose painting by numbers is still the subject of intense debate. There are major concerns about the spatial resolution of hypoxic PET when compared with the distribution and fluctuation of tumor cell hypoxia at the microregional level. In this context, Troost *et al.* investigated ten different head and neck carcinoma xenograft tumor lines using FMISO autoradiography and pimonidazole immunohistochemistry (Fig. 4) [64]. They found that the pattern of the FMISO signal depended on the distribution of hypoxia at the microregional level. In five xenograft tumor lines, a significant correlation between the mean FMISO and pimonidazole signal intensities was found, and this correlation depended on the underlying micro-architecture. This finding indicates that one should be cautious when studying small tumor subvolumes for dose escalation [64]. Apart from different distribution patterns of hypoxia at the microregional level, one has to consider that the oxygenation status changes during the course of radiotherapy, making repetitive PET imaging before and during treatment compulsory [65]. Finally, the question on the radiation dose levels required for effective elimination of the radioresistant subpopulations remains unanswered.

Tumor cell proliferation

The major limitations of FDG-PET in oncology are false-positive readings due to tracer uptake in inflammatory tissue or reactive lymph nodes. Therefore, PET tracers that more specifically image DNA synthesis are being developed and tested.

Tumor cell proliferation during the course of therapy adversely affects radiation treatment outcome and prognosis in squamous cell carcinomas of the head and neck [66]. 3'-deoxy-3'-¹⁸F-fluorothymidine (FLT) is a tracer that reflects the activity of thymidine kinase 1, a

Fig. 4 Gray-value images after immunohistochemical staining of nitroimidazole hypoxia marker pimonidazole (left) and FMISO autoradiography images (right) of SCCNij3 (A), SCCNij153 (B), and SCCNij86 (C) xenografted human squamous cell carcinomas of head and neck. Corresponding staining patterns and signal intensities are seen for pimonidazole and FMISO in SCCNij3 and SCCNij153, but not in SCCNij86.



principal enzyme in the salvage pathway of DNA synthesis [67]. The FLT-PET signal is more specific for actively dividing tumor cells than is the FDG-PET signal. Inflammatory cells near the tumor consume glucose and thus cause false-positive FDG-PET readings. However, as these immune response cells are terminally differentiated, the DNA synthesis rate and therefore the FLT uptake are not increased. FLT-PET was validated against histopathology in a variety of solid tumors including breast, lung, and sarcoma [68-70]. In soft-tissue sarcoma, Cobben *et al.* found a significant correlation between the SUVs and labeling index of the proliferation marker Ki-67. In addition, FLT-PET was able to distinguish low-grade from high-grade soft-tissue sarcomas [68]. In breast tumors, Kenny *et al.* reported a strong correlation between SUVs and the fully quantitative net irreversible plasma to tumor transfer constant (K_t) parameter of dynamic FLT-PET and staining of Ki-67 [69]. Finally, Yap *et al.* also observed a significant correlation between FLT uptake in non-small cell lung cancer lesions and the Ki-67 labeling index [70]. In primary head and neck tumors, this promising compound has thus far been applied only to primary laryngeal tumors [71]. Validation of FLT-PET in seventeen patients with a squamous cell carcinoma of the oral cavity by comparison with iododeoxyuridine (a proliferation marker) uptake and thymidine kinase 1 (TK-1, the key enzyme in FLT phosphorylation) expression has been performed by Troost *et al.* [72]. They demonstrated only a weak correlation between FLT uptake and iododeoxyuridine staining intensity, and found no correlation between FLT uptake and TK 1 staining. Troost *et al.* have also analyzed the role of FLT-PET in detecting metastatic lymph nodes in head and neck cancer patients [73]. They found a high rate of false-positive findings caused by FLT uptake in the germinal centers of reactive lymph nodes resulting in a low specificity and a low positive predictive value (17% and 38%, respectively) and concluded that FLT-PET was not suitable for detection of cervical lymph node metastases in head and neck cancer patients.

Until now, adaptive image-guided radiotherapy has been based on repetitive PET scanning using FDG [30]. As the treatment course progresses, the obtained FDG-PET signal is heavily influenced by the inflammatory response of tumor-surrounding tissues, leading to an increased background activity. As a result, segmentation of the PET signal for tumor delineation purposes becomes increasingly difficult. The use of a proliferation-specific PET tracer, such as FLT, may be a solution to this problem. During the course of therapy, the reduction in the proliferative activity of the primary tumor can be accurately imaged by FLT, not disturbed by increased tracer uptake in surrounding inflammatory tissue. Hoeben *et al.* performed consecutive FLT-PET/CT scans in 48 patients with head and neck cancer undergoing primary (chemo)radiotherapy [74]. Scans were obtained before and during the second and the fourth week of treatment, and PET parameters were correlated with outcome. They demonstrated FLT uptake in all studied cancers, a significant decrease in uptake during the first four weeks of treatment and that a greater decrease in the second week of treatment predicted a more favorable long-term outcome

[74]. Given the notion that tumor cell proliferation is a mechanism of therapy resistance, FLT-PET can provide an effective tool for decisions on early and personalized treatment adaptation, using anti-proliferative treatments such as accelerated radiotherapy or cetuximab, in patients with head and neck cancer (Fig. 5).

Perfusion, Protein Synthesis, and Others

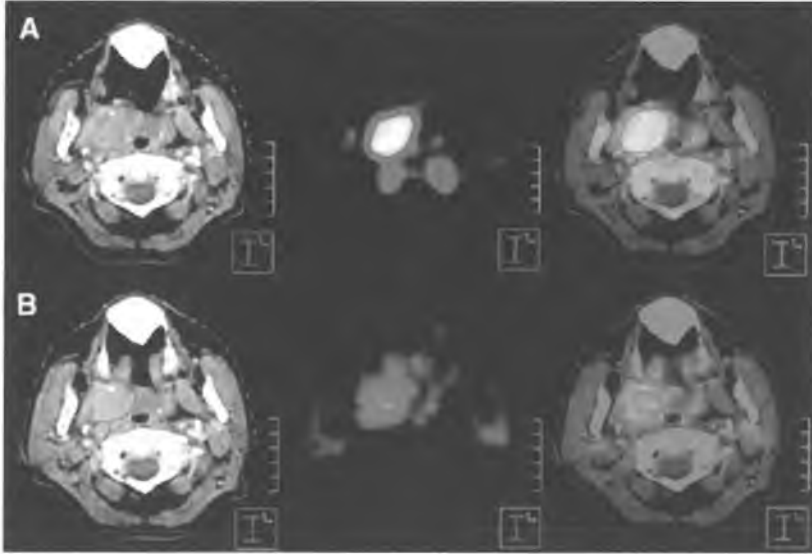
Another significant tumor characteristic strongly related to tumor cell hypoxia is tumor blood perfusion. Hypoxia is a strong stimulus for neovascularization, but many newly formed vessels are of poor quality and have severe structural and functional abnormalities. Despite increased vascular density, the impaired functionality of blood vessels may result in deprivation of oxygen and nutrients. Therefore, an imaging tool for assessment of tumor blood flow may provide important information relevant for radiotherapy responsiveness. Lethio *et al.* used ^{15}O -labeled water and FETNIM for imaging of perfusion and hypoxia in 21 patients with head and neck cancer [56]. Preliminary results from this small study indicated an association between tumor perfusion and radiation treatment outcome [57].

O-2- ^{18}F -fluoroethyl-L-tyrosine (FET) and L-methyl- ^{11}C -methionine (MET) are amino acid analogs used to visualize cellular amino acid uptake or protein synthesis. FET may be useful in differentiating tumor from posttreatment inflammatory tissue, as it is not taken up by inflammatory cells. Several studies compared FET with FDG-PET in squamous cell carcinomas of the head and neck and histopathologically confirmed the specific uptake of FET by malignant cells [75-77]. The specificity of FET-PET was found to be superior to that of FDG-PET (95% - 100% vs. 63% - 79%), but the sensitivity of the amino acid tracer was significantly lower (64% - 75% vs. 93% - 95%, respectively). Because the SUVs for FET-PET were significantly lower than those for FDG-PET, the new tracer will probably not replace FDG-PET as a diagnostic tool but can provide complementary information for discrimination between tumor and inflammatory tissue [75-77].

MET-PET is similar in sensitivity and specificity to FDG-PET [78]. In a delineation study, MET was compared with FDG-PET and CT. Although, compared with CT, FDG-PET yielded significantly smaller GTVs, GTVs based on MET-PET were not different from GTVs based on CT, probably because uptake by surrounding normal mucosa and salivary gland tissue. The authors concluded that MET has no additional value for target volume delineation in head and neck tumors [79].

1- ^{11}C -acetate is suggested to preferentially metabolize to the membrane lipids in tumor cells. In a staging and radiotherapy planning study for head and neck cancer, 1- ^{11}C -acetate PET detected all primary tumors and 95% of the metastatic lymph nodes, more than FDG-PET and CT/MRI [80]. However, the GTVs derived by 1- ^{11}C -acetate PET were 51% larger

Fig. 5 CT (left), FLT-PET (middle), and fused FLT-PET/CT (right) images of patient with T3N0M0 oropharyngeal cancer before radiation therapy (A) and after 8 fractions of 2 Gy (B). Red = GTV_{CT} . In B, significant reduction in FLT-PET signal intensity is already seen, whereas only a modest decrease of GTV_{CT} is seen at this dose level.



than those based on FDG-PET. Before ^{11}C -acetate PET can be introduced in the radiotherapy planning process, further studies are needed to explain this discrepancy and to clarify the mechanism of tumor uptake.

Finally, noninvasive methods to assess the uptake and biodistribution of biologic modifiers will be of great value to direct new targeted therapies. Radiolabeled antibodies and small molecules for PET are currently being developed and tested in preclinical and early clinical studies [81,82].

In conclusion, PET tracers that image specific biologic tumor characteristics offer potential for tailor-made radiation therapy. However, they remain in the research arena until proper clinical validation has occurred.

Technical innovations

Several challenges regarding PET scanning remain, of which some may be resolved or improved whereas others cannot. For example, resolution is limited by the distance a positron travels before it annihilates. This distance is a given fact for a certain radionuclide positron emitter and therefore unchangeable. Furthermore, various developments regarding an increase in spatial and temporal resolution are ongoing. Currently, the spatial resolution for human PET scanners is in the order of 5-7mm, compared with 1-3mm for small-animal scanners. New developments in the size of the detector crystal, the coincidence timing window, and signal processing have achieved a resolution of 2mm for human applications. These developments reduce image distortion and blurring and may increase the precision of tumor delineation.

Integrated PET/MRI scanners combine anatomic with functional imaging and may have a specific impact on the staging and treatment of head and neck cancer [83]. The potential benefits of integrated PET/CT for the planning of radiotherapy have been discussed. However, for particular subsites of the head and neck region, such as oropharyngeal and oral cavity tumors, MRI is the diagnostic imaging modality of choice. In these tumor sites, integrated PET/MRI scanners may further improve the accuracy of GTV delineation. In addition, dynamic MRI studies such as dynamic contrast-enhanced MRI and blood oxygen level dependent MRI, as well as MR spectroscopy and diffusion-weighted MRI, may add complementary functional information.

Conclusions

FDG-PET is the gold standard for noninvasive functional imaging in oncology. In head and neck tumors, FDG-PET is not recommended for detection of the primary tumor (except for unknown primary head and neck tumors), though it is recommended for nodal staging when conventional imaging is equivocal or where treatment may be significantly modified. FDG-PET may influence the treatment decision if distant metastases or second primary tumors are detected.

For radiotherapy planning in head and neck cancer, FDG-PET can provide important information complementary to CT. On the basis of PET information, sparing of normal structures and the escalation of dose to FDG-avid subvolumes is facilitated. However, additional histologic validation studies and properly designed clinical studies are needed to address the clinical value and possible shortcomings of this concept. Several PET tracers that image biologic tumor characteristics reflecting radiation resistance mechanisms are available and offer potential for tailored radiation therapy especially when PET is performed

both before and during treatment as is recently shown by the exciting publications of Zips *et al.* and Hoeben *et al.*, which were discussed earlier [50,74]. These PET tracers should be restricted to research purposes until proper clinical validation has occurred. In this context, the use of more than one tracer may open new horizons in the future. Finally, technical developments in PET scanning in general and in the field of head and neck cancer in particular may increase the precision of radiotherapy planning and thus improve tumor control and reduce treatment-related morbidity.

Currently, a European multicentre randomized fase II study, Artforce, is a good example of a clinical study trying to find reliable PET parameters in order to individualize treatment. Twofold randomization will be performed: radiation combined with cisplatin versus radiation combined with cetuximab and standard dose radiation (70Gy) versus PET-based redistribution of the dose within the gross tumor volume (64Gy in the FDG-PET negative tumor area, 84Gy in the tumor area containing the highest FDG-PET activity). Cetuximab is an epidermal growth factor receptor inhibitor which is able to counteract tumor cell proliferation. In Artforce, pretreatment ⁸⁹Zirconium labeled cetuximab-PET will be acquired and correlated with outcome.

In the coming years PET will be increasingly used in early response assessment study protocols. More robust data will be acquired, and treatment modifications on the basis of PET facilitated treatment response will be performed in the research setting.

The dream of individualized treatment for patients with head and neck cancer eligible for primary radiotherapy will become reality, and PET will most certainly play a key role.

References

- [1] Feng FY, Kim HM, Lyden TH, *et al.* Intensity modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures. *Int J Radiat Oncol Biol Phys* 2007;68:1289-1298.
- [2] Hunt MA, Zelefsky MJ, Wolden S, *et al.* Treatment planning and delivery of intensity-modulated radiation therapy for primary nasopharynx cancer. *Int J Radiat Oncol Biol Phys* 2001;49:623-632.
- [3] Kam MK, Chau RM, Suen J, *et al.* Intensity-modulated radiotherapy in nasopharyngeal carcinoma: dosimetric advantage over conventional plans and feasibility of dose escalation. *Int J Radiat Oncol Biol Phys* 2003;56:145-157.
- [4] Kam MK, Leung SF, Zee B, *et al.* Prospective randomized study of intensity-modulated radiotherapy on salivary gland function in early-stage nasopharyngeal carcinoma patients. *J Clin Oncol* 2007;25:4873-4879.
- [5] Kapanen M, Collan J, Saarilahti K, *et al.* Accuracy requirements for head and neck intensity-modulated radiation therapy based on observed dose response of the major salivary glands. *Radiat Oncol* 2009;93:109-114.
- [6] Fletcher JW, Djulibegovic B, Soares HP, *et al.* Recommendations on the use of 18F-FDG PET in oncology. *J Nucl Med* 2008;49:480-508.
- [7] Yoo J, Henderson S, Walker-Dilks C. Evidence-based guideline recommendations on the use of positron emission tomography imaging in head and neck cancer. *Clin Oncol (R Coll Radiol)* 2013;25:e33-e66.
- [8] Kyzas PA, Evangelou E, Axa-Kyza D, *et al.* 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. *J Natl Cancer Inst* 2008;21:712-720.
- [9] Gregoire V, Jeraj R, Lee JA, *et al.* Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? *Lancet Oncol* 2012;13:e292-e300.
- [10] Schinagl DAX, Kaanders JH, Oyen WJ. From anatomical to biological target volumes: the role of PET in radiation treatment planning. *Cancer Imaging* 2006;6:S107-S116.
- [11] Steenbakkers RJ, Duppen JC, Fitton I, *et al.* Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. *Int J Radiat Oncol Biol Phys* 2006;64:435-448.
- [12] Ciernik IF, Dizendorf E, Baumert BG, *et al.* Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853-863.
- [13] Riegel AC, Berson AM, Destian S, *et al.* Variability of gross tumor volume delineation in head and neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys* 2006;65:726-732.
- [14] Daisne JF, Duprez T, Weynand B, *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.
- [15] Heron DE, Andrade RS, Flickinger J, *et al.* Hybrid PET-CT simulation for radiation treatment planning in head and neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys* 2004;60:1419-1424.
- [16] Nishioka T, Shiga H, Shirato H, *et al.* Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* 2002;53:1051-1057.
- [17] Daisne JF, Sibomana M, Bol A, *et al.* Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiat Oncol* 2003;69:247-250.
- [18] Nestle U, Kremp S, Schaefer-Schuler A, *et al.* Comparison of different methods for delineation of 18F-FDG PET positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342-1348.
- [19] Paulino AC, Koshy M, Howell R, *et al.* Comparison of CT and FDG-PET defined gross tumor volume in intensity modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385-1392.
- [20] Schinagl DAX, Vogel WV, Hoffmann AL, *et al.* Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:1282-1289.
- [21] Geets X, Lee JA, Bol A, *et al.* A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging* 2007;34:1427-1438.
- [22] van Dalen JA, Hoffmann AL, Dicken V *et al.* A novel iterative method for lesion delineation and volumetric quantification with FDG-PET. *Nucl Med Commun* 2007;28:485-493.
- [23] Hatt M, Cheze JC, Descourt P, *et al.* Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys* 2010;77:301-308.

- [24] Schinagel DAX, Hoffmann AL, Vogel WV, *et al.* Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer? *Radioth Oncol* 2009;91:95-100.
- [25] Schwartz DL, Ford EC, Rajendran J, *et al.* FDG-PET/CT-guided intensity modulated head and neck radiotherapy: a pilot investigation. *Head Neck* 2005;27:478-487.
- [26] Vanderstraeten B, Duthoy W, De GW, *et al.* [18F]fluoro-deoxy-glucose positron emission tomography ([18F] FDG-PET) voxel intensity-based intensity-modulated radiation therapy (IMRT) for head and neck cancer. *Radioth Oncol* 2006;79:249-258.
- [27] Madani I, Duthoy W, Derie C, *et al.* Positron emission tomography-guided, focal-dose escalation using intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;68:126-135.
- [28] Rothschild S, Studer G, Seifert B, *et al.* PET/CT staging followed by Intensity-Modulated Radiotherapy (IMRT) improves treatment outcome of locally advanced pharyngeal carcinoma: a matched-pair comparison. *Radiat Oncol* 2007;2:22.
- [29] Vernon MR, Maheshwari M, Schultz CJ, *et al.* Clinical outcomes of patients receiving integrated PET/CT-guided radiotherapy for head and neck carcinoma. *Int J Radiat Oncol Biol Phys* 2008;70:678-684.
- [30] Geets X, Tomsej M, Lee JA, *et al.* Adaptive biological image-guided IMRT with automatic and functional imaging in pharyngo-laryngeal tumors: Impact on target volume delineation and dose distribution using helical tomography. *Radioth Oncol* 2007;85:105-115.
- [31] Brizel DM, Sibley GS, Prosnitz LR, *et al.* Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys* 1997;38:285-289.
- [32] Hockel M, Schlenger K, Mitze M, *et al.* Hypoxia and radiation response in human tumors. *Semin Radiat Oncol* 1996;6:3-9.
- [33] Thomlinson RH, Gray LH. The histological structure of some human lung cancers and the possible implications for radiotherapy. *Br J Cancer* 1955;9:539-549.
- [34] Kaanders JH, Pop LA, Marres HA, *et al.* ARCON: experience in 215 patients with advanced head and neck cancer. *Int J Radiat Oncol Biol Phys* 2002;52:769-778.
- [35] Overgaard J, Hansen HS, Overgaard M, *et al.* A randomized double-blind phase III study of nimorazole as a hypoxic radiosensitizer of primary radiotherapy in supraglottic larynx and pharynx carcinoma. Results of the Danish Head and Neck Cancer Study (DAHANCA) Protocol 5-85. *Radioth Oncol* 1998;46:135-146.
- [36] Bussink J, Kaanders JH, van der Kogel AJ. Tumor hypoxia at the micro-regional level: clinical relevance and predictive value of exogenous and endogenous hypoxic cell markers. *Radioth Oncol* 2003;67:3-15.
- [37] Kaanders JH, Wiffels KJ, Marres HA, *et al.* Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer. *Cancer Res* 2002;62:7066-7074.
- [38] Nordsmark M, Loncaster J, Chou SC, *et al.* Invasive oxygen measurements and pimonidazole labeling in human cervix carcinoma. *Int J Radiat Oncol Biol Phys* 2001;49:581-586.
- [39] Koh WJ, Rasey JS, Evans ML, *et al.* Imaging of hypoxia in human tumors with [F-18]fluoromisonidazole. *Int J Radiat Oncol Biol Phys* 1992;22:199-212.
- [40] Rasey JS, Koh WJ, Evans ML, *et al.* Quantifying regional hypoxia in human tumors with positron emission tomography of [18F]fluoromisonidazole: a pretherapy study of 37 patients. *Int J Radiat Oncol Biol Phys* 1996;36:417-428.
- [41] Valk PE, Mathis CA, Prados MD, *et al.* Hypoxia in human gliomas: demonstration by PET with fluorine-18-fluoromisonidazole. *J Nucl Med* 33;2133-2137.
- [42] Eschmann SM, Paulsen F, Reimold M, *et al.* Prognostic impact of hypoxia imaging with 18F-misonidazole PET in non-small cell lung cancer and head and neck cancer before radiotherapy. *J Nucl Med* 2005;46:253-260.
- [43] Gagel B, Reinartz P, Dimartino E, *et al.* pO₂ polarography versus positron emission tomography ([18F]-2-fluoro-2'-deoxyglucose). An appraisal of radiotherapeutically relevant hypoxia. *Strahlenther Onkol* 2004;180:616-622.
- [44] Hicks RJ, Rishin D, Fisher R, *et al.* Utility of FMISO PET in advanced head and neck cancer treated with chemoradiation incorporating a hypoxia-targeting chemotherapy agent. *Eur J Nucl Med Mol Imaging* 2005;32:1384-1391.
- [45] Rajendran JG, Mankoff DA, O'Sullivan F, *et al.* Hypoxia and glucose metabolism in malignant tumors: evaluation by [18F]fluoromisonidazole and [18F]fluorodeoxyglucose positron emission tomography imaging. *Clin Cancer Res* 2004;10:2245-2252.

- [46] Rishin D, Hicks RJ, Fisher Rea. Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a substudy of Trans-Tasman radiation Oncology Group Study 98.02. *J Clin Oncol* 2006;24:2098-2104.
- [47] Thorwarth D, Eschmann SM, Paulsen F, *et al.* A kinetic model for dynamic [18F]-Fmiso PET data to analyse tumor hypoxia. *Phys Med Biol* 2005;50:2209-2224.
- [48] Thorwarth D, Eschmann SM, Holzner F, *et al.* Combined uptake of [18F]FDG and [18F]FMISO correlates with radiation therapy outcome in head and neck cancer patients. *Radiother Oncol* 2006;80:151-156.
- [49] Thorwarth D, Eschmann SM, Scheiderbauer J, *et al.* Kinetic analysis of dynamic 18F-fluoromisonidazole PET correlates with radiation treatment outcome in head and neck cancer. *BMC Cancer* 2005;5:152.
- [50] Zips D, Zophel K, Abolmaali N, *et al.* Exploratory prospective trial of hypoxia-specific PET imaging during radio-chemotherapy in patients with locally advanced head and neck cancer. *Radiother Oncol* 2012;105:21-28.
- [51] Lee NY, Mechalakos JG, Nehmeh S, *et al.* Fluorine-18-labeled fluoromisonidazole positron emission and computed tomography-guided intensity-modulated radiotherapy for head and neck cancer: a feasibility study. *Int J Radiat Oncol Biol Phys* 2008;70:2-13.
- [52] Rajendran JG, Hendrickson KR, Spence AM, *et al.* Hypoxia imaging-directed radiation treatment planning. *Eur J Nucl Med Mol Imaging* 2006;33 Suppl1:44-53.
- [53] Thorwarth D, Eschmann SM, Paulsen F, *et al.* Hypoxia dose painting by numbers: a planning study. *Int J Radiat Oncol Biol Phys* 2007;68:291-300.
- [54] Lewis JS, McCarthy DW, McCarthy TJ, *et al.* Evaluation of 64Cu-ATSM in vitro and in vivo in a hypoxic tumor model. *J Nucl Med* 1999;40:177-183.
- [55] Chao KS, Bosch WR, Mutic S, *et al.* A novel approach to overcome hypoxic tumor resistance: Cu-ATSM-guided intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys* 2001;49:1171-1182.
- [56] Lethio K, Oikonen V, Gronroos T, *et al.* Imaging of blood flow and hypoxia in head and neck cancer: initial evaluation with [(15)O]H(2)O and [(180F)]fluoroerythronitroimidazole PET. *J Nucl Med* 2001;42:1643-1652.
- [57] Lethio K, Eskola O, Viljanen T, *et al.* Imaging perfusion and hypoxia with PET to predict radiotherapy response in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2004;59:971-982.
- [58] Kumar P, Wiebe LJ, Mannan RH, *et al.* [99mTc]Technetium labeled PnAo-azomycin glucuronides: a novel class of imaging markers of tissue hypoxia. *Appl Radiat Isot* 2002;57:719-728.
- [59] Souvatzoglou M, Grosu AL, Roper B, *et al.* Tumor hypoxia imaging with [18F]FAZA PET in head and neck cancer patients: a pilot study. *Eur J Nucl Med Mol Imaging* 2007;34:1566-1575.
- [60] Grosu AL, Souvatzoglou M, Roper B, *et al.* Hypoxia imaging with FAZA-PET and theoretical considerations with regard to dose painting for individualization of radiotherapy in patients with head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:541-551.
- [61] Mahy P, Geets X, Lonnew M, *et al.* Determination of tumor hypoxia with [18F]EF3 in patients with head and neck tumours: a phase I study to assess the tracer pharmacokinetics, biodistribution and metabolism. *Eur J Nucl Med Mol Imaging* 2008;35:1282-1289.
- [62] Mortensen LS, Johansen J, Kallehauge L, *et al.* FAZA PET/CT hypoxia imaging in patients with squamous cell carcinoma of the head and neck treated with radiotherapy: results from the DAHANCA 24 trial. *Radiother Oncol* 2012;105:14-20.
- [63] Troost EG, Laverman P, Kaanders JH, *et al.* Imaging hypoxia after oxygenation-modification: comparing [18F] FMISO autoradiography with pimonidazole immunohistochemistry in human xenograft tumors. *Radiother Oncol* 2006;80:157-164.
- [64] Troost EG, Laverman P, Philippens ME, *et al.* Correlation of [18F]FMISO autoradiography and pimonidazole [corrected] immunohistochemistry in human head and neck carcinoma xenografts. *Eur J Nucl Med Mol Imaging* 2008;35:1803-1811.
- [65] Nehmeh SA, Lee NY, Schroder H, *et al.* Reproducibility of intratumor distribution of (18)F-fluoromisonidazole in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2008;70:235-242.
- [66] Kim JJ, Tannock IF. Repopulation of cancer cells during therapy: an important cause of treatment failure. *Nat Rev Cancer* 2005;5:516-525.
- [67] Shields AF, Grierson JR, Dohmen BM, *et al.* Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. *Nat Med* 1998;4:1334-1336.

- [68] Cobben DC, Elsinga PH, Suurmeijer AJ, *et al.* Detection and grading of soft tissue sarcomas of the extremities with (18)F-3'-fluoro-3'-deoxy-L-thymidine. *Clin Cancer Res* 2004;10:1685-1690.
- [69] Kenny LM, Vigushin DM, Al-Nahhas A, *et al.* Quantification of cellular proliferation in tumor and normal tissues of patients with breast cancer by [18F]fluorothymidine-positron emission tomography imaging: evaluation of analytical methods. *Cancer Res* 2005;65:10104-10112.
- [70] Yap CS, Czernin J, Fishbein MC, *et al.* Evaluation of thoracic tumors with 18F-fluorothymidine and 18F-fluoro-deoxyglucose-positron emission tomography. *Chest* 2006;129:393-401.
- [71] Cobben DC, van der Laan BF, Maas B, *et al.* 18F-FLT PET for visualization of laryngeal cancer: comparison with 18F-FDG PET. *J Nucl Med* 2004;45:226-231.
- [72] Troost EG, Bussink J, Slootweg PJ, *et al.* Histopathologic validation of 3'-deoxy-3'-18F-fluorothymidine PET in squamous cell carcinoma of the oral cavity. *J Nucl Med* 2010;51:713-719.
- [73] Troost EG, Vogel WV, Merks MA, *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;48:726-735.
- [74] Hoeben BA, Troost EG, Span PN, *et al.* 18F-FLT PET during radiotherapy of chemoradiotherapy in head and neck squamous cell carcinoma is an early predictor of outcome. *J Nucl Med* 2013;54:532-540.
- [75] Balogova S, Perie S, Kerrou K, *et al.* Prospective comparison of FDG and FET PET/CT in patients with head and neck squamous cell carcinoma. *Mol Imaging Biol* 2008;10:364-373.
- [76] Pauleit D, Stoffels G, Schaden W, *et al.* PET with O-(2-18F-fluoroethyl)-L-Tyrosine in peripheral tumors: first clinical results. *J Nucl Med* 2005;46:411-416.
- [77] Pauleit D, Zimmermann A, Stoffels G, *et al.* 18F-FET PET compared with 18F-FDG PET and CT in patients with head and neck cancer. *J Nucl Med* 2006;47:256-261.
- [78] Leskinen-Kallio S, Lindholm P, Lapela M, *et al.* Imaging of head and neck tumors with positron emission tomography and [11C]methionine. *Int J Radiat Oncol Biol Phys* 1994;30:1195-1199.
- [79] Geets X, Daisne JF, Gregoire V, *et al.* Role of 11-C-methionine positron emission tomography for the delineation of the tumor in pharyngo-laryngeal squamous cell carcinoma: comparison with FDG-PET and CT. *Radiother Oncol* 2004;71:267-273.
- [80] Sun A, Sorensen J, Karlsson M, *et al.* 1-[11C]-acetate PET imaging in head and neck cancer - a comparison with 18F-FDG-PET: implications for staging and radiotherapy planning. *Eur J Nucl Med Mol Imaging* 2007;34:651-657.
- [81] Cai W, Chen K, He L, *et al.* Quantitative PET of EGFR expression in xenograft-bearing mice using 64Cu-labeled cetuximab, a chimeric anti-EGFR monoclonal antibody. *Eur J Nucl Med Mol Imaging* 2007;34:850-858.
- [82] Wang JQ, Gao M, Miller KD, *et al.* Synthesis of [11C]Iressa as a new potential PET cancer imaging agent for epidermal growth factor receptor tyrosine kinase. *Bioorg Med Chem Lett* 2006;16:4102-4106.
- [83] Judenhofer MS, Wehrli HF, Newport DF, *et al.* Simultaneous PET-MRI: a new approach for functional and morphological imaging. *Nat Med* 2008;14:459-465.

9 |

Summary

Summary

Progress in radiation oncology requires a re-evaluation of the methods of target volume delineation beyond anatomical localization. New molecular imaging techniques for tumor visualization such as positron emission tomography (PET) provide insight into tumor characteristics and can be complementary to the anatomical data of computed tomography (CT) or magnetic resonance imaging. In this thesis, the role of PET in radiation treatment planning of patients with head and neck carcinoma is studied.

In **Chapter 2**, the results of a literature review are presented in which three issues are discussed: First, can PET identify a tumor more accurately? Second, can biological tumor characteristics be visualized? Third, can intratumoral heterogeneity of these characteristics be identified?

Chapter 3 deals with the selection and validation of the optimal method for software fusion of dedicated PET and CT. In this study, fifteen patients with head and neck cancer underwent separate CT and ^{18}F -fluorodeoxyglucose-PET (FDG-PET), both in a custom-moulded rigid mask fitted with four multimodality fiducial markers. Five image registration methods were applied, and the error of each method was determined by evaluation of the markers. The operator-independent iterative closest point method performed best and its accuracy permits implementation of dedicated PET images in CT-based radiation therapy planning.

Implementing PET-information in the definition of the gross tumor volume (GTV) of the primary tumor of head and neck cancer patients eligible for primary (chemo)radiotherapy is not straightforward. In **Chapter 4**, CT and physical examination-based primary tumor delineation is compared with five methods of PET-based tumor delineation: visual interpretation PET_{VIS} , applying fixed thresholds at a standardized uptake value (SUV) of 2.5 and at 40% and 50% of the maximum signal intensity of the primary tumor ($\text{PET}_{2.5}$, $\text{PET}_{40\%}$, $\text{PET}_{50\%}$) and applying a variable threshold based on the signal-to-background ratio (PET_{SBR}). The primary tumor of 78 patients with stages II-IV squamous cell carcinoma of the head and neck area, who underwent co-registered CT and FDG-PET, was delineated on CT and five PET-based GTVs were obtained. Absolute volumes were compared and overlap analyses were performed. $\text{PET}_{2.5}$ failed to provide successful delineation in almost half of cases. For the other PET delineation tools, volume and shape of the GTV were influenced heavily by the choice of tool. On average, all threshold-based PET-GTVs were smaller than on CT. Nevertheless, PET frequently detected significant tumor extension outside the GTV delineated on CT (15 – 34% of PET volume).

The question “can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head and neck cancer?” is evaluated in **Chapter 5**. The seventy-eight co-registered CT- and FDG-PET scans as used in **Chapter 4** were studied again. Cervical lymph nodes were classified as “enlarged” if the shortest axial diameter on CT was ≥ 10 mm, and as “marginally enlarged” if it was 7-10 mm. Subsequently, lymph nodes were assessed on FDG-PET by applying eight segmentation methods: PET_{VIS}, PET_{2.5}, PET_{40%}, PET_{50%}, PET_{SBR}, and PET_{40%N}, PET_{50%N}, PET_{SBRN}. The latter three methods use the maximum signal of the lymph node (instead of the primary tumor) as the threshold reference. Of 108 nodes classified as “enlarged” on CT, 75% were also identified by PET_{VIS}, 59% by PET_{40%}, 43% by PET_{50%} and 43% by PET_{SBR}. Of 100 nodes classified as “marginally enlarged”, only a minority were visualized by FDG-PET. The respective numbers were 26%, 10%, 7% and 8% for PET_{VIS}, PET_{40%}, PET_{50%}, and PET_{SBR}. PET_{40%N}, PET_{50%N}, and PET_{SBRN}, respectively, identified 66%, 82% and 96% of the PET_{VIS}-positive nodes. The conclusions of this chapter were that many nodes that are enlarged and considered metastatic by standard CT-based criteria appear to be negative on FDG-PET scan. Alternately, a small proportion of marginally enlarged nodes are positive on FDG-PET scan. However, the results are largely dependent on the PET segmentation tool used, and until proper validation FDG-PET is not recommended for target volume definition of metastatic lymph nodes in routine clinical practice.

In head and neck cancer various treatment strategies have been developed to improve outcome, but selecting patients for these intensified treatments remains difficult. There are indications that pretreatment tumor FDG uptake in head and neck cancer may be an independent prognostic factor. In **Chapter 6**, the prognostic value of FDG uptake and primary tumor volume measurements (both CT-based and FDG-PET-based) were assessed in 77 patients with squamous cell carcinoma of the head and neck area treated with primary (chemo)radiotherapy. GTV of the primary tumor was determined on CT (GTV_{CT}) and FDG-PET scans. Five segmentation methods were applied: PET_{VIS}, PET_{2.5}, PET_{40%}, PET_{50%}, and PET_{SBR}. The maximum intratumoral FDG activity (SUV_{MAX}) was recorded. The mean FDG uptake for each PET-based volume was also recorded (SUV_{mean}). Subsequently, to determine the metabolic volume, the integrated SUV was calculated as the product of PET-based volume and SUV_{mean}. All these variables were analyzed as potential predictors of local control (LC), regional recurrence-free survival (RRFS), distant metastasis-free survival (DMFS), disease-free survival (DFS) and overall survival (OS). In oral cavity / oropharynx tumors PET_{VIS} was the only volume-based method able to predict LC. Both PET_{VIS} and GTV_{CT} were able to predict DMFS, DFS and OS in these subsites. In oral cavity / oropharynx tumors integrated SUVs were associated with LC, DMFS, DFS and OS, while SUV_{max} and SUV_{mean} were not. In hypopharyngeal / laryngeal tumors none of the variables was associated with outcome. The conclusion of chapter 6 is that there is no role for pretreatment FDG-PET as a predictor of (chemo)radiotherapy outcome in head and neck

cancer in the daily routine. However, this potential application needs further exploration, focusing on FDG-PET-based primary tumor volume, integrated SUV and SUV_{MAX} of the primary tumor.

There are limited data on the accuracy of radiotherapy target volume delineation by FDG-PET. Therefore, in **Chapter 7**, a validation study was performed by comparing FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer with the pathological standard. Twelve head and neck cancer patients eligible for therapeutic neck dissection underwent preoperative FDG-PET/CT. Twenty-eight metastatic lymph nodes were delineated on CT ($Node_{CT}$) and ten PET segmentation tools were used to assess FDG-PET-based nodal volumes: PET_{VIS} , $PET_{2.5}$, $PET_{40\%}$, $PET_{50\%}$, PET_{SBR} , and an iterative background-subtracted relative threshold level (PET_{RTL}). The latter four tools were applied with the primary tumor as reference and also with the lymph node itself as reference. Nodal volumes were compared with the true volumes as determined by pathological examination. Both $Node_{CT}$ and PET_{VIS} showed good correlations with the pathological volume. PET segmentation tools using the metastatic node as reference all performed well, but not better than PET_{VIS} . The tools using the primary tumor as reference correlated poorly with pathology. The conclusions of **Chapter 7** are: FDG-PET accurately estimates metastatic lymph node volume, but beyond the detection of lymph node metastases (staging), it has no added value over CT alone for the delineation of routine radiotherapy target volumes. If FDG-PET is used in radiotherapy planning, treatment adaptation or response assessment, we recommend an automated segmentation method for purposes of reproducibility and inter-institutional comparison.

In conclusion, integration of PET-information into computer tomography based radiation treatment planning of head and neck cancer patients allows consideration of biological tumor characteristics in the treatment. In this thesis several image registration method was evaluated. Furthermore, several FDG-PET segmentation methods were analyzed for the purpose of delineating the primary tumor and also the metastatic lymph nodes. The role of pretreatment FDG-PET as an outcome predictor of primary (chemo)radiotherapy was prospectively investigated. Finally, the accuracy of FDG-PET was studied by comparing FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer with the true volume using the pathological standard.

10 |

Summary in Dutch *(Nederlandse Samenvatting)*

Dankwoord

List of publications

Curriculum Vitae

Summary in Dutch (*Nederlandse Samenvatting*)

Technologische vooruitgang binnen de radiotherapie vereist een herevaluatie van de methoden van doelvolumebepaling. Nieuwe moleculaire beeldvormende technieken zoals positron emissie tomografie (PET) kunnen inzicht geven in tumoreigenschappen. PET is een techniek waarmee radioactieve stoffen binnen in het menselijk lichaam kunnen worden afgebeeld, gevolgd en gekwantificeerd. Met de meest gebruikte radioactieve stof, fluor-18 (^{18}F) gekoppeld aan een variant van suiker (FDG), kan het energieverbruik (metabolisme) van de verschillende organen en weefsels binnen een patiënt in beeld worden gebracht. PET kan complementair zijn aan anatomische data die gegenereerd worden door computer tomografie (CT) of magneet resonantie imaging (MRI). In dit proefschrift wordt de rol van PET in de bestralingsbehandeling van patiënten met een kwaadaardige tumor uitgaande van het hoofd-hals gebied onderzocht.

In **hoofdstuk 2** worden de resultaten van een literatuur overzicht gepresenteerd, waarbij de nadruk ligt op de volgende drie onderwerpen: Ten eerste, is PET beter in staat een kwaadaardige tumor te herkennen? Ten tweede, kunnen biologische tumoreigenschappen gevisualiseerd worden? Ten derde, kan intratumorale heterogeniteit van deze eigenschappen worden geïdentificeerd?

Hoofdstuk 3 beschrijft de selectie en validatie van verschillende methoden van software-gebaseerde PET en CT beeldfusie. In dit onderzoek ondergingen vijftien patiënten met een kwaadaardige tumor uitgaande van het hoofd-hals gebied zowel een CT als een ^{18}F -fluorodeoxyglucose-PET (FDG-PET). Beide scans werden verricht in een bestralingsmasker en op de masker waren vier markeringen bevestigd die zowel met CT als met FDG-PET afgebeeld konden worden. Vijf beeldregistratie methoden werden toegepast, waarbij de nauwkeurigheid van elke methode werd bepaald door de afwijkingen ter plaatse van de markeringen te meten. De meest geschikt methode kon de onnauwkeurigheid in de beeldfusie terugbrengen tot slechts 3 mm. Daarmee werd het combineren van PET en CT gevalideerd voor toepassing bij geavanceerde 3-dimensionale intensiteits-gemoduleerde bestraling van kanker in het hoofd-hals gebied.

Het gebruiken van PET-informatie bij het bepalen van het tumorvolume (gross tumor volume = GTV) van de primaire tumor van patiënten met kanker uitgaande van het hoofd-hals gebied die in aanmerking komen voor primaire (chemo)radiotherapie is niet vanzelfsprekend. In **hoofdstuk 4** werd tumorvolume-bepaling (delineatie) van de primaire tumor op basis van CT en lichamelijk onderzoek vergeleken met vijf methoden van PET-gebaseerde tumordelineatie: visuele interpretatie PET_{vis}, toepassen van een vaste drempelwaarde gebaseerd op een standaard opname hoeveelheid (standardized uptake value = SUV) van 2.5 en drempels van 40% en 50% van de maximale signaalsterkte ter

plaats van de primaire tumor (PET_{25} , $PET_{40\%}$, $PET_{50\%}$) en toepassen van een variabele drempelwaarde gebaseerd op de signaal-achtergrond verhouding (signal-to-background ratio = $SBR(PET_{SBR})$). De primaire tumor van 78 patiënten met stadium II-IV plaveiselcelcarcinoom van het hoofd-hals gebied die geregistreerde CT en FDG-PET ondergingen werd ingetekend op CT en vijf PET-gebaseerde GTV's werden verkregen. Absolute volumina werden vergeleken en overlap analyses werden verricht. PET_{25} gaf slechts bij iets meer dan de helft van de tumoren een succesvolle delineatie. Bij de overige PET-gebaseerde intekenmethoden werd vastgesteld dat de vorm en het volume van de GTV sterk van elkaar konden verschillen. Over het algemeen waren alle drempel-gebaseerde PET-GTV's kleiner dan de GTV's zoals op CT bepaald. Desalniettemin, kon PET frequent tumoruitbreiding vaststellen buiten het GTV zoals op CT bepaald (15% - 34% van het PET volume).

In **hoofdstuk 5** werd onderzocht of FDG-PET behulpzaam was bij het intekenen van lymfklieruitzaaiingen bij plaveiselcelcarcinoom van het hoofd-hals gebied. De 78 geregistreerde CT- en FDG-PET scans, beschreven in **hoofdstuk 4**, werden opnieuw gebruikt. Halslymfklieren werden geclassificeerd als "vergroot" indien de kortste axiale diameter op CT ≥ 10 mm was, en als "marginaal vergroot" bij een kortste axiale diameter 7 – 10 mm. Vervolgens werden de halslymfklieren beoordeeld door FDG-PET door middel van acht verschillende segmentatiemethoden: PET_{VIS} , PET_{25} , $PET_{40\%}$, $PET_{50\%}$, PET_{SBR} , $PET_{40\%N}$, $PET_{50\%N}$, PET_{SBRN} . De drie laatstgenoemde methoden gebruiken het maximale PET-sigitaal van de klier (in plaats van de primaire tumor) als drempelreferentie. Van de 108 klieren die als "vergroot" geclassificeerd waren, werd 75% ook geïdentificeerd door PET_{VIS} , 59% door $PET_{40\%}$, 43% door $PET_{50\%}$ en 43% door PET_{SBR} . Van de 100 klieren die geclassificeerd waren als "marginaal vergroot" werd slechts een minderheid gevisualiseerd door FDG-PET. Respectievelijk, 26%, 10%, 7% en 8% voor PET_{VIS} , $PET_{40\%}$, $PET_{50\%}$ en PET_{SBR} . $PET_{40\%N}$, $PET_{50\%N}$ en PET_{SBRN} identificeerden respectievelijk 66%, 82% en 96% van de PET_{VIS} -positieve lymfklieren. Geconcludeerd werd, dat veel klieren die vergroot waren en op basis van standaard CT-criteria metastatische ziekte zouden bevatten, negatief waren op de FDG-PET scan. Daarbij werd tevens vastgesteld, dat een klein deel van de marginaal vergrootte klieren positief waren op de FDG-PET scan. De resultaten waren echter erg afhankelijk van de PET-segmentatiemethode.

Er zijn diverse behandelstrategieën ontwikkeld die bij patiënten met plaveiselcelcarcinoom van het hoofd-hals gebied de genezingskans vergroten, het blijft echter moeilijk om de juiste patiënten te selecteren voor deze intensieve therapieën. Er waren aanwijzingen dat de FDG-opname van de primaire tumor voor start van de behandeling een onafhankelijke prognostische factor zou kunnen zijn. De prognostische waarde van FDG-opname en van volumemetingen (zowel op basis van CT als op basis van FDG-PET) van de primaire tumor van 77 patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied die primaire (chemo)radiotherapie ondergingen werd onderzocht. Het GTV van

de primaire tumor werd bepaald op CT (GTV_{CT}) en op FDG-PET scans. Vijf segmentatiemethoden werden gebruikt: PET_{VIS} , $PET_{2.5}$, $PET_{40\%}$, $PET_{50\%}$ en PET_{SBR} . De maximale intratumorale FDG-activiteit (SUV_{MAX}) werd vastgesteld. Tevens werd de gemiddelde FDG-opname binnen elk PET-gebaseerde GTV vastgesteld (SUV_{mean}). Vervolgens werd om het metabole volume te bepalen, de geïntegreerde SUV berekend: dit is het product van het PET-gebaseerde volume en de corresponderende SUV_{mean} . Al deze variabelen werden onderzocht als potentiële voorspellers van lokale controle (LC), regionale ziekte-vrije overleving (RRFS), afstandsmetastasen-vrije overleving (DMFS), ziekte-vrije overleving (DFS) en totale overleving (OS). Bij tumoren uitgaande van cavum oris / oropharynx was PET_{VIS} de enige volume-gebaseerde methode die de LC kon voorspellen. Zowel PET_{VIS} als GTV_{CT} konden bij deze tumorsites de DMFS, DFS en OS voorspellen. Bij tumoren uitgaande van cavum oris / oropharynx waren de geïntegreerde SUV's geassocieerd met LC, DMFS, DFS en OS, terwijl SUV_{MAX} en SUV_{mean} niet geassocieerd waren met deze uitkomstmaten. Bij tumoren uitgaande van hypopharynx / larynx waren geen van de onderzochte variabelen gerelateerd aan behandeluitkomst. De conclusie van hoofdstuk 6 was dat FDG-PET voorafgaand aan de therapie thans geen rol heeft als voorspeller van therapie-succes bij patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied die primaire (chemo)radiotherapie. Echter, deze rol vereist verder onderzoek, met als focus de variabelen: PET-gebaseerd tumorvolume, geïntegreerde SUV en SUV_{MAX} van de primaire tumor.

Er zijn weinig gegevens over de accuraatheid van FDG-PET bij radiotherapie doelvolumebepaling. Derhalve werd in **hoofdstuk 7** een validatiestudie verricht door FDG-PET gebaseerde volumina van halsklieruitzaaiingen bij patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied te vergelijken de volumina verkregen bij weefselonderzoek. Twaalf patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied die in aanmerking kwamen voor een therapeutische halsklierdissectie ondergingen een preoperatieve FDG-PET/CT scan. Achtentwintig halsklieruitzaaiingen werden ingetekend op CT ($Node_{CT}$) en tien PET segmentatiemethoden werden gebruikt om PET-gebaseerde kliervolume te bepalen: PET_{VIS} , $PET_{2.5}$, $PET_{40\%}$, $PET_{50\%}$, PET_{SBR} en een iteratieve 'background-subtracted' relatieve drempel waarde (PET_{RTL}). De vier laatstgenoemde methoden werden toegepast met zowel de primaire tumor als de te segmenteren klieruitzaaiing als referentie. De verkregen volumina werden vergeleken met de echte volumina zoals gemeten patholoog-anatomisch onderzoek. Zowel $Node_{CT}$ als PET_{VIS} toonden goede correlatie met het echte volume. De PET-segmentatiemethoden die de te segmenteren klier als referentie gebruikten presteerden goed, maar niet beter dan PET_{VIS} . De methoden die de primaire tumor als referentie gebruikten, toonden een zwakke correlatie met het echte volume. Er werd geconcludeerd dat FDG-PET nauwkeurig het volume van een lymfklieruitzaaiing kan bepalen, maar dat het geen meerwaarde heeft boven de CT bij doelvolumebepaling voor radiotherapie daar $Node_{CT}$ een goede correlatie toonde met het echte volume.

Indien FDG-PET gebruikt wordt voor radiotherapie planning, behandeladaptatie, of response evaluatie, adviseren wij een automatische segmentatiemethode omwille van de reproduceerbaarheid en de inter-institutionele vergelijkbaarheid.

Het toevoegen van PET-informatie aan CT-gebaseerde bestralingsplanning van patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied biedt de mogelijkheid om rekening te houden met de biologische tumoreigenschappen. In dit proefschrift werden diverse beeldfusiemethoden beoordeeld. Verder werden verschillende FDG-PET segmentatiemethoden onderzocht in hun rol bij het bepalen van het bestralingsvolume van zowel de primaire tumor als de lymfklieruitzaaiingen. De betekenis van FDG-PET voorafgaand aan de behandeling als voorspeller van behandeluitkomst na primaire (chemo)radiotherapie werd prospectief onderzocht. Tenslotte werd de accuraatheid van FDG-PET bestudeerd door FDG-PET gebaseerde volumina van halsklieruitzaaiingen bij patiënten met een plaveiselcelcarcinoom van het hoofd-hals gebied te vergelijken de volumina verkregen bij weefselonderzoek.

Dankwoord

Graag wil ik als eerste mijn twee promotoren bedanken.

Beste Hans, je bent als wetenschapper en als arts een voorbeeld voor mij geweest.

Beste Wim, dank voor je uitstekende begeleiding.

Aswin, Esther, Frank, Jorn, Paul, Piet, Suzan, Thijs en Wouter, allen co-auteur van dit proefschrift, wil ik bedanken voor de fijne samenwerking tijdens de afgelopen jaren.

De patiënten die deelgenomen hebben aan dit onderzoek wil ik danken voor hun inzet en vertrouwen.

De medewerkers (past and present) van de afdeling radiotherapie van het UMC St Radboud wil ik danken voor hun aandeel in mijn werkplezier.

Dank aan de medewerkers van de afdelingen nucleaire geneeskunde, keel- neus en oor-heelkunde, mondkaa- en aangezichtschirurgie, en pathologie van het UMC St Radboud, voor de inspanningen die nodig waren voor dit proefschrift.

Leden van de manuscriptcommissie, beste Pieter, Bram en Dirk, hartelijk dank voor de nauwkeurige beoordeling van het manuscript.

Ik dank alle collega's in UMC St Radboud, Canisius Wilhelmina Ziekenhuis, Maasziekenhuis Boxmeer, Bernhoven Ziekenhuis Uden, en Rivierenland Ziekenhuis Tiel met wie ik gezamenlijke patiëntenzorg verricht heb, in het bijzonder de leden van de mammacarcinoom- en long-carcinoomketens in UMC St Radboud en Canisius Wilhelmina Ziekenhuis. Bij iedere collega was het verlenen van optimale zorg het primaire eindpunt.

Als laatste enkele woorden tot mijn familie, mijn dierbare vrienden, en Samuel, Simon en Carla: dank voor jullie belangstelling, kameraadschap en liefde.

List of publications

Meijer HJM, Dorresteijn LDA, **Schinagl DAX**, van Laarhoven HWM. Recurrent stroke after low dose whole brain radiotherapy for brain metastases of breast cancer. *IJCR*; *accepted for publication*

Schinagl DAX, Span PN, Slootweg P, van den Hoogen FJ, Merks MA, Oyen WJG, Kaanders JHAM. Pathology-based validation of FDG-PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer. *Eur J Nucl Med Mol Imaging*; *accepted for publication*

Arens AI, Troost EGC, **Schinagl DAX**, Kaanders JHAM, Oyen WJG. FDG-PET/CT in radiation treatment planning of head and neck squamous cell carcinoma. *Q J Nucl Med Mol Imaging* 2011;55:512-528.

Schinagl DAX, Span PN, Oyen WJG, Kaanders JHAM. Can FDG-PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. *Eur J Nucl Med Mol Imag* 2011; 38:1449-1458.

Mandigers CM, van de Warrenburg BP, Strobbe LJ, Kluijdt I, Molenaar AH, **Schinagl DAX**. Ataxia telangiectasia: The consequences of a delayed diagnosis. *Radiother Oncol* 2011;99:97-98.

van Wely BJ, Teerenstra S, **Schinagl DAX**, Aufenacker TJ, de Wilt JH, Strobbe LJ. Systematic review of the effect of external beam radiation therapy to the breast on axillary recurrence after negative sentinel lymph node biopsy. *Br J Surg* 2011;98:326-333.

Troost EGC, **Schinagl DAX**, Bussink J, Oyen WJG, Kaanders JHAM. Clinical evidence on PET-CT for radiation therapy planning in head and neck tumours. *Radiother Oncol* 2010;96:328-334.

Schinagl DAX, van Tienhoven G. Breast Cancer: disentangling the intricate web. *J Clin Oncol* 2010;28:281-283.

Schinagl DAX, Marres HA, Kappelle AC, Merks MA, Pop LA, Verstappen SMM, Kaanders JHAM. External beam radiotherapy with endocavitary boost for nasopharyngeal cancer: treatment results and late toxicity after extended follow-up. *Int J Radiat Oncol Biol Phys* 2010;78:989-695.

Troost EGC, **Schinagl DAX**, Bussink J, Kaanders JHAM, Oyen WJG, Boerman OC. Correlation of segmented metabolic tumor volume with outcome. *Clin Cancer Res* 2010;16:1968-1969.

Troost EGC, **Schinagl DAX**, Bussink J, Boerman OC, van der Kogel AJ, Oyen WJG, Kaanders JHAM. Innovations in radiotherapy planning of head and neck cancers: role of PET. *J Nucl Med* 2010;51:66-76.

Jones HA, Antonini N, Hart AA, Peterse JL, Horiot JC, Collin F, Poortmans PM, Oei SB, Collette L, Struikmans H, van den Bogaert WF, Fourquet A, Jager JJ, **Schinagl DAX**, Wårlám-Rodenhuis CC, Bartelink H. Impact of pathological characteristics on local relapse after breast-conserving therapy: a subgroup analysis of the EORTC boost versus no boost trial. *J Clin Oncol* 2009;27:4939-4947.

Schinagl DAX, Troost EGC, Kaanders JHAM. Combined 18F-FDG-PET/CT imaging in radiotherapy target delineation for head and neck cancers: in regard to Guido et al. (*Int J Radiat Oncol Biol Phys* 2009;73:759-763). *Int J Radiat Oncol Biol Phys* 2009;74:1629

Schinagl DAX, Hoffmann AL, Vogel WV, van Dalen JA, Verstappen SMM, Oyen WJG, Kaanders JHAM. Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head and neck cancer? *Radiother Oncol* 2009;91:95-100

Schinagl DAX, Venderink DJ, Strobbe LJ. Occult breast cancer discovered following breast reduction-letter. *Ned Tijdschr Geneeskd.* 2009;153:B314 (Dutch).

Vogel WV, **Schinagl DAX**, van Dalen JA, Kaanders JHAM, Oyen WJG. Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. *Q J Nucl Med Mol Imaging* 2008;52:74-83

Schinagl DAX, Vogel WV, Hoffmann AL, van Dalen JA, Oyen WJG, Kaanders JHAM. Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography based target volume definition in head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007;69:1282-1289

Schinagl DAX, Kaanders JHAM, Oyen WJG. From anatomical to biological target volumes: the role of PET in radiation treatment planning. *Cancer Imaging* 2006; 6:S107-116

Vogel WV, van Dalen JA, **Schinagl DAX**, Kaanders JHAM, Huisman H, Corstens FH, Oyen WJG. Correction of an image size difference between positron emission tomography (PET) and computed tomography (CT) improves image fusion of dedicated PET and CT. *Nucl Med Commun* 2006;27:515-519.

Schinagl DAX, Kappelle AC, van der Maazen RW, Bussink J. The importance of a complete diagnostic workup in patients with nontraumatic (partial) paraplegia. *Ned Tijdschr Geneeskd* 2003;147:2565-2569 (Dutch).

Curriculum Vitae

Dominic Schinagl werd geboren op 24 maart 1972 te Heerlen. In 1990 behaalde hij zijn gymnasium diploma aan het Katholiek Gymnasium Rolduc in Kerkrade. Hierna startte hij met de studie geneeskunde aan de Vrije Universiteit Amsterdam en in 1998 behaalde hij zijn artsexamen. In de periode 1998-2000 was hij werkzaam als arts-assistent heekunde in Hoorn en Amsterdam. In 2000 startte de opleiding tot radiotherapeut-oncoloog aan het UMC St Radboud in Nijmegen. Sinds 2005 is hij als stafid vebonden aan de afdeling radiotherapie van het UMC St Radboud. Dominic is getrouwd met Carla Verstappen en samen hebben zij twee zonen, Samuel en Simon.



Appendix: Color Figures

Chapter 2

Fig. 1 Planning CT scan (A), corresponding FDG-PET scan (B) and fusion image (C) show differences in target volume definition. Volume GTV_{CT} (red) = 47.5 cm³, GTV_{VIS} (green) = 43.8 cm³, $GTV_{40\%}$ (yellow) = 20.1 cm³, $GTV_{2.5}$ (orange) = 32.6 cm³, GTV_{SBR} (blue) = 15.7 cm³. Note that GTV_{SBR} is significantly smaller than GTV_{CT} and GTV_{VIS} .

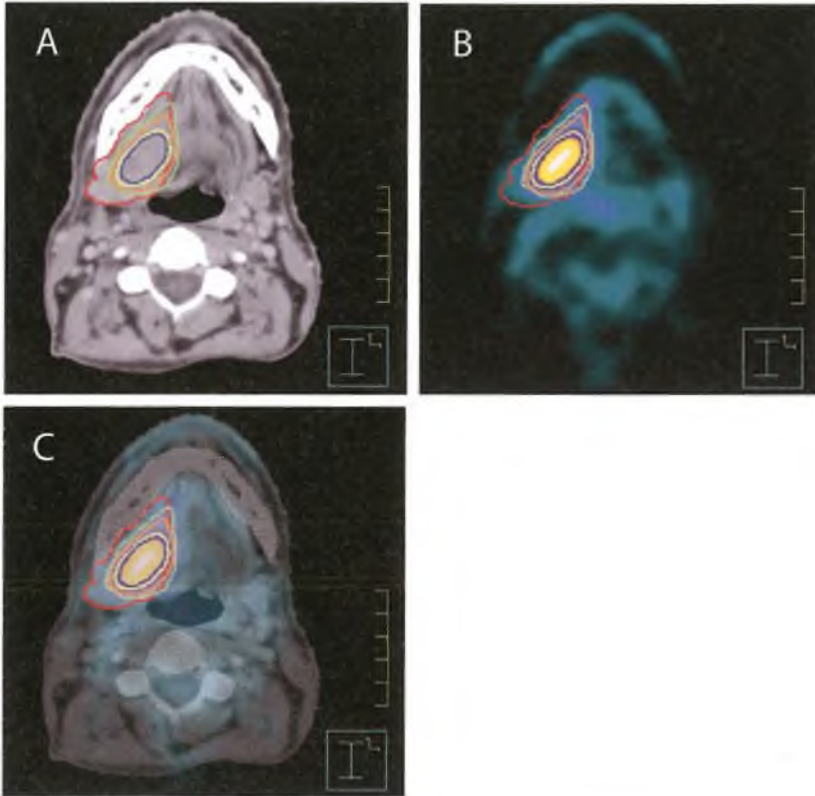


Fig. 2 Planning CT (A), corresponding FDG-PET (B) and fusion image (C) show differences in target volume definition. It also illustrates PET activity in the air cavity.

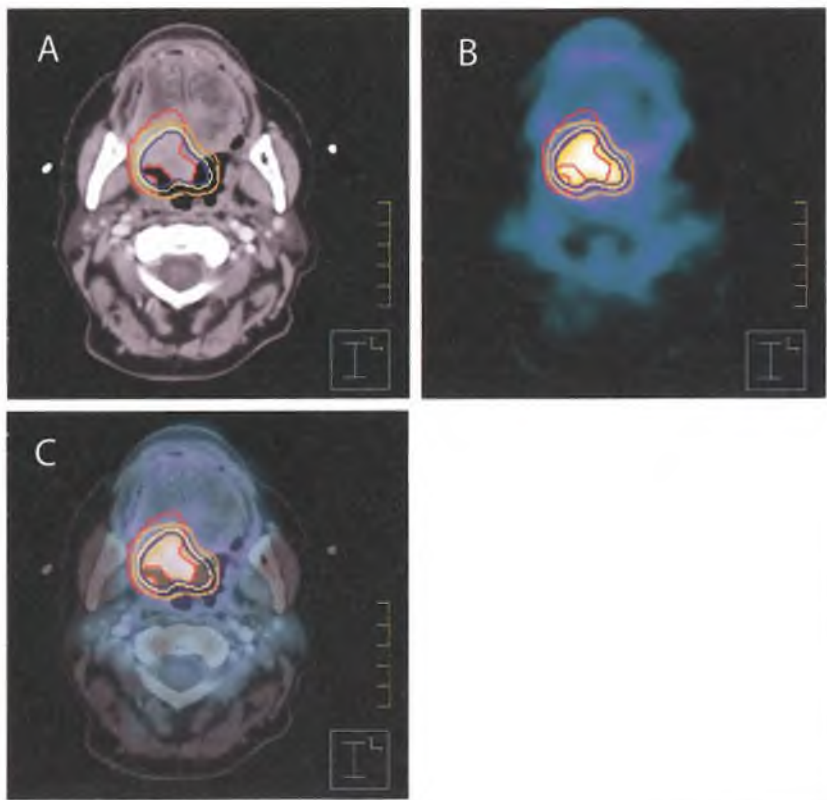
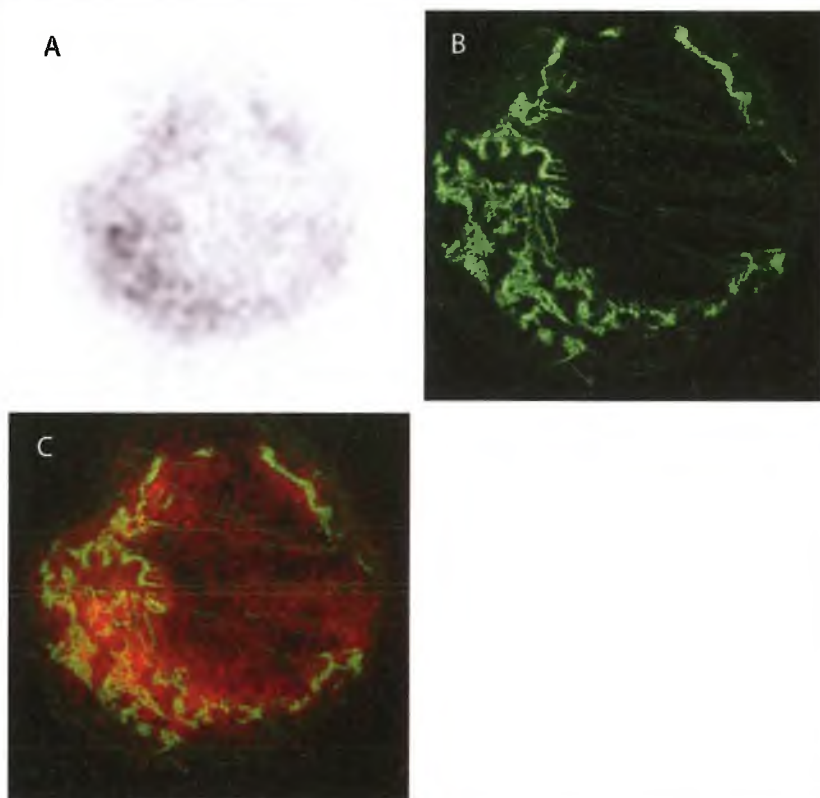


Fig. 4 FMISO autoradiography (A), pimonidazole immunohistochemistry (B) and fusion image (C) of a xenografted human head and neck cancer. Courtesy of E.Troost, Dept. Radiation Oncology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.



Chapter 3

Fig. 1 Multimodality markers. Example of the multi-modality markers used for landmark registration, and image fusion accuracy evaluation. (A) Markers (red arrows) as placed on the mask. (B) Marker as seen on CT filled with iodine-containing contrast. (C) Marker as seen on PET filled with FDG solution. The center of the markers could be determined with an inter-operator variability well below 1 mm on both CT and PET images.

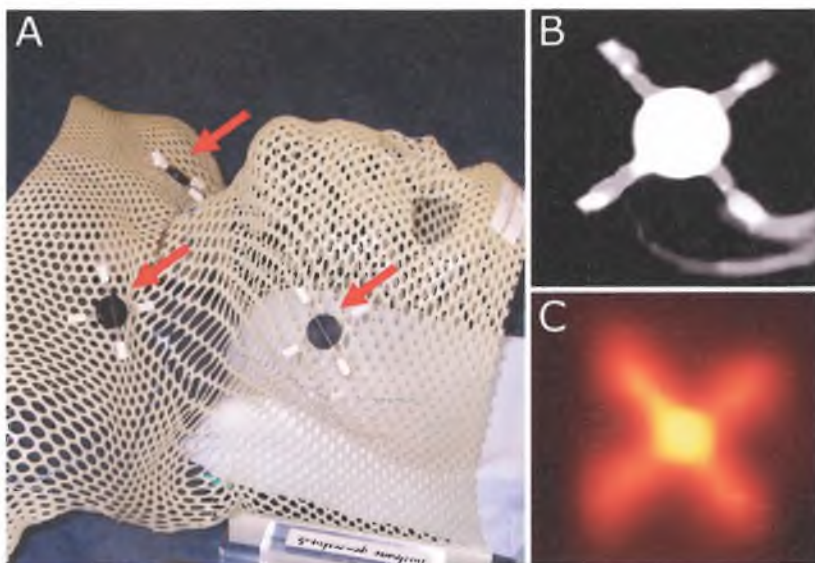


Fig. 2 Iterative closest point registration (ICP) of PET transmission images to CT. (A) On both CT and the PET transmission scan, a surface model of the body contour is generated using a thresholding technique. (B) Both surface models are represented as a 3-dimensional structure. (C) Automatic image registration is performed by iteratively minimizing the distance between the surface models.

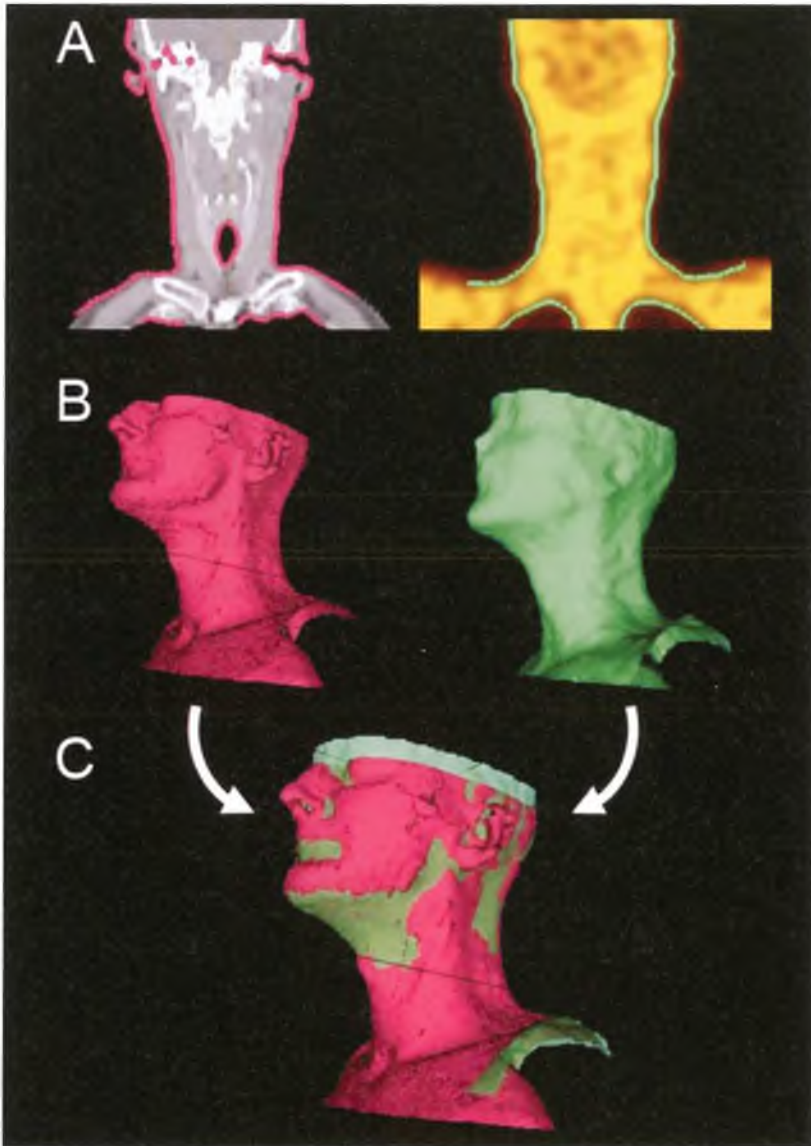
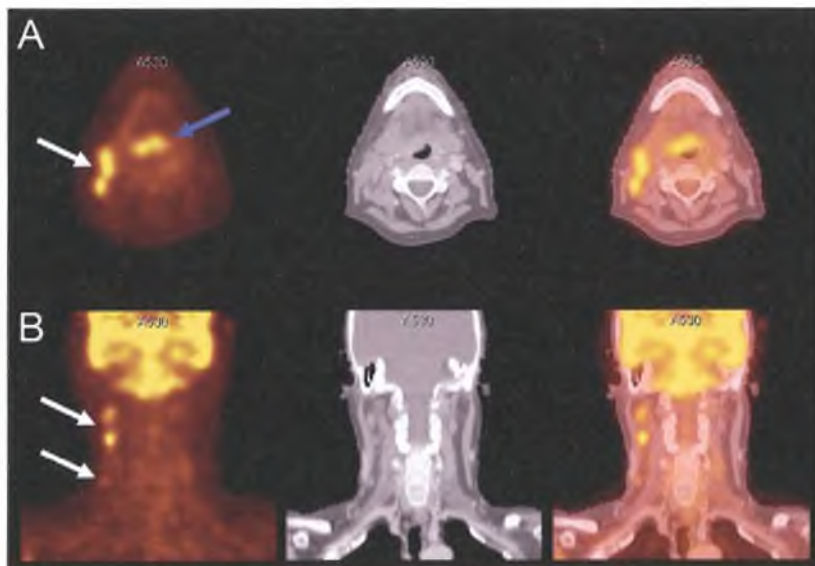


Fig. 3 Example of image fusion using method ICP Transverse (A) and coronal (B) sections of FDG-PET (left), CT (middle) and fused PET-CT (right) in a patient with carcinoma at the base of the tongue (blue arrow) and multiple pathological lymph nodes on the right side of the neck (white arrows). The extent of the malignancy was difficult to appreciate on the CT, while fused PET/CT images allowed clear delineation of the primary tumor, as well as identification of pathological lymph nodes.



Chapter 4

Fig. 1 CT scan (A), corresponding FDG-PET scan (B) and fused image (C) of a patient with a T4N2M0 tongue carcinoma, showing differences in target volume definition. Indicated are GTV_{CT} (red), GTV_{PT} (light green), GTV_{SUV} (orange), GTV_{40%} (yellow), GTV_{50%} (blue), and GTV_{50%} (dark green). GTV_{SUV} is unsuccessful in this case, due to inclusion of large areas of normal background tissue. Note that all other PET-based delineations indicate greater tumor extension towards lateral side and less towards medial side compared to CT delineation. Also note that on this transversal slice GTV_{50%} and GTV_{50%} are indistinguishable.

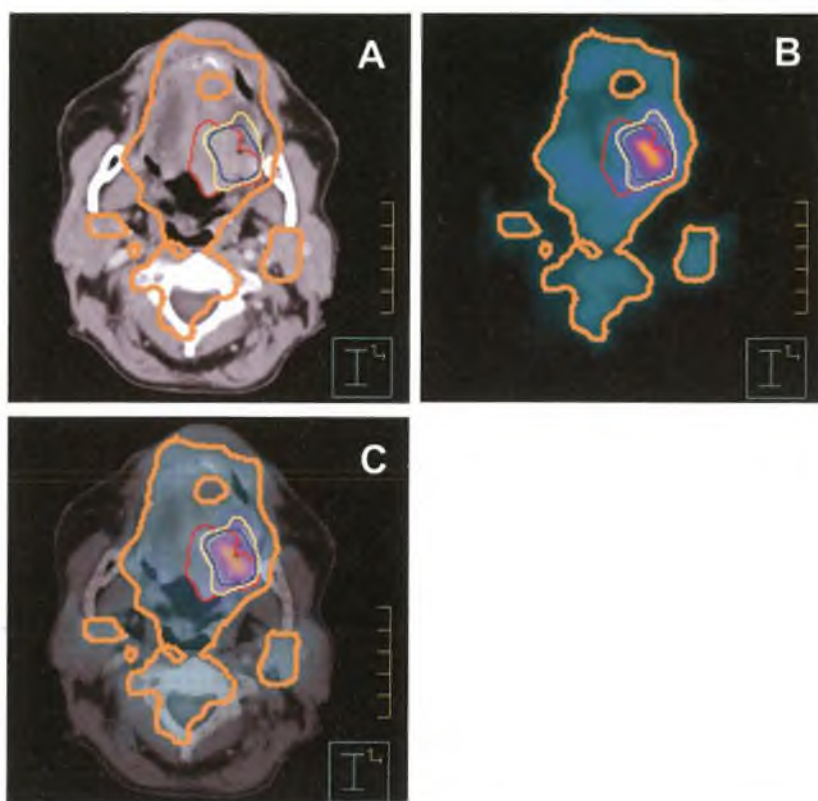
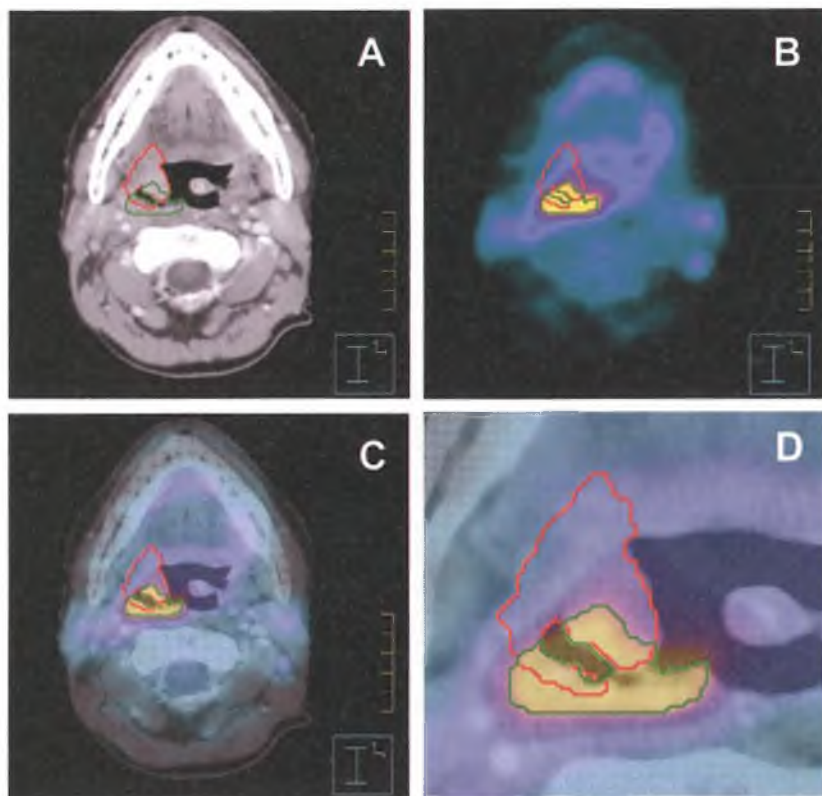
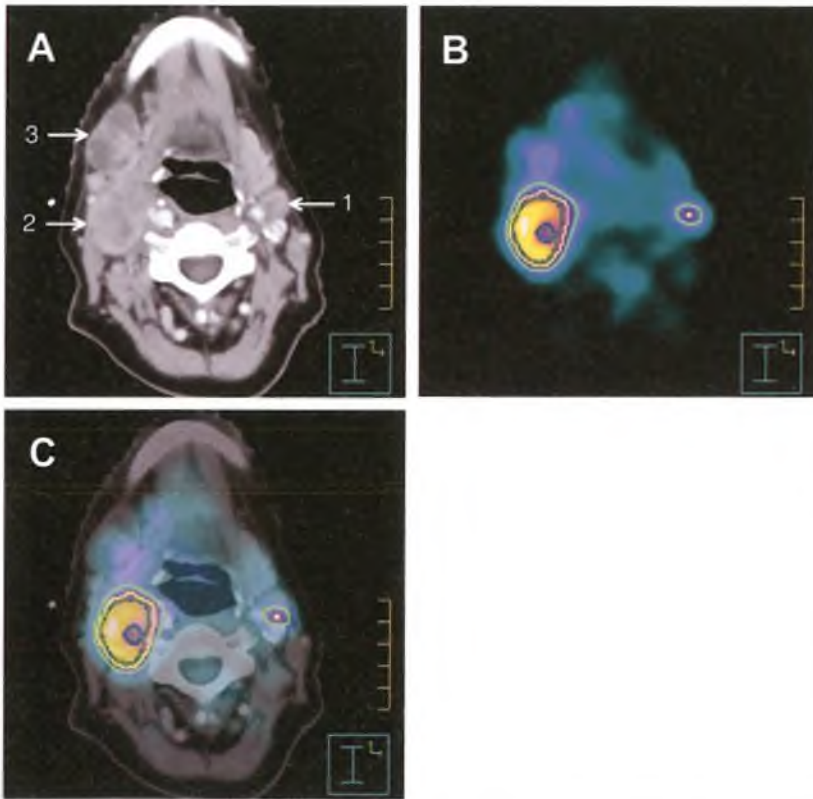


Fig. 3 CT scan (A), corresponding FDG-PET scan (B), fused image (C) and enhanced detail of fused image (D) of a patient with a T3N2cM0 oropharyngeal carcinoma, showing tumor tissue delineated by FDG-PET, but not by CT. Indicated are GTV_{CT} (red) and GTV_{SBR} (dark green).



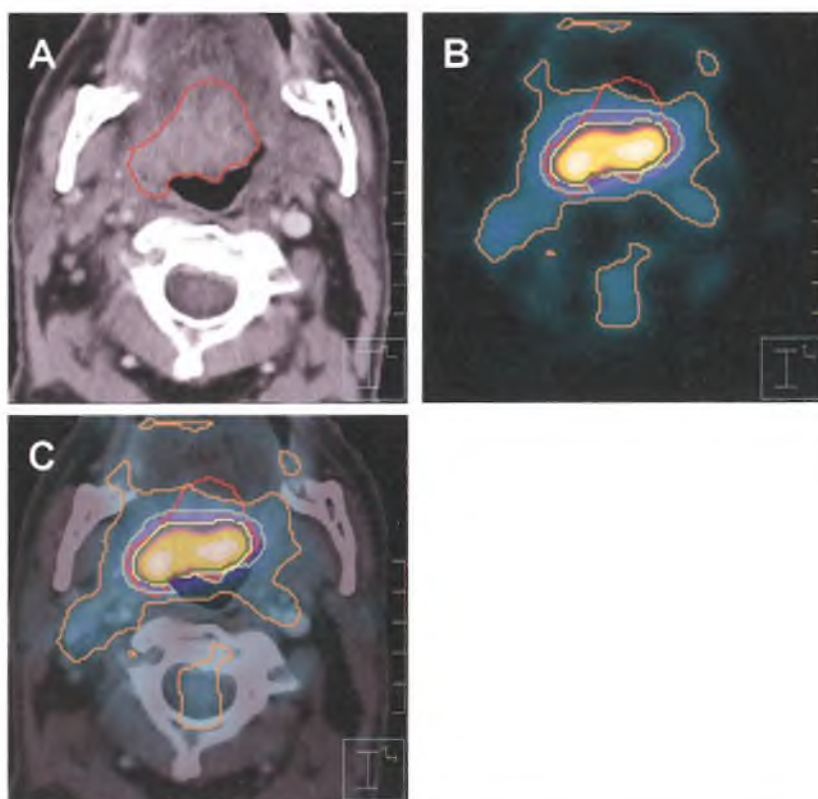
Chapter 5

Fig. 2 (A) CT scan, (B) corresponding FDG-PET scan, and (C) fused image of a patient with a T4 oropharyngeal carcinoma show three lymph nodes. Left (arrow1) a node that is 13mm in the shortest axial diameter that was identified by PET_{V15} (light green) and $PET_{40\%}$ (yellow). Right (arrow2) a node of 23mm that was identified by PET_{V15} , $PET_{40\%}$, $PET_{50\%}$ (blue) and PET_{SBR} (dark green). The pattern of FDG-uptake in this particular node suggests intranodal tumor heterogeneity. Right (arrow3) an overtly metastatic node of 21mm with central necrosis that was not identified by FDG-PET.



Chapter 6

Fig. 1 Planning CT (A), corresponding FDG-PET (B) and fusion image (C) of patient with T3N2bM0 oropharyngeal carcinoma show differences in target-volume definition. Indicated are gross tumour volume (GTV) delineated on CT (GTV_{CT} , red, absolute volume of 34.0cm^3) and PET-based GTVs obtained by visual interpretation (PET_{VIS} , light green, volume 33.8cm^3), applying an isocontour of a standardized uptake value (SUV) of 2.5 ($PET_{2.5}$, orange), using a fixed threshold of 40% ($PET_{40\%}$, yellow, volume 14.0cm^3) and 50% ($PET_{50\%}$, blue, volume 13.4cm^3) of the maximum signal intensity, applying an adaptive threshold based on the signal-to-background ratio (SBR, PET_{SBR} , dark green, volume 15.0cm^3). $GTV_{2.5}$ was unsuccessful in this case because of inclusion of large areas of normal background tissue. Note that on this transversal slice $PET_{50\%}$ and PET_{SBR} are indistinguishable.



Chapter 7

Fig. 2a-1 correlation of CT- and FDG-PET based (PET_{V15} , PET_{SUV} , $PET_{40\%}$, $PET_{40\%N}$, $PET_{50\%}$, and $PET_{50\%N}$) metastatic lymph node volumes in relation to the corresponding pathological volume. The line in the correlation plots represents identity ($y = x$).

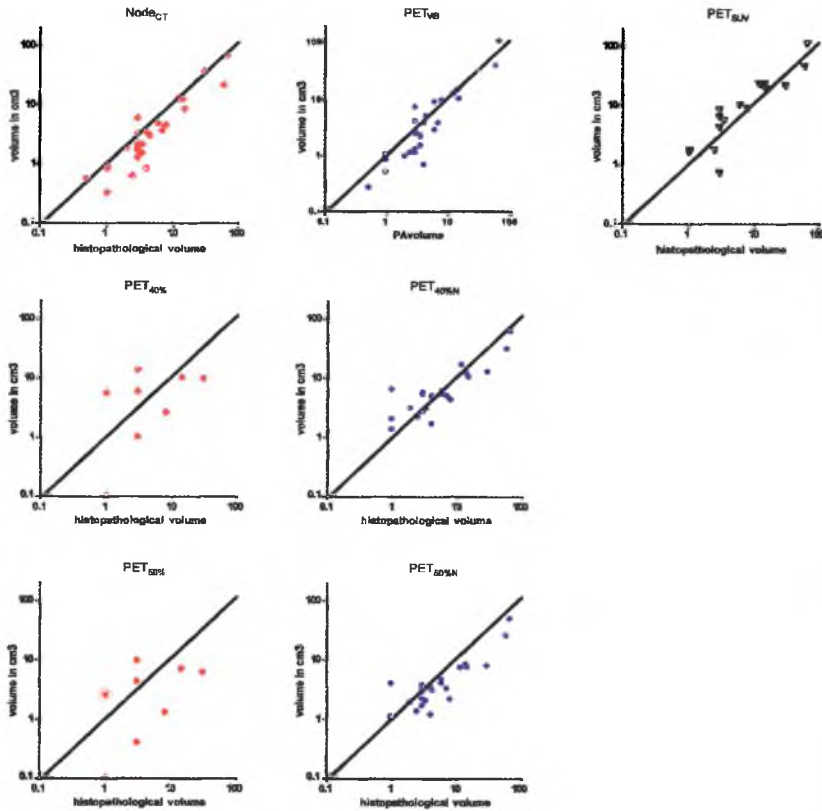


Fig. 2a-2 correlation of CT- and FDG-PET based (PET_{SBR} , PET_{SDRN} , PET_{RTL} , and PET_{RTLN}) metastatic lymph node volumes in relation to the corresponding pathological volume. The line in the correlation plots represents identity ($y = x$).

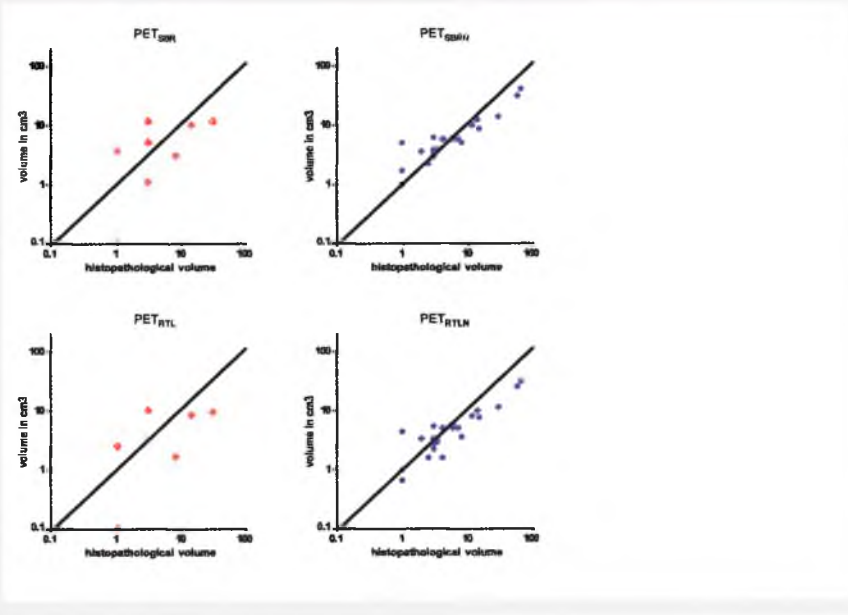
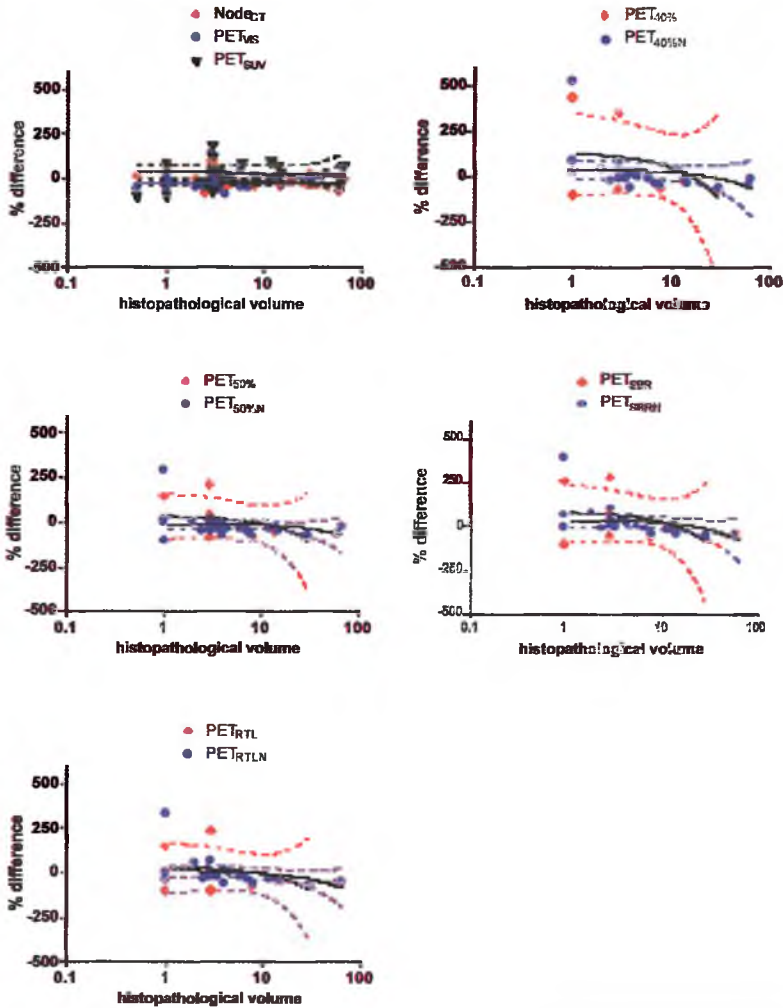


Fig. 2b Difference plots of CT- and FDG-PET based metastatic lymph node volumes in relation to the corresponding pathological volume. The 95% confidence intervals are shown as dotted lines



Chapter 8

Fig. 1 FLT PET/CT scan for image-guided high-precision radiation treatment planning in oropharyngeal cancer.

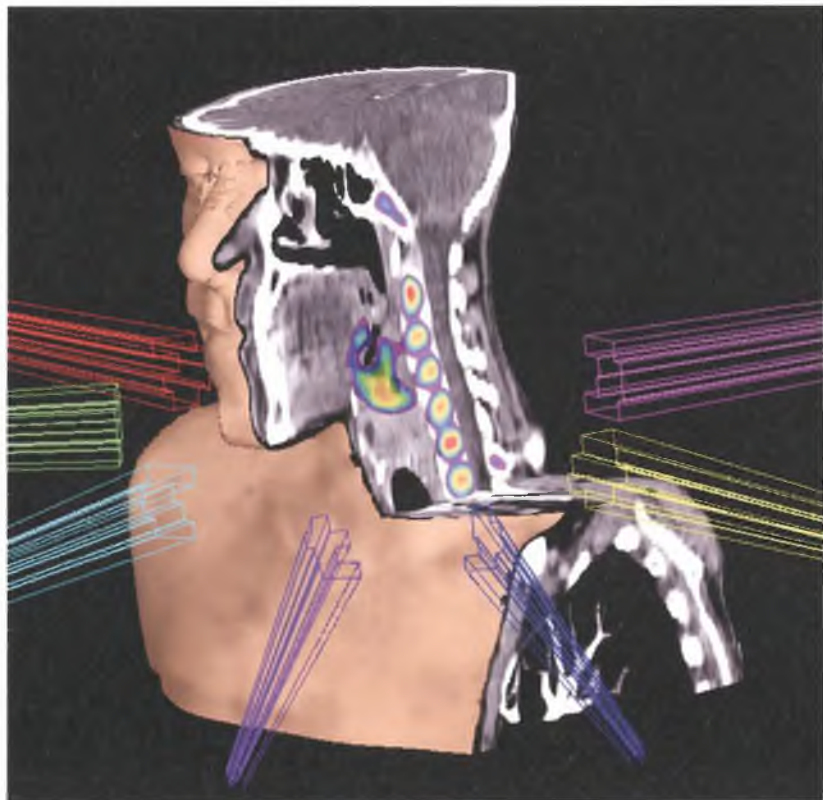


Fig. 2 Planning CT scan (A), corresponding FDG-PET scan (B) and fusion image (C) of a patient with T4N2cM0 oropharyngeal cancer show differences in target volume delineation. Indicated are GTV delineated on CT (GTV_{CT} , red), and PET-based GTVs obtained by visual interpretation (GTV_{vis} , light green), using fixed threshold of 40% ($GTV_{40\%}$, yellow) and 50% ($GTV_{50\%}$, blue) of the maximum signal intensity, applying an adaptive threshold based on the signal-to-background ratio (GTV_{SBR} , dark green, largely covered by $GTV_{50\%}$ in blue) and applying an isocontour of a standardized uptake value (SUV) of 2.5 ($GTV_{2.5}$, orange). The respective volumes ranged from 15.1cm³ (GTV_{SBR}) to 59.7cm³ ($GTV_{2.5}$).

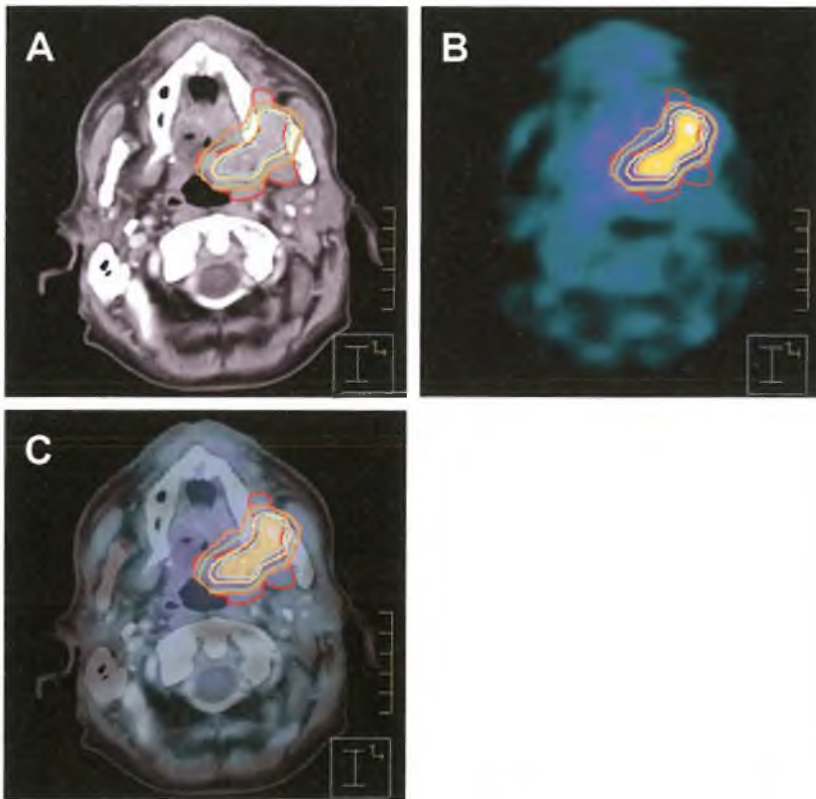


Fig. 3 Planning CT scan (A), corresponding FDG-PET scan (B), fusion image (C) and calculated dose distribution (D) of a patient with T3N2M0 hypopharyngeal cancer. Red = GTV_{CT} (absolute volume of 39 cm³), light blue = GTV_{SBR} (absolute volume of 13.1 cm³). Also illustrated are planning target volume to 50.3 Gy (pink), to 68.0 Gy (dark blue). Additional dose of 4.0 Gy is directed to GTV_{SBR} using IMRT with integrated boost technique in accelerated scheme. A subvolume of GTV_{CT} thus receives a total dose of 72.0 Gy.

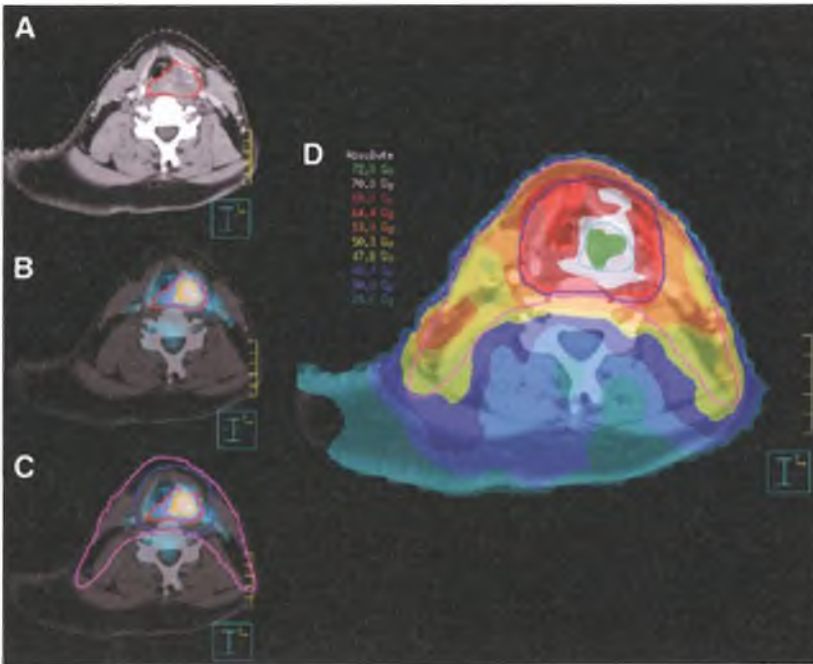


Fig. 4 Gray-value images after immunohistochemical staining of nitroimidazole hypoxia marker pimonidazole (left) and FMISO autoradiography images (right) of SCCNij3 (A), SCCNij153 (B), and SCCNij86 (C) xenografted human squamous cell carcinomas of head and neck. Corresponding staining patterns and signal intensities are seen for pimonidazole and FMISO in SCCNij3 and SCCNij153, but not in SCCNij86.

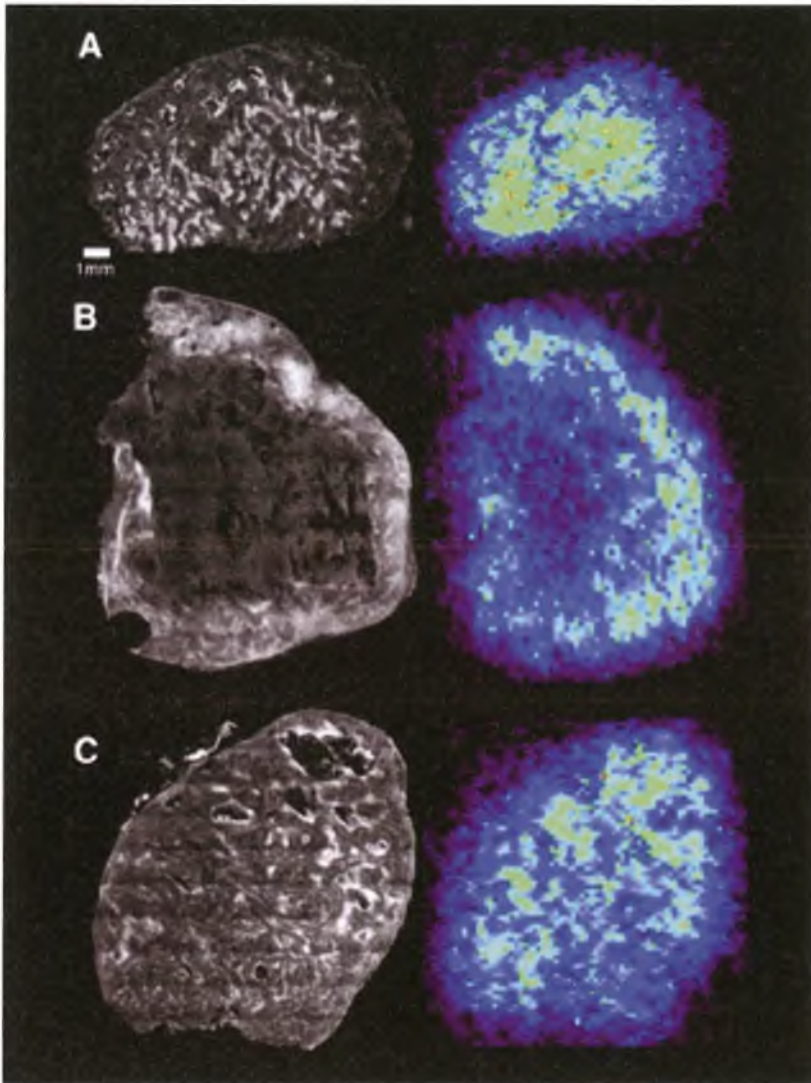


Fig. 5 CT (left), FLT-PET (middle), and fused FLT-PET/CT (right) images of patient with T3N0M0 oropharyngeal cancer before radiation therapy (A) and after 8 fractions of 2 Gy (B). Red = GTV_{CT}. In B, significant reduction in FLT-PET signal intensity is already seen, whereas only a modest decrease of GTV_{CT} is seen at this dose level.

