DIFFUSION-WEIGHTED MRI AND ^{18}F-FDG-PET-CT IN HEAD AND NECK CANCER:

PREDICTION AND RESPONSE EVALUATION

CHARLOTTE SOPHIE SCHOUTEN

COLOFON

Cover and lay-out by: C.S. Schouten
Printed by: Ipskamp printing │ Enschede
ISBN: 978-94-028-0674-8

The research described in this thesis was performed at the Department of Otolaryngology / Head and Neck Surgery of the VU University Medical Center, Amsterdam, the Netherlands. The research was partly funded by a research grant of The Netherlands Organisation for Health Research and Development (ZonMw).

VRIJE UNIVERSITEIT

DIFFUSION-WEIGHTED MRI AND ^{18}F-FDG-PET-CT IN HEAD AND NECK CANCER:

PREDICTION AND RESPONSE EVALUATION

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op maandag 6 november 2017 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

Door

Charlotte Sophie Schouten

Geboren te Oranjestad, Aruba

promotoren:	prof.dr. R. de Bree
	prof.dr. J.A. Castelijns
copromotoren:	prof.dr. O.S. Hoekstra
	prof.dr. C.R. Leemans

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND OUTLINE OF THESIS

HEAD AND NECK SQUAMOUS CELL CARCINOMA

Etiology and staging

Head and neck squamous cell carcinoma (HNSCC) refers to malignant tumors that originate from the epithelium of the upper aerodigestive tract including the oral cavity, pharynx, larynx, nasal cavity and paranasal sinuses (Figure 1). Squamous cell carcinoma consists of malignant cells with squamous differentiation. HNSCC is the sixth most common cancer, accounting for approximately 4% of all malignant tumors worldwide. The estimated incidence is 630.000 per year with a mortality-rate of 350.000 [1,2]. In the Netherlands approximately 3000 new patients are diagnosed with HNSCC every year [3,4]. Tobacco smoking and excessive alcohol consumption are well-known risk factors [5]. Several other factors are also associated with the development of HNSCC, like human papillomavirus (HPV) [6,7], Epstein-Barr virus (EBV) [8] and genetic and immunological predisposition.

The stage at diagnosis is an important parameter as it mostly determines the treatment plan and prognosis of patients with HNSCC. Staging is performed according to the TNM system of the International Union Against Cancer which consists of the extent of the primary tumor (T), the involvement of regional cervical lymph node metastasis (N) and the presence of distant metastasis (M). Based on this system, clinical stages I-IV can be established; tumors localised within the organ of origin are stage I and II, stage III represents local extension or with cervical lymph node metastasis and stage IV means more advanced cancer, *i.e.* primary tumor extension beyond the organ of origin and extensive nodal disease or distant metastasis [9]. About one third of the patients present with early staged disease (stage I and II), while two third presents with advanced staged disease (stage III and IV), especially patients with oropharyngeal cancer [10].

Oropharyngeal squamous cell carcinoma and HPV

The oropharyngeal region is a part of the pharynx and comprises the tonsils, base of tongue, vallecula, posterior pharyngeal wall and the soft palate (including uvula) (Figure 1). Over the last decades, it has been well established that HPV is an important etiological factor in the development of particularly squamous cell carcinomas in the oropharynx (OPSCC), most commonly in the base of tongue and tonsils [6,7,11,12]. Approximately 570 new patients with OPSCC are diagnosed every year in the Netherlands [4] of which 175 patients

have HPV-positive carcinomas [3]. The incidence of oropharyngeal cancer is increasing in North-western Europe and Northern America [13-17], despite successful attempts to control tobacco and alcohol consumption, suggesting an increased contribution of HPV. Prevalence-rates of HPV-positive OPSCC vary around the world. As mentioned before, prevalence rates of HPV-positive OPSCC in the Netherlands are approximately 30% [18], while prevalences of 72% [19] are reported in the United States. HPV-positive OPSCC characteristically occurs in younger (<50 year) white male patients without a history of smoking and alcohol abuse [7,13]. HPV-positive and HPV-negative OPSCC are distinct disease entities [20]; HPV-related tumors have distinct histological features [21,22] and have a different genetic route to cancer [23,24].

The HPV genome consists of double-stranded DNA and produces two oncoproteins E6 and E7. These oncoproteins play a role in oncogenesis by inactivation of apoptotic pathways, disruption of cell cycle control and activation of cellular proliferation. The E6 oncoprotein inactivates and promotes degradation of the tumor suppressor p53, thereby disrupting normal apoptosis [25,26]. The E7 oncoprotein blocks the function of retinoblastoma (RB) pocket proteins, which leads to cell cycle entry and hereby allowing the virus to replicate [27,28].

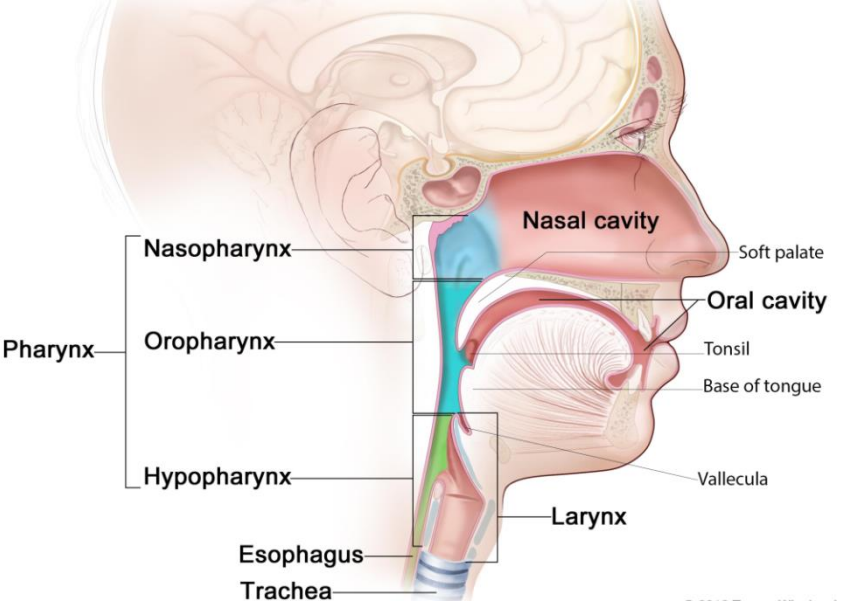


Figure 1.
Anatomy of the head and neck region, with sites of origin of tumors in the head and neck region including oropharyngeal subsites.

Therapy and prognosis

Three modalities are used in the treatment of HNSCC: surgery, radiotherapy and chemotherapy, which can be used separately but often a combination of these therapies is needed. Formerly, patients with advanced but resectable HNSCC were treated with a combination of surgery and postoperative radiotherapy, often leading to functional and cosmetic morbidity, inducing a diminished quality of life [29]. Unresectable HNSCC was treated by radiotherapy with or without chemotherapy. Nowadays, non-surgical protocols are applied in the treatment regimen of resectable tumors as well, intending to preserve organ function and to maintain quality of life. This application of organ preservation protocols resulted in considerable locoregional control rates and it appeared that intensified radiotherapy schemes (accelerated or hyperfractionated) and combinations of chemotherapy and radiotherapy (CRT) all contribute to an increased control rate [30-32]. At present, in the Netherlands most commonly used concomitant CRT with cisplatinum comprises 7 weeks of radiotherapy (70 Gray in 35 fractions) combined with three courses of intravenous cisplatinum (100 mg/kg) on day 1, 22 and 43 [33].

Non-surgical treatment is associated with acute and long-term side effects; common complaints with radiotherapy are xerostomia, impairment in speech and difficulties in swallowing. Fibrosis of the neck with reduced mobility and pain are also not unusual [34,35]. Chemoradiotherapy with cisplatinum may be accompanied by serious toxic side effects as cisplatinum is oto-, neuro- and nephrotoxic. Therefore, chemoradiotherapy is only appropriate for patients with a good health performance and normal kidney function. Accelerated radiotherapy (6 fractions a week) in combination with cetuximab (400 mg/m^2, followed by 250 mg/m^2 a week) is an alternative for elderly patients or patients with a poor performance status or decreased creatinine clearance. Adding cetuximab to radiotherapy alone leads to a significantly better locoregional control and survival [31,36]. Cetuximab appears to be well endured, however, acute and chronic toxic skin effects do occur, such as a rash, discomfort and itching [37].

Prognosis is linked to the clinical stage at diagnosis. Patients with early stages HNSCC (stadium I and II) generally have a fairly good prognosis (80-90% 5-year survival rate), in contrast to patients with advanced disease (stadium III and IV), who have a much worse prognosis (40-50% 5-year survival rate). However, patients with HPV-positive OPSCC show more favourable treatment responses and survival rates as compared with HPV-negative OPSCC, despite the fact that patients with HPV-positive OPSCC more often present with regionally advanced disease [38-40].

Response evaluation

Fairly low locoregional residual and recurrent tumour rates indicate that CRT is an acceptable treatment option for patients with HNSCC [30,32]. However, after failure of CRT, surgical salvage treatment may be curative if residual locoregional disease of initially resectable HNSCC is detected timely. In case of late detection and delayed salvage surgery, locoregional control and survival rates rapidly decrease [41]. An examination under general anesthesia (EUA) (with eventual biopsies) is considered to be the most reliable procedure to detect local residual disease and since salvage surgery is associated with considerable morbidity and a high risk of complications, it will only be performed after histopathological confirmation of viable tumor cells obtained during EUA. Consequently, current clinical practice in the VU University Medical Center is to perform stringent response evaluation, *i.e.* EUA, routinely 12 weeks after the end of CRT. However, due to the acceptable locoregional control rates after CRT, many patients are exposed to unnecessary anaesthesia and biopsies in irradiated areas, potentially inducing inflammation and pain [42].

Timely detected residual neck disease after CRT for HNSCC can often be successfully salvaged with a neck dissection. A neck dissection is a relatively safe procedure with generally good therapeutic results. Nevertheless, it is associated with certain risks and complications, especially after previous radiation therapy. Surgical complications, such as bleeding, infection and chylous leakage can occur [43,44]. Moreover, cutaneous nerves are sacrificed during the operation and can result in insensibility of the neck or a dull neck pain. Also, the accessory nerve can be damaged, leading to shoulder impairment. Clinical palpation of the neck is most commonly used as diagnostic tool to assess the neck after CRT. If the presence of a residual lymph node is suspected, an ultrasound-guided fine needle aspiration cytology (FNAC) can be performed [45]. On the other hand, planned neck dissection after CRT in all patients with initial N2-N3 disease is also advocated because the presence of residual neck disease cannot be safely predicted with currently used clinical methods.

The assessment of local and regional treatment response varies widely between institutions. It mostly includes a combination of physical examination and imaging. Depending on the initial TNM-stage and local guidelines, outpatient follow-up visits and imaging, such as computed tomography (CT), magnetic resonance imaging (MRI) or ultrasound-guided FNAC are performed. However, discrimination of malignant and aspecific changes after CRT with CT and conventional MRI is unreliable since post-treatment changes, including

oedema, fibrosis and necrosis, hamper accurate assessment [46,47]. Manikantan et al. recommended to perform imaging (CT or MRI) 3-6 months after CRT for the purpose of establishing a new baseline situation for later reference [48]. During recent years, functional imaging with positron emission tomography-computed tomography (PET-CT) and diffusion-weighted MRI (DW-MRI) became available. PET-CT has been gradually incorporated as a part of the post-treatment assessment. DW-MRI is a relatively newer technique and is mostly used for scientific research but in some institutions clinicians already rely on DW-MRI for response evaluation after CRT.

Treatment prediction

HNSCCs have a heterogeneous response to treatment [38,39] and the possible value of imaging parameters as predictive factors for treatment outcome are under investigation. Current traditional prognostic factors, such as TNM-stage, HPV-status, tobacco smoking and alcohol consumption are insufficient to classify patients into risk groups. Identification of other prognostic characteristics could lead to patient selection for tailored treatment regimens and possibly result in higher responses to treatment and less treatment-induced side effects. Tumor characteristics can be assessed by imaging prior to treatment or changes can be evaluated early during treatment. For PET-CT, several studies have shown that patients with high pretreatment standardized uptake values (SUV)-values generally have a less favourable outcome, so it seems that a high rate of glucose metabolism is a characteristic of aggressive tumor behaviour [49-52]. For DW-MRI, it has been suggested that HNSCC with relatively low pretreatment apparent diffusion coefficient (ADC)-values respond better to CRT than tumors with higher pretreatment ADC-values [53-56]. However, it is unknown if these reported associations are partly or even completely attributed to the HPV-status, as patients with an HPV-positive OPSCC respond better to CRT than HPV-negative patients.

In patients with resectable HNSCC, CRT with eventual salvage surgery for residual disease in reserve is preferred. However, salvage surgery after CRT is associated with considerable risks and a high incidence of complications [57]. Another important disadvantage of salvage surgery after (chemo)radiotherapy is the fact that, although it may be indicated on the basis of the pathology results of the surgical specimen, postoperative radiotherapy is rarely possible thus limiting the outcome of this treatment. Therefore, a primary non-surgical treatment may not be the choice of treatment in all patients. Prediction of therapy outcome early during treatment could select patients who are likely not to benefit from CRT. This selection of patients may spare a

certain number of patients from ineffective treatment and may allow a treatment switch to surgery in these patients, preserving postoperative radiotherapy (if indicated) as an option. Several clinical studies have indicated the potential of functional imaging with PET-CT or DW-MRI early during CRT as a predictor of treatment response [53,58-62].

PET-CT

General principles

Conventional imaging methods such as CT and MRI are primarily anatomical imaging techniques using the physical properties of tissues (absorption of X-rays and proton densities, respectively) to differentiate abnormal from normal tissue. PET-CT is a functional imaging technique which detects the distribution of specific radioactive tracers and combines imaging of metabolically active tissues with anatomical information (Figure 2).

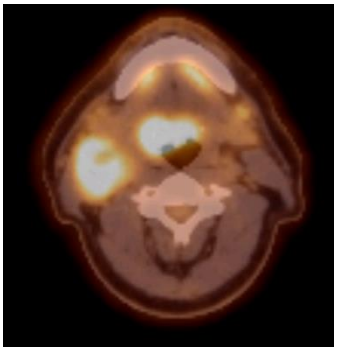


Figure 2.
An example of a ^{18}F-FDG-PET-CT image depicting high ^{18}F-FDG uptake in a right base of tongue cancer and a large ipsilateral lymph node metastasis.

PET imaging uses the administration of a molecule which is labelled with a positron emitter (PET tracer). Positron emitters are radioactive atoms with a positive charge. These radionuclides undergo positron decay through emission of a positron. Depending on the characteristics of the original isotope, the positron has kinetic energy which defines the distance of motion the positron travels in matter. After travelling a distance or range in matter, the positrons lose theirs kinetic energy through ionisation and interactions with electrons. After losing its kinetic energy, the process of annihilation will take place and the positron will combine with a free electron. This results in the emission of two photons with an equally shared energy of 511 keV in opposite direction (180º±25º) from the annihilation site (Figure 3) [63].

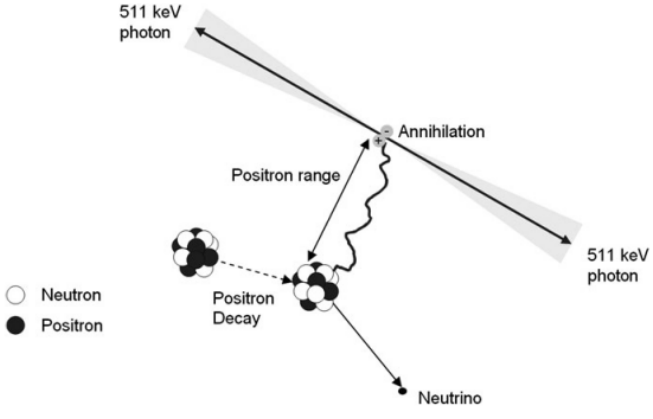


Figure 3. [63]
Positron decay and annihilation

The distribution of the PET tracer within the body can be imaged by a PET scanner, which can be integrated with CT or MRI [64]. Positron imaging uses multiple rings of detectors placed surrounding a patient to detect the directional relationship of two simultaneously emitted photons. Coincidence imaging is used by two opposing detectors to detect a pair of photons, thereby assuming that the point of annihilation occurred within the volume between the ring of detectors. The source of emission (detected event) can be localized along a straight line joining the two detectors and passing through the point of annihilation; the line of response (LOR) (Figure 4). Time-of-flight information is used to further localize the position of annihilation. By accurately measuring the detection times, the time difference in the detection times can provide information about where along the LOR the annihilation happened [63].

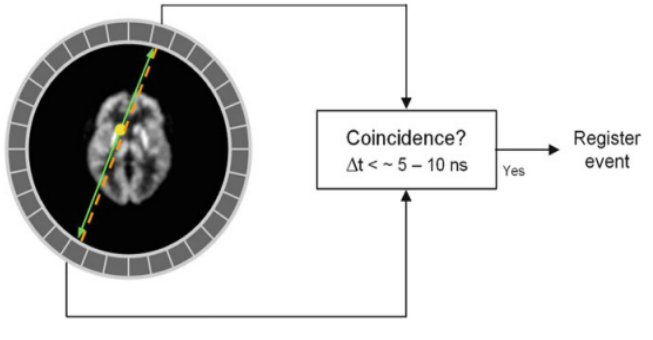


Figure 4. [63]
Coincidence detection of annihilation photons

Currently, mostly PET-CT scanners are employed in the clinical field; a combination of a PET camera and a CT is used in which the patients move sequentially on a common bed from CT to PET acquisition. The CT scan aids in localizing the lesions visualised by PET and is also used for attenuation correction. Due to photoelectric absorption in a patient, the number of detected photons is reduced, which is called attenuation. However, the relatively poor soft-tissue contrast of CT can be a disadvantage using CT as the supplementary anatomical modality with PET, and coupling PET with MRI can be relevant in clinical situations in which the soft-tissue contrast of MRI outperforms that of CT, for example after radiotherapy in the head and neck area. The first generation of hybrid-PET-MRI systems are emerging into the clinical field permitting combined PET and MRI evaluation.

^{18}Fluoro-2-deoxyglucose

In the last decades, several positron emitting radioactive isotopes have been developed worldwide, but ^{18}Fluoro-2-deoxyglucose (^{18}F-FDG) is the most widely used tracer in oncological PET studies. ^{18}F-FDG contains the positron emitter fluorine-18 (^{18}F, $t_{1/2}$= 110 minutes). ^{18}F-FDG is a glucose analogue and can measure glucose metabolism in malignant tissue. Higher concentrations of ^{18}F-FDG accumulate in malignant tumors than in normal tissue, due to increased glycolytic activity of the cancer cells. However, physiological uptake occurs in the brain, myocardium and sometimes the colon and the tracer is cleared by the renal system [65]. Also, unwanted uptake may be seen in the neck musculature and vocal folds, due to movement or talking and uncomfortable waiting conditions [66,67]. Brown adipose tissue (BAT) is closely localized to lymph nodes in the head and neck region and ^{18}F-FDG uptake in BAT can be another source of false-positive results. The administration of β-adrenergic antagonists, such as propranolol, prior to ^{18}F-FDG injection can reduce tracer uptake in BAT [68].

Acquisition

PET-CT examinations are started by the intravenously administration of an amount of ^{18}F-FDG, depending on the body mass index [69], after 4-6 hours of fasting. Blood glucose levels are measured prior the ^{18}F-FDG administration and should be <10 mmol/L. Hereafter, the patients rests for 60 minutes in a dark room to avoid unwanted uptake in muscles. Low-dose CT scanning is performed prior to emission scanning for attenuation correction and anatomical localization. During PET scanning, images are acquired for one to four minutes

per bed position and the scan trajectory usually extends from mid femur to cranial vault. In the VU Medical Center a dedicated head and neck protocol is used for head and neck cancer patients, consisting of a scan trajectory from jugular notch to orbit with 4 min emission scans/bed position while the patients are scanned with the arms alongside their body. This protocol is performed prior to the whole body protocol (2 min emission scan/bed position, arms up).

Applications in head and neck oncology

Over the last decades, ^{18}F-FDG-PET-CT has been increasingly employed in HNSCC for a number of clinical applications [70]. First, PET-CT appeared to be of additional value in tumor staging; *i.e.* the detection of distant metastasis and simultaneous secondary primary tumors in advanced staged HNSCC [71-73]. Secondly, in case of cervical lymph node metastasis from an unknown primary tumor, PET-CT has been reported to locate occult primary tumors in up to 44% of these patients [74,75]. Thirdly, PET-CT has a high accuracy in monitoring response to (chemo)radiotherapy (*i.e.* response evaluation) and for the detection of residual or recurrent tumor during follow-up, both at the primary site and the neck. In a meta-analysis, the pooled sensitivity and specificity of PET (with or without CT) for the detection of local residual/recurrent disease were 79.9% and 87.5%, respectively and 72.7% and 87.6% for the detection of regional recurrent disease [76,77]. Newer implementations of ^{18}F-FDG-PET-CT concern radiotherapy treatment planning and prediction of response to CRT. PET-CT may improve and facilitate delineation of the primary tumor and lymph node metastasis in HNSCC for radiation treatment purposes [78]. PET-CT also has shown to be of predictive value in HNSCC; high pretreatment ^{18}F-FDG uptake is a poor prognostic factor [50-52,79,80]. Another approach is predicting treatment outcome early during chemoradiotherapy; clinical studies report associations between decline in ^{18}F-FDG uptake in the early phase of CRT and a positive treatment outcome [60-62,81].

Qualitative and quantitative interpretation

PET-CT scans can be interpreted in a qualitative (*i.e.* visual) or a quantitative manner. Visual assessment of images is clinically practicable for both nuclear medicine physicians and referring physicians because it provides instantaneous evaluation. However, qualitative interpretation is a subjective task and interobserver variation can occur. A validated interpretation system assists in clarifying equivocal findings found and can improve systematic and reproducible visual review. Until recently there had been no established

interpretation-system for response evaluation in head and neck cancer. In 2014, the Hopkins criteria were validated with substantial interobserver agreement in a retrospective series of 214 patients [82]. These criteria use a qualitative 5-point scale in which scores 1, 2 and 3 are considered negative for tumor and scores 4 and 5 are considered positive for tumor (Table 1).

Table 1. Hopkins Criteria for Head and Neck PET/CT [82]

Score	***^{18}F-FDG-uptake pattern***	***Response category***
1	^{18}F-FDG uptake at the primary site and nodes less than IJV	Complete metabolic response
2	Focal ^{18}F-FDG uptake at the primary site and nodes greater than IJV but less than liver	Likely metabolic response
3	Diffuse ^{18}F-FDG uptake at the primary site or nodes is greater than IJV or liver	Likely postradiation inflammation
4	Focal ^{18}F-FDG uptake at the primary site or nodes greater than liver	Likely residual tumor
5	Focal and intense ^{18}F-FDG uptake at the primary site or nodes	Residual tumor

Quantification of ^{18}F-FDG uptake can also be used and possibly further improve diagnostic accuracy. The SUV is a semi-quantitative measurement of tracer uptake measured by drawing a Volume of Interest (VOI) using a 3-dimensional region-growing algorithm. For each VOI, the pixel with maximum SUV within the VOI (SUV_{max}), the average of a few pixels around the pixel with SUV_{max} (SUV_{peak}), the average SUV using an adaptive threshold of 50% (SUV_{mean}), ^{18}F-FDG-avid tumor volume (metabolic active tumor volume; MATV) and Total Lesion Glycolysis (TLG), calculated as product of SUV_{mean} and MATV, can be obtained.

The standardization of the imaging procedure, the interpretation and report on the results of ^{18}F-FDG-PET-CT in oncologic imaging has led to the applicability of ^{18}F-FDG-PET-CT in multicentre studies and increased the value to evidence-based medicine. For this harmonisation of ^{18}F-FDG-PET-CT, procedure guidelines were formed in 2010 and an updated version was published by the European Association of Nuclear Medicine (EANM) [83]. In these guidelines, a common quality control/quality assurance procedure is described to support the accuracy of quantification, in which repeatability and reproducibility are important factors. In this manner, consistency between scanners and institutions is secured in case these guidelines are pursued.

DIFFUSION-WEIGHTED MRI

General principles

DW-MRI is a relatively new MR imaging technique within the oncologic applications. DW-MRI has several advantages; it can be included on most MRI scanners in addition to existing clinical imaging protocols, taking only a few extra minutes, and without the need for contrast administration. DW-MRI characterizes micro-environment and cellularity of tissues based on differences in the random Brownian motion of water protons, which is mainly influenced by the amount of extracellular space and the presence of cell membranes [84]. In malignant tumors a higher cell density with limited extracellular space is seen, leading to a reduction in mobility of the water protons and consequently a restriction of Brownian displacement. Tissues containing inflammation or necrosis will have lower cell density leading to higher proton mobility [85] (Figure 5 and 6).

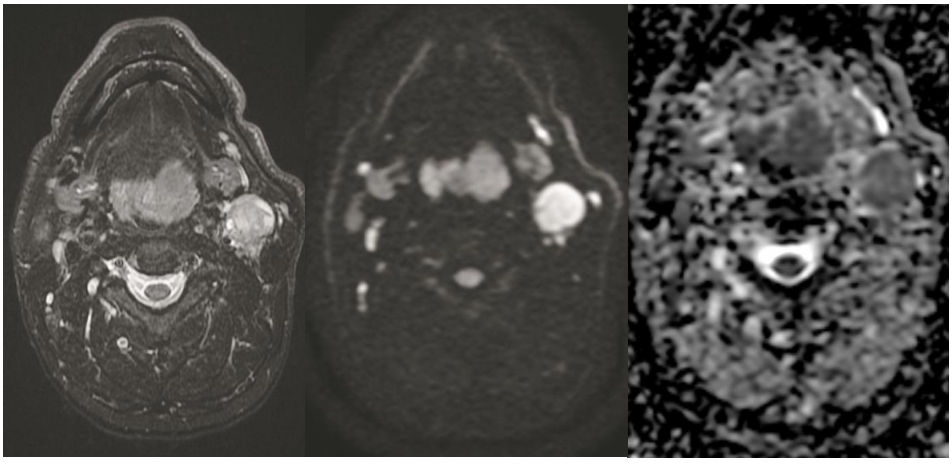


Figure 5.
An example of a DW-MRI image depicting a left base of tongue carcinoma with an ipsilateral lymph node metastasis. Axial STIR MR image on the left, axial DWI b=1000-image in the middle and axial DWI ADC-map on the right.

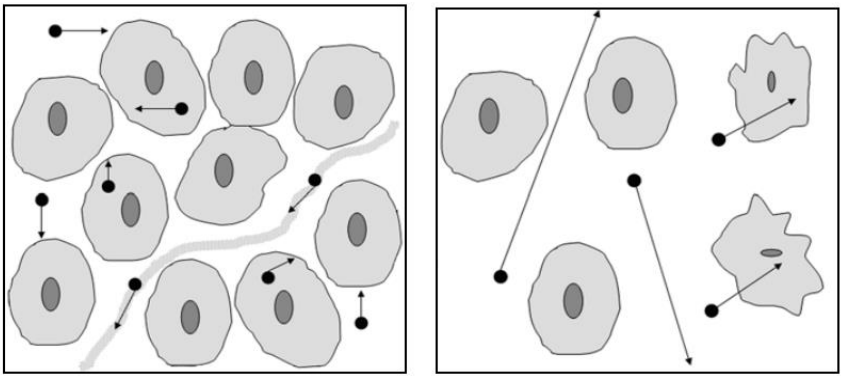


Figure 6. Diffusion of water protons [86]
In a highly cellular environment (left), water diffusion is restricted because of reduced extracellular space and by cell membranes, which act as barrier to water movement. In a less cellular environment (right), relative increase in extracellular space and defective cell membranes allows water diffusion.

DW-MRI uses two additional equally large but opposite motion-probing gradients. The first gradient dephases each water proton which is rephased by the second gradient in case of stationary water protons. Proton movement between the first and second pulse results in incomplete rephasing which leads to signal loss in the DWI images related to the amount and speed of the diffusion motion. The b-value of a DWI sequence measures the amount of diffusion weighting applied, determined by the strength, time of applied gradients and duration between the paired gradients. By repeating the sequences and measuring signal intensities with increasing b-values, the progressive signal decay over the images can be quantified using the ADC, showing an inverse correlation with tissue cellularity. Hypercellular tissue will show limited signal decay with increasing b-value, with persistent high signal intensity on DW-images with a high b-value, resulting in low ADC-values. On the contrary, hypocellular tissue will show signal decay with increasing b-value with low signal intensity on DW-images with a high b-value, resulting in high ADC (Figure 7) [84,86].

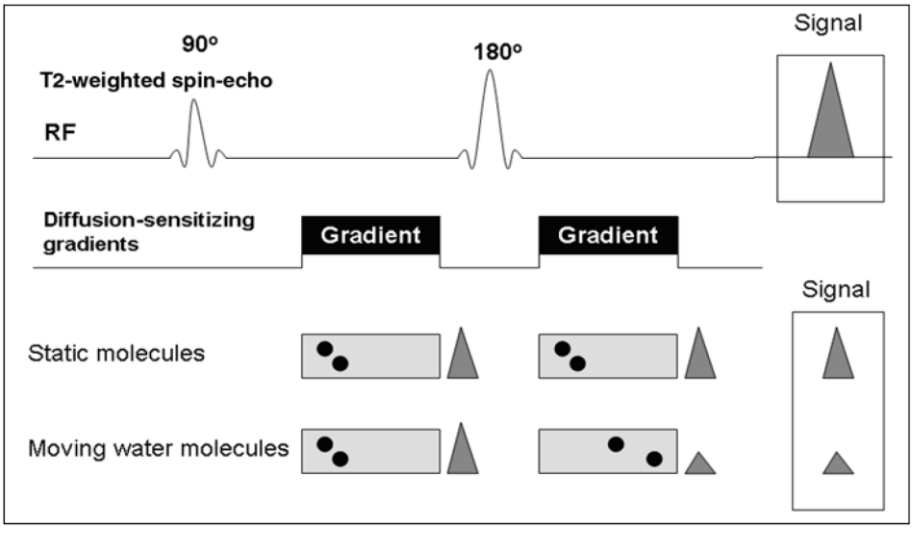


Figure 7. Principle of DW-MRI [86].
Stationary molecules are unaffected by gradients and measured signal intensity is preserved. By contrast, moving water molecules acquire phase information from first gradient, which is not entirely rephased by second gradient, thereby leading to signal loss. Hence, water diffusion is detected as attenuation of measured MR signal intensity. RF= radiofrequency pulse.

Acquisition

Up to now, no uniform imaging protocol for DW-MRI is available. The comparability and quality of DWI-images is influenced by various acquisition parameters such as type of scanner, field strength, choice of b-values and DWI-method (pulse sequence). The selection of b-values is pivotal to implement standardized head and neck DW-MRI protocols. The application of high b-values (b-values between 500 and 1000 sec/mm^2) enables the detection of remote water molecule movements and should be chosen carefully as (too) high b-values result in lower signal-to-noise-ratio (SNR). Using low b-values (b-values <100 sec/mm^2), fast moving water molecules are visualised, *i.e.* intravascular molecules and fast moving water molecules will show the strongest signal loss [87]. In clinical practice, a combination of three b-values is most often used. This combination of b-values is also important; the use of only low b-values will result in high ADC-values and represents both diffusion and perfusion effects. The use of mainly high b-values results in lower ADC-values and better depicts the true diffusion in the tissue [85].

DW-imaging is susceptible to artefacts, particularly in the inhomogeneous head and neck region, containing a variety of tissues including bone, fat, muscle, glandular tissue and air. Moreover, movement-related problems, like swallowing, breathing, coughing, speaking and jaw movements impede imaging of the head and neck. DW-MRI protocols in HNSCC can be performed with an Echo-planar Imaging (EPI)-sequence or non-EPI sequence. With EPI-DWI, imaging of a relatively large volume with multiple b-values is possible due to a short acquisition time. However, EPI-DWI is sensitive to geometric distortions, which is especially strong near interfaces between soft tissue and air or bone. If artefacts are too detrimental, a non-EPI technique, such as half-fourier acquisition single-shot turbo spin-echo (HASTE) or periodically rotated overlapping parallel lines with enhanced reconstruction (PROPELLER), may be an alternative. However, non-EPI-DWI protocols are time-consuming, especially with the inclusion of multiple b-values. Furthermore, detection of tumoral lesions is more difficult due to a lower SNR. DW-MRI protocols in HNSCC are most commonly performed with EPI-DWI and this is probably the most promising technique in HNSCC [53,59,88].

Applications in head and neck oncology

In the past, mainly anatomical imaging methods such as CT and MRI were performed in the diagnostic work-up of patients with HNSCC because they provide detailed information on the extent of the primary tumor and regional nodal metastases. However, small lesions may remain undetected due to morphological and size-related criteria on conventional imaging. Moreover, differentiation between malignant and aspecific post-treatment changes can pose a difficult dilemma since postradiotherapy changes obscure accurate assessment on these conventional images [46,77]. Although relatively little experienced is gained with DW-MRI so far, clinicians more and more rely on functional imaging with DWI, which can be employed for several clinical purposes. First, a few studies studied the diagnostic accuracy of DW-MRI for detection of primary tumors in HNSCC [89,90], although DWI does not seem to be superior to other techniques such as PET-CT to detect a primary tumor [91]. Second, DW-MRI is increasingly used in staging of the neck for differentiating metastatic from benign lymph nodes and may be more accurate than nuclear imaging, especially in the detection of subcentimeter metastatic lymph nodes [92-94]. As with PET-CT, several studies investigated the value of DW-MRI in the prediction of treatment outcome before or during treatment. ADC-values seem to be a prognostic factor in HNSCC; several studies have shown that tumors with relatively low baseline ADC-values have better outcome than

tumors with high ADC-values [53-56]. For predictive imaging during treatment, DW-MRI is suggested as a predictive factor for tumor response during CRT. An early treatment-induced ADC increase is seen in tumors with complete remission [53,58,59]. Finally, DW-MRI is employed in post-treatment imaging for the purpose of monitoring treatment response and detecting tumor recurrence. Literature suggests that ADC-values are significantly lower in residual tumor [59,95] and tumor recurrences [88,96,97] than in postradiotherapy changes.

Qualitative and quantitative interpretation

The acquisition of adequate conventional images prior to DWI imaging is a prerequisite for performing DW-MRI, ideally with the same slices at the same slice position and orientation to provide anatomical co-registration, as well as observing morphological findings. As with PET-CT, DW-MRI scans can be interpreted qualitatively by means of visual assessment and quantitatively by measuring ADC-values. Qualitative interpretation has similar advantages and disadvantages as with PET-CT. However, no established visual assessment system for DW-MRI is available. In the current literature, qualitatively interpretation is mostly performed by visual assessment of the signal intensity on high b-value images and their corresponding ADC-maps [85,96]. Absence of hyperintensity on b=1000 images are regarded negative for malignant tumor. Hypercellular tissue, like malignant tumor, is characterized by a high signal intensity on b=1000 DW-images compared to the surrounding tissue corresponding to low ADC-values. However, restricted diffusion can also be seen in normal structures in the head and neck area such as Waldeyer's ring because these structures also have high cellularity [98]. Areas of high signal intensity on high b-values images should always be assessed with their corresponding ADC-maps to filter out areas of T2 shine-through. The 'T2 shine through' effect indicates a hyperintensity related to the intrinsic T2-weighting of DW-images and is a result from the long tissue T2-relaxation times. Tissue changes induced by CRT (oedema, necrosis) can be seen as high signal intensity but also as low signal intensity on b=1000 images and usually show high values on the corresponding ADC-map (Table 2) [99] .

Table 2. Qualitative interpretation guidelines for DW-MRI [99]

Signal intensity on high b-value images	***Signal intensity on ADC-map***	***Interpretation***
High	Low	Generally, high cellular tumor; rarely abscess, viscous fluids or blood products
High	High	T2 shine-through; liquefactive necrosis
Low	High	Fluid; necrosis; low cellularity; tumor with gland formation
Low	Low/intermediate	Fibromuscular issues; fat; fibrous tissue with low water content

Quantitative evaluation can be performed through the calculation of ADC-values and is probably less prone to inter- and intra-observer variability compared to visual assessment [87]. Currently, ADC-values are acquired by manually drawing a region of interest (ROI) on the native DWI images over the primary tumor or lymph node metastasis on one section or multiple ROIs on several sections to calculate ADC from the entire lesion. Necrotic components of the lesion of interest should be excluded from the ROI. Basic descriptives, such as mean, median, minimum, maximum, standard deviation and range are computed over the delineations.

Up to now, literature on DWI in oncologic head and neck imaging is conducted in single institutions without modifications in MR imaging systems or imaging protocols. Quantitative imaging parameters, such as ADC-values, can be affected by the selected MR imaging systems, sequences and imaging protocols. However, the reproducibility of ADC-values among different MR imaging systems and among sequences has not been established. Reproducibility should be validated across different systems and centers before multicentre DWI studies can be undertaken.

AIM AND OUTLINE OF THESIS

As outlined in this introduction, *treatment prediction* prior to or *during* CRT in patients with HNSCC is an area of interest in scientific research. Besides treatment prediction, there is also need for improvement of current *response evaluation* after treatment. The purposes of the research described in this thesis were to evaluate the accuracy of two imaging techniques (namely: ^{18}F-FDG-PET-CT and DW-MRI) to predict locoregional residual HNSCC *before* and *during* CRT and to detect locoregional residual disease *after* CRT.

This thesis is composed of three parts. The aim of **part 1 (prediction prior to treatment)** was to investigate an association between pretreatment imaging parameters and the presence of HPV and to explore whether these are independent prognostic factors. The identification of additional pretreatment prognostic tumor characteristics may lead to customized treatment regimens and thereby higher treatment responses in individual patients. Differences in glycolytic characteristics between HPV-negative and HPV-positive OPSCC, as measured with pretreatment ^{18}F-FDG-PET-CT, were evaluated in **chapter 2.** In **chapter 3**, we analysed the association between ADC-values, derived from DW-MRI, and the presence of biologically active HPV in patients with OPSCC.

The potential predictive value of these imaging techniques *early during* CRT on locoregional outcome is explored in **part 2 (prediction during treatment)**. **Chapter 4** describes the evaluation of the potential predictive value of EPI- and HASTE-DWI and ^{18}F-FDG-PET-CT early during chemoradiotherapy on locoregional outcome in HNSCC.

The objective of **part 3 (response evaluation)** was to evaluate current response evaluation strategies and the accuracy of ^{18}F-FDG-PET-CT and DW-MRI to detect locoregional residual disease *after* CRT in HNSCC. A questionnaire on the clinical practice concerning response evaluation after CRT for advanced OPSCC in all eight head and neck cancer institutions of the Dutch Head and Neck Oncology Cooperative Group is presented in **chapter 5**. Thereafter, the accuracy of response evaluation using ^{18}F-FDG-PET-CT, DW-MRI and combined PET-CT/DW-MRI for advanced staged OPSCC was prospectively researched in **chapter 6.** In **chapter 7,** we retrospectively evaluated the accuracy and interobserver variation of these imaging techniques to detect residual lymph node metastases after CRT in advanced staged HNSCC. Finally, an estimation of the cost-effectiveness of the imaging techniques in the selection of patients for response evaluation in comparison with an examination under general anesthesia, is performed in **chapter 8**.

In **chapter 9**, a general discussion and future prospects about the topics in this thesis are presented.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J. Clin.* 2011;61(2):69-90.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer.* 2010;127(12):2893-2917.
3. Braakhuis BJ, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. *Oral Oncol.* 2014;50(7):670-675.
4. Dutch Cancer Registration (Nederlandse Kanker Registratie), Netherlands Comprehensive Cancer Organisation (Integraal Kankercentrum Nederland). Head and neck cancer. 2014.
5. Blot WJ, McLaughlin JK, Winn DM, et al. Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer Res.* 1988;48(11):3282-3287.
6. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol. Biomarkers Prev.* 2005;14(2):467-475.
7. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* 2010;11(8):781-789.
8. Hording U, Nielsen HW, Albeck H, Daugaard S. Nasopharyngeal carcinoma: histopathological types and association with Epstein-Barr Virus. *Eur. J. Cancer B Oral Oncol.* 1993;29B(2):137-139.
9. Sobin LH, Gospodarowicz MK, Wittekind C. Head and Neck Tumours. *TNM Classification of Malignant Tumors by International Union Against Cancer (UICC).* 7th ed. Oxford, United Kingdom: Wiley-Black-well; 2009.
10. Dutch Head and Neck Oncology Cooperative Group, Nederlandse Werkgroep Hoofd-Hals Tumoren (NWHHT). Oral cavity and oropharyngeal cancer. National guideline, version 1.4, 2004. Available at: http://www.nvpc.nl/uploads/stand/3702_02_04_Hoofd.pdf.
11. Klussmann JP, Weissenborn SJ, Wieland U, et al. Prevalence, distribution, and viral load of human papillomavirus 16 DNA in tonsillar carcinomas. *Cancer.* 2001;92(11):2875-2884.
12. D'Souza G, Kreimer AR, Viscidi R, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N. Engl. J. Med.* 2007;356(19):1944-1956.
13. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J. Clin. Oncol.* 2008;26(4):612-619.
14. Johnson-Obaseki S, McDonald JT, Corsten M, Rourke R. Head and neck cancer in Canada: trends 1992 to 2007. *Otolaryngol. Head Neck Surg.* 2012;147(1):74-78.
15. Blomberg M, Nielsen A, Munk C, Kjaer SK. Trends in head and neck cancer incidence in Denmark, 1978-2007: focus on human papillomavirus associated sites. *Int. J. Cancer.* 2011;129(3):733-741.
16. Doobaree IU, Landis SH, Linklater KM, El-Hariry I, Moller H, Tyczynski J. Head and neck cancer in South East England between 1995-1999 and 2000-2004: An estimation of incidence and distribution by site, stage and histological type. *Oral Oncol.* 2009;45(9):809-814.

17. Braakhuis BJ, Visser O, Leemans CR. Oral and oropharyngeal cancer in The Netherlands between 1989 and 2006: Increasing incidence, but not in young adults. *Oral Oncol.* 2009;45(9):e85-e89.
18. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int. J. Cancer.* 2013;132(7):1565-1571.
19. D'Souza G, Zhang HH, D'Souza WD, Meyer RR, Gillison ML. Moderate predictive value of demographic and behavioral characteristics for a diagnosis of HPV16-positive and HPV16-negative head and neck cancer. *Oral Oncol.* 2010;46(2):100-104.
20. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat. Rev. Cancer.* 2011;11(1):9-22.
21. Chernock RD, El-Mofty SK, Thorstad WL, Parvin CA, Lewis JS, Jr. HPV-related nonkeratinizing squamous cell carcinoma of the oropharynx: utility of microscopic features in predicting patient outcome. *Head Neck Pathol.* 2009;3(3):186-194.
22. Thariat J, Badoual C, Faure C, Butori C, Marcy PY, Righini CA. Basaloid squamous cell carcinoma of the head and neck: role of HPV and implication in treatment and prognosis. *J. Clin. Pathol.* 2010;63(10):857-866.
23. Smeets SJ, Brakenhoff RH, Ylstra B, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol.* 2009;31(4):291-300.
24. Braakhuis BJ, Snijders PJ, Keune WJ, et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J. Natl. Cancer Inst.* 2004;96(13):998-1006.
25. Hubbert NL, Sedman SA, Schiller JT. Human papillomavirus type 16 E6 increases the degradation rate of p53 in human keratinocytes. *J. Virol.* 1992;66(10):6237-6241.
26. Klingelhutz AJ, Foster SA, McDougall JK. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature.* 1996;380(6569):79-82.
27. Boyer SN, Wazer DE, Band V. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res.* 1996;56(20):4620-4624.
28. Dyson N, Howley PM, Munger K, Harlow E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science.* 1989;243(4893):934-937.
29. Borggreven PA, Aaronson NK, Verdonck-de Leeuw IM, et al. Quality of life after surgical treatment for oral and oropharyngeal cancer: a prospective longitudinal assessment of patients reconstructed by a microvascular flap. *Oral Oncol.* 2007;43(10):1034-1042.
30. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
31. 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(6):567-578.
32. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.

33. Dutch Head and Neck Oncology Cooperative Group, Nederlandse Werkgroep Hoofd-Hals Tumoren (NWHHT). Head and Neck Cancer. Available at: http://richtlijnendatabase.nl/richtlijn/hoofd-halstumoren/chemoradiatie_en_bio-radiatie_van_hoofd-hals.html#uitgangsvraag, 2015.
34. Gupta T, Agarwal J, Jain S, et al. Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: a randomized controlled trial. *Radiother. Oncol.* 2012;104(3):343-348.
35. Nutting CM, Morden JP, Harrington KJ, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. *Lancet Oncol.* 2011;12(2):127-136.
36. Bonner JA, Harari PM, Giralt J, et al. Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. *Lancet Oncol.* 2010;11(1):21-28.
37. Russi EG, Moretto F, Rampino M, et al. Acute skin toxicity management in head and neck cancer patients treated with radiotherapy and chemotherapy or EGFR inhibitors: Literature review and consensus. *Crit Rev. Oncol. Hematol.* 2015.
38. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* 2010;363(1):24-35.
39. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J. Natl. Cancer Inst.* 2008;100(4):261-269.
40. Rietbergen MM, Brakenhoff RH, Bloemena E, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann. Oncol.* 2013;24(11):2740-2745.
41. Yom SS, Machtay M, Biel MA, et al. Survival impact of planned restaging and early surgical salvage following definitive chemoradiation for locally advanced squamous cell carcinomas of the oropharynx and hypopharynx. *Am. J. Clin. Oncol.* 2005;28(4):385-392.
42. Valentino J, Spring PM, Shane M, Arnold SM, Regine WF. Interval pathologic assessments in patients treated with concurrent hyperfractionated radiation and intraarterial cisplatin (HYPERRADPLAT). *Head Neck.* 2002;24(6):539-544.
43. Lavertu P, Bonafede JP, Adelstein DJ, et al. Comparison of surgical complications after organ-preservation therapy in patients with stage III or IV squamous cell head and neck cancer. *Arch. Otolaryngol. Head Neck Surg.* 1998;124(4):401-406.
44. Newman JP, Terris DJ, Pinto HA, Fee WE, Jr., Goode RL, Goffinet DR. Surgical morbidity of neck dissection after chemoradiotherapy in advanced head and neck cancer. *Ann. Otol. Rhinol. Laryngol.* 1997;106(2):117-122.
45. Van der Putten L, Van den Broek GB, De Bree R, et al. Effectiveness of salvage selective and modified radical neck dissection for regional pathologic lymphadenopathy after chemoradiation. *Head Neck.* 2009;31(5):593-603.
46. Nomayr A, Lell M, Sweeney R, Bautz W, Lukas P. MRI appearance of radiation-induced changes of normal cervical tissues. *Eur. Radiol.* 2001;11(9):1807-1817.
47. Klabbers BM, Lammertsma AA, Slotman BJ. The value of positron emission tomography for monitoring response to radiotherapy in head and neck cancer. *Mol. Imaging Biol.* 2003;5(4):257-270.

48. Manikantan K, Khode S, Dwivedi RC, et al. Making sense of post-treatment surveillance in head and neck cancer: when and what of follow-up. *Cancer Treat. Rev.* 2009;35(8):744-753.
49. Suzuki H, Kato K, Fujimoto Y, et al. 18F-FDG-PET/CT predicts survival in hypopharyngeal squamous cell carcinoma. *Ann. Nucl. Med.* 2013;27(3):297-302.
50. Zhang B, Li X, Lu X. Standardized uptake value is of prognostic value for outcome in head and neck squamous cell carcinoma. *Acta Otolaryngol.* 2010;130(7):756-762.
51. Machtay M, Natwa M, Andrel J, et al. Pretreatment FDG-PET standardized uptake value as a prognostic factor for outcome in head and neck cancer. *Head Neck.* 2009;31(2):195-201.
52. Allal AS, Dulguerov P, Allaoua M, et al. Standardized uptake value of 2-[(18)F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy. *J. Clin. Oncol.* 2002;20(5):1398-1404.
53. Kim S, Loevner L, Quon H, et al. Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. *Clin. Cancer Res.* 2009;15(3):986-994.
54. Srinivasan A, Chenevert TL, Dwamena BA, et al. Utility of pretreatment mean apparent diffusion coefficient and apparent diffusion coefficient histograms in prediction of outcome to chemoradiation in head and neck squamous cell carcinoma. *J. Comput. Assist. Tomogr.* 2012;36(1):131-137.
55. Chawla S, Kim S, Dougherty L, et al. Pretreatment diffusion-weighted and dynamic contrast-enhanced MRI for prediction of local treatment response in squamous cell carcinomas of the head and neck. *AJR Am. J. Roentgenol.* 2013;200(1):35-43.
56. Hatakenaka M, Nakamura K, Yabuuchi H, et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81(2):339-345.
57. Eerenstein SEJ, Van der Putten L, De Bree R. Salvage surgery following chemo-radiation for head and neck cancer: does it yield satisfactory results? *Radiotherapy Oncology.* 2007.
58. Vandecaveye V, Dirix P, De Keyzer F, et al. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. *Eur. Radiol.* 2010;20(7):1703-1714.
59. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
60. Brun E, Kjellén E, Tennvall J, et al. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. *Head Neck.* 2002;24(2):127-135.
61. Farrag A, Ceulemans G, Voordeckers M, Everaert H, Storme G. Can 18F-FDG-PET response during radiotherapy be used as a predictive factor for the outcome of head and neck cancer patients? *Nucl. Med. Commun.* 2010;31(6):495-501.
62. Hentschel M, Appold S, Schreiber A, et al. Early FDG PET at 10 or 20 Gy under chemoradiotherapy is prognostic for locoregional control and overall survival in patients with head and neck cancer. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(7):1203-1211.

63. Reiser MF, Hricak H, Knauth M. In: Peller P, Subramaniam R, Guermazi A, eds. *PET-CT and PET-MRI in oncology - a practical guide*: Springer; 2012.
64. Buchbender C, Heusner TA, Lauenstein TC, Bockisch A, Antoch G. Oncologic PET/MRI, part 1: tumors of the brain, head and neck, chest, abdomen, and pelvis. *J. Nucl. Med.* 2012;53(6):928-938.
65. Cook GJ, Fogelman I, Maisey MN. Normal physiological and benign pathological variants of 18-fluoro-2-deoxyglucose positron-emission tomography scanning: potential for error in interpretation. *Semin. Nucl. Med.* 1996;26(4):308-314.
66. Lindholm H, Johansson O, Jonsson C, Jacobsson H. The distribution of FDG at PET examinations constitutes a relative mechanism: significant effects at activity quantification in patients with a high muscular uptake. *Eur. J. Nucl. Med. Mol. Imaging.* 2012;39(11):1685-1690.
67. Boellaard R, Oyen WJ, Hoekstra CJ, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur. J. Nucl. Med. Mol. Imaging.* 2008;35(12):2320-2333.
68. Parysow O, Mollerach AM, Jager V, Racioppi S, San RJ, Gerbaudo VH. Low-dose oral propranolol could reduce brown adipose tissue F-18 FDG uptake in patients undergoing PET scans. *Clin. Nucl. Med.* 2007;32(5):351-357.
69. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(1):181-200.
70. Johnson JT, Branstetter BF. PET/CT in head and neck oncology: State-of-the-art 2013. *Laryngoscope.* 2014;124(4):913-915.
71. Senft A, De Bree R, Hoekstra OS, et al. Screening for distant metastases in head and neck cancer patients by chest CT or whole body FDG-PET: a prospective multicenter trial. *Radiother. Oncol.* 2008;87(2):221-229.
72. Krabbe CA, Pruim J, Dijkstra PU, et al. 18F-FDG PET as a routine posttreatment surveillance tool in oral and oropharyngeal squamous cell carcinoma: a prospective study. *J. Nucl. Med.* 2009;50(12):1940-1947.
73. Haerle SK, Schmid DT, Ahmad N, Hany TF, Stoeckli SJ. The value of (18)F-FDG PET/CT for the detection of distant metastases in high-risk patients with head and neck squamous cell carcinoma. *Oral Oncol.* 2011;47(7):653-659.
74. Johansen J, Buus S, Loft A, et al. Prospective study of 18FDG-PET in the detection and management of patients with lymph node metastases to the neck from an unknown primary tumor. Results from the DAHANCA-13 study. *Head Neck.* 2008;30(4):471-478.
75. Zhu L, Wang N. 18F-fluorodeoxyglucose positron emission tomography-computed tomography as a diagnostic tool in patients with cervical nodal metastases of unknown primary site: a meta-analysis. *Surg. Oncol.* 2013;22(3):190-194.
76. Gupta T, Master Z, Kannan S, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(11):2083-2095.
77. De Bree R, Van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncol.* 2009;45(4-5):386-393.
78. Troost EG, Schinagl DA, Bussink J, et al. Innovations in radiotherapy planning of head and neck cancers: role of PET. *J. Nucl. Med.* 2010;51(1):66-76.

79. Dibble EH, Alvarez AC, Truong MT, Mercier G, Cook EF, Subramaniam RM. 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J. Nucl. Med.* 2012;53(5):709-715.
80. Tang C, Murphy JD, Khong B, et al. Validation that metabolic tumor volume predicts outcome in head-and-neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;83(5):1514-1520.
81. Castaldi P, Rufini V, Bussu F, et al. Can "early" and "late" 18F-FDG PET-CT be used as prognostic factors for the clinical outcome of patients with locally advanced head and neck cancer treated with radio-chemotherapy? *Radiother. Oncol.* 2012;103(1):63-68.
82. Marcus C, Ciarallo A, Tahari AK, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins Criteria)-interreader reliability, accuracy, and survival outcomes. *J. Nucl. Med.* 2014;55(9):1411-1416.
83. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2015;42(2):328-354.
84. Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol.* 2003;45(3):169-184.
85. Thoeny HC, De Keyzer F, King AD. Diffusion-weighted MR imaging in the head and neck. *Radiology.* 2012;263(1):19-32.
86. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. *AJR Am. J. Roentgenol.* 2007;188(6):1622-1635.
87. Vandecaveye V, De Keyzer F, Dirix P, Lambrecht M, Nuyts S, Hermans R. Applications of diffusion-weighted magnetic resonance imaging in head and neck squamous cell carcinoma. *Neuroradiology.* 2010;52(9):773-784.
88. Vandecaveye V, De Keyzer F, Nuyts S, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):960-971.
89. Wang J, Takashima S, Takayama F, et al. Head and neck lesions: characterization with diffusion-weighted echo-planar MR imaging. *Radiology.* 2001;220(3):621-630.
90. Srinivasan A, Dvorak R, Perni K, Rohrer S, Mukherji SK. Differentiation of benign and malignant pathology in the head and neck using 3T apparent diffusion coefficient values: early experience. *AJNR Am. J. Neuroradiol.* 2008;29(1):40-44.
91. Driessen JP, van Kempen PM, van der Heijden GJ, et al. Diffusion-weighted imaging in head and neck squamous cell carcinomas: a systematic review. *Head Neck.* 2015;37(3):440-448.
92. Vandecaveye V, De Keyzer F, Vander Poorten V, et al. Head and neck squamous cell carcinoma: value of diffusion-weighted MR imaging for nodal staging. *Radiology.* 2009;251(1):134-146.
93. Dirix P, Vandecaveye V, De Keyzer F, et al. Diffusion-weighted MRI for nodal staging of head and neck squamous cell carcinoma: impact on radiotherapy planning. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;76(3):761-766.
94. De Bondt RB, Hoeberigs MC, Nelemans PJ, et al. Diagnostic accuracy and additional value of diffusion-weighted imaging for discrimination of malignant cervical lymph nodes in head and neck squamous cell carcinoma. *Neuroradiology.* 2009;51(3):183-192.

95. Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1098-1107.
96. Tshering Vogel DW, Zbaeren P, Geretschlaeger A, Vermathen P, De Keyzer F, Thoeny HC. Diffusion-weighted MR imaging including bi-exponential fitting for the detection of recurrent or residual tumour after (chemo)radiotherapy for laryngeal and hypopharyngeal cancers. *Eur. Radiol.* 2013;23(2):562-569.
97. Abdel Razek AA, Kandeel AY, Soliman N, et al. Role of diffusion-weighted echo-planar MR imaging in differentiation of residual or recurrent head and neck tumors and posttreatment changes. *AJNR Am. J. Neuroradiol.* 2007;28(6):1146-1152.
98. Kolff-Gart AS, Pouwels PJ, Noij DP, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. *AJNR Am. J. Neuroradiol.* 2015;36(2):384-390.
99. Padhani AR. Diffusion magnetic resonance imaging in cancer patient management. *Semin. Radiat. Oncol.* 2011;21(2):119-140.

Part 1: Prediction prior to treatment

CHAPTER 2

INTERACTION OF QUANTITATIVE ^{18}F-FDG-PET-CT IMAGING PARAMETERS AND HUMAN PAPILLOMAVIRUS STATUS IN OROPHARYNGEAL SQUAMOUS CELL CARCINOMA

C.S. Schouten
S. Hakim
R. Boellaard
E. Bloemena
P.A. Doornaert
B.I. Witte
B.J.M. Braakhuis
R.H. Brakenhoff
C.R. Leemans
O.S. Hoekstra
R. de Bree

Head Neck 2016; 38: 529-5

ABSTRACT

Background. Patients with human papillomavirus (HPV)-positive oropharyngeal squamous cell carcinoma (OPSCC) have a better survival than HPV-negative OPSCC. ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography (^{18}F-FDG-PET-CT) may also provide prognostic information. We evaluated glycolytic characteristics in HPV-negative and HPV-positive OPSCC.

Methods. Forty-four patients underwent pretreatment ^{18}F-FDG-PET-CT. Standardized Uptake Values (SUV) and metabolic active tumor volume (MATV) were determined for primary tumors. HPV-status was determined with p16 immunostaining, followed by high-risk HPV DNA detection on the positive cases.

Results. Twenty-seven patients were HPV-positive (61.4%). Median MATV was 2.8 millilitres (ml) [1.6-5.3] for HPV-positive and 6.1 ml [4.5-21.2] for HPV-negative tumors (p<0.001). SUV-values are volume dependent (partial volume effect), therefore, MATV was included as covariate in multivariate analysis. The maximum SUV in HPV-positive was 3.9 units lower than in HPV-negative tumors (p=0.01).

Conclusion. ^{18}F-FDG-PET-CT parameters are lower in HPV-positive than in HPV-negative patients. Low pretreatment SUV-values in HPV-positive OPSCC may be explained by HPV-induced tumor changes.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide [1]. Tobacco smoking and alcohol consumption are known risk factors for the development of HNSCC, but over the past several years, human papillomavirus (HPV) has been recognized as a major etiological factor in the development of particularly squamous cell carcinomas in the oropharynx (OPSCC). HPV-positive and HPV-negative OPSCC are distinct disease entities [2]. Patients with HPV-positive OPSCC show more favourable treatment response rates and prognosis as compared with HPV-negative OPSCC, despite presenting with regionally more advanced disease [3,4]. In addition, HPV-related tumors have distinct histological features [5,6] and have a different genetic route to cancer [7,8]. In this context, traditional prognostic factors such as tumor size and lymph node invasion provide insufficient risk classification. Identification of other prognostic tumor characteristics may lead to individually customized treatment and thereby higher responses to treatment and less treatment-induced side effects.

Positron emission tomography-computed tomography (PET-CT) combines functional imaging with anatomical localisation. ^{18}Fluoro-2-deoxyglucose (^{18}F-FDG), a radiolabelled glucose analogue, is the most widely used tracer in oncologic PET studies and enables imaging of metabolically active tissues. Due to increased glycolytic activity, higher concentrations of ^{18}F-FDG accumulate in malignant tumors. The standardised uptake value (SUV) is a semi-quantitative measurement of tracer uptake. It has been suggested that patients with high pretreatment SUV-values generally have a less favourable outcome [9-12], so it seems that a high rate of glucose metabolism is a characteristic of aggressive tumor behaviour. New functional imaging parameters, such as metabolic active tumor volume (MATV) and total lesion glycolysis (TLG), also have been shown to be of prognostic value in OPSCC and may be additional biomarkers [13-16]. MATV is the ^{18}F-FDG-avid tumor volume and TLG combines SUV_{mean} and MATV.

HPV-positive tumors are genetically different from HPV-negative OPSCC [7,8] but it is largely unknown how they behave with regard to glucose metabolism and whether this plays a role in prognosis [13,17,18]. Some conflicting results have been reported on the relation between ^{18}F-FDG-PET-CT parameters and HPV-status [13,17-19]. We studied pretreatment ^{18}F-FDG-PET-CT with a uniform dedicated head and neck protocol in a group of well-defined HPV-positive tumors and compared this with a group of HPV-negative OPSCC. HPV-positivity was defined according a validated algorithm, p16 staining followed by high-risk HPV-DNA detection [20,21]. The objective of this study was to provide supplemental information about differences in glycolytic

characteristics as measured with pretreatment ^{18}F-FDG-PET-CT between HPV-negative and HPV-positive OPSCC.

MATERIALS AND METHODS

Patients and study design

Between January 2010 and December 2013, all consecutive patients with a histopathologically proven OPSCC were retrospectively screened for the following inclusion criteria: patients who had undergone a dedicated head and neck ^{18}F-FDG-PET-CT imaging for tumor staging and patients with a >T1 oropharyngeal tumor because it was expected that a reliable volume of interest (VOI) could not be drawn in patients with a T1 tumor. This retrospective study was approved by the institutional review board of the VU University Medical Center, with waiver of informed consent (data were analysed anonymously). Forty-seven patients with a pretreatment ^{18}F-FDG-PET-CT could be evaluated. Medical records were reviewed for clinical characteristics, including smoking and alcohol intake, TNM-stage and oropharyngeal subsite. Three patients were excluded because the glucose level could not be retrieved. Thus, 44 patients were enrolled in the analysis.

HPV analysis

HPV testing was performed with our previously defined and validated test algorithm for HPV detection [20,21]. In short, formalin-fixed and paraffin-embedded tumor tissue was stained by immunohistochemistry for p16 (product of the *CDKN2A*) and on the p16-immunopositive cases high-risk HPV-DNA was detected with GP5+/6+ polymerase chain reaction (PCR). Only the cases that were positive in the latter assay as well, were classified as HPV-positive.

^{18}F-FDG-PET-CT

All patients fasted for at least 6 hours. Mean serum glucose levels were 5.9 mmol/l (range from 4.1 to 10.2 mmol/l). PET-CT was started 63.8 ± 5.6 minutes after intravenous injection of 149-386 mega Becquerel of ^{18}F-FDG, depending on the body mass index [22]. PET-CT was performed using a dedicated head and neck protocol (scan trajectory jugular notch-orbit; arms down), using an integrated PET-CT system (Gemini TF-64CT; Philips Healthcare, Cleveland, OH, USA; 3D-mode; 4 min emission scans/bed position). Low-dose CT

scanning was performed with 120 kV and 50 mAs prior to emission scanning for attenuation correction and anatomical localisation of ^{18}F-FDG avid lesions. PET-CT data were reconstructed using a time of flight row-action maximum likelihood algorithm, as implemented by the vendor. Final image matrix size equals 288x288 with a voxel size of 2x2x2 mm. Post reconstruction image resolution equalled 5 mm FWHM.

The primary tumor VOIs for quantitative parameters measurements were drawn using a 3-dimensional (3D) region-growing algorithm, implemented with software developed in-house at the VU University Medical Center [23], by C. Schouten supervised by a nuclear medicine physician (O. Hoekstra). For each VOI, maximum SUV (SUV_{max}), peak SUV (SUV_{peak}), average SUV using an adaptive threshold of 50% (SUV_{mean}), MATV and TLG (calculated as product of SUV_{mean} and MATV) were obtained. SUVs were normalized for lean body mass and serum glucose. Cumulative SUV-volume histograms (CSH) were investigated, representing % of total volume with an SUV above a threshold from 0 to 100% of SUV_{max}. The area under the curve (AUC) is a quantitative index of uptake heterogeneity, with lower values corresponding to higher degrees of tumor heterogeneity [24].

Tumor volume

Primary tumor volume was obtained from CT scans (Discovery CT590RT, General Electric, Milwaukee, Wisconsin, USA) (with 95 mL intravenous contrast injected at rate of 1 mL/s), used for radiotherapy treatment planning. Time between the staging ^{18}F-FDG-PET-CT and the radiotherapy planning CT was on average 1-2 weeks. Patients underwent CT scanning in the supine position and were positioned in a 5 point fixation head mask (Posicast® Thermoplastics, Civco Medical Solutions, Reeuwijk, the Netherlands) and scanned in the RT position. A scan was made from the top of the skull to the aortic arch with a slice thickness of 2.5 mm. The gross tumor volume (GTV) using the CT images was delineated, slice by slice, by head and neck radiation oncologists from our hospital. This CT scan was co-registrated with a diagnostic MRI scan to ensure optimal delineation. All delineated contours were added and a resulting volume was calculated (Eclipse treatment planning system, Varian Medical Systems, Palo Alto, CA, USA).

In three patients, the time interval between the staging ^{18}F-FDG-PET-CT and radiotherapy planning CT was delayed due to the administration of induction chemotherapy (n=2) or due to a complication during percutaneous endoscopic gastrostomy (PEG) tube placement (n=1). One patient died before

planning CT was made. In these patients, tumor volume was obtained from axial STIR MRI-images with 4 mm sections. Contours were manually drawn by S. Hakim, under supervision of a head-and-neck radiologist (P. de Graaf), around the border of the primary tumor at each slice position to measure total tumor volume.

Statistical analysis

Statistical analyses were performed using SPSS software package (version 20.0; IBM Corp., Armonk, NY, USA). The level of significance was set at $p<0.05$ and hypotheses were tested two-sided. Analysis for differences in patient characteristics between the HPV-positive and HPV-negative group were performed with the Pearson Chi-square (χ^2) test for categorical data. Bonferroni correction was used to compare subgroups. The Student's *t* test was used to compare continuous data in case of a normal distribution and the Mann-Whitney *U* test was used in case of a non-normal distribution. We tested whether there was a significant association between PET-CT parameters and HPV-status in a multivariate analysis to look at the association, after adjusting for other explanatory variables.

RESULTS

Patient characteristics

The total study group consisted of 39 men and 5 women, with a mean age at the time of diagnosis of 61.0 years (range 46-78). Patient characteristics by HPV-status are shown in Table 1. HPV-positive patients were more likely to present with a T2 primary tumor and N-positive disease. HPV-negative patients were more likely to have a history of heavy smoking and excessive alcohol consumption.

Table 1. General patient characteristics

Characteristic	*HPV-positive* *n=27 (61.4%)*	*HPV-negative* *n=17 (38.6%)*	*p*-value
Gender			*p*=0.95
Male	24 (88.9)	15 (88.2)	
Female	3 (11.1)	2 (11.8)	
Age at diagnosis, years			*p*=0.14
Mean ± SD	59.7 ± 7.5	63.0 ± 6.4	
Oropharyngeal subsite			*p*=0.54
Tonsil	12 (44.4)	6 (35.3)	
Base of tongue	12 (44.4)	7 (41.2)	
Oropharynx nos	3 (11.1)	4 (23.5)	
Smoking			*p*=0.002
Never (0-5 pack years)	12 (44.4)	0 (0)	
Moderate (6-24 pack years)	6 (22.2)	3 (17.6)	
Heavy (>24 pack years)	9 (33.3)	14 (82.4)	
Alcohol consumption			*p*=0.005
Never (0)	4 (14.8)	0 (0)	
Moderate (1-149 unit years)	20 (74.1)	8 (47.1)	
Heavy (>149 unit years)	3 (11.1)	9 (52.9)	
T-classification			*p*<0.001
T2	18 (66.7)	1 (5.9)	
T3-4	9 (33.3)	16 (94.1)	
N-classification			*p*=0.04
N0	0 (0)	3 (17.6)	
N1-3	27 (100)	14 (82.4)	

Smoking was defined in pack years (1 pack year=20 cigarettes a day during 1 year)
Alcohol consumption was defined in unit years (1 unit year=one alcohol-containing consumption a day during 1 year).
Abbreviations: HPV; human papillomavirus, oropharynx nos; oropharynx not otherwise specified.

Association between HPV-status and imaging parameters

Representative ^{18}F-FDG-PET-CT images from an HPV-negative and an HPV-positive patient are shown in Figures 1 and 2. Table 2 summarizes ^{18}F-FDG-PET-CT imaging parameters in relation to HPV-status. HPV-positive patients had significantly smaller tumor volumes (median 9.6 cm^3 [interquartile range 6.4-22.2]) than HPV-negative patients (median 30.6 cm^3 [22.2-85.5]) (p<0.001). The median SUV_{max} measurements of the primary tumor site for the HPV-positive and HPV-negative groups were 8.0 [5.9-11.8] and 13.6 [9.8-17.2], respectively (p=0.002) (Figure 3). Median MATV was 2.8 [1.6-5.3] ml for HPV-positive and 6.1 [4.5-21.2] ml for HPV-negative tumors (p<0.001). The median

total lesion glycolysis for HPV-positive and HPV-negative was 17.8 [7.1-40.1] and 55.8 [42.3-211.7], respectively (p<0.001). Comparing AUC-CSH, a marker for tumor heterogeneity, HPV-positive primary tumors are significantly less heterogeneous than HPV-negative tumors, 0.91 [0.84-0.96] versus 0.82 [0.78-0.86] (p=0.003).

Some ^{18}F-FDG-PET-CT parameters are volume dependent because of partial volume effects (PVE). Therefore, a more objective assessment of the glucose metabolism parameters is obtained after including MATV in the statistical analysis. MATV was found to be significantly associated with SUV_{max} (spearman's rho= 0.64, p<0.001), SUV_{mean} (spearman's rho= 0.63, p<0.001) and SUV_{peak} (spearman's rho= 0.68, p<0.001). In univariate analysis, the SUV_{max} in HPV-positive was 5.1 units lower than in HPV-negative ones. For SUV_{peak}, a mean difference of 4.1 was found. In multivariate analysis, including MATV as covariate, the SUV_{max} in HPV-positive tumors was 3.9 units (p=0.01) lower than in HPV-negative ones. The mean difference in SUV_{peak} in multivariate analysis was 3.3 (p=0.008).

Table 2. Imaging parameters in relation to HPV-status

^{18}F-FDG-PET-CT	***HPV-positive*** *n=27 (61.4%)*	***HPV-negative*** *n=17 (38.6%)*	***p-value***
Tumor volume* (cm^3)	9.6 [6.4-22.2]	30.6 [22.2-85.5]	p<0.001
SUV_{max}	8.0 [5.9-11.8]	13.6 [9.8-17.2]	p=0.002
SUV_{mean}	5.9 [4.5-8.5]	9.9 [6.9-11.9]	p=0.003
SUV_{peak}	7.1 [5.6-9.6]	11.7 [9.1-14.3]	p=0.001
MATV (ml)	2.8 [1.6-5.3]	6.1 [4.5-21.2]	p<0.001
TLG	17.8 [7.1-40.1]	55.8 [42.3-211.7]	p<0.001
AUC-CSH (%)	0.91 [0.84-0.96]	0.82 [0.78-0.86]	p= 0.003

* Measured with contrast-enhanced CT

Values are presented as median [interquartile range].

Abbreviations: AUC-CSH; area under the curve- cumulative SUV-volume histograms, HPV; human papillomavirus , MATV; metabolic active tumor volume SUV; standardized uptake value, TLG; total lesion glycolysis; ^{18}F-FDG-PET-CT; ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography.

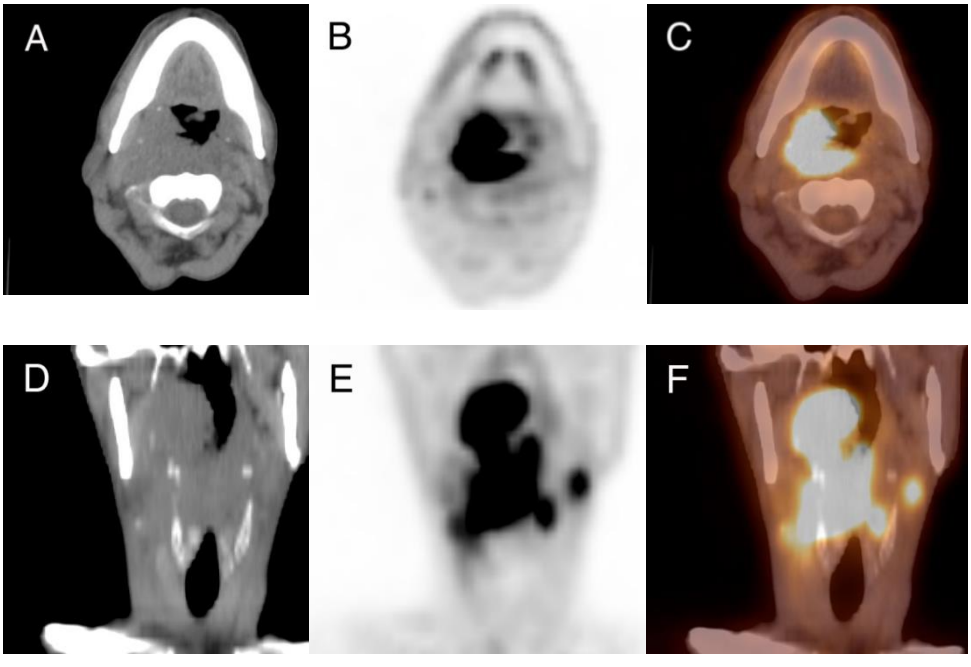


Figure 1.
A 54-year old patient was diagnosed with an HPV-negative OPSCC in the right palatine tonsil. Axial and coronal ^{18}F-FDG-PET-CT images before treatment showing increased ^{18}F-FDG uptake in the primary tumor. **(A)** Axial CT, **(B)** axial PET, **(C)** axial fused, **(D)** coronal CT, **(E)** coronal PET and **(F)** coronal fused. A VOI was drawn for the primary lesion; SUV_{mean} 12.5, SUV_{max} 18.6, SUV_{peak} 13.6, MATV 42.2 and TLG 528.3.

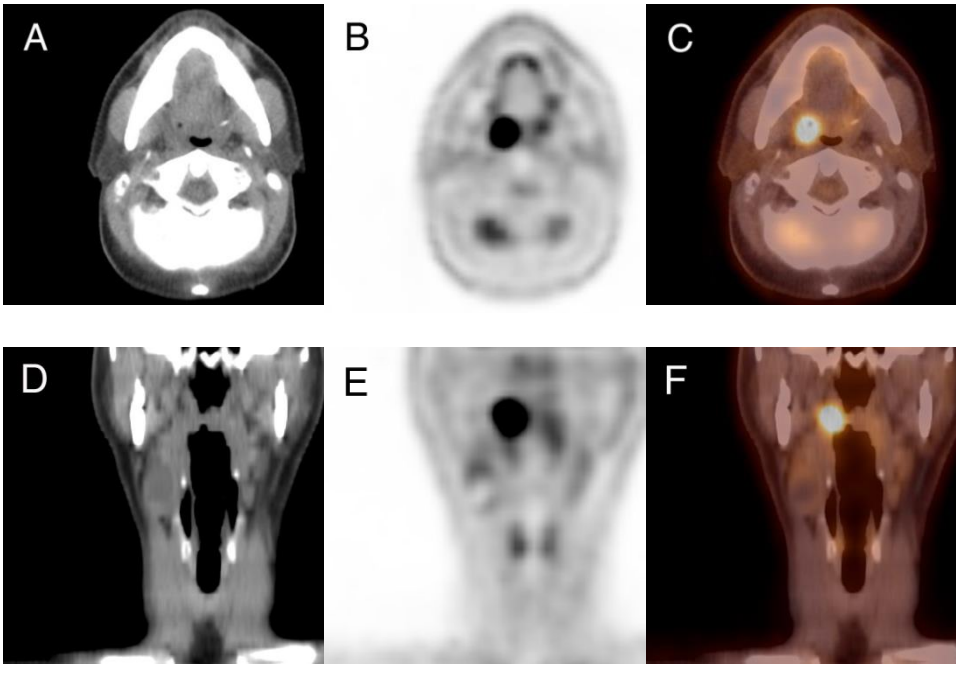


Figure 2.
A 48-year old patient was diagnosed with an HPV-positive OPSCC in the right palatine tonsil. Axial and coronal ^{18}F-FDG-PET-CT images before treatment showing increased ^{18}F-FDG uptake in the primary tumor. **(A)** Axial CT, **(B)** axial PET, **(C)** axial fused, **(D)** coronal CT, **(E)** coronal PET and **(F)** coronal fused. A VOI was drawn for the primary lesion; SUV_{mean} 7.6, SUV_{max} 10.3, SUV_{peak} 9.6, MATV 2.8 and TLG 21.6.

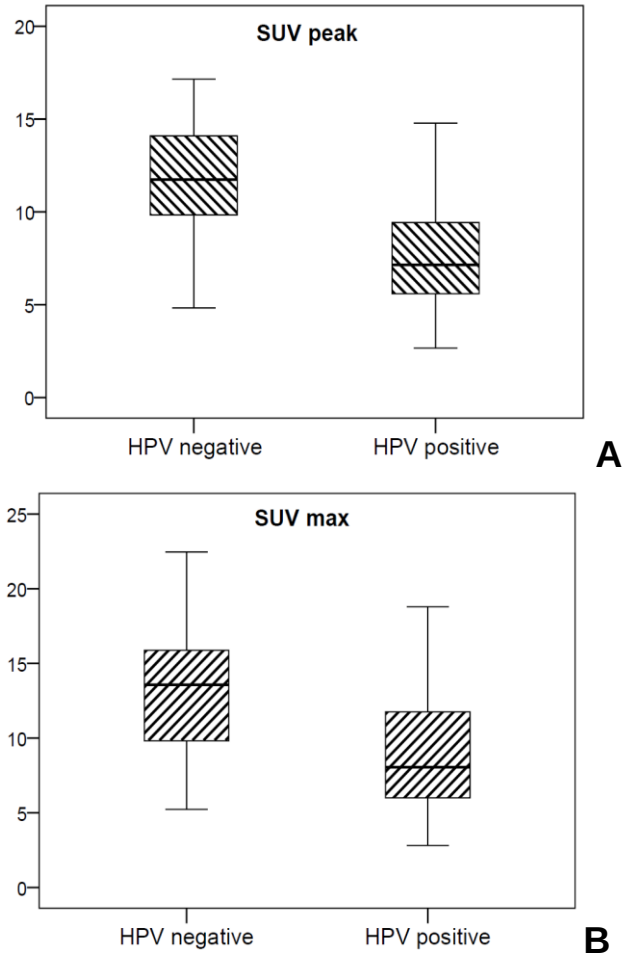


Figure 3.
Association between SUV_{peak} **(A)** and SUV_{max} **(B)** with HPV-status in patients with OPSCC. SUV_{peak} in HPV-positive OPSCCC= 7.1 [5.6-9.6] and 11.7 [9.1-14.3] in HPV-negative OPSCC (p=0.001). SUV_{max} in HPV-positive OPSCC is 8.0 [5.9-11.8] and 13.6 [9.8-17.2] in HPV-negative OPSCC (p=0.002).

DISCUSSION

HPV-related squamous cell carcinoma of the oropharynx is a clinically and histopathologically distinct disease entity [2,25,26]. Patients with HPV-positive OPSCC show higher response rates to chemoradiotherapy and have a better overall survival compared to HPV-negative OPSCC [3,4]. Traditional prognostic factors such as tumor size and lymph node invasion are inadequate to classify patients into risk groups. Therefore, the identification of other independent tumor characteristics that may select patients for tailored treatment regimens becomes of interest. Our results showed significantly lower glycolytic parameters and decreased tumor heterogeneity as measured with ^{18}F-FDG-PET-CT in HPV-positive OPSCC as compared to HPV-negative ones.

A better overall survival after CRT is seen in patients with an HPV-positive OPSCC compared to HPV-negative patients [3,4], mainly due to its increased radiosensitivity [27]. Besides, patients with a complete response to CRT exhibited relatively low pretreatment SUV-values than patients with a partial or non-response to non-surgical treatment [9-12]. Distinct histological features in HPV-positive and HPV-negative OPSCC may be responsible for the different pretreatment SUV-values, potentially due to a different distribution of hypoxic regions within the tumor as a result of other metabolic demands [28]. ^{18}F-FDG-PET-CT reflects metabolic activity which is indirectly related to tumor oxygenation as hypoxia increases glucose metabolism [29]. Considering the radiosensitivity of HPV-related OPSCC, hypothetically the amount of hypoxia would be low and HPV-positive patients should have relatively low pretreatment SUV-values compared to HPV-negative patients. Indeed, in this study we have found that ^{18}F-FDG-PET-CT parameters are significantly lower in patients with HPV-positive OPSCC compared with HPV-negative patients. This hypothesis can be best confirmed by a PET-CT study with a specific hypoxia tracer, such as ^{18}F-fluoroazomycin arabinoside (^{18}F-FAZA).

Heterogeneous histopathology or morphology may be more accurately reflected by a quantitative index of uptake heterogeneity, such as AUC-CSH, than other metabolic PET-CT parameters. Probably, factors such as necrosis and hypoxia within a tumor contribute to tumor heterogeneity in ^{18}F-FDG-uptake [24,30]. As HPV-positive and HPV-negative OPSCC are morphologically distinct diseases, it is of interest to quantify tumor heterogeneity in ^{18}F-FDG uptake. We found that primary tumors of HPV-positive patients are more homogeneous than tumors of HPV-negative patients, using AUC-CSH as a heterogeneity index. A previous study by Tahari et al. [19] in OPSCC also showed that HPV-negative primaries are more heterogeneous by comparison of the ratio SUV_{max} by SUV_{mean}. These results suggest that HPV-negative

OPSCC contain more areas with necrosis and hypoxia, contributing to the reduced radiosensitivity of HPV-negative tumors.

Joo et al. found a significant difference between primary tumor SUV_{max} and HPV-status in 78 patients with OPSCC and concluded that ^{18}F-FDG-PET-CT could be used as an adjuvant diagnostic tool for determination of HPV-status [17]. Tahari et al. also showed that HPV-negative primary tumors have significantly higher metabolic rates, such as SUV_{max}, SUV_{peak} and SUV_{mean}, as compared with HPV-positive patients. In nodal disease, statistical significance was not found for the same parameters [19]. Our results are in accordance with results of these previous studies. However, two methodological comments can be placed. First, accurate quantification of ^{18}F-FDG-uptake is size dependent, because of the partial volume effect (PVE). This process is affected by tumor size and shape, spatial resolution of the scanner, background activity in surrounding tissue, image sampling and voxel size. The PVE effect is most prominent in small tumors and results in an underestimation of the accumulation of radioisotope in tumor [31]. HPV-positive tumors typically present with a small primary tumor at diagnosis [3,4]. Previous mentioned studies did not take the PVE into account. Statistical analysis in our study showed a significant association of MATV with SUV-values. We therefore tested the association between HPV-status and PET-CT imaging parameters using a multivariate analysis as well, including MATV as covariate. In multivariate analysis, SUV_{max} and SUV_{peak} remained significantly associated with HPV-status. Second, in our institute we use a validated algorithm for HPV-detection: p16 immunostaining, followed by a GP 5+/6 PCR on the p16-immunopostitive cases. This algorithm showed an accuracy of 98% [21]. Joo et al. and Tahari et al. assessed HPV status by high risk in situ hybridization. This technique shows a sensitivity and specificity of 88% [32]. HPV analysis by in situ hybridization could have affected their results, with both an underrepresentation of HPV-positive cases and more false positive patients.

^{18}F-FDG-PET-CT is a diagnostic modality used in head and neck oncology for pretreatment staging, treatment planning and assessment of treatment response and has additional prognostic value. ^{18}F-FDG is often increased in malignant tissues, when compared to healthy tissue, caused by an higher glucose metabolism coinciding with an up-regulated glucose uptake through overexpression of glucose transporters (glut). Increased SUV-values of the primary tumor, using SUV-values, is a poor prognostic factor in HNSCC treated with definitive radiotherapy or concomitant with chemotherapy [10-12]. Also, several studies have reported on MATV and TLG as prognostic imaging markers [14,15]. Recent studies show conflicting results whether ^{18}F-FDG-PET-CT parameters and HPV-status are independently associated with survival outcomes. Cheng et al. studied 60 patients with OPSCC and multivariate

analysis revealed that TLG and HPV-positivity were independently associated with overall survival [13]. However, only 12 patients (20%) were HPV-positive. Garsa et al. showed that MATV is a significant predictor of disease free survival and overall survival in subgroups of patients (p16-positive (n=25) and p16-negative (n=18)) [18]. Although p16 immunohistochemistry is a surrogate marker for an HPV-infection, a considerable number of patients are p16-positive but HPV-negative [20]. On the contrary, in a multivariate analysis with 78 patients from Joo et al., only HPV-positivity and not SUV_{max} was associated with 5-year disease-specific survival [17]. Further long-term follow-up studies with sufficient OPSCC patients who undergo ^{18}F-FDG-PET-CT and HPV analysis can answer the question if ^{18}F-FDG-PET-CT imaging parameters and HPV-status are independent prognostic factors. However, this is unlikely if there is a strong association between these tumor characteristics. This study shows that imaging parameters derived from ^{18}F-FDG-PET-CT are significantly associated with HPV-status in patients with OPSCC.

We acknowledge several limitations to this study. First, due to the retrospective study design that only included patients with a dedicated head and neck ^{18}F-FDG-PET-CT, selection bias may have occurred. Secondly, a limited number of patients with OPSCC was included. Despite these limitations, we found a significant association between HPV-status and ^{18}F-FDG-PET-CT parameters after adjusting for MATV and we believe our study provides additional information to previously reported studies on this subject.

CONCLUSION

Glycolytic parameters from ^{18}F-FDG-PET-CT are significantly lower in HPV-positive as compared with HPV-negative patients. Besides, HPV-positive OPSCC show lower degrees of tumor heterogeneity. In multivariate analysis, including MATV as a covariate, SUV_{max} and SUV_{peak} remained significantly associated with HPV-status. Low pretreatment ^{18}F-FDG-PET-CT parameters, which are associated with a more favourable prognosis, may be explained by HPV-induced tumor changes.

ACKNOWLEDGEMENTS

We thank dr. M.M. Rietbergen (Department of Otolaryngology-Head and Neck Surgery, VU University Medical Center, Amsterdam) for her assistance in retrieving HPV-statuses and the authors thank dr. P. de Graaf (Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam) for supervision of tumor volume measurements.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer.* 2010;127(12):2893-2917.
2. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat. Rev. Cancer.* 2011;11(1):9-22.
3. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J. Natl. Cancer Inst.* 2008;100(4):261-269.
4. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* 2010;363(1):24-35.
5. Thariat J, Badoual C, Faure C, Butori C, Marcy PY, Righini CA. Basaloid squamous cell carcinoma of the head and neck: role of HPV and implication in treatment and prognosis. *J. Clin. Pathol.* 2010;63(10):857-866.
6. Chernock RD, El-Mofty SK, Thorstad WL, Parvin CA, Lewis JS, Jr. HPV-related nonkeratinizing squamous cell carcinoma of the oropharynx: utility of microscopic features in predicting patient outcome. *Head Neck Pathol.* 2009;3(3):186-194.
7. Braakhuis BJ, Snijders PJ, Keune WJ, et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J. Natl. Cancer Inst.* 2004;96(13):998-1006.
8. Smeets SJ, Brakenhoff RH, Ylstra B, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol.* 2009;31(4):291-300.
9. Suzuki H, Kato K, Fujimoto Y, et al. 18F-FDG-PET/CT predicts survival in hypopharyngeal squamous cell carcinoma. *Ann. Nucl. Med.* 2013;27(3):297-302.
10. Zhang B, Li X, Lu X. Standardized uptake value is of prognostic value for outcome in head and neck squamous cell carcinoma. *Acta Otolaryngol.* 2010;130(7):756-762.
11. Allal AS, Dulguerov P, Allaoua M, et al. Standardized uptake value of 2-[(18)F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy. *J. Clin. Oncol.* 2002;20(5):1398-1404.
12. Machtay M, Natwa M, Andrel J, et al. Pretreatment FDG-PET standardized uptake value as a prognostic factor for outcome in head and neck cancer. *Head Neck.* 2009;31(2):195-201.
13. Cheng NM, Chang JT, Huang CG, et al. Prognostic value of pretreatment 18F-FDG PET/CT and human papillomavirus type 16 testing in locally advanced oropharyngeal squamous cell carcinoma. *Eur. J. Nucl. Med. Mol. Imaging.* 2012;39(11):1673-1684.
14. Dibble EH, Alvarez AC, Truong MT, Mercier G, Cook EF, Subramaniam RM. 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J. Nucl. Med.* 2012;53(5):709-715.
15. Lim R, Eaton A, Lee NY, et al. 18F-FDG PET/CT metabolic tumor volume and total lesion glycolysis predict outcome in oropharyngeal squamous cell carcinoma. *J. Nucl. Med.* 2012;53(10):1506-1513.

16. Tang C, Murphy JD, Khong B, et al. Validation that metabolic tumor volume predicts outcome in head-and-neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;83(5):1514-1520.
17. Joo YH, Yoo IR, Cho KJ, Park JO, Nam IC, Kim MS. Preoperative F-FDG PET/CT and high-risk HPV in patients with oropharyngeal squamous cell carcinoma. *Head Neck.* 2013.
18. Garsa AA, Chang AJ, Dewees T, et al. Prognostic value of F-FDG PET metabolic parameters in oropharyngeal squamous cell carcinoma. *J. Radiat. Oncol.* 2013;2(1):27-34.
19. Tahari AK, Alluri KC, Quon H, Koch W, Wahl RL, Subramaniam RM. FDG PET/CT Imaging of Oropharyngeal Squamous Cell Carcinoma: Characteristics of Human Papillomavirus-Positive and -Negative Tumors. *Clin. Nucl. Med.* 2013.
20. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int. J. Cancer.* 2007;121(11):2465-2472.
21. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int. J. Cancer.* 2013;132(7):1565-1571.
22. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(1):181-200.
23. Krak NC, Boellaard R, Hoekstra OS, Twisk JW, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur. J. Nucl. Med. Mol. Imaging.* 2005;32(3):294-301.
24. Van Velden FH, Cheebsumon P, Yaqub M, et al. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(9):1636-1647.
25. Gillison ML, Koch WM, Capone RB, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl. Cancer Inst.* 2000;92(9):709-720.
26. El-Mofty SK, Patil S. Human papillomavirus (HPV)-related oropharyngeal nonkeratinizing squamous cell carcinoma: characterization of a distinct phenotype. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 2006;101(3):339-345.
27. Gray LH, Conger AD, Ebert M, Hornsey S, Scott OC. The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br. J. Radiol.* 1953;26(312):638-648.
28. Bose P, Brockton NT, Dort JC. Head and neck cancer: from anatomy to biology. *Int. J. Cancer.* 2013;133(9):2013-2023.
29. Wijsman R, Kaanders JH, Oyen WJ, Bussink J. Hypoxia and tumor metabolism in radiation oncology: targets visualized by positron emission tomography. *Q. J. Nucl. Med. Mol. Imaging.* 2013;57(3):244-256.
30. Koyasu S, Nakamoto Y, Kikuchi M, et al. Prognostic Value of Pretreatment (18)F-FDG PET/CT Parameters Including Visual Evaluation in Patients With Head and Neck Squamous Cell Carcinoma. *AJR Am. J. Roentgenol.* 2014;202(4):851-858.
31. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J. Nucl. Med.* 2007;48(6):932-945.

32. Schache AG, Liloglou T, Risk JM, et al. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clin. Cancer Res.* 2011;17(19):6262-6271.

CHAPTER 3

QUANTITATIVE DIFFUSION-WEIGHTED MRI PARAMETERS AND HUMAN PAPILLOMA-VIRUS STATUS IN OROPHARYNGEAL SQUAMOUS CELL CARCINOMA

C.S. Schouten
P. de Graaf
E. Bloemena
B.I. Witte
B.J.M. Braakhuis
R.H. Brakenhoff
C.R. Leemans
J.A. Castelijns
R. de Bree

American Journal of Neuroradiology 2015; 36: 763-7

ABSTRACT

Background. Patients with human papillomavirus (HPV)-positive oropharyngeal squamous cell carcinomas (OPSCC) have a better survival than HPV-negative OPSCC. Diffusion-weighted MRI (DW-MRI) characterizes biologically relevant tumor features and the generated Apparent Diffusion Coefficient (ADC) may also provide prognostic information. It is unknown if HPV-status and ADC-values are independent tumor characteristics.

Methods. Forty-four patients underwent pretreatment DW-MRI. ADC for the primary tumors were determined using three b-values in a ROI containing the largest area of solid tumor on a single slice of an axial DW-image. HPV-status was determined with p16 immunostaining, followed by high-risk HPV DNA detection on the positive cases.

Results. Twenty-two patients were HPV-positive (50.0%). ADC-values were not significantly different between HPV-negative (ADC_{mean}= 1.56 [1.18-2.18] x 10^3 mm^2/s) and HPV-positive tumors (ADC_{mean}= 1.46 [1.07-2.16] x 10^3 mm^2/s).

Conclusion. No significant association between ADC and HPV-status was found in OPSCC. Apparently, differences in genetical and histological features between HPV-positive OPSCC and HPV-negative ones do not translate into different ADC-values. Long-term follow-up studies are needed to establish whether ADC has prognostic value and whether this is independent from the HPV-status.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide [1]. Over the last decades, it has been well established that besides tobacco smoking and alcohol consumption, human papillomavirus (HPV) is an important etiological factor in the development of HNSCC, in particular squamous cell carcinomas in the oropharynx (OPSCC). HPV-positive and HPV-negative OPSCC are different disease entities [2]. Patients with HPV-positive OPSCC show higher response rates to treatment and have a better overall survival compared with HPV-negative OPSCC, despite the fact that patients with HPV-positive OPSCC more often present with regionally advanced disease [3,4]. In addition, the genetic route to cancer is different for HPV-positive OPSCC [5,6] and HPV-related tumors have distinct histological features [7,8]. In this context, traditional prognostic factors such as tumor size and lymph node invasion may be insufficient to classify patients into risk groups. Identification of other prognostic tumor characteristics may lead to an improved patient selection and as a result, higher responses to treatment and probably less treatment-induced side effects.

Diffusion-weighted MRI (DW-MRI) is a non-invasive functional imaging modality that characterizes tissue based on the random motion of water molecules, which is mainly influenced by the amount of extracellular space and the presence of cell membranes. Differences in water mobility can be quantified with the apparent diffusion coefficient (ADC): hypocellular tissue with necrosis is characterized by a high ADC, whereas hypercellular tissue (*e.g.* malignancy) is characterized by a low ADC. Hence, parameters from DW-MRI, such as ADC, could indicate biologic dissimilarities. Furthermore, studies in breast cancer have demonstrated that the ADC-value significantly correlates with the expression of biological markers of disease, such as oestrogen receptor and progesterone receptor [9,10]. ADC is a possible prognostic factor in HNSCC; several studies have shown that HNSCC with relatively low pretreatment ADC-values respond better to chemoradiotherapy than tumors with higher pretreatment ADC-values [11-14]. HPV-positive OPSCC often have a different histology, characterized by ovoid to spindle-shaped hyperchromatic cells without keratinisation and without a stromal response [15] and it can by hypothesized that these histological characteristics translates in low pretreatment ADC-values; if so, this would support the general notion that low ADC correlates with a better response to chemoradiotherapy. Limited information is available on the relation between HPV-involvement and ADC-values [16].

In this study we investigated the possible association between ADC-values, derived from DW-MRI, and the presence of biologically active HPV in patients with OPSCC.

METHODS

Patients and study design

The study was approved by the institutional review board of VU University Medical Center. We identified patients with a histopathologically proven OPSCC who had undergone DW-MRI imaging for diagnosis and treatment planning between January 2010 and December 2013. Patients with a T1 oropharyngeal tumor were excluded, because a reliable region of interest (ROI) could not be drawn. Fifty-seven patients could be evaluated. Medical records were reviewed for clinical characteristics, including smoking and alcohol intake, TNM-stage and oropharyngeal subsite. Three patients were excluded due to insufficient quality of the DWI. In two other patients the primary tumor was outside the range of obtained MR diffusion images and therefore excluded. Eight more patients were excluded because DW-MRI was obtained with only two b-values. Thus, 44 patients were enrolled in the DW-MRI analysis.

HPV analysis

HPV testing was performed on all tumors with our previously defined and validated algorithm for HPV detection [17,18]. In short, formalin-fixed and paraffin-embedded tumor tissue was stained by immunohistochemistry for p16 (product of the cyclin-dependent kinase inhibitor 2A (*CDKN2A)*) and on the p16-immunopositive cases high-risk HPV-DNA was detected with general primer (GP) 5+/6+ polymerase chain reaction (PCR). Only the cases that were positive in the latter assay as well, were classified as HPV-positive.

Diffusion-weighted MRI

All MR examinations were performed using a 1.5 Tesla MR system (Signa HDx; GE Healthcare, Milwaukee, Wisconsin, USA) with a head coil combined with a phased array spine and neck coil. After an axial short TI inversion-recovery (STIR)-series with 7-mm sections covering the entire neck area, subsequent images were centered on the area of interest containing the primary tumor and enlarged lymph nodes. Axial images (22 slices of 4-mm

slice thickness and 0.4-mm gap) were obtained with STIR (TR/TE/T1=6600/60/160 ms, 2 averages, in-plane pixel size of 0.7x1.1 mm) and T1-weighted (T1WI) spin-echo (TR/TE=500/14 ms, 2 averages, no fat saturation, in-plane pixel size of 0.5x0.5 mm) before and after the intravenous injection of contrast material: Dotarem (0.2 mL/kg of gadoteric acid; Guerbet, Aulnay-sous Bois, France) or Gadovist (0.1 mL/kg of gadobutrol; Bayer Schering Pharma, Berlin-Wedding, Germany).

DWI using a periodically rotated overlapping parallel line with enhanced reconstruction (PROPELLER)-technique was obtained for 16 slices at the same slice positions as the axial STIR and T1WI. Parameters for PROPELLER-DWI were the following: TR/TE=3500/83.87 ms, in-plane pixel size= 2x2 mm, and b-values=0, 750 and 1000 s/mm^2 (3 averages). ADC maps were calculated on-line or off-line, respectively, by using the software of the scanner.

DW-MRI scans were analysed with AEGIS (Aegis^{TM} Web, version 3.2.4, Hologic, Bedford, MA, USA) that allowed viewing of multiple MRI scans. The primary tumors were first identified on conventional MR images. ADC-values were measured by drawing a ROI on a single slice of an axial high b-value DW-image, containing the largest area of solid tumor, excluding any necrotic regions with the aid of post-contrast MR-images. If an artefact within the lesion was present, a smaller ROI was placed only over the undistorted area of the lesion. Subsequently, the ROIs were copied to the corresponding ADC maps. Two examples (HPV-positive and HPV-negative) of a ROI area are illustrated in Fig. 1 and 2. ADC was obtained as ADC_{min} (lowest tumor voxel value within the ROI) and ADC_{mean} (mean ADC within the ROI). The ROIs were outlined by C. Schouten and confirmed by a board certified head and neck radiologist with five years of experience in head and neck radiology (P. de Graaf). Clinical information was known about TNM stage, but the interpreters were blinded to HPV-status.

Statistical analysis

Statistical analyses were performed using SPSS software package (Version 20.0; IBM Corp., Armonk, NY, USA). The level of significance was set at $p<0.05$ and hypotheses were tested two-sided. Analysis for differences in patient characteristics between the HPV-positive and HPV-negative group were performed with the Pearson Chi-square (χ^2) test for categorical data and Fisher's exact test where appropriate. Bonferroni correction was used to compare subgroups. The Mann-Whitney *U* test was used to compare continuous data.

RESULTS

The total study group consisted of 33 men and 11 women, with a mean age at the time of diagnosis of 58.8 years (range 45–78). Patient characteristics by HPV-status are shown in Table 1. HPV-positive patients were more likely to present with a T2 primary tumor and N-positive disease. HPV-negative patients were more likely to have a history of heavy smoking and excessive alcohol consumption.

With DW-MRI, ROI areas for OPSCC ranged from 20.0–770.0 mm^2. Median ROI area for HPV-positive and HPV-negative OPSCCs was 165.0 [range 40-450] mm^2 and 160.0 [20-770] mm^2, respectively. The ROI area did not significantly differ between HPV-positive and HPV-negative patients (p=0.77). Examples of a DW-MRI scan from an HPV-negative and an HPV-positive patient are shown in Figures 1 and 2, respectively. The results did not demonstrate significant differences between ADC_{mean} and ADC_{min} in the HPV-negative and HPV-positive groups: ADC_{mean} and ADC_{min} were 1.46 [1.07-2.16] x 10^{-3} mm^2/s (median [range]) and 1.01 [0.62-1.55] x 10^{-3} mm^2/s for HPV-positive patients and 1.56 [1.18-2.18] x 10^{-3} mm^2/s and 1.07 [0.62-1.51] x 10^{-3} mm^2/s for HPV-negative patients (p=0.51 and p=0.67, respectively) (Table 2).

Table 1. General patient characteristics

Characteristic	***HPV-positive*** *n=22 (50%)*	***HPV-negative*** *n=22 (50%)*	***p*-value**
Gender			*p*=0.30
Male	18 (81.8%)	15 (68.2%)	
Female	4 (18.2%)	7 (31.8%)	
Age at diagnosis, years			*p*=0.92
Median [range]	57 [47-78]	60.5 [45-71]	
Oropharyngeal subsite			*p*=0.29
Tonsil	12 (54.4%)	10 (45.5%)	
Base of tongue	8 (36.4%)	6 (27.3%)	
Oropharynx nos	2 (9.1%)	6 (27.3%)	
Smoking			*p*=0.01
Never (0-5 pack years)	8 (36.4%)	0 (0%)	
Moderate (6-24 pack years)	5 (22.7%)	5 (22.7%)	
Heavy (>24 pack years)	9 (40.9%)	17 (77.3%)	
Alcohol consumption			*p*=0.02
Never (0)	3 (13.6%)	0 (0%)	
Moderate (1-149 unit years)	15 (68.2%)	10 (45.5%)	
Heavy (>149 unit years)	4 (18.2%)	12 (54.5%)	
T-stage			*p*=0.03
T2	12 (54.5%)	5 (22.7%)	
T3-4	10 (45.5%)	17 (77.3%)	
N-stage			*p*=0.04
N0	0 (0%)	5 (22.7%)	
N1-3	22 (100%)	17 (77.3%)	
M-stage			*n.a.*
M0	22 (100%)	22 (100%)	
M1	0 (0%)	0 (0%)	

Smoking was defined in pack years (1 pack year=20 cigarettes a day during 1 year)
Alcohol consumption was defined in unit years (1 unit year=one alcohol-containing consumption a day during 1 year).
Abbreviations: HPV: Human papillomavirus, n.a.: not applicable, oropharynx nos: oropharynx not otherwise specified

Table 2. Imaging parameters in relation to HPV-status

DW-MRI parameter	***HPV-positive*** *n=22*	***HPV-negative*** *n=22*	***p*-value**
ADC_{min} ($x\ 10^3\ mm^2/s$)	1.01 [0.62-1.55]	1.07 [0.62-1.51]	p=0.67
ADC_{mean} ($x\ 10^3\ mm^2/s$)	1.46 [1.07-2.16]	1.56 [1.18-2.18]	p=0.51

Values are presented as median [range]
Abbreviations: HPV: Human papillomavirus

Figure 1.
Axial MR images of an HPV-positive OPSCC. A 48-year old patient was diagnosed with an OPSCC in the right palatine tonsil. **(A)** STIR MR image shows a hyperintense mass in the palatine tonsil (arrow), **(B)** b750 diffusion-weighted image and **(C)** ADC map show decreased ADC. A ROI was drawn on the b750 diffusion-weighted image and copied to the ADC map. In this ROI, ADC_{min} measured 0.992 x 10^{-3} mm^2/s.

Figure 2.
Axial MR images of an HPV-negative OPSCC. A 54-year old patient was diagnosed with an OPSCC in the right palatine tonsil. **(A)** STIR MRI image shows a large hyperintense mass in the oropharynx (arrow), **(B)** b750 diffusion-weighted image and **(C)** ADC map show decreased ADC. A ROI was drawn on the b750 diffusion-weighted image and copied to the ADC map. In this ROI, ADC_{min} measured 0.826 x 10^{-3} mm^2/s.

DISCUSSION

OPSCCs may have a heterogeneous response to treatment and survival varies among patients [3,4,19]. Tobacco smoking, alcohol consumption and HPV play a role in this heterogeneity and in particular an HPV infection is a strong independent prognostic factor for survival in OPSCC [4]. Since traditional prognostic factors, such as tumor size and lymph node involvement may be inadequate to classify patients into risk groups, identification of new prognostic tumor characteristics that may select patients for more or less aggressive treatment regimens becomes of interest. DW-MRI may have prognostic value in HNSCC [11-14]. However, it is not known if the relatively low pretreatment ADC-values, which are associated with better response to CRT in these studies, could be attributed to HPV-status. This study shows that imaging parameters derived from DW-MRI are not significantly associated with the OPSCC HPV-status. Therefore, HPV-status and ADC-values may be factors that independently determine the prognosis of a patient with OPSCC.

HPV-related OPSCC is a clinically and histopathologically distinct disease entity [20,21]. Indeed, in this study we found significant differences in clinical presentation of HPV-positive patients with a T2 primary tumor, but positive N-stage. Histopathologically, HPV-related tumors often have distinct features, characterized by a nonkeratinizing morphology, showing excessive mitoses and comedo-type necrosis, while HPV-negative OPSCC show large polygonal cells and typical keratin formation [8,15,20-22]. In general, HPV-positive OPSCC are more poorly differentiated [23]. Characteristics of poorly differentiated squamous cell carcinomas, such as a higher cell density and an increased nuclear-to-cytoplasmic ratio, reduce extracellular space and thereby the diffusion space of water protons which could result in relatively low ADC-values [24,25]. Indeed, Wang et al. found lower ADC-values in poorly differentiated carcinomas compared with highly or moderately differentiated carcinomas [24]. It can be hypothesized that the distinct histological features of HPV-positive and HPV-negative OPSCC, may translate in different pretreatment ADC-vales. If this is the case, this would support the general opinion that low ADC-values correlate with a better response to chemo-radiotherapy [11-14]. Apparently, viable cells in a highly proliferating tumor (with low pretreatment ADC-value) respond more favourable to CRT, probably related to the better vascularisation that increases the exposure to chemotherapeutic agents [12,26]. This is in contrast to tissues with high levels of necrosis (with high ADC-value), leading to a detrimental effect on treatment efficacy [27].

In previous studies [11-14], an association was reported between treatment response and pretreatment ADC-values. This reported association could partly or even completely be attributed to the HPV-status as patients with an HPV-positive OPSCC respond better to CRT than HPV-negative patients. Unfortunately, HPV-status of the OPSCC studied was not included in the analyses. As outlined before, HPV-positive OPSCC have a different histology and HPV-involvement could translate into different ADC-values. We found no association between pretreatment ADC and HPV-status and this suggests that ADC-value and HPV-status are independent characteristics in patients with OPSCC, each with their own prognostic value.

Currently, limited data are available regarding the association between HPV-status and ADC in patients with OPSCC. Our data are in contrast to those of Nakahira et al. who found significant lower pretreatment mean and minimum ADC-values for HPV-positive than for HPV-negative OPSCCs [16]. In their series with 26 patients, they used p16 protein overexpression as a surrogate marker for HPV-infection. However, p16 immunohistochemistry is not an ideal surrogate marker for an HPV-infection, as a considerable percentage of patients are p16-positive but HPV-negative [17,28]. This could have affected their results. In our institute, we use a validated algorithm for HPV-detection: p16 immunostaining, followed by a GP 5+/6 PCR on the p16-immunopostitive cases. This algorithm showed an accuracy of 98% [18]. In addition, the conflicting data may also be explained by DWI factors; the choice of DWI-technique, the use of different b-values and pulse sequences.

We acknowledge several limitations to this study. First, due to the retrospective study design that only included patients with a MRI with PROPELLER-DWI, selection bias may have occurred. Secondly, the number of patients included is relatively small and the study is performed in a single center. Finally, DWI-studies were performed with a PROPELLER-technique. DWI-studies in HNSCC are most commonly performed with an Echo-planar Imaging (EPI)-sequence[11,29,30]. DW-imaging of the head and neck area is particularly difficult, because this region is very inhomogeneous and susceptible to artefacts. If artefacts are too detrimental, a non-EPI technique, such as PROPELLER, may be a better alternative. Nevertheless, in some primary lesions, artefacts were present. Consequently, a smaller ROI was placed only over the undistorted area of the lesion.

CONCLUSION

Patients with HPV-positive OPSCC have a better survival compared to HPV-negative OPSCC. ADC-values of DWI may also have prognostic value. In this study we investigated a possible association between HPV-status and ADC-

values in OPSCC. No significant associations were found between quantitative imaging parameters from DW-MRI and HPV-status in OPSCC. Apparently, the differences in histological features between HPV-positive and HPV-negative OPSCC do not translate into different pretreatment ADC-values in our study cohort. Long-term follow-up studies with DW-MRI and HPV are needed to investigate if ADC-values and HPV-status are independent prognostic factors in OPSCC patients treated by chemoradiotherapy.

ACKNOWLEDGEMENTS

We thank dr. M.M. Rietbergen and drs. S. Hakim (Department of Otolaryngology-Head and Neck Surgery, VU University Medical Center, Amsterdam) for their assistance in retrieving HPV-statuses and clinical characteristics. The authors thank drs. Daniel P. Noij (Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam) for assisting with ADC measurements.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer.* 2010;127(12):2893-2917.
2. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat. Rev. Cancer.* 2011;11(1):9-22.
3. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J. Natl. Cancer Inst.* 2008;100(4):261-269.
4. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* 2010;363(1):24-35.
5. Braakhuis BJ, Snijders PJ, Keune WJ, et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J. Natl. Cancer Inst.* 2004;96(13):998-1006.
6. Smeets SJ, Brakenhoff RH, Ylstra B, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol.* 2009;31(4):291-300.
7. Thariat J, Badoual C, Faure C, Butori C, Marcy PY, Righini CA. Basaloid squamous cell carcinoma of the head and neck: role of HPV and implication in treatment and prognosis. *J. Clin. Pathol.* 2010;63(10):857-866.
8. Chernock RD, El-Mofty SK, Thorstad WL, Parvin CA, Lewis JS, Jr. HPV-related nonkeratinizing squamous cell carcinoma of the oropharynx: utility of microscopic features in predicting patient outcome. *Head Neck Pathol.* 2009;3(3):186-194.
9. Martincich L, Deantoni V, Bertotto I, et al. Correlations between diffusion-weighted imaging and breast cancer biomarkers. *Eur. Radiol.* 2012;22(7):1519-1528.
10. Choi SY, Chang YW, Park HJ, Kim HJ, Hong SS, Seo DY. Correlation of the apparent diffusion coefficiency values on diffusion-weighted imaging with prognostic factors for breast cancer. *Br. J. Radiol.* 2012;85(1016):e474-e479.
11. Kim S, Loevner L, Quon H, et al. Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. *Clin. Cancer Res.* 2009;15(3):986-994.
12. Srinivasan A, Chenevert TL, Dwamena BA, et al. Utility of pretreatment mean apparent diffusion coefficient and apparent diffusion coefficient histograms in prediction of outcome to chemoradiation in head and neck squamous cell carcinoma. *J. Comput. Assist. Tomogr.* 2012;36(1):131-137.
13. Chawla S, Kim S, Dougherty L, et al. Pretreatment diffusion-weighted and dynamic contrast-enhanced MRI for prediction of local treatment response in squamous cell carcinomas of the head and neck. *AJR Am. J. Roentgenol.* 2013;200(1):35-43.
14. Hatakenaka M, Nakamura K, Yabuuchi H, et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81(2):339-345.
15. Chernock RD, Lewis JS, Jr., Zhang Q, El-Mofty SK. Human papillomavirus-positive basaloid squamous cell carcinomas of the upper aerodigestive tract: a distinct clinicopathologic and molecular subtype of basaloid squamous cell carcinoma. *Hum. Pathol.* 2010;41(7):1016-1023.

16. Nakahira M, Saito N, Yamaguchi H, Kuba K, Sugasawa M. Use of quantitative diffusion-weighted magnetic resonance imaging to predict human papilloma virus status in patients with oropharyngeal squamous cell carcinoma. *Eur. Arch. Otorhinolaryngol.* 2013.
17. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int. J. Cancer.* 2007;121(11):2465-2472.
18. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int. J. Cancer.* 2013;132(7):1565-1571.
19. Lassen P, Eriksen JG, Hamilton-Dutoit S, Tramm T, Alsner J, Overgaard J. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. *J. Clin. Oncol.* 2009;27(12):1992-1998.
20. Gillison ML, Koch WM, Capone RB, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl. Cancer Inst.* 2000;92(9):709-720.
21. El-Mofty SK, Patil S. Human papillomavirus (HPV)-related oropharyngeal nonkeratinizing squamous cell carcinoma: characterization of a distinct phenotype. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 2006;101(3):339-345.
22. El-Mofty SK, Zhang MQ, Davila RM. Histologic identification of human papillomavirus (HPV)-related squamous cell carcinoma in cervical lymph nodes: a reliable predictor of the site of an occult head and neck primary carcinoma. *Head Neck Pathol.* 2008;2(3):163-168.
23. Mendelsohn AH, Lai CK, Shintaku IP, et al. Histopathologic findings of HPV and p16 positive HNSCC. *Laryngoscope.* 2010;120(9):1788-1794.
24. Wang J, Takashima S, Takayama F, et al. Head and neck lesions: characterization with diffusion-weighted echo-planar MR imaging. *Radiology.* 2001;220(3):621-630.
25. Sumi M, Sakihama N, Sumi T, et al. Discrimination of metastatic cervical lymph nodes with diffusion-weighted MR imaging in patients with head and neck cancer. *AJNR Am. J. Neuroradiol.* 2003;24(8):1627-1634.
26. Kim S, Loevner LA, Quon H, et al. Prediction of response to chemoradiation therapy in squamous cell carcinomas of the head and neck using dynamic contrast-enhanced MR imaging. *AJNR Am. J. Neuroradiol.* 2010;31(2):262-268.
27. Tatum JL, Kelloff GJ, Gillies RJ, et al. Hypoxia: importance in tumor biology, noninvasive measurement by imaging, and value of its measurement in the management of cancer therapy. *Int. J. Radiat. Biol.* 2006;82(10):699-757.
28. Rietbergen MM, Snijders PJ, Beekzada D, et al. Molecular characterization of p16-immunopositive but HPV DNA-negative oropharyngeal carcinomas. *Int. J. Cancer.* 2013.
29. Vandecaveye V, Dirix P, De Keyzer F, et al. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. *Eur. Radiol.* 2010;20(7):1703-1714.
30. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.

PART 2:
PREDICTION
DURING TREATMENT

CHAPTER 4

DIFFUSION-WEIGHTED EPI- AND HASTE-MRI AND ^{18}F-FDG-PET-CT EARLY DURING CHEMORADIOTHERAPY IN ADVANCED HEAD AND NECK CANCER

C.S. Schouten
R. de Bree
L. van der Putten
D.P. Noij
O.S. Hoekstra
E.F.I. Comans
B.I. Witte
P.A. Doornaert
C.R. Leemans
J.A. Castelijns

Quantitative Imaging in Medicine and Surgery 2014; 4: 239-50

ABSTRACT

Background. Diffusion-weighted MRI (DW-MRI) has potential to predict chemoradiotherapy (CRT) response in head and neck squamous cell carcinoma (HNSCC) and is generally performed using echo-planar imaging (EPI). However, EPI-DWI is susceptible to geometric distortions. Half-fourier acquisition single-shot turbo spin-echo (HASTE)-DWI may be an alternative. This prospective pilot study evaluates the potential predictive value of EPI- and HASTE-DWI and ^{18}F-fluorodeoxyglucose PET-CT (^{18}F-FDG-PET-CT) early during CRT for locoregional outcome in HNSCC.

Methods. Eight patients with advanced HNSCC (7 primary tumors and 25 nodal metastases) scheduled for CRT, underwent DW-MRI (using both EPI- and HASTE-DWI) and ^{18}F-FDG-PET(-CT) pretreatment, early during treatment and three months after treatment. Median follow-up time was 38 months.

Results. No local recurrences were detected during follow-up. Median Apparent Diffusion Coefficient $(ADC)_{EPI}$-values in primary tumors increased from 77 x 10^{-5} mm^2/s pretreatment, to 113 x 10^{-5} mm^2/s during treatment ($p=0.02$), whereas ADC_{HASTE} did not increase (74 and 74 mm^2/s, respectively). Two regional recurrences were diagnosed. During treatment, ADC_{EPI} tended to be higher for patients with regional control (117.3 ± 12.1 x 10^{-5} mm^2/s) than for patients with a recurrence (98.0 ± 4.2 x 10^{-5} mm^2/s). This difference was not seen with ADC_{HASTE}. No correlations between ΔADC_{EPI} and ΔSUV (Standardized Uptake Value) were found in the primary tumor or nodal metastases.

Conclusion. HASTE-DWI seems to be inadequate in early CRT response prediction, compared to EPI-DWI which has potential to predict locoregional outcome. EPI-DWI and ^{18}F-FDG-PET-CT potentially provide independent information in the early response to treatment, since no correlations were found between ΔADC_{EPI} and ΔSUV.

INTRODUCTION

Patients with resectable advanced staged head and neck squamous cell carcinomas (HNSCC) are currently often treated with non-surgical protocols to preserve organ function and to maintain quality of life [1,2]. Although chemoradiotherapy (CRT) results in acceptable locoregional control rates, recurrence rates remain considerable [2,3]. If residual or recurrent disease is detected after CRT, surgical 'salvage' treatment may be an option, but 'salvage' surgery is often associated with substantial morbidity and complications [4,5]. Prediction of treatment outcome early during treatment might avoid ineffective treatment in certain patients and would allow a treatment switch to surgery in these patients [6].

Diffusion-weighted magnetic resonance imaging (DW-MRI) has been suggested as a predictive factor for response of tumor to CRT [7]. DW-MRI characterizes tissue based on differences in water mobility, which is related to cellularity [8]. These differences can be quantified with Apparent Diffusion Coefficient (ADC): hypercellular tissue (*e.g.* malignancy) is characterized by a low ADC, whereas hypocellular tissue with necrosis or apoptosis is characterized by a high ADC [9]. Conceptually, response to treatment should correspond to an increase in ADC, because treatment-induced loss of tumor cells increases water mobility at the microscopic level. In contrast, residual tumor cells might be detected as decreased ADC-values [10]. Several studies have indicated the potential of DW-MRI as a predictor of treatment response in HNSCC [11-13].

DW-MRI in HNSCC is most commonly performed with an echo-planar imaging (EPI)-sequence [11-13]. It may be difficult to perform DWI of the head and neck area, because this region is very inhomogeneous and susceptible to artefacts. EPI-DWI is particularly prone to geometric distortions due to susceptibility artefacts [14]. DW-MRI is a potential technique for tumor definition in radiotherapy planning, but accurate target definition is essential. Also, with PET/MRI spreading in the clinical field, geometrical accuracy is crucial for fusing PET-images with DW-MRI images. If artefacts with EPI-DWI are too detrimental, a non-EPI method such as half-fourier acquisition single-shot turbo spin-echo (HASTE), may be a better alternative [15]. Verhappen et al. compared EPI- with HASTE-DWI in HNSCC and concluded that EPI images showed more geometric distortions [15]. A comparative study between EPI- and HASTE-DWI in HNSCC for prediction of locoregional control after CRT has not been performed previously.

Tumor metabolism is another potential predictor and can be studied with positron emission tomography (PET). ^{18}F-fluorodeoxyglucose (^{18}F-FDG), a radiolabeled glucose analogue, is used to measure glucose metabolism in

malignant tissues. Clinical studies report associations between decline in ^{18}F-FDG uptake in the early phase of CRT and a positive therapy outcome [16-18].

The aim of this pilot study was twofold. First, the purpose was to compare HASTE-DWI with EPI-DWI and ^{18}F-FDG-PET(-CT) early during CRT for their potential to predict locoregional outcome in patients with HNSCC. Secondly, we wanted to correlate changes in ADC- and SUV-values between pretreatment and early during treatment.

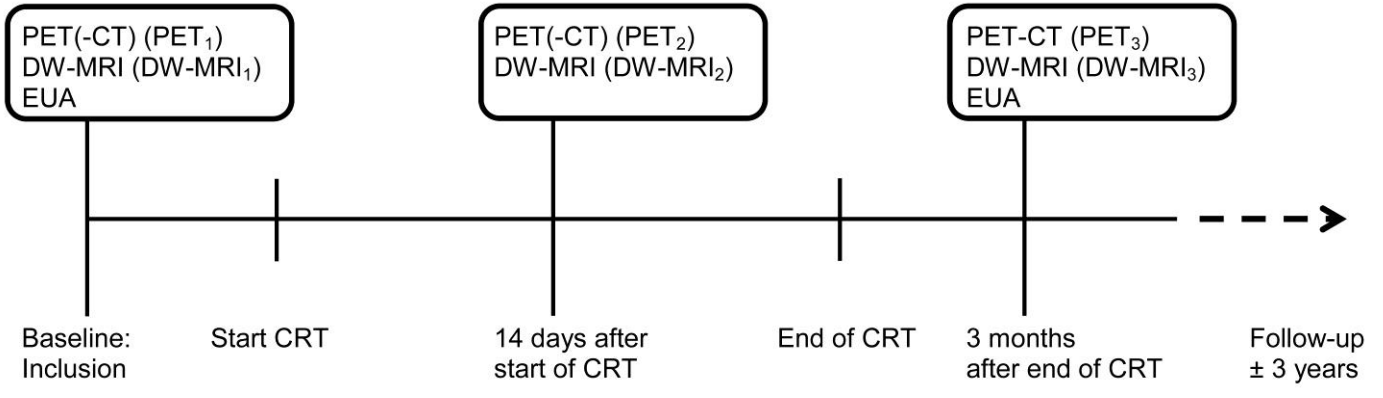
METHODS

Patients and study design

Eight patients with histological proven advanced (T2, T3 or T4) oro- or hypopharyngeal carcinoma (with a total of 7 primary tumors and 25 lymph node metastases) scheduled for primary CRT with curative intent, were enrolled in this prospective pilot study (Table 1). The study was approved by the institutional ethics committee and complies with the Declaration of Helskini. Informed consent was obtained in all patients. Routine pretreatment examinations included ^{18}F-FDG-PET(-CT) (PET_1), MRI and an examination under anesthesia (EUA) with biopsies. For study-purposes, EPI- and HASTE-DWI were added (DW-MRI_1). A second MRI with additional DW-MRI (DW-MRI_2) and a second ^{18}F-FDG-PET(-CT) (PET_2) were performed 14 days (±1 day) after the start of radiotherapy (20 Gy±2 Gy). DW-MRI_2 and PET_2 were not used for clinical assessment. All patients received cisplatin-based CRT (n=6) or cetuximab-based CRT (n=2). A radiation dose of 70 Gray (Gy) in 2 Gy/fraction was delivered and elective nodal regions received a dose of 54.25 Gy-57.75 Gy in 1.55-1.65 Gy/fraction. All patients completed radiotherapy, but toxicity precluded complete cisplatin-CRT in one patient.

During follow-up, patients were regularly examined according to our standard head-and-neck oncology protocol. Routine response evaluation was performed three months after CRT, using DW-MRI (DW-MRI_3), ^{18}F-FDG-PET(-CT) (PET_3) and an examination under general anaesthesia. Median follow-up was 38 months (range 17-60 months). Additional investigations during follow-up were performed at the discretion of the attending physician. Locoregional control was defined as persistent complete regression of the primary tumor and lymph nodes during follow-up. A timeline illustrating the consecutive methodological steps in the study is shown in Figure 1.

Figure 1. Timeline illustrating the consecutive methodological steps in the study



Abbreviations:
EUA: examination under general anesthesia
CRT: chemoradiotherapy

Table 1. Patient and tumor characteristics.

Patient No.	*Gender*	*Age*	*Primary site*	*T*	*N*	*Treatment method*	*Locoregional recurrence*	*Salvage surgery*	*Follow-up*
1	M	51	Palatine tonsil	3	2c	Cisplatin-based CRT	LNM[a]	Yes	37 months DM, DOD
2	M	68	Palatine tonsil	2	2b	Cisplatin-based CRT	No	No	60 months NED
3	M	63	Base of tongue	4	2c	Cisplatin- based CRT	No	No	46 months NED
4	M	56	Palatine tonsil	4	3	Cisplatin- based CRT	No	No	39 months NED
5	F	55	Palatine tonsil	2	2a	Cisplatin-based CRT[b]	No	No	37 months NED
6	M	63	Vallecula	3	2c	Cetuximab-based CRT	LNM	No	17 months DM, DOD
7	F	63	Palatine tonsil	2	2b	Cisplatin-based CRT	No	No	35 months NED
8	M	68	Piriform sinus	3	1	Cetuximab-based CRT	No	No	30 months NED

[a] Histopathologically proven
[b] Toxicity precluded complete chemotherapy
Abbreviations:
Gender: M, Male; F, Female
Age: age at diagnosis (in years)
Locoregional recurrence: LNM, lymph node metastasis
Follow-up: DM, distant metastasis; DOD, dead of disease; NED, no evidence of disease

DW-MRI

MRI was performed using a 1.5 Tesla MR imaging system (Sonata; Siemens, Erlangen, Germany) with a head coil combined with a phased array spine and neck coil. After an axial short TI inversion-recovery (STIR)-series with 7-mm sections covering the entire neck area, subsequent images were centered on the area of interest containing the primary tumor and enlarged lymph nodes. Axial images (22 slices of 4-mm slice thickness and 0.4-mm gap, in-plane pixel size of 0.9x0.9 mm) were obtained with STIR (TR/TE/T1=5500/26/150 ms, 2 averages) and T1-weighted (T1WI) spin-echo (TR/TE=390/140 ms, 2 averages, no fat saturation) before and after the injection of contrast material. Gadovist (0.1 mL/kg of gadobutrol), Magnevist (0.2 mL/kg gadopentetate dimeglumine; both Bayer Schering Pharma, Berlin-Wedding, Germany) or Dotarem (0.2 mL/kg of gadoteric acid; Guerbet, Aulnay-sous Bois, France), was intravenously administered to obtain contrast-enhanced T1WI.

DWI with both EPI- and HASTE-techniques was obtained for the same 22 slices at the same slice position as the axial STIR and T1WI. Parameters for EPI were the following: TR/TE=5000/105 ms, in-plane pixel size= 2x2 mm, and b-values=0, 500 and 1000 s/mm^2 (3 averages). Parameters for HASTE were: TR/TE=900/ 110 ms, in-plane pixel size=1.1 x 1.1 mm, and b-values=0 s/mm^2 (3 averages) and 1000 s/mm^2 (12 averages). ADC maps of both EPI- and HASTE-DWI were calculated on-line or off-line, respectively, by using the software of the scanner.

^{18}F-FDG-PET-CT

All patients fasted for at least 6 hours. Mean serum glucose levels were 6.5 mmol/l, with a range from 4.3 to 11.2 mmol/l. 186-367 MBq of ^{18}F-FDG, depending on the body mass index and PET system used, was intravenously injected. PET_1 consisted of at least a whole-body PET (mid-femur to cranial vault) in all patients plus head and neck imaging (jugulum to orbit) in 4 patients, whereas PET_2 and PET_3 only comprised PET images of the head and neck region. In two patients, PET imaging was performed using a full-ring BGO PET scanner (ECAT EXACT HR+, CTI/Siemens, Erlangen, Germany; 2D-mode; 5 min emission scans/bed position, 2-min transmission scans using Ge-68 rod sources). PET-scanning started at sixty minutes (±15) post-injection (p.i). of ^{18}F-FDG. The PET-images were reconstructed using ordered subset expectation maximisation (OSEM) with 2 iterations and 16 subsets, an image matrix size of 128x128, resulting in voxel sizes of 5x5 mm. A 5-mm FWHM (full width at half maximum) Gaussian post-reconstruction filter was applied,

resulting in a final image resolution of 7 mm FWHM. During reconstruction all corrections needed for quantification were applied, such as decay, attenuation, scatter, dead time and normalisation corrections. In the other patients, PET-imaging was performed using an integrated PET-CT system (Gemini TF, Philips Medical Systems, Best, the Netherlands; 3D-mode; 2 min emission scans/bed position). Low dose CT scanning was performed with 120 kV and 50 mAs prior to emission scanning and used for attenuation correction of the emission scan and for anatomical localisation of FDG-avid lesions. In three patients, PET imaging was performed 60 minutes (±15) post injection and in three patients PET was performed 90 minutes (±15) post injection. PET-CT data were reconstructed using a time of flight row-action maximum likelihood algorithm, as implemented by the vendor. Final image matrix size equals 170x170 with a voxel size of 4x4x4 mm. Final image resolution equalled 7 mm FWHM. Serial PET-CT studies in a single patient were performed using the same scanner, uptake time, acquisition and reconstruction protocols.

Analysis of MRI data

DW-MRI scans were analysed by a radiologist (J. Castelijns) with 29 years of experience in head and neck radiology. Clinical information was provided about TNM stage, but the interpreter was blinded to clinical outcome. DW-MRI_1, DW-MRI_2 and DW-MRI_3 were simultaneous analysed on PACS (Sectra RIS/PACS version 12, Sectra Imtec AB, Linköping, Sweden) that allowed viewing of multiple MRI scans. All primary tumor and metastatic lymph nodes with a minimal axial diameter >5 mm were included. A lymph node was considered metastatic if proven by fine needle aspiration cytology or indicated by increased ^{18}F-FDG uptake on PET(-CT) scan. All included lesions were identified on baseline images and corresponding lesions on DW-MRI_2 and DW-MRI_3 were identified by visual and slice position-based correlation. For each lesion, contours were manually drawn on the conventional MR images by J. Castelijns around the lesional border at each slice position to measure total tumor volume. The volume of the lesions was calculated as the sum of the surfaces at each slice position multiplied by slice thickness and the interslice gap. Volume changes (ΔV_X) in % in relation to DW-MRI_1 were calculated using the formula:

$$\Delta V_X = [(V_X - V_B)/ V_B]*100$$

where V_B represents baseline volume and V_X represents volume on the X^{th} time point during or after treatment. A composite of all included lymph nodes was used to calculate the change in nodal volume. Thereafter, ADC-values were

calculated by drawing a region of interest (ROI) on a single slice of an axial EPI- and HASTE-ADC maps, containing the largest available tumor area. The sets of DWI were evaluated independently from each other. For solid lesions, ROIs were drawn encompassing the entire lesion. In case of necrotic components, ROIs were drawn in that area of the lesion that showed contrast-enhancement in the corresponding post-contrast T1WI. ADC was measured before, during and after treatment in those patients with a residual enlarged lymph node. It was impossible to reliably draw a ROI if lymph node metastases had strongly shrunk due to the treatment. The lowest ADC-value of all pathologic lymph nodes in one patient (ADC_{low}) was considered a representative measure for follow-up, as suggested by Wahl et al. for PET [19]. ADC-changes (ΔADC_X) in % in relation to baseline were calculated, similar to changes in volume

Analysis of PET(-CT) data

PET images were independently interpreted by two nuclear medicine physicians with each 15 years PET experience (O. Hoekstra and E. Comans) in head and neck oncology. PET-images were assessed on the presence of foci of increased activity within the tumor greater than surrounding background. PET readers had access to clinical information and $DW\text{-}MRI_1$ for anatomic correlation, but were blinded to the report of the radiologist and clinical outcome. PET(-CT) images were displayed on a standard workstation allowing simultaneous viewing of coronal, sagittal and transverse planes, with cross-referencing, as well as a 3 dimensional rotation projection. In case of discrepant interpretations a consensus was reached after discussion.

Standardized uptake values (SUV) were calculated as SUV_{max} (highest tumor voxel value within the lesion) and SUV_{mean} (average SUV within the lesion) by C. Schouten, under supervision of O. Hoekstra, measured in the primary tumors and in the (up to 3) largest lymph nodes, using previously described methodology [20]. SUVs were normalized for body weight and serum glucose. If, after treatment, no lesions with increased ^{18}F-FDG uptake were visible, a ROI of 3x3x3 voxels was drawn at the initial location of the primary tumor and/or lymph nodes. SUV-changes (ΔSUV_X) in % in relation to baseline were calculated.

Statistical analysis

Statistical analyses were performed using SPSS software package (version 20.0; IBM Corp., Armonk, NY, USA). The level of significance was set at $p<0.05$. A two-sided nonparametric exact Wilcoxon signed rank test was used for paired data comparisons between primary tumor parameters from the first and second DW-MRI or ^{18}F-FDG-PET(-CT). A two-sided Mann-Whitney U test was used for group comparisons; regional control versus regional recurrent disease. To evaluate correlations between ΔADC and ΔSUV, a Spearman's correlation coefficient was used.

RESULTS

Imaging

DW-imaging before and during treatment was conducted in all patients according to the study protocol. PET(-CT) imaging and reconstruction was not correctly performed in all patients due to different logistic difficulties, as indicated in Table 2. All primary tumors were detected with DW-MRI (both EPI- and HASTE-technique) and PET(-CT) except in one patient, in whom the primary tumor had been resected transorally in another hospital. ADC- and ΔSUV-values of the primary tumor and nodal metastases at baseline and during treatment are shown in Table 2a and 2b.

Treatment outcome

Six out of eight patients remained disease-free during follow-up. In two patients an isolated regional recurrence was diagnosed, 17 and 29 months post-CRT. No local recurrences were detected. One salvage neck dissection was performed with histopathologically proven lymph node metastases in the surgical specimen. In the other patient, regional recurrence was presumed on the basis of clinical examination and ultrasound imaging. This patient died (of a carotid blow-out) before histopathological diagnosis was obtained.

Table 2a. ADC_{EPI}, ADC_{HASTE}, ΔSUV_{mean} and ΔSUV_{max} for primary tumors at baseline and early during treatment

Patient No.	ADC_{EPI} MRI_1 ($x\ 10^{-5}\ mm^2/s$)	ADC_{EPI} MRI_2 ($x\ 10^{-5}\ mm^2/s$)	ADC_{HASTE} MRI_1 ($x\ 10^{-5}\ mm^2/s$)	ADC_{HASTE} MRI_2 ($x\ 10^{-5}\ mm^2/s$)	ΔSUV_{mean} PET_{1-2} (%)	ΔSUV_{max} PET_{1-2} (%)
1	84	117	114	111	15.9	15.8
2	85	102	106	128	NA[1]	NA[1]
3	104	134	70	73	NA[2]	NA[2]
4	77	143	58	73	-74.5	-79.5
5	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
6	56	57	85	74	-69.1	-69.4
7	77	98	74	54	-54.4	-54.9
8	66	113	65	78	NA[4]	NA[4]

[1] PET_1 was performed without a transmission scan
[2] PET_1 was reconstructed with an aberrant voxel size
[3] No primary tumor
[4] PET_2 was not performed
Abbreviations: NA, not applicable

Table 2b. ADC_{EPI}, $ADC_{HASTE,}$ ΔSUV_{mean} and ΔSUV_{max} for nodal metastases at baseline and early during treatment * ‡

Patient No.	ADC_{EPI} MRI_1 ($x\ 10^{-5}\ mm^2/s$)	ADC_{EPI} MRI_2 ($x\ 10^{-5}\ mm^2/s$)	ADC_{HASTE} MRI_1 ($x\ 10^{-5}\ mm^2/s$)	ADC_{HASTE} MRI_2 ($x\ 10^{-5}\ mm^2/s$)	ΔSUV_{mean} PET_{1-2} (%)	ΔSUV_{max} PET_{1-2} (%)
1	93	101	101	107	39.1	28.7
2	80	121	103	136	NA[1]	NA[1]
3	109	124	84	68	NA[2]	NA[2]
4	67	93	41	74	-51.1	-62.2
5	89	121	66	89	-39.4	-42.7
6	78	95	67	71	-25.7	-30.4
7	72	125	71	93	-37.9	-38.2
8	108	120	91	110	NA[3]	NA[3]

* The lowest ADC-value of all included lymph nodes in one patient
‡ The highest SUV-value of all included lymph nodes in one patient
[1] PET_1 was performed without a transmission scan
[2] PET_1 was reconstructed with an aberrant voxel size
[3] PET_2 was not performed
Abbreviations: NA, not applicable

Primary tumor

Figure 2a and 2b represents the pattern of change in ADC_{EPI} and ADC_{HASTE}. With EPI-DWI, six patients showed a substantial ADC-increase from $DW\text{-}MRI_1$ to $DW\text{-}MRI_2$, whereas ADC_{EPI} increased with only 1.8 % in patient 6 on $DW\text{-}MRI_2$. With HASTE-DWI, three patients showed a substantial ADC-increase on $DW\text{-}MRI_2$ compared to $DW\text{-}MRI_1$. ADC-values in the other four patients did not show a substantial increase or showed a decrease. Volume, ADC-, ΔADC- and ΔSUV-values of the primary tumors are listed in Table 3. Median pretreatment ADC_{EPI} was 77 x 10^{-5} (SD 15.2) mm^2/s, significantly increasing to 113 x 10^{-5} (SD 27.8) mm^2/s (p=0.02) early during treatment. Median ADC_{HASTE} values were 74 x 10^{-5} (SD 21.1) mm^2/s and 74 x 10^{-5} (SD 25.6) mm^2/s.

Visual interpretation of PET_2 still showed a focus of increased activity within the tumor in four patients. ΔSUV_{max} decreased with 62.1%±43.1 (median±SD) and ΔSUV_{mean} with 61.7±41.8 from PET_1 to PET_2.

Lymph node metastases

At baseline, median ADC-values of patients with regional control (ADC_{EPI}: 87.5 x 10^{-5} mm^2/s and ADC_{HASTE}: 76.7 x 10^{-5} mm^2/s) and those with recurrent disease (ADC_{EPI}: 85.5 x 10^{-5} mm^2/s and ADC_{HASTE}: 84.0 x 10^{-5} mm^2/s) were similar (p=0.89 and p=0.74, respectively). At $DW\text{-}MRI_2$, ADC_{low} with EPI tended to be (not statistically significant, p=0.18) higher for six patients with regional control (117.3±12.1 x 10^{-5} mm^2/s) than for the patients with a recurrence (98.0±4.2 x 10^{-5} mm^2/s). With HASTE-DWI this difference was not seen (93.5±26.7 x 10^{-5} mm^2/s versus 89.0 ±25.5 x 10^{-5} mm^2/s, p=0.74) (Figure 3a). $\Delta ADC_{low\text{-}2weeks}$ with EPI tended to be higher for patients with regional control than for recurrences (37.4±23.5% versus 15.2±9.3%, p=0.18). $\Delta ADC_{low\text{-}2weeks}$ with HASTE also tended to be higher for patients with regional control (27.4±27.1% versus 6.0±0.02%, p=0.18) (Figure 3b). An example of $DW\text{-}MRI_1$ and $DW\text{-}MRI_2$ in a patient with a regional recurrence is shown in Figure 4.

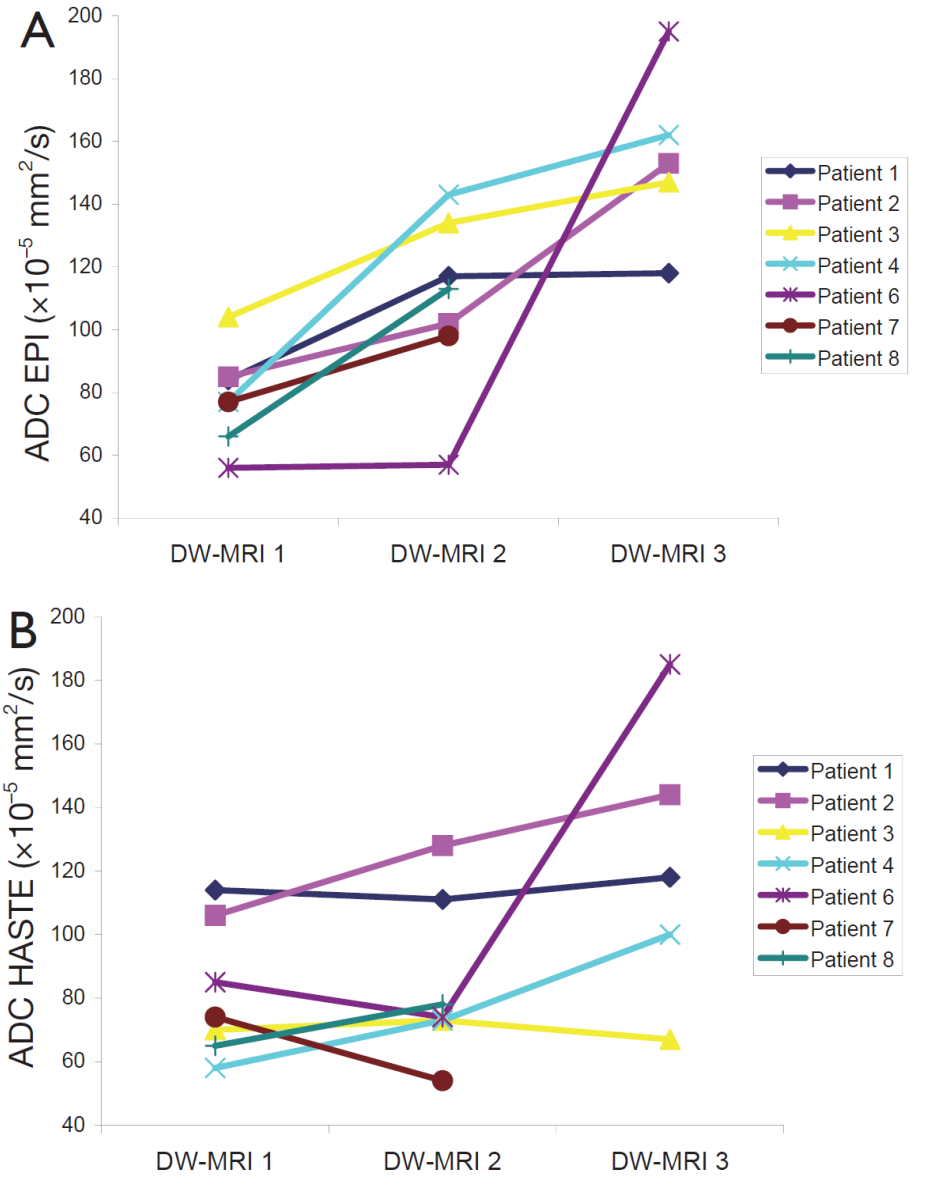


Figure 2.
Patterns of change in ADC_{EPI} **(A)** and ADC_{HASTE} **(B)** between $DW\text{-}MRI_1$, $DW\text{-}MRI_2$ and $DW\text{-}MRI_3$ of the primary tumor. The DW-MRI after treatment was not conducted using study protocol in patient 7 and patient 8.

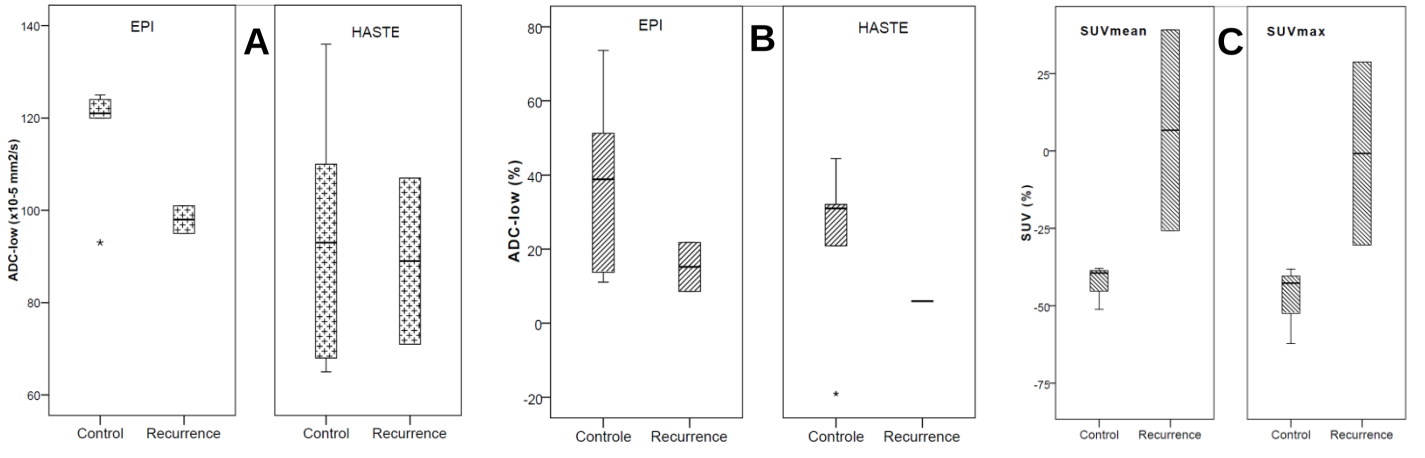


Figure 3.
Comparison of lymph node **(A)** ADC_{low} at $DW\text{-}MRI_{2,}$ **(B)** $\Delta ADC_{low\text{-}2weeks}$ (in %) and **(C)** ΔSUV_{2weeks} (in %), in six patients with regional control and two patients with recurrent disease. Box-whisker plots are presented with median (-), interquartile range (box), and range (-).

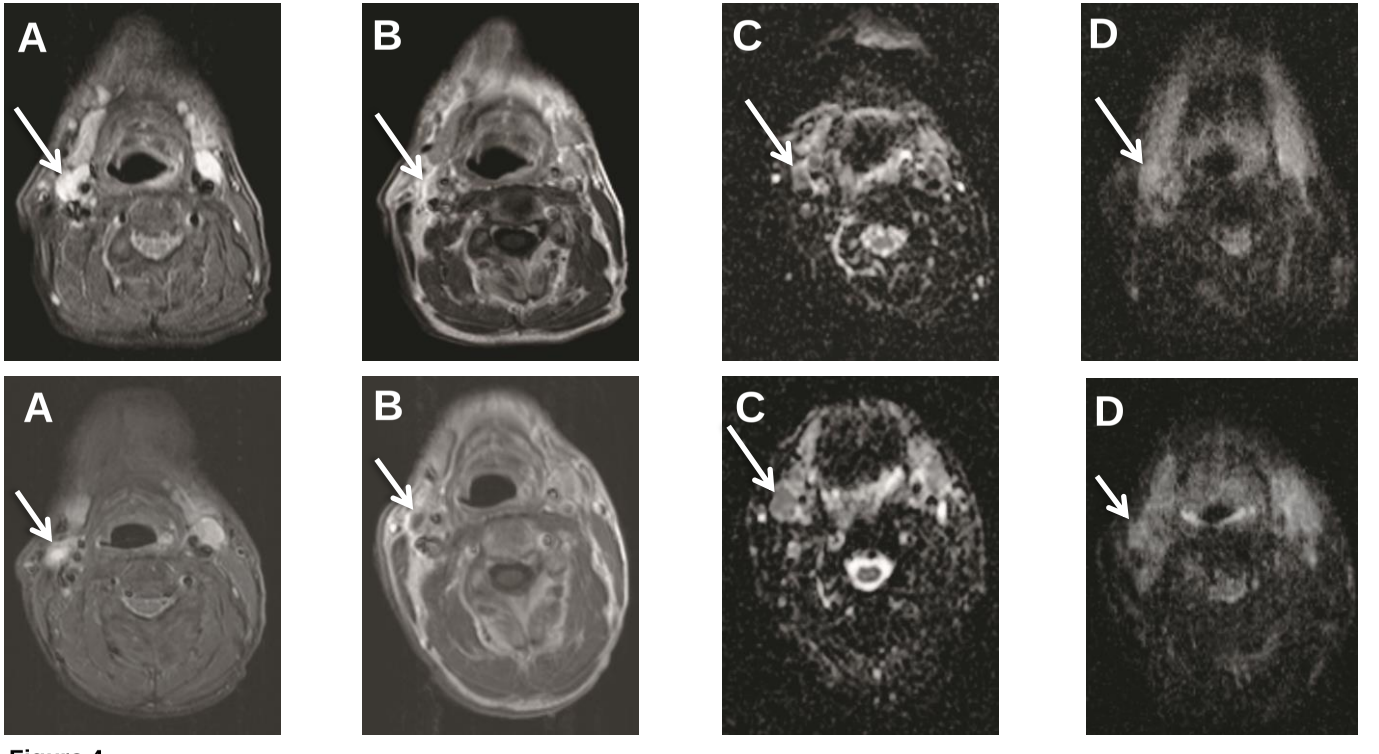


Figure 4.
Axial images showing a metastatic node (arrows) in patient number 1 in whom recurrent viable squamous cell carcinoma was diagnosed histopathologically in level II right during follow-up. DW-MRI_1 (top row) and DW-MRI_2 (bottom row): **(A)** STIR, **(B)** contrast-enhanced T1WI, **(C)** ADC maps with EPI technique and **(D)** ADC maps with HASTE technique. ADC_{EPI}-values of the lymph node (arrow) are 99 and 102 mm^2/s for DW-MRI_1 and DW-MRI_2, respectively. ADC_{HASTE}-values are 106 and 118 mm^2/s. Four years after completion of CRT this patient died due to lung metastases.

$\Delta Volume_{2weeks}$ in six patients with regional control was -38.9%±42.5 (mean±SD) and 13.0%±9.2 in the two patients with a lymph node recurrence (p=0.74). Both patients with a regional recurrence were visually interpreted as a non-complete response on PET_2. Of the patients with regional control, in two patients no focus of increased activity within the lymph nodes was seen, whereas in three patients a focus was still seen. A trend was seen for more change in SUV_{max} in patients with regional control than in patients with a regional recurrence. $\Delta SUV_{max\text{-}2weeks}$ in regional control was -47.7%±12.7 and -0.8%±41.8 in regional recurrences. $\Delta SUV_{mean\text{-}2weeks}$ in patients with regional control was -42.8%±7.2 and 6.7%±45.8 in patients with a recurrence (p=0.08).

Correlation between ADC and SUV

For the primary tumors, no correlation were found between $\Delta ADC_{EPI\text{-}2weeks}$ and $\Delta SUV_{mean\text{-}2weeks}$ or $\Delta SUV_{max\text{-}2weeks}$ (p=0.80) or between $\Delta ADC_{HASTE\text{-}2weeks}$ and $\Delta SUV_{mean\text{-}2weeks}$ or $\Delta SUV_{max\text{-}2weeks}$ (p=0.60). For the lymph node metastases, no correlation was seen in $\Delta ADC_{EPI\text{-}2weeks}$ and $\Delta SUV_{mean\text{-}2weeks}$ (spearman's rho=-0.70, p=0.19) or $\Delta SUV_{max\text{-}2weeks}$ (spearman's rho=-0.40, p=0.6). A significant negative correlation was found between $\Delta ADC_{HASTE\text{-}2weeks}$ and $\Delta SUV_{max\text{-}2weeks}$ (spearman's rho=-0.90, p=0.04) and $\Delta SUV_{mean\text{-}2weeks}$ (spearman's rho=-1.0, p=0.01) (Figure 5).

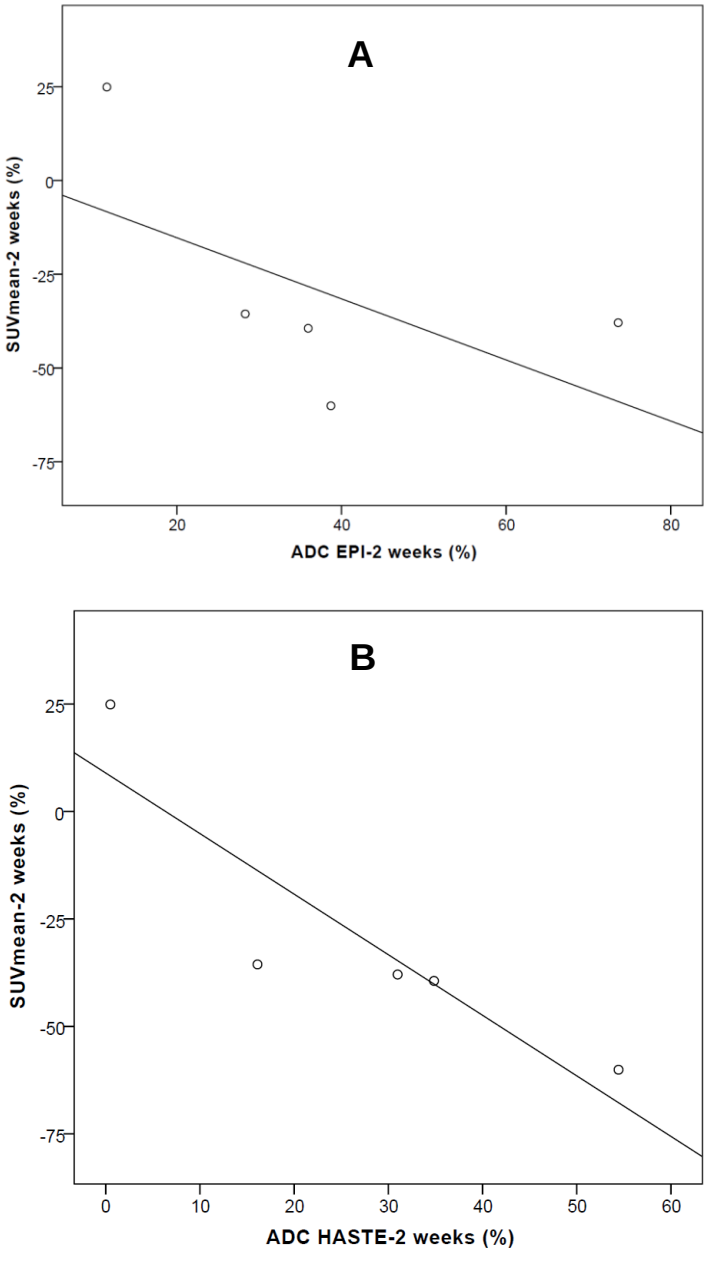


Figure 5.
Correlation for the lymph node metastases between **(A)** $\Delta ADC_{EPI\text{-}2weeks}$ and $\Delta SUV_{mean\text{-}2weeks}$ and **(B)** $\Delta ADC_{HASTE\text{-}2weeks}$ and $\Delta SUV_{mean\text{-}2weeks}$.

DISCUSSION

CRT is a standard therapeutic option for patients with advanced stage HNSCC, also if technically resectable. Identification of non-responders early during CRT may spare a number of patients from a futile extensive treatment. Several results in HNSCC studies suggest that changes in ADC measured with an EPI-DWI technique early during CRT are associated with locoregional response [11-13]. However, EPI-DWI suffers from geometrical distortions, especially in regions with air-tissue transitions such as in the head and neck area. Consequently, the use of EPI-DWI in radiotherapy planning and in simultaneous PET/MRI imaging may be limited. In this pilot study, we wanted to explore the use of a non-EPI DWI method, because such DWI sequences are more robust concerning geometric accuracy. We compared EPI-DWI with HASTE-DWI early during CRT for their potential to predict locoregional outcome. Our preliminary results suggest that EPI-DWI seems to have greater potential in predicting locoregional outcome early after start of CRT than HASTE-DWI. Although HASTE-DWI has a lower incidence of geometric distortions as compared to an EPI-DWI [15], this technique seems to fail in early CRT response prediction in HNSCC.

CRT induces loss of tumor cells and thus increases water mobility at the microscopic level. Response to treatment corresponds to an increase in ADC. This treatment-induced ADC-increase has been confirmed in several HNSCC-studies. Kim et al. showed a significant ADC increase in responding, compared to non-complete responding metastatic lymph nodes from HNSCC, one week after initiation of radiotherapy [11]. Unlike the study of Kim et al., our DWI sequences covered the neck completely instead of only the metastatic lymph node and analysis also included the primary tumor. In a second study, King et al. analysed primary tumors and lymph nodes together, without differentiating these entities and showed that serial changes in tumor ADC, obtained over the course of treatment, provided a marker for treatment response. A fall in ADC during treatment correlated with locoregional failure [13]. In another study with 30 patients, Vandecaveye et al. concluded that ADC-changes of the primary tumor and lymph nodes at 2 and 4 weeks after the start of CRT were significantly associated with locoregional response, in contrast to the change in volume [12]. In the head and neck region, DWI is generally performed with an EPI-sequence, as in previous described studies. Our findings using EPI-DWI are compatible with these HNSCC-studies. In contrary to previous studies, we evaluated DWI as a technique to predict treatment response with both EPI- and HASTE-DWI, to explore the application of a non-EPI sequence in this area of research.

DWI is particularly difficult in this region, because it contains a variety of tissues, including bone, fat, muscle, glandular tissue and air. Moreover, movement-related problems, like swallowing, breathing, coughing, speaking and jaw movements impede imaging of the head and neck. This can produce images with strong susceptibility artefacts. EPI-DWI is sensitive to geometric distortions, which is especially strong near interfaces between soft tissue and air or bone. Functional imaging has a crucial role in accurate tumor delineation and defining the targets for radiotherapy planning. ^{18}F-FDG-PET-CT is commonly used for treatment planning. DW-MRI might has additional value in treatment planning, because DW-MRI can distinguish between reactively enlarged lymph nodes and metastatic lymph nodes [21]. Therefore, DW images without geometric distortions are important for fusing PET images with DWI. If artefacts are too detrimental, a non-EPI technique can be used instead of an EPI-technique. MR images performed with an EPI- or non-EPI method, vary concerning contrast, SNR (signal-to-noise ratio) and artefarcts. In HNSCC, Verhappen et al. showed that primary tumors and metastatic lymph nodes are more easily visualized on EPI-DWI compared to HASTE-DWI due to a higher SNR. However, EPI-DWI demonstrated more frequent susceptibility artefacts resulting in geometric distortions[14,15]. In the present study, we performed both EPI- and HASTE-DWI. As stated above, EPI-DWI might have greater potential in predicting locoregional outcome and HASTE-DWI seems to provide inadequate information. Up to now, it is uncertain which DWI-technique is most appropriate in head-and-neck imaging. However, our study contributes to the general opinion that EPI-DWI probably is the most promising technique in oncologic imaging in the head and neck region. Therefore, further optimization of the EPI-DWI sequence is necessary to reduce image distortions and in order to make this technique useful in radiotherapy planning and simultaneous PET/MRI imaging.

^{18}F-FDG-PET-CT is another possible modality for treatment prediction. Controversial results on the accuracy of PET to predict treatment outcome in HNSCC patients have been reported. Several authors have concluded that changes in FDG-uptake levels during non-surgical treatment are associated with tumor response, locoregional control and overall survival [16-18]. However, Castaldi et al. could not confirm a predictive role for PET-CT performed after 2 weeks of CRT [22]. Ceulemans et al. found a low sensitivity for FDG-PET after 47 Gy [23]. The interpretation of PET-images can be difficult because of false positive findings, as tracer uptake can also occur in normal tissues, inflammatory tissue or reactive lymph nodes. Besides, optimal timing to assess response with PET-CT during radiotherapy remains a matter of debate, since increases in ^{18}F-FDG-uptake early during treatment have been reported due to radiation-induced inflammatory responses and repair processes [24]. We

performed PET(-CT) after 20 Gy. At this time, radiotherapy-induced inflammation and ^{18}F-FDG accumulation in the activated macrophages is assumed to be low [25]. Most aforementioned studies are performed with stand-alone PET, while PET-CT is the current 'state of the art'. In the present study PET-CT was performed in most patients, using CT to improve the optimal delineation of the primary tumor and lymph node metastases (ROI).

DW-MRI and ^{18}F-FDG-PET-CT are both imaging techniques used in oncology and have similar clinical applications. However, both modalities represent different aspects of tumor biology; ADC representing tissue cellularity and SUV representing glucose metabolism. A few studies in HNSCC assessed the correlation between pretreatment ADC-values and SUV-values. Nakajo et al. demonstrated a significant inverse correlation between primary tumor SUV_{max} and ADC in 26 patients [26]. Nakamatsu et al. demonstrated this negative correlation between SUV_{max} and ADC_{min} also in 41 metastatic lymph nodes [27]. Opposite, Fruehwald-Pallamar et al. and Varoquaux et al. did not find a correlation between primary tumor ADC and SUV [28,29]. Our present pilot study is the first study to compare changes in ADC and SUV between pretreatment and early during treatment. For the primary tumor, no correlations between ΔADC (with EPI- and HASTE-DWI) and ΔSUV were found. The results for the nodal metastases demonstrated no correlation between ΔADC_{EPI} and ΔSUV. A significant negative correlation was found between ΔADC_{HASTE} and ΔSUV. Our results suggest that both EPI-DWI and ^{18}F-FDG-PET-CT might provide independent information in the early response to treatment, since no correlations were found. Both techniques could play a different role in clinical assessment, in contrast to HASTE-DWI which seems to provide the same information as ^{18}F-FDG-PET(-CT), since significant correlations were found between ΔADC_{HASTE} and ΔSUV. Therefore, a combination of EPI-DWI and PET might be promising in predictive and follow-up studies of HNSCC and with simultaneous PET/MRI imaging spreading in the clinical field, both techniques can be combined in one single scanner.

We acknowledge several limitations to this study. First, this pilot study had an exploratory character and was conducted with a small number of patients. Although a limited number of patients was included, this is the first study to evaluate the potential predictive value of two DWI-techniques and ^{18}F-FDG-PET(-CT) with follow-up. Multiple tumors (primary and metastases) in a single patient were analysed independently to offset this small number of patients, resulting in 32 tumors. Second, in our patient cohort no local recurrences occurred during follow-up. Therefore, it is difficult to investigate the role of DW-MRI and ^{18}F-FDG-PET(-CT) as an imaging tool to predict local outcome. However, in our preliminary results ADC_{EPI} from the primary tumors showed the expected significant increase from baseline to early during

treatment, in contrast to ADC_{HASTE}. Third, we evaluated the DWI-images by whole-tumor mean ADC-value. The accuracy of DW-MRI can possibly be improved by expanding these calculations with data on skewness and kurtosis, taking tumor heterogeneity into account [30] or a voxel-wise approach with parametric response maps [31]. Finally, we were not able to report absolute SUV-values since we applied two different uptake intervals (60 and 90 minutes); changes of SUV are more robust provided that the same conditions prevailed during serial imaging per patient, and this was the case in our study.

CONCLUSION

Our preliminary results suggest that HASTE-DWI seems to be inadequate in early response prediction, compared to EPI-DWI which has greater potential to predict locoregional outcome after CRT. EPI-DWI and ^{18}F-FDG-PET-CT could potentially provide independent information in the early response to treatment, since no correlations were found between ΔADC_{EPI} and ΔSUV. This is in contrast to ADC-HASTE, which appears to provide similar information as ^{18}F-FDG-PET-CT, because significant correlations were found between ΔADC_{HASTE} and ΔSUV.

ACKNOWLEDGMENTS

We would like to thank prof. dr. R. Boellaard (Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam) for his assistance in reconstructing PET-data.

REFERENCES

1. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
2. 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(6):567-578.
3. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.
4. Lavertu P, Bonafede JP, Adelstein DJ, et al. Comparison of surgical complications after organ-preservation therapy in patients with stage III or IV squamous cell head and neck cancer. *Arch. Otolaryngol. Head Neck Surg.* 1998;124(4):401-406.
5. Newman JP, Terris DJ, Pinto HA, Fee WE, Jr., Goode RL, Goffinet DR. Surgical morbidity of neck dissection after chemoradiotherapy in advanced head and neck cancer. *Ann. Otol. Rhinol. Laryngol.* 1997;106(2):117-122.
6. De Bree R. Functional imaging to predict treatment response after (chemo) radiotherapy of head and neck squamous cell carcinoma. *Quant. Imaging Med. Surg.* 2013;3(5):231-234.
7. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. *AJR Am. J. Roentgenol.* 2007;188(6):1622-1635.
8. Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol.* 2003;45(3):169-184.
9. Ross BD, Moffat BA, Lawrence TS, et al. Evaluation of cancer therapy using diffusion magnetic resonance imaging. *Mol. Cancer Ther.* 2003;2(6):581-587.
10. Hamstra DA, Lee KC, Moffat BA, Chenevert TL, Rehemtulla A, Ross BD. Diffusion magnetic resonance imaging: an imaging treatment response biomarker to chemoradiotherapy in a mouse model of squamous cell cancer of the head and neck. *Transl. Oncol.* 2008;1(4):187-194.
11. Kim S, Loevner L, Quon H, et al. Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. *Clin. Cancer Res.* 2009;15(3):986-994.
12. Vandecaveye V, Dirix P, De Keyzer F, et al. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. *Eur. Radiol.* 2010;20(7):1703-1714.
13. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
14. Schakel T, Hoogduin JM, Terhaard CH, Philippens ME. Diffusion weighted MRI in head-and-neck cancer: geometrical accuracy. *Radiother. Oncol.* 2013;109(3):394-397.
15. Verhappen MH, Pouwels PJ, Ljumanovic R, et al. Diffusion-weighted MR imaging in head and neck cancer: comparison between half-fourier acquired single-shot turbo spin-echo and EPI techniques. *AJNR Am. J. Neuroradiol.* 2012;33:1239-1246.
16. Brun E, Kjellén E, Tennvall J, et al. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. *Head Neck.* 2002;24(2):127-135.

17. Farrag A, Ceulemans G, Voordeckers M, Everaert H, Storme G. Can 18F-FDG-PET response during radiotherapy be used as a predictive factor for the outcome of head and neck cancer patients? *Nucl. Med. Commun.* 2010;31(6):495-501.
18. Hentschel M, Appold S, Schreiber A, et al. Early FDG PET at 10 or 20 Gy under chemoradiotherapy is prognostic for locoregional control and overall survival in patients with head and neck cancer. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(7):1203-1211.
19. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *J. Nucl. Med.* 2009;50 Suppl 1:122S-150S.
20. Krak NC, Boellaard R, Hoekstra OS, Twisk JW, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur. J. Nucl. Med. Mol. Imaging.* 2005;32(3):294-301.
21. Vandecaveye V, De Keyzer F, Vander P, V, et al. Head and neck squamous cell carcinoma: value of diffusion-weighted MR imaging for nodal staging. *Radiology.* 2009;251(1):134-146.
22. Castaldi P, Rufini V, Bussu F, et al. Can "early" and "late" 18F-FDG PET-CT be used as prognostic factors for the clinical outcome of patients with locally advanced head and neck cancer treated with radio-chemotherapy? *Radiother. Oncol.* 2012;103(1):63-68.
23. Ceulemans G, Voordeckers M, Farrag A, Verdries D, Storme G, Everaert H. Can 18-FDG-PET during radiotherapy replace post-therapy scanning for detection/demonstration of tumor response in head-and-neck cancer? *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81(4):938-942.
24. Bussink J, van Herpen CM, Kaanders JH, Oyen WJ. PET-CT for response assessment and treatment adaptation in head and neck cancer. *Lancet Oncol.* 2010;11(7):661-669.
25. Hentschel M, Appold S, Schreiber A, et al. Serial FDG-PET on patients with head and neck cancer: implications for radiation therapy. *Int. J. Radiat. Biol.* 2009;85(9):796-804.
26. Nakajo M, Nakajo M, Kajiya Y, et al. FDG PET/CT and diffusion-weighted imaging of head and neck squamous cell carcinoma: comparison of prognostic significance between primary tumor standardized uptake value and apparent diffusion coefficient. *Clin. Nucl. Med.* 2012;37(5):475-480.
27. Nakamatsu S, Matsusue E, Miyoshi H, Kakite S, Kaminou T, Ogawa T. Correlation of apparent diffusion coefficients measured by diffusion-weighted MR imaging and standardized uptake values from FDG PET/CT in metastatic neck lymph nodes of head and neck squamous cell carcinomas. *Clin. Imaging.* 2012;36(2):90-97.
28. Fruehwald-Pallamar J, Czerny C, Mayerhoefer ME, et al. Functional imaging in head and neck squamous cell carcinoma: correlation of PET/CT and diffusion-weighted imaging at 3 Tesla. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(6):1009-1019.
29. Varoquaux A, Rager O, Lovblad KO, et al. Functional imaging of head and neck squamous cell carcinoma with diffusion-weighted MRI and FDG PET/CT: quantitative analysis of ADC and SUV. *Eur. J. Nucl. Med. Mol. Imaging.* 2013;40(6):842-852.
30. King AD, Chow KK, Yu KH, et al. Head and neck squamous cell carcinoma: diagnostic performance of diffusion-weighted MR imaging for the prediction of treatment response. *Radiology.* 2013;266(2):531-538.

31. Galban CJ, Mukherji SK, Chenevert TL, et al. A feasibility study of parametric response map analysis of diffusion-weighted magnetic resonance imaging scans of head and neck cancer patients for providing early detection of therapeutic efficacy. *Transl. Oncol.* 2009;2(3):184-190

Part 3: Response evaluation

CHAPTER 5

RESPONSE EVALUATION AFTER CHEMORADIOTHERAPY FOR ADVANCED STAGED OROPHARYNGEAL SQUAMOUS CELL CARCINOMA: A NATIONWIDE SURVEY IN THE NETHERLANDS

C.S. Schouten
O.S. Hoekstra
C.R. Leemans
J.A. Castelijns
R. de Bree

European Archives of Otorhinolaryngology 2015; 272: 3507-13

ABSTRACT

Background. Following failure of chemoradiotherapy (CRT) for advanced staged oropharyngeal squamous cell carcinomas (OPSCC), residual tumor can often be treated successfully with salvage surgery, if detected early. Current clinical practice in the VU University Medical Center is to perform routine response evaluation, i.e. examination under general anaesthesia (EUA), 12 weeks after treatment. However, in the Netherlands there is no consensus on response evaluation in patients with advanced oropharyngeal cancer.

Methods. Questionnaire on current clinical practice concerning response evaluation after CRT for advanced OPSCC in all eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group.

Results. The response rate was 100%. Response evaluation was routinely performed with various methods in five institutions (62.5%) and in one institute (12.5%) only if clinical evaluation was difficult. Two centers (25%) did not perform response evaluation. In case of suspicion of residual disease during follow-up, six centers (75%) performed imaging prior to EUA and two centers (25%) only if clinical evaluation was difficult. Diagnostic techniques used prior to EUA were MRI (87.5%), diffusion-weighted MRI (37.5%), ^{18}Fluoro-2-deoxyglucose positron emission tomography-computed tomography (75-87.5%) and CT (37.5%).

Conclusion. This survey shows a substantial variation in the diagnostic policy concerning response evaluation after CRT for advanced OPSCC in the Netherlands. There is a need for guidelines for response evaluation in patients with advanced oropharyngeal cancer.

INTRODUCTION

Head and neck cancer is the sixth most common cancer worldwide [1] and in the Netherlands, each year 270 new patients with oropharyngeal squamous cell carcinoma (OPSCC) are diagnosed. Advanced stage disease accounts for 80% of all oropharyngeal carcinomas [2]. Formerly, unresectable OPSCC was treated by radiotherapy with or without chemotherapy and patients with resectable OPSCC were treated with a combination of surgery and radiotherapy. Surgery often led to functional and cosmetic morbidity, leading to a diminished quality of life. Nowadays, resectable OPSCC is also treated with chemoradiotherapy (CRT) with the intention to preserve organ function and to maintain quality of life, resulting in acceptable locoregional control rates [3-5].

After failure of CRT, residual tumor of initially resectable OPSCC can be treated successfully with salvage surgery, if detected timely. In case of late detection and delayed salvage surgery, local control and survival rates rapidly decrease [6]. However, salvage surgery is associated with considerable risks [7] and will only be performed after histopathological confirmation of viable tumor cells. Therefore, current clinical practice in the VU University Medical Center is to routinely perform response evaluation, i.e. examination under general anaesthesia (EUA), 12 weeks after the end of treatment, to detect residual disease as soon as possible. However, the incidence of residual disease after CRT is low. In this manner, many patients are exposed to unnecessary biopsies in treated areas, inducing pain, inflammation and wound healing problems [8]. Besides, due to sampling errors within the residual mass, biopsies in previously irradiated areas may be false negative [6]. Moreover, an EUA requires hospital stay and operating facilities. Current clinical practice to try and identify patients who may benefit from salvage surgery needs to be improved. This requires a diagnostic strategy which accurately selects patients who should undergo invasive EUA, without compromising the benefit of timely detection of residual disease.

Functional imaging techniques such as ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography (^{18}F-FDG-PET-CT) and diffusion-weighted magnetic resonance imaging (DW-MRI) may be useful in this setting [9]. PET-CT combines imaging of metabolically active tissues with anatomical information. Due to increased glycolytic activity, higher concentrations of ^{18}F-FDG accumulate in malignant tumors than in normal tissue [10]. DW-MRI characterizes tissue based on differences in water mobility, which is related to cellularity [11]. Post-treatment nontumoral tissue changes are expected to show less cellularity than viable tumor tissue [12]. However, at present, there are no national guidelines for the diagnostic policy concerning response evaluation after CRT for advanced OPSCC.

The aim of this study was to evaluate the current clinical practice. A questionnaire was sent to head and neck surgeons (as representatives) in the eight head and neck cancer centres of the Dutch Head and Neck Oncology Cooperative Group (Nederlandse Werkgroep Hoofd-Halstumoren).

MATERIALS & METHODS

In 2014, a questionnaire on current clinical practice concerning response evaluation after CRT for advanced OPSCC was sent to clinicians in all eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group. The questionnaire (Table 1) was accompanied by an explanatory letter.

Table 1. Questionnaire on current practice concerning response evaluation after CRT for advanced OPSCC

Q.1	Who performs the follow-up in patients treated with CRT for OPSCC? *(more than 1 answer allowed)*
	(a) Otolaryngologist/ head and neck surgeon (b) Radiotherapist (c) Medical oncologist
Q.2	Do you routinely perform response evaluation in these patients?
	(a) Yes, always (b) Yes, but only in patients with an initially resectable tumor (c) Sometimes, please explain: …… (d) No, only in case of suspicion of residual disease/ complaints *(move to question 5)*
Q.3	Which technique(s) do you use for response evaluation? *(more than 1 answer allowed)*
	(a) Physical examination (b) CT (c) MRI (d) Ultrasound (with FNAC) (e) ^{18}F-FDG-PET-CT (f) EUA (g) Other, i.e.: ……
Q. 4	When do you perform response evaluation? *(open question)*
Q. 5	Do you routinely perform imaging in patients with OPSCC after CRT to establish the new situation (which can be of value later on, in patients with suspected residual disease during follow-up)?
	(a) Yes, always (b) Yes, but only in patients with an initially resectable tumor (c) Sometimes, please explain: …… (d) No *(move to question 8)*

Q. 6	Which imaging technique(s) do you use to establish to new situation? *(more than 1 answer allowed)*
	(a) CT (b) MRI (c) Ultrasound (with FNAC) (d) ^{18}F-FDG-PET-CT (e) Other, i.e.: ……
Q. 7	When do you perform these imaging techniques of the new situation? *(open question)*
Q. 8	Do you perform imaging before EUA in case of suspicion of residual disease?
	(a) Yes (b) Sometimes, please explain: …… (c) No *(go to question 10)*
Q. 9	Which imaging technique(s) do you use in case of suspicion of residual disease? *(more than 1 answer allowed)*
	(a) CT (b) MRI (c) Ultrasound (with FNAC) (d) ^{18}F-FDG-PET-CT (e) Other, i.e.: ……
Q. 10a	Is diffusion-weighted imaging performed during MRI?
	(a) Yes (b) No *(move to question 11)*
Q. 10b	Do you take the results of DW-MRI into consideration?
	(a) Yes (b) No
Q. 10c	Which b-values are used to perform DW-MRI? *(open question)*
Q. 10d	Are quantitative ADC-values determined?
	(a) Yes (b) No
Q. 11a	Are dedicated head and neck images made during ^{18}F-FDG-PET-CT?
	(a) Yes (b) No, just a whole body PET-CT
Q. 11b	The PET-CT is made after …. minutes post injection of ^{18}F-FDG *(open question)*
Q. 11c	Is a beta-blocker used to reduce the uptake in brown adipose tissue?
	(a) Yes (b) No
Q. 11d	Is a benzodiazepine used to avoid unwanted uptake in muscles?
	(a) Yes (b) No
Q. 11e	Are quantitative SUV-values determined?
	(a) Yes (b) No *(go to question 12)*
Q. 11f	Which SUV-values are determined? *(more than 1 answer allowed)*
	(a) SUV_{mean} (b) SUV_{max} (c) Other, i.e.:……

Table continues on the next page

Q. 12	Do you standard take a biopsy during the EUA?
	(a) Yes (b) No, only in case of clinical suspicion
Q. 13	Do you look at the results of the imaging prior to the EUA? I.e. is your decision to perform a biopsy based on imaging?
	(a) Yes (b) No, my decision to perform a biopsy is not based on imaging (c) No, I did not perform imaging prior to EUA

Abbreviations:
ADC; apparent diffusion coefficient, CT; computed tomography, CRT; chemoradiotherapy, DW-MRI; diffusion-weighed magnetic resonance imaging, EUA; examination under general anaesthesia, FNAC; fine needle aspiration cytology, OPSCC; oropharyngeal squamous cell carcinoma, PET-CT; positron emission tomography, SUV; standardized uptake value

RESULTS

All questionnaires were returned completed. In seven institutions (87.5%), both the head and neck surgeon (otolaryngologist or maxillofacial surgeon) and the radiotherapist performed the clinical follow-up in OPSCC patients after CRT. In one of these centers (12.5%), the medical oncologist was also standard involved in the clinical follow-up. In one center (12.5%) only the radiotherapist performed the clinical follow-up. In these patients, response evaluation was performed with a variety of methods in 5 institutions (62.5%); 4 centers performed response evaluation in all patients and 1 center performed response evaluation only in patients with an initially resectable tumor. One center (12.5%) performed response evaluation if clinical evaluation was difficult and two centers (25%) did not perform response evaluation at all. Response evaluation, if implemented, was performed varying from 8-12 weeks after end of CRT, but most centers (5 out of 6) performed response evaluation 12 weeks after end of CRT. The techniques which are routinely used for response evaluation, are shown by institution in Table 2A.

Besides performing imaging for response evaluation, in three centers (37.5%), routine imaging after CRT was also performed to establish a new baseline situation for future comparisons, in patients with suspected residual disease during follow-up. Four centers (50%) did not perform routine imaging after CRT and one (12.5%) performed imaging after CRT to establish the new situation if clinical evaluation was difficult. The methods used for establishing the new situation are shown in Table 2B.

In case of clinical suspicion of residual disease during follow-up, six centers (75%) performed imaging prior to EUA and two centers (25%) only performed imaging prior to EUA in case clinical evaluation was difficult or in

case no definable tumor was manifest. Table 2C shows the different diagnostic imaging techniques that were used in case of suspected residual disease. A large variation is reported by the institutes which techniques and in what combination these techniques are used for response evaluation, establishing the new anatomy or detection of residual disease: CT, MRI, ^{18}F-FDG-PET-CT, ultrasound (with fine needle aspiration cytology (FNAC)), chest CT and EUA.

If ^{18}F-FDG-PET-CT was performed, seven centers (87.5%) applied dedicated high resolution protocols for the head and neck area (instead of whole body imaging). One center (12.5%) just performed whole body PET-CT imaging. Two centers sometimes used beta-blockers (such as propranolol) to reduce the uptake in brown fat in a research protocol or if the nuclear medicine physician considers it useful. Another center reported the use of benzodiazepines to avoid unwanted uptake in neck musculature. Five centers (62.5%) reported standard quantitative Standardized Uptake Values (SUV_{mean} and/or SUV_{max}) and in one center (12.5%) SUV-values were determined on request. In the other two centers (25%), SUV-values were not determined for clinical purposes.

If MRI was performed after CRT, this included DW-MRI in 5 centers (62.5%). However, only three centers (37.5%) based clinical decisions on the result of the DW-MRI. In the other two centers, DW-MRI had recently become available and the scan protocol was still under development. In the three centers with DW-MRI experience, different b-values and number of b-values were used. All centers used the b-values 0 and 1000, two centers added b750 or b800 to the MRI-protocol. At the time of the questionnaire, three centers (37.5%) did not have experience with DW-MRI.

In seven centers (87.5%), the head and neck surgeons looked into the results of the imaging techniques, if available, prior to the EUA to determine the location of an eventual biopsy. All head and neck surgeons only took a biopsy in case of clinical suspicion of residual disease during EUA.

Table 2. Techniques routinely used by institution for (A) response evaluation, (B) establishing the new situation after therapy to facilitate follow-up and (C) in case of suspected of residual disease

Centers	*A. Response evaluation*	*B. Establishing new situation*	*C. Suspicion residual disease*
1	MRI, ultrasound (with FNAC)[a]	None	MRI, PET-CT, ultrasound (with FNAC)[b]
2	CT or MRI, ultrasound (with FNAC)[a]	CT or MRI, ultrasound (with FNAC)[a]	CT or MRI, PET-CT*, ultrasound (with FNAC)[a]
3	None	None	CT, MRI, PET-CT, ultrasound (with FNAC)[a]
4	MRI, ultrasound (with FNAC)[a]	MRI[a]	MRI, PET-CT[a]
5	None	None	PET-CT[b]
6	CT, MRI[a]	None	CT, MRI, chest CT[c]
7	MRI, PET-CT[b]	MRI[b]	MRI, PET-CT[a]
8	MRI, PET-CT, EUA[c]	MRI, PET-CT[c]	MRI, PET-CT, ultrasound (with FNAC) [a]

[a] In every patient
[b] Only if clinical evaluation is difficult
[c] Only in patients with an initially resectable tumor
[d] in case of planned salvage surgery
Abbreviations: CT= computed tomography, MRI= magnetic resonance imaging, EUA= examination under general anaesthesia, FNAC= fine needle aspiration cytology,
PET-CT= positron emission tomography-computed tomography

DISCUSSION

It is of great importance that CRT can be implemented without compromising locoregional disease control. Early detection in case of residual tumor is an important prognostic factor, since survival rates decline with delayed salvage surgery [6]. In this survey we found a large variation among the institutes which examinations and in what combination these examinations were used for response evaluation. Most institutes offensively pursue potential residues, probably leading to a high rate of futile diagnostic procedures due to the fairly good response rates after CRT. Alternatively, other institutes perform careful clinical observation throughout the course of treatment and during follow-up without performing diagnostic procedures with potential morbidity and costs. However, this harbours the risk of delaying the diagnosis of residual tumor and potentially reduces the chances for surgical cure and survival [6,13]. There is a need for national guidelines for response evaluation in patients with advanced oropharyngeal cancer.

The yield of EUA in the detection of residual OPSCC after CRT may be improved by patient selection based on imaging techniques. Based on a series of 46 patients with OPSCC treated with radiotherapy, Ojiri et al. showed a high negative predictive value for CT. CT findings with grade 0 (no focal lesion and no asymmetry) and grade 1 (anatomic asymmetry or discrete mass < 10 mm) always suggested good control at the primary site [14]. Van den Broek et al. applied these Ojiri criteria to MRI in 82 patients with mostly oral and oropharyngeal cancer who were treated with CRT. The authors concluded that one can refrain from EUA in patients with MRI findings grade 0 or 1 [15]. However, these conventional anatomy-based investigations have limitations and can be difficult to interpret, since post-treatment changes including fibrosis, oedema and necrosis obscure accurate assessment [16]. In a retrospective study at our institute, the interpretation of MR imaging performed 3 months after chemoradiotherapy showed only 'moderate' interobserver variability [17] between two radiologists (Cohen's Kappa, κ=0.52) for the primary tumor in 86 patients [manuscript in preparation].

Potential methods to improve the distinction between residual tumor and aspecific postradiation tissue are functional imaging with ^{18}F-FDG-PET-CT and DW-MRI [9]. For ^{18}F-FDG-PET performed with a mean of 38 days after treatment, Kitagawa et al. showed a sensitivity of 100% and a specificity of 89.5% in a prospective study with 23 oral cancer patients treated with CRT (with an incidence of residual disease of 17%) [18]. Based on a series of 92 HNSCC patients treated with (chemo)radiation, Moeller et al. found a sensitivity of 70% and a specificity of 94% for ^{18}F-FDG-PET-CT conducted 8 weeks after end of treatment [19]. In another prospective study, Krabbe et al. demonstrated in

48 patients with oral or oropharyngeal cancer treated with curative intent (surgery, radiotherapy or a combination), that 73% of the futile EUAs could be avoided with 8% of the local residues missed when ^{18}F-FDG-PET-CT was performed three months after treatment [20]. To summarize, ^{18}F-FDG-PET(-CT) can be used as an additional tool in response evaluation to select patients who should undergo EUA without a high risk of missing residual disease. However, there is a lack of clear response criteria for test positivity. The guideline for standardization of ^{18}F-FDG-PET-CT by Boellaard et al. enables the comparison of study results, since the examinations should be consistent between institutes that acquire the data [21]. Optimization of response criteria should be addressed in larger multicenter studies.

PET-CT imaging is usually conducted as whole body imaging, with a relatively low resolution. Hence, for PET-CT imaging of a complex anatomic region like the head and neck area, whole body imaging is less appropriate. Moreover, the CT-part is performed as a low-dose non-contrast CT, which is adequate for anatomic correlation of the PET images and attenuation correction, but only a high dose CT scan with contrast offers diagnostic image quality. A few studies recommend dedicated high resolution protocols for the head and neck area, especially for the detection of small lymph node metastases [22,23]. The present survey demonstrates the use of dedicated head and neck PET-CT protocols in 87.5% of the head and neck cancer centers in the Netherlands.

In ^{18}F-FDG-PET-CT imaging, unwanted physiological uptake of ^{18}F-FDG may occur in the neck musculature and vocal folds, due to movement or talking, but also due to uncomfortable waiting conditions [24,25]. Benzodiazepines, which have muscle relaxant activity, can be used prior to administration of ^{18}F-FDG to avoid unwanted uptake. Using the protocol for the standardization of multicenter PET studies by Boellaard et al., muscle uptake can be minimized by means of patient instructions and optimal resting conditions [25].

Another source of false-positive results is caused by tracer uptake in areas of brown adipose tissue (BAT). BAT is often closely related to important lymph node groups in the neck and supraclavicular region. Its metabolism is influenced through activation of β-adrenergic receptors. Administering β-adrenergic antagonists (e.g. propranolol) prior to ^{18}F-FDG injection can reduce tracer uptake in BAT and can lead to improved assessment of PET-CT imaging [26,27]. Currently, only two centers in the Netherlands sometimes use propranolol in head and neck oncology patients.

DW-MRI can also be used to monitor treatment response. King et al. showed that DW-MRI with Apparent Diffusion Coefficients performed 6 weeks after (chemo)radiation was an early marker for locoregional failure in a post-treatment mass [28]. Vandecaveye et al. reported significantly lower ΔADC in

lesions with later tumor recurrence than in lesions with complete remission, in a study with 29 HNSCC patients and DW-MRI performed three weeks after CRT [29]. DW-MRI is also increasingly used for the detection of recurrent HNSCC after treatment, in patients with a clinically suspected recurrence. In a pilot study with 30 patients, Abdel Razek et al. showed that DW-MRI provided promising results for discriminating recurrent tumors from postoperative or postradiation changes [30]. Tshering Vogel et al. included 46 patients for MR imaging with a median of 14 months after (chemo)radiotherapy and achieved good diagnostic accuracy [31]. Vandecaveye et al. concluded in a study with 26 patients that DW-MRI can differentiate between persistent or recurrent HNSCC and nontumoral tissue changes after CRT [32]. Taken together, DW-MRI might have additional value to differentiate between residual or recurrent tumor and post-treatment changes. DW-MRI is non-invasive and takes only a few minutes extra for patients who already undergo a MRI. However, only a few studies with small patient groups have been performed. These results needs to be validated in larger studies. In addition, a wide variation in ADC-values of recurrent tumor and postradiation effects have been reported. The use of different scan protocols (e.g. DWI sequence, b values, field strength) can explain this variation. Future research should address developing uniform DWI protocols.

The optimal timing of post-treatment ^{18}F-FDG-PET-CT and DW-MRI imaging is a subject of debate. Response evaluation in the Netherlands is performed varying from 8-12 weeks after end of CRT, but 82.5% of the centers perform response evaluation 12 weeks after CRT. ^{18}F-FDG-PET-CT performed prior to 10 weeks after end of CRT is related to high rates of false negative results, perhaps because residual viable tumor cells did not have adequate time to repopulate to a level that can be detected by PET-CT [33,34]. False positive findings also occur due to inflammation and postradiation soft tissue effects, particularly present early after CRT [34]. Early detection of residual disease is important to allow for prompt salvage treatment when the size of residual tumor is still limited. DW-MRI seems to have prognostic value in this early post-treatment phase, as previous described results from the literature suggest [28,29]. There is currently no consensus for optimal timing of post-treatment imaging in head and neck cancer patients.

CONCLUSION

This survey shows a substantial variation in the diagnostic policy concerning response evaluation after CRT for advanced OPSCC between the eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group. There is a need for guidelines for response evaluation in patients with advanced oropharyngeal cancer. Functional imaging, such as ^{18}F-FDG-PET-

CT and DW-MRI, may be helpful in diagnosing residual disease and avoiding futile EUA. Further multicenter prospective studies and optimization of response criteria are warranted.

ACKNOWLEDGEMENTS

We would like to thank prof. dr. M.W.M. van den Brekel, drs. A. Navran, dr. R.P. Takes, dr. G.B. Halmos, dr. L.A. van der Velden, drs. F. Smit, drs. W.W. Braunius, dr. C.A. Meeuwis, dr. K.W. Kross for completing and returning the questionnaire, as representatives of their respective head and neck cancer centers.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer.* 2010;127(12):2893-2917.
2. Dutch Head and Neck Oncology Cooperative Group. Guidline Oral Cavity- and Oropharyngeal Squamous Cell Carcinoma. 2004. http://www.nvpc.nl/uploads/stand/3702_02_04_Hoofd.pdf.
3. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
4. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.
5. 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(6):567-578.
6. Yom SS, Machtay M, Biel MA, et al. Survival impact of planned restaging and early surgical salvage following definitive chemoradiation for locally advanced squamous cell carcinomas of the oropharynx and hypopharynx. *Am. J. Clin. Oncol.* 2005;28(4):385-392.
7. Roosli C, Studer G, Stoeckli SJ. Salvage treatment for recurrent oropharyngeal squamous cell carcinoma. *Head Neck.* 2010;32(8):989-996.
8. Valentino J, Spring PM, Shane M, Arnold SM, Regine WF. Interval pathologic assessments in patients treated with concurrent hyperfractionated radiation and intraarterial cisplatin (HYPERRADPLAT). *Head Neck.* 2002;24(6):539-544.
9. De Bree R, Van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncol.* 2009;45(4-5):386-393.
10. Klabbers BM, Lammertsma AA, Slotman BJ. The value of positron emission tomography for monitoring response to radiotherapy in head and neck cancer. *Mol. Imaging Biol.* 2003;5(4):257-270.
11. Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol.* 2003;45(3):169-184.
12. Ross BD, Moffat BA, Lawrence TS, et al. Evaluation of cancer therapy using diffusion magnetic resonance imaging. *Mol. Cancer Ther.* 2003;2(6):581-587.
13. Wong LY, Wei WI, Lam LK, Yuen AP. Salvage of recurrent head and neck squamous cell carcinoma after primary curative surgery. *Head Neck.* 2003;25(11):953-959.
14. Ojiri H, Mendenhall WM, Mancuso AA. CT findings at the primary site of oropharyngeal squamous cell carcinoma within 6-8 weeks after definitive radiotherapy as predictors of primary site control. *Int. J. Radiat. Oncol. Biol. Phys.* 2002;52(3):748-754.
15. Van den Broek GB, Rasch CR, Pameijer FA, Peter E, van den Brekel MW, Balm AJ. Response measurement after intraarterial chemoradiation in advanced head and neck carcinoma: magnetic resonance imaging and evaluation under general anesthesia? *Cancer.* 2006;106(8):1722-1729.
16. Nomayr A, Lell M, Sweeney R, Bautz W, Lukas P. MRI appearance of radiation-induced changes of normal cervical tissues. *Eur. Radiol.* 2001;11(9):1807-1817.

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
18. Kitagawa Y, Nishizawa S, Sano K, et al. Prospective comparison of 18F-FDG PET with conventional imaging modalities (MRI, CT, and 67Ga scintigraphy) in assessment of combined intraarterial chemotherapy and radiotherapy for head and neck carcinoma. *J. Nucl. Med.* 2003;44(2):198-206.
19. Moeller BJ, Rana V, Cannon BA, et al. Prospective risk-adjusted [18F]Fluorodeoxyglucose positron emission tomography and computed tomography assessment of radiation response in head and neck cancer. *J. Clin. Oncol.* 2009;27(15):2509-2515.
20. Krabbe CA, Pruim J, Dijkstra PU, et al. 18F-FDG PET as a routine posttreatment surveillance tool in oral and oropharyngeal squamous cell carcinoma: a prospective study. *J. Nucl. Med.* 2009;50(12):1940-1947.
21. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(1):181-200.
22. Yamamoto Y, Wong TZ, Turkington TG, Hawk TC, Coleman RE. Head and neck cancer: dedicated FDG PET/CT protocol for detection--phantom and initial clinical studies. *Radiology.* 2007;244(1):263-272.
23. Rodrigues RS, Bozza FA, Christian PE, et al. Comparison of whole-body PET/CT, dedicated high-resolution head and neck PET/CT, and contrast-enhanced CT in preoperative staging of clinically M0 squamous cell carcinoma of the head and neck. *J. Nucl. Med.* 2009;50(8):1205-1213.
24. Lindholm H, Johansson O, Jonsson C, Jacobsson H. The distribution of FDG at PET examinations constitutes a relative mechanism: significant effects at activity quantification in patients with a high muscular uptake. *Eur. J. Nucl. Med. Mol. Imaging.* 2012;39(11):1685-1690.
25. Boellaard R, Oyen WJ, Hoekstra CJ, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur. J. Nucl. Med. Mol. Imaging.* 2008;35(12):2320-2333.
26. Parysow O, Mollerach AM, Jager V, Racioppi S, San RJ, Gerbaudo VH. Low-dose oral propranolol could reduce brown adipose tissue F-18 FDG uptake in patients undergoing PET scans. *Clin. Nucl. Med.* 2007;32(5):351-357.
27. Agrawal A, Nair N, Baghel NS. A novel approach for reduction of brown fat uptake on FDG PET. *Br. J. Radiol.* 2009;82(980):626-631.
28. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
29. Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1098-1107.
30. Abdel Razek AA, Kandeel AY, Soliman N, et al. Role of diffusion-weighted echo-planar MR imaging in differentiation of residual or recurrent head and neck tumors and posttreatment changes. *AJNR Am. J. Neuroradiol.* 2007;28(6):1146-1152.
31. Tshering Vogel DW, Zbaeren P, Geretschlaeger A, Vermathen P, De Keyzer F, Thoeny HC. Diffusion-weighted MR imaging including bi-exponential fitting for the detection of recurrent or residual tumour after (chemo)radiotherapy for laryngeal and hypopharyngeal cancers. *Eur. Radiol.* 2013;23(2):562-569.

32. Vandecaveye V, De Keyzer F, Nuyts S, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):960-971.
33. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. *Clin. Otolaryngol.* 2008;33(3):210-222.
34. Kapoor V, Fukui MB, McCook BM. Role of 18FFDG PET/CT in the treatment of head and neck cancers: posttherapy evaluation and pitfalls. *AJR Am. J. Roentgenol.* 2005;184(2):589-597.

CHAPTER 6

LOCAL RESPONSE EVALUATION AFTER CHEMORADIOTHERAPY FOR ADVANCED OROPHARYNGEAL SQUAMOUS CELL CARCINOMA USING ^{18}F-FDG-PET-CT AND DIFFUSION-WEIGHTED MRI: THE REACTION STUDY

C.S. Schouten
O.S. Hoekstra
P. de Graaf
E.F.I. Comans
B.I. Witte
C.R. Leemans
J.A. Castelijns
R. de Bree

Manuscript in preparation

ABSTRACT

Background. Evaluation of the accuracy of ^{18}Fluoro-2-deoxyglucose positron emission tomography - computed tomography (^{18}F-FDG-PET-CT), diffusion-weighted MRI (DW-MRI) and combined PET-CT/DW-MRI reading to detect local residual disease after chemoradiotherapy (CRT) for oropharyngeal squamous cell carcinoma (OPSCC) and to select patients for an examination under general anesthesia (EUA).

Methods. Prospectively, we included 46 patients scheduled for CRT for advanced staged OPSCC. They underwent response evaluation with ^{18}F-FDG-PET-CT, DW-MRI and EUA three months after completion of CRT. Imaging was assessed by two observers per modality and combined PET-MRI reading was performed. Histopathology and a follow-up of 9 months after CRT served as reference standard.

Results. Local residual disease was diagnosed in 5 patients. EUA had a sensitivity of 60.0% (95% confidence interval (CI): 14.7-94.7%). For PET-CT and DW-MRI, sensitivity was 75.0% (95% CI: 19.4-99.4%) and 60.0% (95% CI: 14.7-94.7%) at a specificity of 82.9% (95% CI: 67.9-92.8%) and 95.1% (95% CI: 83.5-99.4%), yielding NPV of 97.1% (95% CI: 85.1-99.9%) and 95.1% (95% CI: 83.5-99.4%), respectively. Interobserver agreement was highest with PET-CT (κ=0.86 versus κ=0.24 for DW-MRI). Combined PET-CT/DW-MRI reading yielded a sensitivity of 100.0% (95% CI: 39.8-100.0%) at a specificity of 92.7% (95% CI: 80.1-98.5%).

Conclusion. PET-CT and DW-MRI can reliably be used after CRT for OPSCC, to select patients for EUA and to avoid routine EUA. Interobserver variation of DW-MRI needs to be reduced in the post-treatment assessment. Combined PET-CT/DW-MRI reading may be superior to PET-CT and DW-MRI alone but further research should assess the exact role of PET-MRI.

INTRODUCTION

Over the last decades, the introduction of concurrent administration of chemotherapy and radiotherapy has led to survival advantages in patients with advanced staged oropharyngeal squamous cell carcinoma (OPSCC), attributed to improved locoregional control rates [1-3]. Upon failure of chemoradiotherapy (CRT), surgical salvage treatment may be curative if residual disease is detected timely; survival rates rapidly decline with late detection and delayed salvage surgery [4]. Consequently, we and others perform stringent response evaluation, *i.e.* examination under general anesthesia (EUA) with taking biopsies, routinely 3 months after the end of treatment. However, with low incidence of residual disease after CRT, many patients are exposed to needless anaesthesia and eventual biopsies in formerly irradiated areas, potentially inducing inflammation and pain [5,6]. Furthermore, biopsies may be false negative due to sampling errors within the residual mass. This is reflected in the number of missed recurrences of 33% by EUA in the study of Yom et al [4]. Improvement of this clinical dilemma requires a diagnostic strategy which reliably identifies patients who do not need to undergo invasive EUA, without compromising early detection and early salvage in case of residual disease.

Discrimination between malignant and nonspecific changes with computed tomography (CT) and conventional magnetic resonance imaging (MRI) is unreliable because post-treatment changes, including oedema, fibrosis and necrosis, can hamper accurate assessment [7-9]. Currently, functional imaging with ^{18}Fluoro-2-deoxyglucose positron emission tomography-computed tomography (^{18}F-FDG-PET-CT) or diffusion-weighted (DW-MRI) is available for post-treatment assessment [8,10,11]. The combined metabolic and anatomic information of ^{18}F-FDG-PET-CT might improve the accuracy of PET alone. Without anatomical co-registration, PET readings can be difficult to interpret in this context due to the abundance of physiological uptake in combination with disease-related altered anatomy. A meta-analysis in 2011 showed a high pooled sensitivity of ^{18}F-FDG-PET (with or without CT) for detection of local residual/recurrent head and neck squamous cell carcinoma after (chemo)radiotherapy of 79.9%. However, false positive results can occur, due to glucose uptake in inflamed tissue after radiotherapy, leading to a pooled specificity of 87.5% [12].

DW-MRI characterizes the restriction of random motion of water molecules in tissue, which is mainly influenced by the amount of extracellular space [13]. The water mobility can be quantified with the apparent diffusion coefficient (ADC). Hypercellular tissue is characterized by a persistent high signal intensity on DW-MR images obtained with a high b-value (b=1000 sec/mm^2), resulting in low ADC. Hypocellular tissue with necrosis or apoptosis

is characterized by low signal intensity on the b=1000 sec/mm^2 images resulting in high ADC [14]. A diagnostic accuracy varying from 87-98% for DW-MRI was shown in a few studies [11,15-18]. More recently, PET-MRI methodology has become available which facilitates research of the potential added value of either modality [19,20].

The aim of the present study was to prospectively evaluate the accuracy of ^{18}F-FDG-PET-CT, DW-MRI and combined PET-CT/DW-MRI assessment to detect residual local disease after CRT in patients with advanced OPSCC and to avoid unnecessary EUA.

MATERIALS AND METHODS

Patients and study design

Eligible patients were patients with advanced (stage III and IV) resectable histopathologically proven OPSCC scheduled for CRT with curative intent. Exclusion criteria were age below 18 years, pregnancy, a physical condition contra-indicating salvage surgery or a contra-indication for PET-CT or DW-MRI. This prospective observational study was approved by the Medical Ethics Committee and written informed consent was obtained from each study participant. Between May 2012 and September 2014, 65 consecutive patients were included at the department of Otolaryngology-Head and Neck Surgery of the VU University Medical Center. HPV testing was performed on all tumors with our previously defined and validated algorithm for HPV detection [21].

Radiotherapy comprised 70 Gray (Gy) in 2 Gy/fraction delivered to the primary tumor and elective nodal regions received a dose of 54.25-57.75 Gy in 1.55-1.65 Gy/fraction. Forty-five patients received concurrent cisplatin-based CRT and 5 patients cetuximab-based CRT. All patients completed radiotherapy, but toxicity precluded complete chemotherapy in twelve patients.

Response evaluation was performed 3 months after completion of CRT using ^{18}F-FDG-PET-CT, DW-MRI and EUA, irrespective of the level of clinical suspicion of residual disease. PET-CT and DW-MRI were performed prior to the EUA to avoid any influence from the biopsies on the images. The results of the scans were known to the head and neck surgeon at the time of EUA to guide the biopsy location. Whether or not to take biopsies during EUA was decided by the treating physician on basis of clinical findings during EUA. Follow-up after treatment was minimal 9 months in all patients. During follow-up, patients were regularly examined according to our standard head-and-neck-oncology protocol by a multidisciplinary team (including head and neck surgeons, radiation oncologists and medical oncologists); in the first year every

6-8 weeks. Additional imaging and EUA were performed at the discretion of the treating physician if there was clinical suspicion of residual disease.

^{18}F-FDG-PET-CT

All patients fasted for at least 6 hours. Mean serum glucose levels were 5.7 mmol/l (SD 0.7). PET-CT was started at 63.5 minutes (SD 3.6) after intravenous injection of 137-405 MBq ^{18}F-FDG, depending on the body mass index [22]. Low-dose CT scanning was performed with 120 kV and 50 mAs prior to emission scanning. PET-CT (Gemini TF-64CT; Philips Healthcare, Cleveland, OH, USA; 3D-mode; 4 min emission scans/bed position) was performed using a dedicated head and neck protocol (scan trajectory jugulum-orbit; arms down). Afterwards, a whole body protocol (scan trajectory mid femur to cranial vault; arms up) was performed (3D-mode, 2 min emission scan/bed position). PET-CT data were reconstructed using a time of flight row-action maximum likelihood algorithm, as implemented by the vendor. Final image matrix size was 288x288 with a voxel size of 2x2x2 mm. Post reconstruction image resolution was 5 mm full width at half maximum (FWHM).

DW-MRI

DW-MRI was performed using either a 1.5 Tesla MR system (Signa HDx; GE Healthcare, Milwaukee, Wisconsin, USA) or a 3T MR system (Achieva; Philips, Eindhoven, the Netherlands) with a head coil combined with a phased array spine and neck coil. Within a patient, baseline and follow-up scans were always performed on the same system. After an axial short TI inversion-recovery (STIR)-series with 7-mm sections covering the entire neck area, subsequent images were centred on the area of interest containing the primary tumor. Axial STIR and pre- and post-gadolinium (gadoteric acid, Dotarem; Guerbet, Roissy, France or gadobutrol, Gadovist; Bayer Schering AG, Berlin, Germany) T1-weighted (T1WI) spin-echo images with 4-mm sections were obtained in all patients. On 3T, additional T2WI spin echo images were acquired.

DWI was obtained with the same slices at the same slice position as the axial anatomical MR-images. With Signa HDx, DWI was obtained with echo-planar imaging (EPI). Parameters for EPI were: TR/TE=5600/105 ms, in-plane pixel size=2x2 mm, and b-values=0, 750 and 1000 s/mm^2 (3 averages). With Achieva, DWI was acquired with EPI using: TR/TE=4692/64 ms, acquisition matrix 120x116 and at least four b-values (0, 50, 500, and 1000 s/mm2) (3 averages). ADC maps were generated by using the software of the scanner using all b-values.

Data analysis

Both DW-MRI and PET-CT were visually assessed, *e.g.* qualitative analysis without quantitative review (ADC or Standardized Uptake Value (SUV)). PET-CT images were interpreted in random order by 2 nuclear medicine physicians (both with 17 years of experience; O. Hoekstra, E. Comans) and DW-MRI images by 2 radiologists (with 21 and 6 years of experience; J. Castelijns, P. de Graaf). PET-CT images were displayed on a standard workstation allowing simultaneous viewing coronal, sagittal and transverse viewing, with cross-referencing, as well as a three dimensional rotation projection. DW-MRI images were analysed on PACS (Sectra RIS/PACS version 12, Sectra Imtec AB, Linköping, Sweden) that allowed simultaneous viewing of multiple examinations. Clinical information was provided about TNM stage and readers had access to the pretreatment scans of the respective modalities, but were blinded to response evaluation scans of the other modality and to clinical outcome.

In a first reading session, the observers individually and independently scored the site of the primary tumor. The PET-CT observers interpreted the images using the Hopkins-criteria [10]; a qualitative 5-point scale, using the ^{18}F-FDG activity in the internal jugular vein as a reference for background blood pool. We dichotomized these criteria considering scores 1, 2 and 3 as tumor negative and scores 4 and 5 as tumor positive. DW-MRI images were assessed in conjunction with the conventional MRI. A residual mass was considered 'positive for residual disease' if a hyperintense signal on the b=1000 sec/mm^2 images was present, in an area of contrast-enhancement on post-contrast T1WI and corresponding to low diffusion coefficients in the corresponding ADC-map. Absence of low diffusion coefficients on the ADC map was considered 'negative for residual disease' [23]. Cases were classified using a five-point ordinal ('Likert'-) scale (0=no lesion, 1=definitively benign, 2=probably benign, 3=equivocal, 4=probably malignant or 5=definitively malignant). This Likert scale was reduced to a binominal 'sensitive' (0-2=negative, 3-5=positive) and a 'conservative' result (0-3=negative, 4-5=positive) to obtain accuracy data as a function of test positivity criteria.

In a second reading session, consensus scores were provided, after discussion between the interpreters, for each modality separately in case DW-MRI or PET-CT readings showed discrepancies during the first reading session at the dichotomy level. Thereafter, in a final reading session, a radiologist and a nuclear medicine physician jointly evaluated discrepant PET-CT and DW-MRI readings in consensus to achieve a combined PET-CT/DW-MRI reading. The analysis was performed with simultaneously visual correlation of PET-CT and

DW-MRI images. Observers did not receive feedback until all readings were completed.

The localisations of the residual local mass were marked on an anatomical template of the head and neck area and compared to the localisation of histopathological specimens. We classified patients as having local residual disease if at 9 months after completion of CRT, histopathological proof of viable cancer in the biopsy had been found or local residual disease was clinically suspected and growth during follow-up or at serial imaging was present. The results were used to investigate the potential yield (number of correctly diagnosed residues and futile indications for EUA, *i.e.* tumor negative on EUA and during follow-up) of several diagnostic strategies: EUA without any diagnostic imaging, EUA based on results of PET-CT, DW-MRI or combined PET-CT/DW-MRI.

Statistics

Statistical analyses were performed using SPSS software package (version 22.0; 2013, IBM Corp., Armonk, NY, USA). The level of significance was set at $p<0.05$ and hypotheses were tested two-sided. Exact 95% confidence intervals (CIs) were computed for the accuracy measures. Cohen's Kappa and a 2x2 table [24] were used to evaluate concordance of agreement between the observers for the dichotomous system and the sensitive and conservative scale. The interpretations of kappa are [25]: <0: less than chance agreement, 0.01-0.19: slight, 0.20-0.39: fair, 0.40-0.59: moderate, 0.60-0.79: substantial and 0.80-0.99: almost perfect agreement.

RESULTS

Patients

Of the 50 eligible patients, response evaluation imaging was not performed in one patient due to his physical condition. Three patients were excluded due to poor quality of the response evaluation DW-MRI. In conclusion, data of 46 patients were used in the analysis. Patient characteristics are shown in Table 1 and a flow diagram of the participants is shown in Figure 1. PET-CT, DW-MRI and EUA were obtained within a mean of 12.2 weeks (range 10-15 weeks), 12.1 weeks (range 10-17 weeks) and 15.0 weeks (range 11-23 weeks) after completing CRT, respectively.

Pretreatment DW-MRI and PET-CT were available in all and 26 patients, respectively. DW-MRI was performed using Signa HDx MR-system in 37 patients and Philips Achieva MR-system in 9. The acquisition of PET-CT after CRT failed in one patient because of his physical condition. Five of the 46 patients (10.9%; all HPV-negative) were diagnosed with residual local disease at a median of 23 weeks after completed CRT (range 13-31 weeks). In four patients, tumor cells were histopathologically confirmed. In the remaining patient with residual disease (clinically suspicious and growth at serial MR imaging), local residual disease was not confirmed by an invasive procedure because of the extent of disease (*i.e.* regional and distant metastases). Mean volume of the residual detected disease was 3.3 cm^3, as measured on MRI.

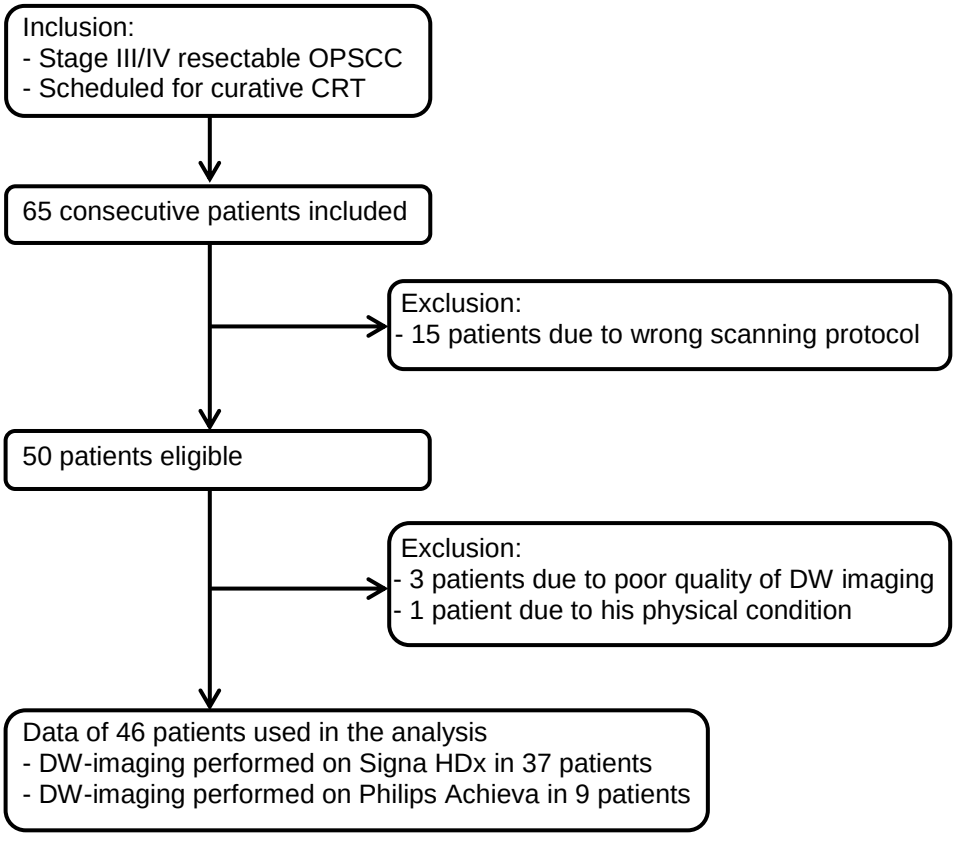


Figure 1.
Flow diagram of study participants

Table 1. Patient and tumor characteristics

Characteristic	*No. of patients (%) (n=46)*
Gender	
Male	35 (76.1%)
Female	11 (23.9%)
Mean age at diagnosis, years (range)	60.4 (44-71)
Oropharyngeal subsite	
Base of tongue	22 (47.8%)
Tonsil	19 (41.3%)
Oropharynx nos	5 (10.9%)
HPV-status	
Positive	20 (43.5%)
Negative	26 (56.5%)
T-stage	
1-2	20 (43.5%)
3	12 (26.1%)
4a	14 (30.4)
N-stage	
0	5 (10.9%)
1	10 (21.7%)
2a	2 (4.3 %)
2b	20 (43.5%)
2c	9 (19.6%)
M-stage	
0	46 (100%)
1	0 (0%)
Smoking	
Never (0-5 pack years)	9 (19.6%)
Moderate (6-24 pack years)	8 (17.4%)
Heavy (>24 pack years)	29 (63.0%)
Alcohol consumption	
Never (0)	3 (6.5%)
Moderate (1-149 unit years)	25 (54.3%)
Heavy (>149 unit years)	18 (39.1%)

Smoking was defined in pack years (1 pack year=20 cigarettes a day during 1 year)
Alcohol consumption was defined in unit years (1 unit year=one alcohol-containing consumption a day during 1 year).
Abbreviations: CRT; chemoradiotherapy, HPV; Human papillomavirus, no; number, nos; not otherwise specified

EUA

EUA for response evaluation showed 3 true-positives (TP), 41 true-negatives (TN), 0 false-positives (FP) and 2 false-negatives (FN) yielding a sensitivity of 60.0% (95% CI: 14.7-94.7%) (Table 2). In both patients with a false negative result, there were no suspicious mucosal lesions in the area of the previous primary tumor. However, in both cases biopsies were taken but were negative. On PET-CT, residual disease was suspected in both these patients and the radiologists considered DW-MRI as 'suspicious' and 'equivocal', respectively. Residual disease was histopathologically confirmed two months later in one of these patients and it became clinically evident three months later in the other.

PET-CT

Consensus readings of PET-CT using the dichotomous system, resulted in 3 TP, 34 TN, 7 FP and 1 FN, respectively (Figures 2-4). These results showed a sensitivity of 75.0% (95% CI: 19.4-99.4%) and a specificity of 82.9% (95% CI: 67.4-92.8%). Accuracy was independent of the availability of baseline PET-CT. In 2 of the 7 false positive readings, the nuclear medicine physicians did not agree; both cases were considered negative for tumor (likely postradiation inflammation) by one observer and positive by the other. In the consensus session, both lesions were classified as likely malignant. In the other five cases, both readers considered the uptake positive for tumor. PET-CT yielded one false negative result. In this case PET-CT showed moderately increased uptake in the oropharynx, which was classified as 'likely postradiation inflammation' by both observers. Residual disease was confirmed with a biopsy 8 months after CRT. PET-CT readers had 'almost perfect' (κ=0.86) and 'fair' agreement (κ=0.35) using the dichotomized and the 5-point Hopkins criteria, respectively (Table 3).

DW-MRI

Consensus dichotomously read DW-MRI showed 2 false negatives and 2 false positives resulting in a sensitivity of 60.0% (95% CI: 14.7-94.7%) at a specificity of 95.1% (95% CI: 83.5-99.4%). The scans of the two false negative patients showed a residual mass (>10 mm), which was considered 'benign' and 'equivocal' by the individual readers and 'likely benign' in the consensus session. Residual disease was confirmed 5 and 7 months after CRT. Using the DW-MRI Likert scale, three and five of the 46 initial readings (per observer)

were equivocal; the diagnostic performance of DW-MRI Likert scale seemed to vary as a function of interpretation of the test positivity criteria. DW-MRI readings had 'fair' interobserver agreement with the dichotomous reading (κ=0.24) as well as with the conservative (κ=0.29) and sensitive readings (κ=0.33).

PET-CT/DW-MRI

Combined diagnostic reading of PET-CT/DW-MRI showed 4 TP, 38 TN, 3 FP and 0 FN resulting in a sensitivity of 100.0% (95% CI: 39.8-100.0%) at a specificity of 92.7% (95% CI: 80.1-98.5%). By using this combination to select patients for EUA after CRT, 84.4% of the EUAs, could be avoided with none of the residual tumors missed. The number of false negative results at single modality assessment decreased in the combined reading because these patients were marked as 'suspicious' by the other imaging modality. Concerning the three false positive results, the original PET-CT readings were at least 'likely residual tumor' versus 'probably benign' on DW-MRI. In the PET-MRI session, the observers decided to consider the examinations 'likely residual disease', because anatomical co-interpretation of the PET-images with MRI confirmed increased ^{18}F-FDG uptake at the site of the original primary tumor. Biopsies were taken during EUA for response evaluation; both showing signs of inflammation but no residual squamous cell carcinoma and follow-up was uneventful.

Table 2. Diagnostic accuracy of PET-CT, DW-MRI, combined PET-CT/DW-MRI reading and EUA.

Scales		*PET-CT*			*DW-MRI*			*PET-CT & DW-MRI*	*EUA*
		Obs 1	*Obs 2*	*Cons*	*Obs 1*	*Obs 2*	*Cons*	*Cons*	
***Dichoto-mous**[≠]*	*Sens*	75.0 [19.4-99.4]	75.0 [19.4-99.4]	75.0 [19.4-99.4]	60.0 [14.7-94.7]	20.0 [0.5-71.6]	60.0 [14.7-94.7]	100.0 [39.8-100.0]	60.0 [14.7-94.7]
	Spec	87.8 [73.8-95.9]	82.9 [67.9-92.8]	82.9 [67.9-92.8]	95.1 [83.5-99.4]	97.6 [87.1-99.9]	95.1 [83.5-99.4]	92.7 [80.1-98.5]	100.0 [91.4-100.0]
	PPV	37.5 [8.5-75.5]	30.0 [6.7-65.2]	30.0 [6.7-65.2]	60.0 [14.7-94.7]	50.0 [1.3-98.7]	60.0 [14.7-94.7]	57.1 [18.4-90.1]	100.0 [29.2-100.0]
	NPV	97.3 [85.8-99.9]	97.1 [85.1-99.9]	97.1 [85.1-99.9]	95.1 [83.5-99.4]	90.9 [78.3-97.5]	95.1 [83.5-99.4]	100.0 [90.7-100.0]	95.3 [84.2-99.4]
Likert conser-vative*	*Sens*	NA	NA	NA	40.0 [5.3-85.3]	20.0 [0.5-71.6]	NA	NA	NA
	Spec	NA	NA	NA	95.1 [83.5-99.4]	97.6 [87.1-99.9]	NA	NA	NA
	PPV	NA	NA	NA	50.0 [6.8-93.2]	50.0 [1.3-98.7]	NA	NA	NA
	NPV	NA	NA	NA	92.9 [80.5-98.5]	90.9 [78.3-97.5]	NA	NA	NA
Likert sensitive*	*Sens*	NA	NA	NA	80.0 [28.4-99.5]	80.0 [28.4-99.5]	NA	NA	NA
	Spec	NA	NA	NA	92.7 [80.1-98.5]	92.7 [80.1-98.5]	NA	NA	NA
	PPV	NA	NA	NA	57.1 [18.4-90.1]	57.1 [18.4-90.1]	NA	NA	NA
	NPV	NA	NA	NA	97.4 [86.5-99.9]	97.4 [86.5-99.9]	NA	NA	NA

Values are presented as point estimates and [95% confidence interval]

[≠] For PET-CT, the Hopkins criteria were reduced to a 2-point scale (1,2,3 are considered tumor negative and 4,5 are considered tumor positive). For DW-MRI, the radiologists were requested to classify the scans dichotomously.

*For DW-MRI, the Likert scale was reduced to a 2-points scale as 'Likert conservative' (1-3=negative, 4-5=positive) and 'Likert sensitive' (1-2=negative, 3-5=positive)

Abbreviations: cons; consensus, EUA; examination under anaesthesia, obs; observer, NA; not applicable, NPV; negative predictive value, PPV; positive predictive value, sens; sensitivity, spec; specificity

Table 3. Interobserver agreement for the dichotomous system and the Likert conservative and sensitive classification

			Dichotomous [≠]			***Likert conservative***			***Likert sensitive***		
			Obs 2			**Obs 2**			**Obs 2**		
			+	**-**	*Kappa*	**+**	**-**	*Kappa*	**+**	**-**	*Kappa*
PET-CT	**Obs 1**	+	8	0	0.86 [0.68-1.0]	NA			NA		
		-	2	35							
DW-MRI	**Obs 1**	+	1	4	0.24 [0.00-0.68]	1	3	0.29 [0.00-0.79]	3	4	0.33 [0.00-0.69]
		-	1	40		1	41		4	35	

Values are presented as the number of patients and as kappa with [95% confidence interval]

[≠] For PET-CT, the Hopkins criteria were reduced to a 2-point scale (1,2,3 are considered tumor negative and 4,5 are considered tumor positive). For DW-MRI, the radiologists were requested to classify the scans dichotomously.

Abbreviations: obs; observer, NA; not applicable

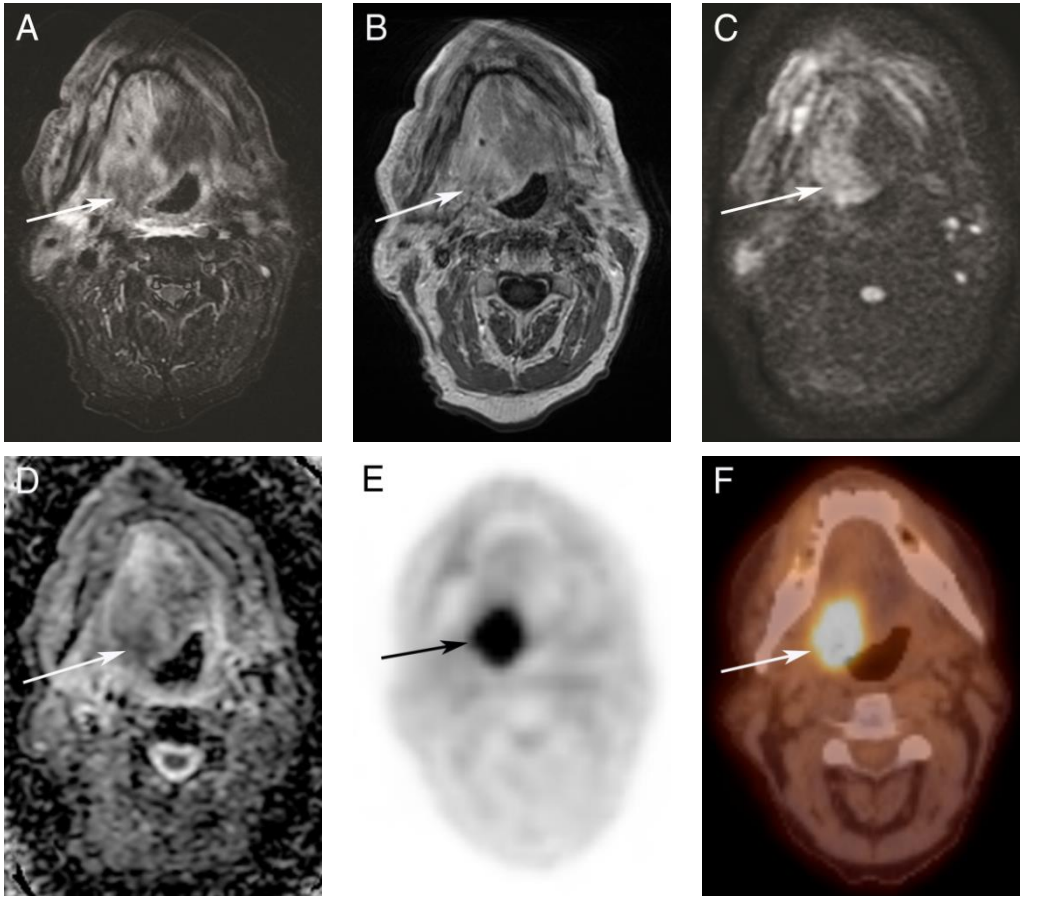


Figure 2.
True positive DW-MRI and ^{18}F-FDG-PET-CT in a 63-year old male patient with a T4aN1 oropharyngeal carcinoma on the right side, 12 weeks after CRT. Residual viable carcinoma was histopathologically proven in a biopsy. Axial STIR MR image **(A)** shows extensive postradiation changes with oedema in the oropharynx and an heterogeneous aspect of the tumor mass on contrast enhanced T1-weighted image **(B)**. DWI b=1000 sec/mm^2 **(C)** shows persistent hyperintensity, corresponding to a hypo intensity on the ADC-map **(D)**. Axial PET (CT-based attenuation correction (CTAC)) **(E)** and fused PET-CT **(F)** shows a focus of intensely increased ^{18}F-FDG uptake in the oropharynx.

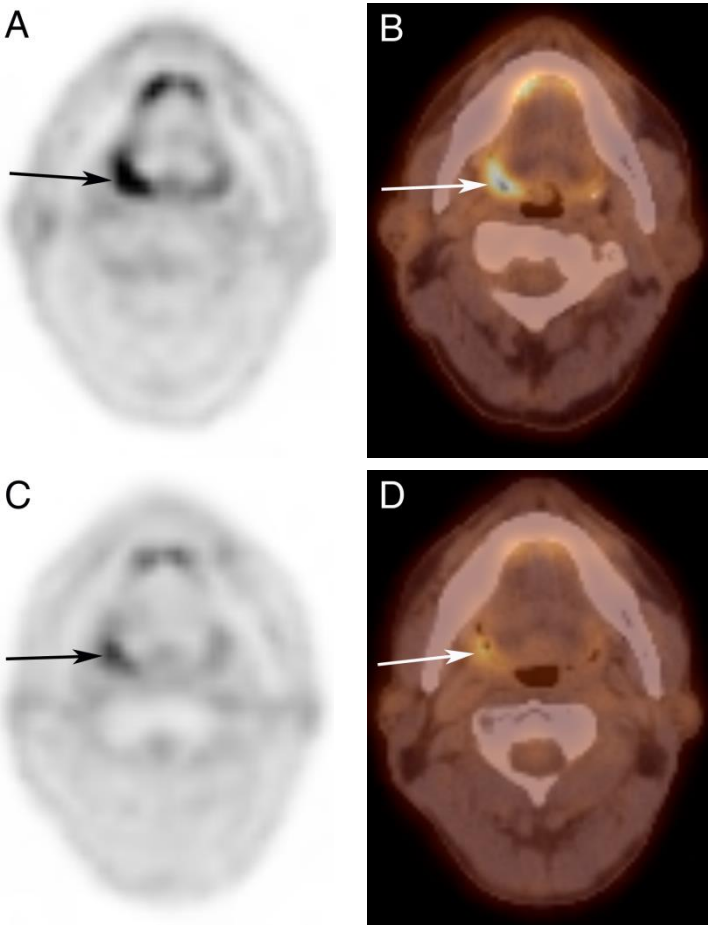


Figure 3.
False positive ^{18}F-FDG-PET-CT at response evaluation in a 54-year old male patient with a T3N2b oropharyngeal carcinoma on the right side. Pre-treatment axial PET (CT-based attenuation correction (CTAC)) **(A)** and fused PET-CT **(B)** shows a focus of intensely increased ^{18}F-FDG uptake in the oropharynx. CTAC **(C)** and PET-CT **(D)** 12 weeks after CRT shows a persistent asymmetrical ^{18}F-FDG uptake higher than the liver which was scored as 'likely residual tumor' by the PET-CT observers. Biopsy showed squamous epithelial cells without malignancy. During follow-up another PET-CT was made which showed complete metabolic response.

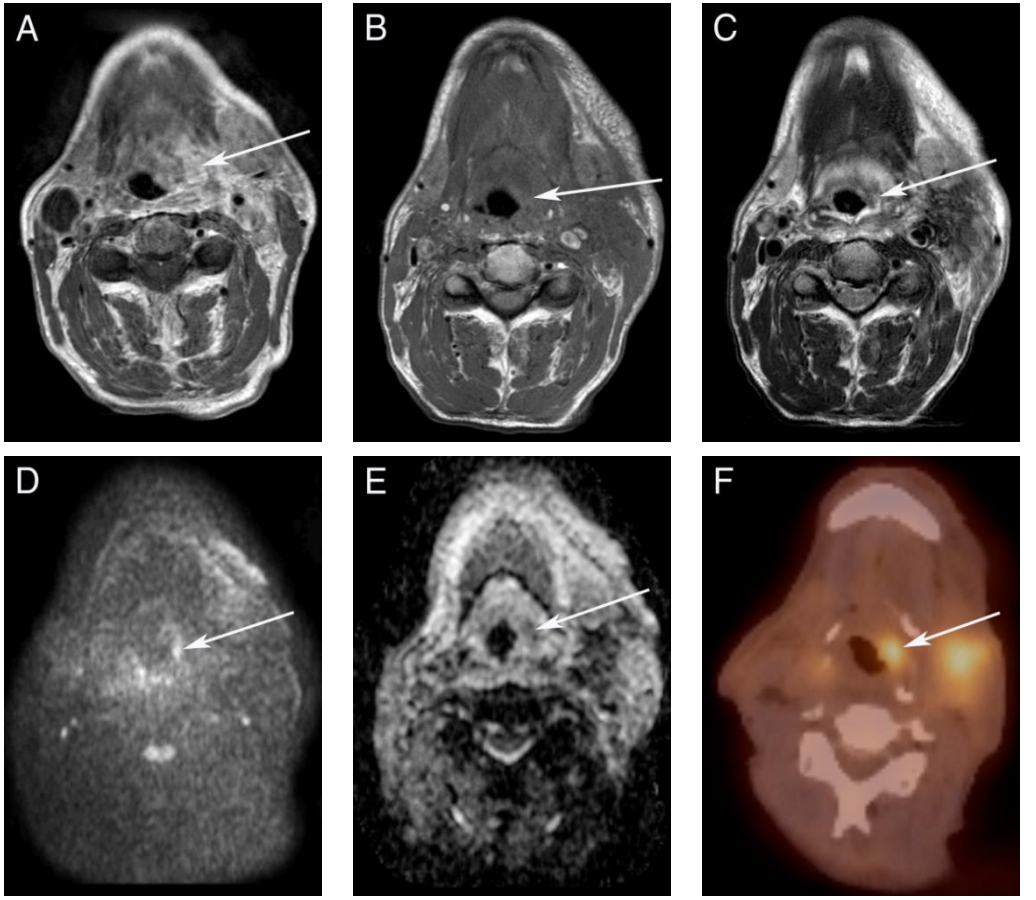


Figure 4.
False negative DW-MRI in a 52-year old patient with a T3N2c vallecula-carcinoma on the left side, 12 weeks after CRT. Pre-treatment axial contrast-enhanced T1-weighted image **(A)** shows a tumor with involvement of the piriform sinus. T1-weighted **(B)** and T2-weighed image **(C)** at response evaluation show a residual mass. DWI b=1000 sec/mm^2 **(D)** shows a small area of hyperintensity, corresponding to low ADC-values on the ADC-map **(E)**. The radiologists did not agree, but ultimately decided the area of hyperintensity on the b=1000 sec/mm^2 was very small and considered the scan as 'likely complete response'. A ^{18}F-FDG-PET-CT **(F)** was performed 2 months after response evaluation, showing a focus of intensely increased ^{18}F-FDG uptake. Residual viable carcinoma was histopathologically proven in a biopsy.

DISCUSSION

A national survey in the Netherlands showed that there is no uniform strategy regarding response evaluation after CRT for OPSCC [26]. EUA is considered to be the most reliable procedure to detect residual disease. However, due to the high response rates after CRT (*i.e.* locoregional control rate 1 year after CRT is reported to be around 90%) many patients are exposed to unnecessary EUA and biopsies [5,6]. Improvement of this situation requires a diagnostic strategy which reliably selects patient for invasive diagnostic procedures without compromising the intended benefit of salvage surgery.

Conventional MRI can be performed as radiologic measurement of tumor response after CRT. Van den Broek et al. reported a sensitivity and specificity of 48% and 85% (yielding PPV and NPV of 55% and 81%, respectively, at a prevalence of 7%) for post-treatment MRI performed 6-8 weeks after CRT in 82 patients with advanced HNSCC [27]. DWI can be incorporated in MRI protocols as a complementary MR sequence, taking only a few minutes extra. In the current study, the implementation of DWI led to an improvement of the accuracy of conventional MRI imaging which resulted in a sensitivity, specificity, PPV and NPV of 60%, 95%, 60% and 95%, respectively.

However, DWI is susceptible to artefacts and geometric distortions, particularly in the inhomogeneous head and neck area. Several quantitatively assessed studies in HNSCC using ADC-values for the purpose of monitoring treatment response [11,16] or detecting tumor recurrence [15,16,18] showed good results with accuracies varying from 80 to 97%. However, only a few studies in head and neck cancer were issued in the literature using qualitative evaluation of the primary site and the level of inter-reader reliability is unknown. Consequently, no established visual interpretation system is described to classify post-treatment DW-MRI. During visual assessment of a scan, a validated interpretation system could assist in clarifying equivocal findings found and could improve systematic and reproducible review. We performed qualitative analysis in analogy with the interpretation system described by Thoeny et al. [28] and performed in this manner by Tshering Vogel et al [18]. Tshering Vogel et al. showed an excellent accuracy in the detection of recurrent/residual hypopharyngeal and laryngeal tumors with a sensitivity of 94% and a specificity of 100%. Compared to this study, our sensitivity is disappointing. This may be explained by the design of the study. We included all patients after CRT, in contrast to Tshering Vogel et al., who included patients with symptoms after CRT. It can be anticipated that symptomatic recurrences are generally larger and more easy to detect as compared to asymptomatic tumors. DWI is probably less appropriate for the recognition of small tumor residues due to its relative low spatial resolution. Besides, our

results showed low interobserver agreement, implying that radiologists are still in the learning curve concerning post-treatment evaluation and future research should address these interpretation criteria for test positivity in DW-MRI.

Post-treatment assessment with ^{18}F-FDG-PET-CT is not routinely performed for response evaluation, but interobserver agreement was much higher for PET-CT than for DW-MRI. Marcus et al. developed the Hopkins Criteria with substantial inter-reader agreement in their retrospective series with 214 HNSCC patients treated with (chemo)radiotherapy [10]. In our hands, the 5-point Hopkins Criteria showed 'fair' agreement, but the dichotomised Hopkins criteria showed 'almost perfect' agreement. These head and neck PET-CT interpretation criteria should be further confirmed in large prospective studies with uniform tumor types and treatment regimens. Also, the optimal timing to assess response with PET-CT after CRT remains a matter of debate, since radiation-induced inflammatory responses can lead to increases in ^{18}F-FDG-uptake and thereby to false positive results. This was also the case in present study, in which PET-CT yielded a considerable amount of false positive results, resulting in a specificity of 82.9%. However, considering salvage surgery as an curative option, early detection of residual disease during follow-up is crucial. Delayed detection may lead to unresectability or distant metastasis, leading to decreased survival rates.

Our data showed that a combined reading with PET-CT and DW-MRI, *i.e.* achieved consensus between the radiologist and nuclear medicine physician, seemed to be able to yield higher sensitivity without compromising the high specificity of DW-MRI alone; 80% of all EUAs could be avoided without losing the option of timely salvage surgery. To our knowledge, these data are the first to suggest that combined PET-CT/DW-MRI in the post-treatment setting in patients with OPSCC may be superior to PET-CT and DW-MRI alone. However, due to the low rate of local residues, these results should be interpreted carefully.

EUA with biopsies for the post-therapy assessment of residual disease in all patients who were treated with CRT for advanced OPSCC had limited value. A negative biopsy is not sufficient to exclude residual tumor as it can miss residual tumor directly under the mucosa due to sampling error. Moreover, the knowledge of the results of the imaging techniques at the start of EUA could have introduced a bias because this may have caused an advantage in favour of EUA to detect local residual disease.

We acknowledge several limitations to this study. First, although we focused on patients with enhanced risk of residual disease (stage III and IV), the prevalence of residual squamous cell carcinoma was only 10.9%. Because of this low incidence of residual tumor, identifying a patient group with additional high risk factors might help to design imaging studies on effective

stratification of patients for EUA. HPV-negativity is an established risk factor [29]. Functional imaging could also provide additional prognostic factors; patients with high pretreatment SUV-values generally have less favorable outcome [30] and patients with relatively low pretreatment ADC-values respond better to CRT than tumors with higher pretreatment ADC-values [31]. Secondly, two MR scanners from different vendors with different field strengths were used due to the simultaneous performance of another research imaging study at the VU University Medical Center. To reduce interobserver variation and possibly further improve diagnostic accuracy, quantification of signal decay in DW-MRI (ADC) may be used. Due to the use different MRI scanners, we could not analyse these quantitative parameters because it has been shown that ADC-values differ significantly between MR imaging systems and sequences [32]. However, visual assessment of post-treatment examinations is clinically practicable for both radiologists and referring physicians because it can provide instantaneous evaluation. Finally, we used a relatively short follow-up period (9 months after end of treatment) as a reference standard. However, it is contemplated that residual tumor will manifest itself within this period and in the literature the same period is used as reference standard [10].

CONCLUSION

Due to the high local control rate after CRT for advanced oropharyngeal cancer, the NPV of an imaging test is bound to be high. Nevertheless, PET-CT and DW-MRI have a very high NPV of 97% and 95%, respectively. Therefore, functional imaging with PET-CT or DW-MRI applied as local response evaluation after CRT in patients with OPSCC seem to be reliable tools to rule out residual primary tumor, possibly making routine EUA superfluous. Hybrid-PET-MRI systems are emerging into the clinical field, permitting combined PET and MRI evaluation. In anticipation of this development, our current results suggest that combined reading of PET-CT/DW-MRI may be superior to PET-CT and DW-MRI alone. These data concerning post-treatment imaging analysis need to be confirmed in complementary research, but due to the low rate of residual disease it is unlikely that this clinical issue will be solved by one research-team. Combining individual results in a meta-analysis will probably be necessary to develop post-treatment interpretation criteria and to validate these criteria.

ACKNOWLEDGEMENTS

The authors thank S. Hakim (Department of Otolaryngology-Head and Neck Surgery, VU University Medical Center, Amsterdam) for her assistance in applying the specific PET-CT protocol for study participants.

REFERENCES

1. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
2. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.
3. 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(6):567-578.
4. Yom SS, Machtay M, Biel MA, et al. Survival impact of planned restaging and early surgical salvage following definitive chemoradiation for locally advanced squamous cell carcinomas of the oropharynx and hypopharynx. *Am. J. Clin. Oncol.* 2005;28(4):385-392.
5. Valentino J, Spring PM, Shane M, Arnold SM, Regine WF. Interval pathologic assessments in patients treated with concurrent hyperfractionated radiation and intraarterial cisplatin (HYPERRADPLAT). *Head Neck.* 2002;24(6):539-544.
6. Sher DJ, Thotakura V, Balboni TA, et al. Treatment of oropharyngeal squamous cell carcinoma with IMRT: patterns of failure after concurrent chemoradiotherapy and sequential therapy. *Ann. Oncol.* 2012;23(9):2391-2398.
7. Nomayr A, Lell M, Sweeney R, Bautz W, Lukas P. MRI appearance of radiation-induced changes of normal cervical tissues. *Eur. Radiol.* 2001;11(9):1807-1817.
8. De Bree R, van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncol.* 2009;45(4-5):386-393.
9. Klabbers BM, Lammertsma AA, Slotman BJ. The value of positron emission tomography for monitoring response to radiotherapy in head and neck cancer. *Mol. Imaging Biol.* 2003;5(4):257-270.
10. Marcus C, Ciarallo A, Tahari AK, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins Criteria)-interreader reliability, accuracy, and survival outcomes. *J. Nucl. Med.* 2014;55(9):1411-1416.
11. Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1098-1107.
12. Gupta T, Master Z, Kannan S, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(11):2083-2095.
13. Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol.* 2003;45(3):169-184.
14. Ross BD, Moffat BA, Lawrence TS, et al. Evaluation of cancer therapy using diffusion magnetic resonance imaging. *Mol. Cancer Ther.* 2003;2(6):581-587.
15. Vandecaveye V, De Keyzer F, Nuyts S, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):960-971.

16. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
17. Abdel Razek AA, Kandeel AY, Soliman N, et al. Role of diffusion-weighted echo-planar MR imaging in differentiation of residual or recurrent head and neck tumors and posttreatment changes. *AJNR Am. J. Neuroradiol.* 2007;28(6):1146-1152.
18. Tshering Vogel DW, Zbaeren P, Geretschlaeger A, Vermathen P, De KF, Thoeny HC. Diffusion-weighted MR imaging including bi-exponential fitting for the detection of recurrent or residual tumour after (chemo)radiotherapy for laryngeal and hypopharyngeal cancers. *Eur. Radiol.* 2013;23(2):562-569.
19. Becker M, Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *Br. J. Radiol.* 2014;87(1036):20130677.
20. Castelijns JA. PET-MRI in the head and neck area: challenges and new directions. *Eur. Radiol.* 2011;21(11):2425-2426.
21. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int. J. Cancer.* 2007;121(11):2465-2472.
22. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(1):181-200.
23. Padhani AR, Koh DM. Diffusion MR imaging for monitoring of treatment response. *Magn Reson. Imaging Clin. N. Am.* 2011;19(1):181-209.
24. De Vet HC, Mokkink LB, Terwee CB, Hoekstra OS, Knol DL. Clinicians are right not to like Cohen's kappa. *BMJ.* 2013;346:f2125.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
26. Schouten CS, Hoekstra OS, Leemans CR, Castelijns JA, De Bree R. Response evaluation after chemoradiotherapy for advanced staged oropharyngeal squamous cell carcinoma: a nationwide survey in the Netherlands. *Eur Arch Otorhinolaryngol.* 2015;272(11):3507-3513.
27. Van den Broek GB, Rasch CR, Pameijer FA, Peter E, Van den Brekel MW, Balm AJ. Response measurement after intraarterial chemoradiation in advanced head and neck carcinoma: magnetic resonance imaging and evaluation under general anesthesia? *Cancer.* 2006;106(8):1722-1729.
28. Thoeny HC, De Keyzer F, King AD. Diffusion-weighted MR imaging in the head and neck. *Radiology.* 2012;263(1):19-32.
29. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* 2010;363(1):24-35.
30. Machtay M, Natwa M, Andrel J, et al. Pretreatment FDG-PET standardized uptake value as a prognostic factor for outcome in head and neck cancer. *Head Neck.* 2009;31(2):195-201.
31. Hatakenaka M, Nakamura K, Yabuuchi H, et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81(2):339-345.
32. Kolff-Gart AS, Pouwels PJ, Noij DP, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. *AJNR Am. J. Neuroradiol.* 2015;36(2):384-390.

CHAPTER 7

RESPONSE EVALUATION AFTER CHEMO-RADIOTHERAPY FOR ADVANCED NODAL DISEASE FROM HEAD AND NECK SQUAMOUS CELL CARCINOMA USING DIFFUSION-WEIGHTED MRI AND ^{18}F-FDG-PET-CT

C.S. Schouten
P. de Graaf
F.M. Alberts
O.S. Hoekstra
E.F.I. Comans
E. Bloemena
B.I. Witte
E. Sanchez
C.R. Leemans
J.A. Castelijns
R. de Bree

Oral Oncology 2015; 51: 541-7

ABSTRACT

Background. Evaluation of accuracy and interobserver variation of diffusion-weighted magnetic resonance imaging (DW-MRI) and ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography (^{18}F-FDGPET-CT) to detect residual lymph node metastases after chemoradiotherapy (CRT) in advanced staged head and neck squamous cell carcinoma (HNSCC).

Methods. Retrospectively, routinely performed DW-MRI (n=73) and ^{18}F-FDG-PET-CT (n=58) 3 months after CRT in HNSCC-patients with advanced nodal disease (N2-N3) were assessed by two radiologists and two nuclear medicine physicians (individually and in consensus). Imaging was scored dichotomously and on a five-point Likert scale. We also explored different scenarios for the potential added value of DW-MRI to PET-CT using the consensus Likert scale. Histopathology and a follow-up of 9 months after CRT served as reference standard.

Results. Five patients (7%) had residual regional disease. DW-MRI showed a sensitivity of 60% and a specificity of 93%, vs 100% and 84% for PET-CT, respectively. DW-MRI and PET-CT observers had 'moderate' and 'substantial' interobserver agreement (κ=0.58 and κ=0.64, respectively) with the dichotomous system. The combination of PET-CT and DW-MRI showed a sensitivity of 100% and a specificity of 95%.

Conclusion. The high sensitivity of PET-CT authorizes a neck dissection in all patients with a positive test result and the high specificity of DW-MRI justifies avoidance of invasive neck dissections if the test is negative. Interobserver agreement varied as a function of test positivity criteria. Adding DW-MRI to PET-CT seemed to increase the specificity of PET-CT alone, thereby ensuring that less patients are exposed to unnecessary neck dissections.

INTRODUCTION

Patients with advanced staged head and neck squamous cell carcinomas (HNSCC) are often treated with combined chemotherapy and radiotherapy (CRT) with acceptable locoregional control rates [1-3]. Timely detected residual neck metastases can be treated successfully with a 'salvage' neck dissection [4-6]. However, post-treatment changes including fibrosis, oedema and necrosis may hamper accurate assessment on computed tomography (CT) and anatomical magnetic resonance imaging (MRI) [7-10]. Therefore, clinicians usually rely on functional imaging techniques such as ^{18}F-fluorodeoxy-glucose positron emission tomography-CT (^{18}F-FDG-PET-CT) [11]. In 2008, a meta-analysis on the accuracy of ^{18}F-FDG-PET to detect residual regional disease after CRT in HNSCC showed a pooled sensitivity of 74% and a specificity of 88% [10]. In combination with results of subsequently published ^{18}F-FDG-PET-CT studies, heterogeneity seems to be considerable with sensitivities ranging from 40-100% and specificities from 25-97% (Table 1) [12-19].

More recently, diffusion-weighted MRI (DW-MRI) has been introduced which characterizes tissue based on differences of the diffusion motion of water protons [20]. Motion-probing gradients with certain strength (b-value) induce a signal loss in the images related to the amount and speed of water proton movement. By measuring signal intensities with different b-values, these differences can be quantified using the apparent diffusion coefficient (ADC). After CRT, hypercellular tissue (e.g. residual tumor) is expected to show persistent high signal intensity on DWI-images with high b-value (b1000 images), resulting in low ADCs vs low signal intensity on b1000 images with high ADCs in hypocellular tissue with necrosis (e.g. post-treatment-changes) [21,22]. DW-MRI is promising for differentiating locoregional residual or recurrent head and neck tumors from postradiation changes [23-25].

Currently, further research is done with a combination of ^{18}F-FDG-PET-CT and DW-MRI, which seems to offer potential [26]. Hybrid-PET-MRI systems are now entering the clinical field allowing for combined PET and MRI reading.

The purpose of our study was to evaluate the accuracy and interobserver variation of DW-MRI and ^{18}F-FDG-PET-CT to detect residual lymph node metastases after CRT in patients with advanced (N2-N3) nodal disease and to explore their potential added value.

Table 1. ^{18}F-FDG-PET-CT studies on the detection of regional residual head and neck cancer

Ref.	*N*	*Inclusion*	*Interval*	*Test positivity criteria*	*Follow-up*	*Prevalence*	*Sensitivity*	*Specificity*	*PPV*	*NPV*	*Therapy*
12	34	N2	3 mos	Focal areas of moderate to intense FDG uptake	39 mos	0%	-	97%	-	100%	CRT
13	48	N2	≥8 wks	Unknown	26 mos	23%	82%	97%	90%	95%	CRT
14	52	N+	Median 12 wks	Presence of FDG avidity	≥9 mos	5%	100%	87.5%	40%	100%	(C)RT
15	65	N+	Median 12 wks	Focal uptake, intensity greater than background	37 mos	11%	71%	89%	38%	97%	CRT
16	43	N2-N3	2-5 mos	Increased metabolic activity and suspicious radiographic characteristics	18 mos, ≥5 mos	19%	88%	91%	70%	97%	CRT
17	30	N0-N3[a]	≥6 wks	Activity above expected background	≥6 mos	27%	100%	70%	36%	100%	CRT
18	17	N2-N3[b]	8-10 wks	Areas of increased uptake more intense than background	10 mos	29%	40%	25%	18%	50%	CRT
19	98	Stage III-IV scc pharynx/ larynx	8 wks	SUV_{max}	92 wks	8%	75%	76%	27%	96%	(C)RT

Ref: reference number, N: number of patients included, inclusion: inclusion criteria, interval: time between end of treatment and PET-CT, test positivity criteria: method of PET-CT assessment for test positivity, follow-up: median follow-up from the date of completion of therapy, prevalence: percentage of patients with regional residual disease

Abbreviations: CRT: chemoradiotherapy, mos: months, NPV: negative predictive value, PPV: positive predictive value, scc: squamous cell carcinoma, SUV: Standardized Uptake Value, RT: radiotherapy, wks: weeks

[a] Clinically symptomatic for suspicious residual tumor

[b] Planned posttreatment neck dissection in all patients

METHODS

Patients and study design

This single institution retrospective study was approved by the institutional review board and was carried out in accordance with the Helsinki Declaration. Between 2009 and 2013, DW-MRI and/or ^{18}F-FDG-PET-CT were routinely performed as response evaluation 3 months after completion of CRT in 84 patients with histopathologically proven HNSCC and advanced nodal disease (N2-N3).

A radiation dose of 70 Gy in 2 Gy/fraction was delivered and elective nodal regions received a dose of 54.25 – 57.75 Gy in 1.55 – 1.65 Gy/fraction. Seventy-five patients received cisplatin-based CRT and 9 patients cetuximab-based CRT. All patients completed radiotherapy, but toxicity precluded to complete chemotherapy in 10 patients (11.9%).

During follow-up, patients were regularly physically examined according to our standard head-and-neck oncology protocol by a multidisciplinary head and neck oncology team; in the first year every 6-8 weeks, in the second year every 2-3 months, in the third year every 4 months and every 6 months thereafter. Additional investigations during follow-up were performed at the discretion of the attending physician.

DW-MRI

DW-MRI was obtained with either a 1.5 Tesla MR system (Sonata (20% of the patients) or Avanto (10%); Siemens, Erlangen, Germany or Signa HDx (50%); GE Healthcare, Milwaukee, Wisconsin, USA) or a 3T MR system (20%) (Achieva; Philips, Eindhoven, the Netherlands) with a head coil combined with a phased array spine and neck coil. After an axial short TI inversion-recovery (STIR)-series with 7-mm sections covering the entire neck area, subsequent images were centred on the area of interest containing the primary tumor and enlarged lymph nodes. Axial images with 4-mm sections were obtained with STIR and T1-weighted (T1WI) spin-echo before and after the intravenous injection of contrast material.

DWI was obtained in all patients for the same slices at the same slice position as the axial conventional images. With Sonata and Avanto, DWI was obtained with a spin-echo echo-planar imaging (EPI). Parameters for EPI were: TR/TE=5000/105 ms, acquisition matrix 128x128, and b-values=0, 500 and 1000 s/mm^2 (3 averages). With Signa HDx, DWI was performed with EPI or periodically rotated overlapping parallel lines with enhanced reconstruction

(PROPELLER). Parameters for EPI-DWI were: TR/TE=5600/76.2 ms, acquisition matrix 86x128, and b-values=0 s/mm^2 (1 average), 750 s/mm^2 (3 averages) and 1000 s/mm^2 (6 averages). Parameters for PROPELLER-DWI were: TR/TE=3500/83.87 ms, acquisition matrix 128x128, and b-values=0, 750 and 1000 s/mm^2 (3 averages). With Achieva, DWI was acquired with EPI using: TR/TE=4692/64 ms, acquisition matrix 120x116 and at least four b-values (0, 50, 500, and 1000 s/mm2) (3 averages). ADC-maps were generated by using the software of the scanner.

^{18}F-FDG-PET-CT

All patients fasted for at least 6 hours. Mean serum glucose levels were 5.9 mmol/l (SD 1.2). PET-CT was started 60±15 minutes after intravenous injection of 125 – 424 MBq ^{18}F-FDG, depending on the body mass index [27]. Low-dose CT was performed with 120 kV and 50 mAs prior to emission scanning. In two-thirds of the patients, PET-CT (Gemini TF-64CT; Philips Healthcare, Cleveland, OH, USA; 3D-mode; 4 min emission scans/bed position) was performed using a dedicated head and neck protocol (scan trajectory jugular notch-orbit; arms down). PET-CT data were reconstructed using a time of flight row-action maximum likelihood algorithm, as implemented by the vendor. Final image matrix size equals 288x288 with a voxel size of 2x2x2 mm. Post reconstruction image resolution was 5 mm full width at half maximum (FWHM). In the remaining patients, PET-CT consisted of a whole body protocol (scan trajectory mid-femur to cranial vault; arms up), using Gemini TF-64CT or Ingenuity TF 128 PET-
CT (Philips Healthcare, Cleveland, OH, USA) (3D-mode; 2 min emission scans/bed position). Final image matrix size was 144x144 with a voxel size of 2x2x2 mm. Post reconstruction image resolution was 7 mm FWHM.

Data analysis

DW-MRI images were retrospectively interpreted by 2 radiologists (P. de Graaf, J. Castelijns) and PET-CT images by 2 nuclear medicine physicians (O. Hoekstra, E. Comans) with ample experience in head and neck imaging. Clinical information was provided about localisation of the primary tumor and TNM stage at diagnosis. DW-MRI was analysed on PACS (Sectra RIS/PACS version 12, Sectra Imtec AB, Linköping, Sweden) that allowed simultaneous viewing of multiple MRI scans. Up to two lymph nodes with a minimal axial diameter of ≥5mm (as measured on T1-weighted images), diagnosed as malignant on pretreatment assessment, were included and were identified by

visual and slice-position based correlation with the baseline examination. PET-CT images were displayed on a standard workstation allowing simultaneous coronal, sagittal and transverse viewing, with cross-referencing, as well as a three dimensional rotation projections.

Readers had access to the pretreatment scans of the respective modalities (if available), but were blinded to the initial reports, the post-treatment scans of the other modality and to clinical outcome. In a first reading session, the observers individually and independently scored the images. The DW-MRI observers scored the regional status according to the Ojiri criteria [28], i.e. minimal axial diameter, presence of focal defect, extracapsular spread. Thereafter, they assessed signal intensity on native b1000 images and the corresponding ADC map. A residual lymph node was considered 'abnormal', if increased signal intensity was present on the b1000 images corresponding to low ADC-values. In case of necrotic intranodal components, the observers assessed the area of contrast-enhancement in the corresponding post-contrast T1WI [29]. PET-CT observers evaluated regional foci of increased ^{18}F-FDG uptake and scored on aspect and intensity. PET-CT was considered 'abnormal' in case of uptake higher than background in areas representing sites of lymph nodes after correlation with the low-dose CT images. Finally, we requested all observers to classify the scans using:

1. a five-points ordinal ('Likert'-) scale (1= definitively benign, 2= probably benign, 3= equivocal, 4= probably malignant or 5= definitively malignant),
2. a dichotomous system, i.e. suspicious for a regional residue or not.

In a second reading session, consensus scores were provided in case DW-MRI or PET-CT readings had been discrepant at the dichotomy level. Observers did not receive feedback until all readings were completed.

The localisations of the residual lymph nodes were marked on an anatomical template of the head and neck area and compared to the localisation of histopathological specimens. Only matching localisations were counted as true-positive. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for all observers separately were evaluated in assessing the presence of regional residual disease, relative to regional lymph node status.

We classified patients as having residual nodal disease if at 9 months after completion of CRT, histopathological proof of viable cancer in the surgical specimen had been found or cytopathologically in nodes that had been biopsied because of growth at serial MR imaging. Patients were classified as node-negative if at 9 months after completed CRT, there was no histopathological proof of regional residual disease and the size of involved nodes had regressed (clinically and/or radiologically).

In addition, we explored the potential added value of DW-MRI to ^{18}F-FDG-PET-CT. Using the consensus Likert scale for each modality, we investigated different scenarios of DW-MRI and PET-CT combination: Likert PET-CT score ≥ '4' in combination with Likert DW-MRI ≥ '3', PET-CT ≥ '4' in combination with DW-MRI ≥ '4', PET-CT ≥ '3' in combination with DW-MRI ≥ '3' and PET-CT ≥ '3' in combination with DW-MRI ≥ '4'.

Statistics

Statistical analyses were performed using SPSS software package (version 20.0; IBM Corp., Armonk, NY, USA). The level of significance was set at $p<0.05$ and hypotheses were tested two-sided. The Likert scale was reduced to a binominal 'sensitive' (1-2=negative, 3-5=positive) and a 'conservative' interpretation scale (1-3=negative, 4-5=positive) to obtain accuracy data. Cohen's Kappa and a 2 x 2 table [30] were used to evaluate concordance of agreement between the observers for the dichotomous system and the sensitive and conservative scale. The interpretations of kappa are [31]: <0: less than chance agreement, 0.01-0.19: slight agreement, 0.20-0.39: fair agreement, 0.40-0.59: moderate agreement, 0.60-0.79: substantial agreement, 0.80-0.99: almost perfect agreement.

RESULTS

Of the 84 potentially eligible patients, 9 were excluded because DW-MRI was degraded by artefacts or its field of view did not comprise the residual nodal mass. One patient died during follow-up (not cancer related) and another was lost to follow-up. Taken together, we included 73 patients with evaluable DW-MRI's and appropriate follow-up (Table 2). DW-MRI was obtained 86.8 days (mean; SD 10.7) after completed CRT. Although ^{18}F-FDG-PET-CT is routinely performed for response evaluation, in clinical practice not all patients underwent PET-CT; Fifty-eight patients (80%) underwent PET-CT; these had been obtained 92.2 days (mean; SD 10.5) after CRT.

Five of the 73 patients (7%) were diagnosed with residual nodal disease during follow-up (at a median of 30 weeks after completed CRT; range 15-38 weeks). Four of them were eligible for neck dissection and viable tumor cells were histopathologically confirmed. The remaining patient with residual disease (positive fine needle aspiration and growth of the lymph node at serial MR imaging) was unresectable because of the extent of the neck disease. Of the 58 patients undergoing PET-CT as well as DW-MRI, three were diagnosed with residual nodal disease. In 44 of these patients, baseline PET-CT scans

were available for comparison. Baseline MRI scans were available in all patients.

During follow-up, 4 patients had a residual palpable cervical lymph node with positive fine needle aspiration cytology (FNAC) for which a neck dissection was performed. Histopathology of the surgical specimen showed no viable squamous cell carcinoma but only reactive lymph nodes or necrotic tissue.

Table 2. Patient and tumor characteristics

Patient characteristics	***No. of patients** (%) (n=73)*
Gender	
Male	52 (71.2%)
Female	21 (28.8%)
Mean age at diagnosis, years (range)	57.6 (23-73)
Location of primary tumor	
Oropharynx	55 (75.3%)
Hypopharynx	7 (9.6%)
Larynx	4 (5.5%)
Nasopharynx	4 (5.5%)
Oral cavity	3 (4.1%)
N-status	
N2a	11 (15.1%)
N2b	40 (54.8%)
N2c	20 (27.4%)
N3	2 (2.7%)
Treatment method	
Cisplatin-based CRT	66 (90.4%)
Cetuximab-based CRT	7 (9.6%)

Abbreviations: CRT= chemoradiotherapy

Consensus readings of DW-MRI and PET-CT using the *dichotomous system* ('suspicious or not'), resulted in 3 and 3 true-positives (TP), 63 and 46 true-negatives (TN), 5 and 9 false-positives (FP) and 2 and 0 false-negatives (FN), respectively. Examples of true and false positive DW-MRIs and PET-CTs are shown in Figure 1 and 2. In the 58 patients, PET-CT had a sensitivity of 100% (95% confidence interval, CI: 29%-100%) and a specificity of 84% (CI 71%-92%) (Table 3). Consensus dichotomously reading PET-CT yielded 9 false positive results. In retrospect, only PET-CT images showing focal areas with at least moderate ^{18}F-FDG uptake proved to be malignant. Using these response criteria could have increased the specificity of PET-CT in our study to

93%. With the PET-CT Likert scale, few readings were equivocal (1 and 2 per observer, respectively), and none were so after the consensus session. Therefore, PET-CT results were similar for all reading systems (i.e. *conservative* and *sensitive* reading scale). PET-CT observers had substantial agreement in the dichotomous and conservative reading scales (κ=0.64) and moderate agreement with the sensitive reading and Likert scale (κ=0.55 and κ=0.53, respectively; Table 4). Accuracy was independent of the availability of baseline PET-CT or scan protocol.

In the 73 patients, DW-MRI had a sensitivity of 60% (CI 15%-95%) at a specificity of 93% (CI 84-98%). Results were similar in the subgroup of 58 patients in whom a head-to-head comparison with PET-CT was feasible (Table 3). Consensus dichotomously read DW-MRI yielded 2 false negative results. DW-MRI showed a residual lymph node in both patients with an axial diameter of 5.5 and 6.3 mm, respectively, which were considered benign by one observer and malignant by the other. In the joint session the observers did not agree but ultimately decided to consider these lymph nodes as benign. Residual nodal disease was confirmed in the surgical specimen, 8 and 9 months after CRT, respectively.

Consensus dichotomously read DW-MRI yielded 5 false positive results. In two cases, the radiologists did not agree; both cases were considered 'benign' by one of the radiologists and 'malignant' by the other. In consensus the observers decided to consider the lymph nodes 'malignant'. In the other three cases, both radiologists considered the lymph node 'malignant'.

For the DW-MRI Likert scale, 5 and 7 of the initial 73 readings (per observer) were equivocal, with 8 equivocals after the consensus session (11%). The diagnostic performance of MRI seemed to vary as a function of the test positivity criteria. DW-MRI readings had 'moderate' interobserver agreement (dichotomously, κ=0.58; conservative reading, κ=0.54; sensitive reading, κ=0.53) and fair agreement with the Likert scale (κ=0.35 for all patients, κ=0.39 for the subgroup) (Table 4). Finally, the Ojiri criteria did not perform well since the radiologists found it difficult to assess the presence of a focal defect or extracapsular spread since in the typically very small residual lymph nodes after CRT. With the consensus PET-CT Likert scale no readings were equivocal. Therefore, PET-CT Likert ≥ '4' in combination with DW-MRI ≥ '3' and PET-CT ≥ '3' in combination with DW-MRI ≥ '3' yielded the same results. These combinations between PET-CT and DW-MRI resulted in a sensitivity of 100% (CI 29%-100%), a specificity of 95% (CI 85%-99%), a PPV of 50% (CI 12%-88%) and a NPV of 100% (CI 93%-100%). PET-CT Likert ≥ '4' in combination with DW-MRI ≥ '4' and PET-CT ≥ '3' in combination with DW MRI ≥ '4' yielded a sensitivity of 66% (CI 9%-99%), a specificity of 95% (CI 85%-99%), a PPV of 40% (CI 5%-85%) and a NPV of 98% (CI 90%-99%).

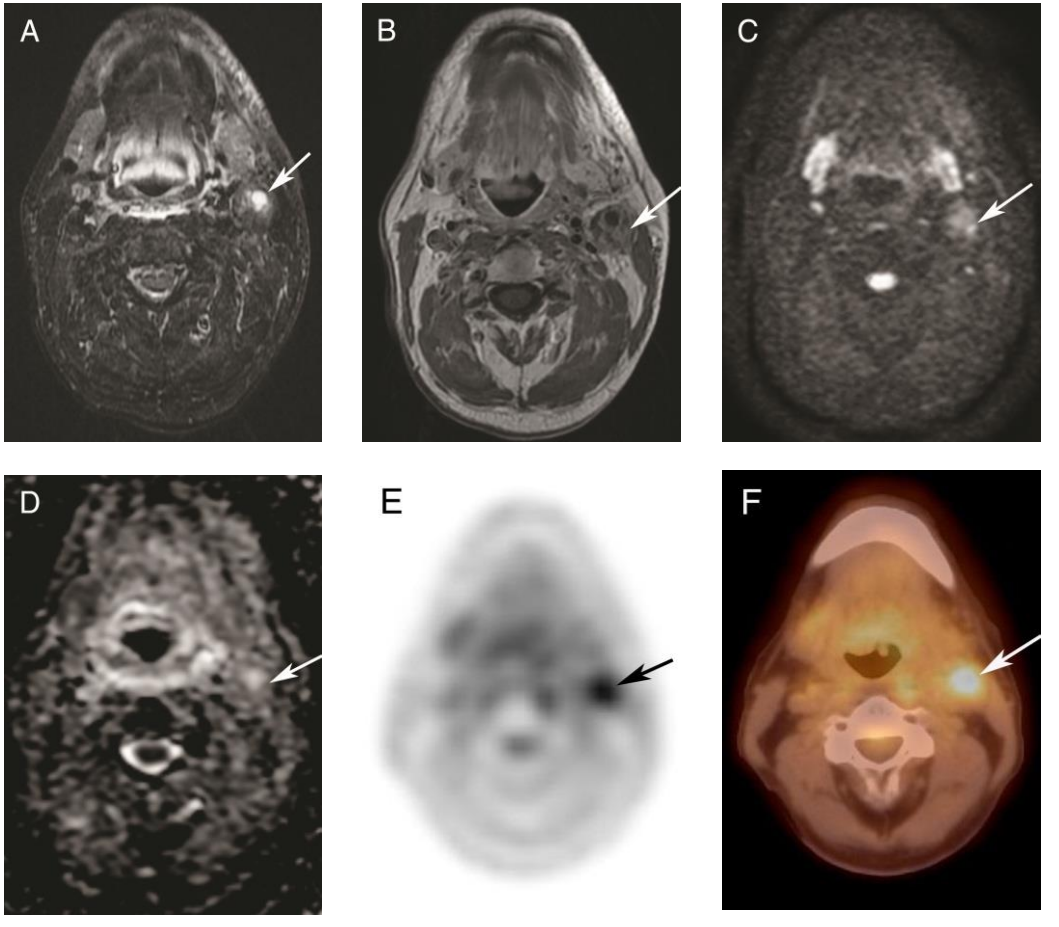


Figure 1.
True positive DW-MRI and ^{18}F-FDG-PET-CT in a patient with a T1N2a oropharyngeal carcinoma on the left side, 12 weeks after CRT. Residual viable tumor cells in the level II lymph node were histopathologically proven in the surgical specimen. Axial STIR MR image **(A)** showing a residual enlarged lymph node (arrow) in level II on the left side and contrast enhanced T1-weighted image **(B)** showing parts of contrast-enhancement in the lymph node. DWI (b1000) **(C)** showing persistent hyperintensity in the contrast-enhanced parts of the lymph node, corresponding to low ADC-values on the ADC-map **(D)**. Axial PET (CT-based attenuation correction (CTAC)) **(E)** and fused PET-CT images **(F)** showing a focus of intensely increased ^{18}F-FDG uptake in the lymph node.

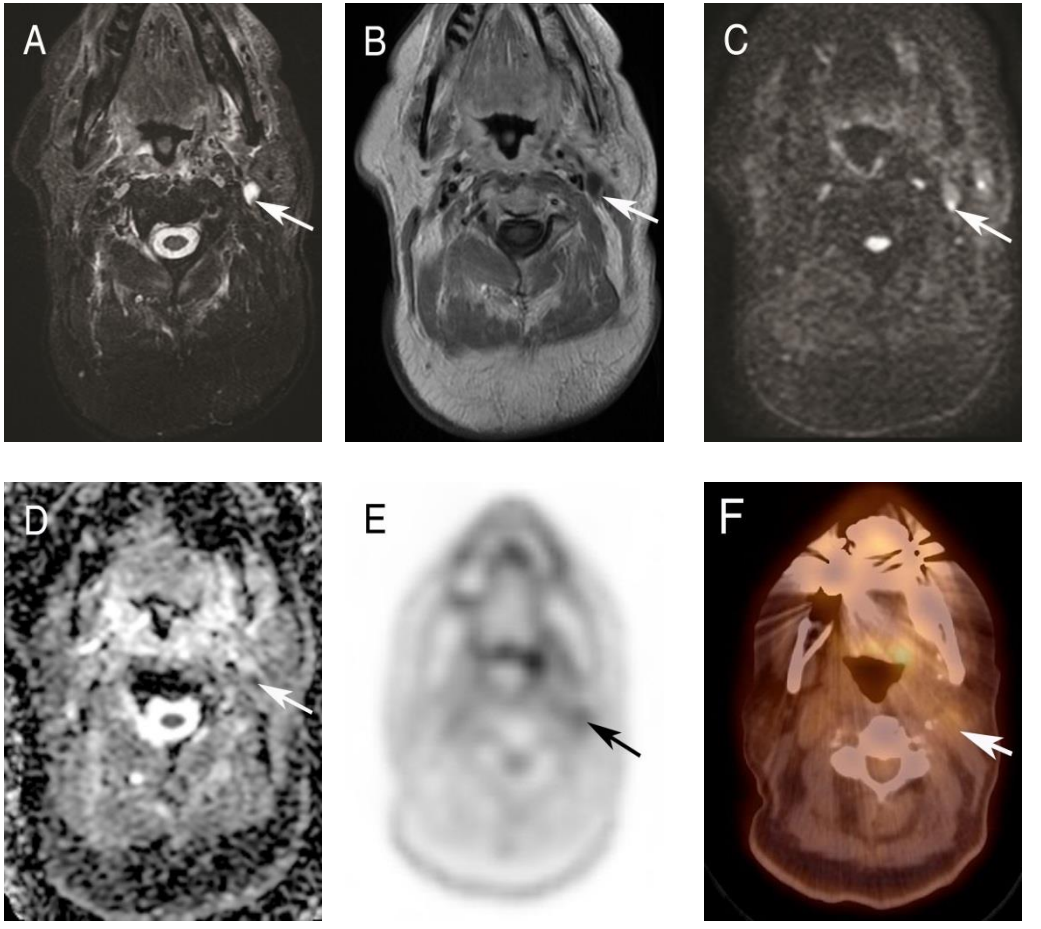


Figure 2.
False positive DW-MRI and ^{18}F-FDG-PET-CT in a patient with a T4bN2c oropharyngeal carcinoma on the left side, 12 weeks after CRT. Repeated tissue sampling with fine needle aspiration showed a reactive lymph node. During follow-up the lymph node further regressed in size. Axial STIR MR image **(A)** showing a residual lymph node (arrow) in level II on the left side and contrast enhanced T1-weighted image **(B)** showing parts of contrast-enhancement in the wall of the lymph node. DWI (b1000) **(C)** showing an area of persistent hyperintensity in the contrast-enhanced parts of the lymph node, corresponding to low ADC-values on the ADC-map **(D)**. Axial PET (CTAC) **(E)** and fused PET-CT images **(F)** showing slightly increased ^{18}F-FDG uptake in the lymph node.

Table 3. Diagnostic performance of PET-CT and DW-MRI as a function of test positivity criteria. The Likert scale was reduced to a 2-points scale as 'Likert conservative' (1-3=negative, 4-5=positive) and 'Likert sensitive' (1-2=negative, 3-5=positive)

Scales		***PET-CT*** *(n=58)*			***DW-MRI*** *(n=58)*			***DW-MRI*** *(n=73)*		
		Obs 1	*Obs 2*	*Cons*	*Obs 1*	*Obs 2*	*Cons*	*Obs 1*	*Obs 2*	*Cons*
Dichotomous	*Sens*	100 [29-100]	100 [29-100]	100 [29-100]	67 [9-99]	100 [29-100]	67 [9-99]	40 [5-85]	60 [15-95]	60 [15-95]
	Spec	91 [80-97]	84 [71-92]	84 [71-92]	86 [73-94]	95 [85-99]	93 [82-98]	85 [75-93]	94 [86-98]	93 [84-98]
	PPV	38 [9-76]	25 [5-57]	25 [6-57]	20 [3-56]	50 [12-88]	33 [4-78]	17 [2-48]	43 [10-82]	38 [9-76]
	NPV	100 [93-100]	100 [92-100]	100 [92-100]	98 [89-100]	100 [93-100]	98 [90-100]	95 [86-99]	97 [90-100]	97 [89-100]
Likert conservative	*Sens*	100 [29-100]	100 [29-100]	100 [29-100]	67 [9-99]	100 [29-100]	67 [9-99]	40 [5-85]	60 [15-95]	40 [5-85]
	Spec	91 [80-97]	84 [71-92]	84 [71-92]	87 [76-95]	95 [85-99]	95 [85-99]	88 [78-95]	94 [86-98]	94 [86-98]
	PPV	38 [9-76]	25 [5-57]	25 [5-57]	22 [3-60]	50 [12-88]	40 [5-85]	20 [3-56]	43 [10-82]	33 [4-78]
	NPV	100 [93-100]	100 [93-100]	100 [93-100]	98 [89-100]	100 [93-100]	98 [90-100]	95 [87-99]	97 [90-100]	96 [88-99]
Likert sensitive	*Sens*	100 [29-100]	100 [29-100]	100 [29-100]	67 [9-99]	100 [29-100]	100 [29-100]	80 [28-100]	60 [15-95]	100 [48-100]
	Spec	89 [78-96]	80 [67-90]	84 [71-92]	82 [69-91]	87 [76-95]	89 [78-96]	81 [70-89]	87 [76-94]	87 [76-94]
	PPV	33 [8-71]	21 [5-51]	25 [6-57]	17 [2-48]	30 [7-65]	33 [8-70]	24 [7-50]	25 [6-57]	36 [13-65]
	NPV	100 [93-100]	100 [92-100]	100 [92-100]	98 [89-100]	100 [93-100]	100 [93-100]	98 [90-100]	97 [89-100]	100 [94-100]

Values are presented as point estimates and [95% confidence interval]

Abbreviations: cons: consensus, obs: observer, NPV: negative predictive value, PPV: positive predictive value, sens: sensitivity, spec: specificity

Table 4. Interobserver agreement for the dichotomous system and the Likert conservative and Likert sensitive classifications

			Dichotomous			*Likert conservative*			*Likert sensitive*		
			Obs 2			Obs 2			Obs 2		
			+	-	*Kappa*	+	-	*Kappa*	+	-	*Kappa*
Subgroup											
PET-CT	**Obs 1**	+	7	1	0.64 [0.38-0.90]	7	1	0.69 [0.43-0.94]	7	2	0.55 [0.28-0.82]
(n=58)		-	5	45		4	46		6	43	
DW-MRI	**Obs 1**	+	5	5	0.57 [0.27-0.87]	4	5	0.47 [0.14-0.80]	7	5	0.55 [0.28-0.83]
(n=58)		-	1	47		2	47		3	43	
All patients											
DW-MRI	**Obs 1**	+	6	6	0.58 [0.31-0.85]	5	5	0.54 [0.24-0.84]	9	8	0.53 [0.29-0.77]
(n=73)		-	1	60		2	61		3	53	

Values are presented as the number of patients and as kappa with [95% confidence interval]
Abbreviations: obs=observer

DISCUSSION

Timely detected residual neck disease after CRT for HNSCC can often be successfully salvaged with a neck dissection. To minimise the number of futile (tumor-negative) neck dissections, reliable diagnostic techniques are warranted. However, the diagnostic management of the post-CRT neck with anatomical modalities like CT and MRI can be difficult [7,9,28]. We studied the accuracy and interobserver variability of DW-MRI and ^{18}F-FDG-PET-CT 3 months after CRT and found that both modalities had a reasonable accuracy with interobserver agreement varying as a function of test positivity criteria. PET-CT tended to have better sensitivity and DW-MRI the better specificity. A combined approach by adding DW-MRI (sensitively scored on a 5-point Likert scale) to PET-CT seemed to yield higher specificity without compromising sensitivity. Obviously, specificity is important if the aim is to reduce futile surgical explorations and associated morbidities and costs.

Observer variation was higher for DW-MRI than for PET-CT. This might be explained by a difference in the extensiveness of the assessment strategies. PET-CT observers have to account for less parameters than DW-MRI observers, which may contribute to less observer variation. Qualitative analysis of DW-MRI is done by visual assessment of the signal intensity on high b-values images and their corresponding ADC maps. Malignant tumor appears as an area of high signal intensity on the b1000 images with low signal intensity on the ADC map. Post-treatment changes demonstrate low signal intensity (necrosis) on b1000 images, but they sometimes manifest as high signal intensity areas (oedema). However, they generally show high signal intensity on the corresponding ADC maps [22,32]. Two quantitatively assessed studies in HNSCC showed that the ADC of residual squamous cell carcinoma is significantly lower than that of a benign post-treatment mass [23,25]. No previous studies were published in HNSCC using qualitative assessment and the level of interobserver agreement is unknown. Our results concerning the interobserver variation imply that future studies should address the criteria for test positivity in DW-MRI.

Interobserver agreement for PET-CT was more robust, but PET-CT results yielded a considerable amount of false positive results, at least in our series with a low prevalence of residual disease. In the current study, PET-CT was considered 'abnormal' if ^{18}F-FDG uptake was above background activity and projected in a lymph node on corresponding CT images. This resulted in a specificity of 84%. Recently, Marcus et al. [33] developed new qualitative therapy response criteria for head and neck PET-CT with substantial inter-reader agreement. In retrospect, in our data only PET-CT images showing focal areas with at least moderate ^{18}F-FDG uptake proved to be malignant. This is in line

with the Hopkins criteria and could have increased the specificity of PET-CT in our study to 93%. However, these new head and neck PET-CT interpretation criteria should be validated in large prospective studies.

In our institute, clinical palpation of the neck in combination with conventional MRI has been employed as diagnostic tool after CRT for HNSCC over the last decades. If one of the tests suggests the presence of a residual lymph node, we perform an ultrasound-guided FNAC. However, Van der Putten et al. [6] reported 18 false positive results after CRT in 46 patients examined with FNAC. In our present study, 4 patients had a false positive FNAC result so that in retrospect they had been exposed to unnecessary surgery. Therefore, reliable imaging modalities are required. In retrospect, in 3 out of these 4 patients the PET-CT was scored 'negative' for residual nodal disease and a futile neck dissection might have been avoided.

Several studies have shown the value of ^{18}F-FDG-PET-CT to reliably exclude residual or recurrent disease in the follow-up of HNSCC after CRT and to avoid futile salvage surgery [12-19]. However, studies differed in inclusion criteria, (leading to highly variable prevalences of persistent disease), treatments (CRT or radiotherapy alone), criteria of test positivity, time interval between end of treatment and PET-CT and the duration of follow-up (Table 1). Although most studies report high NPV in the assessment of the post-treatment neck, sensitivity and specificity can vary substantially (40-100% and 25-97%, respectively). Moreover, despite the seemingly high sensitivity of PET-CT, the use of CT as the complementary anatomical imaging modality has certain drawbacks with the relatively poor soft-tissue contrast being an import one post-treatment. MRI obtains excellent soft-tissue contrast and has evolved towards functional imaging capabilities, including DW-MRI. Even though the prevalence of disease in our series was low, our data suggest that accuracy may be improved by using a combined strategy with DW-MRI added to PET-CT. The optimal scoring strategy for considering a patient positive with the combined strategy is if both PET-CT and DW-MRI are equivocal, probably malignant or definitively malignant. *Sensitive* Likert scoring for DW-MRI resulted in best accuracy for the combination; thus an 'equivocal' assessment or a disagreement between radiologists should be interpreted as positive. Further studies with a joint integrated multidiagnostic modality reading with radiologists and nuclear medicine physicians would be informative and has yet to be investigated in this clinical setting.

We acknowledge several limitations to this retrospective study. First, the prevalence of residual disease was low (7%) and as a consequence the PPV will also likely be low, even using a test with high sensitivity and specificity. To increase the PPV of a test, a program could target the screening test to those at high risk of developing disease. We therefore only included

patients with advanced pretreatment nodal disease (N2-N3), nevertheless did we achieve good regional control in this patient cohort. Secondly, visual interpretation of imaging is a subjective task. To overcome subjective measurements and to reduce observer variation, quantification with Standardized Uptake Values (SUV) and ADC-values can be used. Due to different protocols and scanners we could not analyse these quantitative parameters. It has been shown that ADC values differed significantly between MR imaging systems and sequences [34]. However, in a study with 65 HNSCC-patients, SUV-based analysis did not improve the diagnostic accuracy of scan interpretation, when assessing the neck after CRT [15]. For DW-MRI, the measured observer variation suggests that additional research is required to arrive at robust test positivity criteria.

CONCLUSION

PET-CT recognised all regional residues and DW-MRI recognised most patients with regional control with substantial and moderate observer agreement, respectively. A combined approach with sensitively scored DW-MRI added to PET-CT seemed to increase the specificity of the diagnostic work-up without compromising sensitivity. If validated (for which PET-MRI scanners are optimally equipped), this could imply that fewer patients are exposed to futile neck dissections.

REFERENCES

1. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
2. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.
3. 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(6):567-578.
4. Lavertu P, Bonafede JP, Adelstein DJ, et al. Comparison of surgical complications after organ-preservation therapy in patients with stage III or IV squamous cell head and neck cancer. *Arch. Otolaryngol. Head Neck Surg.* 1998;124(4):401-406.
5. Newman JP, Terris DJ, Pinto HA, Fee WE, Jr., Goode RL, Goffinet DR. Surgical morbidity of neck dissection after chemoradiotherapy in advanced head and neck cancer. *Ann. Otol. Rhinol. Laryngol.* 1997;106(2):117-122.
6. Van der Putten L, Van den Broek GB, De Bree R, et al. Effectiveness of salvage selective and modified radical neck dissection for regional pathologic lymphadenopathy after chemoradiation. *Head Neck.* 2009;31(5):593-603.
7. Liauw SL, Mancuso AA, Amdur RJ, et al. Postradiotherapy neck dissection for lymph node-positive head and neck cancer: the use of computed tomography to manage the neck. *J. Clin. Oncol.* 2006;24(9):1421-1427.
8. Langerman A, Plein C, Vokes EE, et al. Neck response to chemoradiotherapy: complete radiographic response correlates with pathologic complete response in locoregionally advanced head and neck cancer. *Arch. Otolaryngol. Head Neck Surg.* 2009;135(11):1133-1136.
9. Nomayr A, Lell M, Sweeney R, Bautz W, Lukas P. MRI appearance of radiation-induced changes of normal cervical tissues. *Eur. Radiol.* 2001;11(9):1807-1817.
10. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. *Clin. Otolaryngol.* 2008;33(3):210-222.
11. De Bree R, Van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncol.* 2009;45(4-5):386-393.
12. Loo SW, Geropantas K, Beadsmoore C, Montgomery PQ, Martin WM, Roques TW. Neck dissection can be avoided after sequential chemoradiotherapy and negative post-treatment positron emission tomography-computed tomography in N2 head and neck squamous cell carcinoma. *Clin. Oncol. (R. Coll. Radiol.).* 2011;23(8):512-517.
13. Cho AH, Shah S, Ampil F, Bhartur S, Nathan CO. N2 disease in patients with head and neck squamous cell cancer treated with chemoradiotherapy: is there a role for posttreatment neck dissection? *Arch. Otolaryngol. Head Neck Surg.* 2009;135(11):1112-1118.
14. Rabalais AG, Walvekar R, Nuss D, et al. Positron emission tomography-computed tomography surveillance for the node-positive neck after chemoradiotherapy. *Laryngoscope.* 2009;119(6):1120-1124.

15. Ong SC, Schoder H, Lee NY, et al. Clinical utility of 18F-FDG PET/CT in assessing the neck after concurrent chemoradiotherapy for Locoregional advanced head and neck cancer. *J. Nucl. Med.* 2008;49(4):532-540.
16. Nayak JV, Walvekar RR, Andrade RS, et al. Deferring planned neck dissection following chemoradiation for stage IV head and neck cancer: the utility of PET-CT. *Laryngoscope.* 2007;117(12):2129-2134.
17. Chen AY, Vilaseca I, Hudgins PA, Schuster D, Halkar R. PET-CT vs contrast-enhanced CT: what is the role for each after chemoradiation for advanced oropharyngeal cancer? *Head Neck.* 2006;28(6):487-495.
18. Gourin CG, Williams HT, Seabolt WN, Herdman AV, Howington JW, Terris DJ. Utility of positron emission tomography-computed tomography in identification of residual nodal disease after chemoradiation for advanced head and neck cancer. *Laryngoscope.* 2006;116(5):705-710.
19. Moeller BJ, Rana V, Cannon BA, et al. Prospective risk-adjusted [18F]Fluorodeoxyglucose positron emission tomography and computed tomography assessment of radiation response in head and neck cancer. *J. Clin. Oncol.* 2009;27(15):2509-2515.
20. Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol.* 2003;45(3):169-184.
21. Ross BD, Moffat BA, Lawrence TS, et al. Evaluation of cancer therapy using diffusion magnetic resonance imaging. *Mol. Cancer Ther.* 2003;2(6):581-587.
22. Thoeny HC, De KF, King AD. Diffusion-weighted MR imaging in the head and neck. *Radiology.* 2012;263(1):19-32.
23. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
24. Vandecaveye V, De Keyzer F, Nuyts S, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):960-971.
25. Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1098-1107.
26. Becker M, Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *Br. J. Radiol.* 2014;87(1036):20130677.
27. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2010;37(1):181-200.
28. Ojiri H, Mendenhall WM, Stringer SP, Johnson PL, Mancuso AA. Post-RT CT results as a predictive model for the necessity of planned post-RT neck dissection in patients with cervical metastatic disease from squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2002;52(2):420-428.
29. Padhani AR, Koh DM. Diffusion MR imaging for monitoring of treatment response. *Magn Reson. Imaging Clin. N. Am.* 2011;19(1):181-209.
30. De Vet HC, Mokkink LB, Terwee CB, Hoekstra OS, Knol DL. Clinicians are right not to like Cohen's kappa. *BMJ.* 2013;346:f2125.
31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
32. Thoeny HC, Ross BD. Predicting and monitoring cancer treatment response with diffusion-weighted MRI. *J. Magn Reson. Imaging.* 2010;32(1):2-16.

33. Marcus C, Ciarallo A, Tahari AK, et al. Head and Neck PET/CT: Therapy Response Interpretation Criteria (Hopkins Criteria)-Interreader Reliability, Accuracy, and Survival Outcomes. *J. Nucl. Med.* 2014;55(9):1411-1416.
34. Kolff-Gart AS, Pouwels PJ, Noij DP, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. *AJNR Am. J. Neuroradiol.* 2015;36(2):384-390.

CHAPTER 8

COST-EFFECTIVENESS OF RESPONSE EVALUATION AFTER CHEMORADIOTHERAPY IN PATIENTS WITH ADVANCED OROPHARYNGEAL CANCER USING ^{18}F-FDG-PET-CT AND/OR DIFFUSION-WEIGHTED MRI

M.J.E. Greuter
C.S. Schouten
J.A. Castelijns
P. de Graaf
E.F.I. Comans
O.S. Hoekstra
R. de Bree
V.M.H. Coupé

BMC Cancer 2017; 17: 256-4

ABSTRACT

Background. Considerable variation exists in diagnostic tests used for local response evaluation after chemoradiotherapy in patients with advanced oropharyngeal cancer. The yield of invasive examination under general anesthesia (EUA) with biopsies in all patients is low and it may induce substantial morbidity. We explored four response evaluation strategies to detect local residual disease in terms of diagnostic accuracy and cost-effectiveness.

Methods. We built a decision-analytic model using trial data of forty-six patients and scientific literature. We estimated for four strategies the proportion of correct diagnoses, costs concerning diagnostic instruments and the proportion of unnecessary EUA indications. Besides a reference strategy, i.e. EUA for all patients, we considered three imaging strategies consisting of ^{18}Fluoro-2-deoxyglucose positron emission tomography - computed tomography (^{18}FDG-PET-CT), diffusion-weighted MRI (DW-MRI), or both ^{18}FDG-PET-CT and DW-MRI followed by EUA after a positive test. The impact of uncertainty was assessed in sensitivity analyses.

Results. The EUA strategy led to 96% correct diagnoses. Expected costs were €468 per patient whereas 89% of EUA indications were unnecessary. The DW-MRI strategy was the least costly strategy, but also led to the lowest proportion of correct diagnoses, i.e. 93%. The PET-CT strategy and combined imaging strategy were dominated by the EUA strategy due to respectively a smaller or equal proportion of correct diagnoses, at higher costs. However, the combination of PET-CT and DW-MRI had the highest sensitivity. All imaging strategies considerably reduced (unnecessary) EUA indications and its associated burden compared to the EUA strategy.

Conclusion. Because the combined PET-CT and DW-MRI strategy costs only an additional €927 per patient, it is preferred over immediate EUA since it reaches the same diagnostic accuracy in detecting local residual disease while leading to substantially less unnecessary EUA indications. However, if healthcare resources are limited, DW-MRI is the strategy of choice because of lower costs while still providing a large reduction in unnecessary EUA indications.

INTRODUCTION

Patients with resectable advanced staged oropharyngeal squamous cell carcinoma (OPSCC) are often treated with chemoradiotherapy (CRT) in order to preserve organ function and quality of life. Low residual and recurrent tumour rates indicate that CRT is an adequate treatment option [1]. Still, thorough follow-up is warranted to detect residual tumour, which can be successfully treated with salvage surgery if detected early. Previous research has shown that early detection of residual tumour is associated with more favourable survival probabilities and better local control [1,2]. Thus, timely detection of residual tumour is essential.

A Dutch survey showed that there is considerable variation in response evaluation after CRT, especially in the diagnostic tests performed [3]. Tests that are used are examination under general anesthesia with taking of biopsies (EUA), computed tomography (CT), magnetic resonance imaging (MRI) and ^{18}F-fluorodeoxy-glucose positron emission tomography combined with CT (^{18}F-FDG-PET-CT). EUA is considered to be the most reliable procedure to detect residual disease but is an invasive procedure during which biopsies are taken in areas treated with CRT. Besides the risk of side-effects such as pain, inflammation and wound healing problems, there is also the possibility of sampling error leading to false-negative results [1]. Furthermore, EUA has considerable impact on scarce resources because hospital stay and operating facilities are required.

In contrast to EUA, imaging tests are not invasive. However, conventional CT and MRI may not be suitable because postradiation effects, e.g. fibrosis and necrosis [4], may hamper accurate interpretation of the images. More advanced tests such as ^{18}F-FDG-PET-CT and diffusion-weighted MRI (DW-MRI) may be more suitable options. These tests do not only assess if any anatomical residual mass is present, but also determine the metabolic activity and cell density, respectively, of the tumour.

PET-CT and DW-MRI cannot completely replace EUA because pathological confirmation is required before further treatment. Nevertheless, they could be used to select those patients with a high risk of residual disease for further diagnostic workup with EUA. This would reduce the number of patients that have to undergo futile invasive diagnostic procedures. Furthermore, the imaging results provide the opportunity to guide biopsy procedures, thereby possibly reducing sampling error. Besides these advantages for patients, imaging to select patients for EUA might also lead to cost reductions. Therefore, this study aimed to assess the effects and costs of four response evaluation strategies to detect local residual disease, namely EUA for all patients, PET-CT-based selection for EUA, DW-MRI-based

selection for EUA and a combination of PET-CT and DW-MRI to select for EUA. All analyses were conducted using a decision-analytic model based on trial data of forty-six patients and scientific literature.

METHODS

Strategies

We developed a decision-analytic model to evaluate the effects and costs of four strategies for local response evaluation in patients with OPSCC who underwent CRT. In the first strategy under consideration, i.e. the reference strategy, all patients undergo EUA three months after completion of CRT for response evaluation. EUA is considered positive if histopathological proof of viable cancer has been found in the biopsy. The indication for EUA is considered unnecessary if six months of follow-up after response evaluation did not show residual disease.

In the three other strategies, response evaluation is performed by either PET-CT (strategy 2), DW-MRI (strategy 3) or a combination of PET-CT and DW-MRI (strategy 4). Only patients with a positive imaging test are referred to EUA. In strategy 4, patients are referred to EUA based on a combined assessment by a radiologist and nuclear medicine physician. The four strategies are depicted by decision trees in Figure 1. Each end node reflects the diagnostic status of a patient who followed that specific branch as well as the associated costs.

Data

Costs of the different diagnostic instruments were based on official tariff lists [5]. As shown in Table 1, combined PET-CT and DW-MRI was with 1313 Euros the most expensive test. Test characteristics were obtained from a prospective trial (2012-2014) in which patients with advanced but resectable OPSCC who underwent CRT were subjected to PET-CT, DW-MRI and EUA for response evaluation three months after end of CRT (REACTION-study). Patient and tumour characteristics are summarized in Table 1 from chapter 6. PET-CT and DW-MRI images were interpreted by respectively two nuclear medicine physicians (O. Hoekstra, E. Comans) and two radiologists (J. Castelijns, P. de Graaf). For PET-CT, we requested the observers individually to score the PET-CT using the Hopkins criteria (a 5-point scale in which scores 1, 2 and 3 are considered negative for tumour and scores 4 and 5 are considered positive) [6].

DW-MRI observers scored signal intensity on b1000 images and the corresponding ADC-map. A residual mass was considered 'suspicious for residual disease' if increased signal intensity on the b1000 images was present, in an area of contrast-enhancement on post-contrast T1WI and corresponding to low ADC-values. For DW-MRI, we requested the observers individually to classify the MRI-examinations using a dichotomous system, i.e. suspicious for a local residue or not, based on signal intensity of the DW-images. In case readings were discrepant, consensus was reached between the observers of either PET-CT or DW-MRI. For strategy 4, combined reading of PET-CT and DW-MRI with visual correlation by a radiologist and nuclear medicine physician was performed in all patients with discrepancies between the scoring of the two modalities. In total, 46 patients were included of whom five (11%) had residual tumour. Test characteristics of each diagnostic test are shown in Table 1. Sensitivities were 60% or higher, with the combined PET-CT and DW-MRI strategy having the highest sensitivity. Specificity was lowest for PET-CT with 83%.

Analyses

Using the decision trees, we calculated the expected health effects of each response evaluation strategy, i.e. the expected proportion of correctly diagnosed patients (true positives and true negatives). Based on the costs of each diagnostic instrument, the average costs accumulated by a patient when following a specific branch in the decision tree were estimated. Subsequently, for each branch costs were multiplied by the probability that a patient will follow this branch. Summing up the expected costs of each branch of the decision tree results in the overall expected costs for that strategy. In addition, the number of EUA indications as well as the number of unnecessary EUA indications per response evaluation strategy were calculated. It was not possible to correctly estimate quality-adjusted life-years in this study. As an alternative we estimated the costs per true-positive case for the different strategies.

Sensitivity analyses

Due to the small sample size of this trial, we repeated the base-case analysis using test characteristics based on the literature. Sensitivity and specificity for PET-CT and DW-MRI were derived from respectively a systematic review by Gupta et al. (2011) [7] and the study of Vaid et al. (2017) [8]. Test characteristics for EUA and combined PET-CT and DW-MRI are not reported in the literature.

We further explored the impact of uncertainty regarding the test characteristics of the diagnostic instruments in univariate sensitivity analyses in which we increased and decreased the sensitivity and specificity of each instrument with 10% (absolute change). In the strategies in which a positive imaging test was followed by EUA, only the test characteristics of the imaging test were varied.

In addition, in clinical practice a positive EUA is usually followed by MRI and PET-CT to assess if surgical intervention is feasible and if any distant metastases are present, respectively. Therefore, in a sensitivity analysis we included the costs of these additional tests in the EUA strategy. Similarly, we added costs of an MRI or whole body PET-CT (in case of a positive EUA) to those of the "head and neck" PET-CT and DW-MRI, respectively. Finally, we assessed the impact of uncertainty surrounding the prevalence of residual tumour by increasing and decreasing the prevalence to 15% and 5%, respectively [9].

Figure 1.
Decision trees of all evaluated strategies namely **(A)** EUA for all patients, **(B)** PET-CT prior to EUA, **(C)** DW-MRI prior to EUA and **(D)** PET-CT as well as DW-MRI prior to EUA. The end node depicts the diagnostic status of the patient, i.e. true positive (TP), false negative (FN), false positive (FP) or true negative (TN), as well as the costs associated with that specific branch.

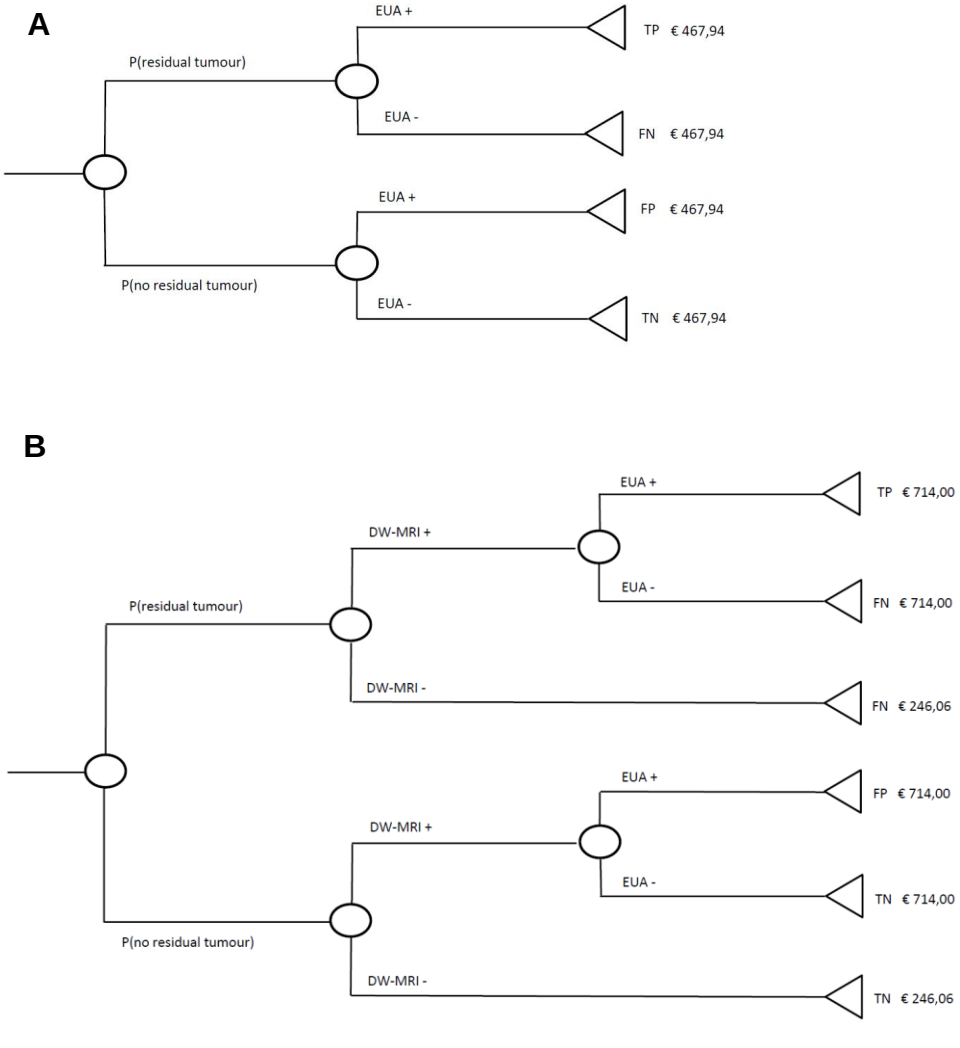
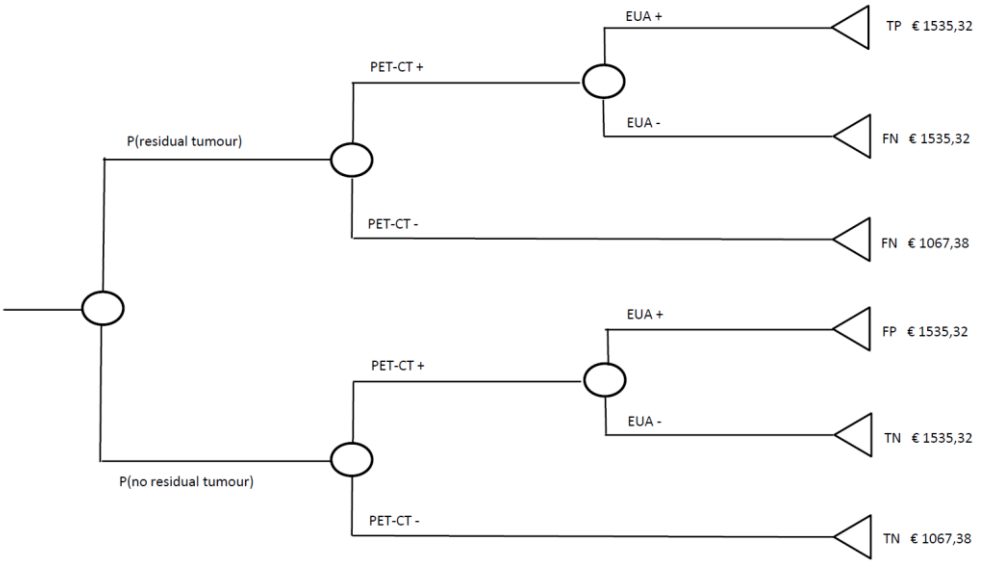


Figure continues on the next page

C


EUA +
TP € 1535,32
PET-CT +
EUA -
FN € 1535,32
P(residual tumour)
PET-CT -
FN € 1067,38
EUA +
FP € 1535,32
PET-CT +
EUA -
TN € 1535,32
P(no residual tumour)
PET-CT -
TN € 1067,38

D

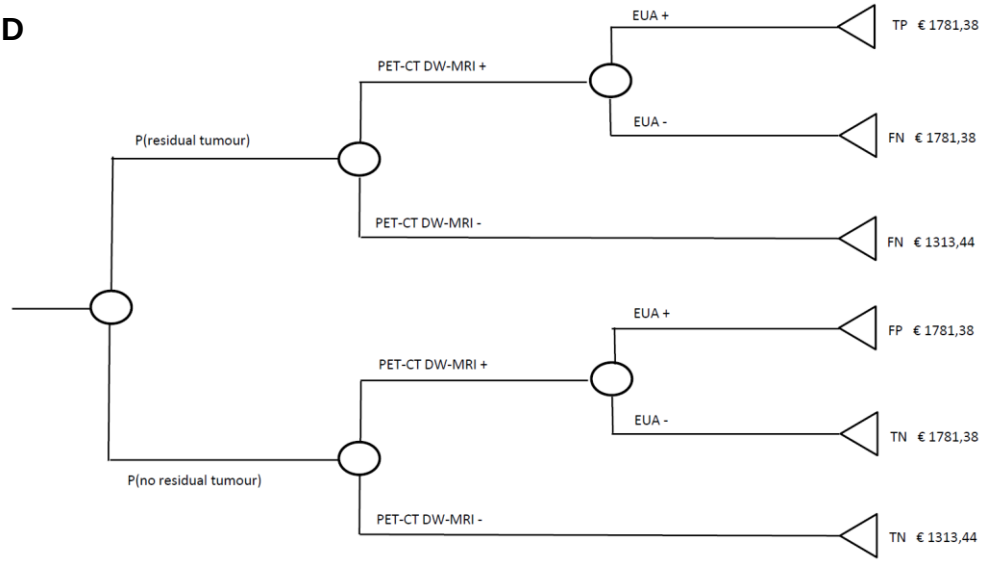

EUA +
TP € 1781,38
PET-CT DW-MRI +
EUA -
FN € 1781,38
P(residual tumour)
PET-CT DW-MRI -
FN € 1313,44
EUA +
FP € 1781,38
PET-CT DW-MRI +
EUA -
TN € 1781,38
P(no residual tumour)
PET-CT DW-MRI -
TN € 1313,44

Table 1. Test characteristics and costs of the diagnostic instruments for response evaluation. In the base-case analysis, trial-reported test characteristics were used.

Diagnostic instrument for response evaluation	*Trial-reported test characteristics*				*Test characteristic based on literature*		*Costs (Euros)*
	Sensitivity [95% CI]	**Specificity** [95% CI]	**Positive predictive value** [95% CI]	**Negative predictive value** [95% CI]	**Sensitivity**	**Specificity**	
EUA	60 [14.7-94.7]	100 [91.4-100.0]	100.0 [29.2-100.0]	95.3 [84.2-99.4]	*	*	468
PET-CT	75 [19.4-99.4]	83 [67.9-92.8]	30.0 [6.7-65.2]	97.1 [85.1-99.9]	79.9[a]	87.5[a]	1067
DW-MRI	60 [14.7-94.7]	95 [83.5-99.4]	60.0 [14.7-94.7]	95.1 [83.5-99.4]	89.0[b]	86.0[b]	246
Combined PET-CT and DW-MRI	100 [39.8-100.0]	93 [80.1-98.5]	57.1 [18.4-90.1]	100.0 [90.7-100.0]	*	*	1313

* Test characteristics for this diagnostic instrument are not reported in the literature.
[a] Based on the systematic review of Gupta et al. (2011) [7]
[b] Based on the study of Vaid et al. (2017) [8]

RESULTS

Base-case analysis

The proportion of correctly diagnosed patients and the expected costs per patient for each strategy are shown in Table 2. The combined PET-CT and DW-MRI strategy and the EUA strategy led to the highest proportion of correctly diagnosed patients, i.e. 96%. However, the expected costs of the EUA strategy were 927 Euros per patient lower than those of the combined PET-CT and DW-MRI strategy. The strategy with DW-MRI led to the lowest expected costs per patient but also to the lowest proportion of correct diagnoses. The PET-CT strategy and the combined PET-CT and DW-MRI strategy were dominated by the EUA strategy because these strategies led to a smaller or equal proportion of correct diagnoses, at higher costs.

Number of EUA indications and missed residual disease

Figure 2 shows the proportion of correct diagnoses as well as the proportion of unnecessary EUA indications per strategy. All imaging strategies led to a considerably lower proportion of unnecessary EUA indications compared to the EUA strategy. The combined PET-CT and DW-MRI strategy led to the lowest proportion of unnecessary EUA indications, i.e. 38%.

In Figure 3, the total number of EUA indications, the number of unnecessary EUA indications and the expected costs per patient are shown for 1,000 patients. In the EUA strategy, all patients underwent EUA, of which 891 indications were unnecessary. All imaging strategies substantially reduced the number of required EUAs as well as the number of unnecessary EUA indications. The DW-MRI required the lowest number of EUAs, i.e. 109 EUAs of which 43 indications were unnecessary. However, this strategy also had the highest number of missed residual disease. Please note that the combined PET-CT and DW-MRI required 65 additional EUAs compared to the DW-MRI strategy, but a slightly lower proportion of these EUA indications was unnecessary.

The costs per true-positive case were for EUA 7175 Euros, for PET-CT 24,058 Euros, for DW-MRI 7588 Euros and for combined PET-CT and DW-MRI 21,387 Euros. Although EUA and DW-MRI are much cheaper, they detect less recurrences than PET-CT leading to delayed salvage treatment with potentially poorer outcome. The addition of DW-MRI to PET-CT lowered the costs per true-positive case.

Table 2. Expected proportion of correctly classified patients, expected costs per patient and proportion of unnecessary EUA indications per strategy.

Strategy	*Expected proportion of correctly classified patients*	*Expected costs*	*Costs per true-positive case*	*Proportion of unnecessary EUA indications*
Base-case analysis				
EUA	0.96	468	7175	0.89
PET-CT	0.94	1177	24,058	0.65
DW-MRI	0.93	297	7588	0.40
PET-CT and DW-MRI	0.96	1395	21,387	0.38
Test characteristics based on the literature				
PET-CT	0.94	1160	22,264	0.56
DW-MRI	0.95	350	6025	0.56
Sensitivity +10%				
EUA	0.97	468	6150	0.89
PET-CT	0.96	1207	18,352	0.60
DW-MRI	0.94	302	6615	0.36
*PET-CT and DW-MRI**	0.96	1395	21,387	0.38
Specificity +10%				
EUA	0.96	468	7175	0.89
PET-CT	0.94	1135	23,205	0.44
DW-MRI	0.93	277	7068	0
PET-CT and DW-MRI	0.96	1364	20,919	0
Sensitivity -10%				
EUA	0.95	468	8610	0.89
PET-CT	0.93	1172	27,639	0.68
DW-MRI	0.92	292	8950	0.44
PET-CT and DW-MRI	0.95	1390	23,677	0.40
Specificity -10%				
EUA	0.87	468	7175	0.89
PET-CT	0.94	1218	24,910	0.75
DW-MRI	0.93	339	8654	0.67
PET-CT and DW-MRI	0.96	1437	22,027	0.59
EUA followed by additional tests				
EUA	0.96	574	8799	0.89
PET-CT	0.94	1189	24,304	0.65
DW-MRI	0.93	351	8966	0.40

Table continues on the next page

Increasing the prevalence of residual disease				
EUA	0.94	9834	5199	0.85
PET-CT	0.92	10,554	17,599	0.56
DW-MRI	0.90	9673	5696	0.32
PET-CT and DW-MRI	0.94	10,778	15,697	0.29
Decreasing the prevalence of residual disease				
EUA	0.98	7290	15,598	0.95
PET-CT	0.97	7982	51,592	0.81
DW-MRI	0.97	7103	15,655	0.61
PET-CT and DW-MRI	0.98	8191	45,645	0.58

*Sensitivity was 100% in base-case analysis

Univariate sensitivity analyses

When the test characteristics of DW-MRI were based on the literature, the proportion of correct diagnoses increased from 0.93 to 0.95. On the other hand, also the expected costs and proportion of unnecessary EUA indications increased. For the PET-CT strategy, the proportion of correct diagnoses remained 0.95 whereas the expected costs and proportion of unnecessary EUA indications decreased when we based the test characteristics on the literature.

In the majority of sensitivity analyses in which we varied test characteristics, additional testing after a positive EUA and the prevalence of residual disease, the ordering of the strategies based on the proportion of correct diagnoses, expected costs and proportion of unnecessary EUA indications did not change, as shown in Table 2. Only in the analysis in which the specificity of the diagnostic instruments was lowered with 10%, the ordering of the strategies changed; the EUA strategy led to the lowest proportion of correct diagnoses.

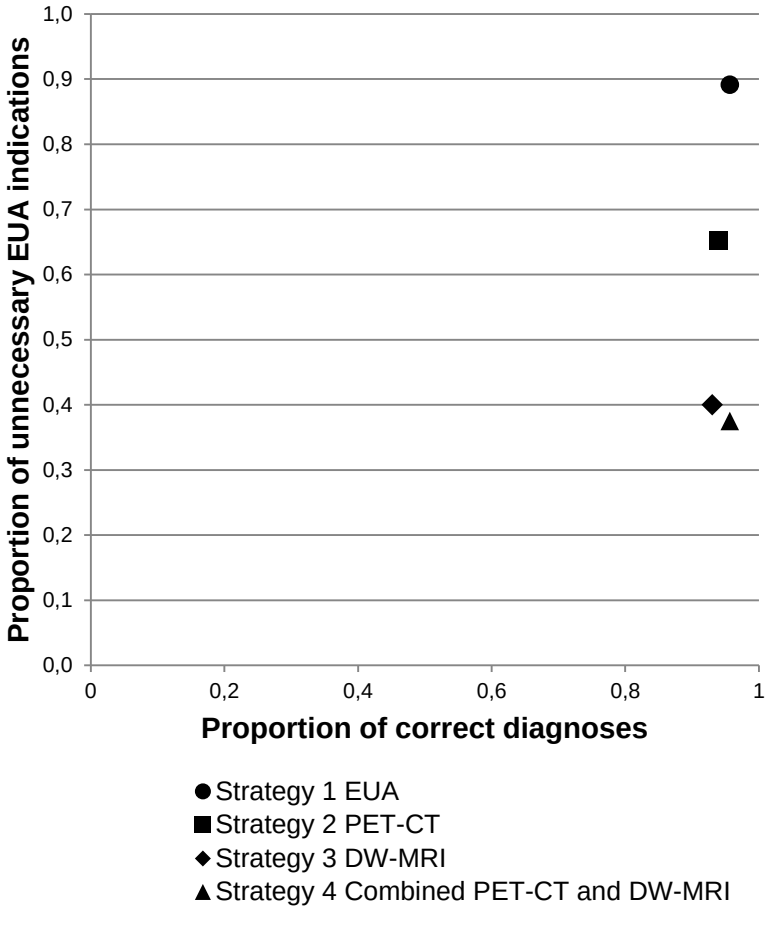


Figure 2.
Proportion of correct diagnoses and unnecessary EUA indications per strategy.

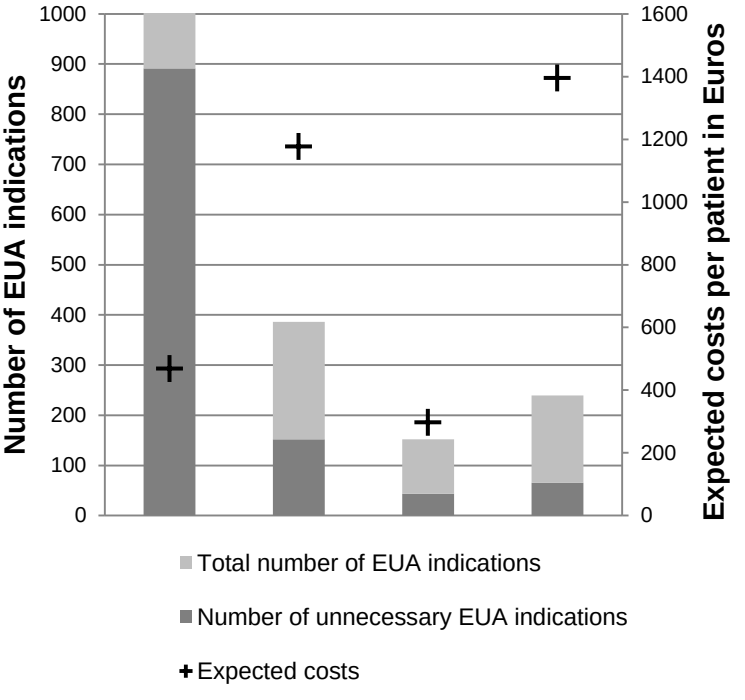


Figure 3.
Total number of EUA indications, number of unnecessary EUA indications and expected costs per patient per strategy for 1,000 patients.

DISCUSSION

This study explored the cost-effectiveness of four strategies used for response evaluation to detect local residual disease after CRT in patients with advanced staged OPSCC. In the EUA strategy, i.e. the reference strategy, 96% of patients were correctly diagnosed. Expected costs were 468 Euros at the expense of 89% unnecessary EUA indications. The DW-MRI strategy was with 297 Euros the least costly strategy. However, this strategy also led to the lowest proportion of correct diagnoses, i.e. 93%. The PET-CT strategy and the combined PET-CT and DW-MRI strategy were dominated by the EUA strategy due to a smaller or equal proportion of correct diagnoses, at higher costs. All imaging strategies considerably reduced the number of EUA indications and unnecessary EUA indications compared to the EUA strategy.

Based on our model results, the combined PET-CT and DW-MRI strategy is preferred over the EUA strategy. This strategy has the same diagnostic accuracy as immediate EUA while considerably reducing the number of EUA indications as well as unnecessary EUA indications. On the other hand, the combined imaging strategy costs an additional 927 Euros. However, only around 220 patients are diagnosed each year with advanced OPSCC in the Netherlands [10]. Thus, these additional costs are negligible on a society level. Furthermore, if all 220 patients were evaluated based on the combined imaging strategy, only 38 EUAs are indicated of which 14 would be unnecessary. When the EUA strategy would be used for response evaluation, 220 EUAs are indicated of which 196 would be unnecessary.

When healthcare resources are limited, our study suggests that DW-MRI prior to EUA is the strategy of choice. This strategy is less costly than both the combined PET-CT and DW-MRI strategy and EUA strategy. Furthermore, it has a lower impact on scarce resources such as hospital stay compared to immediate EUA. Moreover, fewer patients are exposed to invasive EUA and of the EUA indications, a lower proportion is unnecessary. However, the proportion of correct diagnoses is lower than in the EUA strategy. This difference is caused by a higher proportion of false-negative test results meaning that residual tumour is missed.

Besides decreasing the number of unnecessary EUA indications, another possible advantage of imaging is that the results of the imaging test can be used to guide the biopsy taking. This may increase the sensitivity of EUA. As the surgeons in the trial were not blinded to the imaging results, it may be possible that the sensitivity of EUA without prior imaging is overestimated. We assessed the impact of this in sensitivity analyses by assuming a 10% lower sensitivity in the EUA strategy. This changed the ordering of the strategies; the combined PET-CT and DW-MRI strategy became the strategy with most correct diagnoses. Nevertheless, EUA detected only 60% of the residual tumours, questioning the degree of bias.

The test characteristics of PET-CT and DW-MRI were based on a trial in which all patients received PET-CT, DW-MRI and EUA three months after CRT. Timing is an important determinant of these test characteristics. If the test is conducted too soon after treatment, postradiation effects can lead to a false-positive test. Also false-negative test results are possible because tumour cells have not yet reached a detectable size. On the other hand, early diagnosis of residual disease increases treatment success of salvage surgery [11]. A study indicated that DW-MRI may be used for response evaluation three weeks after CRT [12]. The optimal timing of imaging has still to be determined, which might influence test characteristics.

Furthermore, DW-MRI is a relative new imaging technique within oncological applications and therefore, little experience is gained with DW-MRI in the post-treatment evaluation so far. Besides, DW-imaging is susceptible to artefacts, particularly in the inhomogeneous head and neck area which contains a variety of tissues. Also, geometrical distortions due to interfaces between soft tissue and air or bone can occur. Although DW-MRI is not yet an established technique for response evaluation, this early assessment of the expected health effects and costs provides more insight in the potential of DW-MRI as a response evaluation strategy. We showed that DW-MRI is the strategy of choice when combined with PET-CT. However, a response evaluation strategy solely based on DW-MRI followed by EUA in individuals with a positive imaging test is only preferred in settings with limited health care resources. Nevertheless, radiologists are still in the learning curve concerning post-treatment evaluation of DW-MRI. With more training and feedback, DW-MRI is expected to obtain a higher accuracy. The impact of increased sensitivity was assessed in sensitivity analyses, but conclusions did not change.

Another emerging technique for response evaluation is PET-MRI. This technique combines the often complementary data from PET and MRI and could lead to improved anatomic localisation of focal uptake compared to PET-CT [13]. This could potentially decrease the number of false-positive test outcomes and as a consequence, reduce the number of unnecessary EUA indications.

The trial included only 46 patients of whom five were diagnosed with residual disease. Due to this small sample size, there was a fair amount of uncertainty regarding test characteristics and the prevalence of residual disease. Repeating the base-case analysis using test characteristics derived from the literature did not change our conclusion. Furthermore, our residual primary tumour rate of 11% was in agreement with the results of Moeller et al. (2009) [14]. On the other hand, Van den Broek et al. (2006) reported a residual primary tumour rate of 7% [1]. We have addressed this issue by varying the prevalence of residual disease in one-way sensitivity analyses. However, the ordering of the strategies based on correct diagnoses, costs and unnecessary EUA indications did not change.

In this cohort of patients, the residual primary tumour rate was only 11%. To improve the yield of routine response evaluation, only patients with high risk factors should undergo this diagnostic procedure. Risk factors which can be used to select patients include T-stage [15], HPV (human papilloma virus) status [16,17] and pre-treatment metabolic tumour volume [18].

We did not differentiate between HPV-related and HPV-unrelated tumours. For HPV-related tumours, a lower residual disease rate is observed which is probably due to higher responsiveness to chemoradiotherapy [19,20]. However, it is unclear whether HPV-related tumours have the same probability of being detected as HPV-unrelated tumours. In this study, 44% of the tumours were HPV positive whereas all the patients with residual disease had a HPV negative tumour. Nevertheless, the small sample size precludes definite conclusions regarding differences in detection.

Outcomes of this study were the proportion of correctly diagnosed patients, costs concerning diagnostic instruments and the number of unnecessary EUA indications. This means that health benefits, i.e. the proportion of correct diagnoses, and treatment burden, i.e. unnecessary EUA indications, were evaluated separately. Moreover, not all health benefits and treatment burden were captured in these outcomes. For example, declined survival probabilities due to false-negative test results as well as side-effects of EUA were not taken into account. Outcomes that encompass all health benefits and treatment burden such as quality-adjusted life-years (QALYs) would therefore be preferable. Comparing costs per QALY would lead to a more comprehensive evaluation of the response evaluation strategies. However, it was not possible to calculate QALYs due to the trial design. Patients in the trial were subjected to all tests for response evaluation meaning that residual disease which may be missed by one test (false-negative) could be detected by one of the other tests. If patients would be subjected to only one test, as in the strategies evaluated in this study, patients with false-negative test results would have become symptomatically detected at a later point of time. Since earlier detection leads to improved survival probabilities [1,2], it was not possible to correctly estimate QALYs. As an alternative, we calculated costs per true-positive case for the different strategies.

We hypothesize that the ordering of strategies would not change when the evaluation was based on costs per QALY. A strategy using an instrument with a high sensitivity and few (unnecessary) EUA indications would be favoured because this strategy would lead to low rates of missed residual disease and low treatment burden. This means that the combined PET-CT and DW-MRI strategy would still be the strategy of choice. To assess this hypothesis, future studies should include costs per QALY.

To our knowledge, few studies have assessed the cost-effectiveness of response evaluation in patients with head and neck cancer treated with CRT. Two studies compared a strategy of PET-CT scanning prior to neck dissection and up-front neck dissection for all patients. These studies showed that the imaging strategy was more cost-effective [9,21]. Although these studies did not compare imaging to EUA, they also indicate that imaging can be more cost-

effective than more invasive strategies.

In clinical practice, PET-CT can be used for response evaluation at the primary site and neck and to detect distant metastases simultaneously. Previous studies have shown that PET-CT is cost-effective for response evaluation after (chemo)radiation of the pretreatment advanced stage positive neck when only patients with a PET-CT positive neck underwent neck dissection compared to planned neck dissection in all patients [21,22]. Moreover, pretreatment screening for distant metastases using PET-CT appeared also to be cost-effective [23]. These evaluations by PET-CT add to the cost-effectiveness of response evaluation of the primary site as reported in the present study. For DW-MRI no data on cost-effectiveness in response evaluation of neck disease is available. Also for response evaluation of advanced nodal neck disease the combination of PET-CT and DW-MRI seems to have the highest sensitivity and specificity [24].

Results of cost-effectiveness studies can be used for the development of a guideline for response evaluation in patients with OPSCC. The need for such a guideline is underlined by a previous study showing that there is substantial variation in the diagnostic tests used for response evaluation [3]. By including studies on cost-effectiveness in the guideline development process, guideline recommendations will not only be based on the most effective strategy, but on the most cost-effective strategy. This will lead to more sensible use of scarce healthcare resources. This is the first cost-effectiveness study evaluating different response evaluation strategies. However, there was considerable uncertainty regarding important model parameters due to the small sample size of the trial. Additional studies, preferably based on trials with a larger sample size, are required to provide a comprehensive evidence base for guideline development.

CONCLUSION

This study suggests that combined PET-CT and DW-MRI is the strategy of choice for local response evaluation. It is preferred over immediate EUA since it costs only an additional €927 while reaching the same diagnostic accuracy and leading to substantially less unnecessary EUA indications and its associated burden. However, if healthcare resources are limited, DW-MRI is the preferred strategy because of lower costs while maintaining the number of unnecessary EUA indications low.

REFERENCES

1. Van den Broek GB, Rasch CR, Pameijer FA, Peter E, Van den Brekel MW, Balm AJ. Response measurement after intraarterial chemoradiation in advanced head and neck carcinoma: magnetic resonance imaging and evaluation under general anesthesia? *Cancer.* 2006;106(8):1722-1729.
2. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
3. Schouten CS, Hoekstra OS, Leemans CR, Castelijns JA, De Bree R. Response evaluation after chemoradiotherapy for advanced staged oropharyngeal squamous cell carcinoma: a nationwide survey in the Netherlands. *Eur Arch Otorhinolaryngol.* 2015;272(11):3507-3513.
4. Pandya JA, Srikant N, Boaz K, Manaktala N, Kapila SN, Yinti SR. Post-radiation changes in oral tissues - An analysis of cancer irradiation cases. *South Asian J Cancer.* 2014;3(3):159-162.
5. Nederlandse zorgautoriteit, available at: www.nza.nl/regelgeving/tarieven.
6. Marcus C, Ciarallo A, Tahari AK, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins Criteria)-interreader reliability, accuracy, and survival outcomes. *J. Nucl. Med.* 2014;55(9):1411-1416.
7. Gupta T, Master Z, Kannan S, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur. J. Nucl. Med. Mol. Imaging.* 2011;38(11):2083-2095.
8. Vaid S, Chandorkar A, Atre A, Shah D, Vaid N. Differentiating recurrent tumours from post-treatment changes in head and neck cancers: does diffusion-weighted MRI solve the eternal dilemma? *Clin Radiol.* 2017;72(1):74-83.
9. Sher DJ, Tishler RB, Annino D, Punglia RS. Cost-effectiveness of CT and PET-CT for determining the need for adjuvant neck dissection in locally advanced head and neck cancer. *Ann Oncol.* 2010;21(5):1072-1077.
10. Dutch Head and Neck Oncology Cooperative Group, Nederlandse Werkgroep Hoofd-Hals Tumoren (NWHHT). Oral cavity and oropharyngeal cancer. National guideline, version 1.4, 2004. Available at: Available at: http://www.nvpc.nl/uploads/stand/3702_02_04_Hoofd.pdf.
11. De Bree R, van der Putten L, Brouwer J, Castelijns JA, Hoekstra OS, Leemans CR. Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging. *Oral Oncol.* 2009;45(4-5):386-393.
12. Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1098-1107.
13. Becker M, Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *Br. J. Radiol.* 2014;87(1036):20130677.
14. Moeller BJ, Rana V, Cannon BA, et al. Prospective risk-adjusted [18F]Fluorodeoxyglucose positron emission tomography and computed tomography assessment of radiation response in head and neck cancer. *J. Clin. Oncol.* 2009;27(15):2509-2515.
15. Been MJ, Watkins J, Manz RM, et al. Tumor volume as a prognostic factor in oropharyngeal squamous cell carcinoma treated with primary radiotherapy. *Laryngoscope.* 2008;118(8):1377-1382.

16. Bledsoe TJ, Noble AR, Hunter GK, et al. Oropharyngeal squamous cell carcinoma with known human papillomavirus status treated with definitive chemoradiotherapy: patterns of failure and toxicity outcomes. *Radiat Oncol.* 2013;8:174.
17. Guo T, Rettig E, Fakhry C. Understanding the impact of survival and human papillomavirus tumor status on timing of recurrence in oropharyngeal squamous cell carcinoma. *Oral Oncol.* 2016;52:97-103.
18. Sager S, Asa S, Yilmaz M, et al. Prognostic significance and predictive performance of volume-based parameters of F-18 FDG PET/CT in squamous cell head and neck cancers. *J Cancer Res Ther.* 2014;10(4):922-926.
19. Worden FP, Moyer J, Lee JS, et al. Chemoselection as a strategy for organ preservation in patients with T4 laryngeal squamous cell carcinoma with cartilage invasion. *Laryngoscope.* 2009;119(8):1510-1517.
20. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J. Natl. Cancer Inst.* 2008;100(4):261-269.
21. Rabalais A, Walvekar RR, Johnson JT, Smith KJ. A cost-effectiveness analysis of positron emission tomography-computed tomography surveillance versus up-front neck dissection for management of the neck for N2 disease after chemoradiotherapy. *Laryngoscope.* 2012;122(2):311-314.
22. Mehanna H, Wong WL, McConkey CC, et al. PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer. *N Engl J Med.* 2016;374(15):1444-1454.
23. Uyl-de Groot CA, Senft A, de Bree R, Leemans CR, Hoekstra OS. Chest CT and whole-body 18F-FDG PET are cost-effective in screening for distant metastases in head and neck cancer patients. *J Nucl Med.* 2010;51(2):176-182.
24. Schouten CS, de Graaf P, Alberts FM, et al. Response evaluation after chemoradiotherapy for advanced nodal disease in head and neck cancer using diffusion-weighted MRI and 18F-FDG-PET-CT. *Oral Oncol.* 2015;51(5):541-547.

CHAPTER 9

SUMMARY AND FUTURE PERSPECTIVES

NEDERLANDSE SAMENVATTING EN TOEKOMSTPERSPECTIEF

SUMMARY

Nowadays, a combination of concurrent chemotherapy and radiotherapy (CRT) is applied to patients with advanced staged head and neck squamous cell carcinomas (HNSCC), with locoregional control and organ preservation as the main treatment goals [1]. Although acceptable locoregional control rates are accomplished, the chance of residual disease remains considerable [2,3]. Prediction of treatment outcome *prior to* CRT could lead to a non-standard treatment choice in an individual patient. **Part 1** of this thesis addresses this topic. The purpose of these chapters was to investigate an association between *pretreatment* quantitative ^{18}Fluoro-2-deoxyglucose positron emission tomography – computed tomography (^{18}F-FDG-PET-CT) and diffusion-weighted MRI (DW-MRI)-values and the presence of biologically active human papillomavirus (HPV) in oropharyngeal cancer.

Traditional prognostic factors such as tumor size and lymph node invasion provide insufficient information concerning treatment outcome. Identification of additional prognostic factors, may lead to individually customized treatment and thereby possible higher responses to treatment and less treatment-induced side effects. The Standardised Uptake Value (SUV) is a semi-quantitative measurement of ^{18}F-FDG tracer uptake in PET-CT and it has been suggested that patients with low pretreatment SUV-values generally have a more favourable outcome [4,5]. However, patients with HPV-positive oropharyngeal squamous cell carcinoma (OPSCC) also show more favourable treatment response rates and prognosis as compared with HPV-negative patients [6,7]. We evaluated SUV-values in HPV-negative and HPV-positive OPSCC using a retrospective study in **chapter 2.** Forty-four patients underwent pretreatment ^{18}F-FDG-PET-CT of which twenty-seven patients were HPV-positive. HPV-positive patients had significantly smaller primary tumor volumes than HPV-negative patients. SUV-values are volume dependent because of the partial volume effect. The SUV-value in HPV-positive was 3.9 units lower than in HPV-negative tumors after correction for tumor volume. Therefore, we concluded that low pretreatment SUV-values in HPV-positive OPSCC, which are associated with a better prognosis, may be explained by HPV-induced tumor changes.

DW-MRI characterizes biologically relevant tumor features and the Apparent Diffusion Coefficient (ADC) may provide prognostic information. A few studies have shown that HNSCC with relatively low pretreatment ADC-values respond better to CRT than tumors with higher ADC-values [8-10]. HPV-related tumors have distinct histological features compared with HPV-negative OPSCC [11,12] and in **chapter 3** we investigated the possible association between ADC-values and the presence of HPV in patients with OPSCC. Forty-

four patients underwent pretreatment DW-MRI of which twenty-two patients were HPV-positive. ADC-values were not significantly different between HPV-negative- and HPV-positive tumors. Apparently, the differences in histological features between HPV-positive and HPV-negative OPSCC do not translate into different pretreatment ADC-values in our study cohort. Long-term follow-up studies with DW-MRI and HPV are needed to investigate if ADC-values and HPV-status are independent prognostic factors in OPSCC patients.

Predicting treatment outcome early *during* CRT is another potential way of treatment prediction and could hypothetically lead to a treatment switch in patients who are not likely to benefit from CRT, thereby preserving postoperative radiotherapy as an option after surgery if indicated. This kind of treatment prediction the subject in **part 2** of this thesis.

Chapter 4 describes a prospective pilot study which explores the predictive value of DW-MRI and ^{18}F-FDG-PET-CT early during CRT on locoregional outcome in patients with HNSCC. DWI is generally performed using echo-planar imaging (EPI), but non-EPI-DWI, such as Half-fourier acquisition single-shot turbo spin-echo (HASTE)-DWI can be an alternative. These techniques are also compared in this chapter. Eight patients with advanced HNSCC underwent DW-MRI (with both EPI- and HASTE-sequences) and ^{18}F-FDG-PET(-CT) pretreatment, two weeks after the start of CRT and three months after CRT. No local recurrences were detected during follow-up. Median ADC_{EPI}-values in primary tumors increased during treatment, whereas ADC_{HASTE} did not increase. SUV decreased with 62% from baseline PET to PET during treatment. Two regional recurrences were diagnosed. Early during treatment, ADC_{EPI} tended to be higher for patients with regional control than for patients with a recurrence. This difference was not seen with ADC_{HASTE}. A trend was seen for a higher change in SUV in patients with regional control than in patients with a regional recurrence. These preliminary results suggest that HASTE-DWI seems to be inadequate in early CRT response prediction, compared to EPI-DWI which has potential to predict locoregional outcome early after the start of CRT.

In case of residual disease after CRT, a 'salvage' surgery may be considered, but this is associated with substantial morbidity and complications [13] and will only be done after viable tumor has been proven histopathologically. Consequently, response evaluation is routinely performed 12 weeks after CRT in the VU University Medical Center with an examination under general anesthesia (EUA) to evaluate the primary tumor (response evaluation). Ultrasound-guided fine needle aspiration cytology (FNAC) is performed in case of residual neck disease. However, false negative biopsies due to sampling

error within the residual mass and false positive results from FNAC regularly occur [14,15]. Moreover, the discrimination between residual tumor and aspecific changes with conventional imaging is unreliable since post-treatment changes, including oedema, fibrosis and necrosis, hamper accurate assessment. **Part 3** of this thesis addresses this clinical dilemma. The purpose of these chapters was to evaluate clinical practice concerning response evaluation in the Netherlands and to evaluate the accuracy of two functional imaging techniques to detect locoregional residual disease after CRT in HNSCC.

In **chapter 5** the current clinical practice on response evaluation after CRT for advanced OPSCC is described, through a questionnaire sent to clinicians in all eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group. Response evaluation was routinely performed with various methods in five institutions (62.5%) and in one institute (12.5%) only if clinical evaluation was difficult. Two centers (25%) did not perform response evaluation. A substantial variation in the diagnostic policy and the methods used for response evaluation after CRT for advanced OPSCC in the Netherlands was found. We concluded that there is a need for national guidelines concerning response evaluation in patients with advanced oropharyngeal cancer.

In **chapter 6** the results of the REACTION-study, a prospective study, are described. The purpose was to evaluate the accuracy of ^{18}F-FDG-PET-CT, DW-MRI and combined PET-CT/DW-MRI assessment to detect local residual disease after CRT, in patients with advanced OPSCC. Forty-six patients were included and underwent response evaluation with ^{18}F-FDG-PET-CT, DW-MRI and an EUA three months after completion of CRT. Imaging was assessed by two observers per modality and combined PET-CT/DW-MRI reading was performed. Local residual disease was diagnosed in 5 patients. For PET-CT and DW-MRI, sensitivity was 75.0% and 60.0% at a specificity of 82.9% and 95.1%, yielding a NPV of 97.1% and 95.1%, respectively. Combined PET-CT/DW-MRI reading yielded a sensitivity of 100.0% at a specificity of 92.7%. This study suggests that PET-CT and DW-MRI seem to be reliable tools to rule out residual primary tumor, possibly making routine EUA superfluous. Combined PET-CT/DW-MRI reading may be superior to PET-CT and DW-MRI alone, but further research should assess this role of PET-MRI.

Chapter 7 evaluates the accuracy of DW-MRI and ^{18}F-FDG-PET-CT to detect residual lymph node metastases after CRT in patients with advanced (N2-N3) pretreatment nodal disease of HNSCC. Retrospectively, routinely performed DW-MRI (n=73) and ^{18}F-FDG-PET-CT (n=58) 3 months after CRT were assessed by two radiologists and two nuclear medicine physicians. Five patients had residual regional disease. DW-MRI showed a sensitivity of 60% and a specificity of 93%, versus 100% and 84% for PET-CT, respectively. A

statistical combination of PET-CT and DW-MRI showed a sensitivity of 100% and a specificity of 95%. PET-CT recognised all regional residues and DW-MRI recognised most patients with regional control. A combined approach seemed to increase the specificity of PET-CT alone, without compromising sensitivity of PET-CT alone.

An exploration of the diagnostic accuracy and cost-effectiveness of four response evaluation strategies to detect local residual disease in 46 oropharyngeal cancer patients treated with CRT, is performed in **chapter 8**. Besides the reference strategy, *i.e.* EUA for all patients, we considered three imaging strategies, namely ^{18}F-FDG-PET-CT-based selection for EUA, DW-MRI-based selection for EUA and a combination of ^{18}F-FDG-PET-CT and DW-MRI to select for EUA. All analyses were conducted using a decision-analytic model based on trial data of the patients from chapter 6 and scientific literature. The EUA strategy led to 96% correct diagnoses and the expected costs were €468 per patient whereas 89% of EUA indications were unnecessary. The DW-MRI strategy was the least costly strategy (€297 per patient), but also led to the lowest proportion of correct diagnoses, i.e. 93%. The PET-CT strategy (€1177 per patient) and combined imaging strategy (€1395 per patient) were dominated by the EUA strategy due to respectively a smaller or equal proportion of correct diagnoses, at higher costs. Based on our model results, the combined PET-CT and DW-MRI strategy is preferred over immediate EUA since it reaches the same diagnostic accuracy while leading to substantially less unnecessary EUA indications. This strategy costs only an additional €927 per patient. However, if healthcare resources are limited, DW-MRI is the strategy of choice because of lower costs while still providing a large reduction in unnecessary EUA indications.

FUTURE PERSPECTIVES

HNSCC is the sixth most common cancer, with a world-wide incidence of approximately 630.000 cases [16]. Over the last decades, it has been confirmed that besides tobacco smoking and alcohol consumption, HPV is an important etiological factor in the development of HNSCC, in particular in OPSCC. Patients with HPV-positive OPSCC more often present with regionally advanced disease, but have better overall survival than HPV-negative OPSCC [6,7]. In this context, traditional prognostic factors such as tumor size and lymph node invasion are insufficient to classify patients into risk groups. Identification of additional prognostic tumor characteristics may lead to an improved patient selection and individualized treatment approaches when primary treatment is selected. However, numerous factors, such as biomarkers of response, molecular characteristics, cancer stem cells, genetic alterations and patient factors need to be considered. Future research should address the identification of these factors, thereby developing more personalized treatment approaches.

Functional imaging methods like PET-CT and DW-MRI could also have prognostic value [4,5,8,10]. For diagnosing an acute cerebral stroke, diffusion-weighted MRI is an established imaging technique [17]. In head and neck cancer there is increasing experience with DW-MRI, for several indications such as primary tumor detection, prediction of treatment outcome and detecting residual or recurrent tumor after treatment [18,19]. ^{18}F-FDG-PET-CT is also frequently used for these indications, but DW-MRI has several advantages over PET-CT. DW-MRI does not use ionising radiation, it produces high soft-tissue contrast and it is less expensive since it can be implemented on most MRI scanners in addition to existing protocols. However, for PET-CT standardised imaging protocols are available to provide accuracy, reproducibility and quantification of the images, also between different institutions in multi-centre research [20,21]. A common standardised procedure increases the value of publications and their contribution to evidence-based medicine. Currently, for DW-MRI no standard acquisition protocols exits, nor have multicentre trials been performed. To be able to implement DW-MRI in multicentre studies, its reproducibility across different institutions and MR systems should be validated and these parameters have not been established up to now. Moreover, the optimal DWI sequence for interpretation of the head and neck area is unclear. An EPI-sequence is most commonly performed and is probably the most promising technique, due to a high signal-to-noise ratio and fast image acquisition [8,22,23]. Kolff-Gart et al. performed a study in healthy subjects to assess the reproducibility of ADC-values of different DWI sequences and

showed than an EPI-DWI sequence had the best reproducibility for tissues in the head and neck area [24]. Besides, several technical issues need to be resolved in the future, such as the prevention of artefacts and geometric distortions. Sufficient experience should be gained in single institution studies concerning above mentioned various acquisition parameters, before starting multicentre research.

DW-MRI can be interpreted qualitatively and quantitatively. Qualitative assessment is done by visual inspection of the signal intensity on high b-value images and the corresponding ADC-maps [25]. However, it is a subjective task and accurate interpretation is dependent on observer experience and is prone to interobserver variation. Up to now, no established scoring system is available. Future studies are needed to develop such a scoring system and to validate interpretation criteria. Quantification with ADC-values is another way to interpret DW-MRI. ADC-values are mostly obtained by drawing a region of interest (ROI) over a specific lesion [18]. However, no uniform method has been defined for ADC determination. The applied b-values, pulse sequences, field strength and delineation methods all influence ADC-values and statistical pooling is therefore not yet performed. Subject of research are more advanced methods of ADC calculation, such as histogram analysis [26] and parametric response map (voxel-by-voxel changes) [27], to provide detailed information about tumor heterogeneity or changes during treatment, respectively, and these methods show potential additional value to simple ADC calculation methods.

Another advanced imaging technique is dynamic contrast enhanced (DCE)-MRI, also referred to as perfusion-weighted MRI. DCE-MRI uses the serial acquisition of multiple images, performed after the intravenous injection of gadolinium, to assess microvascular properties of tissue perfusion [28,29]. The contrast medium enters the tumor circulation and diffuses into the extravascular space. Quantitative analysis can be performed using different parameters such as K^{trans} (transfer coefficient; rate of volume transfer between plasma and extravascular space), κ_{ep} (rate constant between plasma and extracellular space) or v_e (volume of extracellular space). Malignancies induce angiogenesis but these vessels are characterized by a lack of normal functionality with high leakage and tortuosity [30]. Tissue perfusion is important for the efficacy of CRT and DCE-MRI may serve as a marker of abnormal tumor perfusion and hypoxia and has potential to predict treatment failure. K^{trans} is most frequently used in diagnostic and prognostic studies in HNSCC and lower K^{trans} seems to predictive of poor prognosis [28,29]. Larger studies are required to confirm this potential predictive role of DCE-MRI.

^{18}F-FDG-PET-CT is routinely used for similar applications as DW-MRI in HNSCC, such as the identification of an occult primary tumor in case of cervical lymph node metastasis, detecting recurrent tumor during follow-up and the detection of distant metastasis [33,34]. ^{18}F-FDG uptake can be evaluated quantitatively and qualitatively. Semi-quantitative assessments include SUV, of which the maximum SUV (SUV_{max}) is most often used, metabolic tumor volume and total lesion glycolysis. Qualitative assessments include visual assessment using five-point Likert scales. Examples of Likert scales are the five-point Deauville scale and the Hopkins scale. The Deauville scale is originally developed for malignant lymphoma and relates residual uptake to uptake in the mediastinal blood pool and liver uptake. This scale was used by Sjovall et al. in head and neck cancer patients [31]. The five-point Hopkins scale is also based on intensity, but relative to the jugular vein rather than the mediastinal blood pool [32]. This scale was used in chapter 6. Further research is needed to determine which interpretation for ^{18}F-FDG-PET-CT performs best in diagnosing residual or recurrent primary tumors and lymph node metastases following CRT in patients with head and neck cancer.

In this thesis, we focused on the role of PET-CT with ^{18}F-FDG. ^{18}F-FDG is a glucose-based tracer and measures glucose metabolism in malignant tissue, but inflammation which regularly occurs in irradiated tissue can also lead to increased ^{18}F-FDG uptake. The use of alternative radioactive tracers might increase the accuracy of PET-CT in the differentiation between post-treatment changes and residual tumor. Several radiopharmaceuticals are under investigation. [^{18}F]-fluoromisonidazole (FMISO) and [^{18}F]-fluoroazomycin-arabinoside (FAZA) are experimentally used for the identification of tumor hypoxia. These tracers might have important consequences for the use of PET-CT in PET-based radiotherapy planning, as hypoxic areas are significantly less sensitive to radiotherapy [35-37]. Increased proliferative activity is another characteristic feature of malignancies and the tracer [^{18}F]-fluorothymidine (FLT) is able to characterise tumor cell proliferation [37]. Also, a radiolabeled amino acid, methyl-[1-11C]-L-methionine (MET), has been studied and shows promising results as the uptake in inflammation tissues seems less than with ^{18}F-FDG [37-39]. Larger studies are needed using the above mentioned tracers in comparison with ^{18}F-FDG in head and neck cancer.

The presented studies in this thesis are mainly performed with PET-CT, except in chapter 4 in which standalone PET was performed. Currently, PET-CT scanners are the 'state of the art', in which CT is used for anatomical correlation and attenuation correction. Without anatomical co-registration, PET readings can be difficult to interpret in the context of response evaluation due

to the abundance of physiological uptake in combination with disease-related altered anatomy. Recently, hybrid-PET-MRI systems entered the clinical field and PET-MRI methodology has become available which facilitates research to the potential added value of either modality. The combination of PET-MR can likely increase accuracy of PET-CT as the number of false-positives, due to nonspecific inflammation, might be decreased [40,41]. However, several issues remain a diagnostic challenge in combined PET-MRI imaging. First, due to motion mismatch between MRI and PET acquisition, susceptibility and attenuation correction artefacts can occur and these can impede the assessment of PET-MRI scans. Second, no interpretation system is available for the assessment of PET-MRI and the way to deal with discrepant findings between conventional MRI, DW-MRI and PET is unclear. At last, quantification using MRI-based attenuation and correction methods is still under investigation. Up to now, data assessing the accuracy of hybrid PET-MRI systems are scarce and only large future studies can resolve the above mentioned issues and answer the question if PET-MRI surpasses PET-CT and DW-MRI.

NEDERLANDSE SAMENVATTING

Een combinatie van gelijktijdige chemotherapie en radiotherapie (CRT) wordt tegenwoordig steeds vaker toegepast bij patiënten met een vergevorderd plaveiselcelcarcinoom in het hoofd-halsgebied. Locoregionale controle en behoud van functie zijn de twee belangrijkste doelstellingen van deze behandeling [1]. Alhoewel een acceptabele locoregionale controle wordt bereikt, blijft de kans op een residu aanzienlijk [2,3]. Het voorspellen van de behandeluitkomst *voorafgaand aan* de CRT zou kunnen leiden tot een niet-standaard behandelvoorstel aan een individuele patiënt. **Deel 1** uit deze dissertatie besteedt aandacht aan dit onderwerp. Het doel van deze hoofdstukken is het onderzoeken van een associatie tussen kwantitatieve waarden van ^{18}Fluoro-2-deoxyglucose positron emissie tomografie – computer tomografie (^{18}F-FDG-PET-CT) en diffusie-gewogen MRI (DW-MRI), gemaakt *voorafgaand aan* de behandeling, en het humaan papillomavirus (HPV) in patiënten met een orofarynxcarcinoom.

Traditionele prognostische factoren zoals tumor grootte en de aanwezigheid van lymfekliermetastasen bieden onvoldoende informatie over de uitkomst van de behandeling. Het identificeren van aanvullende prognostische factoren zou kunnen leiden tot het aanpassen van de behandeling aan een individuele patiënt en mogelijk kunnen leiden tot een hogere locoregionale controle en minder bijwerkingen. ^{18}F-FDG opname kan op een PET-CT gekwantificeerd worden met behulp van SUV-waarden ('Standardized Uptake Value'). Lage SUV-waarden voorafgaand aan de behandeling zijn geassocieerd met een betere respons op de behandeling [4,5]. Maar patiënten met een HPV-positief orofarynxcarcinoom laten ook een betere respons op CRT zien, vergeleken met HPV-negatieve patiënten [6,7]. **Hoofdstuk 2** beschrijft een retrospectieve studie waarin SUV-waarden worden vergeleken tussen HPV-negatieve en HPV-positieve orofarynxcarcinoom-patiënten. Vierenveertig patiënten werden geanalyseerd die een PET-CT voorafgaand aan de behandeling hadden ondergaan. Hiervan waren zevenentwintig patiënten HPV-positief. De SUV-waarden in HPV-positieve patiënten waren significant gemiddeld 3.9 punten lager dan in HPV-negatieve patiënten, na correctie voor tumor-volume, want SUV-waarden zijn afhankelijk van tumor volume door het partiële volume-effect. De relatief lage ^{18}F-FDG opname in HPV-positieve tumoren, die geassocieerd is met een betere prognose, kan waarschijnlijk verklaard worden door HPV-geïnduceerde tumorveranderingen.

DW-MRI karakteriseert biologisch relevante tumor-kenmerken en de afgeleide ADC-waarde (Apparent Diffusion Coefficient) heeft mogelijk prognostische waarde. Verschillende studies hebben laten zien dat hoofd-halstumoren met een relatief lage ADC-waarde beter reageren op CRT dan

tumoren met een hoge ADC-waarde [8-10]. HPV-gerelateerde tumoren hebben verschillende histologische kenmerken vergeleken met HPV-negatieve tumoren [11,12] en in **hoofdstuk 3** onderzoeken we een mogelijk verband tussen de HPV-status en ADC-waarden in patiënten met een orofarynxcarcinoom. Vierenveertig patiënten hebben een DW-MRI voorafgaand aan de behandeling ondergaan en werden geanalyseerd. Tweeëntwintig patiënten hadden een HPV-positieve tumor. De ADC-waarden waren niet significant verschillend tussen HPV-negatieve en HPV-positieve tumoren. Schijnbaar leiden verschillen in histologische kenmerken tussen HPV-positieve en HPV-negatieve orofarynxcarcinomen niet tot verschillende ADC-waarden in onze studiepopulatie. Lange termijn follow-up studies zijn nodig om vast te stellen of de ADC-waarde een onafhankelijke prognostische factor is.

Het evalueren van de CRT *vroeg tijdens* behandeling is een andere mogelijke manier om de behandeluitkomst te voorspellen en zou kunnen leiden tot een aanpassing van de behandeling bij patiënten die geen baat hebben bij deze behandeling. Deze manier van voorspellen van de behandelrespons wordt onderzocht in **deel 2** van dit proefschrift.

In **hoofdstuk 4** wordt een prospectieve pilot studie beschreven waarin de potentieel voorspellende waarde van DW-MRI en ^{18}F-FDG-PET-CT vroeg tijdens de CRT op de locoregionale controle van patiënten met een hoofd-halscarcinoom wordt onderzocht. DW-MRI wordt meestal uitgevoerd met een EPI-sequentie, maar een non-EPI sequentie zoals HASTE-DWI, kan een alternatief zijn. Deze twee sequenties worden ook met elkaar vergeleken in dit hoofdstuk. Acht patiënten met een vergevorderd hoofd-halscarcinoom ondergingen een DWI-MRI (met EPI- en HASTE-sequenties) en een ^{18}F-FDG-PET-CT voorafgaand aan, 2 weken na start van de CRT en 3 maanden na het einde van de CRT. Geen van de patiënten ontwikkelden een lokaal recidief. De mediane ADC_{EPI}-waarden van de primaire tumoren stegen tijdens de behandeling, terwijl de ADC_{HASTE}-waarden gelijk bleven. De ΔSUV daalde tussen de PET-CT voorafgaand aan de behandeling en de PET-CT tijdens behandeling. Twee patiënten ontwikkelden een regionaal recidief. In patiënten met regionale controle werd een tendens gezien met hogere ADC_{EPI}-waarden tijdens de behandeling dan in patiënten met een regionaal recidief. Dit verschil werd niet gezien bij ADC_{HASTE}. Tevens werd een tendens gezien met een groter verschil in SUV-waarden in patiënten met regionale controle dan in patiënten met een regionaal recidief. Onze preliminaire resultaten suggereren dat EPI-DWI vroeg tijdens CRT potentie lijkt te hebben om de locoregionale controle te voorspellen terwijl HASTE-DWI deze potentie niet laat zien in de huidige studie.

Indien een residu gediagnosticeerd wordt na de CRT, kan een 'salvage' operatie overwogen worden, maar een dergelijke chirurgische ingreep gaat gepaard met aanzienlijke risico's op bijwerkingen en complicaties en zal alleen uitgevoerd worden als vitaal tumorweefsel histopathologisch wordt aangetoond. Daarom wordt in het VU Medisch Centrum routinematig responsevaluatie uitgevoerd, 12 weken na het einde van de CRT met een onderzoek onder narcose (OON) voor de primaire tumor en een echogeleide cytologische punctie voor klieren in de hals. Echter, fout negatieve biopsieën en fout positieve cytologische uitslagen komen regelmatig voor [14,15]. Daarnaast is het maken van een betrouwbaar onderscheid tussen postradiatie-effecten (oedeem, fibrose en laesies van de mucosa) en een residu op conventionele beeldvorming (CT en/of MRI) vaak lastig. **Deel 3** van deze dissertatie richt zich op dit klinische probleem. Het doel van deze hoofdstukken was het evalueren van de huidige klinische praktijk betreffende responsevaluatie en het onderzoeken van de mogelijkheden van functionele beeldvormingstechnieken om een residu van een hoofd-halscarcinoom na eerdere CRT te detecteren.

Om de huidige manier van responsevaluatie na chemoradiatie voor orofarynxcarcinomen te evalueren, werd in **hoofdstuk 5** een enquête uitgestuurd naar hoofd-halschirurgen in de 8 centra die hoofd-halskanker behandelen in Nederland. Alle centra retourneerden de enquête. Responsevaluatie wordt routinematig uitgevoerd in 5 centra (62.5%), 2 centra voeren geen responsevaluatie uit (25%) en in 1 centrum alleen als de klinische beoordeling moeilijk is (12.5%). Er werd een behoorlijke variatie gevonden tussen de methoden en beeldvormingstechnieken die worden gebruikt voor responsevaluatie. Concluderend, er is behoefte aan een duidelijke richtlijn over responsevaluatie in patiënten met een orofarynxcarcinoom.

In **hoofdstuk 6** worden de resultaten van de prospectieve REACTION studie beschreven. Hierin worden de waarde van ^{18}F-FDG-PET-CT, DW-MRI en een gecombineerde beoordeling geëvalueerd om een lokaal residu na CRT bij patiënten met een orofaryxncarcinoom te detecteren. Zesenveertig patiënten werden geïncludeerd en ondergingen een ^{18}F-FDG-PET-CT, DW-MRI en een OON 3 maanden na het einde van CRT. De beeldvorming werd geëvalueerd door twee beoordelaars per modaliteit en tevens werd een gecombineerde PET en MRI beoordeling uitgevoerd. Histopathologische uitslagen en een follow-up van 9 maanden na het einde van de CRT werden gebruikt als referentiestandaard. Een lokaal residu werd gediagnosticeerd bij 5 patiënten. Voor PET-CT en DW-MRI was de sensitiviteit 75.0% en 60.0%, bij een specificiteit van 82.9% en 95.1%, leidend tot een negatief voorspellende waarde (NPV) van 97.1% en 95.1% respectievelijk. De gecombineerde PET-CT/DW-MRI beoordeling leidde tot een sensitiviteit van 100% bij een specificiteit van 92.7%. Concluderend lijken zowel PET-CT als DW-MRI

betrouwbare testen om een lokaal residu uit te sluiten, wat een OON mogelijk overbodig maakt. Een gecombineerde PET-CT/DW-MRI beoordeling is mogelijk superieur ten opzichte van PET-CT en DW-MRI alleen, maar toekomstige studies moeten deze rol verder onderzoeken.

Hoofdstuk 7 evalueert de waarde van ^{18}F-FDG-PET-CT en DW-MRI bij de detectie van een regionaal residu na CRT. Routinematig uitgevoerde DW-MRI (n=73) en PET-CT (n=58) 3 maanden na het einde van de CRT, werden geëvalueerd door twee radiologen en twee nucleair geneeskundigen (individueel en in consensus). De referentiestandaard bestond uit een follow-up van 9 maanden na het einde van de CRT en eventuele histopathologische uitslagen. Vijf patiënten hadden een regionaal residu. Een sensitiviteit van 60% en een specificiteit van 93% werden gevonden voor DW-MRI versus 100% en 84% voor PET-CT. Een modelmatige combinatie van PET-CT en DW-MRI liet een sensitiviteit van 100% en een specificiteit van 95% zien. Concluderend identificeerde PET-CT alle patiënten met een regionaal residu en DW-MRI alle patiënten met regionale controle. Een gecombineerde benadering (PET-CT en DW-MRI) zorgde voor een verhoging van de specificiteit van een PET-CT alleen, zonder de sensitiviteit nadelig te beïnvloeden.

Een exploratie van de diagnostische accuraatheid en kosteneffectiviteit van vier verschillende responsevaluatie strategieën om een lokaal residu na CRT te detecteren in 46 orofarynxcarcinoom-patiënten, wordt beschreven in **hoofdstuk 8**. Naast een referentiestrategie, te weten het OON bij alle patiënten, werden drie andere strategieën onderzocht, namelijk ^{18}F-FDG-PET-CT, DW-MRI en een PET-MRI combinatie, gevolgd door een OON in het geval van een positieve testuitslag. Alle analyses zijn uitgevoerd gebruik makende van een 'besliskundig'-model gebaseerd op de data van de 46 patiënten uit hoofdstuk 6 en de wetenschappelijke literatuur. De referentiestrategie leidde tot 96% juiste diagnoses, de verwachte kosten waren €468 per patiënt, waarbij 89% van de OON overbodig waren. De DW-MRI strategie was de goedkoopste strategie (€297 per patiënt), maar leidde ook tot de laagste proportie correcte diagnoses, namelijk 93%. De PET-CT strategie (€1177 per patiënt) en de gecombineerde strategie (€1395 per patiënt) werden gedomineerd door de OON-strategie door een kleiner of gelijk aantal correcte diagnoses tegen hogere kosten. Gebaseerd op bovenstaande model, wordt de gecombineerde PET-MRI strategie verkozen boven de referentiestrategie omdat hiermee gelijke accuraatheid met substantieel minder OON wordt bereikt, terwijl deze strategie maar €927 meer kost per patiënt. DW-MRI lijkt de strategie van keuze te zijn als de zorgmiddelen beperkt zijn, omdat hierbij de kosten minder zijn, terwijl er wel een behoorlijke reductie in het aantal onnodige OON plaatsvindt.

TOEKOMSTPERSPECTIEF

Plaveiselcelcarcinomen in het hoofd-halsgebied zijn de 6e meest voorkomende kankersoort met een wereldwijde incidentie van ongeveer 630.000 patiënten per jaar [16]. Gedurende de laatste decennia is bekend geworden dat naast roken en het drinken van alcohol, het HPV virus een belangrijke rol speelt in de ontwikkeling van met name orofarynxcarcinomen. Patiënten met een HPV-positieve tumor presenteren zich vaker met lymfekliermetastasen in de hals, maar hebben toch een betere overlevingskans dan HPV-negatieve patiënten [6,7]. In dit licht zijn traditionele risicofactoren, zoals tumor-grootte en lymfeklier-metastasering, ontoereikend om patiënten in risicogroepen in te delen. Het identificeren van additionele risicofactoren zou kunnen leiden tot een individueel behandelplan. Echter, veel verschillende factoren zoals biomarkers, moleculaire karakteristieken, kankerstamcellen, genetische veranderingen en patiëntgebonden factoren spelen hierin een rol. Toekomstig onderzoek moet zich richten op het identificeren van deze factoren die de respons op een behandeling zouden kunnen voorspellen.

Functionele beeldvormingstechnieken, zoals PET-CT en DW-MRI, kunnen prognostische waarde hebben [4,5,8,10]. DW-MRI is een gevestigde techniek die gebruikt wordt bij het diagnosticeren van een acuut herseninfarct [17]. Ook in de hoofd-halsoncologie wordt steeds meer ervaring opgedaan met DW-MRI; onder andere bij de detectie van primaire tumoren, het voorspellen van de therapie-respons en het detecteren van een residu na CRT [18,19]. ^{18}F-FDG-PET-CT wordt ook voor deze indicaties gebruikt, maar DW-MRI heeft een aantal voordelen boven PET-CT; DW-MRI maakt geen gebruik van ioniserende straling en het is goedkoper omdat het geïmplementeerd kan worden in bestaande MRI-protocollen. Echter, voor PET-CT zijn gestandaardiseerde protocollen beschikbaar wat zorgt voor reproduceerbaarheid tussen verschillende ziekenhuizen in multicenter-onderzoek [20,21]. Tevens, vergroten gestandaardiseerde protocollen de waarde van wetenschappelijke publicaties. Tot op heden zijn er voor DW-MRI geen gestandaardiseerde protocollen beschikbaar en heeft multicenter-onderzoek nog niet plaatsgevonden. Om dit te bewerkstelligen moet de reproduceerbaarheid van DW-MRI in verschillende instituten en op verschillende scanners bekend zijn. Daarnaast is de optimale MRI-sequentie voor de hoofd-halsoncologie gebied nog niet bekend. Een EPI-sequentie wordt het meest gebruikt omdat deze een hoge ‘signal-to-noise’ ratio heeft en binnen enkele minuten gemaakt kan worden [8,22,23]. Ook heeft Kolff-Gart et al. laten zien dat een EPI-sequentie de beste reproduceerbaarheid heeft voor weefsels in het hoofd-halsgebied [24]. Daarnaast moeten verschillende technische problemen in de toekomst opgelost worden, zoals de

preventie van artefacten en geometrische distorsies. Eerst moet voldoende ervaring opgedaan worden in individuele instituten voordat multicenter studies kunnen worden opgestart.

Een DW-MRI kan kwalitatief en kwantitatief beoordeeld worden. Een kwalitatieve beoordeling wordt gedaan door visuele interpretatie van de signaalintensiteit op een diffusie-beeld met een hoge b-waarde en de corresponderende ADC-map [25]. Echter, dit is een subjectieve manier van beoordelen en is afhankelijk van de ervaring van de beoordelaar. Er is nog geen standaard scoringssysteem beschikbaar. Toekomstige studies zijn noodzakelijk om een scoringssysteem te ontwikkelen en de interpretatiecriteria te valideren. Het kwantitatief beoordelen van een DW-MRI is mogelijk met behulp van ADC-waarden. ADC-waarden worden verkregen door een 'Region of Interest' (ROI) over een specifieke laesie te tekenen [18]. Maar er is nog geen uniforme methode gedefinieerd om ADC-waarden te bepalen en verschillende factoren zijn van invloed op de ADC-waarden; zoals de gebruikte b-waarden, de veldsterkte en de methode van delineatie van de tumor. Op dit moment wordt onderzoek gedaan naar geavanceerde methoden voor het bepalen van de ADC-waarden, zoals histogram-analyses [26] en een 'parametrische respons map' [27], om gedetailleerde informatie over tumor-heterogeniteit en veranderingen in ADC-waardes gedurende de behandeling te verkrijgen.

Een Dynamische Contrast Enhanced MRI (DCE-MRI) is een andere geavanceerde beeldvormingstechniek, ook wel perfusie-MRI genoemd. DCE-MRI maakt gebruik van de acquisitie van meerdere beelden, die gemaakt worden na de intraveneuze injectie met contrast. Hiermee worden de microvasculaire eigenschappen van weefsels in kaart gebracht [28,29]. Het contrast komt de circulatie van de tumor binnen en diffundeert naar de extravasculaire ruimte. Een kwantitatieve analyse kan gedaan worden met behulp van verschillende parameters; K^{trans} (een maat voor transfer van volume tussen plasma en de extracellulaire ruimte), κ_{ep} (snelheidsconstante tussen plasma en de extracellulaire ruimte) of v_e (het volume van de extracellulaire ruimte). Maligniteiten hebben een sterke angiogenese maar deze bloedvaten hebben een grotere mate van lekkage en kronkeligheid [30]. Daarnaast is weefsel perfusie belangrijk voor de effectiviteit van CRT en dus kan DCE-MRI mogelijk fungeren als een marker voor abnormale tumor-perfusie. K^{trans} wordt het meest gebruikt in diagnostische en prognostische studies naar hoofd-halskanker en een lage K^{trans} lijkt geassocieerd te zijn met een slechte prognose [28,29]. Grotere studies zijn nodig om deze mogelijke voorspellende rol van DCE-MRI te bevestigen.

^{18}F-FDG-PET-CT wordt routinematig gebruikt voor vergelijkbare klinische toepassingen als DW-MRI in de hoofd-halsoncologie, zoals het opsporen van een primaire tumor, het detecteren van een recidief tijdens de follow-up of het diagnosticeren van afstand metastasen [33,34]. Een ^{18}F-FDG-PET-CT kan kwantitatief en kwalitatief geëvalueerd worden. Een semi-kwantitatieve beoordeling maakt gebruik van SUV-waarden waarvan de maximale SUV-waarde (SUV_{max}) het meest wordt gebruikt. Een visuele beoordeling met behulp van een 5-punts Likert-schaal is een manier van een kwalitatieve beoordeling. Voorbeelden van een Likert-schaal zijn de 'Deauville'- en 'Hopkins'-schalen. De Deauville-schaal is oorspronkelijk ontwikkeld voor maligne lymfomen en relateert traceropname aan opname in de lever. Deze schaal werd ook gebruikt door Sjovall et al. bij hoofd-halskankerpatiënten [31]. De Hopkins-schaal is gebaseerd op de traceropname in een tumor in relatie tot opname in de vena jugularis interna [32]. Deze schaal is gebruikt in hoofdstuk 6. Meer onderzoek is nodig om te bepalen welke schaal de beste resultaten laat zien bij het detecteren van een residu na CRT bij hoofd-halskankerpatiënten.

In dit proefschrift hebben we de klinische toepasbaarheid van PET-CT met ^{18}F-FDG onderzocht. ^{18}F-FDG is een op glucose gebaseerde tracer en meet het glucosemetabolisme in weefsel. Echter, zowel maligne weefsel als inflammatie, kan ^{18}F-FDG opnemen en dit maakt het onderscheid lastig met ^{18}F-FDG-PET-CT. Het gebruik van alternatieve tracers kan mogelijk de nauwkeurigheid van PET-CT verhogen bij het onderscheid tussen postradiatie-effecten en residu tumor. Verschillende radiofarmaceutica worden onderzocht. [^{18}F]-fluoromisonidazole (FMISO) en [^{18}F]-fluoroazomycin-arabinoside (FAZA) worden experimenteel gebruikt voor het identificeren van tumor hypoxie, wat klinisch relevant kan zijn bij het gebruik van deze tracers tijdens PET-gebaseerde radiotherapie-planning. Hypoxische delen binnen een tumor zijn namelijk significant minder sensitief voor radiotherapie [35-37]. Verhoogde proliferatieve activiteit is een ander kenmerk van maligniteiten en de tracer [^{18}F]-fluorothymidine (FLT) kan tumorcel proliferatie aantonen [37]. Daarnaast wordt een radioactief gelabeld aminozuur, methyl-[1-11C]-L-methionine (MET), onderzocht die veelbelovende resultaten laat zien, met name omdat de opname in inflammatie minder zou zijn in vergelijking met ^{18}F-FDG [37-39]. Meer onderzoek met bovenstaande tracers, in vergelijking met ^{18}F-FDG, is nodig om de klinische toepasbaarheid hiervan te bepalen.

De studies in dit proefschrift zijn voornamelijk uitgevoerd met PET in combinatie met CT, behalve in hoofdstuk 4. Tegenwoordig zijn PET-CT scanners 'state of the art', waarin CT wordt gebruikt voor anatomische correlatie en attenuatie-correctie. Zonder deze anatomische co-registratie kan de interpretatie van PET in de context van responsevaluatie moeilijk zijn door

fysiologische ^{18}F-FDG-opname en postradiatie-effecten. Meer recent zijn hybride PET-MRI scanners ontwikkeld waardoor studies naar de potentieel aanvullende waarde van PET met MRI gefaciliteerd kunnen worden. De combinatie van PET en MRI kan mogelijk de accuraatheid van PET-CT verhogen doordat het aantal fout-positieven, als gevolg van traceropname in post-radiatie effecten, gereduceerd kan worden [40,41]. Verschillende problemen bestaan echter nog in de beeldvorming met PET-MRI. Ten eerste kunnen co-registratie-artefacten als gevolg van een mismatch tussen de PET en MRI acquisitie voorkomen. Ook kunnen attenuatie-correctie artefarcten voorkomen, waardoor de beoordeling van PET-MRI scans nadelig wordt beïnvloedt. Ten tweede is er nog geen interpretatiesysteem beschikbaar voor het beoordelen van PET-MRI scans en is het nog onduidelijk op welke manier moet worden omgegaan met discrepanties tussen de uitkomsten van conventionele MRI, DW-MRI en PET. Tot slot, een kwantitatieve beoordeling met behulp van op MRI-gebaseerde attenuatie-correctie methoden wordt nog onderzocht. Tot dusver zijn data over de nauwkeurigheid van hybride PET-MRI systemen schaars en zijn toekomstige grotere studies nodig om bovenstaande kwesties op te lossen en om te bepalen of PET-MRI PET-CT kan overtreffen.

REFERENCES

1. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695-1709.
2. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet.* 2000;355(9208):949-955.
3. 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(6):567-578.
4. Allal AS, Dulguerov P, Allaoua M, et al. Standardized uptake value of 2-[(18)F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy. *J. Clin. Oncol.* 2002;20(5):1398-1404.
5. Machtay M, Natwa M, Andrel J, et al. Pretreatment FDG-PET standardized uptake value as a prognostic factor for outcome in head and neck cancer. *Head Neck.* 2009;31(2):195-201.
6. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J. Natl. Cancer Inst.* 2008;100(4):261-269.
7. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* 2010;363(1):24-35.
8. Kim S, Loevner L, Quon H, et al. Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. *Clin. Cancer Res.* 2009;15(3):986-994.
9. Chawla S, Kim S, Dougherty L, et al. Pretreatment diffusion-weighted and dynamic contrast-enhanced MRI for prediction of local treatment response in squamous cell carcinomas of the head and neck. *AJR Am. J. Roentgenol.* 2013;200(1):35-43.
10. Hatakenaka M, Nakamura K, Yabuuchi H, et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81(2):339-345.
11. Thariat J, Badoual C, Faure C, Butori C, Marcy PY, Righini CA. Basaloid squamous cell carcinoma of the head and neck: role of HPV and implication in treatment and prognosis. *J. Clin. Pathol.* 2010;63(10):857-866.
12. Chernock RD, El-Mofty SK, Thorstad WL, Parvin CA, Lewis JS, Jr. HPV-related nonkeratinizing squamous cell carcinoma of the oropharynx: utility of microscopic features in predicting patient outcome. *Head Neck Pathol.* 2009;3(3):186-194.
13. Eerenstein SEJ, Van der Putten L, De Bree R. Salvage surgery following chemo-radiation for head and neck cancer: does it yield satisfactory results? *Radiotherapy Oncology.* 2007.

14. van der Putten L, van den Broek GB, de Bree R, et al. Effectiveness of salvage selective and modified radical neck dissection for regional pathologic lymphadenopathy after chemoradiation. *Head Neck.* 2009;31(5):593-603.
15. Yom SS, Machtay M, Biel MA, et al. Survival impact of planned restaging and early surgical salvage following definitive chemoradiation for locally advanced squamous cell carcinomas of the oropharynx and hypopharynx. *Am. J. Clin. Oncol.* 2005;28(4):385-392.
16. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer.* 2010;127(12):2893-2917.
17. Buckley BT, Wainwright A, Meagher T, Briley D. Audit of a policy of magnetic resonance imaging with diffusion-weighted imaging as first-line neuroimaging for in-patients with clinically suspected acute stroke. *Clin Radiol.* 2003;58(3):234-237.
18. Thoeny HC, De Keyzer F, King AD. Diffusion-weighted MR imaging in the head and neck. *Radiology.* 2012;263(1):19-32.
19. Vandecaveye V, De Keyzer F, Dirix P, Lambrecht M, Nuyts S, Hermans R. Applications of diffusion-weighted magnetic resonance imaging in head and neck squamous cell carcinoma. *Neuroradiology.* 2010;52(9):773-784.
20. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur. J. Nucl. Med. Mol. Imaging.* 2015;42(2):328-354.
21. Boellaard R, Oyen WJ, Hoekstra CJ, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur. J. Nucl. Med. Mol. Imaging.* 2008;35(12):2320-2333.
22. King AD, Mo FK, Yu KH, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. *Eur. Radiol.* 2010;20(9):2213-2220.
23. Vandecaveye V, De Keyzer F, Nuyts S, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):960-971.
24. Kolff-Gart AS, Pouwels PJ, Noij DP, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. *AJNR Am. J. Neuroradiol.* 2015;36(2):384-390.
25. Padhani AR. Diffusion magnetic resonance imaging in cancer patient management. *Semin. Radiat. Oncol.* 2011;21(2):119-140.
26. Srinivasan A, Chenevert TL, Dwamena BA, et al. Utility of pretreatment mean apparent diffusion coefficient and apparent diffusion coefficient histograms in prediction of outcome to chemoradiation in head and neck squamous cell carcinoma. *J. Comput. Assist. Tomogr.* 2012;36(1):131-137.
27. Galban CJ, Mukherji SK, Chenevert TL, et al. A feasibility study of parametric response map analysis of diffusion-weighted magnetic resonance imaging scans of head and neck cancer patients for providing early detection of therapeutic efficacy. *Transl. Oncol.* 2009;2(3):184-190.

28. Bernstein JM, Homer JJ, West CM. Dynamic contrast-enhanced magnetic resonance imaging biomarkers in head and neck cancer: potential to guide treatment? A systematic review. *Oral Oncol.* 2014;50(10):963-970.
29. Noij DP, de Jong MC, Mulders LG, et al. Contrast-enhanced perfusion magnetic resonance imaging for head and neck squamous cell carcinoma: a systematic review. *Oral Oncol.* 2015;51(2):124-138.
30. Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature.* 2000;407(6801):249-257.
31. Sjovall J, Chua B, Pryor D, et al. Long-term results of positron emission tomography-directed management of the neck in node-positive head and neck cancer after organ preservation therapy. *Oral Oncol.* 2015;51(3):260-266.
32. Marcus C, Ciarallo A, Tahari AK, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins Criteria)-interreader reliability, accuracy, and survival outcomes. *J. Nucl. Med.* 2014;55(9):1411-1416.
33. Johnson JT, Branstetter BF. PET/CT in head and neck oncology: State-of-the-art 2013. *Laryngoscope.* 2014;124(4):913-915.
34. Bussink J, van Herpen CM, Kaanders JH, Oyen WJ. PET-CT for response assessment and treatment adaptation in head and neck cancer. *Lancet Oncol.* 2010;11(7):661-669.
35. Halmos GB, Bruine de BL, Langendijk JA, van der Laan BF, Pruim J, Steenbakkers RJ. Head and neck tumor hypoxia imaging by 18F-fluoroazomycin-arabinoside (18F-FAZA)-PET: a review. *Clin. Nucl. Med.* 2014;39(1):44-48.
36. Lin Z, Mechalakos J, Nehmeh S, et al. The influence of changes in tumor hypoxia on dose-painting treatment plans based on 18F-FMISO positron emission tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2008;70(4):1219-1228.
37. Wedman J, Pruim J, Roodenburg JL, et al. Alternative PET tracers in head and neck cancer. A review. *Eur. Arch. Otorhinolaryngol.* 2013;270(10):2595-2601.
38. Lindholm P, Leskinen-Kallio S, Grenman R, et al. Evaluation of response to radiotherapy in head and neck cancer by positron emission tomography and [11C]methionine. *Int. J. Radiat. Oncol. Biol. Phys.* 1995;32(3):787-794.
39. Chesnay E, Babin E, Constans JM, et al. Early response to chemotherapy in hypopharyngeal cancer: assessment with (11)C-methionine PET, correlation with morphologic response, and clinical outcome. *J. Nucl. Med.* 2003;44(4):526-532.
40. Becker M, Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *Br. J. Radiol.* 2014;87(1036):20130677.
41. Buchbender C, Heusner TA, Lauenstein TC, Bockisch A, Antoch G. Oncologic PET/MRI, part 1: tumors of the brain, head and neck, chest, abdomen, and pelvis. *J. Nucl. Med.* 2012;53(6):928-938.

APPENDICES

ACKNOWLEDGMENTS │ DANKWOORD

Tijdens mijn oudste co-schap in het LUMC kwam ik via dr. L.A. van der Velden in contact met prof. dr. R. de Bree (waarvoor dank!), over een vacature voor een onderzoekstraject in het VU Medisch Centrum op de afdeling KNO-Heelkunde & Hoofd-Halschirurgie. Mede dankzij een eerdere ervaring op het gebied van kosten-effectiviteitsonderzoek kon ik beginnen met dit project. Met veel plezier kijk ik terug op de periode in het VUmc. Dankzij de hulp van velen is mijn promotie bijna afgerond en daar ben ik verschillende mensen dank voor verschuldigd. Een aantal personen zou ik hier graag in het bijzonder noemen.

Allereerst, alle patiënten die bereid zijn geweest belangeloos deel te nemen aan een van de prospectieve studies van dit proefschrift.

Prof. dr. R. de Bree. Beste Remco, ik heb bewondering voor jouw gedrevenheid en overgave in het doen en begeleiden van wetenschappelijk onderzoek. Bijzonder knap hoe je dit met klinische werkzaamheden weet te combineren en dan ook nog in je vrije tijd razendsnel gaat hardlopen. Gedurende het uitvoeren van dit onderzoek heb ik veel vrijheid van je gekregen, maar ik kon laagdrempelig aan de bel trekken want jouw reactie volgde bijna altijd dezelfde dag nog. Bedankt voor het vertrouwen dat je in me hebt gehad.

Prof. dr. J.A. Castelijns. Beste Jonas, via jou ben ik geïntroduceerd in de wereld van de radiologie en heb ik de mogelijkheid gekregen om ervaring op te doen met het interpreteren van met name MRI scans van het hoofd-halsgebied. We hebben veel en prettig contact gehad in de afgelopen jaren, onder andere bij het beoordelen van de vele scans; altijd had je wel een uurtje tussendoor tijd om samen MRI's te bekijken, ook al moest je dat doen volgens bepaalde interpretatiecriteria waar je het niet altijd mee eens was. Bedankt daarvoor!

Prof. dr. O. S. Hoekstra. Beste Otto, dank voor de kans die je me hebt gegeven om dit project te doen. Ik heb je supervisie, waardevolle reflecties en kritische noot, vanaf de eerste ideeën tot aan het einde van een manuscript, zeer gewaardeerd.

Prof. dr. C.R. Leemans, bedankt dat u me de mogelijkheid heeft geboden om dit onderzoek uit te voeren binnen de afdeling Keel-, Neus-, Oorheelkunde & Hoofd-Halschirurgie.

Prof. dr. E.F.I. Comans en dr. P. de Graaf. Beste Emile en Pim, bedankt voor de vele uren samen scans beoordelen. Ik heb medisch-inhoudelijk veel van jullie geleerd. Pim, jouw tomeloze inzet en enthousiasme voor het vak werkt stimulerend, bedankt voor de fijne samenwerking.

Geachte leden van de leescommissie, hartelijk dank voor het kritisch beoordelen van het manuscript en het plaats willen nemen in de leescommissie.

Prof. dr. R.H. Brakenhoff en dr. B.J.M. Braakhuis. Beste Ruud en Boudewijn, doordat het koffiezetapparaat in de kamer van de promovendi staat, kwamen jullie elke dag even koffie halen. Wetenschappelijk vraagstukken konden altijd tussendoor aan jullie voorgelegd worden, met enthousiaste reacties en verhelderende antwoorden tot gevolg. Bedankt ook voor de gezelligheid tijdens de lunchpauzes in het restaurant of de hortus botanicus.

Graag wil ik alle co-auteurs bedanken voor hun waardevolle aandeel in het tot stand komen van dit proefschrift. Vijf co-auteurs in het bijzonder. Dr. B.I. Witte, beste Birgit, bedankt voor jouw ondersteuning in het statistieke werk van de stukken in dit proefschrift. Prof. dr. E. Bloemena, fijn dat ik u mocht storen voor HPV nabepalingen van studie-patiënten. Dr. Doornaert, beste Patricia, het includeren van nieuwe patiënten in een van de prospectieve studies was een logistieke uitdaging, bedankt voor jouw hulp tijdens dit proces. Dr. V.M.H. Coupé en dr. M.J.E. Greuter, beste Veerle en Marjolein, we hebben samen uiteindelijk een mooi artikel gepubliceerd, veel dank voor jullie inspanningen hierbij.

De medewerkers van de afdeling Radiologie & Nucleaire Geneeskunde, waaronder de röntgen laboranten en MNW'ers, in het bijzonder Judith van Es, Femke Jongsma en Nasserah Sais. Bedankt voor jullie inzet bij het plannen van de scans volgens studieprotocol en hulp bij het verzamelen van de studiedata.

Team 'Samen leven met kanker'; beste dr. N. van Uden-Kraan, drs. H. Melissant, drs. F. Jansen. Beste Nelly, Heleen en Femke. Vele uren hebben jullie met patiënten getelefoneerd over hun visie op responsevaluatie en samen hebben we vervolgens deze interviews gecodeerd. Veel dank voor jullie ondersteuning van dit onderdeel van de REACTION-studie.

De multicentergroep van de REACTION-studie; prof. dr. M.W.M. van den Brekel, drs L. Lansaat, drs. T.J. Warmerdam, prof. dr. A. van der Lugt. Bedankt dat u dit project met ons heeft willen starten en voor het gastvrije ontvangst in de betreffende ziekenhuizen.

Alle overige mensen met wie ik als arts-onderzoeker heb samengewerkt in het VUmc; de hoofd-halschirurgen, AIOS KNO, medewerkers van de KNO-afdeling en polikliniek, radiotherapeuten, medisch oncologen, bedankt voor de prettige samenwerking. Ton Houffelaar, Sandra Biemans, Hanneke Tielens, Gerrie Buys, Vanessa Buijs en Maarten Regtuit, bedankt voor jullie ondersteuning.

Oud-onderzoekscollega's; Derrek Heuveling, Annette van Nieuwenhuizen, Laura Korsten, Inge Braspenning, Steven Mes, Vicky de Boer en Maarten van Loon. Mede dankzij jullie is mijn tijd in het VUmc voorbij gevlogen, bedankt voor de vele leuke momenten samen. Annette, jarenlang zijn we directe collega's geweest, dank voor het wegwijs maken op de afdeling en vrolijke gesprekken tijdens de lunch. Laura, zo gezellig dat jij de onderzoekskamer kwam versterken! Veel succes met het afronden van je onderzoek en de start van de opleiding volgend jaar. Hopelijk tot snel tijdens een koffietje in de stad.

'De Sao Paulo congres groep' zoals ook in het proefschrift van Sophie vermeld. Hester, Caroline, Simone, Sophie, Ellen, Saar, Steven, Baris en Pim. Dit was echt een week om niet te vergeten!

De stafleden van de afdeling Keel-, Neus- en Oorheelkunde & Hoofd-Halschirurgie van het RadboudUMC, in het bijzonder prof. dr. H.A.M. Marres en dr. F.J. A. van den Hoogen. Bedankt dat jullie het vertrouwen hebben gehad en mij de mogelijkheid hebben geboden om aan het RadboudUMC de opleiding tot KNO-arts te kunnen volgen.

Collega (oud-) arts-assistenten en arts-onderzoekers KNO uit het RadboudUMC: Lisette, Thijs, Josephine, Saskia, Anne, Rik, Ruud, Eline, Ingrid, Jasmijn, Henrieke, Caroline, Annemarie, Joost, Erik, Hubert, Richard, Chrisje, Corinne, David, Stijn, Luuk, Bas, Brechje, Floris, Mayke, Ivo, Machteld, Mieke, Jeroen, Coosje en Cindy. Ondanks dat ik direct in de kliniek aan het werk ging, werd ik welkom ontvangen binnen de groep. Bedankt voor de fijne samenwerking in het ziekenhuis, maar zeker ook voor de sociale activiteiten buiten het Radboud. Jasmijn, jij hebt mij overtuigd van het prettige opleidingsklimaat in Nijmegen, waarvoor veel dank. Je bent een topprofessional met dik verdiend fellowship! Ingrid, David en Bas, bijzonder dat het zo gemakkelijk en gezellig verloopt als collega's samenwonend in een huis.

Caroline, gedurende mijn eerste jaar heb jij mij meegenomen in het Nijmeegse leven. Drankjes, naar de bioscoop, hardlopen, enzovoort. Dit heb ik enorm gewaardeerd. Dames van de bootcampclub, mooi nieuw initiatief! Hopelijk volgen nog vele uren springen en kruipen door het Valkhofpark.

Dank aan de maatschap KNO van het Jeroen Bosch Ziekenhuis. In mijn perifere jaar bij jullie heb ik de ruimte gekregen om aan dit proefschrift te werken. Bedankt voor alles wat ik in mijn tweede jaar bij jullie heb geleerd. Dames van de poli, Cynthia en Karlijn in het bijzonder, bedankt voor de gezellige tijd in Den Bosch.

Mijn paranimfen, lieve Sara en Michelle. 'Charlies Angels' nogmaals samen tijdens een verdediging! De vriendschap met jullie is een van de meest waardevolle aspecten die deze promotie me heeft gebracht. Lieve Saar, omdat we vergelijkbaar onderzoek deden hebben we lief en leed met elkaar gedeeld vanaf het begin. Ik hoop dat ik jou op moeilijke momenten heb kunnen bijstaan. Bewonderenswaardig hoe jij sterker hieruit bent gekomen en het geluk nu weer toelacht. Lieve Mitch, als kamergenoot hebben we jarenlang een top tijd gehad. Bovendien heb je me inhoudelijke tips en waardevolle adviezen gegeven over 'het promoveren'. Jouw ambitie, doortastendheid en besluitvaardigheid is inspirerend.

Erik-Jan Bosman, dank voor jouw hulp bij de lay-out van het binnen- en buitenwerk van dit boekje!

Hurley dames 8 (dames 11, dames 9, dames …). Al jaren het leukste en gezelligste (maar ook fanatieke team) in het Amsterdamse bos! Bedankt voor alle sportieve momenten op het hockeyveld en mooie momenten tijdens wintersport, weekendjes weg, toernooien, kaarten of in de kroeg. Lieve Dominique, naast teamgenoot ook huisgenoot. Bedankt voor de fijne gesprekken en mooie tijd die we hebben gehad in ons appartement in Oost.

De KIP-groep. 'Op Koninginnedag swipete ik het Kippetje naar rechts', het begin van de KIP-groep; Suus & Guus, Louk & Patriek, Fem & Jor, de ABtjes, Birnie & RJ. Bijzonder om met jullie in deze samenstelling zo leuk contact te hebben. Ik hoop op nog vele borrels en barbecues samen.

Oud-huisgenoten van de Singel. Lieve Rinkel, ik waardeer jouw luisterend oor, onvoorwaardelijke steun, zorgzaamheid en mening enorm. Fijn dat we voorlopig dicht bij elkaar blijven wonen. Lieve Venlo, Miss Lisse en Miss Venlo uit het 2004, door dik en dun! Dank voor de nodige afleiding, ook de afgelopen jaren, tijdens een van onze borrels of weekenden weg.

Lieve Pip, wij kennen elkaar sinds je mijn trouwe sjaarsje was. Ik heb genoten van onze tijd in Leiden. Je vond alles mooi en ging overal mee naar toe, en nog steeds! Lieve Teun, zowel op de Singel als in Amsterdam huisgenoten, je staat altijd klaar voor je vrienden. Lieve Nout, Ali en CoCo, dat er nog veel mooie borrels en weekendjes Kerstshoppen mogen komen.

Mijn oudste en lieve vriendinnetjes sinds de basis- en middelbare school. Lieve Céline, Suzanne en Madeleine, wat ben ik blij dat we elkaar nog regelmatig zien. Met alle gezelligheid tijdens borrels, diners, festivals, terrasjes of sporten hebben jullie mij geholpen het promoveren even los te laten. Lieve Céline, van mijn vriendinnen ken ik jou het langst, vanaf het moment op de glijbaan, al bijna 30 jaar. Je bent me enorm dierbaar.

Lieve papa en mama, Omi, broer en kleine broertje, Denise en Noor. Papa en mama, mede door jullie heb ik dit moment bereikt. Jullie geven me onvoorwaardelijke steun en liefde. Het is fijn te weten dat we altijd bij jullie terecht kunnen. ‘A’, met veel plezier denk ik terug aan onze gezamenlijke tijd in Amsterdam: standaard woensdagavond hardlopen in het Vondelpark en daarna samen eten met Denise. Wat een lef van jou om een eigen praktijk te beginnen, en met succes! Lieve Pauli, alhoewel je in Japan een fantastische tijd hebt gehad, ben ik blij dat jij en Noor weer terug zijn. Ik miste het kuieren, borrelen op het terras, gesprekken tijdens het koffiedrinken, enz. Jij hebt het direct in de gaten als er iets is. Ik geniet ervan samen met jou te ‘zijn’ en ben enorm trots op je!

Lieve Arnoud, ‘pas’ twee jaar in mijn leven, maar niet meer weg te denken. Je hebt de laatste fase van dit proefschrift meegekregen, maar zorgde vooral voor een hoop belevenissen en mooie reizen samen, naast het promoveren. Ik ben zo blij dat ik je ontmoet heb; jij helpt mij ontspannen als ik druk ben, begeleid mij in mijn keuzes als ik twijfel en stimuleert mijn ambities. Ik kijk enorm uit naar onze toekomst samen, waar dit ook zal zijn.

CURRICULUM VITAE

Charlotte Sophie Schouten, auteur van dit proefschrift, werd op 8 februari 1986 geboren te Oranjestad, Aruba. Ze groeide op in Lisse als middelste in het gezin met 2 broers en behaalde in 2004 haar gymnasium diploma op het Fioretti College. Dat jaar verhuisde zij naar Leiden om te beginnen met de studie Rechtsgeleerdheid aan de Universiteit van Leiden. In 2005 startte zij met de studie Geneeskunde, tevens aan de Universiteit van Leiden. Voorafgaand aan de co-schappen deed zij wetenschappelijk onderzoek binnen de afdelingen Kindergeneeskunde en KNO-Heelkunde van het LUMC naar de kosten-effectiviteit van de neonatale gehoorscreening. Ze deed twee buitenlandse co-schappen; in 2009 een co-schap tropengeneeskunde in het Holy Family Medical Center, Kameroen en in 2010 het co-schap dermatologie in het Academisch Ziekenhuis Paramaribo, Suriname. In maart 2012 behaalde zij haar artsexamen.

In mei 2012 begon ze onder begeleiding van prof. dr. R. de Bree (KNO-Heelkunde), prof. dr. J.A. Castelijns (Radiologie) en prof. dr. O.S. Hoekstra (Nucleaire Geneeskunde) met wetenschappelijk onderzoek op de afdeling Keel-, Neus-, Oorheelkunde & Hoofd-Halschirurgie in het VU Medisch Centrum, waarvan de resultaten in dit proefschrift beschreven staan.

In december 2014 startte zij met de opleiding tot Keel-, Neus- en Oorarts in het RadboudUMC te Nijmegen (onder prof. dr. H.A. Marres en dr. F.J van den Hoogen), waar ze tot op heden mee bezig is. Tijdens de opleiding heeft ze ook gewerkt in het Jeroen Bosch ziekenhuis te Den Bosch (onder begeleiding van drs. A.V.M. de Visscher en dr. S.G.T. Balter).

LIST OF PUBLICATIONS

M.J.E. Greuter, **C.S. Schouten**, J.A. Castelijns, P. de Graaf, E.F.I. Comans, O.S. Hoekstra, R. de Bree, V.M.H. Coupé
Cost-effectiveness of response evaluation after chemoradiotherapy in patients with advanced oropharyngeal cancer using ^{18}F-FDG-PET-CT and diffusion-weighted MRI
BMC Cancer. 2017; 17:256-4

C.S. Schouten, S. Hakim, R. Boellaard, E. Bloemena, P.A. Doornaert, B.I. Witte, B.J.M. Braakhuis, R.H. Brakenhoff, C.R. Leemans, O.S. Hoekstra, R. de Bree
Interaction of quantitative ^{18}F-FDG-PET-CT imaging parameters and human papillomavirus status in oropharyngeal squamous cell carcinoma
Head Neck. 2016; 38: 529-35

C.S. Schouten, O.S. Hoekstra, C.R. Leemans, J.A. Castelijns, R. de Bree
Response evaluation after chemoradiotherapy for advanced staged oropharyngeal squamous cell carcinoma: a nationwide survey in the Netherlands
Eur Arch Otorhinolaryngol. 2015; 272: 3507-13

C.S. Schouten, P. de Graaf, F.M. Alberts, O.S. Hoekstra, E.F.I. Comans, E. Bloemena, B.I. Witte, E. Sanchez, C.R. Leemans, J.A. Castelijns, R. de Bree
Response evaluation after chemoradiotherapy for advanced nodal disease from head and neck squamous cell carcinoma using diffusion-weighted MRI and ^{18}F-FDG-PET-CT
Oral Oncol. 2015; 51: 541-7

C.S. Schouten, P. de Graaf, E. Bloemena, B.I. Witte, B.J.M. Braakhuis, R.H. Brakenhoff, C.R. Leemans, J.A. Castelijns, R. de Bree
Quantitative diffusion-weighted MRI parameters and human papillomavirus status in oropharyngeal squamous cell carcinoma
AJNR Am J Neuroradiol. 2015; 36: 763-7

C.S. Schouten, R. de Bree, L. van der Putten, D.P. Noij, O.S. Hoekstra, E.F.I. Comans, B.I. Witte, P.A. Doornaert, C.R. Leemans, J.A. Castelijns
Diffusion-weighted EPI- and HASTE-MRI and ^{18}F-FDG-PET-CT early during chemoradiotherapy in advanced head and neck cancer
Quant Imaging Med Surg. 2014; 4: 239-50

C.S. Schouten, R.A. Akrum, L.A. van der Velden
Cystic neck mass. Cervical cleft cyst or lymph node metastasis?
Ned Tijdschr Geneeskd. 2013; 157 (3): A5087

C.S. Schouten, J.A. Castelijns, O.S. Hoekstra, C.R. Leemans, R. de Bree
Response evaluation after chemoradiation for advanced oropharyngeal cancer using PET-CT and MRI: the REACTION study
Ned Tijdschr Oncol. 2012; 9:369-71

ABBREVIATIONS

ADC	Apparent diffusion coefficient
AUC	Area under the curve
BAT	Brown adipose tissue
CI	Confident interval
CRT	Chemoradiotherapy
CT	Computed tomography
CTAC	CT-based attenuation correction
DCE	Dynamic contrast enhanced
DWI	Diffusion-weighted imaging
EPI	Echo-planar imaging
EUA	Examination under general anesthesia
FLT	^{18}F-fluorothymidine
FMISO	^{18}F-fluormisonidazole
FN	False negative
FNAC	Fine needle aspiration cytology
FP	False positive
FWHM	Full width at half maximum
HASTE	Half-fourier acquisition single shot spin-echo
HNSCC	Head and neck squamous cell carcinoma
HPV	Human papillomavirus
MATV	Metabolic active tumor volume
MBq	Mega Becquerel
MET	Methyl-[1-11C]-L-methionine
MRI	Magnetic resonance imaging
NPV	Negative predictive value
OPSCC	Oropharyngeal squamous cell carcinoma
OSEM	Ordered subset expectation maximisation
PCR	Polymerase chain reaction
PEG	Percutaneous endoscopic gastrostomy
PET	Positron emission tomography

PPV	Positive predictive value
PROPELLER	Periodically rotated overlapping parallel lines with enhanced reconstruction
PVE	Partial volume effect
QALY	Quality-adjusted life-year
ROI	Region of interest
SD	Standard deviation
SNR	Signal-to-noise-ratio
STIR	Short T1 iversion-recovery
SUV	Standardized uptake value
TE	Echo time
TLG	Total lesion glycolysis
TN	True negative
TNM	Tumor node metastasis
TP	True positive
TR	Repetition time
T1WI	T1-weighted
VOI	Volume of interest
^{18}F-FDG	^{18}Fluoro-2-deoxyglucose
^{18}F-FAZA	^{18}F-fluorazomycin arabinoside

NEDERLANDSE AFKORTINGEN

OON	Onderzoek onder narcose
NPV	Negatief voorspellende waarde
PPV	Positief voorspellende waarde