## HEAD AND NECK RADIOLOGY

# $^{18}$F-Fluorodeoxyglucose-PET/CT to evaluate tumor, nodal disease, and gross tumor volume of oropharyngeal and oral cavity cancer: comparison with MR imaging and validation with surgical specimen

Oliver Seitz · Nicole Chambron-Pinho ·
Markus Middendorp · Rober Sader · Martin Mack ·
Thomas J. Vogl · Sotirios Bisdas

**Abstract**

*Introduction* The purpose of this paper is to evaluate the impact of adding combined $^{18}$F-PET/CT to MRI for T and N staging of the oral and oropharyngeal cancer and calculation of the gross tumor volume (GTV) having histopathology as reference standard.

*Methods* PET/CT and MRI were performed in 66 patients with suspected oral and oropharyngeal cancer (41 primary tumors/25 recurrent tumors) and nodal disease (114 nodes). Statistical analysis included the McNemar test, sensitivity, specificity for the diagnostic modalities as well as regression analysis, and Bland–Altman graphs for calculated tumor volumes.

*Results* There was no statistically significant difference between the two modalities compared to pathological findings regarding detection of disease ($P \geq 0.72$). The sensitivity/specificity for tumor detection were 100/80% and 96.72/60% for MRI and PET/CT, respectively. The sensitivity/specificity for nodal metastases were 88.46/75% and 83.81/73.91% for MRI and PET/CT, respectively. In 18% of cases, the MRI-based T staging resulted in an overestimation of the pathologic tumor stage. The corresponding rate for PET/CT was 22%. Regarding the treated necks, both modalities showed 100% sensitivity for detection of the recurrent lesions. In necks with histologically N0 staging, MRI and PET/CT gave 22% and 26% false positive findings, respectively. The mean tumor volume in the pathologic specimen was 16.6±18.6 ml, the mean volume derived by the MR imaging was 17.6± 19.1 ml while the estimated by PET/CT volume was 18.8± 18.1 ml ($P \leq 0.007$ between the three methods). The Bland–Altman analysis showed a better agreement between PET/CT and MRI.

*Conclusion* The diagnostic performance of FDG-PET/CT in the local staging of oral cancer is not superior to MRI.

**Keywords** PET/CT · MRI · Oral cancer · Oropharynx · Gross tumor volume

Oliver Seitz and Sotirios Bisdas have equally contributed to the study design and the preparation of the manuscript.

O. Seitz · N. Chambron-Pinho · R. Sader
Department of Oromaxillofacial Surgery, JW Goethe University,
Frankfurt, Germany

M. Middendorp
Department of Nuclear Medicine, JW Goethe University,
Frankfurt, Germany

M. Mack · T. J. Vogl
Department of Radiology, JW Goethe University,
Frankfurt, Germany

S. Bisdas (✉)
Department of Neuroradiology, Eberhard Karls University,
Hoppe-Seyler-Str. 3,
72076 Tübingen, Germany
e-mail: Sotirios.Bisdas@med.uni-tuebingen.de

## Introduction

The behavior of head and neck cancer reflects a complexity of genetic, biological, histopathological, and clinical heterogeneity [1]. Surgery and combined radio-chemotherapy are recognized as key treatment modalities [2,3] and proper management requires accurate information about the site

and the extent of the tumor and nodal disease. Current non-invasive staging imaging techniques include computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), usually combined with CT (PET/CT). There is evidence that ^{18}F-fluorodeoxyglucose (FDG)-PET or combined PET/CT may be superior to the conventional CT and MRI in the evaluation of the primary and recurrent head and neck disease, especially the nodal one [4–6]. In the oropharynx and oral cavity, the primary concern of the head and neck surgeon as well as of the radio-oncologist is the exact anatomical detail that may optimize the therapy outcome. There is accumulating evidence that PET/CT in this region may be used even in patients with dental artifacts on CT or MR images [7,8] or may enhance the effectiveness of the radiotherapy planning especially by delineating the gross tumor volume (GTV) in intensity-modulated radiation therapy (IMRT) [9–11]. Nevertheless, the accurate delineation of the primary tumor, the detection of infiltration of key neighboring structures, the detection of recurrent disease (in the extent that this is achievable), and the characterization of the dignity of the enlarged nodes are in the majority of the centers facilitated by the MRI.

The aim of this study was to evaluate the impact of adding combined ^{18}F-PET/CT to the standard MR imaging protocol for oral tumor and nodal staging having as reference standard the histopathological grading. As a secondary endpoint, we sought to evaluate the gross tumor volume (GTV) assessment by the two imaging modalities by comparing the results to the histopathological tumor volume calculation.

## Materials and methods

After institutional review board approval, a retrospective study was performed on a group of 66 patients (39 males, 27 females; mean age, 63±14 years; age range, 25–89 years) who fulfilled the following criteria: (a) clinical suspicion of a primary or recurrent carcinoma in the oropharynx and oral cavity; (b) surgical excision of the tumor with histological diagnosis and calculation of the specimen volume; and (c) baseline whole-body ^{18}F-FDG-PET/CT scan followed by head and neck MRI within 5 days and up to 5 days before surgery. Tumor stage was determined according to the TNM classification. All patients completed both exams and received a histological diagnosis which was positive in 61 patients (92.4%) and negative in five patients (7.6%). Squamous cell carcinoma (SCCA) was diagnosed in 55 patients, adenoid cystic carcinoma in five patients, post-therapeutic changes with inflammation in four patients and mandibular osteoradionecrosis in one patient. Primary tumor was diagnosed in 41 patients and locoregional recurrent

tumor was histologically confirmed in 20 patients. The primary tumors were localized in the floor of the mouth (18 patients), in the tongue (16 patients) and in the oropharynx (seven patients). The recurrent tumors were diagnosed in the floor of the mouth (six patients), in the tongue (ten patients), and in the oropharynx (four patients). The distribution of the histologically diagnosed primary T stages was: 12 subjects with T1 tumor, nine patients with T2, 14 patients with T3, and six patients with T4 tumors. The histologically diagnosed recurrent T stages were nine T2-, four T3-, and seven T4-tumors. The histologically determined N stage of the primary tumors was as follows: N0 ($n$=23), N1 ($n$=2), N2a ($n$=5), N2b ($n$=5), N2c ($n$=4), N3 ($n$=2). The N stage of the recurrent tumors were N1 ($n$=5), N2a ($n$=7), N2b ($n$=4), N2c ($n$=4). Eighteen patients with primary tumors (T3 and T4 stage) received neoadjuvant chemoradiation before salvage surgery, while patients who experienced a recurrence had received either radiotherapy (12 patients) or combined chemoradiation (eight patients) in the past (median time after treatment 10 months, range 5–14 months). The patients underwent tumor resection at the floor of the mouth in 15 cases and tongue in 12 cases, combined resection of the floor of the mouth and tongue in 19 cases, mandibulectomy in eight cases, maxillectomy in five cases, combined buccopharyngectomy in five cases, while four subjects underwent an additional tonsillectomy. The performed neck dissection was radical in 19 subjects, modified in ten subjects, unilateral in seven cases.

### Whole-body ^{18}F-FDG-PET/CT technique

After a fasting period of at least 6 h and with the precondition of a blood glucose level below 130 mg/dl patients were intravenously injected 3 MBq [^{18}F]FDG per kilogram body weight (maximum of 370 MBq) and 20 mg furosemide for reduction of radiation exposure to the bladder wall. Patients stayed in a semi-Fowler´s position in a quiet and darkened room for 60 min. [^{18}F] FDG imaging was performed on an integrated PET-CT scanner (Biograph 6 TruePoint PET-CT, Siemens Medical Solutions, Erlangen, Germany). Contrast-enhanced CT scans were initially obtained from the top of the skull to the proximal thighs (120 mAs, 130 kV, 5 mm slice collimation) followed by PET scans in 3-D mode with an acquisition time of 3 min per bed position (axial FOV 16.2 cm). Images were reconstructed by an attenuation-weighted ordered-subsets expectation maximization algorithm (four iterations, eight subsets) and a post-reconstruction smoothing Gaussian filter. For comparison with MRI data, additional image reconstruction was done for the head and neck region (slice thickness of 4 mm, matrix of 128×128).

## Head and neck MRI technique

In all patients, MRI was performed at a 1.5-T MR unit with a 40 mT/m gradient system (Avanto, Siemens Medical Solutions) using a combined head and neck phased-array coil. Axial TIRM- (turbo inversion recovery magnitude), T1- [$T_R$ (repetition time), 467 ms; $T_E$ (echo time), 12 ms], and T2-weighted imaging ($T_R$, 4,140 ms; $T_E$, 79 ms), as well as coronal T1-weighted imaging were acquired before contrast agent administration. After gadolinium (Magnevist®, Bayer Schering Pharma, Berlin, Germany) injection (0.2 mmol/kg; mean dose of the administered gadolinium, 15±1.9 ml; range, 12–19 ml), using a power injector, fat-saturated, flow-rephased axial and coronal T1-weighted images ($T_R$, 600 ms; $T_E$, 17 ms) were acquired. All slices were 4-mm-thick in order to facilitate comparison with the PET/CT sections.

## Data analysis

Two observers, an experienced head-and-neck imaging senior radiologist and a nuclear medicine specialist consensually reviewed tumor and lymph node findings on FDG-PET/CT images. The primary or recurrent tumor was assessed for the presence of invasion in the nasopharynx, nasal cavity, parapharyngeal region, skull base, paranasal sinus, and hypopharynx. The head and neck was assessed for the presence of any node from the retropharyngeal region down to the supraclavicular fossa. Tumor tissue was identified as any voxel in the three-dimensional dataset with counts greater than a fixed threshold fraction of the peak activity in the tumor [12]. The threshold level for tumor characterization was selected as standardized uptake value (SUV) equal to 3.5 [8]. In each of the PET/CT images, the volumes of the abnormal uptake areas over the cutoff point as well as the maximum diameter of the tumor were calculated automatically by the software (TrueD, Siemens Medical Systems). Lymph nodes were not included in the calculation of GTV for simplification of the volumetric analysis. The SUVs were calculated automatically by the computer using the body weight method: SUV = decay corrected tissue activity (kBq ml$^{-1}$)/injected $^{18}$F-FDG dose per body weight (kBq g$^{-1}$). The maximum SUV value (SUV$_{max}$) was used instead of average SUV as this is more accurate and consistent. The SUV$_{max}$ of the tumor lesions were also recorded. The shortest axial diameter and SUV$_{max}$ of a lymph node, recorded at each nodal level (I–V), were chosen in order to facilitate the nodal staging. The size-based SUV$_{max}$ cutoff values were 1.9, 2.5, and 3.0 for lymph nodes <10 mm, 10–15 mm, and >15 mm, respectively [13].

Two senior radiologists, blinded to the PET/CT findings, consensually reviewed tumor and lymph node findings on MR images. The primary or recurrent tumor was assessed for its dimensions and infiltration of neighboring structures. For MR imaging, GTVs were initially determined from the post-contrast T1-weighted images, but comparison with the pre-contrast T1-weighted as well as T2-weighted images was performed when necessary. Criteria for malignancy were unilateral change in anatomy when compared with the normal contralateral side, the mass effect, nodular or infiltrative abnormal tissue, fat replacement, and pathologic enhancement in the post-contrast series. The tumor volume was calculated by the summation of areas technique based on the ROI outline using the provided software features. By MRI, nodes were considered to be metastatic if they had (a) a shortest axial diameter of ≥10 mm in the jugulodigastric and subdigastric region, >5 mm in the retropharyngeal region and >10 mm in all other regions of the neck; (b) groups of three or more lymph nodes in the lymph node drainage regions of the tumor that were borderline in size; (c) all nodes, irrespective of size, that showed irregular enhancement on MRI and that were surrounded by a rim of enhancing viable tumor were considered metastatic; (d) any nodes with necrosis or extracapsular spread irrespective of size [14]. The lymph nodes in MRI and PET/CT were assessed individually and the comparison with histology was done level by level.

For the surgical specimens, GTVs were determined from visualization of gross tumor infiltration in the fresh specimens which were collected en bloc in the operation room and were placed in a polystyrene cast marked longitudinally in the three dimensions and filled with gelatin solution in order to avoid shrinkage of the specimen (i.e., using formalin solution). The specimen was cut in the transverse plane to generate parallel, contiguous slices. The delineation was performed prospectively without any knowledge of the imaging data.

## Statistical analysis

The values are presented as mean±standard deviation (SD). Data normality was tested by Kolmogorov–Smirnov test. Statistical differences for the tumor identification rates among the imaging modalities were analyzed by the McNemar test. Tumor volume as estimated by all three methods was not normally distributed ($P \leq 0.018$). To facilitate a comparison between the pathologic volume and the MRI as well as PET/CT estimated volumes, a linear regression analysis as well as Bland–Altman analysis after log transformation (to warrant a normal distribution) was performed. Any differences in the tumor volume estimations between the three methods were determined by Wilcoxon rank sum test and by paired-samples $t$ test for their logarithmically transformed values. SPSS software (release 15.0.1; SPSS Inc., Chicago, Ill.) and MedCalc

software (version 10.0.0.0, Mariakerke, Belgium) were used for the statistical analyses. A two-tailed *P* value of less than 0.05 was considered significant.

## Results

The McNemar test revealed no statistically significant difference between the two modalities compared to the pathological findings regarding the detection of the disease ($P \geq 0.72$). Figure 1 shows the MRI and PET/CT findings in two patients.

The sensitivity, specificity, positive and negative likelihood ratio, positive and negative predictive value for both modalities in detecting the tumor ($n=66$ tumor sites) and the nodal metastases ($n=114$ examined nodes) are summarized in Tables 1 and 2. Compared to the histological results, MRI did not miss the diagnosis in any tumor. Falsely, based on MRI, one patient was diagnosed with recurrence; however, it proved to be only therapy-induced tissue changes. In four cases, the negative MRI findings were in concordance with the histological findings. Regarding PET/CT in two cases with T1 tumors, PET/CT missed the diagnosis (in one case the tumor was obscured by dental artifacts), while according to PET/CT findings

two cases were diagnosed as recurrent tumors but the histology was negative. In three cases the negative PET/CT findings were in concordance with the findings in the pathology specimen. In 18% of the cases, the MRI-based T staging resulted in an overestimation of the actual tumor stage as defined by the pathologic specimen; specifically, T3 tumors were incorrectly classified as T4 tumors and one T2 tumor was misclassified as T3. In 8% of the cases, MRI underestimated the actual T stage including three T4 tumors and two T3 tumors. The corresponding rates for PET/CT were 22% and 12%, respectively. Regarding the nodal stage, MRI led to an overestimation in 26% of the cases and in an underestimation in 12% of the patients. On the contrary, an overestimation of the N stage as evaluated by PET/CT was observed in 14% of the patients and an underestimation in 17% of the cases. None of the above-mentioned differences between the two imaging modalities were statistically significant ($P>0.1$). Bone invasion was histologically confirmed in eight cases and was correctly diagnosed by MRI in six patients while PET/CT could diagnose it only in three cases.

The T and N staging was further analyzed in two important categories of patients in our population, namely the patients with treated neck and suspected recurrence and the patients with no histologically proven nodal disease (23



**Fig. 1 a–d** MR and $^{18}$F-FDG-PET/CT imaging in a 23-year-old female patient with T1N0 squamous cell carcinoma in the left tongue (**a–b**) and in a 46-year-old male patient with squamous cell carcinoma in the maxilla (**c–d**). Post-contrast T1-weighted fat-saturated MR image (**a**) shows the pathologically enhancing tumor in the left border of the tongue (*arrow*). The corresponding PET/CT image (**b**) shows no significantly hypermetabolic focus on this site ($SUV_{max}<3.5$) and the lesion was missed at the initial diagnosis. Post-contrast T1-weighted fat-saturated MR image (**c**) shows a mass (*arrows*) diffusely infiltrating the maxilla and the adjacent soft-tissue. The PET/CT (**d**) shows increased FDG-metabolism in this site (*arrows*) but the lesion is overspilled by the dental artifacts
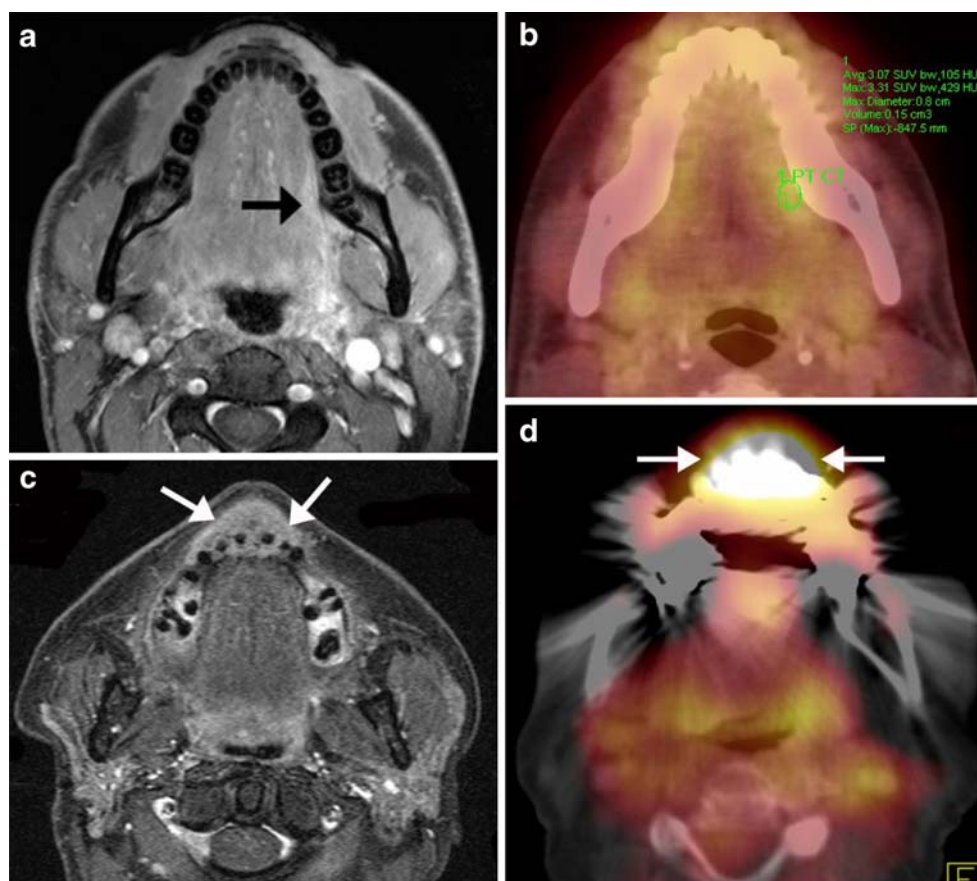
**Table 1** Sensitivity, specificity, positive and negative likelihood ratios, and positive and negative predictive values of MRI and FDG-PET/CT exams for tumors.

| | Pooled tumors | | Primary tumors | | Recurrent tumors | |
|---|---|---|---|---|---|---|
| | MRI | PET/CT | MRI | PET/CT | MRI | PET/CT |
| Sensitivity[a] | 100% | 96.72% | 100% | 95.12% | 100% | 100% |
| Specificity[a] | 80% | 60% | | | 80% | 60% |
| Positive likelihood ratio[c] | 5 | 2.42 | | | 5 | 2.5 |
| Negative likelihood ratio[d] | 0 | 0.05 | | | 0 | 0 |
| Positive predictive value[e] | 98.39% | 96.72% | 100% | 100% | 95.24% | 90.91% |
| Negative predictive value[f] | 100% | 60% | | 0 | 100% | 100% |

True positive (TP); false positive (FP); true negative (TN); false negative (FN)

MRI in suspected primary tumors ($n=41$): TP, 41; FP, 0; FN, 0; TN, 0

MRI in suspected recurrent tumors ($n=25$): TP, 20; FP, 1; FN, 0; TN, 4

PET/CT in suspected primary tumors ($n=41$): TP, 39; FP, 0; FN, 2; TN, 0

MRI in suspected recurrent tumors ($n=25$): TP, 20; FP, 2; FN, 0; TN, 3

[a] Sensitivity: TP/(TP+FN)

[b] Specificity: TN/(FP+TN)

[c] Positive likelihood ratio=sensitivity/(1−specificity)

[d] Negative likelihood ratio=(1−sensitivity)/specificity

[e] Positive predictive value=TP/(TP+FP)

[f] Negative predictive value=TN/(FN+TN)

cases). Concerning the treated necks, both modalities showed 100% sensitivity for detection of the recurrent lesions. MRI misclassified a case with post-therapeutic changes as recurrent lesion while PET/CT misclassified two cases as recurrent tumors. In 23 necks with histologically N0 staging, five cases (22%) were misdiagnosed by the MRI based on the bright T2-signal and the contrast agent enhancement of the nodes which simulated a metastasis: in four cases, the surgeons could avoid a bilateral neck dissection and in one case a unilateral neck dissection.

**Table 2** Sensitivity, specificity, positive and negative likelihood ratios, and positive and negative predictive values of MRI and FDG-PET/CT imaging for nodal disease.

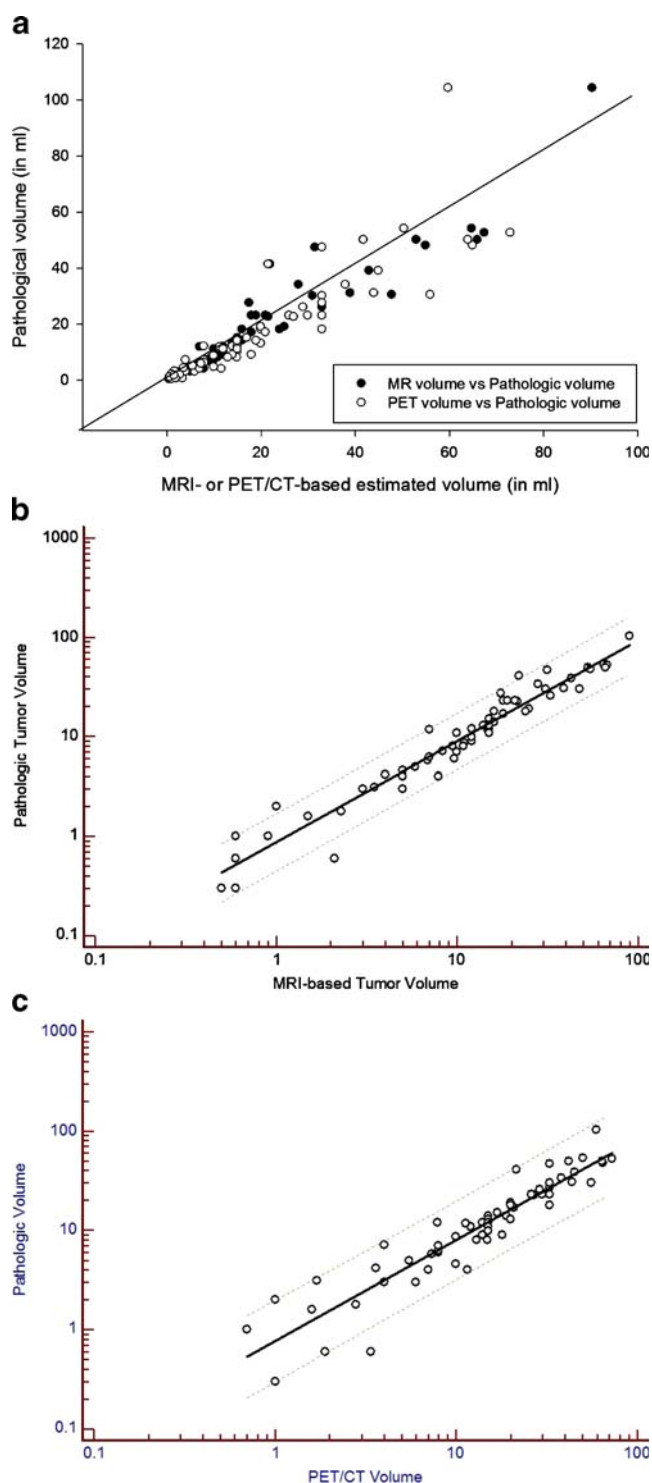| | Nodal disease | |
|---|---|---|
| | MRI | PET/CT |
| Sensitivity | 88.46% | 83.81% |
| Specificity | 75% | 73.91% |
| Positive likelihood ratio | 3.54 | 3.21 |
| Negative likelihood ratio | 0.15 | 0.22 |
| Positive predictive value | 93.88% | 93.62% |
| Negative predictive value | 60% | 50% |

True positive (TP); false positive (FP); true negative (TN); false negative (FN)

MRI in suspected nodes ($n=114$): TP, 92; FP, 6; FN, 12; TN, 18

PET/CT in suspected nodes ($n=114$): TP, 88; FP, 6; FN, 17; TN, 17

Based on the PET/CT findings, six cases with nodes <10 mm (26%) were misclassified as positive due to the increased SUV (>1.9), three patients received a bilateral neck dissection and three patients a unilateral neck dissection.

Dental artifacts, which slightly compromised the GTV in some cases on MRI, were present in 20 patients (30.3%). The mean GTV in the pathologic specimen was 16.6± 18.6 ml, the mean volume derived by the MR imaging was 17.6±19.1 ml, while the estimated by PET/CT volume was 18.8±18.1 ml. Graph analysis showed the larger the tumor volume, the more pronounced the difference with the histopathologic specimen (Fig. 2). The mean $SUV_{max}$ of the tumors was 9.4±5.2. The regression analysis of the log-transformed tumor volume values showed: $log_{pathologic\ volume}=-0.06+1.01\ log_{MRI\ volume}$ ($R^2=0.94$) while for PET/CT was $log_{pathologic\ volume}=-0.11+1.01\ log_{PET/CT\ volume}$ ($R^2=0.86$; Fig. 2). The Bland–Altman analysis (Fig. 3) showed that the arithmetic mean of the difference between the pathologic and the MRI tumor volume estimation was −0.05±0.14 (95% CI −0.08 to −0.009). The mean difference between the pathologic and the PET/CT-based tumor volume calculation was −0.1±0.19 (95% CI −0.14 to −0.04). Wilcoxon rank sum test showed significant differences between the volume estimations between the three methods as follows: pathologic volume–MRI volume, $P=0.007$; pathologic volume–PET/CT volume, $P<0.0001$; MRI volume–PET/CT volume, $P=0.002$. Generally, MRI and PET/CT overestimated the
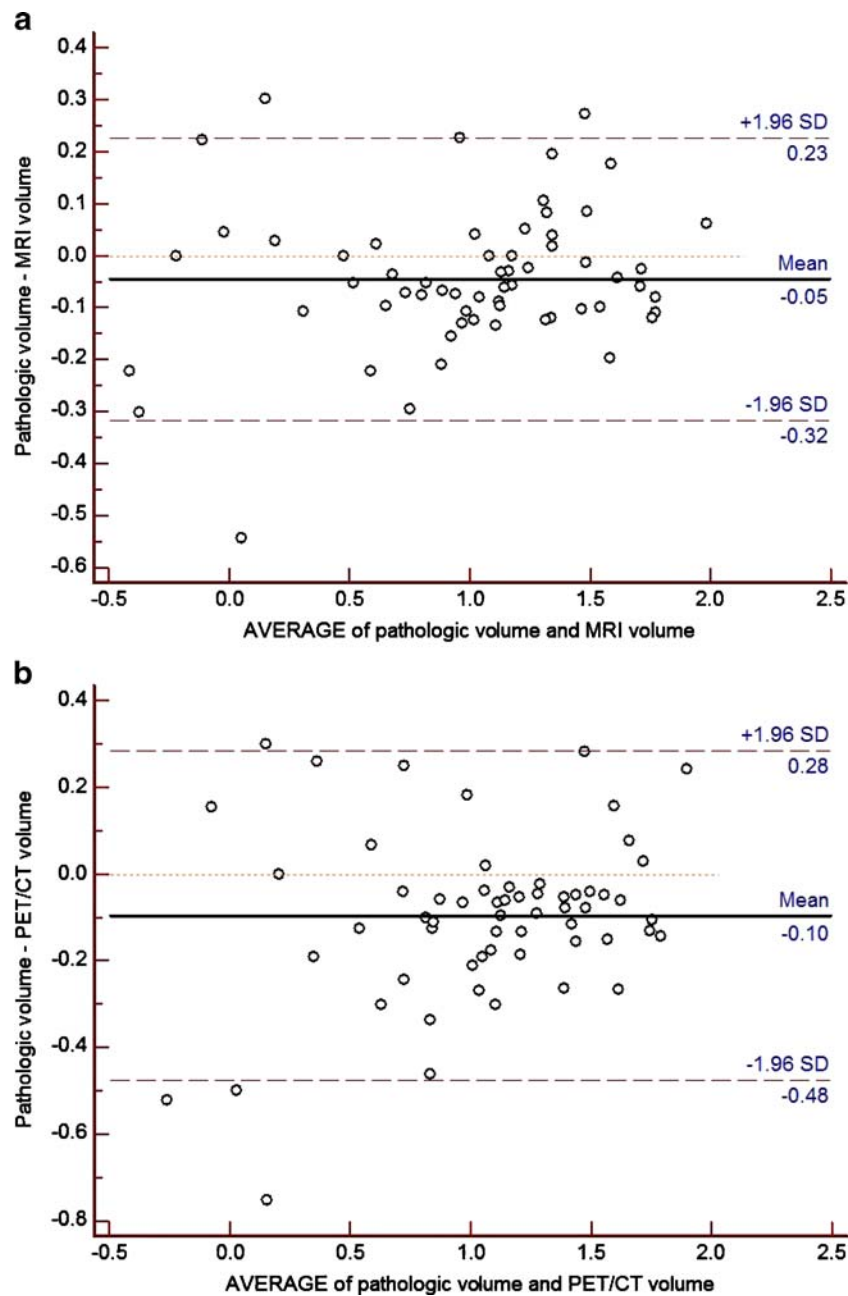
## Discussion

Precise information about the extent of tumors and nodal disease is critical for treatment decision in patients with oral carcinoma. Traditionally, conventional CT and MRI is the workhorse for the evaluation of nodal and tumor extent involving the surrounding tissues such as the base of the tongue, mandible, pharyngeal walls, and nasopharynx. However, X-ray-beam-hardening artifacts and low resolution of the soft-tissue structures limit the accuracy of CT scans compared to MRI. The latter may be, in turn, limited by ferromagnetic dental fillings and moving artifacts due to longer acquisition time. With the advent of the combined PET/CT imaging, the combination of functional and anatomical information proved to be very sensitive in detecting primary and recurrent malignant tumor in the baseline setting as well as in the post-treatment regimen in the head and neck [5,15]. Besides T stage and N stage, of major importance for the radiotherapy planning and monitoring is also the GTV, which may be calculated in the fused PET/CT images [16]. To our best knowledge, there is limited evidence of the impact of different imaging modalities on the pathologically calculated tumor volume [8,17], on the histological T staging [6] and nodal staging of oral cancer [18–20], the latter with partly conflicting results. Moreover, there is the first study to compare combined PET/CT and MRI under these aspects in a large patient population.

In this study, the positive predictive value for PET/CT (96.7%) was slightly higher than the one reported (92.7%) by Fleming et al. in an almost same number of patients with head and neck cancer [15]. Though the hybrid imaging results in improved anatomic localization and accuracy compared with PET alone [21], the sensitivity and specificity of the PET/CT were worse than those of MRI. Our results are not in agreement with those of a recent study of Roh et al. [18] in a comparable patient population presumably due to the other tumor sites (apart from oropharynx and oral cavity) included in their study. Similar to our work, PET/CT in the study of Roh et al. [18] as well as in the study of Pentenero et al. [19] failed to detect the primary tumor when the latter was superficial or located in tonsils. These findings highlight the intrinsic weakness of

tumor volume compared to the pathologic specimens. Moreover, PET/CT delivered higher volume results than MRI measurements. After logarithmic transformation, paired-samples *t* test showed slightly lower significant differences: pathologic volume–MRI volume, $P=0.013$; pathologic volume–PET/CT volume, $P<0.0001$, MRI volume–PET/CT volume, $P=0.012$.

**Fig. 3 a–b** Bland–Altman analysis of the normally distributed logarithmically transformed volume values shows narrower confidence intervals (*dashed lines*) for the mean difference in MRI (**a**) compared to PET/CT (**b**). The values of the mean difference in graph **b** appear also scattered (around the line of mean difference) with four outliers indicating low agreement between the pathologic volume values and the PET/CT-based calculations

PET/CT to detect small or superficial lesions due to inadequate spatial resolution or dental artifacts. A further reason for this shortcoming may be the scattering of the focal FDG uptake in tumor by other neighboring structures, while the normal FDG uptake may mask small tumors with moderate glucose metabolism. Such cases are common in the oropharyngeal and oral cavity cancer where tumors often rise in the vicinity of teeth amalgams and tonsils, and occasionally present superficially with complex anatomical relationships, which are more pronounced in the post-treatment regimen. Under these circumstances, the utility of PET/CT in T1 tumors is limited only to the detection of pathological nodes, which, however, due to their expected small size, may not be correctly evaluated for micro-metastases or low tumor load. Thus, PET/CT may not abrogate the need for MR imaging in this patient group.

Noteworthy in the performed comparison between the two modalities was the negative predictive value of 100%

in the case of MRI compared to the 60% of PET/CT. Obviously, performing MRI in these patients gives confidence that a false negative result is unlikely. On the other side, in 40% of the cases evaluated only with PET/CT a negative result needs to be re-evaluated by conventional physical examination or follow-up imaging. This implies additional healthcare costs and may be due to the limited diagnostic accuracy of CT, low metabolic rates in tumors, tumor size below the available resolution, and tumor obscuration by artifacts or adjacent highly FDG-uptaking structures. Nevertheless, the positive predictive value (PPV) of MRI and PET/CT are satisfactorily high and comparable with the results of Fleming et al. (92.7% PPV for PET/CT) [15] and may indicate a prompt therapy without additional diagnostic procedures.

Though the critical determinant of an imaging modality is to detect the presence or absence of tumor, the correct T staging is also a prerequisite. Interestingly, in about 30% of the cases the T stage was misclassified by the PET/CT and MRI in a comparable fashion. This has implications for the extent of the operation or the radiotherapy and was mostly due to the well-known limitations of both modalities like the diagnostic uncertainty regarding the bone infiltration [19], the unspecific findings of the peritumoral inflammation and granulation/fibrosis as well as extrinsic factors that hamper the diagnosis like the beam-hardening or the movement artifacts. The quote of misclassified tumors was lower (not significantly) in MRI, presumably due to the better spatial resolution in all planes (compared with the limited resolution of about 5 mm in combined PET/CT) and the possibility of performing fat-saturation which leads to an enhancement of the diagnostic accuracy.

Besides the patients with primary tumors in the oropharynx and oral cavity, the patients in the post-treatment surveys have benefited in the last years from the combined PET/CT [5] as early diagnosis of recurrent cancer in this region and a subsequent salvage surgery implies a substantially better prognosis than a delayed identification [22]. MRI and PET/CT proved to be highly sensitive (100% sensitivity) and specific (98% and 95%, respectively) in the patients with suspected recurrent tumor in our population. These results are in concordance or even better with the sensitivity/specificity found in other studies of FDG-PET/CT in the treated head and neck after at least 8 weeks of (chemo)radiotherapy [23–25]. Undoubtedly, the differentiation between post-therapeutic changes, including (non)infectious inflammation, and tumor recurrence is difficult by FDG-PET/CT scan and MRI in many cases and is the reason for the few missed cases in our study. Both modalities showed an equal effectiveness, which in MRI has also to do with the experience of the reader, while the application of diffusion- or perfusion-weighted as well

as proton spectroscopy MR-sequences may enhance in the future the ability to detect persistent or recurrent disease after (chemo)radiotherapy [26,27].

Another determinant of the diagnostic usefulness of an imaging modality in the head and neck is its ability to detect the presence or absence of metastatic nodal disease, particularly where this is not evident by physical examination. FDG-PET/CT might detect small metastases in slightly enlarged nodes; however, it has resolution limitations and is not likely to be able to detect small metastatic nodes (i.e., <4 mm) and micrometastases in normal-sized nodes. An inflammation in the reactive nodes may be also misinterpreted false-negatively as metastatic deposit. Second, nodal necrosis may cause false negative findings on PET because of the low glycolytic activity of the necrotic material. Third, false results may be anchored in the art of PET/CT evaluation (visual or semi-quantitative). A recent study showed that ROC curves of SUV of 3.25 yielded satisfactory sensitivity and specificity but did not surpass the physician's performance with visual interpretation [28]. This finding is likely related to several factors: SUV is dependent on the physician and experienced nuclear medicine physicians may calculate SUV only in areas that are considered suspicious or equivocal by visual analysis; SUV may be affected by many different factors, including body fat, time of measurement after injection, the choice of mean or maximum values and the size of ROI [29]. Regarding MRI, a level-specific evaluation concerning the size criterion is widely accepted while other criteria, like the central necrosis, may be difficult to evaluate in small nodes despite the high soft-tissue resolution. The reported sensitivity of MRI for lymph node metastases ranges from 64% to 92%, whereas the reported specificity ranges from 40% to 81% [30]. In the present study, the sensitivity was 85% and the specificity 76% which are in concordance with the majority of the published studies [30]. The reported sensitivity and negative likelihood ratio for PET/CT in our study were in almost perfect agreement with the overall sensitivity (79%) and negative likelihood ratio (0.24) in cN0 and cN+ patients across 32 studies with FDG-PET surveys [31]. The specificity of PET/CT in detecting the pathological lymph nodes was fairly low in the present study unlikely to the recent study of Ng et al. [6]. According to the latter study, in 124 patients with SCCA of the oral cavity the specificity of PET alone was 93% and together with CT/MRI was 94.5%. Our lower-than-expected result may be due to the smaller patient population and the different way of evaluating the lymph nodes (visual vs. semi-quantitative evaluation). Notably, the sensitivity ranged from 47–84% and the specificity from 88–100% in the study of Murakami et al. whose criteria for nodal staging in PET/CT were applied in our study. Our sensitivity rates are not contradicting their results, whereas

the discrepancy in specificity may arise from the higher proportion on negative necks (83%) compared to our study (34.8%) and the smaller population. Our results are closer to the ones published by Stuckensen et al. who reported 82% specificity in 106 patients with SCCA of the oral cavity [32] as well as to the ones of Pentereno et al. in 23 patients with oral SCCA (76% specificity)[19]. The reason for low specificity of PET/CT in nodal staging has been among others attributed to the almost equal tracer uptake in reactive nodes as that in metastatic lesions [33]. This fact was also influencing our study as our SUV measurements were based on the relatively low threshold of 1.9 for nodes smaller than 10 mm. Moreover, the lack of CT criteria for defining nodal metastases (except of the nodal size), due to the non-tissue-specific acquisition parameters, may be a further reason for the low specificity. This methodology introduced a bias for the benefit of MRI where the readers had a wide choice of evaluating criteria for the nodal staging.

A clinically relevant result in our work was the impact of false positive findings in both modalities which led to unnecessary neck dissections in necks without nodal disease. The rate of incorrectly performed dissections based on the imaging findings was slightly higher in PET/CT than in MRI (26% and 22%, respectively). Thus, the routine clinical application of $^{18}$F-FDG-PET in the pre-treatment evaluation of the lymph node status in patients with oral cancer is not supported as other imaging methods appear to have similarly limited diagnostic performance [31]. Of course, PET/CT imaging in these patients may be performed for a variety of additional reasons besides N staging (e.g., baseline for future follow-up or identification of distant metastases). Anyhow, the decision of neck dissection in the routine clinical practice is more complicated as it may be coupled with the local extent of the tumor, the potential complications, and the cost of the examination. Therefore, the choice of PET/CT or MRI may be made on a case-by-case basis.

The last part of the present work concerned the calculation of the GTV, which has an increased impact on the radiotherapy-treatment planning. The tumor areas were differently contoured on MR and PET/CT images and, subsequently, showed significant differences with the pathologic tumor estimations. Consistently, the MRI-based estimations were closer to the histopathologic ones which resulted in better fitting of the regression analysis and lower significant differences compared to the PET/CT-based estimations. The overestimation of the tumor tissue was a key finding and is in concordance with similar studies in patients with laryngeal or hypopharyngeal tumors [17]. The disagreement between the three methods may be caused by the different ways of volume calculation; whereas in PET/CT, the applied method is three-dimensional and semi-

automated (minimizing any inter-observer bias); the MRI- and histopathology-based calculations are manual and have intrinsic limitations. On the other hand, PET/CT-based delineation suffers from inaccuracies due to pre-defined thresholds, low spatial resolution, and heterogeneous distribution of radioactivity in complex-shaped tumors or recurrent lesions [16]. Although it is reported that dental artifacts may not affect the GTV calculation [8], lesions in our study were partly obscured and the only choice in this case was the evaluation of the non-corrected PET images with its subsequent inherent limitations. It has to be proven whether new technologies in PET detectors, advanced segmentation algorithms, and combined PET/MR imaging may correct such overestimations. At first sight, this overestimation may add confidence in the security margins of the radiation-treatment planning; however, the effect on the dose distribution after radiation therapy is unknown. Furthermore, the inherent shortcoming of PET/CT to detect perineural malignancy, which may alter the radiation planning, is crucial for the treatment. On the other hand, the incorporation of functional MRI data to the morphological images may strengthen the role of MRI in the radiation-treatment planning.

Nevertheless, the present study has certain limitations, which have to been taken into account. First, the design of the study is retrospective and the evaluation as well as the discussion of the results is mostly based on the total number of the tumors and metastatic nodes, which include primary and recurrent tumors altogether. Second, the patient population included, in a small percentage (7%), adenoid cystic carcinoma, which may influence the SUV findings.

Our study sheds light on the role of FDG-PET/CT and MRI in three aspects of the oral cancer evaluation in a large patient population: T staging, N staging, and GTV. The role of FDG-PET/CT in the local staging of the oral cancer seems not significantly better than that of MRI. Both methods suffer from low specificity, whereas their sensitivity, especially for the recurrent disease, is very good. The GTV overestimation is also a significant drawback of both modalities; however, PET/CT is enhanced by the functional information that offers to the radiation oncologists and thus, it has gained popularity among them. Notwithstanding the presented results in T and N staging, the utilization of PET-CT in the clinical routine is justified by its ability to detect distant metastasis. In conclusion, this work suggests the use of either PET/CT or MRI in an individual-based manner acknowledging the shortcomings and benefits of both methods.

# References

1. Haddad RI, Shin DM (2008) Recent advances in head and neck cancer. N Engl J Med 359:1143–1154
2. Pignon JP, Bourhis J, Domenge C et al (2000) 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 355:949–955
3. Bourhis J, Calais G, Lapeyre M et al (2004) Concomitant radiochemotherapy or accelerated radiotherapy: analysis of two randomized trials of the French head and neck cancer group (GORTEC). Semin Oncol 31:822–826
4. Kim MR, Roh JL, Kim JS et al (2007) Utility of 18F-fluorodeoxyglucose positron emission tomography in the preoperative staging of squamous cell carcinoma of the oropharynx. Eur J Surg Oncol 33:633–638
5. Abgral R, Querellou S, Potard G et al (2009) Does 18F-FDG PET/CT improve the detection of posttreatment recurrence of head and neck squamous cell carcinoma in patients negative for disease on clinical follow-up? J Nucl Med 50:24–29
6. Ng SH, Yen TC, Liao CT et al (2005) 18F-FDG PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of 124 patients with histologic correlation. J Nucl Med 46:1136–1143
7. Nahmias C, Lemmens C, Faul D et al (2008) Does reducing CT artifacts from dental implants influence the PET interpretation in PET/CT studies of oral cancer and head and neck cancer? J Nucl Med 49:1047–1052
8. Baek CH, Chung MK, Son YI et al (2008) Tumor volume assessment by 18F-FDG PET/CT in patients with oral cavity cancer with dental artifacts on CT or MR images. J Nucl Med 49:1422–1428
9. Newbold KL, Partridge M, Cook G et al (2008) Evaluation of the role of 18FDG-PET/CT in radiotherapy target definition in patients with head and neck cancer. Acta Oncol 47:1229–1236
10. Vernon MR, Maheshwari M, Schultz CJ et al (2008) Clinical outcomes of patients receiving integrated PET/CT-guided radiotherapy for head and neck carcinoma. Int J Radiat Oncol Biol Phys 70:678–684
11. Lee N, Xia P, Fischbein NJ et al (2003) Intensity-modulated radiation therapy for head-and-neck cancer: the UCSF experience focusing on target volume delineation. Int J Radiat Oncol Biol Phys 57:49–60
12. Nestle U, Kremp S, Schaefer-Schuler A et al (2005) Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-Small cell lung cancer. J Nucl Med 46:1342–1348
13. Murakami R, Uozumi H, Hirai T et al (2007) Impact of FDG-PET/CT imaging on nodal staging for head-and-neck squamous cell carcinoma. Int J Radiat Oncol Biol Phys 68:377–382
14. Vogl T, Bisdas S (2007) Lymph node staging. Top Magn Reson Imaging 18:303–316
15. Fleming AJ Jr, Smith SP Jr, Paul CM et al (2007) Impact of [18F]-2-fluorodeoxyglucose-positron emission tomography/computed tomography on previously untreated head and neck cancer patients. Laryngoscope 117:1173–1179
16. Greco C, Nehmeh SA, Schoder H et al (2008) Evaluation of different methods of 18F-FDG-PET target volume delineation in the radiotherapy of head and neck cancer. Am J Clin Oncol 31:439–445
17. Daisne JF, Duprez T, Weynand B et al (2004) Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 233:93–100
18. Roh JL, Yeo NK, Kim JS et al (2007) Utility of 2-[18F] fluoro-2-deoxy-D-glucose positron emission tomography and positron emission tomography/computed tomography imaging in the preoperative staging of head and neck squamous cell carcinoma. Oral Oncol 43:887–893
19. Pentenero M, Cistaro A, Brusa M et al (2008) Accuracy of 18F-FDG-PET/CT for staging of oral squamous cell carcinoma. Head Neck 30:1488–1496
20. Yoon DY, Hwang HS, Chang SK et al (2009) CT, MR, US, 18F-FDG PET/CT, and their combined use for the assessment of cervical lymph node metastases in squamous cell carcinoma of the head and neck. Eur Radiol 19:634–642
21. Schoder H, Yeung HW, Gonen M et al (2004) Head and neck cancer: clinical usefulness and accuracy of PET/CT image fusion. Radiology 231:65–72
22. Goodwin WJ Jr (2000) Salvage surgery for patients with recurrent squamous cell carcinoma of the upper aerodigestive tract: when do the ends justify the means? Laryngoscope 110:1–18
23. Ong SC, Schoder H, Lee NY et al (2008) Clinical utility of 18F-FDG PET/CT in assessing the neck after concurrent chemo-radiotherapy for locoregional advanced head and neck cancer. J Nucl Med 49:532–540
24. Andrade RS, Heron DE, Degirmenci B et al (2006) Posttreatment assessment of response using FDG-PET/CT for patients treated with definitive radiation therapy for head and neck cancers. Int J Radiat Oncol Biol Phys 65:1315–1322
25. Kubota K, Yokoyama J, Yamaguchi K et al (2004) FDG-PET delayed imaging for the detection of head and neck cancer recurrence after radio-chemotherapy: comparison with MRI/CT. Eur J Nucl Med Mol Imaging 31:590–595
26. Zbaren P, Weidner S, Thoeny HC (2008) Laryngeal and hypopharyngeal carcinomas after (chemo)radiotherapy: a diagnostic dilemma. Curr Opin Otolaryngol Head Neck Surg 16:147–153
27. Bisdas S, Baghi M, Huebner F et al (2007) In vivo proton MR spectroscopy of primary tumours, nodal and recurrent disease of the extracranial head and neck. Eur Radiol 17:251–257
28. Fakhry N, Lussato D, Jacob T et al (2007) Comparison between PET and PET/CT in recurrent head and neck cancer and clinical implications. Eur Arch Otorhinolaryngol 264:531–538
29. Wong RJ, Lin DT, Schoder H et al (2002) Diagnostic and prognostic value of [(18)F]fluorodeoxyglucose positron emission tomography for recurrent head and neck squamous cell carcinoma. J Clin Oncol 20:4199–4208
30. de Bondt RB, Nelemans PJ, Hofman PA et al (2007) Detection of lymph node metastases in head and neck cancer: a meta-analysis comparing US, USgFNAC, CT and MR imaging. Eur J Radiol 64:266–272
31. Kyzas PA, Evangelou E, Denaxa-Kyza D et al (2008) 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. J Natl Cancer Inst 100:712–720
32. Stuckensen T, Kovacs AF, Adams S et al (2000) Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J Craniomaxillofac Surg 28:319–324
33. Schoder H, Carlson DL, Kraus DH et al (2006) 18F-FDG PET/CT for detecting nodal metastases in patients with oral cancer staged N0 by clinical examination and CT/MRI. J Nucl Med 47:755–762