

Published in final edited form as:

Radiother Oncol. 2011 December ; 101(3): 356–361. doi:10.1016/j.radonc.2011.05.040.

Correlation between metabolic tumor volume and pathologic tumor volume in squamous cell carcinoma of the oral cavity

James D. Murphy, M.D.¹, Karen M. Chisholm, M.D., Ph.D.², Megan E. Daly, M.D.¹, Ellen A. Wiegner, M.D.¹, Daniel Truong¹, Andre Iagaru, M.D.³, Peter G. Maxim, Ph.D.¹, Billy W. Loo Jr, M.D., Ph.D.¹, Edward E. Graves, Ph.D.¹, Michael J. Kaplan, M.D.⁴, Christina Kong, M.D.², and Quynh-Thu Le, M.D.¹

¹Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA

²Department of Pathology, Stanford University School of Medicine, Stanford, CA

³Department of Nuclear Medicine, Stanford University School of Medicine, Stanford, CA

⁴Department of Otolaryngology - Head and Neck Surgery

Abstract

Purpose—To explore the relationship between pathologic tumor volume and volume estimated from different tumor segmentation techniques on ¹⁸F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in oral cavity cancer.

Materials and Methods—Twenty-three patients with squamous cell carcinoma of the oral tongue had PET-CT scans before definitive surgery. Pathologic tumor volume was estimated from surgical specimens. Metabolic tumor volume (MTV) was defined from PET-CT scans as the volume of tumor above a given SUV threshold. Multiple SUV thresholds were explored including absolute SUV thresholds, relative SUV thresholds, and gradient-based techniques.

Results—Multiple MTV's were associated with pathologic tumor volume; however the correlation was poor (R^2 range 0.29–0.58). The ideal SUV threshold, defined as the SUV that generates an MTV equal to pathologic tumor volume, was independently associated with maximum SUV ($p=0.0005$) and tumor grade ($p=0.024$). MTV defined as a function of maximum SUV and tumor grade improved the prediction of pathologic tumor volume ($R^2 = 0.63$).

Conclusions—Common SUV thresholds fail to predict pathologic tumor volume in head and neck cancer. The optimal technique that allows for integration of PET-CT with radiation treatment planning remains to be defined. Future investigation should incorporate biomarkers such as tumor grade into definitions of MTV.

Keywords

head-and-neck cancer; positron emission tomography; metabolic tumor volume

Introduction

¹⁸F-fluorodeoxyglucose (FDG) positron emission tomography combined with computed tomography (PET-CT) has emerged as a prominent tool for staging [20,26] and has shown potential as an independent prognostic factor [1,2] in head and neck cancer. Research at our

Correspondence to: Quynh-Thu Le, M.D. Department of Radiation Oncology Stanford University Cancer Center 875 Blake Wilbur Drive Stanford, CA 94305 Tel: (650) 498-5032 Fax: (650) 725-8231 qle@stanford.edu.

Conflict of Interest Statement: no conflicts of interest exist

institution has shown that metabolic tumor volume (MTV) – defined on PET-CT as the volume of hypermetabolic tumor above a given SUV threshold – predicts disease progression and death in head and neck cancer [19,21].

In addition to its role in staging and as a prognostic factor, PET-CT has proven useful in radiation treatment planning [7,10,11,13,14,16,25,28]. Technical advances in radiation oncology allow for the delivery of radiation with continually increasing precision, which permits increased radiation dose to tumors and decreased dose to adjacent normal tissue. The ability to construct radiation treatment plans with steep dose gradients increases the therapeutic ratio of radiotherapy; however this technology emphasizes the importance of target delineation. Tools such as PET-CT, which can visualize the metabolic and anatomic components of cancer, have the capacity to improve the distinction between normal tissue and cancer.

Despite the promise of PET-CT in target definition, the optimal method of how to incorporate this imaging modality into radiation treatment planning remains unclear. The ideal tumor delineation technique with PET-CT would employ an SUV threshold or other segmentation method that distinguishes between pathologic tumor and normal tissue. Several SUV threshold techniques and different tumor segmentation methods have been reported, however authors typically compare PET-CT volumes to phantoms [3,12] or other imaging modalities [10,17,22,23] that may not adequately reflect the pathologic tumor volume [7,11]. Few authors have compared PET-CT segmentation techniques to pathologic tumor specimens in head and neck cancer [7,11,13], and the small sample sizes in these studies preclude definitive conclusions. The purpose of this study was to explore different PET-CT tumor delineation techniques and compare these PET-CT volumes to pathologic tumor volumes in a group of patients with head and neck cancer, specifically oral tongue cancer. We chose to focus on oral tongue cancer in this study since these tumors are often removed as a single specimen, thus facilitating accurate three dimensional pathologic tumor measurement. This study serves as a foundation to identify the optimal metabolic tumor volume for segmentation, which then can be validated against other head and neck tumor types in the future.

Methods

Patients

After Institutional Review Board approval, we reviewed the medical records of all patients with cancer of the oral tongue who underwent surgery or had a PET-CT scan at Stanford University between April 2003 and July 2010. Patients were included if they had histologically confirmed squamous cell cancer of the oral tongue, and a PET-CT scan within 6 weeks of definitive surgery. Patients were excluded if they had recurrent disease, received chemotherapy or radiotherapy prior to surgery, or underwent surgery with palliative intent. Thirty patients met the above criteria: however, the resection specimen was not available in two patients, and five patients had very small T1 primary tumors that were not identifiable on PET-CT. The remaining 23 patients made up the cohort analyzed in this study. Patient and treatment characteristics are provided in Table 1.

FDG PET-CT imaging

All patients underwent PET-CT scans for staging purposes prior to surgery. Patients fasted for at least 6 hours, and plasma glucose was confirmed to be less than 200 mg/dL before injection with the prescribed dose of 15 mCi of FDG (range: 10 to 18 mCi of FDG). PET and CT images were acquired approximately 60 minutes after FDG administration. CT images were acquired first for attenuation correction and anatomical localization of FDG

activity. Two-dimensional (2-D) PET imaging was obtained over 3–5 minutes of acquisition time per bed position. The 2-D PET data were reconstructed with an ordered set expectation maximization (OSEM) algorithm, and reviewed on a dedicated workstation.

Pathology

All patients received a partial- (13%) or hemi-glossectomy (87%) as surgical treatment for their primary oral tongue tumor. Specimens from the operating room were submitted to pathology either fresh (n = 10) or in 10% buffered formalin (n = 13). Those that were submitted fresh from the operating room had three dimensional gross measurements taken from the tumor before being placed into formalin. Otherwise, three dimensional measurements of the tumor were obtained after the tumor had been in formalin for up to 24 hours. To assess the degree of shrinkage from formalin fixation, tumor measurements from the fresh specimens were compared to measurements taken from formalin fixed histology slides. The tumor type (squamous cell carcinoma), grade, keratinization score [24], presence or absence of perineural invasion, margins, and lymph node status were also determined from the histology slides.

PET-CT volumetric endpoints

Metabolic tumor volume (MTV_x) was defined as the volume of hypermetabolic tissue within the region of the gross tumor with an SUV greater than a defined threshold x (Figure 1). While MTV depends explicitly on an SUV threshold, the optimal SUV threshold remains to be defined. Therefore, this study evaluated multiple MTVs with different SUV thresholds including the following: absolute SUV thresholds; relative SUV thresholds expressed as a percentage of the maximum SUV (SUV_{max}); and a gradient-based technique. The absolute threshold levels we explored included SUVs that ranged from 2.0 to 6.0 in increments of 1.0 ($MTV_{2.0} - MTV_{6.0}$). For example, $MTV_{2.0}$ was the volume tumor with an SUV greater than 2.0. The relative SUV thresholds ranged from 30% to 70% of the maximum SUV in increments of 10% ($MTV_{30\%} - MTV_{70\%}$). For example, $SUV_{30\%}$ was the volume of tumor with an SUV greater than 30% of SUV_{max} . In prior studies [19,21], the MTV included the primary tumor and involved nodal groups, however in this study MTV included only the primary oral tongue tumor since this was compared to the primary tumor pathology specimen. The MTV was trimmed off normal tissues such as the mandible (not FDG avid), and the sublingual glands (often increased baseline level of physiologic FDG uptake [5,29]). The absolute and relative MTVs were determined with the MIM® Software Suite along with the MIMfusion® and MIMcontouring® packages (MIMvista Corporation, Cleveland, OH).

The gradient-based technique ($MTV_{gradient}$) defined the boundary of MTV from the gradient between the high SUV in tumor cells and the lower SUV in adjacent normal tissues. $MTV_{gradient}$ was determined with RT Image (version 0.7β), which is an open source software package designed at Stanford to analyze functional imaging data [15]. The following settings for the gradient ROI tool in RT Image were used: initial threshold relative; limit minimum; search maximum; range 20mm; tolerance 0.5; alpha 0.001; iterations 500; unsample 4; downsample 4; and smoothing 3.

Analysis

The pathologic volume of the oral tongue tumor (V_{path}) was estimated from the volume of an ellipsoid:

$$V_{path} = \frac{\pi}{6} * (x_{path} * y_{path} * z_{path})$$

where x_{path} , y_{path} , and z_{path} were the three orthogonal diameters obtained from the resected tumor specimen.

In addition to directly comparing the pathologic tumor volume to MTV, a more clinically relevant measure is the margin of expansion or contraction between the two volumes. The circumferential marginal expansion or contraction (d) was determined by solving the following equation for d :

$$MTV = \frac{\pi}{6} * (x_{path} - 2d) * (y_{path} - 2d) * (z_{path} - 2d)$$

For example, if d were 2mm, then MTV would overestimate the pathologic tumor volume by an approximate 2mm circumferential margin.

In addition to using pre-defined SUV thresholds with the multiple MTV endpoints defined above, we also determined ideal SUV thresholds. The ideal SUV threshold was defined as the SUV value for each patient that yielded an MTV equal to the pathologic tumor volume.

Our initial definitions of the threshold for MTV depended solely on SUV_{max} . With this study we sought to determine if accounting for potential confounding factors could improve the definition of MTV and thus improve the ability of PET to predict pathologic tumor volume. This was accomplished by searching for significant predictors of the ideal SUV threshold, because if one can predict the ideal SUV threshold then one could define an MTV equal to the pathologic tumor volume. First, we noted that SUV_{max} only partially predicted the ideal SUV threshold, however there was unexplained variation which could be attributable to confounding factors. Next, we used multivariate linear regression models to examine the role potential confounders along with SUV_{max} to determine if accounting for the confounders could improve the correlation with the ideal SUV threshold. With the small number of patients in this study the multivariate linear regression model was limited to two predictors, the first was always SUV_{max} , and the second was the potential confounder. The individual potential confounders we examined included tumor grade, tumor keratinization score, tumor stage, nodal stage, perineural invasion status, and the elapsed time between PET-CT and surgery. Of all the confounders we examined, only tumor grade, in addition to SUV_{max} , were significant predictors of the ideal SUV threshold. A multivariate linear regression model including SUV_{max} and tumor grade was constructed to generate a prediction of the ideal SUV threshold ($SUV_{predicted\ threshold}$). Finally, we incorporated the $SUV_{predicted\ threshold}$ into the definition of MTV by introducing $MTV_{tumor\ grade}$ which was defined as the volume of hypermetabolic tissue above each patient's $SUV_{predicted\ threshold}$.

Statistical analysis was done with SAS version 9.2 (SAS Institute Inc., Cary, NC).

Results

Relationship between MTV and pathology volume

The median time between each patient's PET-CT scan and surgery was 10 days (range 3–36 days). The median estimated pathologic tumor volume was 3.1 cm³ (range 0.013–38 cm³). MTV's with absolute SUV thresholds ($MTV_{2.0}$ – $MTV_{6.0}$), relative SUV thresholds ($MTV_{30\%}$ – $MTV_{70\%}$), and MTV with the gradient-based technique ($MTV_{gradient}$) all roughly correlated with pathologic tumor volume ($p < 0.05$), and are demonstrated in Figures 2A–2K. However, the strength of the correlation between MTV and pathologic tumor volume was relatively poor (R^2 range 0.29–0.58). The diagonal gray bands in Figure 2 represent the region where MTV was within $\pm 50\%$ of the pathologic tumor volume, and with each MTV endpoint there were multiple patients that fell outside of the bands,

suggesting a significant over- or underestimation of pathologic tumor volume in several patients. One patient seen on the left side of the plots in Figure 2 had a very small pathologic tumor volume pathology and a larger tumor on PET. Excluding this outlier from the analysis did not significantly change the correlation between PET and pathologic volumes.

Table 2 illustrates the difference between pathologic tumor volume and MTV when the relationship is expressed as the circumferential marginal difference. Similar to the direct volumetric comparison, MTV often under- or overestimated the pathologic volume by large margins.

Predictors of the ideal SUV threshold

The ideal SUV threshold was defined as the SUV cutoff that generated a metabolic tumor volume equal to the gross pathology tumor volume. The median ideal SUV threshold was 5.2 (range 2.1–12.7). Figure 3 demonstrates the relationship between the ideal SUV threshold and SUV_{max} . The ideal SUV threshold increases with SUV_{max} ($p = 0.002$), suggesting that the definition of MTV should in part depend on SUV_{max} . Despite this significant relationship, the overall correlation between the ideal threshold and SUV_{max} was relatively poor ($R^2 = 0.37$), which suggests that other factors may confound this relationship. Indeed, the multivariate linear regression found that both SUV_{max} ($p = 0.0005$) and tumor grade ($p = 0.024$) independently predict the ideal SUV threshold. This analysis yielded the following regression model for the ideal SUV threshold:

$$SUV_{predicted\ threshold} = 4.9 + (0.36 * SUV_{max}) - (1.9 * tumor\ grade)$$

where the *tumor grade* equals 1, 2 or 3 for well-, moderately- or poorly-differentiated squamous cell carcinoma, respectively. This multivariate regression analysis suggests that for a given SUV_{max} , the ideal SUV threshold for a moderately differentiated tumor would be 1.9 units lower than a well differentiated tumor. Additionally, a poorly differentiated tumor's ideal SUV threshold would be 1.9 units lower than a moderately differentiated tumor. This multivariate analysis indicates that higher tumor grades require larger margins on PET-CT to adequately estimate the pathologic tumor volume. Neither T-stage ($p = 0.53$), N-stage ($p = 0.25$), degree of tumor keratinization ($p = 0.95$), perineural invasion ($p = 0.95$), nor elapsed time between PET and surgery ($p = 0.11$) were associated with the ideal SUV threshold.

Finally, $MTV_{tumor\ grade}$ (Figure 2L) slightly improved MTV's ability to predict pathologic tumor volume ($R^2 = 0.63$). This improvement was also seen when the relationship between $MTV_{tumor\ grade}$ and pathologic tumor volume was expressed as the circumferential marginal difference (bottom of Table 2). $MTV_{tumor\ grade}$ was within 2mm of the pathologic tumor volume in 16 patients (70%), and no patient had an $MTV_{tumor\ grade}$ that was more than 5mm away from the pathologic tumor volume.

Tumor shrinkage from formalin fixation

To assess the degree of shrinkage from formalin fixation, the available tumor measurements from the fresh specimens were compared to measurements taken from the associated formalin fixed histology slides. After tumor fixation, the linear tumor dimension decreased an average of 0.22 cm (range 0.1–0.3 cm), which corresponds to an average 14% (range 5–23%) relative decrease in tumor dimension length.

Discussion

The key finding in this study relates to the observation that no single SUV threshold gives a metabolic tumor volume that adequately captures pathologic tumor volume in patients with cancer of the oral tongue. The ability of PET-CT to distinguish between tumor and normal tissue makes it an attractive tool in tumor delineation with radiation treatment planning. Unfortunately, the optimal method to combine this technology with tumor segmentation remains unclear.

When contouring a tumor on a PET-CT, the most frequently used method involves the radiation oncologist visually inspecting PET images, and consciously distinguishing tumor from normal tissue. This method depends substantially on the windowing of the PET and CT scans, as well as the judgment and experience of the treating physician. Ciernik *et al.* [10] noted that PET-CT tumor volumes differ by an average of 9.1cm^3 between independent expert observers. Because of the bias involved with gross visual interpretation, several investigators have studied more objective approaches to tumor delineation.

Burri *et al.* [7] concluded that a threshold of 40% of the maximum SUV ($\text{MTV}_{40\%}$) provided the best compromise between accuracy and avoiding underestimation of the pathologic tumor volume. However, from the raw data reported in the Burri *et al.* study, only four of the twelve patients with surgery specimens had $\text{MTV}_{40\%}$ values that were within $\pm 50\%$ of the pathologic tumor volume. Daisne *et al.* [11] found that an SUV threshold defined as a function of the tumor-to-background ratio most accurately predicted pathologic tumor volume compared to CT or MRI. However, the raw data presented in the Daisne *et al.* investigation revealed that five of the nine patients with tumor specimens had PET volumes more than 50% greater than the pathologic volume. These findings mimic the observations of this current study in that PET-CT volumes fail to accurately predict pathologic tumor volume.

Further work by Geets *et al.* [13] suggested that gradient based segmentation methods outperform the tumor-to-background threshold technique. Indeed, all of the seven analyzed patients in the Geets *et al.* study had gradient based PET volumes within $\pm 50\%$ of the pathologic tumor volume. These findings differ from our results. Although our gradient-based MTV performed relatively better than other MTV volumes, it still underestimated the pathology volume by greater than 5 mm in one patient. Potential explanations for these discordant results could relate to the technical differences in measuring pathologic tumor volume, or the differing techniques for determining the gradient-based MTV. While the gradient-based method in the Geets *et al.* study appears superior, one limitation relates to the requirement that each PET-CT scanner undergo individual calibration prior to implementing their gradient-based technique.

Even though this current study failed to identify a tight relationship between MTV and pathologic tumor volume, an interesting finding relates to the observation that the MTV prediction of pathologic tumor volume improved when it integrated both SUV_{max} and histologic grade. This indicates that the underlying tumor biology, as reflected by tumor grade in the present study, can influence FDG uptake and may be an important factor in defining the optimal SUV threshold. In oral cavity cancer, higher grade tumors have been associated with deeper and more invasive fronts [18], which may not be large enough to be adequately visualized on FDG PET-CT. These invasive fronts may explain why high grade tumors require larger MTV margins. Further research into this hypothesis is warranted.

Although combining the PET parameters with tumor grade makes sense and may improve PET volume delineation, a major drawback of this model relates to the potential variability of tumor cell differentiation within the entire tumor specimen. For example, a single tumor

may have both well differentiated and poorly differentiated regions, and a biopsy from the well differentiated region would conclude the tumor was low grade. Since the standard three-tiered grading system has not consistently correlated with prognosis in oral SCC, proposals have been made for grading the invasive front which is presumed to represent the most aggressive portion of the tumor [6,8]. Another factor that requires consideration relates to the fact that grade is subject to significant inter-observer and intra-observer variation. Prior studies have shown poor interobserver agreement for the grading of oral squamous cell carcinomas [6,24]. For our study, grade concordance was reached between two independent pathologists, who graded the tumor based on the highest grade present within the entire tumor specimen. However, analysis of the entire tumor specimen is not applicable in head and neck cancer cases treated with definitive radiation, which is the most common scenario where PET-CT is used for treatment planning. In addition to sampling variability, tumor grade is often not available in cases when nodal fine needle aspiration has been used to establish diagnosis. Finally, the influence of grade in this model, which focuses on oral cavity cancer, may not translate to other tumor subsites in the head and neck.

Other limitations of this current study are also worth mentioning. First, our pathologic tumor size estimation assumed an ellipsoid tumor shape. The pathologic tumor volume of an oblong or stellate shaped tumor could vary significantly from our estimated pathologic volume, which could significantly bias our conclusions. On the other hand, one reason we chose oral tongue cancer in this analysis relates to the hypothesis that oral tongue tumors should grow relatively unimpeded by hard anatomic structures, and therefore their shape would theoretically imitate an ellipsoid. Regardless, we accept the drawbacks of volumetric estimation with ellipsoid approximation [27], and therefore in future prospective studies we plan to use slice-by-slice measurement of fresh tumor to validate these initial findings. Second, irregularly shaped or heterogeneous tumors could suffer from inaccuracies related the relatively low resolution of PET images [9], which could affect our PET volume estimation. Third, the non-standardized retrospective nature of the pathology specimen handling and measurement process could add bias to our tumor volume measurements. Fourth, this study included both fresh and formalin fixed tumor specimens. While we determined that the overall magnitude of formalin shrinkage was relatively small, this could lead to an underestimation of the pathologic tumor volume. Fifth, the methods used in this study did not allow us to evaluate the spatial location of the PET volume compared to the actual *in vivo* tumor volume. Therefore, the possibility of a geographic miss remains unaddressed. Finally, the presence of dental fillings in the oral cavity can lead to streak artifacts on CT, which would alter the attenuation corrected PET [4], and could unpredictably disturb our MTV estimation.

Despite these limitations, this study demonstrates that several commonly used tumor segmentation techniques fail to accurately predict pathologic tumor volume. The optimal method of integrating PET-CT into radiotherapy treatment planning has yet to be defined, and needs further study.

Acknowledgments

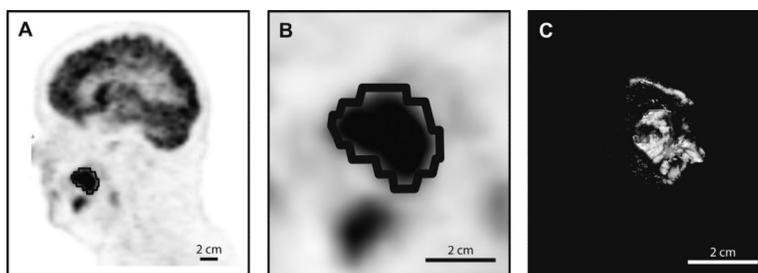
Supported in part by: R01 CA118582-04 (QTL, CK, EEG) & P01 CA67166 (QTL, EEG)

REFERENCES

- [1]. 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; 20:1398–1404. [PubMed: 11870185]

- [2]. Allal AS, Slosman DO, Kebdani T, Allaoua M, Lehmann W, Dulguerov P. Prediction of outcome in head-and-neck cancer patients using the standardized uptake value of 2-[18F]fluoro-2-deoxy-D-glucose. *Int J Radiat Oncol Biol Phys.* 2004; 59:1295–1300. [PubMed: 15275712]
- [3]. Black QC, Grills IS, Kestin LL, et al. Defining a radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys.* 2004; 60:1272–1282. [PubMed: 15519800]
- [4]. Bockisch A, Beyer T, Antoch G, et al. Positron emission tomography/computed tomography--imaging protocols, artifacts, and pitfalls. *Mol Imaging Biol.* 2004; 6:188–199. [PubMed: 15262234]
- [5]. Bogsrud TV, Karantanis D, Lowe VJ. Normal uptake of F-18 FDG in the sublingual gland. *Clin Nucl Med.* 2006; 31:50. [PubMed: 16374130]
- [6]. Bryne M, Koppang HS, Lilleng R, Stene T, Bang G, Dabelsteen E. New malignancy grading is a better prognostic indicator than Broders' grading in oral squamous cell carcinomas. *J Oral Pathol Med.* 1989; 18:432–437. [PubMed: 2607465]
- [7]. Burri RJ, Rangaswamy B, Kostakoglu L, et al. Correlation of positron emission tomography standard uptake value and pathologic specimen size in cancer of the head and neck. *Int J Radiat Oncol Biol Phys.* 2008; 71:682–688. [PubMed: 18258379]
- [8]. Chang YC, Nieh S, Chen SF, Jao SW, Lin YL, Fu E. Invasive pattern grading score designed as an independent prognostic indicator in oral squamous cell carcinoma. *Histopathology.* 2010; 57:295–303. [PubMed: 20659175]
- [9]. Christian N, Lee JA, Bol A, De Bast M, Jordan B, Gregoire V. The limitation of PET imaging for biological adaptive-IMRT assessed in animal models. *Radiother Oncol.* 2009; 91:101–106. [PubMed: 19097661]
- [10]. Ciernik IF, Dizendorf E, Baumert BG, et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. *Int J Radiat Oncol Biol Phys.* 2003; 57:853–863. [PubMed: 14529793]
- [11]. Daisne JF, Duprez T, Weynand B, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology.* 2004; 233:93–100. [PubMed: 15317953]
- [12]. Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol.* 2003; 69:247–250. [PubMed: 14644483]
- [13]. Geets X, Lee JA, Bol A, Lonneux M, Gregoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging.* 2007; 34:1427–1438. [PubMed: 17431616]
- [14]. Geets X, Tomsej M, Lee JA, et al. Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: impact on target volume delineation and dose distribution using helical tomotherapy. *Radiother Oncol.* 2007; 85:105–115. [PubMed: 17562346]
- [15]. Graves EE, Quon A, Loo BW Jr. RT_Image: an open-source tool for investigating PET in radiation oncology. *Technol Cancer Res Treat.* 2007; 6:111–121. [PubMed: 17375973]
- [16]. Henriques de Figueiredo B, Barret O, Demeaux H, et al. Comparison between CT- and FDG-PET-defined target volumes for radiotherapy planning in head-and-neck cancers. *Radiother Oncol.* 2009; 93:479–482. [PubMed: 19880207]
- [17]. Heron DE, Andrade RS, Flickinger J, et al. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. *Int J Radiat Oncol Biol Phys.* 2004; 60:1419–1424. [PubMed: 15590173]
- [18]. Kaneoya A, Hasegawa S, Tanaka Y, Omura K. Quantitative analysis of invasive front in tongue cancer using ultrasonography. *J Oral Maxillofac Surg.* 2009; 67:40–46. [PubMed: 19070746]
- [19]. La TH, Filion EJ, Turnbull BB, et al. Metabolic tumor volume predicts for recurrence and death in head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2009; 74:1335–1341. [PubMed: 19289263]
- [20]. Lonneux M, Hamoir M, Reyckler H, et al. Positron emission tomography with [18F]fluorodeoxyglucose improves staging and patient management in patients with head and

- neck squamous cell carcinoma: a multicenter prospective study. *J Clin Oncol.* 2010; 28:1190–1195. [PubMed: 20124179]
- [21]. Murphy JD, La TH, Chu K, et al. Post-Radiation Metabolic Tumor Volume Predicts Outcome in Head-and-Neck Cancer. *Int J Radiat Oncol Biol Phys.* 2010 in press.
- [22]. Nishioka T, Shiga T, Shirato H, et al. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys.* 2002; 53:1051–1057. [PubMed: 12095574]
- [23]. Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2005; 61:1385–1392. [PubMed: 15817341]
- [24]. Sawair FA, Irwin CR, Gordon DJ, Leonard AG, Stephenson M, Napier SS. Invasive front grading: reliability and usefulness in the management of oral squamous cell carcinoma. *J Oral Pathol Med.* 2003; 32:1–9. [PubMed: 12558952]
- [25]. Schinagl DA, Hoffmann AL, Vogel WV, et al. Can FDG-PET assist in radiotherapy target volume definition of metastatic lymph nodes in head-and-neck cancer? *Radiother Oncol.* 2009; 91:95–100. [PubMed: 19285354]
- [26]. 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; 87:221–229. [PubMed: 18410977]
- [27]. Somville J, De Beuckeleer L, De Schepper A, Verstreken J, Taminiau A. Reliability of measuring volume by different methods for tumors of the musculoskeletal system. *Acta Orthop Belg.* 2001; 67:338–343. [PubMed: 11725564]
- [28]. Troost EG, Schinagl DA, Bussink J, Oyen WJ, Kaanders JH. Clinical evidence on PET-CT for radiation therapy planning in head and neck tumours. *Radiother Oncol.* 2010; 96:328–334. [PubMed: 20708808]
- [29]. Wang Y, Chiu E, Rosenberg J, Gambhir SS. Standardized uptake value atlas: characterization of physiological 2-deoxy-2-[18F]fluoro-D-glucose uptake in normal tissues. *Mol Imaging Biol.* 2007; 9:83–90. [PubMed: 17225983]

**Figure 1. Metabolic tumor volume**

PET scan (A and B) and pathology specimen (C) of an oral tongue squamous cell carcinoma. The black contour encircling the dark hypermetabolic lesion represents an example of metabolic tumor volume (MTV_{50%}).

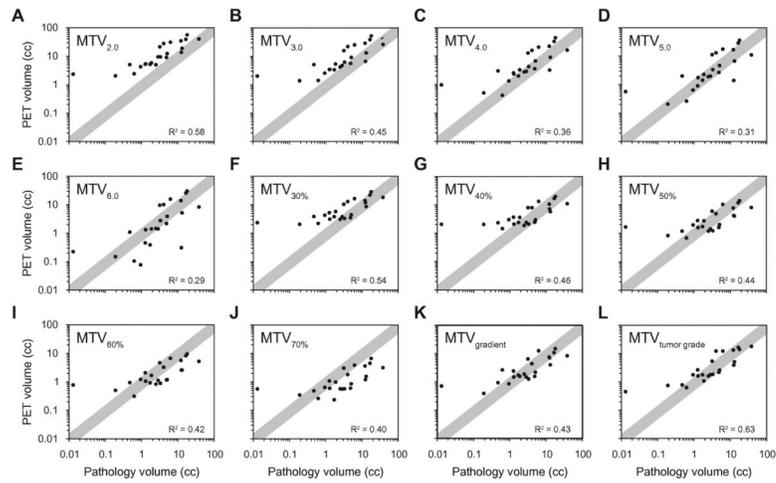


Figure 2. Relationship between Pathology volume and PET volume

Plots represent the correlation between oral tongue tumor volume and metabolic tumor volume (MTV). MTV was defined as the following: volume of tumor on PET-CT above an SUV threshold of 2.0–6.0 (MTV_{2.0} – MTV_{6.0}, A–E); the volume of tumor above 30%–70% of the SUV_{max} (MTV_{30%} - MTV_{70%}, F–J); the volume of tumor defined by the gradient of change in SUV (MTV_{gradient}, K); and an SUV threshold that depends on SUV_{max} and tumor grade (MTV_{tumor grade}, L). The black dots represent individual patients, and gray bands represent the region where MTV is within +/-50% of the pathologic tumor volume.

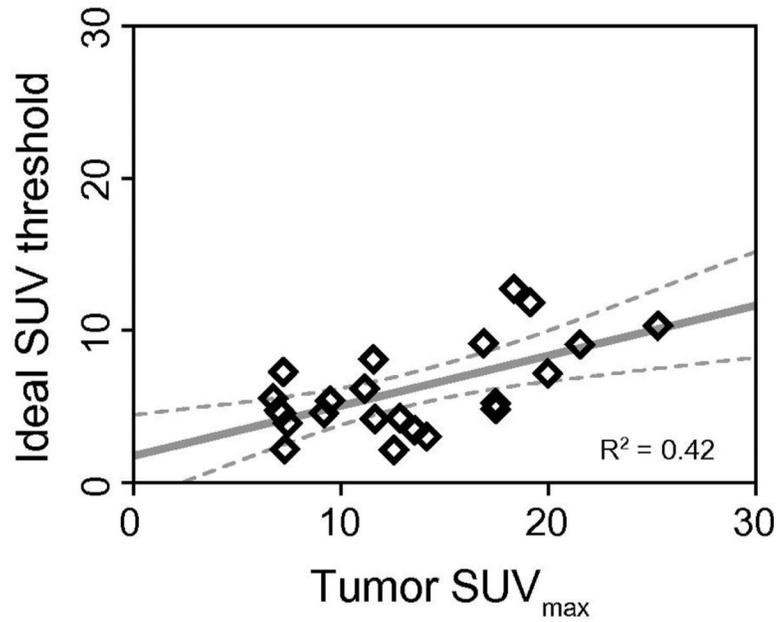


Figure 3. Relationship between the ideal SUV threshold and maximum SUV
Metabolic tumor volume was defined as the volume of tumor above an SUV threshold. The ideal SUV threshold was defined as the SUV cutoff that generates a metabolic tumor volume equal to the pathologic tumor volume. This plot demonstrates the correlation between the ideal SUV threshold and maximum SUV (SUV_{max}) in the 23 patients in this study. The open black diamonds represent individual patients. The solid gray line represents a linear regression through the data, and the dotted gray lines represent the 95% confidence bands for the linear regression.

Table 1

Patient and treatment characteristics

Characteristic	n (%)
Gender	
Male	10 (43)
Female	13 (57)
Age (median [range])	52y (30y–81y)
Tumor stage	
T1	7 (30)
T2	9 (39)
T3	4 (17)
T4	3 (13)
Nodal stage	
N0	16 (70)
N1	1 (4)
N2	6 (26)
Tumor grade	
Well differentiated	5 (22)
Moderately differentiated	15 (65)
Poorly differentiated	3 (13)
Tumor keratinization	
High (>50%)	12 (52)
Moderate (20–50%)	7 (30)
Minimal (5–20%)	4 (17)
None (<5%)	0 (0)
Perineural invasion	
Yes	11 (48)
No	4 (17)
Not available	8 (35)
Surgical margin status	
Negative	21 (91)
Positive	2 (9)
Treatment	
Surgery only	15 (65)
Surgery followed by postoperative radiotherapy	8 (35)

Table 2

Agreement between metabolic tumor volume and pathology volume.

PET volume	Number of patients – n (%)		
	MTV underestimates pathology volume >5mm	2–5mm	MTV overestimates pathology volume >5mm
Absolute SUV thresholds			
MTV _{2.0}	0 (0)	0 (0)	3 (13)
MTV _{3.0}	0 (0)	2 (9)	7 (30)
MTV _{4.0}	1 (4)	1 (4)	13 (57)
MTV _{5.0}	2 (9)	1 (4)	13 (57)
MTV _{6.0}	2 (9)	5 (22)	10 (43)
Relative SUV thresholds			
MTV _{30%}	0 (0)	1 (4)	9 (39)
MTV _{40%}	1 (4)	2 (9)	12 (52)
MTV _{50%}	1 (4)	5 (22)	14 (61)
MTV _{60%}	3 (13)	8 (35)	11 (48)
MTV _{70%}	3 (13)	9 (39)	10 (43)
Gradient-based			
MTV _{gradient}	1 (4)	5 (22)	14 (61)
Tumor grade			
MTV _{tumor grade}	0 (0)	4 (17)	16 (70)

This table represents the circumferential marginal difference between PET volume and pathology volume, which was defined as the circumferential distance in millimeters required to expand or shrink the metabolic tumor volume (MTV) such that the MTV and pathology volumes were equal (see Methods for details).