Validation that Metabolic Tumor Volume Predicts Outcome in Head and Neck Cancer

Chad Tang, MS[*,†], **James D. Murphy, MD, MS**[*,†], **Brian Khong, BA**[†], **Trang H. La, MD**[†], **Christina Kong, MD**[††], **Nancy J. Fischbein, MD**[§], **A. Dimitrios Colevas, MD**[&], **Andrei H. Iagaru, MD**[§], **Edward E. Graves, PhD**[†], **Billy W. Loo Jr, MD, PhD**[†], and **Quynh-Thu Le, MD**[†]

[†]Department of Radiation Oncology, Stanford University, Stanford, CA, USA

[††]Department of Pathology, Stanford University, Stanford, CA, USA

[§]Department of Radiology, Stanford University, Stanford, CA, USA

[&]Division of Medical Oncology, Department of Medicine, Stanford University, Stanford, CA, USA

Abstract

Purpose—We have previously reported that metabolic tumor volume (MTV) obtained from pre-treatment FDG PET/CT predicted outcome in patients with head-and-neck cancer (HNC). The purpose of this study is to validate these results on an independent dataset, determine if the primary tumor or nodal MTV drives this correlation, and explore the interaction with $p16^{INK4a}$ status as a surrogate marker for HPV.

Methods and Materials—The validation dataset in this study included 83 patients with squamous cell HNC who had a FDG PET/CT scan prior to definitive radiotherapy. MTV and SUV_{max} were calculated for the primary tumor, involved nodes, and the combination of both. The primary endpoint was to validate that MTV predicted progression-free survival and overall survival. Secondary analyses included determining the prognostic utility of primary tumor versus nodal MTV.

Results—Similar to our prior findings, an increase in total MTV of 17 cm^3 (difference between 75th and 25th percentile) was associated with a 2.1 fold increase in the risk of disease progression (p=0.0002), and a 2.0 fold increase in the risk of death (p=0.0048). SUV_{max} was not associated with either outcome. Primary tumor MTV predicted progression-free (HR=1.94; p<0.0001) and overall (HR=1.57; p<0.0001) survival, whereas nodal MTV did not. In addition, MTV predicted progression-free (HR=4.23; p<0.0001) and overall (HR=3.21; p=0.0029) survival in patients with $p16^{INK4a}$ positive oropharyngeal cancer.

Conclusions—This study validates our previous findings that MTV independently predicts outcomes in HNC. MTV should be considered as a potential risk stratifying biomarker in future studies of HNC.

Correspondence to Dr. Quynh-Thu Le, 875 Blake Wilbur Drive, MC 5847, Stanford, CA 94305-5847, qle@stanford.edu; Tel: 650-498-5032; Fax: 650-725-8231.
[*]equal contributions.

Conflict of Interest Notification: None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Strategies combining chemotherapy and radiotherapy in head-and-neck cancer (HNC) have reduced rates of disease progression and increased survival(1). Unfortunately, these improvements in outcome come at the cost of increased toxicity. In HNC subsites such as oropharyngeal cancers, one of the most significant prognostic factors to emerge has been human papilloma virus (HPV), where positive HPV status has been found to portend favorable prognosis(2). In light of such findings, there has been recent motivation to de-intensify treatment to reduce toxicity while maintaining efficacy. To this end, an ongoing Eastern Cooperative Oncology Group (ECOG) clinical trial is evaluating a dose-reduced radiation schedule in HPV positive oropharyngeal cancer patients who have had a good response to induction chemotherapy(3). Biomarkers such as HPV, which stratify patients according to risk of disease progression, will become crucial as we venture towards risk-adapted therapy.

Our institution has previously reported on one such risk-stratifying biomarker--metabolic tumor volume (MTV)--which is defined from ^{18}F-fluorodeoxyglucose positron emission tomography (PET) scans combined with computed tomography (CT)(4). Overall, studies evaluating the prognostic utility of PET/CT in HNC have been mixed(4, 5); however, most have focused solely on the maximum standardized uptake value (SUV_{max}). Unlike SUV_{max}, MTV does not rely solely on a single point, but rather provides a quantification of metabolic tumor burden. Indeed, we previously demonstrated that increased MTV was associated with higher rates of disease progression and death(6). While our original study demonstrated promise for MTV as a risk-stratifying biomarker, the retrospective study design limits its generalizability. The purpose of the current study was to validate these results on an independent dataset. Additionally, we sought to determine if primary tumor or nodal MTV drives any correlations with outcome. Finally, we explored the relationship between MTV and p16^{INK4a} status, which was used as a surrogate marker for HPV.

Methods and Materials

Patients

Following institutional review board approval, we reviewed medical charts of patients with biopsy proven squamous cell HNC and a PET/CT scan conducted at most 2 months prior to radiotherapy. All patients received radiation treatment between April 2003 and December 2009. Patients were excluded if they had evidence of distant metastatic disease, received prior definitive surgery, radiation, or chemotherapy, or were treated with palliative intent. Patients with salivary gland, paranasal sinus, thyroid, and skin primaries were also excluded.

Our original analysis included 85 patients that we reported on previously(6). The validation dataset in this study consisted of 83 new patients who were accrued after the original dataset. The complete dataset combines both the original and validation datasets (total 168 patients). Validation analysis was conducted on the validation dataset alone, and all subsequent analyses were conducted on the total dataset to increase statistical power.

FDG-PET/CT imaging

After fasting for at least 8 hours and ensuring blood sugars were less than 180 mg/dL, patients were injected with 10 to 18 mCi of FDG followed by PET imaging 45 to 60 minutes

afterwards. CT data was collected in helical acquisition mode. PET images covering the same field of view were acquired in two-dimension mode. The PET images were reconstructed with an ordered-subset expectation algorithm using the CT data for attenuation correction.

SUV$_{max}$ and MTV measurement

The metabolic volumes of interest were retrospectively delineated on all PET/CT scans. The definition of MTV from our original report(6) included both the primary tumor and involved lymph nodes. In this study we divided MTV into either primary tumor MTV or nodal MTV. Primary tumor MTV was defined as the primary tumor volume above 50% of the primary tumor SUV$_{max}$. Nodal MTV was defined as the nodal tumor volume above 50% of the nodal SUV$_{max}$. The total MTV was the sum of the primary tumor and nodal MTV. The total SUV$_{max}$ was defined as the maximum of the primary tumor and nodal SUV$_{max}$. Representative images in Figure 1 show nodal and primary tumor MTVs. PET volumetric analysis was conducted with MIM Software (MIMvista Corporation, Cleveland OH).

p16^{INK4a} staining

Tumor tissue was available in 47% of oropharynx cancer patients. Immunoperoxidase stains for p16^{INK4a} (clone E6H4, Dako) were performed on tissue sections as previously described(7). Weak cytoplasmic staining in < 5% of the cells was interpreted as negative. Focal strong nuclear and/or cytoplasmic staining (5–80% of cells), and diffuse strong staining (>80% of cells) were both considered positive.

Follow-up and Neck Management Policy

Two months after completion of chemoradiotherapy, all patients underwent detailed head-and-neck physical examination and had imaging studies (either CT or MRI). Patients with clinically palpable node(s) proceeded to have an ultrasound-guided fine needle aspiration (FNA) of the largest node. If the FNA detected cancer a salvage neck dissection was performed, at which time an evaluation of the primary site was conducted intraoperatively. Patients with a negative FNA, and those with a clinically negative neck and a complete response at the primary site underwent a follow up PET/CT 3 months after radiotherapy. If the PET/CT showed persistent uptake at the primary site, a biopsy was performed and surgical salvage was carried out based on the biopsy results. Additional follow-up included a detailed HN evaluation every 2 months for the first 2 years, every 3 months for the third year, every 6 months for the fourth and fifth years, and yearly thereafter. Chest x-rays were obtained annually.

Data analysis

The validation dataset (n=83) was used to validate our original findings. For the validation analysis we used identical analytic methods as employed in our original study(6). Survival curves were generated from the method of Kaplan and Meier. The outcomes, progression-free survival and overall survival, were defined from the date of diagnosis. Events for progression-free survival included disease progression or death from any cause. Events for overall survival were death from any cause. Cox proportional hazard models were used to evaluate the prognostic utility of the PET endpoints. Prior to entry into the proportional hazards model, each PET endpoint was normalized to its interquartile range (difference between 1st and 3rd quartiles). As with our original report(6), the validation PET endpoints of interest included total MTV and total SUV$_{max}$.

After the validation analysis, all subsequent analyses were conducted on the entire dataset (n=168) defined above. We evaluated relationships between outcomes (local-regional

control and distant metastatic failure) and total MTV with cumulative incidence plots (to account for competing risks), and differences between strata were evaluated with Gray's test(8). The association between primary tumor MTV, nodal MTV and outcome was assessed with Cox proportional hazard models. Further subset analyses were conducted focusing on tumor subsite and p16^{INK4a} status. The correlations between PET endpoints and tumor characteristics were assessed with a Pearson correlation coefficient. Statistical analysis was done with SAS version 9.2 (SAS Institute Inc., Cary, NC).

Results

Patient characteristics

Demographics of the validation dataset compared to the original dataset are shown in Table 1. The validation dataset included a higher percentage of oropharyngeal primaries and lower T-classification when compared to the original dataset. There was no difference in the distribution of gender, nodal stage, or histology grade. More oropharynx tumors in the validation subset were tested for p16^{INK4a} status, however the proportion of tumors with p16^{INK4a} positivity was similar. The median follow-up time was significantly longer in the original dataset compared to the validation dataset (38 vs 20 months; p<0.0001), likely owing to the fact that patients in the validation set were accrued later. The median follow-up for the entire cohort was 24 months (range 1.4–85 months), and the median follow-up in living patients was 25 months (range 1.4–85 months).

The majority of patients (96%) received concurrent chemotherapy with definitive radiation, consisting of cisplatin- or cetuximab-based regimens. Radiation doses ranged from 66 to 72 Gy. All patients were treated with 3-dimensional conformal radiation (5%) or intensity-modulated radiation therapy (IMRT) (95%). Fifteen patients (9%) had post-chemoradiation neck dissections, of which 11 were performed because of persistently enlarged nodes on clinical evaluation at 2 months or concerning nodal uptake on PET/CT at 3 months after chemoradiation, and four were done as planned neck dissections, pre-declared for patients treated on 2 different clinical trials. Of the 15 patients undergoing neck dissection, six were found to have residual cancer.

Validation of MTV as a prognostic factor

In the validation dataset the 2-year progression-free and overall survival rates were 80% and 86%, respectively. As with the original analysis, total MTV predicted progression-free and overall survival (Table 2, and Figure 2). The median total MTV was 10.5 cm^3 (range 0.8–70 cm^3). An increase in total MTV of 17 cm^3 (difference between 1st and 3rd quartiles) increased the risk of disease progression by 107% (p=0.0002) and increased the risk of death by 99% (p=0.005). When compared to our original analysis (Table 2), the hazard ratios remained relatively unchanged, despite different underlying patient characteristics. Similar to our original analysis, SUV$_{max}$ failed to predict progression-free (hazard ratio [HR]=1.04; p=0.88) or overall (HR=1.10; p=0.70) survival in these new patients.

Local and distant failure

Among the entire dataset, there were 10 local-regional only failures, 17 distant metastatic only failures and 5 patients who failed both local-regionally and distantly. The 2-year cumulative incidence of local progression and distant metastases were 8% and 12%, respectively. Total MTV was a significant predictor of local-progression (p=0.014) and distant metastatic failure (p=0.023) (Figure 3). Total SUV$_{max}$ failed to predict local-regional progression (p=0.54), but did predict distant-metastatic failure (p=0.026).

Primary tumor and nodal MTV

To further explore the driving forces behind the relationship between MTV and outcome, we divided MTV into its primary tumor and nodal components as described above. When analyzing the entire dataset, the primary tumor MTV predicted progression-free (HR=1.94; p<0.0001) and overall (HR=1.57; p<0.0001) survival. However, nodal MTV predicted neither progression-free (HR=1.08; p=0.41) nor overall (HR=1.05; p=0.66) survival. Of note, there was no correlation between primary tumor MTV and nodal MTV (Pearson R^2<0.01).

When evaluating the different components of SUV_{max}, primary tumor SUV_{max} failed to predict progression-free (HR=1.20; p=0.25) or overall (HR = 1.27; p=0.11) survival, whereas nodal SUV_{max} predicted both progression-free (HR=1.42; p=0.015) and overall (HR=1.34; p=0.04) survival. There was minimal correlation between primary tumor SUV_{max} and nodal SUV_{max}(R^2=0.19).

Prognostic value of MTV by p16[INK4a] status

We next sought to determine if MTV predicts outcome across different HNC subsites and specifically for oropharyngeal tumor when stratifying by p16[INK4a] status (9). Within the oropharygeal carcinoma group, there was a trend towards decreased MTV in p16[INK4a] positive tumors (median MTV=11 cm^3; range 1.6–47 cm^3) compared to p16[INK4a] negative tumors (median MTV=18 cm^3; range 4.4–80 cm^3), but this did not reach statistical significance (p=0.25). In the p16[INK4a] positive oropharynx subset (n=64), total MTV remained a robust predictor of progression-free (HR=4.23; p<0.0001) and overall (HR=3.21; p=0.0029) survival (Figure 4). The limited number of p16[INK4a] negative oropharyngeal carcinoma (n=10) precluded an adequately powered analysis in this subset.

In the combined subset of patients with hypopharynx or larynx cancer (n=23) there was a trend towards worse outcomes with higher MTV similar in magnitude to the entire dataset, however this did not reach statistical significance with either progression-free (HR=1.9; p=0.40) or overall (HR=1.9; p=0.46) survival. On the other hand, with nasopharyngeal carcinoma (n=30), increased MTV appeared less associated with outcome, and failed to predict progression-free (HR=1.42; p=0.39) or overall (HR=0.78; p=0.77) survival. Of note, the limited number of patients with nasopharyngeal or laryngeal/hypopharyngeal carcinomas reduces the power of this subset analysis; thus, caution should be taken when interpreting these non-significant p-values. Similarly, the small numbers of patients with oral cavity (n=4) or unknown primary (n=2) prevented further analysis on these subsets.

MTV as an independent prognostic factor

Finally, we assessed potential confounding variables and their effect on the prognostic utility of MTV. Given that MTV estimates tumor burden, one could argue that total MTV is a surrogate for tumor or nodal classifications. Despite this argument, the correlation between primary tumor MTV and T-classification was weak (R^2=0.24), and there was no correlation between nodal MTV and N-classification (R^2=0.07). Additionally, when controlling for T- and N-classification, total MTV remained an independent significant predictor of progression-free (HR=1.85; p=0.0002) and overall (HR=1.70; p=0.0048) survival on multivariate analysis. In fact, we also found that when controlling for N-classification and total MTV, T-classification was not an independent significant predictor of progression-free (p=0.10) or overall (p=0.48) survival. Similarly, when controlling for T-classification and total MTV, N-classification was not an independent significant predictor of progression-free (p=0.68) or overall (p=0.46) survival.

Discussion

The key finding of this study relates to the validation of MTV as an independent and robust biomarker easily obtainable from pre-treatment PET/CT imaging. The patient characteristics of our validation dataset differed compared to the original dataset. The current dataset contained younger patients, with lower T-classification, and a higher frequency of oropharyngeal primaries, which reflects the shifting demographics of HNC(10). In recent years we have observed a surge in p16^{INK4a} positive oropharyngeal cancer, for which the patients are often younger and have smaller primary tumors. The nearly identical relative risk of disease progression and death between our original study and this validation study (Table 2) in the face of differing patient characteristics only adds to the generalizeability of our findings.

The most well studied PET/CT metric in HNC is SUV$_{max}$. Although some groups found predictive utility from SUV$_{max}$, many studies including this one have struggled to reproduce this finding(5, 6). Recently, Moeller and colleagues conducted the first prospective trials evaluating the prognostic impact of SUV$_{max}$ in 98 patients with locally advanced HNCs(11). They found SUV$_{max}$ to outperform CT in only a small subset of patients at high risk of treatment failure, mainly those with HPV negative cancer, non-oropharyngeal primaries, or with a history of tobacco abuse. Similar to Moeller's findings, we found that SUV$_{max}$ was not a useful prognostic factor for most patients.

Echoing our findings, other groups have found correlations between high pre-treatment MTV and outcome(12). Furthermore, others have also divided MTV into primary tumor and nodal components. For example, Hu and colleagues found a trend towards a higher rate of metastatic disease with increasing nodal MTV(13). On the contrary, we found that primary tumor MTV, not nodal MTV, drove the correlation between MTV and outcome. While the mechanism behind this observation remains unclear, these results could reflect that disease burden of the primary tumor has more prognostic value than that in the lymph nodes.

The response to combined modality chemoradiation in HNC remains variable, even in contemporary series(14, 15). Despite this variability, certain patient subsets have more favorable prognoses, such as p16^{INK4a} positive oropharynx cancer patients(2). Biomarkers such as HPV and p16^{INK4a} have led to risk stratifying protocols(3) aimed at reducing treatment intensity in patients with favorable risk profiles. Validated biomarkers such as MTV will allow for further risk stratification, even when applied in conjunction with other biomarkers, namely with p16^{INK4a} positive oropharyngeal cancers.

While our findings show the robust prognostic utility of MTV, certain study limitations exist that are worth mentioning. Our study cohort included a heterogeneous group of patients harboring squamous cell carcinoma of different head-and-neck mucosal subsites who received treatment with different chemoradiotherapy regimens. The patient and treatment heterogeneity makes it difficult to confirm the patient subset for which MTV best predicts outcome. On the other hand, heterogeneity in our study increases the generalizeability of our findings, which is important since HNC is inherently a heterogeneous disease. Another limitation relates to the reduced statistical power when comparing smaller subsets. For example, our analysis of patients with non-oropharyngeal subsets was underpowered, therefore the question of whether MTV predicts outcome in these tumors remains unaddressed. Additionally, the small number of post-chemoradiotherapy neck-dissections limited our ability to answer the important question of whether nodal MTV predicts for the presence of residual nodal disease. Another limitation relates to the relatively short follow up in our cohort. While our survival curves show a wide separation between patients with high and low MTV, the curves could theoretically converge over time. Longer follow up is

required to determine the durability of MTV's prognostic capability. Ultimately, evaluating MTV on a larger population in a prospective setting will overcome several of these limitations, and will allow for further validation of these current findings.

Other limitations relate to the single-institution study design. Our institution used a standard protocol with PET imaging and a single software system for image analysis. Different imaging protocols, PET scanners, and image processing techniques could potentially affect the results and impact the effectiveness of MTV, especially when utilized across different institutions. Given these limitations, the relationship between MTV and survival in patients with locally advanced HNC should be validated in prospective multi-institution study, which incorporates a rigorous standardized protocol for PET/CT imaging(16).

Finally, the question of how MTV compares to tumor volume derived from other imaging modalities remains unaddressed. Moeller et al. has made the argument that contrast-enhanced CT or MRI has equal predictive value compared to SUV_{max}(11). One could extend this argument to other PET metrics, including MTV. Ultimately, we feel that a standardized approach to defining MTV should significantly reduce inter-user variability inherent in other radiologic metrics. Indeed, PET in combination with CT has been shown to reduce inter-user variability in the definition of tumor volume compared to CT alone(17). Additionally, tumor volumes measured with PET-CT have been shown to estimate true pathologic tumor volume more precisely than CT or MRI(18), which lends support to the hypothesis that MTV defined on PET-CT could be a superior surrogate of tumor burden.

In summary, this study validates our previous findings that MTV independently predicts HNC outcomes. This finding holds true in p16^{INK4a} positive oropharyngeal carcinomas, and it appears that primary tumor MTV and not nodal MTV drives the correlation between MTV and outcome. Additionally, we found that MTV was a better predictor of outcomes compared to the more widely used PET metric, SUV_{max}. MTV should be considered as a potential risk stratifying biomarker in future studies.

Acknowledgments

References

1. Pignon JP, le Maitre A, Maillard E, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol. 2009; 92:4–14. [PubMed: 19446902]

2. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010; 363:24–35. [PubMed: 20530316]

3. ECOG. Institute NC. ClinicalTrials.gov. 2011. Paclitaxel, Cisplatin, and Cetuximab Followed By Cetuximab and Intensity-Modulated Radiation Therapy in Treating Patients With HPV-Associated Stage III or Stage IV Cancer of the Oropharynx That Can Be Removed By Surgery.

4. 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]

5. Murphy JD, La TH, Chu K, et al. Postradiation Metabolic Tumor Volume Predicts Outcome in Head-and-Neck Cancer. Int J Radiat Oncol Biol Phys. 2010

6. Tang C, Kligman F, Larsen CC, et al. Platelet and endothelial adhesion on fluorosurfactant polymers designed for vascular graft modification. J Biomed Mater Res A. 2009; 88:348–358. [PubMed: 18286624]

7. Kong CS, Narasimhan B, Cao H, et al. The relationship between human papillomavirus status and other molecular prognostic markers in head and neck squamous cell carcinomas. Int J Radiat Oncol Biol Phys. 2009; 74:553–561. [PubMed: 19427557]

8. Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. Annals of Statistics. 1988; 16:1141–1154.

9. Rischin D, Young RJ, Fisher R, et al. Prognostic significance of p16INK4A and human papillomavirus in patients with oropharyngeal cancer treated on TROG 02.02 phase III trial. J Clin Oncol. 2010; 28:4142–4148. [PubMed: 20697079]

10. Marur S, D'Souza G, Westra WH, et al. HPV-associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol. 2010; 11:781–789. [PubMed: 20451455]

11. 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; 27:2509–2515. [PubMed: 19332725]

12. Xie P, Yue JB, Zhao HX, et al. Prognostic value of 18F-FDG PET-CT metabolic index for nasopharyngeal carcinoma. J Cancer Res Clin Oncol. 2010; 136:883–889. [PubMed: 19936788]

13. Hu, K.; Urdaneta, A.; Young, A., et al. Value of pre-treatment PET and post-treatment pathologic nodal status to predict risk for distant metastasis in N2-3 patients treated with concurrent chemotherapy. Multidisciplinary Head and Neck Cancer Symposium; Chandler, AZ. 2010.

14. Daly ME, Le QT, Jain AK, et al. Intensity-modulated radiotherapy for locally advanced cancers of the larynx and hypopharynx. Head Neck. 2011; 33:103–111. [PubMed: 20848427]

15. Daly ME, Le QT, Maxim PG, et al. Intensity-modulated radiotherapy in the treatment of oropharyngeal cancer: clinical outcomes and patterns of failure. Int J Radiat Oncol Biol Phys. 2010; 76:1339–1346. [PubMed: 19540068]

16. Juweid ME, Stroobants S, Hoekstra OS, et al. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. J Clin Oncol. 2007; 25:571–578. [PubMed: 17242397]

17. Murakami R, Uozumi H, Hirai T, et al. Impact of FDG-PET/CT fused imaging on tumor volume assessment of head-and-neck squamous cell carcinoma: intermethod and interobserver variations. Acta Radiol. 2008; 49:693–699. [PubMed: 18568563]

18. 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]
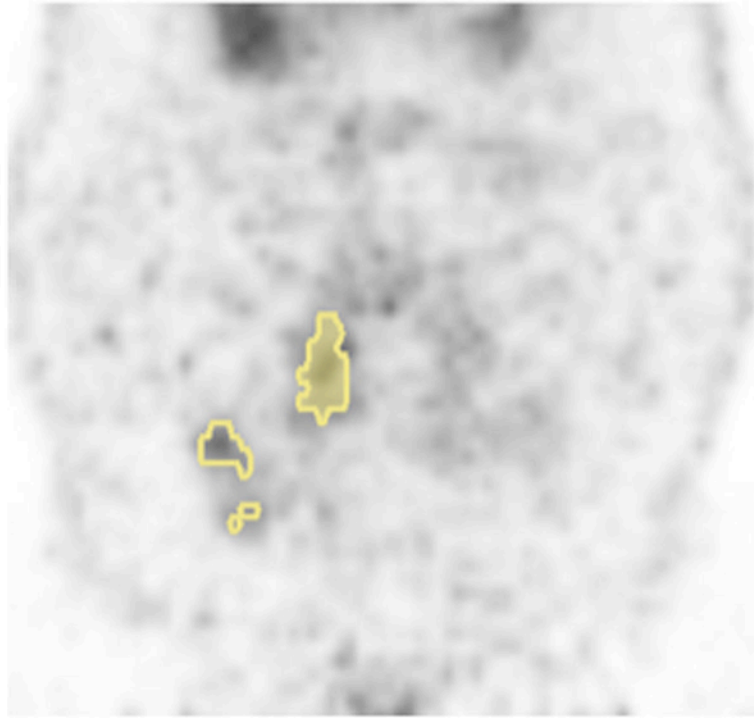
Figure 1.
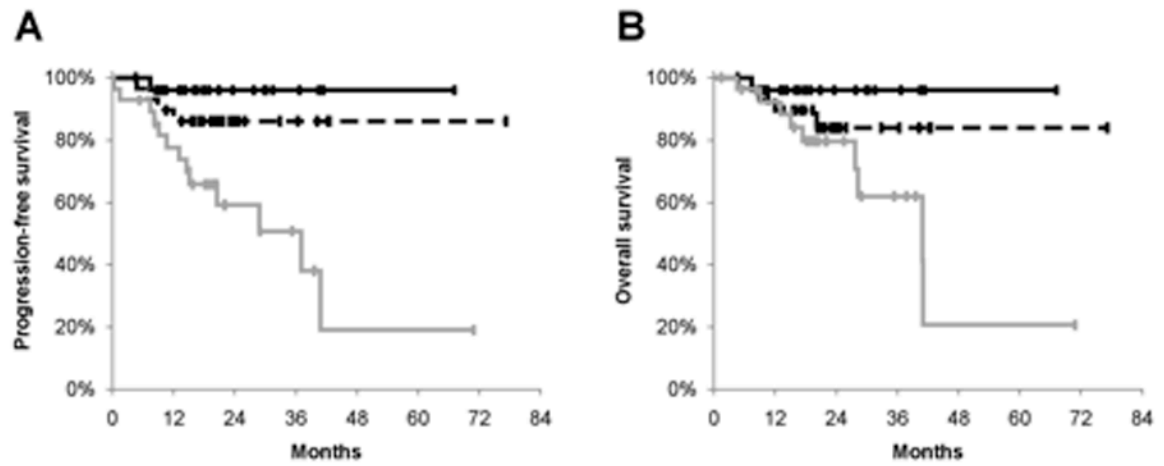Representative PET image with primary tumor MTV (shaded fill) and nodal MTV (open fill).

Figure 2.
Plots represent Kaplan Meier analysis of progression-free survival (A), and overall survival (B) by total MTV for the validation dataset (n=83). Lines represent total metabolic tumor volume (MTV) divided into bottom (solid black), middle (dashed black), and top (solid grey) tertiles.
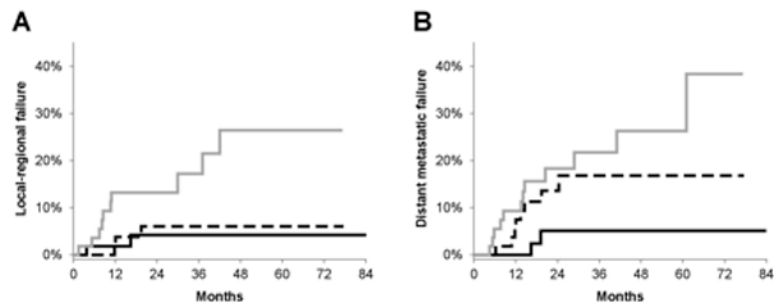
Figure 3.
Plots represent cumulative incidence of local-regional failure (A) and distant metastatic failure (B) by total MTV for the entire dataset (n=168). Lines represent total metabolic tumor volume (MTV) divided into bottom (solid black), middle (dashed black), and top (solid grey) tertiles.
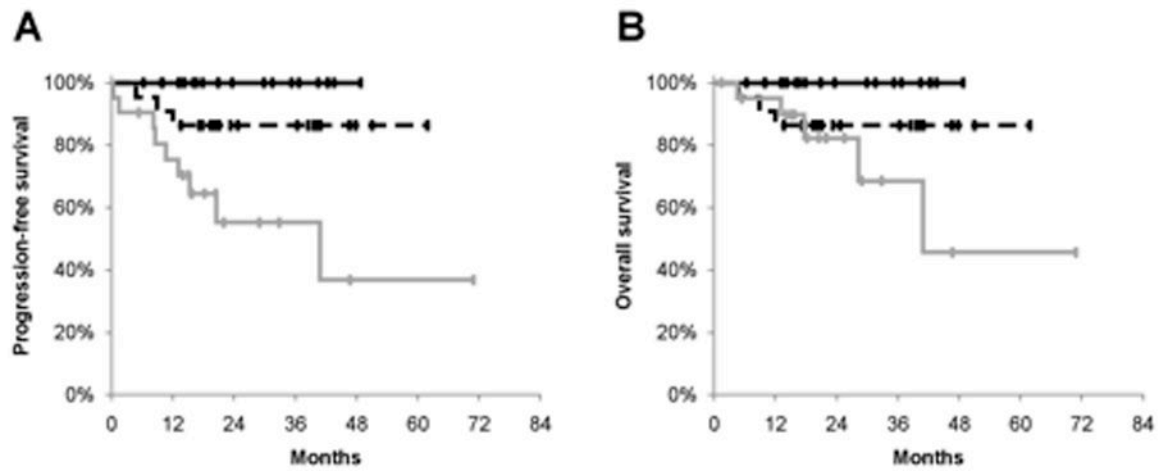
Figure 4.
Plots represent Kaplan Meier analysis of progression-free (A) and overall (B) survival for the subset of patients with HPV/p16INK4a positive oropharyngeal cancer (n=64). Lines represent total metabolic tumor volume (MTV) divided into bottom (solid black), middle (dashed black), and top (solid grey) tertiles.

Table 1

Characteristics of Original and Validation Datasets

Characteristic	Original patients (n=85)	Validation patients (n=83)	p-value[*]	All patients (n=168)
Age - median (range)	57 (14–86)	59 (35–89)	0.06	58 (14–89)
Gender				
Female	12 (14)	7 (8)		19 (11)
Male	73 (86)	76 (92)	0.33	149 (89)
T-classification				
T0	2 (2)	-		2 (1)
T1	12 (14)	20 (24)		32 (19)
T2	25 (29)	26 (31)		51 (30)
T3	18 (21)	24 (29)		42 (25)
T4	28 (33)	13 (16)	0.05	41 (24)
N-classification				
N0	10 (12)	9 (11)		19 (11)
N1	6 (7)	11 (13)		17 (10)
N2	59 (69)	56 (67)		115 (68)
N3	10 (12)	7 (8)	0.37	17 (10)
Tumor site				
Oropharynx	45 (53)	64 (77)		109 (65)
Larynx	4 (5)	6 (7)		10 (6)
Hypopharynx	8 (9)	5 (6)		13 (8)
Nasopharynx	22 (26)	8 (10)		30 (18)
Oral Cavity	4 (5)	-		4 (2)
Unknown primary	2 (2)	-	0.003	2 (1)
p16^{INK4a} status (oropharynx subset only)				
Positive	17 (33)	47 (73)		64 (59)
Negative	2 (4)	8 (13)	1.00 [**]	10 (9)
Unknown	33 (63)	9 (14)		35 (32)
Pathology grade				
Low	2 (2)	2 (2)		4 (2)
Medium	29 (34)	31 (37)		60 (36)
High	44 (52)	35 (42)		79 (47)
Unknown	10 (12)	15 (18)	0.4	25 (15)
Karnofsky performance status				
60	2 (2)	1 (1)		3 (2)
70	9 (11)	7 (8)		16 (10)
80	14 (16)	22 (27)		36 (21)
90	59 (69)	43 (52)		102 (61)
100	1 (1)	3 (4)		4 (2)
Unknown	-	7 (8)	0.45	7 (4)

**
this comparison refers only to patients with known p16^{INK4a} status.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Validation Analysis

Inter-dataset validation		
	Total MTV	
Dataset and outcomes	p-value	HR (95% CI)
Original dataset		
Progression-free survival	0.00012	1.86 (1.36 – 2.56)
Overall survival	0.00051	2.06 (1.37 – 3.09)
Validation dataset		
Progression-free survival	0.00017	2.07 (1.42 – 3.03)
Overall survival	0.0048	1.99 (1.23 – 3.2)

Abbreviations: MTV=metabolic tumor volume, HR=hazard ration, CI=confidence interval