18F-FDG Metabolic Tumor Volume and Total Glycolytic Activity of Oral Cavity and Oropharyngeal Squamous Cell Cancer: Adding Value to Clinical Staging

Elizabeth H. Dibble1, Ana C. Lara Alvarez1, Minh-Tam Truong2, Gustavo Mercier1, Earl F. Cook3, and Rathan M. Subramaniam1,2

1Department of Radiology, Boston University School of Medicine, Boston, Massachusetts; 2Department of Radiation Oncology, Boston University School of Medicine, Boston, Massachusetts; and 3Department of Epidemiology, Harvard School of Public Health and Brigham and Women’s Hospital, Boston, Massachusetts

18F-FDG metabolic tumor volume (MTV) and total glycolytic activity (TGA) have been proposed as potential prognostic imaging markers for patient outcome in human solid tumors. The purpose of this study was to establish whether MTV and TGA add prognostic information to clinical staging in patients with oral and oropharyngeal squamous cell carcinomas (SCCs).

Methods: The Institutional Review Board approved this Health Insurance Portability and Accountability Act–compliant single-institution retrospective study. Forty-five patients with histologically proven oral or oropharyngeal SCC underwent PET/CT for initial cancer staging and were included in the study. MTV was measured using a gradient-based method (PET Edge) and fixed-threshold methods at 38%, 50%, and 60% of maximum standardized uptake value (SUV). The TGA is defined as MTV × mean SUV. Bland–Altman analysis was used to establish the reliability of the methods of segmentation. Outcome endpoints were overall survival (OS) and progression-free survival. Cox proportional hazards univariate and multivariate regression analyses were performed. Results: In Cox regression models, MTV and TGA were the only factors significantly associated with survival outcome after adjusting for all other covariates including American Joint Committee on Cancer (AJCC) stage, with hazards ratio of 1.06 (95% confidence interval, 1.01–1.10; \( P = 0.006 \)) and 1.00 (95% confidence interval, 1.00–1.01; \( P = 0.02 \)). The model fit was significantly better when MTV was added to AJCC stage in model I (\( \chi^2 \) value change, 1.16–6.71; \( P = 0.01 \)) and when TGA was added to AJCC stage in model II (\( \chi^2 \) value change, 1.16–4.37; \( P = 0.04 \)). The median cutoff point of 7.7 mL for primary tumor MTV was predictive of time to OS (log rank \( P = 0.04 \)). The median cutoff point of 55 g for PET Edge primary tumor TGA was predictive of time to OS (log rank \( P = 0.08 \)), though the result was not statistically significant.

Conclusion: Gradient-based segmentations of primary tumor MTV and TGA are potential 18F-FDG markers for time to survival in patients with oral and oropharyngeal SCC and may provide prognostic information in addition to AJCC stage. These exploratory imaging markers need validation in larger cohort studies.

Key Words: metabolic tumor volume; total glycolic activity; oral and oropharyngeal SCC; staging

DOI: 10.2967/jnumed.111.099531

Approximately 50,000 new cases of squamous cell carcinomas (SCCs) of the head and neck (HNSCC) are diagnosed each year, more than 39,000 of which are oral and oropharyngeal cancers with an expected mortality of about 7,900 in 2011 in the United States (1). Despite sharing a common histologic classification, HNSCC includes a heterogeneous mix of cancers with different natural histories that are historically defined best by site of origin and TNM staging (2). PET/CT has become increasingly important in localizing and staging HNSCC, identifying unknown primaries, detecting synchronous primaries, assessing therapy response, and monitoring for cancer recurrence (3–5).

Various treatment strategies are used to improve outcome in patients with HNSCC. Selecting appropriate treatment strategies and predicting patients’ prognoses remain difficult for clinicians, despite careful evaluation of clinical factors, TNM staging, and anatomic subsite. Identification of novel pretreatment imaging biomarkers that potentially predict long-term outcome is of great interest. PET/CT standardized uptake value (SUV) measurements are reproducible imaging biomarkers that have diagnostic and prognostic value in HNSCC in general (5,6) and in oral and oropharyngeal SCC specifically (7). Recently, 18F-FDG metabolic tumor volume (MTV) and total glycolytic activity (TGA) have been reported as additional diagnostic and prognostic imaging biomarkers in various human solid tumors (8–11).

The purpose of this study was to establish whether MTV and TGA add prognostic information to clinical staging in patients with oral cavity and oropharyngeal SCCs.
MATERIALS AND METHODS

Study Population

We conducted a retrospective study of patients with histologically proven oral and oropharyngeal SCC who underwent PET/CT between 2007 and 2009 at a single institution. The Institutional Review Board approved this Health Insurance Portability and Accountability Act–compliant study, and informed consent was waived. All patients who had a biopsy-proven oral or oropharyngeal SCC and who had a baseline PET/CT examination at our institution were included in the study. All patients who had undergone local or systemic therapy or surgical intervention before the baseline PET/CT examination were excluded. All surviving patients had at least a 12-mo follow-up. Forty-five patients (33 men and 12 women; age range, 39–91 y) were eligible for inclusion in the study.

PET/CT Protocol

All PET/CT studies were performed on a Discovery STE 16 (GE Healthcare) PET/CT scanner according to the institutional standard clinical protocol. For all patients, a dedicated head and neck protocol was instituted. Patients were scanned from skull base to aortic arch with the arms down, and then from clavicle to mid thigh with the arms up. The average patient blood glucose level was 105 ± 21.8 mg/dL. Patients were injected with an average of 525.4 ± 111 MBq (14.2 ± 3 mCi) of 18F-FDG and incubated for an average period of 111 ± 22 min.

The dedicated head and neck imaging protocol consisted of 2-dimensional PET scans obtained from the skull base to the arch of the aorta with a 30-cm field of view and 128 × 128 matrix. The emission scan lasted for 5 min per bed position. The remainder of the body (down to the mid thighs) was imaged using a weight-emission scan lasted for 5 min per bed position. The remainder of the aorta with a 30-cm field of view and 128 · 128 matrix. The SUV was normalized to lean body mass, mean SUV (SUVmean), MTV, and TGA from PET. Both SUVmax and SUVmean were measured from the tumor volume with 18F-FDG uptake, normalized to lean body mass, MTV, and TGA from PET. Both SUVmax and SUVmean were measured from the tumor volume segmented by a gradient-based method (PET Edge). MTV was defined as the tumor volume with 18F-FDG uptake segmented by the PET Edge method and fixed-threshold methods at 38%, 50%, and 60% of SUVmax (12). The TGA was defined as (MTV) × (SUVmean). The commercially available MIMvista software analysis suite (MIM Software Inc.) included a contouring suite for radiation therapy planning and PET/CT fusion suite. The edges of the primary tumor were automatically calculated and outlined in both segmentation methods. Once the primary tumor (target) was segmented, SUVmax, SUVmean, MTV, and TGA were automatically calculated by the MIMvista software.

Segmentation Methods

Gradient Segmentation Method. The gradient segmentation of tumor volume depends on the identification of tumor based on a change in count level at the tumor border. Complex methods have been previously proposed, including denoising, deblurring, gradient estimation, and watershed transformation (13). The gradient segmentation method used in MIMvista (version 4.1) has been previously described (12) and is simple and easy to use. It calculates spatial derivatives along the tumor radii and then defines the tumor edge on the basis of derivative levels and continuity of the tumor edge. The software relies on an operator-defined starting point near the center of the lesion. As the operator drags out from the center of the lesion, 6 axes extend out, providing visual feedback for the starting point of gradient segmentation. Spatial gradients are calculated along each axis interactively, and the length of an axis is restricted when a large spatial gradient is detected along that axis. The 6 axes define an ellipsoid that is then used as an initial bounding region for gradient detection (12). The reader added regions until he was visually satisfied that the entire primary tumor was included in the contour.

Fixed-Percentage Threshold Segmentation Method. The contouring method using a fixed SUVmax threshold relies on including all voxels that are greater than a defined percentage of the maximum voxel within an operator-defined sphere (in this study, 38%, 50%, and 60% of SUVmax). Cross-sectional circles were displayed in all 3 projections (axial, sagittal, and coronal) to ensure 3-dimensional coverage of the primary tumor (12).

Outcome Endpoints

The primary endpoints were to establish whether the exploratory imaging markers, MTV and TGA, added prognostic information—overall survival (OS) and progression-free survival (PFS)—to clinical staging. OS is defined as the time from therapy initiation to death or to most recent inpatient or outpatient follow-up through March 31, 2011. PFS is defined as the time from initiation of therapy to the first documented progression at the primary site, at regional nodes, or at distant metastatic sites through March 31, 2011. Death from the primary cancer without a documented site of recurrence or progression or death from an unknown cause is considered death from local regional disease. Electronic medical records, imaging records, office visits at our institution, and the Social Security Administration Web-based mortality database (14) were used to establish the OS and PFS.

Statistical Methods

We present our summary statistics as the mean ± SD for continuous variables or frequency and percentage for categoric variables. The association between clinical variables, imaging parameters, and survival was examined with Cox proportional hazards regression. Crude and adjusted Cox regression relative risks were estimated. Analyses were also performed with a bootstrap method using 1,000 simple bootstrap samples with a 95% confidence interval (CI). Multicollinearity between variables was
established using the Pearson correlation coefficient. We also used the Pearson correlation coefficient to establish the relationship between different segmentation methods, and we used Bland–Altman analysis between the 2 best-correlated segmentation methods for MTV and TGA to establish the reliability of measuring MTV and TGA by these methods. Receiver operating characteristic analysis was used to determine area under the curve (AUC) to estimate the accuracy and predictive ability of various imaging biomarkers. Kaplan–Meier curves with median cutoff points for MTV and TGA were generated for survival analysis and compared using the Mantel Cox log rank test. We used the Prism 5 (GraphPad Software Inc.) and SPSS (version 19; SPSS Inc.) statistical packages for all analyses. All hypothesis tests were 2-sided, with a significance level of 0.05.

RESULTS

Patients

Forty-five patients met the eligibility criteria. Patient characteristics, including sex, age, ethnicity, pack-years, primary site, American Joint Committee on Cancer (AJCC) stage, tumor grade, months of follow-up, disease progression, and OS, are listed in Table 1. Thirty-three patients were men and 12 were women; the average age of patients was 62.1 y, with an age range of 39–91 y. The distribution of tumors by AJCC stage was 9% stage I (n = 4), 18% stage II (n = 8), 11% stage III (n = 5), and 62% stage IV (n = 28). Mean follow-up was 23.7 mo (range, 1.7–46.5 mo). Twenty-five patients had surgery, 11 had concurrent chemoradiotherapy, 2 had radiotherapy, 2 had chemotherapy, and 5 had no treatment for oral or oropharyngeal cancer.

Sixteen patients (36%) died during follow-up, and 12 patients (27%) experienced disease progression during follow-up. A total of 20 patients died or had progression of disease during the follow-up period.

Cox Proportional Hazards Univariate and Multivariate Analysis

Cox proportional hazards regression was performed to assess the impact of clinical and imaging parameters on the likelihood of predicting OS or PFS for patients with oral and oropharyngeal SCC. The initial model contained 5 clinical and 3 imaging variables (AJCC stage, smoking pack-years, age, sex, tumor grade, SUVmax, primary tumor MTV, and primary tumor TGA). MTV, TGA, and tumor grade were the most statistically significant parameters (P < 0.05) associated with PFS or OS (Table 2) in the univariate Cox regression analysis. There was multicollinearity between MTV and TGA (r = 0.94; r² = 0.89), as expected, a priori, given that TGA is compounded using MTV and SUVmean. MTV and TGA were incorporated in 2 separate models adjusting for all other parameters. MTV (P = 0.006) and TGA (P = 0.02) were the only statistically significant parameters at baseline associated with time-dependent event-free (defined as either OS or PFS) survival.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B coefficient</th>
<th>Hazards ratio</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJCC stage</td>
<td>0.248</td>
<td>1.28</td>
<td>0.79–2.07</td>
<td>0.31</td>
</tr>
<tr>
<td>Smoking pack-years</td>
<td>0.007</td>
<td>1.01</td>
<td>0.92–1.02</td>
<td>0.37</td>
</tr>
<tr>
<td>Age</td>
<td>0.019</td>
<td>1.02</td>
<td>0.98–1.06</td>
<td>0.37</td>
</tr>
<tr>
<td>Sex</td>
<td>0.406</td>
<td>1.50</td>
<td>0.50–4.50</td>
<td>0.47</td>
</tr>
<tr>
<td>Tumor grade</td>
<td>−0.776</td>
<td>0.56</td>
<td>0.19–1.63</td>
<td>0.04</td>
</tr>
<tr>
<td>SUVmax</td>
<td>0.019</td>
<td>1.02</td>
<td>0.96–1.09</td>
<td>0.56</td>
</tr>
<tr>
<td>MTV</td>
<td>0.051</td>
<td>1.05</td>
<td>1.02–1.09</td>
<td>0.002</td>
</tr>
<tr>
<td>TGA</td>
<td>0.003</td>
<td>1.00</td>
<td>1.00–1.01</td>
<td>0.006</td>
</tr>
</tbody>
</table>
To address the limited sample size, a bootstrap procedure was used. MTV and TGA were the only parameters that were significantly associated with an event-free survival after bootstrap simulation of 1,000 samples. Proportional hazards assumptions were tested for both MTV and TGA with Schoenfeld residuals and were non-significant, with zero curves.

The final models consisted of AJCC stage (a priori) and MTV in model I and AJCC stage (a priori) and TGA in model II. The model fit was significantly better when MTV was added to AJCC stage in model I ($\chi^2$ value changed from 1.16 to 6.71; $P = 0.01$) and when TGA was added to AJCC stage in model II (the $\chi^2$ value changed from 1.16 to 4.37; $P = 0.04$).

### SUVmax, MTV, and TGA Segmentation Methods

The primary tumor MTV measured by the PET Edge method strongly correlated with MTV as measured by fixed SUVmax threshold segmentations (38% SUVmax $r = 0.98$; 50% SUVmax $r = 0.96$; 60% SUVmax $r = 0.94$; $P < 0.0001$) (Fig. 1). Bland–Altman analysis between the PET Edge MTV and 38% SUVmax MTV resulted in a bias of 0.49, with an SD of 2.97 (Fig. 2). There was also a strong correlation between the TGA as measured by PET Edge and by fixed SUVmax threshold segmentations (38% SUVmax $r = 0.97$; 50% SUVmax $r = 0.97$; 60% SUVmax $r = 0.94$; $P < 0.0001$) (Fig. 2). Bland–Altman analysis between the PET Edge TGA and 38% SUVmax TGA resulted in a bias of $-0.67$, with an SD of 16.5 (Fig. 2). There was only a fair correlation between SUVmax and MTV or TGA, with correlation coefficients varying between 0.47 (PET Edge segmentation) and 0.44 (60% SUVmax segmentation).

### Primary Tumor SUVmax, SUVmean, and Outcome

The mean SUVmax measurements for those alive and deceased were 14.31 (95% CI, 11.9–16.8) and 14.63 (95% CI, 10.9–18.4), respectively ($P = 0.9$). The mean SUVmean measurements for those alive and deceased were 8.3 (95% CI, 6.3–9.8) and 8.3 (95% CI, 6.1–10.4), respectively ($P = 1.0$). The AUC for the primary tumor SUVmax and SUVmean, calculated by the gradient-based method, for predicting OS were 0.50 (95% CI, 0.32–0.68) ($P = 0.97$) and 0.53 (95% CI, 0.32–0.72) ($P = 0.74$), respectively.

### MTV and Outcome

The AUC for predicting OS with primary tumor MTV by gradient-based, 38%, 50%, and 60% SUVmax methods were 0.71 ($P = 0.02$), 0.70 ($P = 0.03$), 0.69 ($P = 0.04$), and 0.71 ($P = 0.02$), respectively. Because there was no significant difference in the AUC for different segmentation methods, we decided to use the gradient-based method for survival analysis. The median cutoff point of 7.7 mL for PET Edge primary tumor MTV was predictive of time-dependent OS (log rank $P = 0.04$) (Fig. 3).

### TGA and Outcome

The AUC for predicting OS with primary tumor TGA using PET Edge, 38%, 50%, and 60% SUVmax methods

---

### TABLE 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B Coefficient</th>
<th>Adjusted hazards ratio</th>
<th>95% CI</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJCC stage</td>
<td>0.37</td>
<td>1.45</td>
<td>0.74–2.84</td>
<td>0.28</td>
</tr>
<tr>
<td>Smoking pack-years</td>
<td>0.015</td>
<td>1.02</td>
<td>1.00–1.05</td>
<td>0.18</td>
</tr>
<tr>
<td>Age</td>
<td>0.021</td>
<td>1.02</td>
<td>0.96–1.09</td>
<td>0.49</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.49</td>
<td>0.64</td>
<td>0.20–2.08</td>
<td>0.46</td>
</tr>
<tr>
<td>Tumor grade</td>
<td>-0.66</td>
<td>0.52</td>
<td>0.21–1.25</td>
<td>0.14</td>
</tr>
<tr>
<td>SUVmax</td>
<td>-0.032</td>
<td>0.97</td>
<td>0.89–1.06</td>
<td>0.51</td>
</tr>
<tr>
<td>MTV</td>
<td>0.055</td>
<td>1.06</td>
<td>1.01–1.10</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Because of multicollinearity ($r = 0.94$) between MTV and TGA, TGA was not included in model I.

### TABLE 4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B coefficient</th>
<th>Adjusted hazards ratio</th>
<th>95% CI</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJCC stage</td>
<td>0.421</td>
<td>1.52</td>
<td>0.79–2.82</td>
<td>0.27</td>
</tr>
<tr>
<td>Smoking pack-years</td>
<td>0.013</td>
<td>1.01</td>
<td>1.00–1.03</td>
<td>0.07</td>
</tr>
<tr>
<td>Age</td>
<td>0.012</td>
<td>1.01</td>
<td>0.95–1.07</td>
<td>0.77</td>
</tr>
<tr>
<td>Sex</td>
<td>0.297</td>
<td>1.35</td>
<td>0.42–4.37</td>
<td>0.67</td>
</tr>
<tr>
<td>Tumor grade</td>
<td>-0.814</td>
<td>0.44</td>
<td>0.18–1.09</td>
<td>0.23</td>
</tr>
<tr>
<td>SUVmax</td>
<td>-0.046</td>
<td>0.96</td>
<td>0.87–1.06</td>
<td>0.47</td>
</tr>
<tr>
<td>TGA</td>
<td>0.004</td>
<td>1.00</td>
<td>1.00–1.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Because of multicollinearity ($r = 0.94$) between MTV and TGA, MTV was not included in model II.
were 0.67 \( (P = 0.06) \), 0.68 \( (P = 0.05) \), 0.61 \( (P = 0.2) \), and 0.66 \( (P = 0.07) \), respectively. The median cutoff point of 55 g for PET Edge primary tumor TGA was predictive of time to OS (log rank \( P = 0.08 \)) (Fig. 4), though the result was not statistically significant.

**DISCUSSION**

Our results established MTV and TGA as potential prognostic markers for event-free survival. These imaging markers add prognostic information to AJCC staging. In addition, MTV and TGA had greater AUC in the receiver operating characteristic analysis than did SUVmax and are thus more accurate predictors of OS. These exploratory findings need further validation in larger cohort studies because our study has a limited number of patients.

The TNM system \((15)\) published by the International Union Against Cancer and the AJCC is the accepted standard method to categorize patients into prognostic groups. However, limitations exist in its ability to predict treatment response in head and neck cancer \((16)\).

PET using the radiotracer \(^{18}\)F-FDG is widely applied as an imaging modality for head and neck cancers. The SUVmax has been used to evaluate the likelihood of aggressive disease, metabolic response to therapy, early detection of disease recurrence, and patient outcome in head and neck cancers \((6,17,18)\). Although convenient to measure and widely used, SUVmax has limitations. It is a single-pixel value representing the most intense \(^{18}\)F-FDG uptake in the tumor and may not be an adequate surrogate marker representing tumor biology. In addition, SUVmax variability increases as lesion matrix size increases and patient size increases \((19)\). There is also a statistical bias of SUVmax in that larger lesions are more intense because of more available counts \((20)\).

A fundamental question in assessing tumor biology is whether it is the total MTV or the maximally active portion of the tumor that is more important in predicting outcome. Though the maximally active portion of the tumor may represent the more aggressive part of the tumor, it may also respond to treatment more effectively and thus have lesser impact on outcome. Our results suggest that MTV and TGA are more accurate predictors of outcome than is the maximally active portion of the tumor as represented by SUVmax. MTV and TGA have been demonstrated to predict therapy assessment, pathologic parameters, and outcome in early studies of various human solid tumor \((8,21–25)\).

MTV defines the volume of the tumor on the basis of the distribution of metabolic activity instead of the traditional x-ray or CT densities that depend on electron density of the target. TGA goes a step further and effectively weighs this volume by its mean metabolic activity. Hence, a large TGA may reflect a small volume with high metabolic activity (high SUVmean) or a large volume with a lower metabolic activity. Both MTV and TGA are potentially better surrogate imaging markers for tumor biology than SUVmax or tumor diameter.

Many methods of segmentation have been proposed to measure MTV and TGA. The visual-analysis gradient-based edge-detection method, the fixed-SUVmax-threshold...
method, and the adaptive threshold method based on signal-to-background ratio were investigated in head and neck cancers (26). The gross tumor volume is influenced by the segmentation method used. We used fixed-SUVmax-threshold methods and a gradient-based edge-detection method in our study. There was excellent correlation and no significant difference in the accuracy of these methods in predicting OS. A recent study has also suggested that MTV contoured by visual analysis with manual contouring is a predictive parameter for local control (27), but this study did not demonstrate an association between MTV segmented using other methods and local control. From our results, it appears that the MTV segmented using various methods is more important in prognostic outcome than are the methods used to derive the volume, in oral and oropharyngeal cancers.

One of the main limitations of our study was the limited number of subjects because this was an exploratory study. In addition, we did not include human papilloma virus (HPV) status in our models. Recent literature suggests HPV/p16 is a prognostic factor in patients with oral and oropharyngeal SCC (28). HPV/p16 status has routinely been determined for oral and oral cavity SCC at our institution only since the beginning of 2010 and thus was not available for this retrospective study from 2007 to 2009. Our results apply only to oral and oropharyngeal cancer because other head and neck cancers, such as thyroid cancer, may demonstrate variable 18F-FDG uptake when SUVmax and MTV are prognostic factors (29). We did not investigate the CT volume of the primary tumor in our study because the performance of CT segmentation algorithms may suffer in soft-tissue tumors in which the background soft-tissue radiodensity is similar to tumors, especially when intravenous contrast is not used in all patients.

We foresee that MTV and TGA will become valuable imaging biomarkers in many human solid tumors, especially for therapy assessment with neoadjuvant chemotherapy and concurrent chemotherapy, and as prognostic biomarkers for short- to intermediate-term survival outcomes, adding value to current clinical staging. Prediction models for short- to intermediate-term survival outcome incorporating these imaging biomarkers will enable physicians to modify management decisions in the future, either to treat patients more aggressively to improve outcome or to treat patients palliatively to save health care costs.

CONCLUSION

Gradient-based segmentations of primary tumor MTV and TGA provide prognostic information in addition to AJCC clinical stage and are potential 18F-FDG imaging markers for survival in patients with oral and oropharyngeal SCC. These imaging markers need validation in larger cohort studies and have the potential to be useful in treatment stratification and clinical care of patients in the future.

DISCLOSURE STATEMENT

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

ACKNOWLEDGMENT

Rathan Subramaniam was supported by a GE-AUR research fellowship. No other potential conflict of interest relevant to this article was reported.

REFERENCES


**MTV, TGA, and Oral and Oropharyngeal SCC • Dibble et al.**