Clinical Applications of Tumor Volume Measurements for Predicting Outcome in Patients with Squamous Cell Carcinoma of the Upper Aerodigestive Tract

Suresh K. Mukherji, Ilona M. Schmalfuss, Jonas Castelijns, and Anthony A. Mancuso

Gross tumor volume (GTV) at the primary site, as derived from pretreatment CT findings, can help predict local control of squamous cell carcinoma at different head and neck subsites after treatment with nonsurgical organ preservation. Local recurrence is more likely with large tumors than with small lesions in the same anatomic subsite, and GTV is often more strongly associated with local control than is tumor stage. This review discusses tumor volume calculation—technique, current literature, and potential clinical applications—and aims to help the reader to understand the role of GTV calculations and to integrate this knowledge into clinical practice.

The treatment for head and neck squamous cell carcinoma (HNSCCA) depends on various factors. Physicians and patients often have the option of both surgical and nonsurgical treatment for tumors of similar stages. The final decision is based on an overall assessment of the risk of local recurrence, 5-year survival rates, treatment-associated morbidity, and institutional and patient preference.

In general, nonsurgical organ preservation is associated with less morbidity and improved functional status after definitive treatment. Proper patient selection is an important factor in weighing the different treatment options for HNSCCA. Cross-sectional imaging has been shown to be more accurate than physical examination for assessing the size and extent of the primary tumor in many instances (1–3). As a result, pretreatment imaging has become an accepted adjunct to staging and, as stated in the Manual for Staging of Cancer from the American Joint Committee on Cancer (AJCC), “Any diagnostic information which contributes to the overall accuracy of the pretreatment assessment should be considered in clinical staging and treatment planning” (4).

Over the past several years, numerous studies have shown that the gross tumor volume (GTV) at the primary site, as derived from pretreatment CT scans, can predict local control of squamous cell carcinoma at different head and neck subsites after treatment with nonsurgical organ preservation (5–24). The likelihood of local recurrence is higher in large-volume tumors than in small-volume lesions arising in the same anatomic subsite (5–24). Threshold volumes have been identified for carcinomas of the T3 glottis and in the nasopharyngeal, oropharyngeal, supraglottic, and pyriform sinuses (5–16). Tumors with a GTV below these thresholds have rates of local control significantly higher than those of tumors with volumes above the thresholds. Further investigations have identified a direct relationship between GTV and the rate of local control for nasopharyngeal, oropharyngeal, and supraglottic carcinomas, where increasing GTV is associated with a progressively lower likelihood of local control (5–24). Several investigators have also shown that the association between GTV and local control is stronger than that between tumor (T) stage and local control (6, 7, 14–24).

This review discusses tumor volume calculation: its technique, current literature, and potential clinical applications. It is our hope that this review article will help the reader understand the role of GTV calculations in HNSCCA and permit them to integrate this knowledge into their clinical practice.

Technique

GTV calculations are based on contouring the area of visible disease on cross-sectional CT or MR imaging studies. Tumor contouring is not a new or unique technique. Radiation oncologists routinely perform tumor contouring on a daily basis by using 3D conformal therapy or intensity-modulated radiation therapy (RT). Individuals experienced in the interpretation of cross-sectional images of the extracranial head and neck region can obtain accurate and reproducible GTV measurements of HNSCCA (25).

GTV is derived from an abnormal soft-tissue mass at the primary site on a cross-sectional image. Contouring can be performed by using a hard- or soft-copy format. Hard-copy images must eventually be transferred to a digital format for contouring. This can be performed by using several commercially available techniques. If the study is in its original digital format, tumor contouring can be performed by using a variety
The exact method for accomplishing this varies with the different software and hardware packages (5). The study must then be corrected for magnification (5–16, 22). This step is typically performed with correction factors unique to each postprocessing unit. GTV must also be corrected for respiratory misregistration before tumor volumes are calculated. To do so, each image of the study must be evaluated to ensure that different anatomic sections are identified at different table positions. Similar anatomic levels at different table positions indicate respiratory misregistration, which will artifactually increase the tumor volume. To do so, each image of the study must be evaluated to ensure that different anatomic sections are identified at different table positions. Similar anatomic levels at different table positions indicate respiratory misregistration, which will artifactually increase the GTV. The repeated image should be eliminated from the final contour analysis. The exact method for accomplishing this varies with the different software and hardware packages (5–16, 22).

Either CT or MR imaging can be used for GTV measurements. Most reports in the published literature used CT scans. MR imaging can also be used; however, to our knowledge, no studies have determined the optimal sequences for contouring different primary sites. A nonenhanced T1-weighted sequence may be optimal because tumors have intermediate signal intensity, and the margins can be distinguished from adjacent fat. Contrast enhancement may possibly reduce the conspicuity of the enhancing tumor from the surrounding enhancing fat or muscle. HNSCCA typically has intermediate signal intensity on T2-weighted images, and the tumor margins may be difficult to distinguish from fat on rapid T2-weighted images that do not suppress fat.

Because of the relatively small size of the structures and their similarity in attenuation, it is difficult to differentiate between deep spread of tumor from peritumoral edema in the extracranial neck, especially the larynx. As a result, no attempt is made to differentiate tumor and edema when the primary site is contoured.

Results have also suggested that GTV can be calculated from PET scans of cervical disease. Miller et al (21) demonstrated that GTV can be accurately measured from PET studies and that GTV is associated with disease-free survival in cervical cancer treated with RT. However, this technique for GTV measurement is uncommonly used.

Previous authors have suggested that volume measurements obtained in this fashion have an intrinsic error rate (5–24). It is generally believed that GTVs calculated in this manner are within 10–20% of the actual tumor volumes (5–24). Furthermore, the exact size of a tumor cannot be reliably measured at surgery, because tumors often decrease in size immediately after resection owing to the effects of specimen drying and pathologic fixatives (5–24). This relatively small error does not appear to affect the validity of the measurements, as the technique is performed in a consistent manner. In fact, if the errors result from a nonsystematic, random process, they diminish the apparent association between GTV and local control, producing a conservative bias. Therefore, the intrinsic measurement error rate associated with measuring GTV may actually cause underestimation of the statistical strength of the published results.

**Primary Sites**

**Larynx**

Several studies have shown that certain CT and MR imaging parameters may help predict the local outcome of laryngeal squamous cell carcinoma treated with definitive RT (5, 9, 15, 26–28). Important imaging parameters include GTV, cartilage abnormalities, and involvement of specific subsites. In most reports, GTV appears to be an important predictor of local control after definitive RT of laryngeal cancer. Parameters reflecting cartilage abnormalities and deep invasion of the laryngeal tissue also appear to be factors that help in predicting local control.

**Supraglottis.**—Specific imaging parameters appear to be strong predictors of local and locoregional outcomes of supraglottic carcinomas treated with definitive RT. GTV appears to be the strongest independent predictor of local failure after RT (5, 9, 15). Pretreatment CT measurements of GTV permit the stratification of local control in patients with supraglottic squamous cell carcinoma treated with RT alone. Mancuso et al (5) reported a local control rate of 89% in tumors smaller than 6 mL and 52% when volumes are ≥6 mL (Fig 2). Kraas at al suggested that the threshold volume might be higher (8 mL). The specific threshold may vary; however, what is of greater importance is the linear relationship between tumor volume and local control. Mancuso et al demonstrated a direct linear relationship between increasing GTV and decreasing local control for supraglottic carcinomas treated with definitive RT (5). On the basis of this relationship, we can obtain estimates of cure for individual tumors that are more accurate...
than general threshold measurements (5, 17). Hermans et al (6) suggested that the degree of involvement of the paraglottic space at the level of the true vocal cord is also a predictor of local recurrence. Involvement of the paraglottic space is defined as obscuration of the paraglottic fat by the adjacent tumor. Increasing tumor burden in this space is associated with a lower local control rate after RT. This effect may also be due to a larger gross tumor burden in larger-volume tumors.

Glottis.—Several imaging-based parameters have been shown to be predictive of local outcome in glottic cancer treated with RT (15). GTV is directly associated with local recurrence, with larger lesions associated with a higher recurrence rate. Other factors, including cartilage abnormalities (ie, sclerosis and invasion), involvement of the anterior commissure, subglottic extension, invasion of the pre-epiglottic space, and involvement of the paraglottic space at the level of the true vocal cords, are also significantly associated with the local recurrence rate. In multivariate analysis, however, only the degree of involvement of the paraglottic space (at the level of the true vocal cords) and the pre-epiglottic space are independent predictors of local recurrence (15). The T category is significantly correlated with local outcome in actuarial analysis but not in multivariate analysis (15). Overall, GTV and other factors may be the most accurate means of risk stratification, which is relatively simple when performed by adequately trained interpreters of the imaging studies.

T3 Glottis.—Pretreatment CT results can be used to stratify patients with T3 glottic carcinoma into groups according to the likeliness of local control with definitive RT. Local control rates for these tumors can be improved by using a CT-based tumor profile; the ideal profile for a radiocurable T3 glottic larynx carcinoma is a volume $<3.5$ mL with normal or a single sclerotic laryngeal cartilage (10, 11). GTV is a significant predictor of local control. For tumors $\leq 3.5$ mL, local control is achieved in $\geq 85\%$ of patients, whereas for tumors $>3.5$ mL, local control is achieved in $<35\%$ (10, 11). (Fig 3) Tumor score, as a measure of anatomic extent (ie, the number anatomic sites, such as subglottic, paraglottic, and ventricle), is also a significant predictor of local control (10). Tumor scores ranged from 1 to 8, and the local control rate for tumors assigned a low score ($<5$) is $86\%$ (11) compared with $14\%$ for tumors assigned a high tumor score ($>5$) (11). A significant decrease in the local control rate is observed for cancers involving the paraglottic space at the level of the false vocal cord, the arytenoid cartilage (on CT scans), and the interarytenoid region. In 12 patients, CT scans of both the ipsilateral arytenoid and the adjacent cricoid cartilage showed sclerosis. These patients had a significantly lower local control rate.}

![Fig 2. Patient with an 8.4-mL supraglottic carcinoma. Pre-RT enhanced CT scan obtained at the lower supraglottic level shows a mass (arrow) infiltrating the paraglottic space on the left. Despite the large tumor volume, the patient elected definitive RT, which failed. The patient then underwent successful salvage resection of the supraglottic larynx (supraglottic laryngectomy).](image)

![Fig 3. Pre-RT enhanced CT scans at the level of the true vocal cords in a patient with a T3 glottic carcinoma and sclerosis of the arytenoid and cricoid cartilages on the left. Calculated tumor volume was 4.3 mL. The patient elected curative RT alone, though the likelihood of control was $<20\%$ based on the CT risk profile. Because of the high risk of recurrence, close post-treatment surveillance with CT was chosen. Treatment failed, and total laryngectomy was performed.](image)
decreased rate of local control (33%) (10), which represents a high likelihood of microscopic invasion of the cartilage (29). However, several investigators suggest that patients with T3 (and even T4) glottic cancer and isolated sclerosis of the laryngeal cartilage can still be cured with definitive RT (29).

T2 and T2 Glottis.—Mukherji et al (30) did not find a statistically significant relationship between the GTV or tumor score and the outcome of T2 glottic tumors treated with definitive RT. Like low-volume supraglottic and T3 glottic carcinomas, T2 glottic squamous cell carcinomas are likely (82%) to be controlled with definitive RT. Failure to control the primary tumor is attributable to factors other than volume that might not be detectable by imaging methods. These factors may include various tumor-host biologic factors, such as certain gene mutations, the status of the host’s immune system, and some tumors that appear to be inherently more aggressive than others. However, Murakami et al (31) demonstrated contrary findings. In a study of T1 and T2 glottic lesions treated with definitive RT, univariate analysis revealed that GTV and other imaging-based parameters were predictive of local control. Imaging findings of involvement of the anterior commissure, ventricle, and thyroid cartilage are significant prognostic factors. Only GTV (>1 mL) and a tumoral location directly adjacent to the thyroid cartilage are independent predictors at multivariate analysis. Lesions separate from the thyroid cartilage have a 95% probability of local control, whereas the lesions adjacent to the cartilage have only a 42% control rate. Because the paraglottic space is directly adjacent to the thyroid cartilage, the prognostic findings associated with tumor directly adjacent to this cartilage may identify a risk for microscopic invasion of the cartilage, similar to the sclerosis factor for T3 glottic tumors.

**Pyriform Sinus**

CT results appear to be predictive of outcomes in patients with T1 or T2 squamous cell carcinomas of the pyriform sinus treated with definitive RT. The imaging findings that appear to be most predictive of local control are GTV and the extent of tumor at the apex of the sinus. Tumors with a GTV <6.5 mL have an 89% likelihood of local control, whereas tumors with GTV >6.5 mL have a local control rate of 25% (P < .05). When involvement of the apex is divided into minimal (<10 mm in largest dimension) and bulk (>10 mm in largest dimension) groups, the local control rates are 83% and 25%, respectively (Fig 4). On the basis of these two criteria, an imaging-based risk profile is calculated. T1 or T2 pyriform sinus tumors with GTV <6.5 mL and minimal apical disease have a 94% likelihood of local control when treated with definitive RT. Tumors with either GTV ≥6.5 mL or bulk apical disease have a <50% local control rate when treated with RT alone (32).

**Nasopharyngeal Carcinoma**

Several studies have demonstrated that GTV is an important prognostic factor for patients with nasopharyngeal carcinoma treated with definitive RT or combined chemotherapy and RT. Chua et al (7) reported that GTV is predictive of local control in patients with nasopharyngeal carcinoma treated with RT. Five-year local control rates are 88% in tumors with GTV ≤20 mL, 80% in those with GTV of 20–60 mL, and 56% in tumors with GTV >60 mL. In fact,
GTV is superior to the T stage as a predictor of local control (7). These data suggest that tumors with a GTV <20 mL can be successfully treated with RT alone, thereby sparing the patient the morbidity of adjuvant chemotherapy. However, the morbidity associated with the addition of chemotherapy is likely warranted in patients with tumor volumes >20 mL, and certainly those with volumes exceeding 60 mL, because of their higher failure rates.

Similar to Chua et al, Chang et al (23) reported that GTV >60 mL is associated with local recurrence rates and disease-free survival rates significantly worse than those associated with GTV <30 mL. Chang et al (24) also reported that GTV is superior to the T stage for predicting local control in patients with nasopharyngeal carcinoma treated with combined chemotherapy and RT. The authors recommend that GTV be incorporated into future refinements of current staging systems.

**Oropharynx**

The AJCC staging criteria is based on a clinical estimate of the size of the tumor, where 2 and 4 cm are the threshold values used to upstage the primary site. Several groups (18, 33, 34) have studied the association of GTV and local control in squamous cell carcinoma treated with definitive RT. No significant correlation was found between GTV and local control after surgical or nonsurgical treatment (32–34). Hermans et al (35) demonstrated that GTV is significantly correlated with local recurrence when all treated oropharyngeal carcinomas are considered. However, GTV is not predictive of local control in the T2, T3, and T4 subcategories (33). Nathu et al (34) reported a large variation in GTV within a given T stage, which was found to be the most important variable associated with local control. GTV was only marginally associated with local control (P = .10). In a surgical series, Keberle et al (18) found no association between GTV or T stage and local control for patients treated with primary resection. The only anatomic factor correlated with local control was extension across the midline.

It is unclear why the relationship between GTV of oropharyngeal squamous cell carcinomas is less pronounced than that identified in nasopharyngeal, hypopharyngeal, and laryngeal carcinomas. One potential cause may be the fact that certain oropharyngeal subsites, such as the faucial (palatine) tonsil, are intrinsically more radiosensitive than other subsites within or outside the oropharynx. Primary subsites in the oropharynx and subsites directly adjacent to it can have markedly varying radiosensitivities, despite the fact that they are separated by only a few centimeters (eg, faucial tonsil vs tongue base, oral tongue vs floor of the mouth). This observation would result in the association between GTV and local control being less pronounced. Another potential explanation may be the growth pattern of the cancer and its radiosensitivity (34). Fletcher and Hamberger (35) suggested that exophytic cancers are more radiosensitive than infiltrating tumors. Variable degrees of deep tissue infiltration by tonsillar carcinomas may possibly reduce the association between GTV and local control.

**Treatment Monitoring at the Primary Site**

An important, often-mentioned concern is the best time to perform post-treatment imaging in attempts to determine the response at the primary site after nonsurgical organ-preservation therapy. Accurate assessment helps in identifying patients with successful responses from those in whom treatment fails and who may benefit from early salvage surgery. Previous studies have suggested that changes in tumor size can predict treatment response. Post-treatment imaging appears to be most reliable for predicting primary-site responses of squamous cell carcinoma when performed approximately 3–4 months after the completion of definitive RT. Complete radiologic resolution of the lesion on the post-treatment study compared with pretreatment study strongly suggests a successfully controlled primary site. Tumors that fail treatment are characterized by a <50% reduction in size following treatment (36, 37). A partial response defined as a persistent mass that has reduced in size by >50% but is still present is deemed “indeterminate” and requires further imaging and close clinical observation. On follow-up studies, interval enlargement of the indeterminate mass is suggestive of recurrent disease or tumor necrosis. However, stability of the mass over a 2-year period suggests fibrosis and scarring. Imaging modalities aimed at measuring metabolic activity, such as PET or thallium SPECT, may prove beneficial in this differentiation. The value of metabolic imaging agents (eg, 2-[fluorine 18]-fluoro-2-deoxy-D-glucose and rhallium-201) in differentiating recurrent disease from radiation changes is promising (38–58).

**Lymph Nodes**

**Association with Locoregional Control and Survival**

Nodal tumor volume (NTV) is defined as the sum of the volumes of all metastatic lymph nodes, as determined with CT scanning. The association between NTV of metastatic squamous cell carcinoma, locoregional, and survival of various subsites is under active investigation. For nasopharyngeal carcinoma, Chua et al (7) reported that NTV is associated with nodal failure and the disease specific survival rate. Their nodal failure and disease-free survival rates were significantly worse in patients that had large tumors. NTV was directly associated with the node (N) stage, and the association between large NTV and lower survival probably reflects the known association between an advanced N stage and reduced survival. No association was found between NTV and distant metastases and local control for nasopharyngeal carcinoma (7).

For oropharyngeal squamous cell carcinomas, NTV is significantly associated with regional control.
In general, patients with larger NTV have a higher rate of locoregional recurrence. However, patients with NTV of 0.5–14.5 mL have better regional control than patients with NTV <0.5 mL; the latter group may have tumors that are intrinsically and biologically more aggressive. Hermans et al (33) did not investigate the association between NTV and survival.

**Treatment Monitoring**

The increasing role of concomitant chemotherapy and RT (chemoradiation) for organ-preservation protocols and for the treatment of unresectable HNSCCA have resulted in investigations to determine if CT scanning can be used noninvasively to predict the response of lymph nodes after therapy. Accurate CT prediction of tumor in lymph nodes after chemoradiation may obviate planned neck dissection in patients with pretreatment nodal metastases. Labadie et al (59) reported that a post-treatment NTV reduction in an individual lymph node by 90%, compared with the pretreatment NTV, is correlated with negative pathologic findings for tumor in that node after planned neck dissection performed 6 weeks after chemoradiation. These initial results suggest that changes in NTV can predict nodal response. Therefore, an argument can be made that a favorable imaging response can preclude planned neck dissection in a recently treated neck, thereby reducing patient morbidity. However, the effect of such an alteration in standard treatment on overall survival needs to be investigated.

**Discussion**

There appear to be imaging parameters specific to certain primary head and neck sites that can be used to predict responses to nonsurgical organ-preservation therapy. GTV appears to be the most consistent parameter for predicting local control, and it is often superior to the T stage for many primary sites in attempts to assess the probability of cure after nonsurgical organ-preservation therapy. This quantitative information permits patients to decide the importance of preserving native laryngeal function and allows physicians to better counsel patients about the relative likelihood of local tumor control with primary surgery or irradiation. By providing quantitative volumetric information about the relative risk of local failure, this approach allows patients to decide what risk they are willing to assume to preserve native function on the basis of objective and reproducible data (33). The specific impact that a GTV calculation will have varies with the institution. Outcome assessments of volume measurements have been predominantly performed on populations treated with definitive nonsurgical organ-preservation therapy. It is generally agreed that GTV calculations provide a measure of the risk of local recurrence at various subsites. However, how this information is integrated into management differs depending on the treatment parameters specific to the institution. Institutions that aggressively use RT often use this data to identify patients treated with definitive RT who are at high risk for local recurrence. These patients may benefit from early surveillance imaging to detect recurrences that may be clinically occult. Early identification of recurrence may increase the success rate of salvage surgery. Another role of GTV measurements is to determine which patients may benefit from the addition of adjuvant chemotherapy. Patients at moderate risk of local recurrence after definitive RT, as determined with GTV measurements, may achieve a higher local control rate, and the additional morbidity and toxicity of adjuvant chemotherapy might be objectively justified. GTV may also be used to identify advanced large-volume tumors that are likely to be unresponsive to any form of combined nonsurgical therapy. Earlier identification may permit earlier surgical intervention and avoid the expense, morbidity, and delay in treatment associated with unsuccessful combined therapy.

Currently, both neuroradiologists and radiation oncologists perform tumor volume measurements, depending on the institution. As demonstrated in the American College of Radiology Imaging Network (ACRIN) protocol 6658, neuroradiologists and radiation oncologists who are knowledgeable in the pathologic and normal anatomic appearances of the head and neck can obtain reproducible GTV measurements. Many authors suggest that GTV should be integrated into future modifications of the current staging systems (23, 24). This information may permit more informed decision making that will account for the desires of the individual patient and for the objective and quantifiable data derived from cross-sectional imaging. We hope that this review helps to improve the reader's understanding of this important concept and its potential effect on patient care.

**References**


51. Paulus S, Sambon A, Vivegnis D, et al. 18FDG PET for the assessment of primary head and neck tumors: clinical, computed tomog-


