



HHS Public Access

Author manuscript

Am J Clin Oncol. Author manuscript; available in PMC 2016 September 07.

Published in final edited form as:

Am J Clin Oncol. 2014 April ; 37(2): 154–161. doi:10.1097/COC.0b013e31826e04d6.

A Prognostic Volumetric Threshold of Gross Tumor Volume in Head and Neck Cancer Patients Treated With Radiotherapy

Paul B. Romesser, MD^{*}, Muhammad M. Qureshi, MBBS^{*}, Rathan M. Subramaniam, MD, PhD^{*,†}, Osamu Sakai, MD, PhD[†], Scharukh Jalisi, MD[‡], and Minh T. Truong, MD^{*,†}

^{*}Department of Radiation Oncology, Boston Medical Center, Boston University School of Medicine, Boston, MA.

[†]Department of Radiology Boston Medical Center, Boston University School of Medicine, Boston, MA.

[‡]Department of Otolaryngology, Boston Medical Center, Boston University School of Medicine, Boston, MA.

Abstract

Objectives—To determine the prognostic utility of a volumetric threshold for gross tumor volume (GTV) of the primary and nodal disease when accounting for the TNM classification in head and neck cancer (HNC) patients treated with definitive radiotherapy (RT).

Materials and Methods—From 2004 to 2011, 79 HNC patients were treated to a median dose of 70 Gy, using intensity-modulated RT in 78.5% and 3-dimensional conformal RT in 21.5% with 83.5% receiving concurrent chemotherapy. Primary (GTV-P) and nodal (GTV-N) GTVs were derived from computed tomography (CT)-based contours for RT planning, of which 89.7% were aided by positron emission tomography-computed tomography. Local (LC), nodal (NC), distant (DC) control, and overall survival (OS) were assessed using the Kaplan-Meier product-limit method.

Results—With a median follow-up of 27.1 months GTV-P, threshold of <32.9 mL (mean value) compared with ≥32.9 mL, correlated with improved 2-year LC (96.2% vs. 63.9%, $P < 0.0001$), NC (100% vs. 69.2%, $P < 0.0001$), DC (87.9% vs. 64.2%, $P = 0.001$), and OS (88.4% vs. 58.6%, $P = 0.001$). GTV-P demonstrated its prognostic utility in multivariate analyses when adjusted for tumor category, cancer site, and chemotherapy regimen. Nodal GTV (mean, 34.0 mL) was not predictive of nodal control and survival.

Conclusions—A volumetric threshold of the primary tumor may be used as an independent prognostic factor in patients with HNC undergoing definitive RT.

Reprints: Minh T. Truong, MD, Department of Radiation Oncology, Boston Medical Center, Boston University School of Medicine, 830 Harrison Ave, Moakley Building LL 238, Boston, MA 02118. mitruong@bu.edu.

Present address: Paul B. Romesser, Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065.

Presented in part at the 53rd Annual Meeting of the American Society for Radiation Oncology, October 2 to 6, 2011, Miami Beach, FL.

The authors declare no conflicts of interest.

Keywords

head and neck cancer; intensity-modulated radiation therapy; gross tumor volume; cancer staging; PET/CT

Locoregional control is of paramount importance in determining long-term outcomes for head and neck cancer (HNC) patients undergoing definitive radiotherapy (RT) with or without concomitant chemotherapy.¹ However, despite modern RT techniques, such as intensity-modulated radiotherapy (IMRT), locoregional failure occurs in 30% to 50% of patients and mainly within the high-dose region of the RT field.²⁻⁵

The most commonly applied systems for the classification of disease and prognosis are the Tumor, Nodal, Metastasis (TNM) system devised by the International Union Against Cancer and the American Joint Committee on Cancer (AJCC) Staging Manual.^{6,7} Although these prognostic systems are widely accepted, much attention has been called to their weaknesses in accurately predicting treatment outcome in HNC patients.⁷⁻¹⁰ One possible reason for this limitation is that staging systems such as the TNM focus on a single dimension size criteria for tumor and nodal categorization in combination with the anatomic extent of disease rather than tumor volume factors.¹

Studies have reported the importance of tumor volume in predicting treatment outcomes in curative surgery,¹¹ definitive radiation therapy,¹²⁻¹⁴ and definitive chemoradiation.^{8-10,15,16} For patients undergoing radiation-based treatment, total dose to the tumor is usually limited to 70 Gy due to adjacent normal tissue constraints, which suggests that HNC with larger tumor burden may require additional treatment intensification.

The objective of this study was to determine the prognostic utility of a volumetric threshold for gross tumor volume (GTV) of the primary and nodal disease when accounting for the TNM classification in HNC patients treated with primary radiation-based treatment.

MATERIALS AND METHODS

Patient Selection

The study was performed as a retrospective review approved by the institutional review board with a waiver of informed consent. Newly diagnosed, nonmetastatic HNC patients treated between December of 2004 and May of 2011 with definitive RT were included. One hundred eight patients underwent definitive RT for biopsy-proven HNC during this interval with 79 patients meeting inclusion criteria defined as no prior history of HNC or other malignancies in the prior 5 years and no prior surgical treatment. In total 29 patients were excluded: 3 patients with recurrent HNC, 11 patients with synchronous malignancies, 5 patients had prior history of malignancy, 2 patients died during or shortly after treatment from noncancer-related causes, 3 patients had <1 month of follow-up, 5 patients were excluded for HNC of unknown primary (n = 3), ear (n = 1), and sinonasal cavity (n = 1). All patients were staged according to the 2002 AJCC classification with history, physical examination, focused head and neck evaluation, panendoscopy with biopsy, and computed

tomography (CT).¹⁷ Sixty-nine patients (89.7%) underwent a staging 2-[¹⁸F] fluoro-2-deoxy-D-glucose positron emission tomography (PET/CT).

Treatment Planning and GTV Measurement

The GTVs were manually contoured, by a single radiation oncologist, for IMRT or 3D-chemoradiation treatment planning. For patients who underwent induction chemotherapy, GTVs were contoured based on the preinduction chemotherapy CT, as the standard practice was to use the prechemotherapy tumor volume to define the GTV. The primary tumor GTV (GTV-P), nodal tumor GTV (GTV-N), and a total combined GTV (GTV-TC), the sum of GTV-P and GTV-N, were automatically calculated from clinical dose volume histogram data (Philips Pinnacle software suite). PET/CT images were fused using Philips Pinnacle software suite or MIMVista (for patients treated after November 2008). Volumetric expansions from GTV were 7 to 15 mm respecting normal tissue planes to create the clinical target volume, followed by a 3 to 5 mm expansion to the planning target volume. The cohort's mean GTV-P was 32.9 mL (n = 79; SD ± 36.0), GTV-N was 34.0 mL (n = 54; SD ± 45.8), and GTV-TC was 56.1 mL (n = 79; SD ± 53.2).

The prescription planning target volume was treated to a median dose of 69.96 Gy (range, 66 to 72 Gy) over 33 fractions (range, 32 to 42 d) and 48 days (range, 38 to 72 d) to encompass gross primary and nodal disease. For patients undergoing IMRT, a simultaneous integrated boost technique was used with elective nodal areas treated to 54 to 56 Gy for low-risk areas and 60 Gy to high-risk areas but without gross disease.

Treatment

Sixty-two patients (78.5%) were treated with IMRT and 17 patients (21.5%) received 3D-conformal RT. Choice of chemotherapy regimen was at the discretion of the medical oncologist. Sixty-six patients (83.5%) received concurrent chemotherapy of which 65 patients had AJCC stage III to IV disease and 1 patient had stage II nasopharyngeal cancer. Three AJCC stage III patients did not receive chemotherapy. The distribution of chemotherapy regimen was as follows: 47 received cisplatin, 10 received carboplatin, and 9 received cetuximab. Twenty of these 66 patients underwent induction chemotherapy before the initiation of chemoradiation: 9 with docetaxel, cisplatin, and 5-fluorouracil (5-FU) (TPF regimen), 5 with carboplatin and docetaxel, 1 with carboplatin and paclitaxel, 4 with cisplatin and docetaxel, and 1 with carboplatin and 5-FU.

Follow-up

Patients were followed up after the conclusion of treatment and continuing until analysis or patient death. Follow-up consisted of serial clinical examinations, a PET/CT after completion of RT at approximately 8 to 12 weeks, and then annually for patients who obtained a complete response as part of routine clinical care. Disease recurrence was defined as site of failure including local failure (LF), nodal failure (NF), or distant failure (DF). All failures were confirmed by biopsy.

Statistical Analysis

Descriptive statistics were used to obtain patient and tumor characteristics. Analysis of variance was conducted to assess association between the GTV and various descriptive characteristics and disease outcomes. Two-year actuarial rates were estimated for local control (LC), nodal control (NC), distant control (DC), and overall survival (OS) (death due to any cause) using the Kaplan-Meier product-limit method.¹⁸ All endpoints were measured from the end of RT until relapse or death with censorship at last follow-up or death. The comparison of rates among the groups was done using the 2-tailed log-rank test.

GTV analysis included categorical division using the overall mean values. Receiver operating characteristics (ROC) curves were constructed for each endpoint with an optimized sensitivity and specificity defined volumetric threshold identified as the cut-point with the greatest percent area under curve (AUC). Multivariate (MV) analyses were performed for the following potential confounding variables: age at diagnosis (y), sex, smoking history (smokers vs. nonsmokers), cancer site (nasopharyngeal/oropharyngeal, oral cavity/hypopharyngeal, and larynx), tumor category (T1 to T2 vs. T3 to T4), nodal category (N0 vs. N1 to N3), AJCC stage (stage I to III vs. stage IV), and chemotherapy regimen (no chemotherapy, induction chemotherapy with concurrent chemoradiotherapy [CCRT], and CCRT only). Crude and adjusted hazard ratios (HR) with 95% confidence intervals (CI) were computed using Cox regression modeling for LF, NF, DF, and overall death (OD).¹⁹ A probability value of <0.05 was considered statistically significant for all analyses. All statistical computations were performed on SAS 9.1 system (SAS Institute, Cary, NC).

RESULTS

Patients and Tumor Characteristics

The patient population comprised of 74.7% male, 87.3% were smokers, and 86.1% were locally advanced stage III or IV. Median follow-up was 27.1 months (range, 3.0 to 78.5 mo) in the overall study population and 30.0 months (range, 4.9 to 78.5 mo) among surviving patients. Patient, tumor, and volume characteristics are described in Table 1.

Treatment Outcome

Local, nodal, and distant recurrences occurred in 12 (15.2%), 8 (10.1%), and 15 (19.0%) patients, respectively. The median time to recurrence for LF was 2.6 months (range, 1.4 to 35.5 mo), NF was 2.1 months (range, 0.3 to 3.3 mo), and for DF was 3.1 months (range, 1.9 to 23.0 mo).

GTV statistics by tumor control and survival are described in Table 2. Patients who failed locally had a significantly larger GTV-P (64.0 vs. 27.3 mL, $P=0.001$) and non-significantly larger GTV-TC (69.9 vs. 53.6 mL, $P=0.332$). GTV-P correlated with all disease outcomes. Oropharyngeal and larynx patients who failed locally had significantly larger GTV-P compared with those who were locally controlled (oropharynx: 70.3 vs. 23.8 mL, $P=0.002$; larynx: 106.6 vs. 16.0 mL, $P=0.0001$). A similar analysis of GTV-P statistics by LC status could not be performed for other cancer sites due to fewer patients.

Correlating T-Category With Primary GTV and Treatment Outcome

Forty-nine patients (62.1%) presented with advanced T-category (T3 to T4). The mean GTV-P for T3 to T4 disease was 43.2 mL compared with 16.1 mL for T1 to T2 disease ($P=0.001$). No significant differences were noted between T1 to T2 and T3 to T4 disease with 2-year LC rates of 93.3% and 81.1%, respectively ($P=0.087$), and corresponding 2-year OS rates of 88.9% and 72.6% ($P=0.138$).

Correlating Mean Primary, Nodal, and Total GTVs With Treatment Outcome

By mean GTV-P, a volume of <32.9 mL significantly correlated with improved LC (96.2% vs. 63.9%, $P<0.0001$) (Fig. 1A), NC (100% vs. 69.2%, $P<0.0001$), DC (87.9% vs. 64.2%, $P=0.001$), and OS (88.4% vs. 58.6%, $P=0.001$) (Fig. 1B, Table 3). 92.3% of patients with GTV-P of >32.9 mL had T3 to T4 disease compared with 47.2% with GTV-P of <32.9 mL ($P=0.0001$).

A GTV-N of >34.0 mL (mean value) did not predict NC, DC, or OS but did predict improved LC. Smaller mean GTV-P (22.3 mL) was seen with larger GTV-N (>34.0 mL) in our patient population ($P=0.032$). Hence, tumor control outcomes according to nodal GTV burden were largely driven by GTV-P burden. By mean GTV-TC, a volume of >56.1 mL predicted NC and DC, but not LC and OS.

Primary GTV Analysis Restricted to Advance T-Category Disease (T3 to T4)

A subset analysis of patients with advanced disease (T3 to T4) demonstrated that smaller mean GTV-P significantly predicted improved LC (Fig. 2A), NC, DC, and OS (Fig. 2B, Table 3). Furthermore, a smaller GTV-P (<32.9 mL) was associated with improved OS regardless of T-category (in both T1 to T2 and T3 to T4 subsets, Figure 3). The 2-year OS was 91.6% and 84.3% in patients with GTV-P of <32.9 mL and T-category T1 to T2 and T3 to T4, respectively, whereas patients with GTV-P of >32.9 mL had 2-year OS of 50.0% and 59.1% for T1 to T2 and T3 to T4 disease respectively.

Primary GTV Analysis by Chemotherapy Regimen

In the locally advanced patient cohort, a subgroup analysis restricted to patients who underwent induction chemotherapy with CCRT and separately to those who were treated with CCRT alone was performed. No differences were noted between the 2 cohorts for age ($P=0.687$), sex ($P=0.760$), smoking ($P=0.156$), cancer site ($P=0.672$), tumor category ($P=0.140$), nodal category ($P=0.197$), and AJCC stage ($P=1.0$). All GTVs were derived from pretreatment CT scans, and there were no significant differences in mean GTV-P (34.7 vs. 38.1 mL, $P=0.742$), GTV-N (31.6 vs. 37.0 mL, $P=0.692$), or GTV-TC (63.2 vs. 65.5 mL, $P=0.875$), for induction chemotherapy followed by CCRT versus CCRT only, respectively. A larger GTV-P of >32.9 mL correlated with worse LC, NC, DC, and OS, for patients undergoing CCRT alone. In patients undergoing induction chemotherapy, GTV-P lost significance for LC and survival, although remained significant for nodal and DC; however, the analysis was limited in the induction chemotherapy cohort by low patient numbers (Table 3).

Univariate and MV Analyses

Univariate analyses confirmed that a larger GTV-P correlated with increased risk of LF, DF, OF, and OD (Table 4). When adjusted individually for T-category and cancer site, GTV-P HRs retained significance. Larger GTV-P, after adjusting for T-category, was associated with an increase in LF (HR, 14.6; 95% CI, 2.3–94.4), DF (HR, 5.0; 95% CI, 1.3–19.2), and OD (HR, 3.8; 95% CI, 1.4–10.1). Because of lack of events in the favorable group for NF a MV model could not be performed. Finally, age, sex, smoking history, nodal category, AJCC stage, and chemotherapy regimen did not affect the crude results, confirming the prognostic utility of GTV-P.

ROCs Curve Optimized Cut-points for Primary GTV

An exploratory analysis of ROC identified a volumetric threshold of 31.9 mL for GTV-P and LC (AUC, 85.2%; 95% CI, 75.4–92.2). When dichotomized by the ROC defined threshold it was noted that a volume of <31.9 mL was associated with improved LC (100.0% vs. 59.4%, $P < 0.0001$), NC (100% vs. 70.5%, $P < 0.0001$), DC (87.7% vs. 65.6%, $P = 0.001$), and OS (90.0% vs. 58.1.6%, $P = 0.0003$).

DISCUSSION

The current study evaluates the prognostic utility of GTV in a population of HNC patients treated with definitive radiation-based treatment. A smaller GTV-P correlated with improved LC, NC, and DC, as well as OS. In a MV analysis, GTV-P retained significance when adjusted individually for T-category, cancer site, chemotherapy regimen, and other patient/tumor factors. In contrast, nodal GTV did not correlate with disease control or survival in this study. The ability of combined total GTV to correlate with nodal and DC was likely an effect of GTV-P, rather than the inclusion of GTV-N. An exploratory ROC analysis determined an optimal GTV-P threshold of 31.9 mL, which was similar to mean GTV-P threshold of 32.9 mL.

Traditionally, tumor and nodal categorization as part of TNM staging have used unidimensional thresholds. The current staging system uses primary tumor size criteria (2, 2 to 4, and > 4 cm cut-point) to distinguish between T1, T2, and T3 tumors, respectively, whereas for nodal categorization nodal sizes of 3, 3 to 6, and > 6 cm are used to categorize nodal disease as N1, N2, and N3 category, respectively. Such categorization makes an assumption of tumors being spherical and such measurements are assumed to be a surrogate for tumor burden. This study shows that in analyses restricted to locally advanced disease patients (T3 and T4), primary site GTV remained significant for all endpoints. Similarly, a smaller tumor volume was associated with improved OS regardless of T categorization thereby suggesting a more important role of volume-based measurements, such as a GTV, as a prognostic factor compared with traditional TNM staging.

Interestingly, in a subgroup analysis performed on patients who received CCRT alone, a larger GTV-P was significant for worse outcome for all endpoints, whereas among the induction chemotherapy patients GTV-P was only significant for NC and DC despite differences in LC and OS, likely in part due to small sample size. The LC in the induction

chemotherapy cohort was lower for patients with GTV < 32.9 mL compared with the CCRT alone group, whereas the DC was higher in the induction chemotherapy group compared with CCRT alone, which suggests there may be a trade-off between LC and DC between these different treatment approaches based on GTV-P.

Our study findings are consistent with previous reports that patients with a significant tumor volume had inferior local, nodal, and overall control as well as OS (Table 5). Previous studies have directly compared the predictive value of GTV and TNM classification. Doweck and colleagues reported that primary tumor volume had the greatest effect on survival ($P=0.0007$) and was superior to the TNM classification system in predicting OS and LC, whereas Studer and colleagues demonstrated that GTV was able to predict risk of LF in a cohort of locally advanced patients (defined as T3/4 or N2c/3). Strongin and colleagues reported that patients with locoregional failure had a larger GTV-P than patients with DFs (58 vs. 31 mL, $P=0.019$), yet the average GTV-P in patients who experienced no treatment failure (35.5 mL) was greater than that of patients who experienced DF (31 mL).¹⁵ In this study, we found that patients with LF, NF, or DF had significantly larger primary tumor volumes (64.0, 65.3, and 50.5 mL, respectively) than patients who were free of disease recurrence (27.3, 29.2, and 28.8 mL, respectively).

As expected, differences in GTV-P were noted by site, including hypopharyngeal and oral cavity cancers, which on average had greater GTV-Ps, and higher T-category, likely representing patients who were not surgical candidates and received primary radiation-based treatment. GTV-P was able to stratify patients into high-risk and low-risk groups independent of T-category, cancer site, and chemotherapy regimen. MV analysis confirmed the prognostic utility of GTV when adjusting for T-category demonstrating a 14.6-fold risk of LF, a 5.0-fold risk for developing distant metastases, and a 3.8-fold risk of death in patients with a GTV-P >32.9 mL.

Primary GTV likely can identify patients eligible for organ preservation approaches using chemoradiation. Chen and colleagues reported that in stage III and IV, nonbulky hypopharyngeal cancer patients with a primary GTV < 30 mL had improved locoregional control rates on MV analysis.²⁰ This may have implications for locally advanced T4 category patients who are not typically considered eligible for laryngeal preservation therapies.²⁰ In larynx cancer, treatment decisions for total laryngectomy versus organ preservation using chemoradiation are primarily based on whether the primary tumor constitutes large-volume T4 disease. However, large volume is currently defined by anatomic extent as tumor penetrating through the cartilage or extending >1 cm into the base of the tongue and at the interpretation of the treating physician. It is possible that such patients may benefit from a more defined volumetric cut-point to define prognosis and to aid in the decision between treatment strategies such as a total laryngectomy or organ-preserving chemoradiation.

The potential utility of GTV-P as a prognostic variable rests largely on the ability to delineate tumors accurately and reproducibly. Similarly, the efficacy of IMRT is largely dependent on the ability to define disease burden given the steep dose gradients between tumor and normal tissue.²¹ Studies have shown that the integration of PET/CT into the GTV

delineation process increases the precision and interobserver reliability in defining a tumor volume.²¹ The majority of our patients received a pretreatment PET/CT, which may have improved accuracy in defining the tumor volumes as compared with CT alone. In our experience, PET defined volumes using autosegmentation often underestimates the tumor burden as compared with CT; and therefore, we did not use auto segmentation methods for RT planning as we relied on a combination of imaging and endoscopic/clinical data to define GTV-P and GTV-N.²² Interobserver variation in GTV contouring is low; as 1 study reported GTV measurements were reliable and reproducible when performed by radiation oncologists and neuroradiologists with an interclass correlation coefficient of 0.81, which was reported as excellent.¹¹ More recent studies have reported on the improving ability of computerized volumetric analyses, which may assist integration of volumetric measurements into clinical algorithms.^{23–25}

Our results are in accordance with published reports that a lower GTV-P correlates with improved control and survival rates, yet several key factors distinguish our study from those previously reported. All GTV values were obtained from clinical treatment plans based on simulation CT with PET/CT in 89.7% of patients, with the majority of our patients receiving IMRT (78.5%). Although other studies have demonstrated the predictive significance of nodal GTV, we did not find that GTV-N correlated with outcome or survival, which may reflect the radiation management of the neck. All nodal gross disease defined by imaging was treated to 70 Gy without planned neck dissection, and evaluation for residual disease at 10 to 12 weeks with PET/CT was performed to evaluate patients who needed salvage neck dissection.

Studies have reported on the improved prognostic value of GTV as compared with AJCC staging and TNM categorization, yet we demonstrate the added utility of GTV-P to further dichotomize patients into favorable and unfavorable risk groups within T-categories signifying GTV-P as an independent variable that is complementary to the current TNM system. Although it is unlikely that GTV alone will supersede the TNM system, our study demonstrates that future TNM staging may incorporate a tumor size criteria based on volume rather than a single dimension in addition to defining anatomic tumor spread for head and neck malignancies.

Limitations of the study included the retrospective design, small patient cohort, exclusion of surgical patients, multiple head and neck sites, and variation in chemotherapy regimen given, although the patient population received uniform radiation treatment with all GTVs receiving a median prescription dose of 70 Gy, at a single center, and all primary/nodal GTVs were determined by a single radiation oncologist using uniform RT planning software.

In conclusion, primary GTV is an independent prognostic variable in patients with HNC undergoing definitive radiation-based treatment. GTV-P retained prognostic significance when accounting for T-category, AJCC stage, cancer site, and in patients with locally advanced disease. GTV-P was significant for all endpoints in patients receiving CCRT alone, although was predictive only for NC and DC in patients receiving induction chemotherapy followed by CCRT. A primary GTV threshold may assist in risk stratification to help

identify patients at high risk of failure who might benefit from various strategies of treatment intensification using combined modality therapy as well as identify lower risk patients eligible for less intensive organ preservation therapies.

REFERENCES

1. Bernier, J. *Head and Neck Cancer Multimodality Management*. New York: Springer; 2011.
2. Dawson LA, Anzai Y, Marsh L, et al. Patterns of local-regional recurrence following parotid-sparing conformal and segmental intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys*. 2000; 46:1117–1126. [PubMed: 10725621]
3. Hauswald H, Simon C, Hecht S, et al. Long-term outcome and patterns of failure in patients with advanced head and neck cancer. *Radiat Oncol*. 2011; 6:70. [PubMed: 21663634]
4. Studer G, Luetolf UM, Glanzmann C. Locoregional failure analysis in head-and-neck cancer patients treated with IMRT. *Strahlenther Onkol*. 2007; 183:417–423. discussion 424–415. [PubMed: 17680220]
5. Schoenfeld GO, Amdur RJ, Morris CG, et al. Patterns of failure and toxicity after intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys*. 2008; 71:377–385. [PubMed: 18164838]
6. UICC. *TNM Classification of Malignant Tumours*. 7th. John Wiley and Sons; 2009.
7. AJCC. *Cancer Staging Manual*. 7th. Chicago, IL: Springer; 2010.
8. Studer G, Lutolf UM, El-Bassiouni M, et al. Volumetric staging (VS) is superior to TNM and AJCC staging in predicting outcome of head and neck cancer treated with IMRT. *Acta Oncol*. 2007; 46:386–394. [PubMed: 17450476]
9. Doweck I, Denys D, Robbins KT. Tumor volume predicts outcome for advanced head and neck cancer treated with targeted chemoradiotherapy. *Laryngoscope*. 2002; 112:1742–1749. [PubMed: 12368607]
10. Chen MK, Chen TH, Liu JP, et al. Better prediction of prognosis for patients with nasopharyngeal carcinoma using primary tumor volume. *Cancer*. 2004; 100:2160–2166. [PubMed: 15139059]
11. Mukherji SK, O'Brien SM, Gerstle RJ, et al. The ability of tumor volume to predict local control in surgically treated squamous cell carcinoma of the supraglottic larynx. *Head Neck*. 2000; 22:282–287. [PubMed: 10748452]
12. Mendenhall WM, Parsons JT, Mancuso AA, et al. Definitive radiotherapy for T3 squamous cell carcinoma of the glottic larynx. *J Clin Oncol*. 1997; 15:2394–2402. [PubMed: 9196155]
13. Mukherji SK, Mancuso AA, Mendenhall W, et al. Can pretreatment CT predict local control of T2 glottic carcinomas treated with radiation therapy alone? *Am J Neuroradiol*. 1995; 16:655–662. [PubMed: 7611018]
14. Lee WR, Mancuso AA, Saleh EM, et al. Can pretreatment computed tomography findings predict local control in T3 squamous cell carcinoma of the glottic larynx treated with radiotherapy alone? *Int J Radiat Oncol Biol Phys*. 1993; 25:683–687. [PubMed: 8454487]
15. Strongin A, Yovino S, Taylor R, et al. Primary tumor volume is an important predictor of clinical outcomes among patients with locally advanced squamous cell cancer of the head and neck treated with definitive chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2012; 82:1823–1830. [PubMed: 21549516]
16. Lok BH, Setton J, Caria N, et al. Intensity-modulated radiation therapy in oropharyngeal carcinoma: effect of tumor volume on clinical outcomes. *Int J Radiat Oncol Biol Phys*. 2012; 82:1851–1857. [PubMed: 21640497]
17. AJCC. *Cancer staging manual*. New York: Springer-Verlag; 2002.
18. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958; 53:25.
19. Cox DR. Regression models and life-tables. *J R Stat Soc*. 1972; 34:34.
20. Chen SW, Yang SN, Liang JA, et al. Prognostic impact of tumor volume in patients with stage III-IVA hypopharyngeal cancer without bulky lymph nodes treated with definitive concurrent chemoradiotherapy. *Head Neck*. 2009; 31:709–716. [PubMed: 19260114]

21. Hermans R. Head and neck cancer: how imaging predicts treatment outcome. *Cancer Imaging*. 2006; 6:S145–S153. [PubMed: 17114069]
22. Thiagarajan A, Caria N, Schoder H, et al. Target volume delineation in oropharyngeal cancer: impact of PET, MRI, and physical examination. *Int J RadiatOncolBiol Phys*. 2012; 83:220–227.
23. Hadjiiski L, Mukherji SK, Gujar SK, et al. Treatment response assessment of head and neck cancers on CT using computerized volume analysis. *Am J Neuroradiol*. 2010; 31:1744–1751. [PubMed: 20595363]
24. Hadjiiski L, Mukherji SK, Ibrahim M, et al. Head and neck cancers on CT: preliminary study of treatment response assessment based on computerized volume analysis. *Am J Roentgenol*. 2010; 194:1083–1089. [PubMed: 20308515]
25. Street E, Hadjiiski L, Sahiner B, et al. Automated volume analysis of head and neck lesions on CT scans using 3D level set segmentation. *Med Phys*. 2007; 34:4399–4408. [PubMed: 18072505]

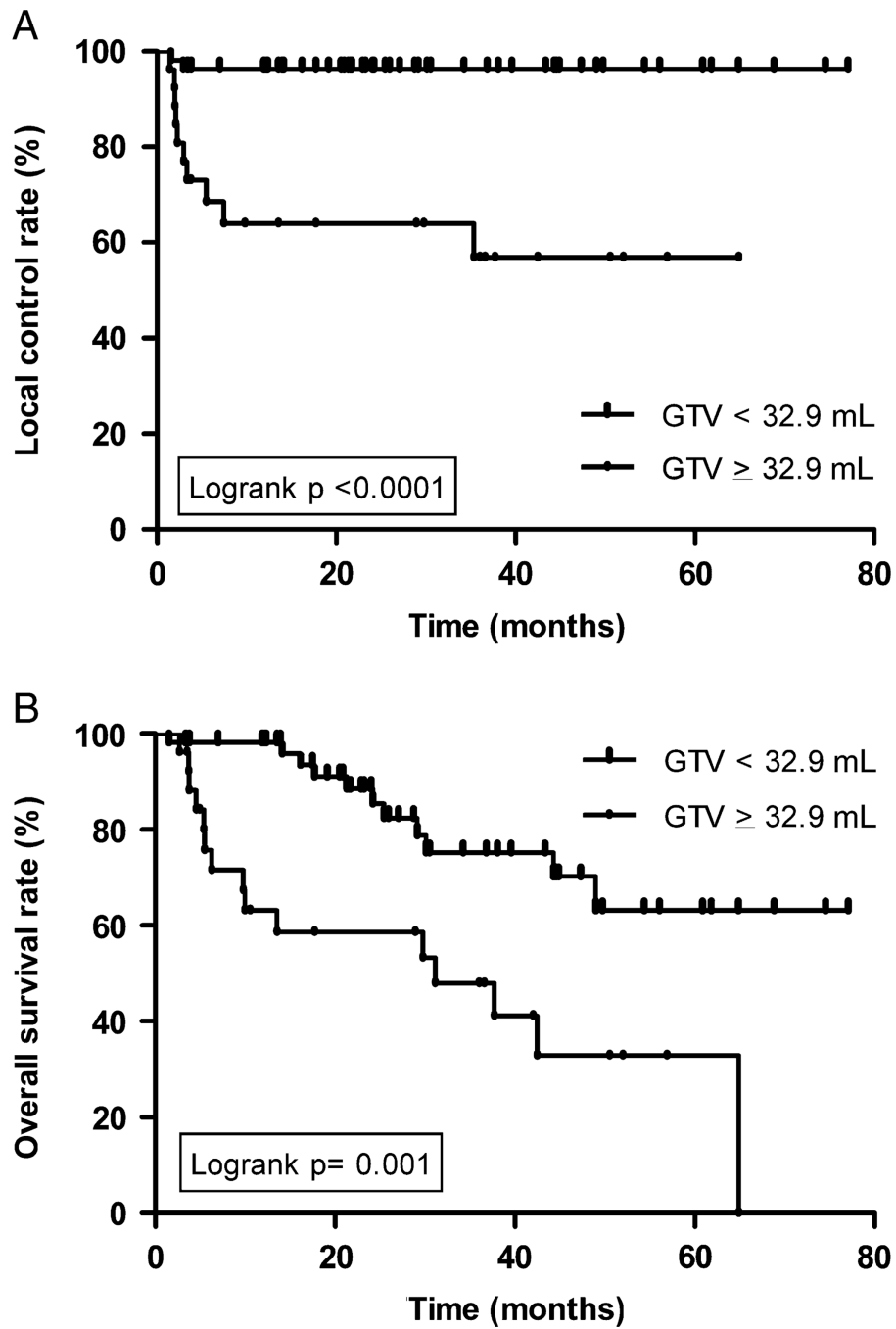
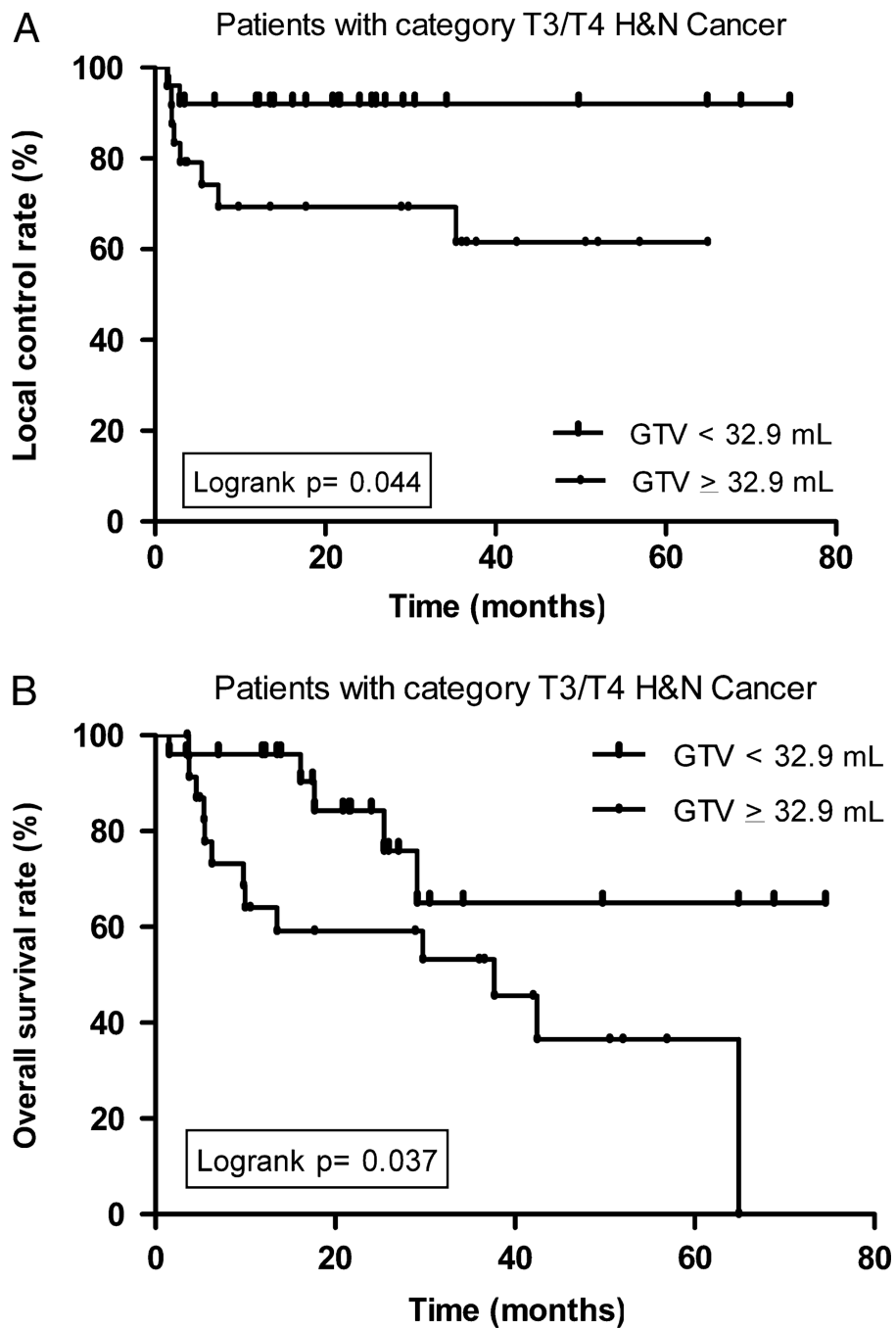


FIGURE 1. A, Local control according to primary GTV. B, Overall survival according to primary GTV. GTV indicates gross tumor volume.

**FIGURE 2.**

A, Local control according to primary GTV in the subset of T3/T4 patients. B, Overall survival according to primary GTV in the subset of T3/T4 patients. GTV indicates gross tumor volume.

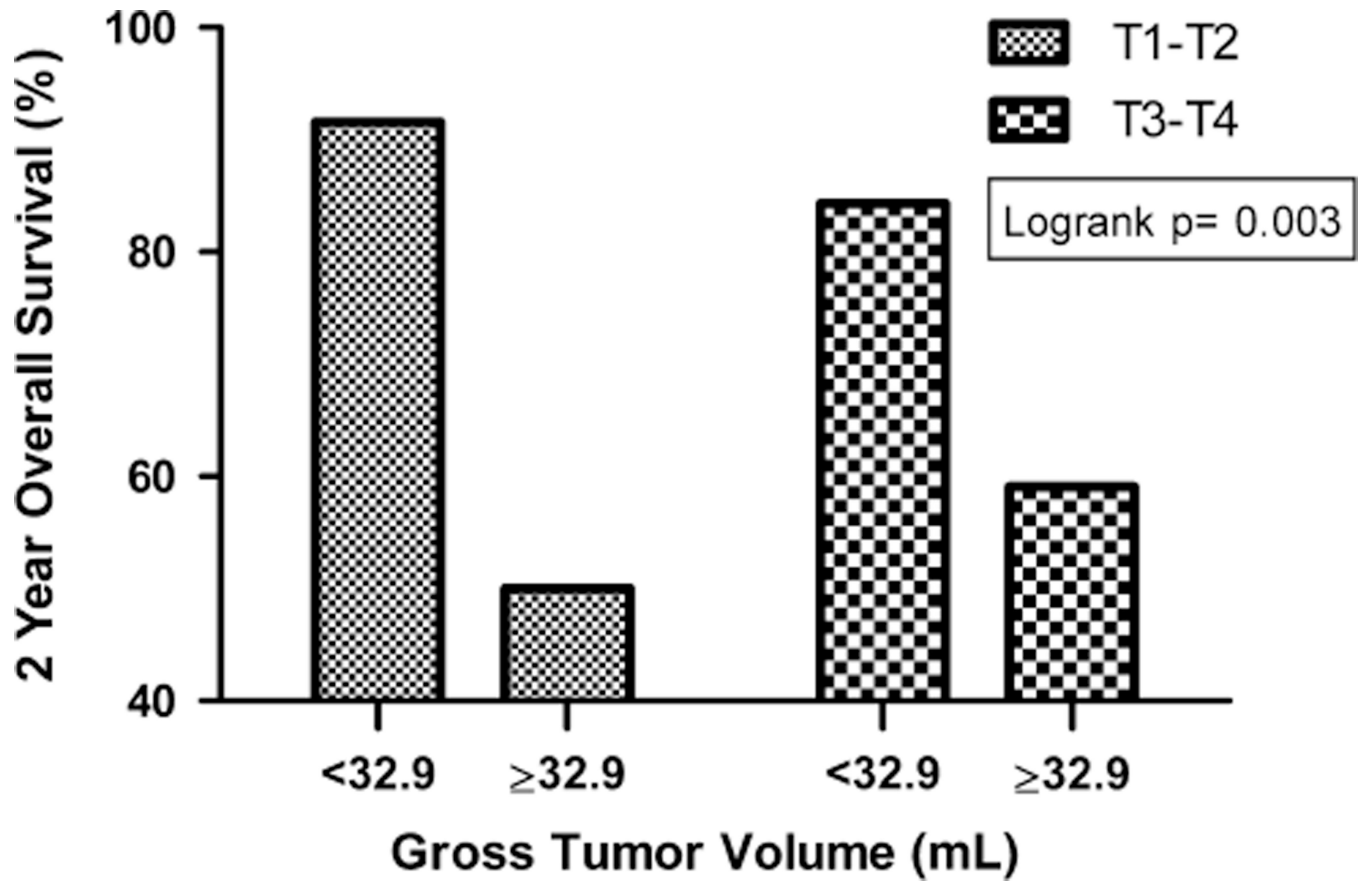


FIGURE 3. Overall survival according to primary gross tumor volume and T-category.

TABLE 1

Patient and Tumor Characteristics of 79 Head and Neck Cancer Patients

	n	Median	Mean ± SD	Range
Age (y)	79	59.0	58.8 ± 9.9	31–86
Smoking (pack-years)	77	35.0	35.6 ± 26.6	0–120
Gross tumor volume (mL), primary				
Overall	79	20.2	32.9 ± 36.0	1.5–177.9
Oropharyngeal	33	17.4	28.1 ± 26.6	1.5–108.2
Larynx	24	11.1	23.5 ± 33.9	3.3–162.5
Hypopharyngeal	9	35.0	42.2 ± 26.5	9.1–99.2
Oral cavity	7	64.6	84.9 ± 60.7	32.0–177.9
Nasopharyngeal	6	22.6	22.3 ± 13.1	8.0–39.5
Gross tumor volume (mL), nodal	54	14.3	34.0 ± 45.8	0.6–214.0
Gross tumor volume (mL), total	79	39.3	56.1 ± 53.2	3.8–258.3
Follow-up (mo)				
Overall study follow-up	79	27.1	30.1 ± 19.5	3.0–78.5
Surviving patients	53	30.0	33.8 ± 19.9	4.9–78.5
			n (%)	
Pathology				
Squamous cell carcinoma			71 (89.9)	
Other			8 (10.1)	
AJCC Stage				
I			6 (7.6)	
II			5 (6.3)	
III			12 (15.2)	
IV			56 (70.9)	
Tumor category				
T1			11 (13.9)	
T2			19 (24.0)	
T3			25 (31.7)	
T4			24 (30.4)	
Nodal category				
N0			25 (31.7)	
N1			7 (8.9)	
N2			39 (49.4)	
N3			8 (10.1)	

AJCC indicates American Joint Committee on Cancer; n, number of patients.

TABLE 2

Gross Tumor Volume (mL) by Disease Control and Survival Status of Patients

	GTV Primary (mL)		GTV Nodal (mL)		GTV Total (mL)	
	n	Mean ± SD	n	Mean ± SD	n	Mean ± SD
Local control						
Local control	67	27.3 ± 32.4	44	40.1 ± 48.7	67	53.6 ± 55.1
Local failure	12	64.0 ± 40.5	10	7.1 ± 6.8	12	69.9 ± 40.1
<i>P</i>		0.001		0.039		0.332
Nodal control						
Nodal control	71	29.2 ± 35.2	48	31.8 ± 40.2	71	50.7 ± 49.1
Nodal failure	8	65.3 ± 26.1	6	51.6 ± 81.1	8	100.4 ± 67.2
<i>P</i>		0.006		0.322		0.006
Distant Control						
Distant control	64	28.8 ± 33.8	41	23.2 ± 25.0	64	43.6 ± 38.2
Distant failure	15	50.5 ± 40.9	13	67.9 ± 74.4	15	109.3 ± 74.0
<i>P</i>		0.035		0.002		< 0.0001
Overall survival						
Alive	53	23.1 ± 21.5	39	31.7 ± 40.7	53	46.5 ± 43.5
Died	26	52.9 ± 49.6	15	39.7 ± 58.2	26	75.8 ± 65.6
<i>P</i>		0.0004		0.570		0.020

GTV indicates gross tumor volume.

TABLE 3

GTV (mL) and Survival/Time to Disease Outcomes

	N	GTV Primary (mL) Mean	T3-T4 n (%)	Events (2 y Actuarial Control and Survival Rates) n (%)				
				Local Control	Nodal Control	Distant Control	Overall Survival	
All subjects	79	32.9	49 (62.1)	12 (85.8)	8 (89.8)	15 (80.0)	26 (78.8)	
GTV primary (mL)								
< 32.9	53	14.4	25 (47.2)	2 (96.2)	0 (100.0)	5 (87.9)	11 (88.4)	
32.9	26	70.7	24 (92.3)	10 (63.9)	8 (69.2)	10 (64.2)	15 (58.6)	
<i>P</i>		<0.0001	0.0001	<0.0001	<0.0001	0.001	0.001	
GTV nodal (mL)								
< 34.0	39	37.3	28 (71.8)	10 (76.1)	4 (89.5)	7 (81.4)	11 (79.0)	
34.0	15	22.3	7 (46.7)	0 (100.0)	2 (86.7)	6 (65.2)	4 (86.7)	
<i>P</i>		0.032	0.083	0.033	0.755	0.253	0.691	
GTV total (mL)								
< 56.1	50	17.3	28 (56.0)	6 (89.6)	1 (98.0)	3 (93.0)	14 (83.8)	
56.1	29	59.7	21 (72.4)	6 (79.3)	7 (75.9)	12 (59.1)	12 (70.7)	
<i>P</i>		<0.0001	0.147	0.290	0.002	0.0002	0.245	
T3-T4 restricted	49	43.2	49 (100.0)	10 (81.1)	6 (87.5)	12 (74.0)	18 (72.6)	
GTV primary (mL)								
< 32.9	25	17.9	25 (100.0)	2 (92.0)	0 (100.0)	3 (80.7)	5 (84.3)	
32.9	24	69.5	24 (100.0)	8 (69.3)	6 (75.0)	9 (66.1)	13 (59.1)	
<i>P</i>		<0.0001	N/A	0.044	0.009	0.043	0.037	
Induction with CCRT	20	34.7	17 (85.0)	6 (68.2)	4 (79.0)	5 (77.7)	3 (83.6)	
GTV primary (mL)								
< 32.9	12	16.9	10 (83.3)	2 (83.3)	0 (100.0)	0 (100.0)	1 (91.7)	
32.9	8	61.6	7 (87.5)	4 (46.9)	4 (50.0)	5 (46.9)	2 (70.0)	
<i>P</i>		<0.0001	1.000	0.136	0.009	0.005	0.331	
CCRT only	46	38.1	31 (67.4)	6 (89.1)	3 (93.5)	9 (78.5)	17 (77.8)	
GTV primary (mL)								
< 32.9	29	15.3	15 (51.7)	0 (100.0)	0 (100.0)	5 (81.2)	5 (88.9)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	N	GTV Primary (mL) Mean	T3-T4 n (%)	Events (2 y Actuarial Control and Survival Rates) n (%)			
				Local Control	Nodal Control	Distant Control	Overall Survival
32.9	17	77.0	16 (94.1)	6 (70.1)	3 (82.4)	4 (75.5)	12 (58.8)
<i>P</i>		<0.0001	0.003	0.0007	0.020	0.379	0.0008

CCRT indicates concurrent chemoradiation treatment; GTV, gross tumor volume; N/A, not applicable.

Cox Regression Analysis of Primary Gross Tumor Volume (mL) and Survival/Time to Disease Outcomes

TABLE 4

		Hazard Ratio (95% CI)		
	n	Events	Univariate	Adding T-Category* Adding Cancer Site [†]
Local failure				
< 32.9	53	2	Reference	Reference
32.9	26	10	11.9 (2.6–54.6)	14.6 (2.3–94.4)
<i>P</i>			0.009	0.005
Distant failure				
< 32.9	53	5	Reference	Reference
32.9	26	10	5.4 (1.8–15.8)	5.0 (1.3–19.2)
<i>P</i>			0.002	0.018
Overall death				
< 32.9	53	11	Reference	Reference
32.9	26	15	3.5 (1.6–7.6)	3.8 (1.4–10.1)
<i>P</i>			0.002	0.008

Analysis of nodal failure could not be run because of 0 events in referent group.

* T-category modeled as T1 to T2 vs. T3 to T4.

[†] Cancer site modeled as nasopharyngeal/oropharyngeal, oral cavity/hypopharyngeal, and larynx.

CI indicates confidence interval.

TABLE 5
 Summary of Select Studies on Head and Neck Cancer Gross Tumor Volume and Disease Control

References	N	Sites	Stage	IMRT (%)	CCRT (%)	Mean GTV (mL)	Follow-up (mo)	GTV Cutoff (mL)	Local Control (%), P	Overall Survival (%), P
Doweck et al ⁹	64	E, H, L, NP, O, OC, PS	III, IV	0	100	Primary: 18.4 Nodal: 22.4 Total: 35.4	Mean: 26.3	19.6	5-y: 93.8 vs. 57.0, P = 0.001	5-y: 41.5 vs. 14.1, P = 0.0018
Studer et al ^{4,8}	172	H, NP, O, OC, PS	All	100	75	Primary: 37.7 Total: 52	Mean: 17	15/70	2-y: 95 vs. 70 vs. 50, P = 0.0001	
Chen et al ²⁰	76	H (nonbulky)	III, IVA	60.5	100	Primary: 33.4	Median: 37	30	3-y: 72 vs. 23*, P = 0.0001	3-y: 75 vs. 20 [†] , P = 0.0001
Strongin et al ¹⁵	78	H, O, L	III, IV	39	100	Primary: 38.7 Nodal: 24.9	Median: 36	35		3-y: 84 vs. 42, P < 0.001
Lok et al ¹⁶	340	O	All	100	95	Primary: 42.53 Nodal: 31.58	Median: 34	32.79	2-y: 98.1 vs. 89.6, P = 0.004	2-y: 94.3 vs. 82.7, P < 0.0001
This study	79	H, L, O, OC, NP	All	78.5	83.5	Primary: 32.9 Nodal: 34.0 Total: 56.1	Median: 27.1	31.9	2-y: 100.0 vs. 59.4, P < 0.0001	2-y: 90.0 vs. 58.1, P = 0.0003

* Primary Tumor relapse-free survival.

[†] Cause-specific survival.

CCRT indicates concurrent chemoradiotherapy; E, esophagus; GTV, gross tumor volume; H, hypopharynx; IMRT, intensity-modulated radiotherapy; L, larynx; N, number of patients; NP, nasopharynx; O, oropharynx; OC, oral cavity; PS, paranasal sinus.