Better Prediction of Prognosis for Patients with Nasopharyngeal Carcinoma Using Primary Tumor Volume

Mu-Kuan Chen, M.D., M.S.¹,²
Tony Hsiu-Hsi Chen, Ph.D.³
Jen-Pei Liu, Ph.D.⁴,⁵
Cheng-Chuan Chang, M.D.¹
Wei-Chu Chie, M.D., Ph.D.³

¹ Department of Otorhinolaryngology, Head and Neck Surgery, Changhua Christian Hospital, Changhua, Taiwan.
² Institute of Medicine, Chung-Shan Medical University, Taichung, Taiwan.
³ Graduate Institute of Preventive Medicine and Department of Public Health, College of Public Health, National Taiwan University, Taipei, Taiwan.
⁴ Division of Biostatistics and Bioinformatics, National Health Research Institutes, Taipei, Taiwan.
⁵ Department of Statistics, National Cheng-Kung University, Tainan, Taiwan.

Supported by the National Science Council, Taiwan (grants NSC89-2320-B-002-096 and NSC90-2320-B-002-177).

Address for reprints: Wei-Chu Chie, M.D., Ph.D., Graduate Institute of Preventive Medicine and Department of Public Health, College of Public Health, National Taiwan University, Room 209, 19 Huachow Road, Taipei, 10020, Taiwan; Fax (011) 886 223920466; E-mail: weichu@episerv.cph.ntu.edu.tw

Received August 11, 2003; revision received November 21, 2003; accepted February 18, 2004.

BACKGROUND. Heterogeneity of primary tumor volume within tumors of the same classification indicates a need to elucidate the effects of primary tumor volume on treatment outcomes in patients with nasopharyngeal carcinoma (NPC).

METHODS. From 1994 through 1996, 129 patients with newly diagnosed NPC who were treated with high-dose radiotherapy were enrolled in the study. Computed tomography-derived primary tumor volume was measured using the summation-of-area technique. Correlations between American Joint Committee on Cancer (AJCC) disease stage, primary tumor volume, and disease-specific survival were assessed using a Cox regression model. Cross-validation based on receiver operating characteristic (ROC) curve also was examined.

RESULTS. Compared with the AJCC staging system and the TNM classification system, primary tumor volume was better at determining cumulative survival for patients with NPC. Hazard ratios increased with tumor volume, ranging from 6.68 (95% confidence interval [95% CI], 1.89–23.67) for tumor volumes between 20–40 mL, 18.03 (95% CI, 4.80–67.75) for tumor volumes between 40–60 mL, and 26.06 (95% CI, 7.70–88.20) for tumor volumes > 60 mL. With both tumor volume and T classification in the same Cox regression model, only tumor volume remained statistically significant in the prognosis of NPC. The validation results with ROC curves also revealed that, in predicting patient outcome, primary tumor volume (area under the ROC = 83.33%) was superior to disease stage (area under the ROC = 66.53%) and TNM classification (area under the ROC = 58.61%).

CONCLUSIONS. The incorporation of primary tumor volume may lead to a further refinement of the current AJCC staging system, particularly for patients with large primary tumor volumes (> 60 mL), who require more aggressive treatment. Cancer 2004;100:2160–6. © 2004 American Cancer Society.

KEYWORDS: nasopharyngeal carcinoma (NPC), primary tumor volume, TNM stage, survival.

Nasopharyngeal carcinoma (NPC) is preponderant in Taiwan, with annual incidence rates of 8.68 per 100,000 for males and 3.54 per 100,000 for females, respectively.¹ It also is ranked as the seventh leading cause of death among males according to a cancer registry report in 1998.¹ Treatment for patients with NPC primarily relied on radiotherapy, partly because of its inaccessible anatomic site and partly because of the high likelihood of early lymph node metastases with possible involvement of the node of Rouviere, which is hard to remove with surgical methods. Like other malignancies, a variety of staging systems were developed for predicting the prognosis of patients and to select appropriate treatment modalities. The most widely used systems include the American Joint Committee of Cancer...
(AJCC) staging system,2 the International Union Against Cancer staging system,3 and the Ho staging system.4 For classifying and staging NPC, all of these systems are based on anatomic location and cranial nerve involvement. Arguably, no system is impeccable, and all have limitations. Consequently, the AJCC has suggested revising T-stage classification with the incorporation of additional (further) clinical information.

It has been reported that tumor bulk, representing the number of tumor clonogens that need to be sterilized, is an important prognostic factor for patients who are treated with primary radiation therapy. Although tumor bulk has been used for predicting the prognosis of patients with other types of malignancies,5–7 to our knowledge its application in patients with NPC barely has been addressed. Our previous study showed that primary tumor volume was heterogeneous within the same disease stage and T-classification, particularly in patients with advanced-stage disease.8 In patients with advanced NPC, we also measured the primary tumor volume, which is regarded as an important prognostic factor for treatment outcome.9 However, whether primary tumor volume is a significant prognostic factor for patients after primary radiation therapy remains unclear. There also is a lack of empiric evidence regarding the comparison between primary tumor volume and the most recent AJCC TNM staging system with respect to clinical prognosis.

The objectives of this study were 1) to correlate primary tumor volume with disease stage and tumor classification according to the 1997 AJCC staging system for patients with NPC and 2) to elucidate the impact of tumor volume on the prognosis of patients with NPC after primary radiotherapy. The results may provide empiric evidence for refining both the TNM staging system and the clinical management of patients with NPC.

MATERIALS AND METHODS

Study Patients

A retrospective cohort was ascertained, including 162 patients with NPC who were treated with high-dose radiotherapy (> 70 grays [Gy]) with or without chemotherapy between 1994 and 1996. Thirty-three of 162 patients were not eligible for analyses because of the presence of distant metastasis at the time of presentation (n = 16 patients), incomplete radiotherapy course (n = 8 patients), artifacts on computed tomography (CT) scans that resulted in difficulty with delineating tumor contours (n = 4 patients), loss to follow-up (n = 3 patients), and incomplete baseline CT information (n = 2 patients).

Tumor Volume Measurement

All patients in this study had pretreatment, contrast-enhanced CT scans that were done along the axial scan plain parallel to the infraorbital-meatal line extending from the skull base to the top of the manubrium with contiguous, 5-mm scanning and a 25-cm field of view. In addition, direct coronal scans also were taken to provide an auxiliary diagnosis. One hundred milliliters of contrast medium were administered, using an injector with an injection rate between 1.0 mL per second and 2.5 mL per second after an initial 5-mL dose. The CT scans were redigitized using a film scanner. Using applied computer software, the images were visualized easily seen on a large-screen monitor. User-defined magnification of the images was possible. The images were offered in a prefixed window/center setting that was accrued from a prior pilot study.8 Users also were able to change this setting according to their own needs. For tumor characteristics, clinicians kept track of the tumor contour manually. The volume was calculated by the summation-of-areas technique, which multiplies the entire areas by the image reconstruction interval. A pilot study10 showed low interobserver and intraobserver variability. In the current study, one author (M.-K.C.) was the major assessor of tumor volumes and was assisted by another author (C.C.C.). A radiologist who specialized in head and neck malignancies participated when the judgments of the two authors differed.

Radiotherapy

All patients completed a course of radiotherapy, with radiation doses of 70–76 Gy. Irradiation was given primarily with megavoltage (6–10 megavolts [MV]) X-rays from linear accelerators. The dose was 1.8–2.0 Gy per fraction. Each patient was administered with 5 fractions per week. All patients were assessed by nasopharyngoscopy at intervals of 10–14 weeks after the completion of radiotherapy. Patients who achieved a complete response to treatment were followed regularly. Postoperative clinical examinations included indirect mirror examinations of the nasopharynx with or without nasopharyngoscopy. Chest X-rays, abdominal sonography, whole-body bone scans, and body CT scans were performed only when they were indicated clinically.

Statistical Methods

Because the primary endpoint was death from NPC, all 129 patients were linked to a mortality registry until December 31, 1999, to ascertain all deaths. Survival was used as the outcome measure and was calculated as the duration from treatment to the date of death.
caused by NPC, or to the date of censoring due to other causes of death, or to loss to follow-up, or to the end of follow-up (December 31, 1999). Cumulative survival rates were calculated by the Kaplan–Meier method. A log-rank test was used to test the difference between cumulative survival rates with respect to risk groups according to primary tumor volume, tumor classification, lymph node classification, and disease stage, respectively. A Cox proportional hazards regression model also was applied to assess the correlations between different staging systems and the prognosis of patients NPC after they received primary radiotherapy while controlling for age, gender, and chemotherapy. Note that, because our major objective was to compare the predictive validity of different staging systems associated with death from NPC, receiver operating characteristic (ROC) analysis also was applied. The procedure for ROC curve analysis was as follows: First, the total samples were allocated randomly into two groups, with two-thirds test samples and one-third validated samples. Regression coefficients were trained using test samples. Parameters trained from test samples were applied further to the validated samples to predict death status for each individual. The numbers of predicted deaths were then compared with the numbers of observed deaths. Whether the predicted values were comparable to the observed values was assessed by goodness of fit based on Pearson chi-square or deviance tests. Sensitivity and false-positive rates were calculated by comparing observed and predicted values using validated samples, and ROC curves were constructed with sensitivity and false-positive rates as the x-axis and the y-axis, respectively. The area under ROC curve was used to assess the predicted validity of different staging systems. Calculations were based on the method of Hanley and McNeil. All stages of disease and tumor volumes were included in the survival analysis and were cross-validated to make comparisons between advanced and early tumors in the prediction of survival.

RESULTS

The mean follow-up (± standard deviation) was 37.9 months (± 20.0 months) for all patients, 53.7 months (± 11.1 months) for survivors, and 24.6 months (± 15.5 months) for patients who died before the end of the study. Patient characteristics and distributions of primary tumor volume associated with tumor classification, lymph node classification, and disease stage are presented in Table 1. The median primary tumor volume was 5.48 mL in patients with T1 tumors, 18.55 mL in patients with T2 tumors, 29.60 mL in patients with T3 tumors, and 54.07 mL in patients with T4 tumors, with a range of 3.23–9.65 mL in T1 tumors, 6.31–131.82 mL in T2 tumors, 8.03–131.82 mL in T3 tumors, and 6.70–223.0 mL in T4 tumors. The median primary tumor volumes in patients with Stage I, Stage II, Stage III, Stage IVA, and Stage IVB disease were 5.48 mL, 17.32 mL, 29.77 mL, 54.60 mL, and 30.54 mL, respectively. Primary tumor volume was heterogeneous for patients in all disease stages, particularly for patients with advanced-stage disease.

The cumulative survival rates by disease stage, TNM classification, and primary tumor volume are shown in Figures 1–3. Figure 1 shows that patients with advanced disease stage showed a poorer cumulative survival rate. Patients with Stage I tumors using the TNM classification system or T-stage had nearly 100% survival after 6 years of follow-up. Cumulative

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total no.</th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
<th>Stage IVA</th>
<th>Stage IVB</th>
<th>Total</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no.</td>
<td>129</td>
<td>5</td>
<td>28</td>
<td>39</td>
<td>28</td>
<td>29</td>
<td>129</td>
<td>—</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>42 (32.6)</td>
<td>0 (0.0)</td>
<td>12 (42.9)</td>
<td>14 (35.9)</td>
<td>8 (28.6)</td>
<td>8 (27.6)</td>
<td>42 (32.6)</td>
<td>0.394</td>
</tr>
<tr>
<td>Male</td>
<td>87 (67.4)</td>
<td>5 (100.0)</td>
<td>16 (57.1)</td>
<td>25 (64.1)</td>
<td>20 (71.4)</td>
<td>21 (72.4)</td>
<td>87 (67.4)</td>
<td></td>
</tr>
<tr>
<td>Age (yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>49.8 ± 11.8</td>
<td>45.1 ± 14.3</td>
<td>47.1 ± 10.2</td>
<td>50.1 ± 12.5</td>
<td>51.7 ± 11.4</td>
<td>51.0 ± 12.6</td>
<td>49.8 ± 11.8</td>
<td>0.629</td>
</tr>
<tr>
<td>Tumor volume (mL)</td>
<td>49.16 ± 39.43</td>
<td>5.86 ± 2.39</td>
<td>20.21 ± 12.26</td>
<td>40.05 ± 29.53</td>
<td>60.52 ± 43.80</td>
<td>49.16 ± 53.16</td>
<td>40.91 ± 39.43</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>≤ 20</td>
<td>46 (35.7)</td>
<td>5 (100.0)</td>
<td>19 (67.9)</td>
<td>8 (28.6)</td>
<td>3 (10.7)</td>
<td>11 (37.9)</td>
<td>46 (35.7)</td>
<td>0.001</td>
</tr>
<tr>
<td>20–40</td>
<td>41 (31.8)</td>
<td>0 (0.0)</td>
<td>7 (25.0)</td>
<td>18 (66.7)</td>
<td>9 (32.1)</td>
<td>7 (24.1)</td>
<td>41 (31.8)</td>
<td></td>
</tr>
<tr>
<td>40–60</td>
<td>17 (13.2)</td>
<td>0 (0.0)</td>
<td>2 (7.1)</td>
<td>5 (17.2)</td>
<td>4 (14.3)</td>
<td>6 (20.7)</td>
<td>17 (13.2)</td>
<td></td>
</tr>
<tr>
<td>&gt; 60</td>
<td>25 (19.4)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>8 (28.6)</td>
<td>12 (42.9)</td>
<td>5 (17.2)</td>
<td>25 (19.4)</td>
<td></td>
</tr>
</tbody>
</table>

SD: standard deviation.
survival curves using the TNM classification system were difficult to separate at the beginning of follow-up. The cross-over phenomenon disappeared using the T classification (Fig. 2). However, cumulative survival rates in the beginning of follow-up still were fraught with heterogeneity. Using 4 categories of primary tumor volume (\(\leq 20\) mL, 20–40 mL, 40–60 mL, and \(> 60\) mL), cumulative survival curves were separated clearly (Fig. 3).

Table 2 shows adjusted hazard ratios for the classifications of TNM stage and primary tumor volume after controlling for age, gender, and chemotherapy. For TNM classification, the T classification had better prediction than the N classification. Taking both into consideration, prediction became better but still was poor in patients with advanced tumors. Compared with TNM classification, adjusted hazard ratios decreased with an increase of tumor volume, indicating a gradient correlation. This suggests that primary tumor volume may be a better predictor of prognosis for patients with NPC. After adjusting for the TNM classification and primary tumor volume in one another, the primary tumor volume remained an independent prognostic factor, but the T classification became insignificant (Table 3).

To verify whether the predictive validity for primary tumor volume was better compared with TNM
classification and disease stage, cross-validation, as discussed above is shown in Figure 4. The results revealed that, in predicting outcomes, the primary tumor volume was superior to disease stage and TNM classification. The area under the ROC curve was 83.33% (95% CI, 76.38–90.29%) for primary tumor volume compared with 66.53% (95% CI, 57.04–76.01%) for disease stage and 58.61% (95% CI, 48.07–69.15%) for TNM classification.

**DISCUSSION**

**Reliability of Primary Tumor Volume Measurement**

The first contribution of the current study is that we assured the reliability of primary tumor volume measurement. Before CT scanning was available in clinical practice, tumor volume measurements were difficult because of intraobserver and interobserver variability. The findings of Hermans et al. regarding the variability of tumor volume measurements for laryngeal tumor suggest that it is mandatory to assure the reliability of primary tumor volume measurement for NPC before elucidating the relation between tumor volume and prognosis for patients with NPC. In the studies of Chua et al. and Willner et al., a correlation between tumor volume and prognosis was reported, but those studies did not analyze the reliability or validity of tumor volume measurement. In our previous pilot study, we assured interobserver and intraobserver reliability. Based on the results from that study, our measurements of primary tumor volumes are reliable.

**Primary Tumor Volume and Prognosis of Patients with NPC after Radiotherapy**

The results of the current study showed that substantial variation in primary tumor volume was observed in patients with all disease stages and overlapped among patients with different stages of disease, especially advanced stages, similar to the results from previous studies. This indicates the limitation of using the current TNM classification system (as revised in 1997) to separate large tumors from small tumors, particularly in patients with advanced NPC, because staging according to the TNM system is subjective and unidimensional, and it fails to define the real, three-dimensional bulk, which is measured by tumor volume.

We also examined the predictive power of survival for primary tumor volume and disease stage by including both variables in the same Cox multiple regression model, adjusting for age, gender, and chemotherapy status. Only primary tumor volume remained statistically significant, whereas the effect of disease stage/T classification on the prognosis of survival disappeared. Our results were congruent with earlier studies.

**Predictive Validity of Primary Tumor Volume**

Another major objective of this study was to verify the adequacy of the model using primary tumor volume and disease stage, respectively, through cross-validation. After the application of parameters estimated from the trained data set to the validated data, we found that the ratios of sensitivity to specificity in ROC curves, given different cut-off points,
were higher in the model using primary volume compared with the model using disease stage or TMN classification. This suggests that the predictive validity associated with the prognosis of patients with NPC on the basis of primary tumor volume is favorable.

Limitations and Future Directions of Study

In the current study, we used CT scans to assess primary tumor volume. In patients with head and neck carcinoma, it has been reported that magnetic resonance imaging (MRI) is superior to CT in detecting tumor extension into soft tissue, separation of tumor from mucous, and bone marrow invasion and is more sensitive in assessing for skull base extension and intracranial spread in patients with advanced disease. Rasch et al. indicated that MRI-derived tumor volume is smaller and has less interobserver variation compared with CT-derived tumor volume; however, CT cannot be neglected and may become complementary to MRI in delineating tumor volume.\(^\text{16}\) CT artifacts inherent to the skull base, such as artifacts from dental hardware, and the overlapping densities between tumor, edema, and normal anatomic structures, such as muscle, make CT less than optimal, although CT is preferred for the visualization of bone cortex invasion.\(^\text{17–19}\) MRI also suffers poorer geometric accuracy of the patient contour and lack of electron density information.\(^\text{20,21}\) In the face of pros and cons for the two imaging techniques, we did not intend to adopt CT as a complement to MRI in delineating tumor volume. A technique for MRI-derived tumor volume assessment was not available in Taiwan when this study was conducted. We will compare the two techniques for tumor volume delineation and prognosis prediction in future studies. Primary tumor volume has been taken into consideration in treatment planning.\(^\text{22,23}\) We will plan radiotherapy and chemotherapy based on the information of primary tumor volume rather than conventional staging, especially for patients with large and advanced tumors.

Conclusions

Tumor volumes in patients with NPC are heterogeneous in the concurrent T-classification. The incorporation of primary tumor volume may lead to a further refinement of the current AJCC staging system, particularly for patients with large primary tumor volumes (> 60 mL), who require more aggressive treatment.

REFERENCES


