A Scoring System for Prediction of Cervical Lymph Node Metastasis in Patients with Head and Neck Squamous Cell Carcinoma

M.S. Chung, Y.J. Choi, S.O. Kim, Y.S. Lee, J.Y. Hong, J.H. Lee and J.H. Baek

AJNR Am J Neuroradiol 2019, 40 (6) 1049-1054
doi: https://doi.org/10.3174/ajnr.A6066
http://www.ajnr.org/content/40/6/1049
A Scoring System for Prediction of Cervical Lymph Node Metastasis in Patients with Head and Neck Squamous Cell Carcinoma

M.S. Chung, Y.J. Choi, S.O. Kim, Y.S. Lee, J.Y. Hong, J.H. Lee, and J.H. Baek

ABSTRACT

BACKGROUND AND PURPOSE: An accurate and comprehensive assessment of lymph node metastasis in patients with head and neck squamous cell cancer is crucial in daily practice. This study constructed a predictive model with a risk scoring system based on CT characteristics of lymph nodes and tumors for patients with head and neck squamous cell carcinoma to stratify the risk of lymph node metastasis.

MATERIALS AND METHODS: Data included 476 cervical lymph nodes from 191 patients with head and neck squamous cell carcinoma from a historical cohort. We analyzed preoperative CT images of lymph nodes, including diameter, ratio of long-to-short axis diameter, necrosis, conglomeration, infiltration to adjacent soft tissue, laterality and T-stage of the primary tumor. The reference standard comprised pathologic results. Multivariable logistic regression analysis was performed to develop the risk scoring system. Internal validation was performed with 1000-iteration bootstrapping.

RESULTS: Shortest axial diameter, ratio of long-to-short axis diameter, necrosis, and T-stage were used to develop a 9-point risk scoring system. The risk of malignancy ranged from 7.3% to 99.8%, which was positively associated with increased scores. Areas under the curve of the risk scoring systems were 0.886 (95% CI, 0.881–0.920) and 0.879 (95% CI, 0.845–0.914) in internal validation. The Hosmer–Lemeshow goodness-of-fit test indicated that the risk scoring system was well-calibrated ($P = .360$).

CONCLUSIONS: We developed a comprehensive and simple risk scoring system using CT characteristics in patients with head and neck squamous cell carcinoma to stratify the risk of lymph node metastasis. It could facilitate decision-making in daily practice.

ABBREVIATIONS: AUC = area under the curve; HNSCC = head and neck squamous cell carcinoma; L/S = long-to-short axis

The presence of lymph node metastasis has a great impact on the treatment and prognosis in patients with head and neck squamous cell carcinoma (HNSCC). Solitary lymph node metastasis from HNSCC has a 5-year survival rate of 50%; additional contralateral lymph node metastasis reduces survival to 33%. Detection of lymph node metastasis by imaging is more accurate than by clinical examination; thus, performing preoperative CT or MR imaging in the initial work-up for HNSCC has become routine.

Imaging assessment of lymph node metastasis in the head and neck can be challenging for the radiologist because there are multiple cervical levels to review and variable suggested criteria for metastatic lymph nodes. Furthermore, multiple image features of lymph nodes (including diameters, shapes, and presence of necrosis), combinations of those features, and characteristics of primary tumor (such as T-stage or location) should be considered to determine the possibility of metastasis in daily practice. In particular, there is greater difficulty in cases that show both benign and probable abnormal features. Therefore, a comprehensive and systematic approach, based on a combination of previously proposed criteria, is needed to maximize the benefits of preoperative CT.

Furthermore, additional image evaluation and pathologic confirmations are not possible in all cases (eg, deep locations, including retropharyngeal or upper mediastinal lymph nodes);
these increase medical costs and procedure-related complications.\textsuperscript{6,7} MR imaging also may provide an option for the diagnosis of lymph node metastasis due to its higher spatial resolution and advanced techniques, such as diffusion-weighted imaging, whereas the real added value of MR imaging combined with other imaging modalities remains controversial despite high costs.\textsuperscript{8,9} PET/CT is known to have potential diagnostic and prognostic roles in patients with HNSCC,\textsuperscript{10-14} and PET/CT has achieved approximately 21% sensitivity improvement in the diagnosis of nodal metastasis per-neck-level compared with conventional images.\textsuperscript{15} However, considering false-positive results caused by inflammatory conditions and false-negative results due to small size, necrosis, or cystic change of metastatic lymph nodes using PET/CT, a proper diagnostic approach based on the conventional image findings should be essential.\textsuperscript{10,16}

Accordingly, a practical and comprehensive prediction model that uses variable CT findings could be useful for assessing the presence of lymph node metastasis in patients with HNSCC in daily practice and might maximize the efficiency of CT by enhancing the understanding of variable CT features. A risk scoring system estimates the probability of the presence/occurrence of a particular event, based on multiple predictors, to facilitate individual diagnostic and therapeutic decision-making.\textsuperscript{17} Therefore, this study aimed to develop a simple prediction model using a risk scoring system in patients with HNSCC, based on CT findings, to stratify the risk of lymph node metastasis.

**MATERIALS AND METHODS**

This retrospective study was approved by our institutional review board at Asan Medical Center, and informed consent was waived for data evaluation. The methods and reporting of results are in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement,\textsuperscript{18} the Strengthening the Reporting of Observational Studies in Epidemiology statement,\textsuperscript{19} and the statistical methods for prediction models.\textsuperscript{17}

**Study Population**

The study population was obtained from a historical cohort of consecutive patients who were newly diagnosed with HNSCC and underwent pretreatment contrast-enhanced neck CT at our large (2700 beds) academic tertiary referral hospital between July 2010 and December 2013. The eligibility criteria were as follows: 1) Patients had newly diagnosed HNSCC; 2) patients underwent a neck CT examination before the treatment of HNSCC; 3) patients underwent lymph node dissection, and the presence of lymph node metastasis was pathologically confirmed at each cervical neck CT examination before the treatment of HNSCC; 3) patients underwent radiation therapy and/or chemotherapy before definite radiation or concurrent chemoradiation therapy; 2) patients were older than 20 years of age.

Exclusion criteria included the following: 1) Patients underwent definite radiation or concurrent chemoradiation therapy; 2) patients underwent radiation therapy and/or chemotherapy before lymph node neck dissection; 3) CT scan slice thickness was >3 mm; 4) CT scans were without coronal reconstruction; and 5) CT scans had severe metal- or motion-related artifacts, making tumor boundary delineation difficult. We also retrospectively collected patients’ demographic and clinical data by review of electronic medical records; these data included age, sex, location of the mass, and surgical and chemoradiotherapy history.

**Image Acquisition**

All patients underwent CT examinations on one of several commercially available CT systems, with the multidetector capability ranging from 64 to 128 channels. The techniques and parameters varied depending on the system used. However, most examinations were performed with a 128-channel CT scanner (Somatom Definition Flash; Siemens Medical Solutions, Erlangen, Germany), and typical imaging parameters were as follows: 120 kV, 200 effective mAs, axial scan mode, 22-cm display FOV, 50-cm large-body scan FOV, pitch of 1, gantry rotation time of 0.5 seconds, detector collimation of 128 × 0.6 mm, and 3-mm axial and coronal reconstructed slice thickness with a soft-tissue algorithm reconstruction. Real-time automatic tube current modulation software (CARE Dose4D; Siemens) was applied to regulate the tube current in accordance with the patient’s anatomic structures. Scan coverage was from the upper margin of the frontal sinus to the top of the aortic arch. Acquisition of CT images began 70 seconds after the injection of 140 mL of intravenous iopamidol (Isovue-370; Bracco, Princeton, New Jersey) at a rate of 2.5 mL/s. A PACS system was used for the review and analysis of all radiologic images.

**Image Analysis**

All CT images were independently reviewed by 2 neuroradiologists, with 5 and 12 years of head and neck oncology imaging experience, respectively. Before evaluation, the 2 neuroradiologists completed a training session with 10 patients to help them reach a consensus regarding the measurement of the diameter and imaging findings of the lymph nodes.

We analyzed possible predictors of lymph node metastasis on preoperative CT images as follows: lymph node diameter (shortest and longest axial diameter and longest coronal diameter);\textsuperscript{3} the long-to-short axis (L/S) ratio;\textsuperscript{2,5,20} presence of necrosis;\textsuperscript{3,21} conglomeration of lymph nodes, infiltration to adjacent soft tissue, laterality to the primary tumor (ipsilateral versus contralateral); and T-stage of primary tumor, following the AJCC Cancer Staging Manual, eighth edition. The largest diameters of each node in the axial and coronal planes were measured and defined as the longest diameter.\textsuperscript{5} The largest diameter perpendicular to the longest axial diameter was measured and defined as the shortest axial diameter.\textsuperscript{5} Slices for the measurement of the diameter of the lymph nodes were independently selected by each observer, and electronic calipers on the PACS system were used for such measurements. The L/S ratio was calculated by using the longest axial diameter/shortest axial diameter.\textsuperscript{5} The presence of necrosis was defined as central low density with irregular or rim-like enhancement of residual lymphatic tissue.\textsuperscript{21,22} The degree of necrosis was classified as none, present, or cystic (demonstrating a rim-like thin enhancing or imperceptible wall with >90% of central low density) by visual analysis.\textsuperscript{1} Infiltration to adjacent soft tissue was defined as poorly defined nodal margins or soft-tissue infiltration or stranding of the muscles or fat in the neck.\textsuperscript{22}
Table 1: Univariable and multivariable regression models of lymph node metastasis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariable Analysis</th>
<th>Multivariable Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>T-stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.13</td>
<td>0.68–1.89</td>
</tr>
<tr>
<td>2</td>
<td>1.59</td>
<td>0.98–2.59</td>
</tr>
<tr>
<td>3 or 4</td>
<td>4.12</td>
<td>18.09–94.27</td>
</tr>
<tr>
<td>Shortest axial diameter (cm)</td>
<td>6.59</td>
<td>4.12–10.56</td>
</tr>
<tr>
<td>Longest axial diameter (cm)</td>
<td>4.87</td>
<td>3.36–7.08</td>
</tr>
<tr>
<td>L/S ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longest longitudinal diameter (cm)</td>
<td>4.87</td>
<td>3.36–7.08</td>
</tr>
<tr>
<td>Necrosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>38.99</td>
<td>17.81–85.37</td>
</tr>
<tr>
<td>Cystic</td>
<td>82.25</td>
<td>24.88–717.89</td>
</tr>
<tr>
<td>Infiltration*</td>
<td>75.54</td>
<td>10.26–556.09</td>
</tr>
<tr>
<td>Latency to primary tumor (ipsilateral vs contralateral)*</td>
<td>2.27</td>
<td>1.37–3.75</td>
</tr>
</tbody>
</table>

*The reference standard in this study comprised no conglomeration, no infiltration, and no contralateral location to the primary tumor.

Statistical Analysis

For the development of a risk scoring system for the discrimination of lymph node metastasis, pathologic results after lymph node dissection were used as a reference standard for the presence or absence of metastasis of lymph nodes. Univariable and multivariable logistic regression analysis was performed to estimate the risk of metastasis associated with CT findings and pathologic results. Variable selection for the multivariable model was achieved by iterative backward elimination in 1000-iteration bootstrap resampling. A 50% relative frequency of selection was the criterion for inclusion in the final model.

A simple scoring system was developed using the penalized maximum likelihood estimates of the covariates in models that followed the method of Sullivan et al.23 After selecting a base value of each variable, we used regression coefficients as weights and distance from the base value to generate each point value (On-line Table). Score 1 was defined as the effect of a 5-mm increase in the minimal axial diameter of the lymph node. To prove the power of discrimination of the risk scoring system, we calculated the area under the curve (AUC) of the receiver operating characteristic curve. The Hosmer–Lemeshow goodness-of-fit statistic was used to evaluate the agreement between the observed and expected number of metastatic lymph nodes across all strata, based on the probabilities of malignancy estimated from the prediction model.

Internal validation was performed using the bootstrap validation algorithm.24-25 Bootstrap resampling began by fitting the logistic model in a bootstrap sample of the same number of nodules as the original sample (n = 476), which was drawn with replacement from the original sample. The optimism-corrected performance for internal validation was calculated as the AUC of the receiver operating characteristic curve with 1000 repetitions. The apparent performance represented the performance in the original sample. The performance in each bootstrap sample represents the apparent performance of the bootstrap model, and the test performance represents the performance of the bootstrap model in the original sample. The difference between these performances is an estimate of the optimism. The optimism performance is subtracted from the apparent performance to estimate the internally validated performance: Optimism-Corrected Performance = Apparent Performance – Average (Bootstrap Performance – Test Performance).24-26 All statistical analyses were performed using SPSS, Version 22.0 (IBM, Armonk, New York) and R, Version 3.0.2 (www.r-project.org), with rms and pROC packages. A 2-sided P value < .05 was considered statistically significant.

RESULTS

This study included 476 lymph nodes from 191 patients (142 men, 49 women; mean age, 61.8 ± 13.0 years; range, 28–91 years). Metastatic lymph nodes were present in 38% (181/476) of lymph nodes. Location of the primary tumor was the oral cavity in 48.2% (92/191), oropharynx in 23.0% (44/191), and larynx in 28.8% (55/191). The T-stage of the primary tumor was stage 1 in 25.7% (49/191), stage 2 in 33.5% (64/191), stage 3 in 29.3% (56/191), and stage 4 in 11.5% (22/191).

By multivariate logistic regression analysis, the shortest axial diameter of the lymph node and necrosis of the lymph node demonstrated significant differences between benign and metastatic lymph nodes. Necrosis of the lymph node showed the highest odds ratio of >15 (Table 1).

A 9-point risk scoring system was developed using the results of multivariable analysis (Table 2, On-line Table, and Fig 1). The T-stage of the primary tumor, shortest axial diameter, L/S ratio, and necrosis of the lymph node were assigned to one or more points, according to their β (regression coefficient) values in the development of the risk scoring system. The presence of necrosis in the lymph nodes, regardless of the extent of necrosis (ie, present versus cystic), and the shortest axial diameter of >2 cm showed relatively high scores in this system: 3 points and 4 points, respectively. Table 2 also shows the risk of metastasis according to the risk scoring system. If a lymph node had a score of zero, the risk of metastasis was <7.0%. Furthermore, the risk of metastasis continuously increased when the risk score increased, with 9 points achieving the highest value of 100% of the estimated probability of metastasis.

The AUC of the risk scoring system was 0.886, with a 95% CI, 0.881–0.920 (Fig 2). The sensitivity and specificity of the risk scoring system were 74.0% (95% CI, 67.2%–79.9%) and 95.3% (95% CI, 92.2%–97.2%), using the cutoff value of 3 points. The optimism-corrected performance was good (AUC = 0.879; 94% CI, 0.845–
Risk scoring systems have been used to stratify the risk of lymph node metastasis in various organs, such as esophageal cancer, gastric cancer, melanoma, endometrial cancers, and thyroid cancer.27–33 Similar to these tumors, the existence and extent of lymph node metastasis change the extent of lymph node dissection and therapeutic plans in patients with HNSCC; consequently, accurate decisions for lymph node metastasis are crucial. However, almost all pre-existing modalities have variable diagnostic accuracy (sensitivities range from 14% to 80% for CT and from 29% to 85% for MR imaging; specificities range from 80% to 100% for both CT and MR imaging) for detecting node metastases in HNSCC.3,5,6,8,34 Our prediction model, which provides risk scores for the evaluation of lymph nodes, may provide objective evidence for diagnosis and might contribute to reduced interobserver variability.

PET/CT is known to have an advantage in staging the nodal status of HNSCC.10–15 A recent meta-analysis reported that the sensitivity, specificity, and AUC of PET/CT were 84% (95% CI, 71%–87%), 96% (95% CI, 94%–97%), and 0.97 (95% CI, 0.95–0.98) with per-neck-level data.15 PET/CT showed the improvement of per-neck-level sensitivity by 21% over conventional imaging.15 The sensitivity, specificity, and AUC of our scoring system are slightly lower than those of PET/CT and similar to those of conventional CT.

However, our risk scoring system may apply to several clinically complex circumstances for radiologists. First, there are several cases in which it is technically impossible to perform ultrasonography and biopsy. Although ultrasonography and ultrasonography-guided biopsy are very powerful techniques for the assessment of indeterminate lymph nodes in patients with cancer, technical limitations exist in deep lymph nodes, such as retropharyngeal, mediastinal, and lower level 6 lymph nodes, as well as in situations that involve poor patient conditions or low patient compliance.6,7 Second, a risk scoring system could provide a more targeted review for radiologists in an additional imaging study. Rather than focusing on high-score (definite malignant) or low-score (definite benign) lymph nodes, the additional evaluation should focus on lymph nodes with intermediate risk scores. Furthermore, our risk scoring system integrates variable pre-existing CT criteria for lymph node metastasis and primary tumors in HNSCC. Given these advantages, our risk scoring system could enhance the efficiency of preoperative CT by supporting reasonable decision-making based on estimated malignancy.

Our present study had several limitations of note. First, some selection bias could have been introduced because our study was performed at a single center with a small number of patients. Further validation studies with larger populations are warranted. Second, we did not evaluate the potential for interobserver variability in interpretations of CT images between radiologists. However, the radiologists who participated in the study had >5 years of clinical experience.
in head and neck oncology imaging, and both had completed an image training session to perform more reproducible and accurate image analysis within the parameters of the study protocol. Third, we focused on the development of the prediction model and did not perform external validation. However, as we previously described, our risk scoring system might be helpful, at least in several clinically complicated circumstances. The clinical implications of our study should be validated in further analyses. Fourth, we enrolled the patients who underwent a curative operation because we used pathologic specimens after neck dissection as a reference standard to match image findings in each lymph node. Therefore, some degree of selection bias was also included because we did not include patients who were treated with definite radiation therapy or concurrent chemoradiotherapy. Finally, we focused on the image findings of enhanced neck CT, without correlation of the results of other modalities. As we previously described, many studies suggest the potential diagnostic and prognostic roles of PET/CT in patients with HNSCC, and some countries and referral centers routinely perform PET/CT for the initial work-up. Further studies may reveal whether adding this scoring system to the results of PET/CT could improve the sensitivity and specificity of determining neck nodal disease.

CONCLUSIONS

A simple 9-point risk scoring system using CT characteristics of lymph nodes and tumors for HNSCC could be feasible to stratify the risk of cervical lymph node metastasis with high diagnostic accuracy. This comprehensive and practical risk scoring system, based on various CT features, could be helpful for decision-making regarding the possibility of lymph node metastasis in patients with HNSCC in daily practice.

Disclosures: Jung Hwan Baek—UNRELATED: Consultancy: RF Medical and StarMed since 2017.

REFERENCES


