JAMA Otolaryngology–Head & Neck Surgery | Original Investigation | AI IN OTOLARYNGOLOGY

Artificial Intelligence Model for Imaging-Based Extranodal Extension Detection and Outcome Prediction in Human Papillomavirus–Positive Oropharyngeal Cancer

Gabriel S. Dayan, MD; Gautier Hénique, MSc; Houda Bahig, MD, PhD; Kristoff Nelson, MD; Coralie Brodeur, BEng; Apostolos Christopoulos, MD, MSc; Edith Filion, MD; Phuc-Felix Nguyen-Tan, MD; Brian O'Sullivan, MD; Tareck Ayad, MD; Eric Bissada, MD, DMD; Paul Tabet, MD, MSc; Louis Guertin, MD; Antoine Desilets, MD, MSc; Samuel Kadoury, PhD; Laurent Letourneau-Guillon, MD, MSc

IMPORTANCE Although not included in the eighth edition of the *American Joint Committee on Cancer Staging System*, there is growing evidence suggesting that imaging-based extranodal extension (iENE) is associated with worse outcomes in HPV-associated oropharyngeal carcinoma (OPC). Key challenges with iENE include the lack of standardized criteria, reliance on radiological expertise, and interreader variability.

OBJECTIVE To develop an artificial intelligence (AI)–driven pipeline for lymph node segmentation and iENE classification using pretreatment computed tomography (CT) scans, and to evaluate its association with oncologic outcomes in HPV-positive OPC.

DESIGN, SETTING, AND PARTICIPANTS This was a single-center cohort study conducted at a tertiary oncology center in Montreal, Canada, of adult patients with HPV-positive cN^{+} OPC treated with up-front (chemo)radiotherapy from January 2009 to January 2020. Participants were followed up until January 2024. Data analysis was performed from March 2024 to April 2025.

EXPOSURES Pretreatment planning CT scans along with lymph node gross tumor volume segmentations performed by expert radiation oncologists were extracted. For lymph node segmentation, an nnU-Net model was developed. For iENE classification, radiomic and deep learning feature extraction methods were compared.

MAIN OUTCOMES AND MEASURES iENE classification accuracy was assessed against 2 expert neuroradiologist evaluations using area under the receiver operating characteristic curve (AUC). Subsequently, the association of AI-predicted iENE with oncologic outcomes—ie, overall survival (OS), recurrence-free survival (RFS), distant control (DC), and locoregional control (LRC)—was assessed.

RESULTS Among 397 patients (mean [SD] age, 62.3 [9.1] years; 80 females [20.2%] and 317 males [79.8%]), AI-iENE classification using radiomics achieved an AUC of 0.81. Patients with AI-predicted iENE had worse 3-year OS (83.8% vs 96.8%), RFS (80.7% vs 93.7%), and DC (84.3% vs 97.1%), but similar LRC. AI-iENE had significantly higher Concordance indices than radiologist-assessed iENE for OS (0.64 vs 0.55), RFS (0.67 vs 0.60), and DC (0.79 vs 0.68). In multivariable analysis, AI-iENE remained independently associated with OS (adjusted hazard ratio [aHR], 2.82; 95% CI, 1.21-6.57), RFS (aHR, 4.20; 95% CI, 1.93-9.11), and DC (aHR, 12.33; 95% CI, 4.15-36.67), adjusting for age, tumor category, node category, and number of lymph nodes.

CONCLUSIONS AND RELEVANCE This single-center cohort study found that an AI-driven pipeline can successfully automate lymph node segmentation and iENE classification from pretreatment CT scans in HPV-associated OPC. Predicted iENE was independently associated with worse oncologic outcomes. External validation is required to assess generalizability and the potential for implementation in institutions without specialized imaging expertise.

JAMA Otolaryngol Head Neck Surg. 2026;152(1):7-17. doi:10.1001/jamaoto.2025.3225
Published online September 30, 2025.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Laurent Letourneau-Guillon, MD, MSc, Centre Hospitalier de l'Université de Montréal, 1051 Rue Sanguinet, Montréal, QC H2X 3E4, Canada (laurent.letourneau-guillon.1@umontreal.ca).

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

Human papillomavirus (HPV)–positive status is a well-established positive prognostic factor in oropharyngeal squamous cell carcinoma (OPSCC).[1] The *American Joint Committee on Cancer (AJCC) Staging Manual, eighth edition*[2] classifies HPV-positive OPSCC as a distinct entity from HPV-negative OPSCC. Current efforts in the head and neck oncology community focus on deintensifying treatments for HPV-positive disease to reduce toxic effects while preserving oncologic outcomes.

Extranodal extension (ENE), traditionally a pathologic feature, is defined as the extension of cancer cells beyond the capsule of a metastatic lymph node. Based on pooled analysis from 2 landmark trials in head and neck oncology, ENE is considered among the most adverse prognostic factors for head and neck cancers (HNC).[3-5]

A major change in the eighth edition of the *AJCC Staging Manual*[2] was the incorporation of ENE status for HPV-negative OPSCC, with the upstaging of the node (N) category to N3b for any clinically overt ENE.[2] Conversely, ENE status was not included in the N classification of HPV-positive OPSCC. However, recent literature suggests that imaging-based ENE (iENE) are significant prognostic factors in both HPV-positive and HPV-negative OPSCC.[6-9] Notably, the International Collaboration on Oropharyngeal Cancer Network for Staging[10] has proposed the integration of iENE into the next edition of the *AJCC Staging Manual* for HPV-positive OPSCC, it is planned to be incorporated into the forthcoming ninth edition of the tumor, node, and metastases (TNM) classification of the Union for International Cancer Control[11] and the *AJCC Staging Manual* scheduled to be released in January 2026.

Hence, iENE detection will likely affect treatment decisions and management strategies in HPV-positive OPSCC in the near future. Key challenges with iENE include the lack of standardized criteria, reliance on radiological expertise that is typically restricted to tertiary centers, and interreader variability.[12,13] Artificial intelligence (AI) has the potential to standardize neuroimaging analysis, democratizing advanced capabilities to institutions without specialized neuroradiologists, while simultaneously providing expert centers with support to reduce misclassification errors.

This study aimed to develop a comprehensive automated pipeline for HPV-positive, clinical nodal positive (cN^+) OPSCCs that incorporates (1) an AI segmentation model that automatically identifies and segments the largest pathologic lymph nodes on pretreatment CT scans, followed by (2) a binary classification model that predicts ENE status. The secondary aims were to evaluate the association between AI-derived iENE classifications and oncologic outcomes, and to assess whether the model improves predictive value over established clinical predictors.

Key Points

Question Can an artificial intelligence (AI)–driven model predict imaging-based extranodal extension (iENE) and oncologic outcomes from pretreatment computed tomography scans of patients with human papillomavirus (HPV)–positive oropharyngeal squamous cell carcinoma (OPSCC)?

Findings In this single-center cohort study of 397 patients with HPV-positive cN^+ OPSCC, an automated pipeline integrating lymph node segmentation and iENE classification achieved an area under the receiver operating characteristic curve of 0.81. AI-predicted iENE was significantly associated with worse distant failure, recurrence-free survival, and overall survival, and outperformed expert radiologist assessment.

Meaning These findings suggest that automated iENE detection using AI models may offer a powerful prognostic tool to complement clinical decision-making in HPV-positive OPSCC and extend iENE interpretation capabilities to centers that lack specialized radiologists.

Methods

The institutional review board of the Centre Hospitalier de l'Université de Montréal (Quebec, Canada) approved this study and waived the need for informed consent due to the use of deidentified data. We followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis using regression or machine learning methods (TRIPOD+AI) guidelines.[14]

Study Design and Participants

This analysis was conducted using a retrospective database of patients with OPSCC at a tertiary oncology center in Montreal, Canada. Included patients were age 18 years or older, diagnosed with HPV-positive OPSCC and cN^+ disease, who underwent up-front radiation or chemoradiation therapy from January 2009 to January 2020, and were followed up until January 2024 (**Figure 1**). Exclusion criteria included HPV-negative tumors, HPV-status not available, and tumors treated by transoral robotic surgery (TORS).

Available demographic and clinical data, including age, sex, smoking status and number of pack-years, Eastern Cooperative Oncology Group (ECOG) Performance Status Scale status, as well as TNM staging (per the *AJCC Staging Manual, eighth edition)* were retrieved from the database and assessed by the treating clinical team. The representativeness of the dataset is supported by the inclusion of all consecutive cases treated during more than a decade at a high-volume tertiary reference center. Missing data were excluded from multivariable model development.

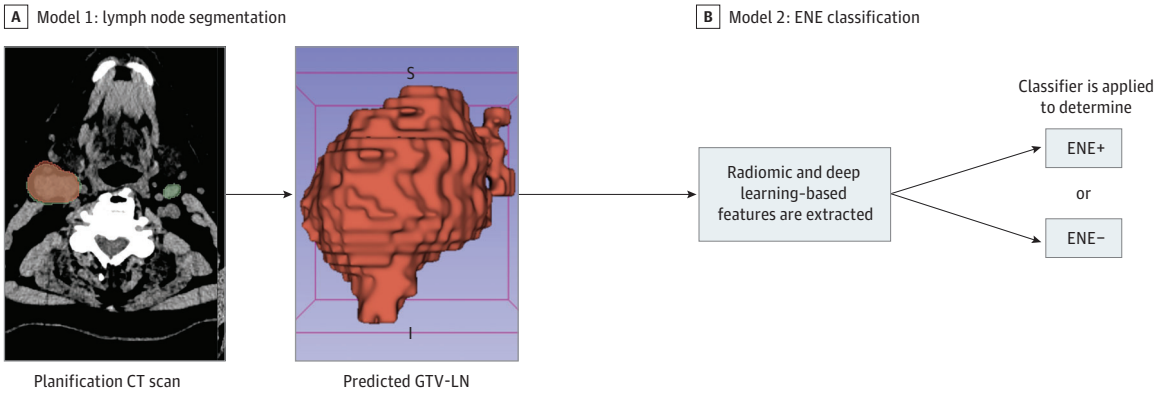
Exposure

Imaging and Segmentation Acquisition

Computed tomography (CT) scans that had been used to plan radiation therapy were retrieved from our center's picture archiving and communication system and deidentified. Lymph node gross tumor volume (GTV-LN) segmentations, performed by expert radiation oncologists during routine treatment planning and augmented by departmental peer review, were exported and used as the reference standard for nodal segmentations. These segmentations were verified by a fellowship-trained expert head and neck radiation oncologist (H.B.) and an otolaryngology–head and neck surgery resident (G.D.).

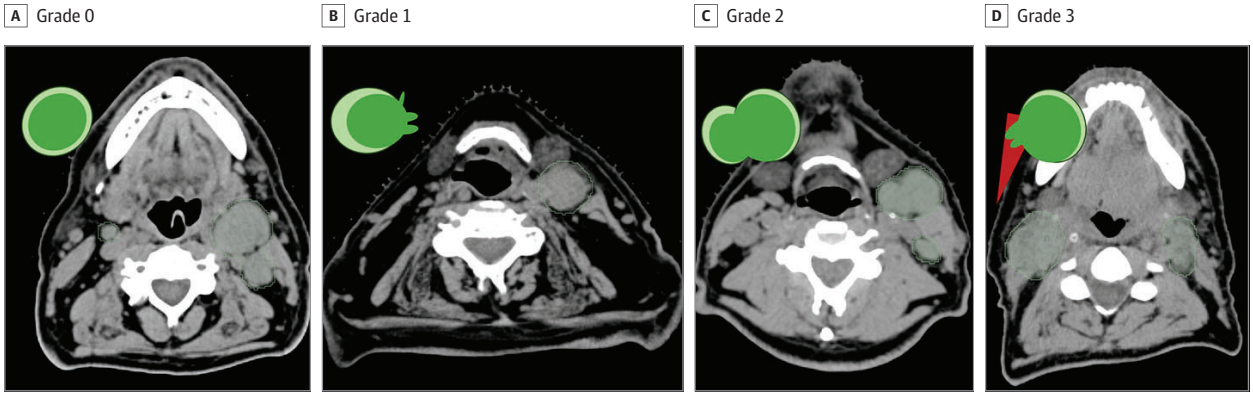
© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

Figure 1. Overview of Artificial Intelligence (AI) Pipeline for Imaging-Based Extranodal Extension (iENE) Prediction



Proposed 2-step pipeline for predicting iENE from radiotherapy planification using CT scans. Model 1, lymph node segmentation (left panel) assesses different segmentation architectures to identify and segment pathologic lymph nodes (lymph node gross tumor volume) from CT scans, generating a 3-dimensional reconstruction of the predicted lymph node volume. Model 2, ENE classification (right panel) extracts radiomic and deep learning-based features from the largest segmented lymph node and applies a classifier to determine the presence (ENE^+) or absence (ENE^-) of iENE. CT indicates computed tomography; ENE, extranodal extension; GTV-LN, lymph node gross tumor volume; $iENE^+$ indicates ENE positive, and $iENE^-$, ENE negative.

Figure 2. Four-Tier Grading System for Imaging-Based Extranodal Extension (iENE) per Consensus Recommendations



Grade 0, no evidence of extranodal extension (ENE); grade 1, clearly irregular or ill-defined nodal margins, characterized by loss of definition between the lymph node capsule and surrounding fat, or extension into perinodal fat; grade 2, invasion involving 2 or more inseparable adjoining lymph nodes forming a coalescent (matted) mass, with loss of internodal planes; grade 3, extension of tumor beyond the perinodal fat into adjacent structures, eg, muscle, skin, glands, or neurovascular structures. Representative computed tomography scans (with automated lymph node gross tumor volume segmentations) are shown for each iENE grade, along with schematic illustrations. Light green indicates the lymph node and capsule; dark green, nodal metastasis; and red, adjacent muscle.

Outcomes

Imaging-Detected Extranodal Extension

The primary outcome of interest was imaging-based extranodal extension (iENE), assessed by expert board-certified neuroradiologists with 11 and 3 years of experience (L.L.G. and K.N., respectively), on pretreatment radiation planning CT scans (rad-iENE). To improve reliability in downstream ratings, both raters evaluated 20 neck CT scans from a similar population not included in the current study, and these were subsequently reviewed in consensus to discuss ambiguous findings and corner cases. iENE was graded using a 4-tier classification system based on consensus recommendations for radiologic diagnosis of ENE in head and neck cancer (**Figure 2**).[15]

The first 60 patients in the cohort were independently evaluated by the aforementioned 2 raters to establish interobserver agreement for iENE assessment (eMethods 1 and eTable 1 in Supplement 1). The remaining patients were assessed by a single rater. Radiologists were blinded to outcomes. For AI model training, iENE gradings were subsequently dichotomized as iENE-negative (grade 0) vs iENE-positive (grades 1 to 3), as recently proposed.[10]

Oncologic Outcomes

The following end points were evaluated: locoregional control (LRC), distant control (DC), and recurrence-free survival (RFS), defined respectively as time to locoregional recurrence, distant recurrence, or any recurrence, with deaths censored. Overall survival (OS) was defined as time to death from any cause. Time to event was calculated from diagnosis.

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

Statistical Analysis

GTV-LN Segmentation and iENE Classification Models

For GTV-LN segmentation, we used the self-configuring nnU-Net model,[16] (available at github.com/MIC-DKFZ/nnUNet) with performance assessed using dice similarity coefficient (DSC) and intersection over union with bootstrapping to calculate 95% CIs. For iENE classification models (AI-iENE), features were extracted from the largest predicted lymph node volume using both radiomics and deep-learning approaches, followed by principal component analysis and least absolute shrinkage and selection operator (LASSO)–based selection and classification via extreme gradient boosting (XGBoost) or multilayer perceptron. Five-fold cross-validation with stratification for iENE was performed, and synthetic minority oversampling technique was applied to address class imbalance on the training set. In each fold, a separate test set was reserved for performance assessment and was not used for hyperparameter tuning. Additional details are provided in eMethods 2 in Supplement 1.

Association of Predicted iENE With Oncologic Outcomes

Statistical analyses were performed using R, version 4.4.2 (R Foundation for Statistical Computing). Survival end points at 3 years were estimated using the Kaplan-Meier method. Univariable Cox regression model was used to identify factors associated with oncologic outcomes. To compare predictive accuracy for oncologic outcomes of AI-iENE vs rad-iENE, difference of concordance index (C-index) with 95% CIs were calculated. Multivariable Cox regression analysis was performed to explore the independent association between iENE and these outcomes. Variables were entered in the multivariable model based on univariable logistic regression ($P < .10$), expert opinion and previously published literature. Backward selection method, with a probability of removing a variable of 5%, was used to produce the final model. Likelihood ratio tests were performed to compare models containing only clinical variables with those expanded to include AI-iENE. The proportional hazards assumption was assessed using Schoenfeld residuals and tested using the survival package. Calibration was assessed using calibration plots.

Interpretability

For the top-performing radiomics-based model, we generated a Shapley additive explanation plot to illustrate feature contributions. For the deep-learning feature extraction approach, we used saliency maps to highlight contributive regions for iENE classification. All data analyses were performed from March 2024 to April 2025, and used Python, version 3.10.18 (Python Software Foundation; packages: scikit-learn, 1.7.1; XGBoost, 3.0.4; imbalanced-learn, 0.13.0; and shap, 0.47.2).

Results

Of 493 patients with HPV-positive OPSCC, 397 patients (80.5%) had cN^+ disease and were included in this analysis. Their median (IQR) age at diagnosis was 62 (56-69) years; 80 were female (20.2%) and 317 were male (79.8%) individuals. Patient and tumor characteristics are summarized in **Table 1**; demographic information was collected by electronic health record review. Regarding N category, 311 patients (78.3%) had N1 disease, 67 (16.9%) had N2, and 19 (4.8%) had N3 per the *AJCC Staging Manual, eighth edition*.[2] Ninety-nine patients (24.9%) presented with at least 5 abnormal lymph nodes. Radiologic assessment by expert neuroradiologists identified 265 CT scans (66.8%) as rad-iENE–negative (grade 0) and 132 (33.2%) that were rad-iENE–positive (grades 1-3).

GTV-LN Automated Segmentations and iENE Predictions

The nnU-Net model achieved a mean DSC of 0.74 (95% CI, 0.72-0.77) for segmenting pathologic lymph nodes (detailed segmentation results are available in eTable 2 in Supplement 1). From the resulting GTV-LN segmentations, the largest lymph node in each scan was selected for downstream analysis. Various feature extraction and classification methods were tested (eTable 3 in Supplement 1). The method that maximized the area under the receiver operating characteristic curve (AUC) for differentiating iENE-positive from iENE-negative cases was the radiomic-based feature extraction with LASSO feature selection and XGBoost classification (AUC, 0.809; 95% CI, 0.764-0.850) (eFigure 1 in Supplement 1). Applying the threshold that maximized the Youden index (0.29) yielded a sensitivity of 74.6% (95% CI, 66.1%-81.9%) and specificity of 73.1% (95% CI, 67.4%-78.3%) for comparison of the AI model to radiologist performance on the entire dataset. Examples of concordant and discordant iENE classifications between expert raters and automated model are shown in eFigures 2 and 3 in Supplement 1.

AI-Predicted iENE vs Oncologic Outcomes

The mean (SD) and median (IQR) follow-up were 3.9 (1.9) and 3.7 (2.7-5.1) years, respectively. At 3 years, patients with AI-iENE–positive had worse OS compared with AI-iENE–negative (83.8% [95% CI, 78.0%-90.0%] vs 96.8% [94.5%-99.2%]), RFS (80.7% [95% CI, 74.7%-87.2%] vs 93.7% [95% CI, 90.5%-96.9%]), and DC (84.3% [95% CI, 78.7%-90.3%] vs 97.1% [95% CI, 94.9%-99.4%]) (all log-rank $P < .001$). However, 3-year LRC was comparable between groups (90.3% [95% CI, 85.7%-95.1%] vs 94.7% [95% CI, 91.8%-97.6%]; log-rank $P = .16$).

Univariable Cox analyses

Complete univariable Cox regression results for all variables tested across oncologic outcomes are summarized in **Table 2**. Probability (continuous) of AI-iENE and AI-iENE dichotomized using the Youden index were both associated with OS, with a hazard ratio (HR) of 4.18 (95% CI, 1.88-9.31) and 4.87 (95% CI, 2.30-10.29), respectively. Rad-iENE approached statistical significance for OS (HR, 1.83; 95% CI, 0.97-3.47). Both the probability of AI-iENE and dichotomized AI-iENE were associated with RFS (HR, 4.82; 95% CI, 2.40-9.70, and HR, 3.02; 95% CI, 1.68-5.40, respectively) and DC (HR, 12.98; 95% CI, 4.96-34.00 and HR, 5.20; 95% CI, 2.36-11.46, respectively). Rad-iENE was also associated with RFS (HR, 2.52; 95% CI, 1.45-4.37) and DC (HR, 4.59; 95% CI, 2.25-9.37). None were associated with LRC (Table 2).

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

Table 1. Baseline Characteristics of Study Cohort, by Artificial Intelligence–Predicted Imaging-Based Extranodal Extension (AI-iENE)

	No. (%)		
		AI-iENE prediction	
Variable	All patients	Negative	Positive
Patients, No.	397	230	167
Age at diagnosis, median (IQR), y	62 (56-69)	64 (57-70)	61 (56-67)
Female	80 (20.2)	51 (22.2)	29 (17.4)
Male	317 (79.8)	179 (77.8)	138 (82.6)
ECOG performance status			
0	69 (17.4)	42 (18.3)	27 (16.2)
1	305 (76.8)	177 (77.0)	128 (76.6)
2	7 (1.8)	1 (0.4)	6 (3.6)
3	0	0	0
Smoking pack-years, median (IQR)	8 (0-30)	6 (0-30)	10 (0-30)
Smoking status			
Never	141 (35.5)	90 (39.1)	51 (30.5)
Prior	73 (18.4)	36 (15.7)	37 (22.2)
Current	183 (46.1)	104 (45.2)	79 (47.3)
Tumor category[a]			
1	102 (25.7)	52 (22.6)	50 (29.9)
2	151 (38.0)	83 (36.1)	68 (40.7)
3	95 (23.9)	60 (26.1)	35 (21.0)
4	49 (12.3)	35 (15.2)	14 (8.4)
Node category[a]			
1	311 (78.3)	189 (82.2)	122 (73.1)
2	67 (16.9)	38 (16.5)	29 (17.4)
3	19 (4.8)	3 (1.3)	16 (9.6)
Abnormal lymph node, No.			
1-4	298 (75.1)	197 (85.7)	101 (60.5)
≥5	99 (24.9)	33 (14.3)	66 (39.5)
Abnormal retropharyngeal lymph node >8 mm	44 (11.1)	19 (8.3)	25 (15.0)
iENE (radiologist)			
0	271 (68.3)	195 (84.8)	70 (41.9)
1	26 (6.5)	12 (5.2)	15 (9.0)
2	75 (18.9)	22 (9.6)	58 (34.7)
3	25 (6.3)	0 (0.0)	24 (14.4)
Concurrent systemic therapy	333 (83.9)	181 (78.7)	152 (91.0)

Abbreviation: ECOG, Eastern Cooperative Oncology Group.

[a] Per the *American Joint Committee on Cancer Staging Manual, eighth edition*.

Oncologic Outcomes of AI-iENE vs Rad-iENE

Three-year event rates and Kaplan-Meier curves comparing AI-iENE and rad-iENE across all oncologic outcomes are presented in eTable 4 in Supplement 1, **Figure 3**, and eFigure 4 in Supplement 1, respectively.

C-indices were compared between AI- and rad-iENE to assess which approach more accurately predicted oncologic outcomes. C-indices were significantly higher for of AI-iENE compared to rad-iENE for OS (0.64 vs 0.55; difference, 0.09; 95% CI, 0 to 0.17; $P = .04$), RFS (0.67 vs 0.60; difference, 0.08; 95% CI, 0 to 0.15; $P = .04$), and DC (0.79 vs 0.68; difference, 0.11; 95% CI, 0.03-0.19; $P < .01$). For LRC, the C-index was 0.58 for AI-iENE vs 0.53 for rad-iENE (difference, 0.05; 95% CI, −0.05 to 0.15; $P = .29$).

Multivariable Cox Analyses

Multivariable regression results, with corresponding C-indices, are presented in Table 2. After backward stepwise selection, the probability of AI-iENE remained independently associated with OS (adjusted HR, [aHR] 2.82; 95% CI, 1.21-6.57) and RFS (aHR, 4.20; 95% CI, 1.93-9.12). For LRC, AI-iENE was nonsignificant (aHR, 1.40; 95% CI, 0.54-3.59). Assessment of the proportional hazards assumption using Schoenfeld residuals revealed no violations for any variables (eTable 5 in Supplement 1).

Models With Only Clinical Variables vs Expanded by AI-iENE

Likelihood ratio tests were performed to compare models with only clinical variables to models expanded by AI-iENE. Adding iENE significantly improved the model fit for OS (log likelihood, −199.3 vs −195.5; $\chi^2 = 7.606$; $P = .006$), RFS (−267.7 vs −259.7; $\chi^2 = 14.864$; $P < .001$), and DC (−174.3 vs −160.9; $\chi^2 = 26.749$; $P < .001$). No significant improvement was observed for LRC (−163.9 vs −163.6; $\chi^2 = 0.472$; $P = .49$).

Results of model calibration are presented in eFigure 5 in Supplement 1 and those of proportional hazards analysis in

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

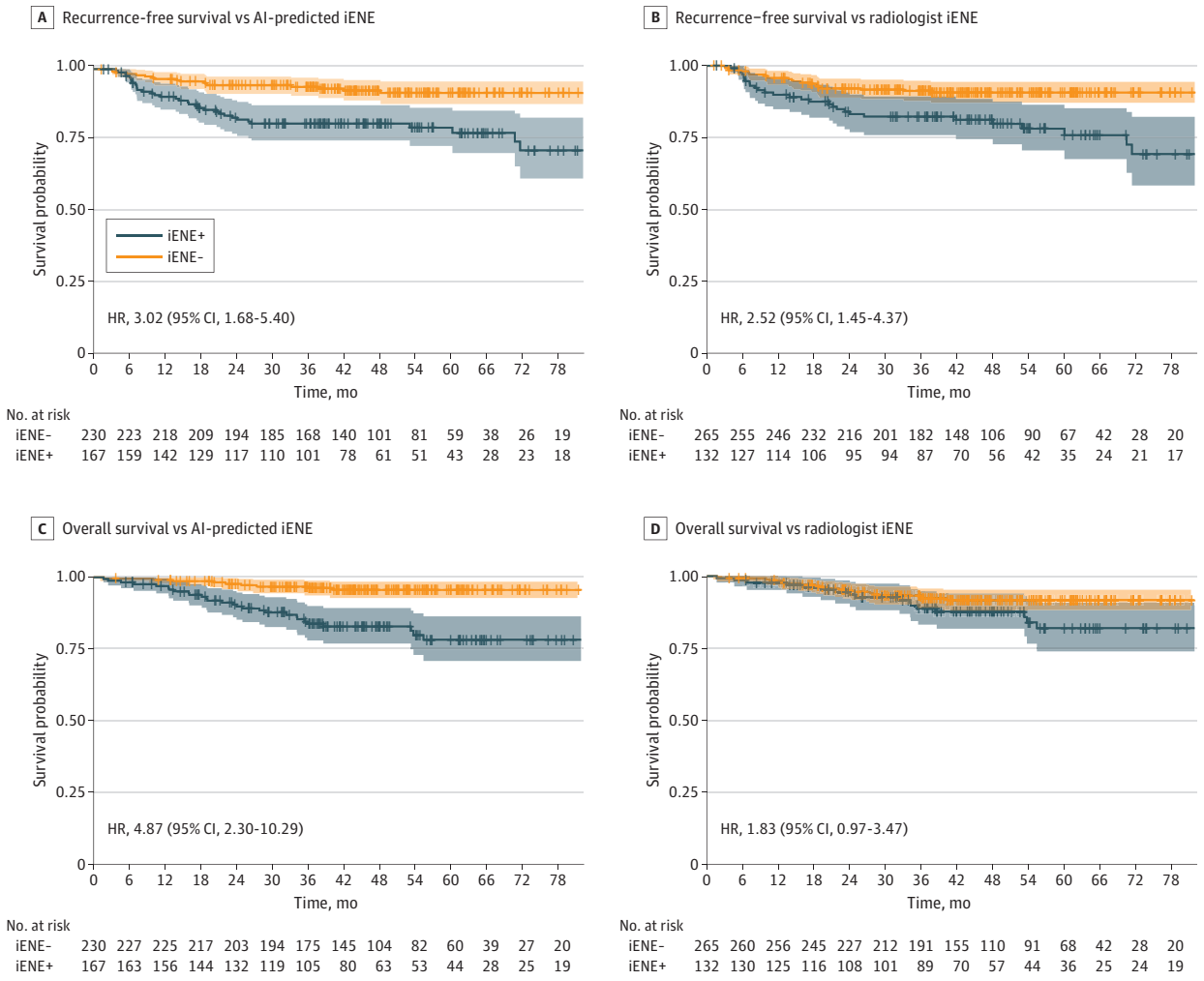
Table 2. Univariate (UV) and Multivariate (MV) Regression Results for Oncologic Outcomes

Variable	Hazard ratio (95% CI)							
	Locoregional control		Distant control		Recurrence-free survival		Overall survival	
	UV	MV	UV	MV	UV	MV	UV	MV
Age at diagnosis	1.04 (1.00-1.08)	NA	1.04 (1.00-1.08)	1.07 (1.03-1.12)	1.04 (1.00-1.07)	1.05 (1.02-1.09)	1.07 (1.03-1.11)	1.08 (1.04-1.12)
Female sex	1.74 (0.80-3.81)	NA	0.64 (0.25-1.64)	NA	0.97 (0.48-1.94)	NA	1.61 (0.80-3.25)	NA
ECOG performance status (0 vs ≥1)	6.27 (0.85-46.11)	NA	3.44 (0.82-14.39)	NA	3.42 (1.06-11.00)	NA	2.41 (0.74-7.87)	NA
Smoking pack-years	1.02 (1.00-1.03)	1.02 (1.00-1.04)	1.01 (1.00-1.02)	NA	1.01 (1.00-1.02)	NA	1.02 (1.00-1.03)	NA
Smoking status								
Never	1 [Reference]	NA	1 [Reference]	NA	1 [Reference]	NA	1 [Reference]	NA
Prior	2.28 (0.83-6.30)	NA	1.71 (0.71-4.12)	NA	1.90 (0.87-4.17)	NA	3.55 (1.40-9.03)	NA
Current	1.73 (0.70-4.24)	NA	1.15 (0.53-2.50)	NA	1.68 (0.86-3.27)	NA	2.31 (0.97-5.50)	NA
Tumor category								
3-4 vs 1-2	2.39 (1.16-4.91)	2.65 (1.27-5.51)	1.98 (1.02-3.85)	2.99 (1.44-6.21)	2.12 (1.22-3.68)	2.55 (1.42-4.59)	1.19 (0.62-2.27)	NA
Node category								
1	1 [Reference]	NA	1 [Reference]	NA	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
2	1.98 (0.86-4.51)	NA	2.83 (1.33-5.99)	1.64 (0.74-3.62)	2.34 (1.26-4.35)	1.39 (0.72-2.67)	1.52 (0.68-3.38)	1.15 (0.51-2.59)
3	3.35 (0.99-11.35)	NA	7.94 (3.14-20.06)	5.49 (2.05-14.69)	4.59 (1.91-11.05)	3.36 (1.35-8.41)	5.38 (2.19-13.21)	3.40 (1.30-8.93)
Total abnormal lymph nodes								
1-4	1 [Reference]	NA	1 [Reference]	NA	1 [Reference]	1 [Reference]	1 [Reference]	NA
≥5	3.41 (1.67-7.00)	3.37 (1.57-7.26)	7.13 (3.54-14.37)	5.95 (2.76-12.79)	4.83 (2.77-8.44)	3.93 (2.18-7.08)	3.24 (1.71-6.14)	2.74 (1.40-5.36)
Abnormal retropharyngeal lymph nodes	2.78 (1.19-6.49)	NA	4.55 (2.22-9.33)	NA	3.41 (1.81-6.42)	NA	2.45 (1.12-5.36)	NA
iENE								
Probability of iENE	1.98 (0.82-4.82)	1.40 (0.54-3.59)	12.98 (4.96-34.00)	12.33 (4.15-36.67)	4.82 (2.40-9.70)	4.20 (1.93-9.11)	4.18 (1.88-9.31)	2.82 (1.21-6.57)
Pred iENE, 0 vs 1-3	1.67 (0.82-3.43)	NA	5.20 (2.36-11.46)	NA	3.02 (1.68-5.40)	NA	4.87 (2.30-10.29)	NA
Radiologist iENE, 0 vs 1-3	1.36 (0.66-2.83)	NA	4.59 (2.25-9.37)	NA	2.52 (1.45-4.37)	NA	1.83 (0.97-3.47)	NA
Concordance index, mean (SD)	NA	0.722 (0.046)	NA	0.868 (0.27)	NA	0.769 (0.040)	NA	0.771 (0.042)

Abbreviations: ECOG, Eastern Cooperative Oncology Group; iENE, imaging-based extranodal extension; NA, not applicable.

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

Figure 3. Kaplan-Meier Curves for Oncologic Outcomes Stratified by Imaging-Based Extranodal Extension (iENE), Grades 0 vs 1 to 3



Recurrence-free and overall survival as determined by AI vs radiologist interpretation. AI indicates artificial intelligence; HR, hazard ratio; iENE$^+$, ENE positive; and iENE$^-$, ENE negative.

eTable 6 in Supplement 1. To improve interpretability of model outputs, a Shapley additive explanations plot is presented in eFigure 6 in Supplement 1 for the top-performing radiomic model (LASSO/XGBoost) and a comparison with a saliency map generated from our deep-learning feature extractor is presented in eFigure 7 in Supplement 1. A post hoc exploratory analysis reclassifying nodal categories according to the proposed ninth edition of the *AJCC Staging Manual* and framework are presented in eTable 6 in Supplement 1.[10,11]

Discussion

In this study, we developed a 2-step AI pipeline for automated pathologic lymph node segmentation and iENE classification from pretreatment CT scans in HPV-positive OPSCC. First, our 3-dimensional nnU-Net model achieved clinically satisfactory accuracy for metastatic lymph node segmentation. For iENE classification, radiomic feature extraction outperformed deep-feature extraction, achieving an AUC of 0.81 for detecting iENE. From a clinical perspective, patients predicted to have iENE by the model had significantly worse DC, RFS, and OS. Notably, AI-based iENE detection demonstrated prognostic significance exceeding expert radiologist assessment. In multivariable analysis, AI-predicted iENE remained an independent predictor of DC, RFS, and OS, even after adjusting for T stage, N burden, and other prognostic variables.

Although clinical ENE influences the N category in HPV-negative OPSCC, the *AJCC Staging Manual, eighth edition* does not account for ENE for HPV-positive tumors. However, emerging evidence suggests that iENE is prognostically significant even in HPV-mediated disease, as reinforced by our findings.[8,9,17] A multicenter cohort showed that radiographic detection of ENE vs pENE improved with expert neuroradiologist assessment compared to generalists, although iENE had limited prognostic value in a mixed-expertise setting.[13] Even though our study compares AI results to radiologists' interpretations rather than pathologic ENE (pENE), these findings

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

highlight the value of radiologic expertise and the potential of our model to assist with standardization in settings that lack such expertise. A 2021 meta-analysis by Benchetrit et al[6] demonstrated that both pathologic pENE and iENE were associated with worse outcomes. Notably, HRs were higher for iENE than for pENE, suggesting that macroscopically visible ENE on imaging may indicate particularly aggressive disease. The strongest association was observed between iENE and distant metastasis risk, just as in the current study. While ENE is more commonly cited as an indication for adjuvant systemic therapy in the postoperative setting, this finding aligns with evidence from the definitive chemoradiotherapy setting supporting the use of chemotherapy to reduce distant failure in patients with high-risk features, including bulky nodal disease and potential ENE.[3,18]

Researchers have advocated for incorporating iENE into future staging frameworks for HPV-positive OPSCC.[1,8,10] iENE is incorporated in the forthcoming ninth edition of TNM classification by the Union for International Cancer Control and the AJCC, scheduled for January 2026, by upstaging the N category by 1 stratum when iENE is present.[11] Hence, iENE identification on pretreatment imaging will likely have a direct impact on clinical management, particularly in recognizing the inherent risk of developing distant metastasis. For example, in HPV-positive OPSCC, patients with iENE may be excluded from de-escalation protocols, whereas those predicted to be iENE-negative may be considered to be better candidates for deintensified treatment protocols. Parenthetically, patients with iENE may benefit from newer treatment approaches and/or trials to mitigate risk of distant metastasis given that the current standard of cisplatin appears insufficient to address this.[10] Future prospective studies are warranted to validate this approach and automated image analysis could improve patient selection in such clinical trials. Additionally, pretreatment detection of ENE can guide patient selection for TORS. Patients with iENE may be steered toward primary chemoradiotherapy rather than TORS given that pENE would necessitate adjuvant treatment and the risk of high morbidity associated with multimodal therapy.[19]

By providing automated lymph node segmentation and iENE classification, our tool offers accessible prognostic data in an interpretable output. Our AI-based approach has the potential to address key challenges in iENE assessment, including interreader variability, reliance on expert radiologists, and standardization. While our interrater evaluation was limited to a subset of studies by 2 expert radiologists, prior studies report a wide range of interreader agreement ranging from low to substantial.[20-23] Using a high-certainty threshold paired with clearly defined nomenclature can improve interrater agreement.[21] However, these strategies have a greater chance for uptake in specialized centers. In other settings, an AI system trained on expert-defined criteria may represent an opportunity to standardize and democratize iENE detection across institutions. This may improve staging consistency for referring clinicians who practice in community centers that lack specialized radiologic expertise, thus improving health equity for patients without access to specialized imaging interpretation.

One of the most striking findings is that although the AI model was trained using radiologist-assessed iENE as the reference standard, its predictions demonstrated stronger associations with oncologic outcomes than did those of the original radiologist assessments. This may be explained by the AI model's ability to detect consistent, high-dimensional imaging patterns across the entire training dataset, even when those patterns only partially overlapped with what human experts subjectively identify as iENE. In contrast, radiologists' assessments may vary due to interobserver variability, shifts in judgment from day to day, and reliance on limited visual cues. Although trained on human labels, the model may have generalized beyond them, learning imaging signatures of aggressive nodal disease that more closely align with prognosis.

Our study builds on the growing literature on AI-driven ENE detection in HNC. Kann et al[24] developed a 3-dimensional convolutional neural network to identify pENE on CT scans, achieving AUCs of 0.91 in their initial study, 0.84-0.90 in external cohorts,[25] and 0.86 in a post hoc analysis of the ECOG-ACRIN E3311 trial,[26] which included surgically treated HPV-associated OPSCC. Key distinctions exist between their approach and ours. While they used pENE as the reference standard, we focused on iENE, which may be more reflective of clinical practice, particularly in nonsurgical patients for whom pENE is unavailable. While pENE represents a more definitive reference standard, some studies[27-30] suggest that pENE does not significantly impact oncologic outcomes in HPV-positive OPSCC, at least in populations where many surgically managed patients also received adjuvant radiotherapy that could mitigate its influence. Conversely, there is growing literature supporting the association between iENE and outcomes, supporting our choice of prediction target.[6,8,11] It should be noted that the ECOG 3311 trial[26] stratified patients by the extent of pENE, with a greater than 1 mm as higher risk. Although we could not directly assess pENE extent (≤ 1 mm vs >1 mm) on imaging with the current study design, it is likely that radiographically detected ENE preferentially captures these larger-extent higher-risk cases. Kann et al[24-26] did not report the association between predicted pENE and oncologic outcomes. Additionally, we evaluated 1 scan per patient, whereas Kann et al assessed individual lymph nodes, including nonmetastatic nodes and patients with cN0 disease. This distinction affects how false negatives and diagnostic performance are interpreted, limiting AUC comparisons. Interestingly, Kann et al[25] found that radiologists' accuracy improved with AI support, suggesting potential value in integrating these models as an assistive device to augment clinical interpretations. Lastly, they relied on manual segmentations, whereas our pipeline incorporated an automated segmentation model. A key objective of our study was to minimize manual labor to address a practical hurdle in radiomics—the time, resources and expertise required for manual contouring. Improving segmentation accuracy may further enhance overall model performance.

Pipelines such as ours, which integrate automated volumetric segmentation and feature extraction, represent an important step toward deployment of prognostic models in the clinical setting. While previous studies have used segmenta-

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.
Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

tion for outcome prediction, ours is, to our knowledge, the first to incorporate ENE as a human-interpretable intermediate imaging biomarker for oncologic outcome prediction. The HECKTOR 2022[31] challenge similarly addressed segmentation and outcome prediction in OPSCC (N = 883; HPV-positive or negative). For lymph node segmentation, the top model achieved a mean DSC of 77.6%[32] comparable to our study's 74.4% for GTV-LN. For RFS prediction, their best radiomics model reached a C-index of 0.68,[33] whereas ours achieved 0.77. Although these results were obtained from different cohorts, the near 10% higher C-index in our study may reflect differences in methods. Our study focused on iENE, a specific biomarker of nodal aggressiveness which has a pathophysiological association with outcome, aligning with clinical knowledge, rather than broad feature extraction with the potential of spurious correlation. Unlike HECKTOR,[31] we did not include primary tumor features. Prior studies, such as by Bogowicz et al,[34] have shown improved LRC prediction when combining primary tumor and nodal radiomics features. Future work should explore if integrating primary tumor features into our models improves outcome prediction for HPV-positive OPSCC.

Limitations

This study has several limitations. First, it is a retrospective, single-center study conducted in a tertiary care center with expert neuroradiologists, which may limit generalizability. External validation with multi-institutional datasets are required to move beyond in silico research, a common limitation in the field.[35] Following successful independent testing, the next step would be to evaluate the added value of a clinical decision support tool, including decision-curve analysis, and ultimately, a clustered randomized clinical trial allocating physicians to make management decisions with or without model support. This requires identifying the best-suited oncologic outcome and its probability threshold that considers the risks of false-positive and negative predictions. Notably, careful consideration is required to balance the potential complications of treatment intensification against the risks of recurrence and poor oncological outcomes associated with inappropriate de-escalation. Given that there is currently a changing landscape in OPSCC-related management strategies it would be premature to attempt such an analysis before performing external and prospective evaluation. As noted, neither the identification of iENE (whether by AI or as incorporated in the upcoming nineth edition of TNM classification[11]) should be used to change treatment approaches at this time, and we must continue to rely on the results of robust clinical trials.[1,10]

Second, the segmentation model was trained using GTV-LN contours delineated for radiation planning, which may not precisely reflect true nodal boundaries and are subject to interclinician variability. Our qualitative review revealed heterogeneous boundary definition sometimes extending beyond strict anatomic boundaries to provide security margins during treatment. Third, our model focused on the largest pathologic node for feature extraction—an oversimplification given that ENE can occur in smaller nodes and may have cumulative effects when present in multiple nodes. However, this approach reflects the biology of nodal metastases because the largest node often represents the dominant tumor burden and carries the highest risk of ENE,[24] while also ensuring consistency and minimizing noise from smaller, nonrepresentative nodes. It also simplifies the image analysis pipeline and the model interpretability. However, our analysis did consider the total number of pathologic nodes, which was strongly associated with all survival end points. Fourth, the ENE classification reference standard that we used in this study was based on radiologist assessment rather than surgical pathologic findings, potentially producing false negatives or false positives. Despite this, our outcome analyses demonstrate that AI-iENE is prognostically relevant, aligning with prior findings that iENE appears more predictive of outcomes than pENE in HPV-positive OPSCC. It also aligns with clinical settings in which pENE is often not available when patients with HPV-positive OPSCC are treated nonsurgically. Finally, the cohort had a sex imbalance and lacked sociodemographic data, limiting assessment of disparities in model performance. Future work should also assess robustness across scan quality and manufacturers.

Conclusions

This single-center cohort study demonstrated that an AI-driven pipeline integrating automated lymph node segmentation and radiomic feature extraction can be used for iENE classification in HPV-positive OPSCC. Furthermore, AI-iENE holds prognostic significance in this population, despite ENE currently not being considered for the staging of these tumors. Future directions include refining segmentation models, validating the approach on external multicenter datasets, and expanding its application to other HNC subsites.

ARTICLE INFORMATION

Accepted for Publication: August 7, 2025.

Published Online: September 30, 2025. doi:10.1001/jamaoto.2025.3225

Author Affiliations: Division of Otolaryngology–Head and Neck Surgery, Centre Hospitalier de l'Université de Montréal (CHUM), Université de Montréal, Quebec, Canada (Dayan, Christopoulos, Ayad, Bissada, Tabet, Guertin); Research Center of CHUM, Université de Montréal, Quebec, Canada (Dayan, Hénique, Bahig, Brodeur, Christopoulos, O'Sullivan, Desilets, Kadoury, Letourneau-Guillon); Department of Computer and Software Engineering, Polytechnique Montréal, Quebec, Canada (Hénique, Brodeur, Kadoury); Division of Radiation Oncology, Department of Radiology, Radiation Oncology and Nuclear Medicine, CHUM, Université de Montréal, Quebec, Canada (Bahig, Filion, Nguyen-Tan, O'Sullivan); Division of Radiology, Department of Radiology, Radiation Oncology and Nuclear Medicine, CHUM, Université de Montréal, Quebec, Canada (Nelson, Letourneau-Guillon); Division of Medical Oncology, Department of Medicine, CHUM, Université de Montréal, Quebec, Canada (Desilets).

Author Contributions: Drs Dayan and Létourneau-Guillon had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Dayan, Hénique, Bahig, Christopoulos, Filion, Bissada, Tabet, Guertin, Kadoury, Letourneau-Guillon.
Acquisition, analysis, or interpretation of data: Dayan, Hénique, Nelson, Brodeur, Nguyen-Tan, O'Sullivan, Ayad, Desilets, Letourneau-Guillon.
Drafting of the manuscript: Dayan, Kadoury, Letourneau-Guillon.
Critical review of the manuscript for important

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

intellectual content: Dayan, Hénique, Bahig, Nelson, Brodeur, Christopoulos, Filion, Nguyen-Tan, O'Sullivan, Ayad, Bissada, Tabet, Guertin, Desilets, Letourneau-Guillon.
Statistical analysis: Dayan, Hénique, Brodeur, Letourneau-Guillon.
Obtained funding: Bahig, Kadoury.
Administrative, technical, or material support: Brodeur, Nguyen-Tan, Kadoury.
Supervision: Bahig, Christopoulos, Filion, Ayad, Bissada, Tabet, Guertin, Desilets, Kadoury, Letourneau-Guillon.

Conflict of Interest Disclosures: Dr Bahig reported serving on the advisory board of AstraZeneca; speaking fees from Bristol Myers Squibb, AstraZeneca, and Pfizer; and a grant from Varian, outside the submitted work. Dr Brodeur reported a research scholarship from Polytechnique Montreal during the conduct of the study. Dr Desilets reported institutional grants from Exelixis, Regeneron (in-kind drug support)and Foundation Saputo; advisory board and/or consulting fees from Boehringer Ingelheim, EMD Serono, Merck, and Johnson & Johnson–Janssen; speaking and/or personal fees from Boehringer Ingelheim, AstraZeneca, AbbVie Exelixis, Onc Live, Regeneron, Adlai Nortye, Bristol Myers Squibb, EMD Serono, Johnson & Johnson–Janssen, Merck, and Novartis, all outside the submitted work. Dr Letourneau-Guillon reported institutional research funding/grants from Fonds de Recherche du Québec en Santé and Fondation de l'Association des Radiologistes du Québec, the Foundation of the Radiological Society of North America (No. RSD2122), and Quebec Bio-Imaging Network; internal funding from the professoral du Département de radiologie, radio-oncologie et médecine nucléaire de l'Université de Montréal/Bayer, Radiology departement Centre Hospitalier de l'Université de Montréal (CHUM) and CHUM Research Center start-up funds; and educational honorarium from the Association des Neurologues du Quebec. No other disclosures were reported.

Funding/Support: The study was supported by the Canadian Institutes of Health Research, with salary support from the Fonds de Recherche Quebec–Sante to Dr Bahig (No. 180404; doi:10.69777/349871) and from the Fonds de Recherche Quebec–Sante/Fondation de l'Association des Radiologistes du Québec to Dr Letourneau-Guillon (doi:10.69777/311203).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Meeting Presentation: This paper was presented at the ASTRO (American Society for Radiation Oncology) 2025 Annual Meeting; September 30, 2025; San Francisco, California.

Data Sharing Statement: See Supplement 2.

REFERENCES

1. O'Sullivan B, Huang SH, Su J, et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal Cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol*. 2016;17(4):440-451. doi:10.1016/S1470-2045(15)00560-4

2. Amin MB, Edge SB, Greene FL, et al, eds. *AJCC Cancer Staging Manual*. 8th ed. Springer; 2018.

3. Bernier J, Cooper JS, Pajak TF, et al. Defining risk levels in locally advanced head and neck cancers: a comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (# 9501). *Head Neck*. 2005;27(10):843-850. doi:10.1002/hed.20279

4. Bernier J, Domenge C, Ozsahin M, et al; European Organization for Research and Treatment of Cancer Trial 22931. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med*. 2004;350(19):1945-1952. doi:10.1056/NEJMoa032641

5. Cooper JS, Pajak TF, Forastiere AA, et al; Radiation Therapy Oncology Group 9501/Intergroup. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2004;350(19):1937-1944. doi:10.1056/NEJMoa032646

6. Benchetrit L, Torabi SJ, Givi B, Haughey B, Judson BL. Prognostic significance of extranodal extension in HPV-mediated oropharyngeal carcinoma: a systematic review and meta-analysis. *Otolaryngol Head Neck Surg*. 2021;164(4):720-732. doi:10.1177/0194599820951176

7. Chin O, Yu E, O'Sullivan B, et al. Prognostic importance of radiologic extranodal extension in nasopharyngeal carcinoma treated in a Canadian cohort. *Radiother Oncol*. 2021;165:94-102. doi:10.1016/j.radonc.2021.10.018

8. Huang SH, O'Sullivan B, Su J, et al. Prognostic importance of radiologic extranodal extension in HPV-positive oropharyngeal carcinoma and its potential role in refining TNM-8 cN-classification. *Radiother Oncol*. 2020;144:13-22. doi:10.1016/j.radonc.2019.10.011

9. Pilar A, Yu E, Su J, et al. Prognostic value of clinical and radiologic extranodal extension and their role in the 8th edition TNM cN classification for HPV-negative oropharyngeal carcinoma. *Oral Oncol*. 2021;114:105167. doi:10.1016/j.oraloncology.2020.105167

10. Huang SH, Su J, Koyfman SA, et al. A proposal for HPV-associated oropharyngeal carcinoma in the ninth edition clinical TNM classification. *JAMA Otolaryngol Head Neck Surg*. 2025;151(7):655-664. doi:10.1001/jamaoto.2025.0848

11. Brierley J, Giuliani M, O'Sullivan B, Rous B, Van Eycken E. *TNM Classification of Malignant Tumours*. 9th ed. 2025.

12. Hoebers F, Yu E, O'Sullivan B, et al. Augmenting inter-rater concordance of radiologic extranodal extension in HPV-positive oropharyngeal carcinoma: a multicenter study. *Head Neck*. 2022;44(11):2361-2369. doi:10.1002/hed.27130

13. Mehanna H, Abou-Foul AK, Henson C, et al. Accuracy and prognosis of extranodal extension on radiologic imaging in human papillomavirus-mediated oropharyngeal cancer: a Head and Neck Cancer International Group (HNCIG) real-world study. *Int J Radiat Oncol Biol Phys*. 2025:S0360-3016(25)00370-0. doi:10.1016/j.ijrobp.2025.04.003

14. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378

15. Henson C, Abou-Foul AK, Yu E, et al. Criteria for the diagnosis of extranodal extension detected on radiological imaging in head and neck cancer: Head and Neck Cancer International Group consensus recommendations. *Lancet Oncol*. 2024;25(7):e297-e307. doi:10.1016/S1470-2045(24)00066-4

16. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203-211. doi:10.1038/s41592-020-01008-z

17. Billfalk-Kelly A, Yu E, Su J, et al. Radiologic extranodal extension portends worse outcome in cN+ TNM-8 stage I human papillomavirus-mediated oropharyngeal cancer. *Int J Radiat Oncol Biol Phys*. 2019;104(5):1017-1027. doi:10.1016/j.ijrobp.2019.03.047

18. Lacas B, Carmel A, Landais C, et al; MACH-NC Collaborative Group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 107 randomized trials and 19,805 patients, on behalf of MACH-NC Group. *Radiother Oncol*. 2021;156:281-293. doi:10.1016/j.radonc.2021.01.013

19. Achim V, Bolognone RK, Palmer AD, et al. Long-term functional and quality-of-life outcomes after transoral robotic surgery in patients with oropharyngeal cancer. *JAMA Otolaryngol Head Neck Surg*. 2018;144(1):18-27.

20. Carlton JA, Maxwell AW, Bauer LB, et al. Computed tomography detection of extracapsular spread of squamous cell carcinoma of the head and neck in metastatic cervical lymph nodes. *Neuroradiol J*. 2017;30(3):222-229. doi:10.1177/1971400917694048

21. Chin O, Alshafai L, O'Sullivan B, et al. Inter-rater concordance and operating definitions of radiologic nodal feature assessment in human papillomavirus-positive oropharyngeal carcinoma. *Oral Oncol*. 2022;125:105716. doi:10.1016/j.oraloncology.2022.105716

22. Sahin O, Kamel S, Wahid KA, et al. International multi-specialty expert physician preoperative identification of extranodal extension n oropharyngeal cancer patients using computed tomography: prospective blinded human inter-observer performance evaluation. *Cancer*. 2025:e35815. doi:10.1002/cncr.35815

23. Tran NA, Palotai M, Hanna GJ, et al. Diagnostic performance of computed tomography features in detecting oropharyngeal squamous cell carcinoma extranodal extension. *Eur Radiol*. 2023;33(5):3693-3703. doi:10.1007/s00330-023-09407-4

24. Kann BH, Aneja S, Loganadane GV, et al. Pretreatment identification of head and neck cancer nodal metastasis and extranodal extension using deep learning neural networks. *Sci Rep*. 2018;8(1):14036. doi:10.1038/s41598-018-32441-y

25. Kann BH, Hicks DF, Payabvash S, et al. Multi-institutional validation of deep learning for pretreatment identification of extranodal extension in head and neck squamous cell carcinoma. *J Clin Oncol*. 2020;38(12):1304-1311. doi:10.1200/JCO.19.02031

26. Kann BH, Likitlersuang J, Bontempi D, et al. Screening for extranodal extension in HPV-associated oropharyngeal carcinoma:

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026

evaluation of a CT-based deep learning algorithm in patient data from a multicentre, randomised de-escalation trial. *Lancet Digit Health*. 2023;5(6):e360-e369. doi:10.1016/S2589-7500(23)00046-8

27. Kharytaniuk N, Molony P, Boyle S, et al. Association of extracapsular spread with survival according to human papillomavirus status in oropharynx squamous cell carcinoma and carcinoma of unknown primary site. *JAMA Otolaryngol Head Neck Surg*. 2016;142(7):683-690. doi:10.1001/jamaoto.2016.0882

28. Maxwell JH, Ferris RL, Gooding W, et al. Extracapsular spread in head and neck carcinoma: impact of site and human papillomavirus status. *Cancer*. 2013;119(18):3302-3308. doi:10.1002/cncr.28169

29. Sinha P, Kallogjeri D, Gay H, et al. High metastatic node number, not extracapsular spread or N-classification is a node-related prognosticator in transorally-resected, neck-dissected p16-positive oropharynx cancer. *Oral Oncol*. 2015;51(5):514-520. doi:10.1016/j.oraloncology.2015.02.098

30. Sinha P, Lewis JS Jr, Piccirillo JF, Kallogjeri D, Haughey BH. Extracapsular spread and adjuvant therapy in human papillomavirus-related, p16-positive oropharyngeal carcinoma. *Cancer*. 2012;118(14):3519-3530. doi:10.1002/cncr.26671

31. Andrearczyk V, Oreiller V, Abobakr M, et al. Overview of the HECKTOR challenge at MICCAI 2022: automatic head and neck tumor segmentation and outcome prediction in PET/CT. *Head Neck Tumor Chall (2022)*. 2023;13626(13626):1-30. doi:10.1007/978-3-031-27420-6_1

32. Sun X, An C, Wang L. A coarse-to-fine ensembling framework for head and neck tumor and lymph segmentation in CT and PET Images. In: Andrearczyk V, Oreiller V, Hatt M, Depeursinge A, eds. *Head and Neck Tumor Segmentation and Outcome Prediction*. HECKTOR 2022. Lecture Notes in Computer Science. 2022:13626. doi:10.1007/978-3-031-27420-6_3

33. Rebaud L, Escobar T, Khalid F, Girum KB, Buvat I. Simplicity is all you need: out-of-the-box nnUNet followed by binary-weighted radiomic model for segmentation and outcome prediction in head and neck PET/CT. In: Andrearczyk V, Oreiller V, Hatt M, Depeursinge A, eds. *Head and Neck Tumor Segmentation and Outcome Prediction*. HECKTOR 2022. Lecture Notes in Computer Science. 2022:13626. doi:10.1007/978-3-031-27420-6_13

34. Bogowicz M, Tanadini-Lang S, Guckenberger M, Riesterer O. Combined CT radiomics of primary tumor and metastatic lymph nodes improves prediction of loco-regional control in head and neck cancer. *Sci Rep*. 2019;9(1):15198. doi:10.1038/s41598-019-51599-7

35. Liu GS, Fereydooni S, Lee MC, et al. Scoping review of deep learning research illuminates artificial intelligence chasm in otolaryngology-head and neck surgery. *NPJ Digit Med*. 2025;8(1):265. doi:10.1038/s41746-025-01693-0

© 2025 American Medical Association. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by University of Montreal user on 06/29/2026