Oral Oncology 167 (2025) 107448

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Assessment of established prognostic factors and artificial intelligence-based evaluation of tumor-infiltrating lymphocytes in oral tongue squamous cell carcinoma

Jongwon Lee [a,1,2], Ming Fan [b,1,3], Donghyeok Jo [c,4], JungBok Lee [d,5], Joon Seon Song [e,6], Hee Jin Lee [e,7], Seung-Ho Choi [f,8], Soon Yuhl Nam [f,9], Yoon Se Lee [f,10], Dong-Joo Kim [c,11], Kyung-Ja Cho [e,*,12]

[a] Department of Pathology, Korea University Guro Hospital, Seoul, Republic of Korea
[b] AIVIS Incorporation, Seoul, Republic of Korea
[c] Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea
[d] Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[e] Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[f] Department of Otolaryngology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Tongue cancer
Artificial intelligence
Tumor-infiltrating lymphocytes
Prognosis
Risk factor

ABSTRACT

Background: Traditional risk assessment for tongue cancer relies on clinicopathological parameters. Although tumor-infiltrating lymphocytes (TILs) are promising prognostic markers, their evaluation lacks standardization. This study aimed to validate established prognostic factors and introduce an artificial intelligence (AI)-based TIL assessment method.

Methods: We analyzed 139 tongue cancer cases from a single institution (2010–2017) to establish prognostic factors and developed an AI model for TIL quantification. Clinicopathological characteristics including AI- and manually assessed TILs were evaluated.

Results: The AI-assessed stromal TIL ratio exerted protective effects across all stages and demonstrated superior discriminative capability compared to manual evaluation (C-index: 0.649 vs. 0.604 for overall survival [OS]), with substantial inter-method agreement (Intraclass Correlation Coefficient = 0.796). In the multivariate analysis, a statistical model incorporating the lymph node ratio, AI-assessed stromal TIL ratio, depth of invasion grade, perineural invasion, lymphovascular invasion, and a close surgical resection margin (<5 mm) showed superior prognostic performance, with excellent discriminative power (OS area under the curve [AUC]: 0.851; recurrence-free survival [AUC]: 0.826). Stage-specific analysis revealed that advanced-stage patients were significantly affected by adverse factors and stromal TIL levels, whereas early stage patients showed trends but no statistically significant associations.

* Corresponding author at: Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea.
E-mail address: kjc@amc.seoul.kr (K.-J. Cho).
[1] These authors contributed equally to this work.
[2] Jongwon Lee: https://orcid.org/0000-0003-3057-7874.
[3] Ming Fan: https://orcid.org/0000-0002-0988-9091.
[4] Donghyeok Jo: https://orcid.org/0009-0005-2328-3474.
[5] JungBok Lee: https://orcid.org/0000-0002-1420-9484.
[6] Joon Seon Song: https://orcid.org/0000-0002-7429-4254.
[7] Hee Jin Lee: https://orcid.org/0000-0002-4963-6603.
[8] Seung-Ho Choi: https://orcid.org/0000-0001-9109-9621.
[9] Soon Yuhl Nam: https://orcid.org/0000-0002-8299-3573.
[10] Yoon Se Lee: https://orcid.org/0000-0001-6534-5753.
[11] Dong-Joo Kim: https://orcid.org/0000-0002-0988-2236.
[12] Kyung-Ja Cho: https://orcid.org/0000-0002-4911-7774.

https://doi.org/10.1016/j.oraloncology.2025.107448
Received 24 February 2025; Received in revised form 3 June 2025; Accepted 19 June 2025
Available online 3 July 2025

Conclusions: AI-based stromal TIL assessment outperformed manual TIL assessment as a prognostic marker. This AI approach robustly predicts survival when combined with factors such as the lymph node ratio and a close resection margin status (<5 mm). Our findings may enhance risk stratification, particularly in advanced-stage disease.

Introduction

Various histopathologic features including worst pattern of invasion (WPOI), tumor budding, lymphovascular invasion (LVI), perineural invasion (PNI), and extranodal extension (ENE) have been suggested as prognostic indicators in oral tongue squamous cell carcinoma (SCC) [1,2]. Among these factors, the lymph node ratio (LNR) has demonstrated superior prognostic value compared to conventional N staging [3,4]. In addition to traditional histopathologic parameters, tumor-infiltrating lymphocytes (TILs) have emerged as critical determinants of prognosis and treatment response. However, current TIL assessment methods lack standardization and suffer from significant inter-observer variability [5,6]. The subjective nature of manual TIL evaluation limits its clinical implementation and reproducibility across institutions.

Recent advances in artificial intelligence (AI) and digital pathology offer promising objective, reproducible quantification of histopathologic features, potentially overcoming the limitations of manual assessment. However, the comparative effectiveness of AI-based versus traditional pathologist-assessed TIL evaluation in oral tongue SCC remains unclear.

This study aimed to: (1) validate prognostic significance of established factors; (2) evaluate AI-based TIL assessment; and (3) develop an integrated prognostic model in 139 tongue SCC cases.

Materials and methods

Patient cohort

We analyzed 139 patients with pathologically confirmed primary oral tongue SCC who underwent surgical resection at Asan Medical Center, Seoul, Korea, between 2010 and 2017. The exclusion criteria included the presence of distant metastasis during the operation, previous treatment for head and neck cancer, undetermined lymph node status, and unavailability of pathological hematoxylin and eosin (H&E)-stained slides. This study was approved by the Ethics Committee of Asan Medical Center (approval number #2022–1429; August 2022). The requirement for written informed consent was waived by the Ethics Committee because of the retrospective nature of the study.

Histopathological assessment

All samples were independently evaluated by two investigators (J. L. and K. C.) using light microscopy, followed by a joint consensus review. Microscopic analysis focused on tumor differentiation, tumor size, depth of invasion (DOI), WPOI, tumor budding, TILs, tumor-stromal ratio (TSR), LVI, PNI, ENE, and LNR [7].

DOI was graded as: 0 (≤5 mm), 1 (>5–10 mm), and 2 (>10 mm). LNR was calculated as the number of positive lymph nodes divided by the total number of harvested lymph nodes [8]. However, raw LNR values are often small, making them potentially difficult to interpret in the context of statistical analysis. To improve interpretability and enable a more meaningful statistical evaluation, the LNR values were multiplied by 100, which rescaled the data without altering the relative ratios. WPOI was dichotomized into groups 1–4 versus 5. (Fig. 1**A–E**) [9].

Lymph node evaluation was performed in 129 patients (92.8 %, 129/139). Ten patients (7.2 %, 10/139) did not undergo lymph node assessment: seven underwent partial glossectomy without lymph node dissection because of clinically N0 disease (n = 6) or comorbidities (n = 1), and three underwent glossectomy with selective lymph node dissection but the lymph nodes could not be harvested during the procedure.

According to AJCC guidelines recommending examination of at least 10 lymph nodes in selective neck dissection, harvested lymph nodes were categorized as follows: not harvested (n = 10, 7.2 %), <10 nodes (n = 4, 2.9 %), 10–19 nodes (n = 17, 12.2 %), and ≥ 20 nodes (n = 108, 77.7 %), as presented in Table 1. Four patients had a suboptimal but adequate lymph node yield (2–8 nodes) because of patient age, surgical complexity, or anatomical factors.

For lymph node-related analyses (LNR, metastatic lymph nodes, ENE, AJCC N category, and overall stage), patients with any lymph node harvest were included (n = 129), while the 10 patients without lymph node assessment were excluded from these specific analyses to maintain analytical validity.

Surgical margin status was evaluated and classified into three categories based on the 2024–2025 National Comprehensive Cancer Network guidelines: positive margin (tumor involvement), close margin (<5 mm from tumor), and clear margin (≥5 mm from tumor). Margin distance was determined by measuring the shortest distance from the invasive tumor edge to the nearest resection margin using standard histopathological methods. Among the 139 surgically resected tongue cancer cases, complete slide sets with comprehensive resection margin assessment were available for 117 cases. All margin assessments were confirmed by consensus meetings between two pathologists (J. L. and K. J. C.).

Tumor budding was assessed at 200x magnification and categorized as low-grade (≤4 buds/HPF), intermediate-grade (5–9 buds/HPF), and high-grade (≥10 buds/HPF) (Fig. 1**F**) [10]. The TSR was calculated with ≤ 50 % categorized as “low” (stroma-rich) and > 50 % as “high” (stroma-poor) (Fig. 1**C**) [11,12].

Manual TIL assessment was performed by experienced pathologists (J. L. and K. C.) using a standardized 10-tier grading system. Stromal TILs at the invasive tumor front (‘Stromal TIL’) and TILs evaluated at tumor center (‘Central TIL’) were evaluated in representative HPFs (×400 magnification) and scored on a 10-tier grading system, corresponding to TIL percentages in 10 % increments (0: 0–10 %, 1: 10–20 %, …, 9: 90–100 %).

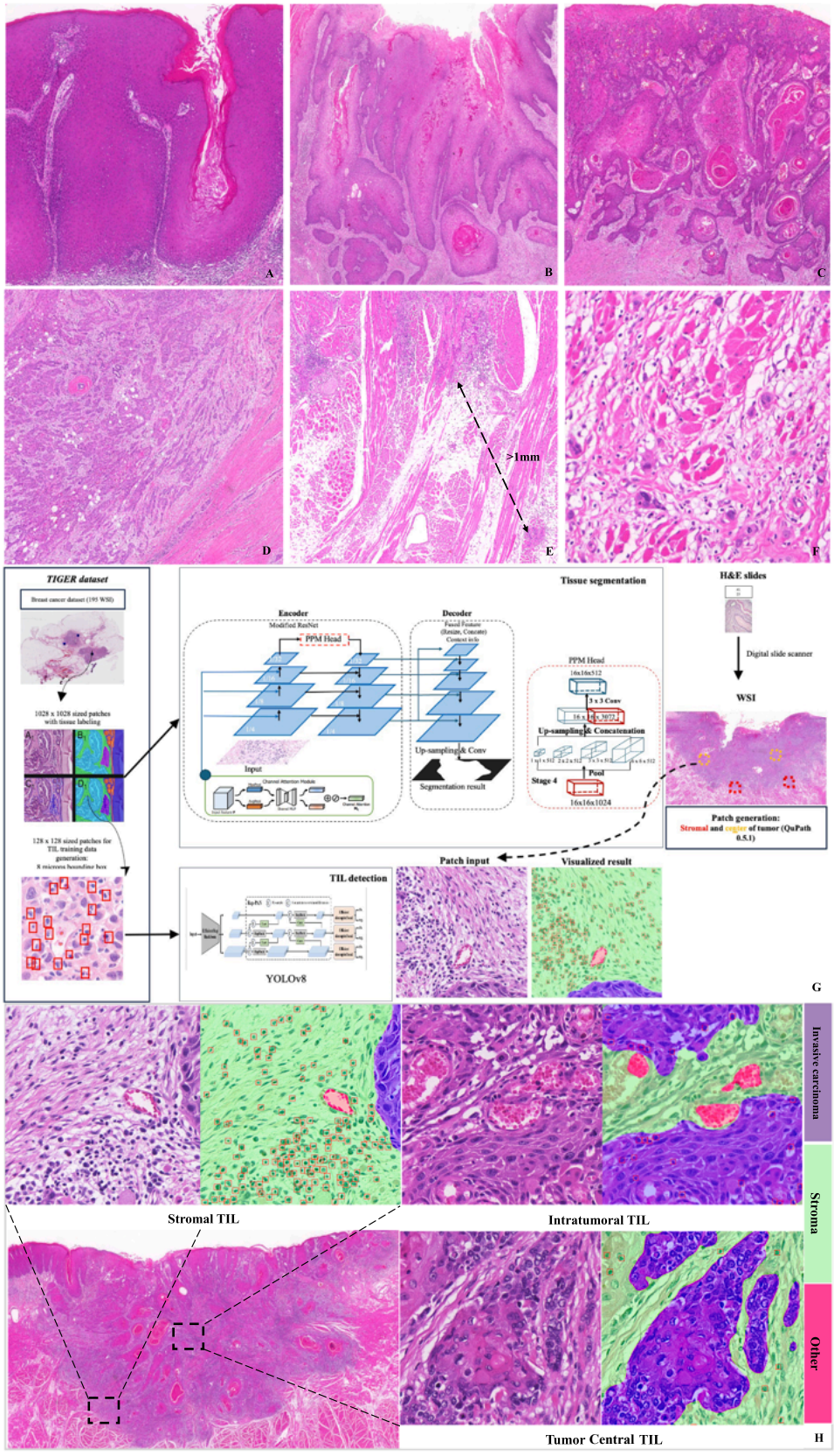
LVI, PNI, and ENE were recorded as present or absent [13]. Tumor size was categorized using 2-cm cut-offs, following the AJCC 8th T stages: group 1: ≤2 cm, group 2: >2 cm and ≤ 4 cm, and group 3: >4 cm [14].

Ai-based TIL Detection system

We developed an AI-based system for automated TIL quantification using transfer learning methodology. Detailed methodology and validation results are provided in Supplementary Methods.

Statistical analysis

Overall survival (OS) was calculated from surgery to death or last follow-up, with survival status confirmed for all 139 patients through national health insurance records. Recurrence-free survival (RFS) was analyzed in 96 patients with available data. Univariate and multivariate Cox proportional hazards regression analyses were performed. Four model versions were developed to compare ENE vs. LNR and AI vs. pathologist TIL assessment. Model performance was evaluated using time-dependent ROC curves and validated using bootstrap resampling. Detailed statistical methodology is provided in Supplementary Methods.



(caption on next page)

Fig. 1. Histopathological parameters and AI model development. (A) WPOI Grade 1: Tumor exhibits expansile invasion (40x). (B) WPOI Grade 2: tumor invading in solid cords and strands (40x). (C) WPOI Grade 3: invasive islands of the tumor with >15 cell clusters (arrows). In this case, the tumor to stromal ratio is high (40x). (D) WPOI Grade 4: invasive tumor islands with <15 cell clusters (40x). (E) WPOI Grade 5: tumor islands more than 1 mm away from the progressive end of the tumor (40x). (F) High tumor budding at the invasive front (200x). (G) End-to-end deep learning framework for tumor-infiltrating lymphocyte (TIL) detection. (H) Representative images demonstrating AI-based TIL assessment in different tumor compartments. The lower left panel shows a whole slide image with annotated regions of interest. Three distinct compartments were analyzed: stromal (invasive front) TILs, intratumoral TILs, and central TILs with original hematoxylin and eosin images (left) and corresponding AI annotations (right). AI annotations highlight tumor areas (purple), stromal regions (green), and identified TILs (red bounding boxes). Stromal TILs show lymphocytic infiltration at the tumor-stroma interface, intratumoral TILs represent lymphocytes within tumor nests, and central TILs indicate lymphocytic infiltration in the central portion of the tumor. Original magnification: whole slide image 10×, detailed images 200×. WPOI, worst pattern of invasion; AI, artificial intelligence; TIL, tumor-infiltrating lymphocyte; H&E, hematoxylin and eosin.

Statistical significance was set at $p < 0.05$.

Results

Patient characteristics and clinicopathological comparison

The clinicopathological characteristics of 139 patients are summarized in Table 1. During follow-up, 75 patients (75/139, 54.0 %) remained alive and 64 patients (64/139, 46.0 %) died. Among the deceased patients, 49 (76.6 % of deaths, 35.3 % of total cohort) died from confirmed cancer-related causes, while 15 (15/139, 23.4 %) died from unknown causes, with 13 of these patients having advanced-stage disease and dying within 20 months of last contact.

Deceased patients had significantly larger tumors (3.3 ± 1.5 cm vs. 2.1 ± 1.1 cm, $p < 0.001$), greater DOI (12.1 ± 7.0 mm vs. 7.2 ± 4.9 mm, $p < 0.001$), and higher prevalence of adverse features: LVI (29.7 % vs. 13.3 %, $p = 0.031$), PNI (39.1 % vs. 13.3 %, $p = 0.001$), and ENE (39.1 % vs. 5.3 %, $p < 0.001$). Nodal disease burden was substantially higher with more metastatic lymph nodes (3.0 ± 3.5 vs. 0.5 ± 1.1, $p < 0.001$) and higher LNRs (6.3 ± 7.1 vs. 0.9 ± 1.9, $p < 0.001$).

Both AI-assessed and pathologist-assessed (manual) TIL parameters demonstrated significant differences between groups. AI-assessed stromal TIL ratios were lower in deceased patients (53.1 ± 15.8 % vs. 63.3 ± 17.6 %, $p = 0.001$), as were central TIL ratios (45.2 ± 12.5 % vs. 52.4 ± 14.9 %, $p = 0.003$). Similarly, the number of pathologist-assessed stromal TILs was lower in deceased patients (4.8 ± 2.2 vs. 5.9 ± 2.0, $p = 0.002$), along with central TILs (3.4 ± 2.2 vs. 4.8 ± 2.1, $p < 0.001$). Intratumoral TIL ratios showed no significant difference between groups ($p = 0.188$).

Advanced disease staging was more prevalent in the deceased group, with higher proportions of T3-4 tumors (59.4 % vs. 17.3 %, $p < 0.001$), N2-3 nodal disease (59.4 % vs. 12.0 %, $p < 0.001$), and overall stages III-IV (65.6 % vs. 30.6 %, $p < 0.001$). No significant differences were observed in tumor budding, the tumor-stroma ratio, WPOI, differentiation grade, age, or sex distribution between groups.

Correlation analysis and variable selection

Correlation analysis was performed to guide variable selection and avoid multicollinearity in multivariate modeling (Fig. 2). The normalized Spearman correlation matrix (Fig. 2**A**) revealed several critical relationships. Among nodal parameters, LNR demonstrated extremely high correlation with metastatic lymph node count ($r = 0.98$), leading to LNR selection as the representative nodal parameter. Given the established clinical importance of ENE as an independent predictor, separate ENE-based and LNR-based model versions were developed.

Key correlations informing variable selection included DOI continuous measurement with DOI grade ($r = 0.94$), leading to DOI grade selection for ordinal prognostic categorization; tumor size showing moderate correlations with invasion parameters ($r = 0.45$–0.71), resulting in exclusion due to multicollinearity concerns; and AI-assessed central TIL ratio with stromal TIL ratio ($r = 0.60$), with stromal TIL selected based on established literature emphasizing invasive front immune infiltration. The intratumoral TIL ratio demonstrated weaker correlations with other immune parameters (<0.30) and showed no significant prognostic associations in univariate analysis, leading to its exclusion.

AI- and pathologist-assessed TIL parameters demonstrated substantial inter-method correlation ($r = 0.661$, Intraclass Correlation Coefficient [ICC] = 0.796, Table 2), validating automated assessment reliability. Point-biserial correlation analysis revealed important associations: surgical margin status (<5 mm) with T category ($r = 0.41$), DOI ($r = 0.30$), and DOI grade ($r = 0.30$) (Fig. 2**B**); ENE with multiple continuous parameters including LNR, metastatic lymph node count, and invasion depth (Fig. 2C); and PNI with DOI ($r = 0.43$), tumor size ($r = 0.37$), and DOI grade ($r = 0.35$) (**Sup.** Fig. 1**A**).

Variables with high correlations ($|r| \geq 0.7$) included metastatic lymph nodes and LNR ($r = 0.975$), DOI and DOI grade ($r = 0.944$), and tumor size with DOI ($r = 0.746$) and DOI grade ($r = 0.740$). Analysis of LNR distribution across harvested lymph node categories (Fig. 2**D**) demonstrated the critical importance of adequate nodal sampling. Traditional prognostic parameters including tumor budding, TSR, and WPOI showed generally weak correlations with other variables, consistent with their lack of statistical significance in subsequent analyses. Based on these findings, four multivariate model versions were constructed: ENE + AI stromal TIL, LNR + AI stromal TIL, ENE + pathologist stromal TIL, and LNR + pathologist stromal TIL.

TIL distribution patterns and risk stratification

Analysis of AI-assessed stromal TIL ratios across AJCC 8th Edition stages demonstrated a progressive decrease with advancing stage (Fig. 3**A**). Early stage tumors (stages I-II) showed higher TIL infiltration with greater variability compared to advanced-stage disease (stages III-IV), indicating heterogeneous immune responses that become more uniformly suppressed in advanced disease. This inverse relationship between stage and TIL infiltration was consistent across both AI-assessed and pathologist-assessed parameters.

Quartile-based risk stratification effectively distinguished prognostic groups for both assessment methods, specifically chosen to enhance reproducibility across different institutions and patient populations, avoiding dataset-specific overfitting inherent in ROC-derived optimal cutoffs (Fig. 3**B,** Table 2). AI-assessed stromal TIL ratios using a median cutoff demonstrated significant survival advantages, with high TIL patients showing superior OS (hazard ratio [HR] = 0.40, 95 % confidence interval [CI]: 0.24–0.66) and RFS (HR = 0.53, 95 % CI: 0.29–0.97). Using the more conservative 75th percentile cutoff (66.0 %), the high TIL group comprised 25.2 % of patients with maintained prognostic significance (OS HR = 0.44, 95 % CI: 0.22–0.87).

Similarly, pathologist-assessed stromal TIL demonstrated comparable prognostic stratification **(**Fig. 3**C).** Using the median cutoff (6.0 on the 10-point scale), patients with high TIL scores showed significant survival benefits for both OS (HR = 0.52, 95 % CI: 0.32–0.86) and RFS (HR = 0.55, 95 % CI: 0.30–1.00). The 75th percentile approach (7.0 cutoff) identified 38.8 % of patients as high-risk with sustained prognostic discrimination.

Comparative performance analysis revealed that AI-based assessment demonstrated superior prognostic discrimination compared to manual assessment (Fig. 3**D,** Table 2). AI achieved higher C-index values for both OS (0.649 vs. 0.604) and RFS (0.631 vs. 0.608), indicating

Table 1
Clinicopathologic Characteristics of Patients with Tongue Cancer Based on Survival Status (n = 139).

Characteristics	Category	Total n = 139 (%)	Alive n = 75 (54.0)	Deceased[7] n = 64 (46.0)	P-value
Demographics					
Age at diagnosis, years	Mean ± SD	60.6 ± 14.4	59.5 ± 14.4	62.0 ± 14.5	0.295
Sex	Female	56 (40.3)	33 (44.0)	23 (35.9)	0.428
	Male	83 (59.7)	42 (56.0)	41 (64.1)	
Tumor Characteristics					
Tumor size, cm	Mean ± SD	2.7 ± 1.4	2.1 ± 1.1	3.3 ± 1.5	<0.001
DOI, mm	Mean ± SD	9.4 ± 6.4	7.2 ± 4.9	12.1 ± 7.0	<0.001
DOI grade	≤5 mm	42 (30.2)	31 (41.3)	11 (17.2)	<0.001
	5 mm–10 mm	47 (33.8)	30 (40.0)	17 (26.6)	
	>10 mm	50 (36.0)	14 (18.7)	36 (56.2)	
Histopathologic Features					
Differentiation	Well differentiated	77 (55.4)	48 (64.0)	29 (45.3)	0.087
	Moderately differentiated	53 (38.1)	23 (30.7)	30 (46.9)	
	Poorly differentiated	9 (6.5)	4 (5.3)	5 (7.8)	
Tumor budding	<5	92 (66.2)	48 (64.0)	44 (68.8)	0.203
	5–9	30 (21.6)	20 (26.7)	10 (15.6)	
	≥10	17 (12.2)	7 (9.3)	10 (15.6)	
Tumor-stroma ratio	≤50 %	36 (25.9)	21 (28.0)	15 (23.4)	0.676
	>50 %	103 (74.1)	54 (72.0)	49 (76.6)	
WPOI	1	1 (0.7)	1 (1.3)	0 (0.0)	0.117
	2	9 (6.5)	7 (9.3)	2 (3.1)	
	3	42 (30.2)	27 (36.0)	15 (23.4)	
	4	68 (48.9)	30 (40.0)	38 (59.4)	
	5	19 (13.7)	10 (13.3)	9 (14.1)	
WPOI grade	1–4	120 (86.3)	65 (86.7)	55 (85.9)	1.000
	5	19 (13.7)	10 (13.3)	9 (14.1)	
LVI	Not identified	110 (79.1)	65 (86.7)	45 (70.3)	0.031
	Present	29 (20.9)	10 (13.3)	19 (29.7)	
PNI	Not identified	104 (74.8)	65 (86.7)	39 (60.9)	0.001
	Present	35 (25.2)	10 (13.3)	25 (39.1)	
ENE[1]	Not identified	100 (71.9)	62 (82.7)	38 (59.4)	<0.001
	Present	29 (20.9)	4 (5.3)	25 (39.1)	
	Not assessed	10 (7.2)	9 (12.0)	1 (1.6)	
TIL Parameters					
AI-assessed Parameters					
Stromal TIL ratio	Mean ± SD	58.6 ± 17.4	63.3 ± 17.6	53.1 ± 15.8	0.001
Central TIL ratio	Mean ± SD	49.2 ± 14.2	52.4 ± 14.9	45.2 ± 12.5	0.003
Intratumoral ratio	Mean ± SD	21.0 ± 8.5	21.9 ± 8.0	20.0 ± 9.1	0.188
Pathologist-assessed Parameters					
Stromal TIL	Mean ± SD	5.4 ± 2.2	5.9 ± 2.0	4.8 ± 2.2	0.002
Central TIL	Mean ± SD	4.2 ± 2.3	4.8 ± 2.1	3.4 ± 2.2	<0.001
Lymph Node Parameters					
Metastatic LNs[1]	Mean ± SD	1.7 ± 2.8	0.5 ± 1.1	3.0 ± 3.5	<0.001
Lymph node ratio1′2	Mean ± SD	3.4 ± 5.6	0.9 ± 1.9	6.3 ± 7.1	<0.001
Harvested LN count1′3	Mean ± SD	41.2 ± 26.4	36.2 ± 25.9	46.9 ± 26.1	0.017
Harvested LN categories	Not harvested	10 (7.2)	9 (12.0)	1 (1.6)	0.079
	<10	4 (2.9)	1 (1.3)	3 (4.7)	
	10–19	8 (5.8)	4 (5.3)	4 (6.2)	
	≥20	117 (84.2)	61 (81.3)	56 (87.5)	
Surgical Approach					
Operation Name	Excision, Partial glossectomy	8 (5.8)	7 (9.3)	1 (1.6)	<0.001
	Glossectomy and SLND	73 (52.5)	54 (72.0)	19 (29.7)	
	Glossectomy and MRND	57 (41.0)	14 (18.7)	43 (67.2)	
	Others[4]	1 (0.7)	0 (0.0)	1 (1.6)	
Surgical Margins					
Surgical margin status	Not involved	133 (95.7)	74 (98.7)	59 (92.2)	0.032
	Involved by carcinoma	5 (3.6)	0 (0.0)	5 (7.8)	
	Involved by LGD	1 (0.7)	1 (1.3)	0 (0.0)	
Margin distance5	<5 mm from tumor	58 (41.7)	26 (34.7)	32 (50.0)	<0.001
	≥5 mm from tumor	53 (38.1)	41 (54.7)	12 (18.8)	
	Not assessed	28 (20.1)	8 (10.7)	20 (31.2)	
Staging					
AJCC 8th T Category	1	36 (25.9)	28 (37.3)	8 (12.5)	<0.001
	2	52 (37.4)	34 (45.3)	18 (28.1)	
	3	18 (12.9)	6 (8.0)	12 (18.8)	
	4	33 (23.7)	7 (9.3)	26 (40.6)	
AJCC 8th N Category[1]	0	67 (48.2)	49 (65.3)	18 (28.1)	<0.001
	1	15 (10.8)	8 (10.7)	7 (10.9)	
	2	25 (18.0)	6 (8.0)	19 (29.7)	
	3	22 (15.8)	3 (4.0)	19 (29.7)	
	Not assessed	10 (7.2)	9 (12.0)	1 (1.6)	
Overall stage[1]	I	24 (17.3)	19 (25.3)	5 (7.8)	<0.001
	II	30 (21.6)	24 (32.0)	6 (9.4)	
	III	20 (14.4)	10 (13.3)	10 (15.6)	
	IV	55 (39.6)	13 (17.3)	42 (65.6)	

(continued on next page)

Table 1 (*continued*)

Characteristics	Category	Total n = 139 (%)	Alive n = 75 (54.0)	Deceased[7] n = 64 (46.0)	P-value
	Not assessed	10 (7.2)	9 (12.0)	1 (1.6)	
Clinical Outcomes					
Survival, months	Mean ± SD	52.2 ± 37.4	74.7 ± 34.0	23.1 ± 28.7	<0.001
Recurrence*[8]	No	53 (38.1)	53 (70.7)	0 (0.0)	<0.001
	Yes	43 (30.9)	7 (9.3)	36 (56.2)	
	Unknown	43 (30.9)	15 (20.0)	28 (43.8)	
Time to recurrence, months	Mean ± SD	31.2 ± 42.8	102.2 ± 31.4	8.1 ± 8.0	<0.001
Follow-up period, months	Mean ± SD	52.2 ± 37.4	74.7 ± 34.0	23.1 ± 28.7	<0.001

[1]Variables with missing data (7.2%, 10/139 cases): ENE, Metastatic LNs, LNR, AJCC 8th N Category, and Overall stage were marked as not assessed in cases where LN harvest was not performed or failed to harvest LNs after dissection, primarily in early stage patients.
[2]Lymph node ratio (LNR) was calculated as the number of positive lymph nodes divided by the total number of harvested lymph nodes. To ensure the validity of LNR as a prognosticator, analyses were performed only on a subgroup with adequate lymph node yield (≥10 nodes).
[3]Harvested LN was categorized into 4 groups according to the AJCC guidelines which indicate that at least 10 lymph nodes should be examined in selective neck dissection for adequate staging.
[4]Others: Glossectomy, mandibulectomy with modified radical neck dissection.
[5]The 111 patients with cancer who had uninvolved surgical margins and available margin status evaluation were classified as having margins ≥5mm or <5mm according to the 2024–2025 updated NCCN guidelines.
[6]Missing values in the statistical analysis were handled by excluding cases with missing data from the specific analyses (complete case analysis). This approach was chosen to maintain the integrity of the survival models while maximizing the use of available data.
[7]Among the 64 deceased patients, 49 (76.6%) died from confirmed cancer-related causes, while 15 (23.4%) died from unknown causes. Of the 15 patients with unknown cause of death, 13 had advanced-stage disease (Stage III-IV) and died within 20 months of last contact, while 2 had early-stage disease (Stage I).
[8]*Recurrence status was determined based on complete clinical and imaging follow-up data. Patients categorized as 'Unknown' (43 out of 139 patients, 31%) had confirmed survival status through national health insurance records but insufficient clinical follow-up to determine recurrence status definitively. To maintain analytical integrity, these patients were excluded from all recurrence-free survival analyses but were retained in overall survival analyses where their vital status was known with certainty.
Abbreviations: DOI, depth of invasion; WPOI, worst pattern of invasion; LVI, lymphovascular invasion; PNI, perineural invasion; ENE, extranodal extension; LN, lymph node; LNR, lymph node ratio; SLND, selective neck dissection; MRND, modified radical neck dissection; TIL, tumor-infiltrating lymphocytes; AJCC, American Joint Committee on Cancer; LGD, low grade dysplasia.
Notes: Data are presented as n (%) for categorical variables and mean ± standard deviation for continuous variables. P-values were calculated using the Chi-square test or Fisher's exact test for categorical variables and the Student's t-test or Mann-Whitney U test for continuous variables.

better discriminative capability. However, 5-year AUC analysis showed comparable predictive performance between methods (AI: 0.682, 95 % CI: 0.573–0.764; manual: 0.636, p = 0.759), suggesting that both approaches provide clinically meaningful prognostic information. Inter-method validation demonstrated substantial agreement between AI and pathologist assessment (r = 0.661, ICC = 0.796, 95 % CI: 0.72–0.85), confirming the reliability of automated quantification while highlighting the complementary value of expert pathologist interpretation.

Univariate and multivariate Cox Regression analysis

Univariate Cox proportional hazard analysis identified multiple significant prognostic factors (**Sup.** Table 1). Traditional tumor characteristics including the DOI grade (OS: HR = 2.67, 95 % CI: 1.86–3.83, p < 0.001), surgical margin < 5 mm (OS: HR = 3.45, 95 % CI: 1.79–6.67, p < 0.001), and histopathologic features such as LVI (OS: HR = 2.45, 95 % CI: 1.41–4.23, p = 0.001) and PNI (OS: HR = 2.63, 95 % CI: 1.59–4.36, p < 0.001) demonstrated strong associations with both OS and RFS.

Among nodal parameters, ENE (OS: HR = 3.91, 95 % CI: 2.33–6.55, p < 0.001; RFS: HR = 3.27, 95 % CI: 1.72–6.23, p < 0.001) and LNR (OS: HR = 1.10, 95 % CI: 1.06–1.13, p < 0.001; RFS: HR = 1.12, 95 % CI: 1.07–1.17, p < 0.001) emerged as particularly strong predictors. The AI-assessed stromal TIL ratio demonstrated significant protective effects for both endpoints (OS: HR = 0.97, 95 % CI: 0.95–0.98, p < 0.001; RFS: HR = 0.97, 95 % CI: 0.94–0.99, p = 0.004).

Multivariate analysis systematically addressed the correlation between ENE and LNR while comparing AI and pathologist stromal TIL assessment approaches. Following comprehensive four-model comparison analysis with OS and RFS (**Sup.** Tables 2 and 3), two representative models were selected: (1) a traditional approach using ENE + manual stromal TIL assessment, and (2) a novel approach employing LNR + AI-based stromal TIL ratio quantification (Table 3).

Using baseline variables with confirmed statistical independence (DOI grade, differentiation, LVI, PNI, and resection margin status), both models demonstrated robust predictive performance across survival endpoints. The novel LNR + AI stromal TIL model demonstrated superior performance for both survival endpoints (OS AUC = 0.851, RFS AUC = 0.826). For OS, the LNR exhibited the strongest prognostic association among all variables (HR = 4.44, 95 % CI: 2.01–9.81, p < 0.001), while the AI-assessed stromal TIL provided robust independent prognostic value (HR = 0.61, 95 % CI: 0.42–0.88, p = 0.009) alongside close surgical margins < 5 mm (HR = 2.13 95 % CI: 1.05–4.35, p = 0.03). For RFS, the model maintained strong performance with the LNR (HR = 3.90, 95 % CI: 1.77–8.59, p = 0.001) and AI stromal TIL (HR = 0.66, 95 % CI: 0.46–0.95, p = 0.02) both achieving independent significance.

In contrast, the traditional ENE + pathologist-assessed stromal TIL model achieved strong but marginally lower discrimination (OS AUC = 0.842, RFS AUC = 0.799). For OS, ENE emerged as a significant predictor (HR = 3.09, 95 % CI: 1.75–5.47, p < 0.001) alongside close surgical margins < 5 mm (HR = 2.22, 95 % CI: 1.09–4.55, p = 0.03), while pathologist-assessed stromal TIL showed borderline significance (HR = 0.71, 95 % CI: 0.50–1.01, p = 0.05). Notably, pathologist-assessed stromal TIL failed to achieve statistical significance in RFS analysis across both ENE- (HR = 0.76, 95 % CI: 0.54–1.07, p = 0.12) and LNR-based models (HR = 0.78, 95 % CI: 0.56–1.10, p = 0.15), while AI-assessed parameters maintained consistent prognostic value across both survival endpoints.

Stage-Specific analysis and t-distributed Stochastic Neighbor Embedding validation

To address potential stage mixture bias affecting prognostic variable interpretation, comprehensive stage-specific analyses were performed (**Sup.** Tables 4 and 5). Stage-specific analyses were conducted on 129 patients with available lymph node data, excluding the 10 patients without adequate lymph node harvest. Early stage disease (stages I-II, n = 54 for OS, n = 39 for RFS) demonstrated markedly different prognostic patterns compared to advanced-stage disease (stages III-IV, n = 75 for OS, n = 50 for RFS). In early stage tumors, univariate analysis revealed minimal prognostic discrimination, with only age emerging as a significant factor for OS (HR = 1.06, 95 % CI: 1.01–1.11, p < 0.05),

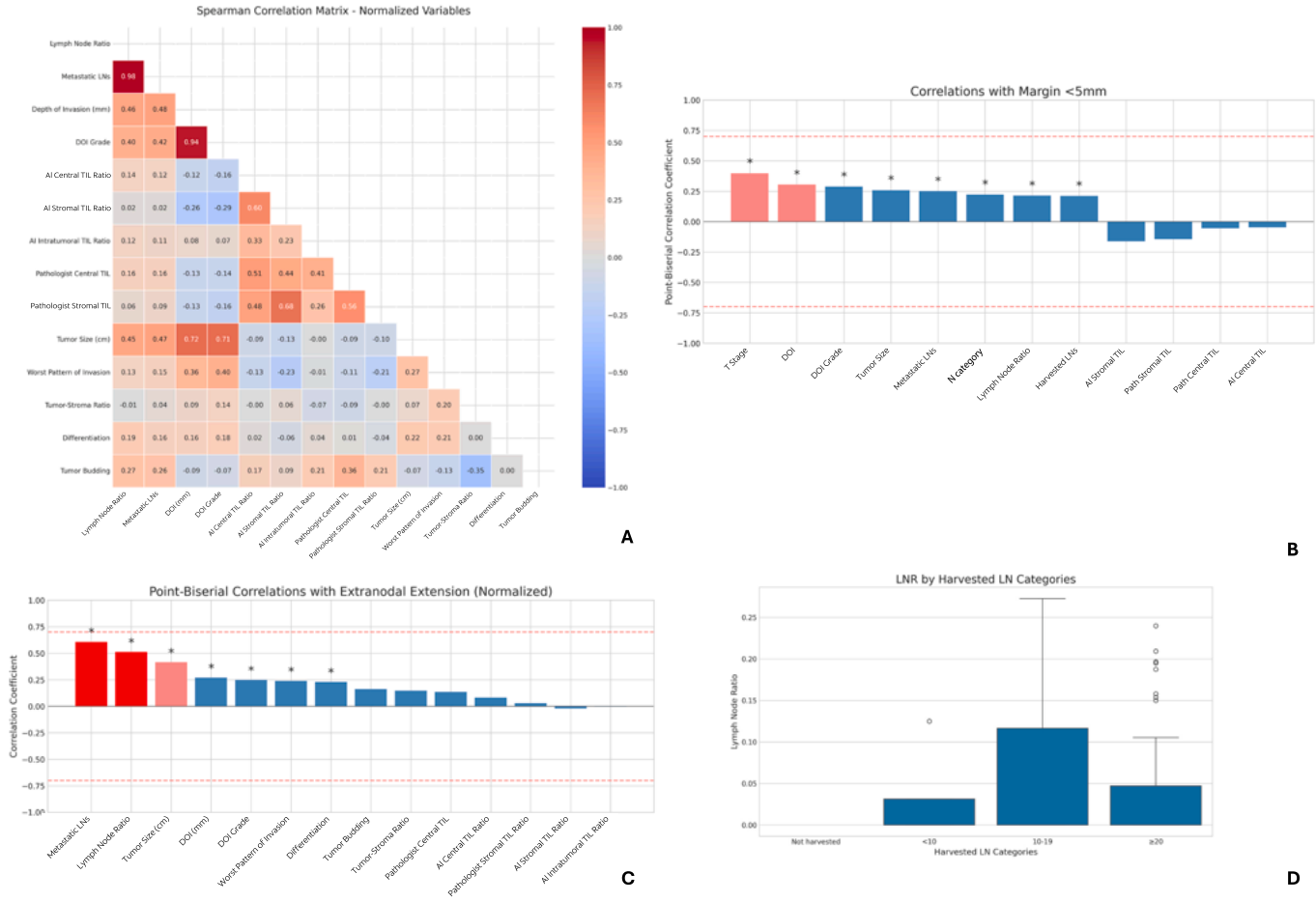


Fig. 2. Correlation Analysis and Variable Selection Strategy. (A) Spearman correlation matrix of normalized continuous variables demonstrating key relationships that informed variable selection for multivariate modeling. High correlations ($|r| \geq 0.7$) between variables such as the LNR and metastatic lymph nodes (r=0.98) and DOI and DOI grade (r=0.94) necessitated careful variable selection to avoid multicollinearity. The color scale represents the correlation strength from -1.0 (blue) to +1.0 (red). (B) Point-biserial correlations with the surgical margin status (<5 mm), showing significant associations with the T category, DOI, and DOI grade. Red bars indicate positive correlations while blue bars represent negative correlations. Asterisks (*) denote statistically significant correlations ($p < 0.05$). (C) Point-biserial correlations with ENE, demonstrating strong associations with multiple continuous parameters including the LNR, metastatic lymph node count, and invasion depth, supporting the interconnected nature of aggressive tumor features. (D) LNR distribution across harvested lymph node categories, validating the critical importance of adequate nodal sampling. Patients with 10-19 harvested nodes showed higher mean LNR values compared to those with ≥20 nodes, supporting analytical restriction to cases with adequate nodal harvest per the AJCC guidelines. AJCC, American Joint Committee on Cancer; DOI, depth of invasion, ENE, extranodal extension; LNR, lymph node ratio; TIL, tumor infiltrating lymphocyte; AI, artificial intelligence; LN, lymph node.

while no variables achieved significance for RFS. Traditional pathological parameters including the tumor size, DOI, and histopathologic features showed no significant associations with either survival endpoint in early stage disease.

Conversely, advanced-stage disease exhibited multiple significant prognostic factors across both survival endpoints. For OS, tumor size (HR = 1.53, 95 % CI: 1.24–1.90, $p < 0.001$), PNI (HR = 2.13, 95 % CI: 1.22–3.74, $p < 0.01$), and AI-assessed stromal TIL parameters (HR = 0.64, 95 % CI: 0.49–0.84, $p < 0.01$) demonstrated significant associations. For RFS, the AI stromal TIL (HR = 0.62, 95 % CI: 0.44–0.86, $p < 0.01$), metastatic lymph node count (HR = 1.12, 95 % CI: 1.03–1.23, $p < 0.05$), and LNR (HR = 2749.25, 95 % CI: 21.22–356117.09, $p < 0.01$) emerged as significant predictors. This stage-specific analysis revealed that TIL-based prognostic value is primarily driven by advanced-stage disease, where immune microenvironment characteristics become critically important for both OS and recurrence prediction.

Five-fold cross-validation of stage-specific multivariate models demonstrated adequate performance for advanced-stage disease (C-index = 0.343 for OS, 0.266 for RFS) with stable coefficient estimates (Coefficient of Variation [CV] < 1.0 for core variables), while early stage disease models showed limited discriminative ability due to low event rates (11 deaths in 54 patients for OS, 11 recurrences in 39 patients for RFS), supporting the clinical observation that early stage tongue cancer generally has favorable outcomes regardless of traditional prognostic markers. Five-fold cross-validation of stage-specific multivariate models demonstrated adequate performance for advanced-stage disease (C-index = 0.343 for OS, 0.266 for RFS) with stable coefficient estimates (CV < 1.0 for core variables), while early stage disease models showed limited discriminative ability because of low event rates (11 deaths in 54 patients), supporting the clinical observation that early stage tongue cancer generally has favorable outcomes regardless of traditional prognostic markers.

To address concerns about the low number of samples and stage mixture bias affecting tumor budding and TSR results as prognostic variables, t-SNE dimensional reduction analysis was performed (**Sup.** Fig. 2). Overall cohort t-SNE analysis (**Sup.** Fig. 2**A**) revealed heterogeneous clustering patterns when analyzed without stage stratification. Stage-specific t-SNE analysis revealed markedly different clustering behaviors between disease stages. Early stage disease (stages I-II, **Sup.** Fig. 2**B**) demonstrated relatively homogeneous clustering patterns, with patients showing minimal separation by traditional risk factors, including the budding level and TSR status. The compact clustering in

Table 2
AI vs Manual Stromal TIL Assessment: Validation and Clinical Implementation (Part A: Method Agreement and Correlation Part B: Quartile-based Risk Stratification Performance Part C: Predictive Performance Comparison).

Assessment Method	Mean ± SD	Correlation (r)	ICC (95 % CI)	Agreement Level
AI Stromal TIL ratio[a]	58.3 ± 17.2	0.661	0.796 (0.72–0.85)	Substantial
Manual Stromal TIL[b]	5.3 ± 2.1			

Method & Cutoff	**High TIL Group**	**OS HR (95% CI)**	**RFS HR (95% CI)**	**Clinical Utility**
AI Stromal TIL ratio				
Median (Q2: 5.6×10^{-4})	70/139 (50.4%)	0.40 (0.24–0.66)	0.53 (0.29–0.97)	Balanced groups, optimal sensitivity
75th percentile (Q3: 6.6×10^{-4})	35/139 (25.2%)	0.44 (0.22–0.87)	0.61 (0.28–1.31)	Conservative approach, high specificity
Manual Stromal TIL				
Median (Q2: 6)	78/139 (56.1%)	0.52 (0.32–0.86)	0.55 (0.30–1.00)	Standard clinical approach
75th percentile (Q3: 7)	54/139 (38.8%)	0.45 (0.26–0.79)	0.51 (0.27–0.98)	Selective high-risk identification

Performance Metric	AI Stromal TIL	Manual Stromal TIL	Statistical Comparison	Clinical Advantage
C-index (OS)	0.649	0.604	AI superior	Better discrimination capability
C-index (RFS)	0.631	0.608	AI superior	Consistent prognostic advantage
5-year AUC (95% CI)	0.682 (0.573–0.764)	0.636	p = 0.759	Comparable predictive performance
Reproducibility	Excellent	Good	AI consistent	Objective, standardized quantification
Clinical interpretation	Objective	Contextual	Complementary	Both methods provide clinical value

Abbreviations
TIL, tumor-infiltrating lymphocytes; OS, overall survival; RFS, recurrence-free survival; HR, hazard ratio; CI, confidence interval; ICC, intraclass correlation coefficient; AUC, area under the curve.
Key Finding: AI demonstrates superior prognostic discrimination compared to manual assessment (C-index: 0.649 vs 0.604 for OS; 0.631 vs 0.608 for RFS).
Clinical Translation: Both methods show significant protective effects with substantial inter-method correlation (r = 0.661, ICC = 0.796).
Quartile Strategy: Percentile-based cutoffs enhance reproducibility across institutions and avoid dataset-specific overfitting.
Implementation: AI provides objective, standardized quantification; manual assessment contributes clinical expertise and contextual interpretation.
External Validation: Independent cohort validation confirms biological relevance and clinical applicability of AI-based measurements.
[a] AI Stromal TIL: Tumor-infiltrating lymphocytes quantified using artificial intelligence-based stromal analysis, values scaled × 100,000 for clinical interpretation.
[b] Manual Stromal TIL: Tumor-infiltrating lymphocytes assessed by pathologist using 1–9 scale at invasive tumor front, values scaled × 10.

early stage disease reflected the generally favorable prognosis and limited prognostic heterogeneity in this population, explaining why these parameters show reduced discriminative value.

Advanced-stage disease (stages III-IV, **Sup.** Fig. 2**C**) exhibited distinct clustering patterns with clear separation between different risk groups. Patients clustered more distinctly based on the budding levels and TSR status, demonstrating meaningful prognostic heterogeneity. The more dispersed clustering pattern in advanced disease validated the increased relevance of traditional prognostic markers in this population, where tumor biology becomes more aggressive and variable.

Comparison of model performance and validation analysis

Comparative analysis of the four multivariate model versions revealed that LNR-based models demonstrated superior predictive performance compared to ENE-based models (Fig. 3**E, F; Sup.** Tables 2 and 3). For OS, the LNR + AI stromal TIL model achieved the highest discrimination (AUC = 0.851), followed by ENE + AI stromal TIL (AUC = 0.850), ENE + pathologist stromal TIL (AUC = 0.842), and LNR + pathologist stromal TIL (AUC = 0.844). For RFS, the LNR + pathologist TIL model performed best (AUC = 0.831), followed by the LNR + AI TIL model (AUC = 0.826).

Bootstrap validation with 2000 replicates confirmed model stability across all versions, with optimism-corrected AUC values maintaining statistical significance. Five-fold cross-validation demonstrated consistent coefficient stability (CV < 1.0) for core prognostic variables, supporting model robustness and generalizability. To address concerns about stage mixture bias affecting prognostic variables, comprehensive stage-specific analyses were performed (**Sup.** Tables 3 and 4). Early stage disease (I-II) showed distinct prognostic patterns with age as the primary predictor, while advanced-stage disease (III-IV) demonstrated significant associations with traditional pathological parameters and TIL ratios.

External validation was conducted using an independent cohort of 21 patients with primary tongue SCC from another institution (Korea University Guro Hospital, **Sup.** Fig. 2**D**). Despite the limited sample size, AI-assessed TIL measurements demonstrated a consistent inverse correlation with the disease stage (Spearman ρ = -0.217), replicating the pattern observed in the primary cohort. t-SNE analysis of the validation cohort showed clustering patterns consistent with the development dataset, supporting the biological relevance and institutional transferability of AI-based TIL quantification.

Discussion

This study demonstrated that AI-based digital pathology can provide objective TIL quantification in tongue cancer, addressing traditional subjective interpretation challenges. LNR and AI-quantified TIL measurements emerged as robust prognostic indicators, outperforming traditionally emphasized parameters such as WPOI, TSR, and tumor budding. These factors demonstrated a significant prognostic value in advanced-stage disease.

The significance of TILs in tongue cancer has been increasingly recognized, with recent studies emphasizing the importance of stromal inflammation [5,6,15]. Our AI-based approach revealed distinct spatial patterns of immune cell infiltration and their prognostic implications. The moderate stromal-central TIL correlation and progressive decrease in advanced stages suggests complex immune dynamics, consistent with multicenter investigations demonstrating stromal TIL prognostic value at the invasive front [5].

Our multivariate analysis yielded a parsimonious prognostic model incorporating AI-assessed stromal TIL ratio and LNR. LNR has emerged as a standardized prognostic marker across malignancies compared to pN category, ENE, and metastatic deposit size [3,4,16–18]. The robust model performance, validated through bootstrap resampling, demonstrated how AI-quantified immune parameters combined with

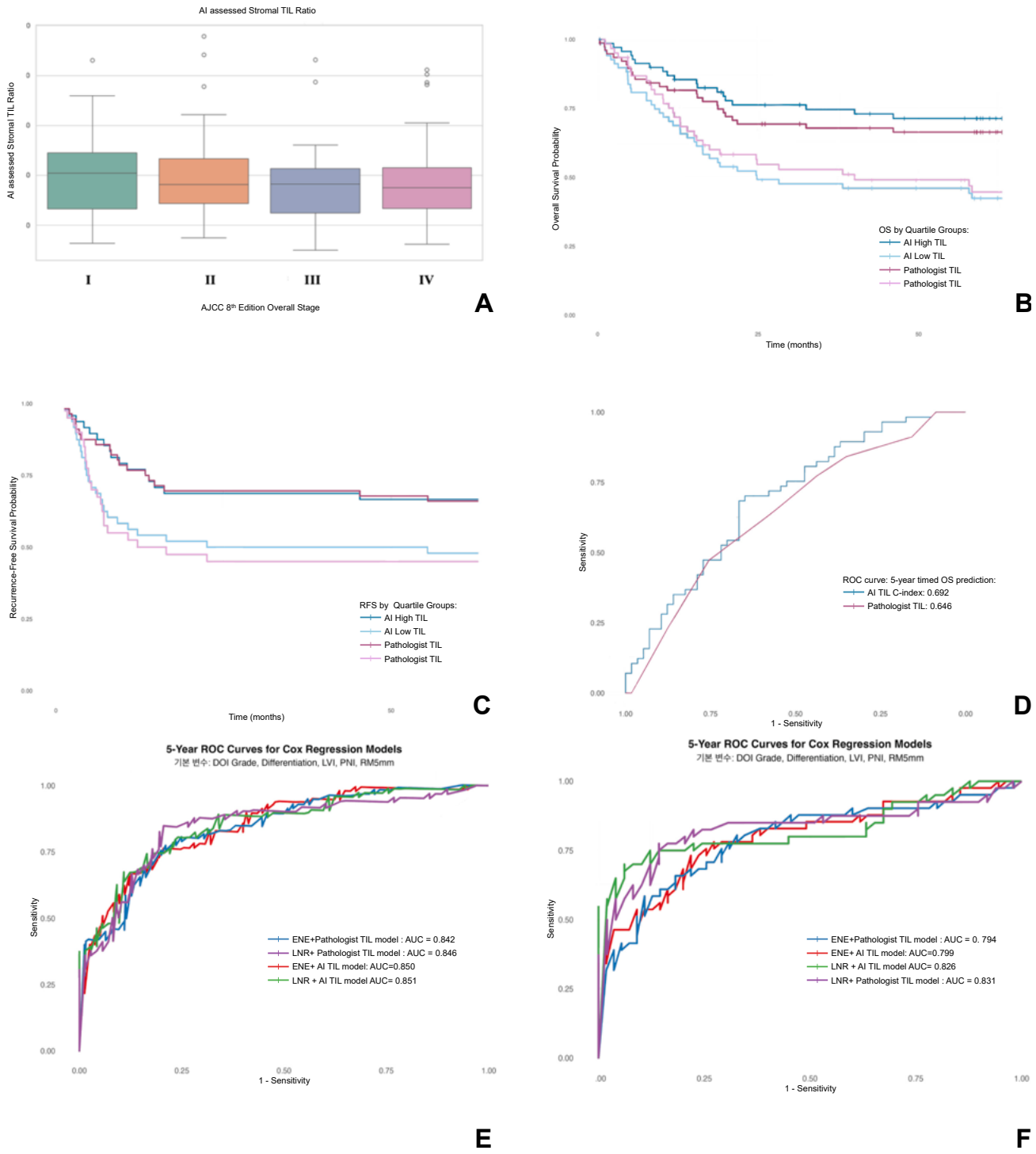


Fig. 3. TIL Distribution Patterns and Prognostic Performance Analysis. (A) Stromal TIL ratio distribution across AJCC 8th Edition stages (I-IV) showing a progressive decrease with advancing stage. Box plots demonstrate considerable variation within each stage, with notable outliers particularly in stages I and IV, indicating heterogeneous immune responses even within same stage groups. (B) Kaplan–Meier survival curves for overall survival stratified by AI-assessed stromal TIL using quartile-based cutoffs. Patients with high TIL infiltration (above median) demonstrated significantly superior survival outcomes compared to low TIL groups, validating the prognostic value of AI-based immune assessment. (C) Kaplan–Meier survival curves for recurrence-free survival stratified by AI-assessed stromal TIL quartile groups. Similar to overall survival, patients with higher TIL levels showed significantly better recurrence-free outcomes, supporting the protective role of immune infiltration. (D) ROC curve comparison between AI-based and manual TIL assessment methods for 5-year survival prediction. AI-based assessment achieved superior discriminative performance (AUC=0.682) compared to manual evaluation (AUC=0.636), demonstrating the enhanced precision of automated quantification. (E) Time-dependent ROC curves at 5 years for overall survival across four multivariate model versions incorporating different combinations of ENE/LNR and AI/manual TIL assessment. The LNR + AI TIL model achieved optimal performance (AUC=0.851), validating the superior prognostic value of this combination. (F) Time-dependent ROC curves at 5 years for recurrence-free survival across the same four model combinations. Consistent with overall survival findings, LNR-based models outperformed ENE-based approaches, with the LNR + AI TIL model achieving excellent discrimination (AUC=0.826). AJCC, American Joint Committee on Cancer; LNR, lymph node ratio; TIL, tumor infiltrating lymphocyte; AI, artificial intelligence; ROC, receiver operating characteristic; ENE, extranodal extension; LVI, lymphovascular invasion; PNI, perineural invasion; DOI, depth of invasion; AUC, area under the ROC curve.

Table 3
Multivariate Cox Regression Analysis of Overall Survival in Patients with Tongue Cancer: Comparison of Best Two Models.

Variables	LNR + AI TIL Model	ENE + Pathologist TIL Model
	HR (95 % CI), P-value	HR (95 % CI), P-value
Tumor Characteristics		
DOI Grade[1]	1.38 (0.97–1.96), 0.07	1.51 (1.07–2.12), 0.02
Differentiation		
Well differentiated	1.0 (reference)	1.0 (reference)
Moderately differentiated	0.92 (0.53–1.59), 0.76	0.96 (0.55–1.68), 0.90
Poorly differentiated	0.97 (0.34–2.78), 0.95	0.93 (0.33–2.64), 0.89
Resection Margin[2]		
<5 mm	2.13 (1.05–4.35), 0.03	2.22 (1.09–4.55), 0.03
≥5 mm	1.0 (reference)	1.0 (reference)
Involved	4.03 (1.34–12.07), 0.83	4.00 (1.33–12.00), 0.71
NA	5.58 (3.04–10.26), 0.53	6.18 (3.27–11.67), 0.49
LVI[3]		
Absent	1.0 (reference)	1.0 (reference)
Present	1.70 (0.95–3.02), 0.07	1.69 (0.95–3.00), 0.07
PNI[4]		
Absent	1.0 (reference)	1.0 (reference)
Present	1.13 (0.63–2.01), 0.69	1.11 (0.62–1.98), 0.72
Nodal Status		
ENE5		
Present	—	3.09 (1.75–5.47), <0.001
NA	—	0.56 (0.07–4.31), 0.58
LNR[6]	4.44 (2.01–9.81), <0.001	—
TIL Assessment[7]		
AI Stromal TIL	0.61 (0.42–0.88), 0.009	—
Pathologist Stromal TIL	—	0.71 (0.50–1.01), 0.05
Model Performance		
Model AUC	0.851	0.842

Multivariate Cox Regression Models
[1]DOI: Depth of invasion grade categorized as ≤5 mm, 5–10 mm, and >10 mm according to AJCC 8th edition guidelines.
[2]Resection margin status: reference category changed to <5 mm to show increased risk with inadequate margins.
[3]LVI: Lymphovascular invasion.
[4]PNI: Perineural invasion.
[5]ENE: Extranodal extension.
[6]LNR: Lymph node ratio (positive lymph nodes/total harvested lymph nodes).
[7]TIL: Tumor-infiltrating lymphocytes assessed via AI-based analysis or pathologist evaluation.
Abbreviations
HR denotes hazard ratio, CI confidence interval, DOI depth of invasion, LVI lymphovascular invasion, PNI perineural invasion, ENE extranodal extension, LNR lymph node ratio, TIL tumor-infiltrating lymphocytes, AUC area under the curve, and NA not applicable.
Notes: OS analysis included n=129 patients with adequate number of harvested lymph nodes. Bold indicates statistically significant results and highest performing model (AUC = 0.851).

clinicopathological factors enhance risk stratification.

Traditionally emphasized parameters such as WPOI did not maintain multivariate significance, contrasting previous reports [12,19]. This discrepancy highlights potential subjective histological assessment limitations and underscores the need for objective, quantifiable parameters, possibly using the help of AI models. AI-quantified parameters combined with established clinical factors may provide more reliable prognostic information in the future.

The limitations of our study include its retrospective nature and single-institutional design. Additionally, 10.1 % of patients had a suboptimal lymph node yield (<10 nodes), which may affect the reliability of the LNR as a prognostic factor according to the AJCC guidelines. However, this reflects real-world clinical practice where patient age, comorbidities, early stage disease, and surgical factors influence the lymph node harvest.

In conclusion, our comprehensive analysis demonstrated that AI-based stromal TIL assessment may be a robust predictive factor for tongue cancer survival, along with the LNR and a very close surgical margin. AI-based TIL assessment performed better than pathologist assessed TILs, as evidenced by multiple survival analyses. Further, although future prospective multicenter studies are needed to establish universally validate the generalizability of the findings, our results demonstrate substantial progress toward more accurate prognostic assessment with the potential to guide personalized treatment strategies and improve patient outcomes through enhanced risk stratification.

Author contributions: JL and MF: Writing – Original Draft, Formal analysis, Software; MF, JL, DJ and DJK: Software, Methodology; JBL: Formal analysis, Methodology; JSS and HJL: Investigation, Validation; SHC, SYN, and YSL: Resources, Investigation; KJC: Conceptualization, Supervision, Writing – Review & Editing, Validation. All authors critically reviewed and approved the final version of the manuscript.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request due to privacy and ethical restrictions.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Claude (Anthropic) to assist with manuscript publication formatting and review. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication

CRediT authorship contribution statement

Jongwon Lee: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ming Fan:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Donghyeok Jo:** Software, Formal analysis. **JungBok Lee:** Methodology, Formal analysis. **Joon Seon Song:** Data curation. **Hee Jin Lee:** Data curation. **Seung-Ho Choi:** Data curation. **Soon Yuhl Nam:** Data curation. **Yoon Se Lee:** Data curation. **Dong-Joo Kim:** Validation, Software. **Kyung-Ja Cho:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the medical staff and pathologists who contributed to data collection and case review. We also acknowledge the patients who participated in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oraloncology.2025.107448.

References

[1] Almangush A, Bello IO, Coletta RD, Mäkitie AA, Mäkinen LK, Kauppila JH, et al. For early-stage oral tongue cancer, depth of invasion and worst pattern of invasion are the strongest pathological predictors for locoregional recurrence and mortality. Virchows Arch. 2015;467:39–46. https://doi.org/10.1007/s00428-015-1758-z.

[2] Bello IO, Soini Y, Salo T. Prognostic evaluation of oral tongue cancer: means, markers and perspectives (II). Oral Oncol. 2010;46:630–5. https://doi.org/10.1016/j.oraloncology.2010.06.008.
[3] Kim SY, Nam SY, Choi SH, Cho KJ, Roh JL. Prognostic value of lymph node density in node-positive patients with oral squamous cell carcinoma. Ann. Surg. Oncol. 2011;18:2310–7. https://doi.org/10.1245/s10434-011-1614-6.
[4] Huang TH, Li KY, Choi WS. Lymph node ratio as prognostic variable in oral squamous cell carcinomas: Systematic review and meta-analysis. Oral Oncol. 2019;89:133–43.
[5] Heikkinen I, Bello IO, Wahab A, Hagström J, Haglund C, Coletta RD, et al. Assessment of tumor-infiltrating lymphocytes predicts the behavior of early-stage oral tongue cancer. Am. J. Surg. Pathol. 2019;43:1392–6. https://doi.org/10.1097/PAS.0000000000001323.
[6] Lundqvist L, Stenlund H, Laurell G, Nylander K. The importance of stromal inflammation in squamous cell carcinoma of the tongue. J. Oral Pathol. Med. 2012;41:379–83. https://doi.org/10.1111/j.1600-0714.2011.01107.x.
[7] World Health Organization, Head and Neck Tumours, fifth ed., beta version ahead of print ed. World Health Organization (IARC), Lyon, France, 2022.
[8] Huang SH, Hwang D, Lockwood G, Goldstein DP, O'Sullivan B. Predictive value of tumor thickness for cervical lymph-node involvement in squamous cell carcinoma of the oral cavity: a meta-analysis of reported studies. Cancer 2009;115:1489–97. https://doi.org/10.1002/cncr.24161.
[9] Li Y, Bai S, Carroll W, Dayan D, Dort JC, Heller K, et al. Validation of the risk model: high-risk classification and tumor pattern of invasion predict outcome for patients with low-stage oral cavity squamous cell carcinoma. Head Neck Pathol. 2013;7:211–23. https://doi.org/10.1007/s12105-012-0412-1.
[10] A. Lugli, R. Kirsch, Y. Ajioka, F. Bosman, G. Cathomas, H. Dawson, H. El Zimaity, J. F. Fléjou, T.P. Hansen, A. Hartmann, S. Kakar, Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016, Mod. Pathol. 30 (2017) 1299–1311. doi: 10.1038/modpathol.2017.46.
[11] Souza da Silva RM, Queiroga EM, Paz AR, Neves FF, Cunha KS, Dias EP. Standardized assessment of the tumor-stroma ratio in colorectal cancer: interobserver validation and reproducibility of a potential prognostic factor. Clin. Pathol. 2021;14:2632010x21989686. https://doi.org/10.1177/2632010X21989686.
[12] Almangush A, Alabi RO, Troiano G, Coletta RD, Salo T, Pirinen M, et al. Clinical significance of tumor-stroma ratio in head and neck cancer: a systematic review and meta-analysis. BMC Cancer 2021;21:480. https://doi.org/10.1186/s12885-021-08222-8.
[13] Brandwein-Gensler M, Smith RV, Wang B, Penner C, Theilken A, Broughel D, et al. Validation of the histologic risk model in a new cohort of patients with head and neck squamous cell carcinoma. Am. J. Surg. Pathol. 2010;34:676–88. https://doi.org/10.1097/PAS.0b013e3181d95c37.
[14] Byrd DR, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, et al. AJCC Cancer Staging Manual. Springer; 2017.
[15] Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nat. Rev. Cancer 2012;12:298–306. https://doi.org/10.1038/nrc3245.
[16] Zhao LY, Li CC, Jia LY, Chen XL, Zhang WH, Chen XZ, et al. Superiority of lymph node ratio-based staging system for prognostic prediction in 2575 patients with gastric cancer: validation analysis in a large single center. Oncotarget 2016;7:51069–81. https://doi.org/10.18632/oncotarget.9714.
[17] Parvathareddy SK, Siraj AK, Qadri Z, Ahmed SO, DeVera F, Al-Sobhi S, et al. Lymph node ratio is superior to AJCC N stage for predicting recurrence in papillary thyroid carcinoma. Endocr. Connect. 2022;11. https://doi.org/10.1530/EC-21-0518.
[18] Talmi YP, Takes RP, Alon EE, Nixon IJ, López F, De Bree R, et al. Prognostic value of lymph node ratio in head and neck squamous cell carcinoma. Head Neck 2018;40:1082–90. https://doi.org/10.1002/hed.25080.
[19] Xu B, Salama AM, Valero C, Yuan A, Khimraj A, Saliba M, et al. The prognostic role of histologic grade, worst pattern of invasion, and tumor budding in early oral tongue squamous cell carcinoma: a comparative study. Virchows Arch. 2021;479:597–606. https://doi.org/10.1007/s00428-021-03063-z.