A transformer-based deep learning model for early prediction of lymph node metastasis in locally advanced gastric cancer after neoadjuvant chemotherapy using pretreatment CT images

Yunlin Zheng,[a,b,i] Bingjiang Qiu,[a,b,c,i] Shunli Liu,[d,i] Ruirui Song,[e,i] Xianqi Yang,[f] Lei Wu,[a,b] Zhihong Chen,[g] Abudouresuli Tuersun,[h] Xiaotang Yang,[e,***] Wei Wang,[f,**] and Zaiyi Liu[a,b,*]

[a]Department of Radiology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, 510080, China
[b]Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangzhou, 510080, China
[c]Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Sciences, Guangzhou, 510080, China
[d]Department of Radiology, The Affiliated Hospital of Qingdao University, 16 Jiangsu Road, Qingdao, Shandong Province, 266000, China
[e]Department of Radiology, Shanxi Cancer Hospital/Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences/Cancer Hospital Affiliated to Shanxi Medical University, Taiyuan, 030013, China
[f]Department of Gastric Surgery, and State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, 510060, China
[g]Institute of Computing Science and Technology, Guangzhou University, Guangzhou, 510006, China
[h]Department of Radiology, The First People's Hospital of Kashi Prefecture, Kashi, 844700, China

Summary
Background Early prediction of lymph node status after neoadjuvant chemotherapy (NAC) facilitates promptly optimization of treatment strategies. This study aimed to develop and validate a deep learning network (DLN) using baseline computed tomography images to predict lymph node metastasis (LNM) after NAC in patients with locally advanced gastric cancer (LAGC).

Methods A total of 1205 LAGC patients were retrospectively recruited from three hospitals between January 2013 and March 2023, constituting a training cohort, an internal validation cohort, and two external validation cohorts. A transformer-based DLN was developed using 3D tumor images to predict LNM after NAC. A clinical model was constructed through multivariate logistic regression analysis as a baseline for subsequent comparisons. The performance of the models was evaluated through discrimination, calibration, and clinical applicability. Furthermore, Kaplan–Meier survival analysis was conducted to assess overall survival (OS) of LAGC patients at two follow-up centers.

Findings The DLN outperformed the clinical model and demonstrated a robust performance for predicting LNM in the training and validation cohorts, with areas under the curve (AUCs) of 0.804 (95% confidence interval [CI], 0.752–0.849), 0.748 (95% CI, 0.660–0.830), 0.788 (95% CI, 0.735–0.835), and 0.766 (95% CI, 0.717–0.814), respectively. Decision curve analysis exhibited a high net clinical benefit of the DLN. Moreover, the DLN was significantly associated with the OS of LAGC patients [Center 1: hazard ratio (HR), 1.789, P < 0.001; Center 2:HR, 1.776, P = 0.013].

Interpretation The transformer-based DLN provides early and effective prediction of LNM and survival outcomes in LAGC patients receiving NAC, with promise to guide individualized therapy. Future prospective multicenter studies are warranted to further validate our model.

*Corresponding author.
**Corresponding author.
***Corresponding author.
    E-mail addresses: liuzaiyi@gdph.org.cn (Z. Liu), wangwei@sysucc.org.cn (W. Wang), yangxt210@126.com (X. Yang).
[i]These authors contributed equally to this work as the first authors.

Research in context

Evidence before this study

We conducted a literature search on Google Scholar and PubMed up to July 30, 2024, utilizing the terms "(advanced gastric cancer OR locally advanced gastric cancer) AND (neoadjuvant chemotherapy) AND (lymph node OR lymph node metastasis) AND (predict OR prediction) AND (radiomics OR deep learning)" without any language restriction. However, previous studies primarily reported lymph node metastasis (LNM) in patients with gastric cancer after radical gastrectomy rather than those after neoadjuvant chemotherapy (NAC), and mainly used two-dimensional tumor sections rather than whole tumors for analysis. These may dissatisfy the requirement of clinical applicability in individualized precision medicine, and further research remains warranted for patients receiving NAC.

Added value of this study

To our best knowledge, this large multicenter study is the first three-dimensional analysis of entire tumors using a transformer-based deep learning network (DLN) for early prediction of LNM after NAC, significantly outperforming the clinical model. Furthermore, visualization of the DLN revealed that intra-tumoral heterogeneity and invasive margins may contribute to predicting LNM. Additionally, the DLN exhibited a strong association with overall survival in the two follow-up centers.

Implications of all the available evidence

The results of our study demonstrated that DLN was able to early predict LNM in patients with locally advanced gastric cancer receiving NAC and had significant prognostic potential, providing a promising method to guide individualized treatment. Future prospective multicenter studies will further validate the predictive performance of our findings in clinical practice.

Introduction

Gastric cancer (GC) is a leading cause of cancer-related deaths worldwide, ranking fifth in both incidence and mortality globally.[1] Most patients are diagnosed with locally advanced gastric cancer (LAGC), and radical gastrectomy with lymphadenectomy is the cornerstone treatment.[2] Unfortunately, the prognosis of LAGC remains poor, with a 5-year survival rate below 40%, mainly due to considerable tumor heterogeneity.[3] Neoadjuvant chemotherapy (NAC) has become the standard therapy for LAGC, improving survival outcomes by downsizing or downstaging the primary tumor, eliminating micro-metastases, and increasing R0 surgical resection rates.[4,5] Regrettably, only about 30% of the NAC recipients exhibit lymph node (LN) regression, and lymph node metastasis (LNM) typically has a profound impact on the therapeutic strategies and prognosis of LAGC patients.[6,7] Therefore, accurate identification of LN involvement after NAC is crucial.

In clinical practice, computed tomography (CT) is the routinely preferred method for diagnosing LNM in LAGC patients receiving NAC.[2,8] Yet, subjective evaluation of LNM poses challenges in terms of diagnostic accuracy and yields unsatisfactory results, especially with a low sensitivity of approximately 57%.[9] Moreover, evaluating LN status after NAC is further complicated by the presence of micro-metastases, inflammatory edema, or fibrosis, leading to a high rate of missed diagnoses of LNM.[10] Several clinicopathological factors have been linked to LNM, such as tumor size, depth of tumor infiltration, histological type, and neutrophil-lymphocyte ratio.[11,12] Nevertheless, these indicators remain controversial, imprecise, and lacking consensus on their applicability.

Deep learning (DL) can automatically learn representative information from original images to comprehensively quantify tumor heterogeneity, and its remarkable performance has garnered significant attention in medical research.[13,14] Numerous studies have showcased the application of DL to predict LNM, determine N-stage, and forecast prognosis for LAGC patients, offering promising clinical advantages.[15–17] Recently, DL employing a transformer framework has emerged as a promising technique for predicting LNM and prognosis.[18,19] Its multi-attention mechanism effectively captures intricate spatial relationships across different levels and locations within the data, focusing on specific details while simultaneously enhancing

model generalization.[20] However, to date, no studies have investigated a DL model utilizing the transformer framework to predict the LN status after NAC in patients with LAGC.

Hence, the main purpose of this study was to develop and validate a transformer framework-based deep learning network (DLN) using baseline CT images for early and accurately prediction of LNM in LAGC patients. Additionally, the study also explored the prognostic value of the DLN.

Methods

Ethics

This multicenter retrospective study obtained ethical approval from the Ethics Committee and Institutional Review Boards of Shanxi Cancer Hospital (NO. KY2023100), the Sun Yat-sen University Cancer Center (NO. B2022-407-01), and the Affiliated Hospital of Qingdao University (NO. QYFY WZLL 28009). Due to the retrospective nature of this study, the requirement for informed consent was waived. The study adhered to the Declaration of Helsinki and its later amendments.

Patients and study design

In this study, we recruited datasets of 1205 LAGC patients (cT2-4NxM0) who underwent NAC across three centers. Detailed information on the enrollment procedure is displayed in Supplementary A1 and Fig. 1. 516 LAGC patients from Center 1 (Shanxi Cancer Hospital) between June 2013 and February 2023 were randomly assigned into a training cohort (TC, n = 361) and an internal validation cohort (IVC, n = 155) at a ratio of 7:3. An external validation cohort 1 (EVC1) of 319 patients was established from Center 2 (the Sun Yat-sen University Cancer Center) between January 2013 and April 2022. The external validation cohort 2 (EVC2), consisting of 370 patients from Center 3 (the Affiliated Hospital of Qingdao University), was enrolled from February 2013 to March 2023. The sample size estimation was calculated by considering the rate of LNM after NAC of approximately 52%–67%,[21,22] as described in Supplementary A1. Furthermore, the follow-up data on LAGC patients from Center 1 (n = 453) and Center 2 (n = 313) were available for survival analysis. The overall study design is depicted in Fig. 2. We performed this study adhering to the Standards for Reporting of Diagnostic Accuracy (STARD) 2015 guidelines.

Neoadjuvant chemotherapy regimens and clinicopathological data collection

All participants received 2–4 cycles of NAC and underwent radical gastrectomy with lymphadenectomy 2–3 weeks after NAC completion. The choices and implementations of all NAC regimens adhered to the latest Chinese clinical guidelines for the diagnosis and treatment of GC[23] (Supplementary A2).

We retrospectively reviewed clinicopathological data including age, sex, body mass index (BMI), carcinoembryonic antigen (CEA) level, carbohydrate antigen (CA) 199 level, tumor location, Borrmann type, Lauren type, tumor differentiation, total cycles of NAC, NAC regimens, clinical T (cT) stage, and clinical N (cN) stage, according to the 8th edition of the American Joint Committee on Cancer TNM staging system.[24] All patients were divided into a LNM group (N+) and a non-LNM group (N0) based on postoperative pathology.

The follow-up data typically included blood tests (CEA and CA19-9 tumor markers) or CT scans, repeated every 3–6 months for the first two years and every 6–12 months for years 3–5. Notably, overall survival (OS) was calculated from the date of surgery to date of death or the last follow-up in October 2023, whichever occurred first.

Image acquisition and preprocessing and 3D semi-automatic tumor segmentation

Contrast-enhanced CT was performed for all patients within two weeks before NAC. Details of the CT scan protocols and image preprocessing are depicted in Supplementary A3 and Table S1. Portal venous-phase CT images were utilized in our analysis due to the excellent contrast between tumor tissue and the adjacent normal gastric wall. To improve the efficiency and accuracy of delineation, we implemented a semi-automated segmentation technique, leveraging our previously developed AILEN.[25] Further details are outlined in Supplementary A4 and Fig. S1. Afterwards, two radiologists (Y.Z. and R.S.) with 6 and 10 years of experience in abdominal CT, respectively, blinded to clinicopathological information, manually adjusted the preliminary segmented volumes of interest (VOIs) in different centers using ITK-SNAP software (version 3.8.0, http://www.itksnap.org/). Subsequently, all VOIs were examined by two senior radiologists (X.Y. and S.L.), and any discrepancies were resolved through consensus.

Development of a deep learning network

We proposed a DLN based on a transformer framework[26] to predict the risk of LNM. The DLN comprised 32-layer transformer blocks for extracting tumor features, alongside convolutional layers with fully connected layers for LNM prediction. The network processes tumor VOIs as input and generates output in the form of probability values for predicting LNM. Details regarding model development and training process are described in Supplementary A5 and Fig. S2.

Construction of a clinical model and nomogram

In the TC, statistically significant clinical predictors were identified through univariate and multivariate logistic regression analyses, with calculated odds ratios
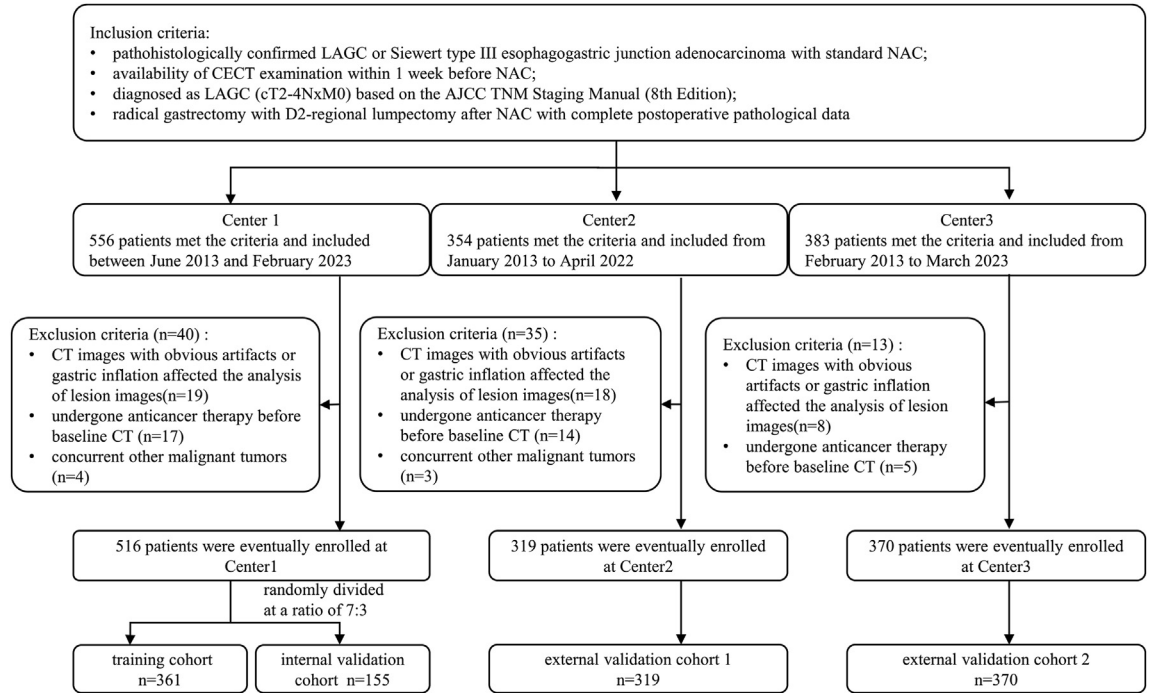
Fig. 1: **Flowchart of the study enrollment process**. Abbreviations: LAGC, locally advanced gastric cancer; NAC, neoadjuvant chemotherapy; CECT, contrast-enhanced computed tomography; AJCC, American Joint Committee on Cancer; CT, computed tomography.

Fig. 2: **The overall flowchart of this study**. Abbreviations: DLN, deep learning network; AILEN, a pre-segmentation model of the tumor volume of interest; ROC, receiver-operator characteristic; DCA, decision curve analysis; K-M, Kaplan–Meier.

(ORs) and 95% confidence intervals (CIs). Subsequently, a clinical model was developed as a baseline for comparison. A nomogram was then constructed by integrating the DLN with independent clinical predictors to further enhance its performanc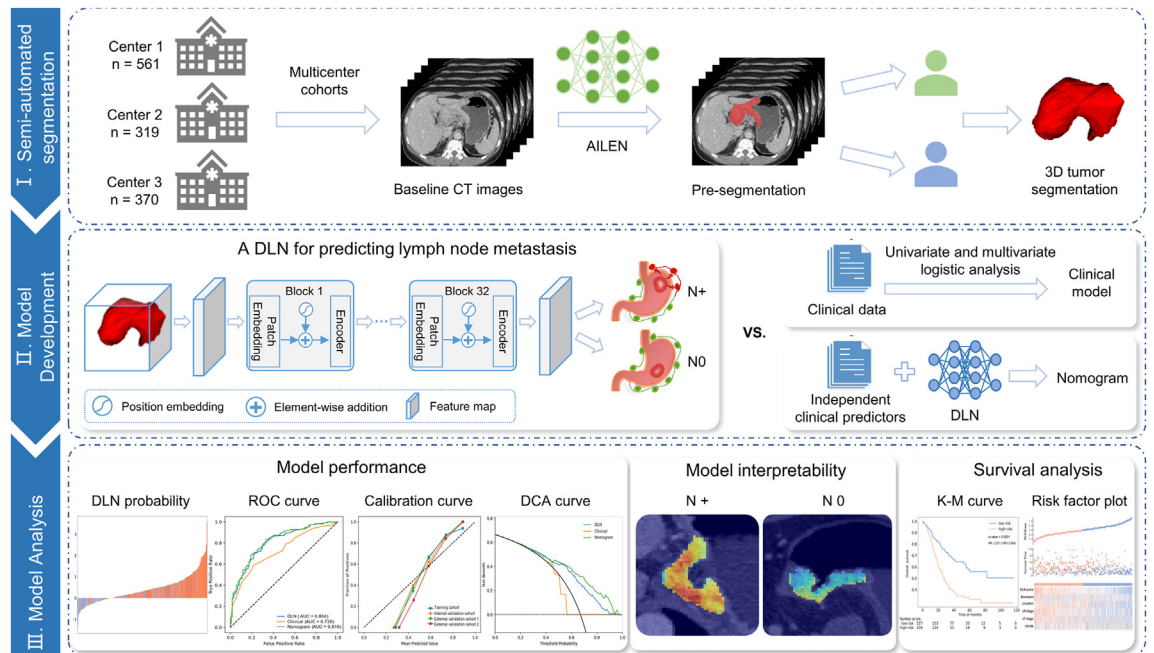e. Net reclassification index (NRI) and integrated discrimination improvement index (IDI) were calculated to compare performance between models. Additionally, we examined the performance of the DLN across all clinicopathological factors.

Model performance assessment and interpretability

The performance of each model in predicting LNM was assessed by the area under the curve (AUC) of the receiver operating characteristic curve (ROC). Additionally, the corresponding accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated. Calibration of all models was evaluated in the training and validation groups through calibration curves derived from 1000 resampling bootstraps and the Hosmer–Lemeshow goodness-of-fit test. Decision curve analysis (DCA) was implemented to estimate the clinical utility of each model by quantifying the net benefit under different threshold probabilities. Notably, gradient-weighted class activation mapping (Grad-CAM)[27] and feature maps were employed to visualize the location and distribution of decision information captured by the DLN in predicting LNM.

Statistics

Statistical analysis was conducted using SPSS software (version 26.0, IBM) and Python software (version 3.8; http://www.python.org) (Supplementary A6). Continuous variables were compared using the Student's t-test or Mann–Whitney U-test, while categorical variables were analyzed using the chi-square test or Fisher's exact test, as appropriate. The AUCs of different models were compared by the DeLong test. The optimal cutoff value was determined via the maximally selected rank statistics approach in the TC, categorizing patients into high-risk or low-risk groups. The prognostic value of the DLN for OS was assessed using Kaplan–Meier survival analysis with the log-rank test. Univariate and multivariate Cox proportional hazards regression analyses were performed to identify independent risk predictors of OS. Adjusted P-values for multiple comparisons were obtained using the Bonferroni correction. A two-tailed P-value <0.05 was considered statistically significant.

Role of the funding source

The funders of this study had no roles in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Clinicopathological characteristics

The baseline clinicopathological characteristics of all 1205 patients with LAGC are outlined in Table 1. 240 (66.5%), 102 (65.8%), 200 (62.7%), and 234 (63.2%) patients had LNM after NAC in the training and validation cohorts, respectively. No significant different was found in the age, BMI, sex, CEA, CA199, cT stage, and NAC regimens between the LNM and non-LNM groups in all validation cohorts (P > 0.05). However, cN stage, tumor location, Borrmann type, Lauren type, differentiation, as well as NAC cycles, showed significant differences in some validation cohorts (P < 0.05). Moreover, most of the above-mentioned indices showed significant differences in TC, and only cN stage and tumor location were retained after the univariate and multivariate logistic regression analyses (Table 2), and were subsequently established as a clinical model. LNM was more likely to be present in a higher cN stage (OR, 11.940; 95% CI, 3.950–36.091) and multisite or whole stomach lesions (OR, 3.587; 95% CI, 1.795–7.165).

Model performances and interpretability

As illustrated in Table 3 and Fig. 3a–d, the DLN displayed robust predictive performance for LNM after NAC in the TC (AUC, 0.804; 95% CI 0.752–0.849) and consistently high efficacy in all validation cohorts, with AUC values of 0.748 (95% CI 0.660–0.830), 0.788 (95% CI 0.735–0.835), and 0.766 (95% CI 0.717–0.814), respectively. Moreover, the performance of DLN, especially in terms of sensitivity, was significantly superior to the clinical model in all cohorts in Table 3 and Fig. 4 (DeLong test, all P < 0.05).

We further integrated DLN into the clinical model for constructing the combined model, and the corresponding nomogram was depicted in Fig. S3. As shown in Fig. 4 and Table 3, the predictive performances of the nomogram were significantly higher than the clinical model (DeLong test, all P < 0.05) and slightly better than the DLN in all cohorts. The calculated IDI and NRI values further proved that the integration of the DLN into the combined model enhanced the accuracy of LNM classification compared to the clinical model (Supplementary Table S2).

Subgroup analyses were conducted to explore the predictive performance of the DLN regarding crucial clinical characteristics, such as cT stage, cN stage, and tumor location. The findings indicated that the DLN also exhibited satisfactory predictive performance in patients with cT4a/4b, cN0, and cN3a/3b across all cohorts, as well as consistent performance across various tumor locations (Supplementary Tables S3–S5). Additionally, differences in the DLN were observed among all clinicopathologic factors, as presented in Fig. S4.

Articles

Variables	Training Cohort (n = 361)			Internal Validation Cohort (IVC) (n = 155)			External Validation Cohort 1 (EVC1) (n = 319)			External Validation Cohort 2 (EVC2) (n = 370)		
	N0 (n = 121)	N+ (n = 240)	P-value	N0 (n = 53)	N+ (n = 102)	P-value	N0 (n = 119)	N+ (n = 200)	P-value	N0 (n = 136)	N+ (n = 234)	P-value
Age(y), median (IQR)	62.0 (54.5, 67.0)	60.0 (53.0, 65.0)	0.048*	60.0 (54.0, 65.0)	61.0 (55.0, 65.3)	0.397	64.0 (56.0, 69.0)	64.5 (53.0, 70.0)	0.888	63.0 (56.0, 68.0)	62.0 (53.8, 67.0)	0.272
BMI, median (IQR)	22.8 (20.3, 24.6)	22.9 (20.8, 25.0)	0.337	23.5 (20.7,25.6)	21.6 (20.3, 24.4)	0.128	21.2 (19.5, 23.9)	22.4 (20.8, 24.3)	0.028*	24.0 (23.9, 24.0)	24.0 (23.4, 24.0)	0.367
Sex (%)			0.047*			0.240			0.367			0.921
Female	17 (14.0%)	55 (22.9%)		13 (24.5%)	17 (16.7%)		39 (32.8%)	56 (28.0%)		32 (23.5%)	54 (23.1%)	
Male	104 (86.0%)	185 (77.1%)		40 (75.5%)	85 (83.3%)		80 (67.2%)	144 (72.0%)		104 (76.5%)	180 (76.9%)	
CEA (ng/mL, %)			0.008*			0.891			0.119			0.077
Normal	100 (82.6%)	167 (69.6%)		39 (73.6%)	74 (72.5%)		95 (79.8%)	144 (72.0%)		113 (83.1%)	176 (75.2%)	
Elevated	21 (17.4%)	73 (30.4%)		14 (26.4%)	28 (27.5%)		24 (20.2%)	56 (28.0%)		23 (16.9%)	58 (24.8%)	
CA199 (U/mL, %)			0.035*			0.292			0.127			0.213
Normal	95 (78.5%)	163 (67.9%)		43 (81.1%)	75 (73.5%)		98 (82.4%)	150 (75.0%)		111 (81.6%)	178 (76.1%)	
Elevated	26 (21.5%)	77 (32.1%)		10 (18.9%)	27 (26.5%)		21 (17.6%)	50 (25.0%)		25 (18.4%)	56 (23.9%)	
cT stage (%)			< 0.001*			0.596			0.270			0.547
cT2	7 (5.8%)	2 (0.8%)		2 (3.8%)	3 (2.9%)		4 (3.4%)	5 (2.5%)		7 (5.1%)	7 (3.0%)	
cT3	66 (54.5%)	98 (40.8%)		28 (52.8%)	46 (45.1%)		54 (45.4%)	74 (37.0%)		46 (33.8%)	77 (32.9%)	
cT4a/4b	48 (39.7%)	140 (58.3%)		23 (43.4%)	53 (52.0%)		61 (51.3%)	121 (60.5%)		83 (61.0%)	150 (64.1%)	
cN stage (%)			< 0.001*			0.015*			0.001*			0.011*
cN0	28 (23.1%)	15 (6.3%)		11 (20.8%)	5 (4.9%)		14 (11.8%)	8 (4.0%)		9 (6.6%)	9 (3.8%)	
cN1	34 (28.1%)	50 (20.8%)		13 (24.5%)	28 (27.5%)		18 (15.1%)	16 (8.0%)		39 (28.7%)	40 (17.1%)	
cN2	52 (43.0%)	106 (44.2%)		20 (37.7%)	40 (39.2%)		61 (51.3%)	97 (48.5%)		59 (43.4%)	107 (45.7%)	
cN3a/3b	7 (5.8%)	69 (28.8%)		9 (17.0%)	29 (28.4%)		26 (21.8%)	79 (39.5%)		29 (21.3%)	78 (33.3%)	
Location (%)			< 0.001*			0.154			0.466			0.026*
Cardia	65 (53.7%)	79 (32.9%)		20 (37.7%)	39 (38.2%)		37 (31.1%)	60 (30.0%)		29 (21.3%)	49 (20.9%)	
Body	19 (15.7%)	35 (14.6%)		11 (20.8%)	10 (9.8%)		14 (11.8%)	27 (13.5%)		27 (19.9%)	78 (33.3%)	
Antrum	22 (18.2%)	49 (20.4%)		7 (13.2%)	25 (24.5%)		35 (29.4%)	45 (22.5%)		76 (55.9%)	98 (41.9%)	
Multiple/whole	15 (12.4%)	77 (32.1%)		15 (28.3%)	28 (27.5%)		33 (27.8%)	68 (34.0%)		4 (2.9%)	9 (3.8%)	
Borrmann (%)			0.002*			0.107			0.002*			< 0.001*
Type I	7 (5.8%)	11 (4.6%)		5 (9.4%)	6 (5.9%)		2 (1.7%)	3 (1.5%)		13 (9.6%)	30 (12.8%)	
Type II	43 (35.5%)	44 (18.3%)		21 (39.6%)	26 (25.5%)		28 (23.5%)	21 (10.5%)		76 (55.9%)	64 (27.4%)	
Type III	68 (56.2%)	170 (70.8%)		27 (50.9%)	66 (64.7%)		84 (70.6%)	150 (75.0%)		44 (32.4%)	108 (46.2%)	
Type IV	3 (2.5%)	15 (6.3%)		0 (0.0%)	4 (3.9%)		5 (4.2%)	26 (13.0%)		3 (2.2%)	32 (13.7%)	
Lauren (%)			0.225			0.011*			0.008*			0.322
Intestinal	36 (29.8%)	58 (24.2%)		22 (41.5%)	20 (19.6%)		57 (47.9%)	61 (30.5%)		28 (20.6%)	55 (23.5%)	
Diffuse	58 (47.9%)	109 (45.4%)		24 (45.3%)	57 (55.9%)		43 (36.1%)	94 (47.0%)		72 (52.9%)	105 (44.9%)	
Mixed	27 (22.3%)	73 (30.4%)		7 (13.2%)	25 (24.5%)		19 (16.0%)	45 (22.5%)		36 (26.5%)	74 (31.6%)	
Differentiation (%)			0.639			0.224			< 0.001*			0.751
Well	3 (2.5%)	8 (3.3%)		2 (3.8%)	4 (3.9%)		0 (0.0%)	0 (0.0%)		2 (1.5%)	2 (0.9%)	
Moderate	36 (29.8%)	61 (25.4%)		20 (37.7%)	25 (24.5%)		41 (34.5%)	34 (17.0%)		26 (19.1%)	40 (17.1%)	
Poor	82 (67.8%)	171 (71.3%)		31 (58.5%)	73 (71.6%)		78 (65.5%)	166 (83.0%)		108 (79.4%)	192 (82.1%)	
NAC regimens (%)			0.502			0.601			0.186			0.178
Doublet-drug	78 (64.5%)	146 (60.8%)		35 (66.0%)	63 (61.8%)		114 (95.8%)	184 (92.0%)		71 (52.2%)	139 (59.4%)	
Triplet-drug	43 (35.5%)	94 (39.2%)		18 (34.0%)	39 (38.2%)		5 (4.2%)	16 (8.0%)		65 (47.8%)	95 (40.6%)	
NAC cycles (%)			0.494			0.281			0.038*			0.594
≤3	84 (69.4%)	158 (65.8%)		38 (71.7%)	81 (79.4%)		110 (92.4%)	169 (84.5%)		67 (49.3%)	122 (52.1%)	
>3	37 (30.6%)	82 (34.2%)		15 (28.3%)	21 (20.6%)		9 (7.6%)	31 (15.5%)		69 (50.7%)	112 (47.9%)	

NOTE: P-values were calculated using the Student's t-test or Mann–Whitney U-test for continuous variables and the chi-square test or Fisher's exact test for categorical variables, as appropriate. * P-value <0.05. Abbreviations: BMI, body mass index; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; NAC, neoadjuvant chemotherapy.

Table 1: **Characteristics of patients with LAGC in the training and validation cohorts.**

Variables	Univariate analysis		Multivariate analysis	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Age(y), median (IQR)	0.978 (0.955–1.002)	0.068		
BMI, median (IQR)	1.046 (0.976–1.121)	0.204		
Sex	0.550 (0.303–0.996)	0.049*	0.672 (0.349–1.293)	0.234
CEA (ng/mL)	2.207 (1.282–3.800)	0.004*	1.617 (0.870–3.006)	0.129
CA199 (U/mL)	1.827 (1.097–3.043)	0.021*	1.366 (0.777–2.401)	0.279
cT stage				
cT2	Ref.		Ref.	
cT3	5.197 (1.047–25.796)	0.044*	4.504 (0.677–29.958)	0.120
cT4a/4b	10.208 (2.050–50.831)	0.005*	4.297 (0.631–29.273)	0.136
cN stage				
cN0	Ref.		Ref.	
cN1	2.745 (1.279–5.891)	0.010*	2.375 (1.042–5.410)	0.040*
cN2	3.805 (1.872–7.736)	<0.001*	3.045 (1.363–6.807)	0.007*
cN3a/3b	18.400 (6.777–49.960)	<0.001*	11.940 (3.950–36.091)	<0.001*
Location				
Cardia	Ref.		Ref.	
Body	1.516 (0.793–2.897)	0.208	1.856 (0.886–3.888)	0.101
Antrum	1.833 (1.005–3.341)	0.048*	1.587 (0.816–3.087)	0.174
Multiple/whole	4.224 (2.220–8.036)	<0.001*	3.587 (1.795–7.165)	<0.001*
Borrmann				
Type I	Ref.			
Type II	0.651 (0.231–1.836)	0.417		
Type III	1.591 (0.592–4.275)	0.357		
Type IV	3.182 (0.668–15.146)	0.146		
Lauren				
Intestinal	Ref.			
Diffuse	1.549 (0.895–2.682)	0.118		
Mixed	1.339 (0.795–2.255)	0.273		
Differentiation				
Well	Ref.			
Moderate	0.945 (0.289–3.091)	0.926		
Poor	1.417 (0.449–4.474)	0.553		
NAC regimens	1.288 (0.825–2.011)	0.265		
NAC cycles	1.134 (0.710–1.809)	0.599		

NOTE: Only variables identified as significant (*Represents $P < 0.05$) in the univariable analyses were entered into the multivariable analysis. Abbreviations: OR, Odds ratio; CI, confidence interval; BMI, body mass index; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; NAC, neoadjuvant chemotherapy.

Table 2: Univariate and multivariable logistic regression analysis for selecting clinical features in the training cohort.

The calibration of the DLN in diverse cohorts was evaluated, and good concordance between predicted and actual probabilities was illustrated by the calibration curves ($P > 0.05$) (Fig. 3e). The results of DCA revealed that the DLN yielded a high net benefit for predicting LNM in all cohorts (Fig. 3f).

The Grad-CAM provided valuable information for predicting LNM, which was deeply mined by the DLN. It visualized the distribution of pixel weights through different colors, highlighting differences in the primary tumor images between the non-LNM and LNM groups. To gain further understanding of the decision given by the DLN, we selected four patients from each group in the TC for observation. DLN-captured tumor images in the LNM group typically exhibited larger activated regions, while fewer regions were activated for the non-LNM group. Notably, the DLN was sensitive to high intra-tumoral heterogeneity and marginal regions (Fig. 5).

Deep learning network connected with survival

We explored the prognostic value of the DLN at two follow-up centers, as detailed characteristics of patients were described in Supplementary Table S6. The median follow-up time was 25.9 months for Center 1 and 22.2 months for Center 2. The optimal DLN score for

Models	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
DLN						
Training	0.804 (0.752–0.849)	0.759 (0.712–0.801)	0.792 (0.738–0.838)	0.694 (0.615–0.769)	0.837 (0.789–0.882)	0.627 (0.542–0.700)
IVC	0.748 (0.660–0.830)	0.703 (0.626–0.774)	0.735 (0.644–0.813)	0.642 (0.509–0.766)	0.798 (0.708–0.879)	0.557 (0.431–0.684)
EVC1	0.788 (0.735–0.835)	0.715 (0.661–0.765)	0.785 (0.722–0.843)	0.597 (0.504–0.685)	0.766 (0.705–0.820)	0.623 (0.531–0.717)
EVC2	0.766 (0.717–0.814)	0.692 (0.649–0.738)	0.692 (0.637–0.755)	0.691 (0.618–0.766)	0.794 (0.742–0.850)	0.566 (0.491–0.639)
Clinical						
Training	0.726 (0.671–0.778)	0.620 (0.568–0.665)	0.521 (0.461–0.581)	0.818 (0.744–0.880)	0.850 (0.785–0.906)	0.463 (0.393–0.545)
IVC	0.603 (0.502–0.700)	0.535 (0.536–0.646)	0.461 (0.364–0.560)	0.679 (0.556–0.800)	0.734 (0.621–0.844)	0.396 (0.297–0.500)
EVC1	0.627 (0.565–0.687)	0.592 (0.536–0.646)	0.670 (0.604–0.732)	0.462 (0.373–0.556)	0.667 (0.612–0.742)	0.455 (0.372–0.545)
EVC2	0.585 (0.522–0.646)	0.557 (0.505–0.608)	0.564 (0.502–0.626)	0.544 (0.459–0.628)	0.680 (0.615–0.747)	0.420 (0.346–0.489)
Nomogram						
Training	0.816 (0.771–0.860)	0.776 (0.731–0.817)	0.829 (0.780–0.872)	0.669 (0.578–0.752)	0.833 (0.785–0.877)	0.664 (0.582–0.741)
IVC	0.756 (0.676–0.832)	0.703 (0.626–0.774)	0.804 (0.728–0.879)	0.509 (0.373–0.642)	0.759 (0.676–0.833)	0.574 (0.422–0.725)
EVC1	0.797 (0.741–0.841)	0.749 (0.705–0.796)	0.885 (0.843–0.925)	0.521 (0.429–0.606)	0.756 (0.670–0.810)	0.729 (0.631–0.816)
EVC2	0.770 (0.713–0.818)	0.714 (0.665–0.760)	0.769 (0.714–0.820)	0.618 (0.524–0.695)	0.776 (0.723–0.829)	0.609 (0.522–0.687)

Abbreviations: DLN, deep learning network; IVC, internal validation cohort; EVC1, external validation cohort 1; EVC2, external validation cohort 2; AUC, area under the curve; PPV, positive prediction value; NPV, negative prediction value; CI, confidence interval.

Table 3: **The performance of models in predicting lymph node metastasis.**

predicting OS in the TC was 0.395, categorizing patients into high-risk and low-risk groups. Kaplan–Meier curves indicated that higher DLN scores were significantly associated with a worse OS [for Center 1, hazard ratio (HR), 2.201 (95% CI, 1.651–2.934), P < 0.0001; for Center 2, HR, 1.829 (95% CI, 1.233, 2.714), P = 0.002]
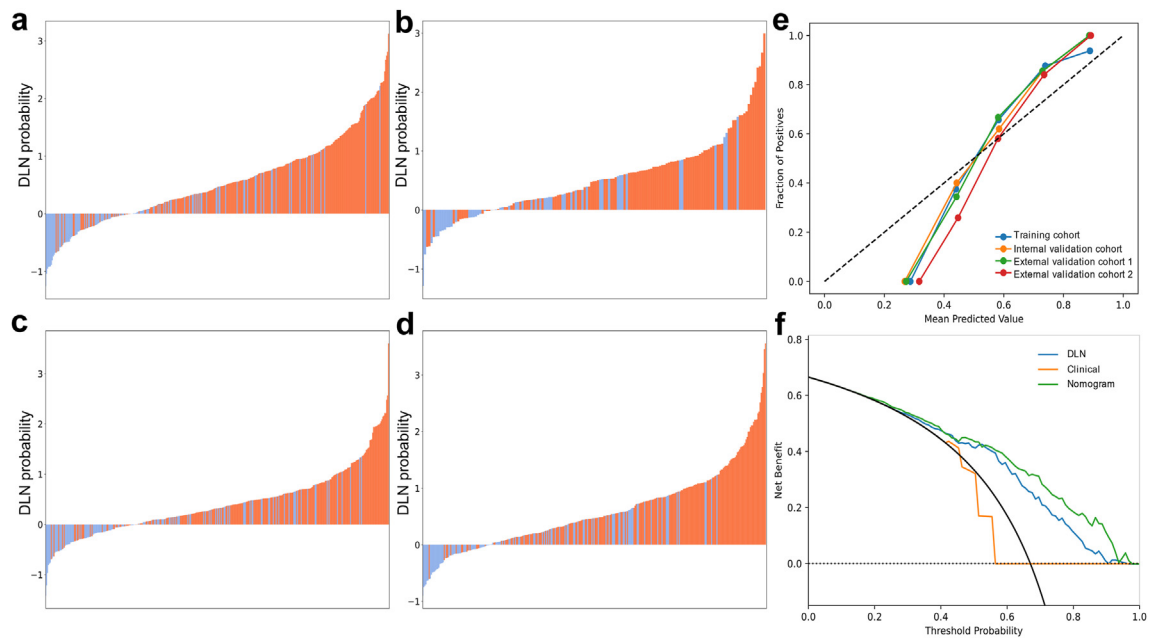

Fig. 3: **Deep learning network (DLN) and its predictive performance**. Waterfall plots of DLN predicted probability in the training cohort (a), internal validation cohort (b), external validation cohort 1 (c), and external validation cohort 2 (d), respectively. The output probability of the DLN was categorized as LNM above 0.561 and as non-LNM below 0.561. The blue below the baseline indicates the correctly predicted non-LNM, the orange above the baseline indicates the correctly predicted LNM, and the cross section is wrongly predicted, with a favorable overall prediction; (e) Calibration curves for DLN in all four cohorts; (f) Decision curve analysis of DLN, clinical, and nomogram models. Abbreviations: DLN, deep learning network; LNM, lymph node metastasis.
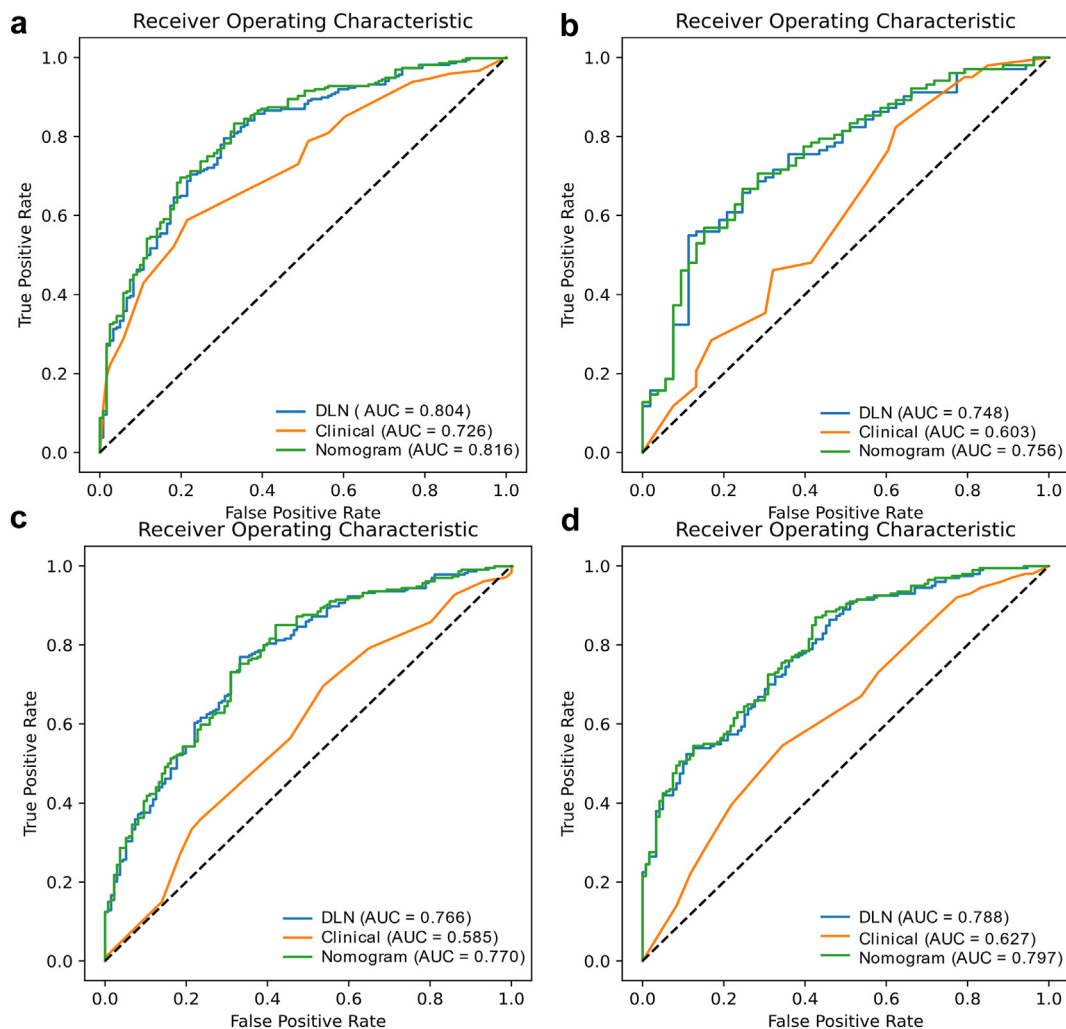
Fig. 4: Predictive performance of lymph node metastasis after neoadjuvant chemotherapy. The receiver operating characteristic (ROC) curves of the DLN, clinical, and nomogram in the training cohort (a), internal validation cohort (b), external validation cohort 1 (c), and external validation cohort 2 (d), respectively. Abbreviations: DLN, deep learning network; ROC, receiver operating characteristic.

(Fig. 6a–b). Furthermore, we visually presented the corresponding DLN scores along with their survival status and survival time for each patient in the two centers (Fig. 6c–d). Table 4 and Fig. 6e–f provided the results of univariate and multivariate Cox regression analyses of predictors of OS in the two centers, demonstrating that the DLN was an independent prognostic factor for OS (for Center 1: HR, 1.789; 95% CI, 1.293–2.476; for Center 2: HR, 1.776; 95% CI, 1.129–2.794; log-rank test, all P < 0.005).

Discussion

Early prediction of LNM in patients with LAGC receiving NAC is essential for refining therapeutic decisions and optimizing patient survival outcomes. In this multicenter study, we developed a DLN based on pre-treatment CT images and further validated its predictive efficacy in independent cohorts. It was discovered in this study that the DLN demonstrated more sensitive and robust performance in predicting LNM compared to the clinical model. Intriguingly, the DLN also manifested a promising prognostic capability for forecasting long-term survival in LAGC patients.

LNM stands as the most common metastasis pattern of LAGC, with an incidence rate surpassing 70%.[2] It is widely recognized that the LNM pathway of GC entails a complex network involving multiple nodes,[28] making the preoperative assessment of LNM by clinicians challenging and controversial. Previous studies have indicated that several clinicopathological biomarkers may contribute to predicting LNM. However, these predictors exhibit limited accuracy and lack consistency
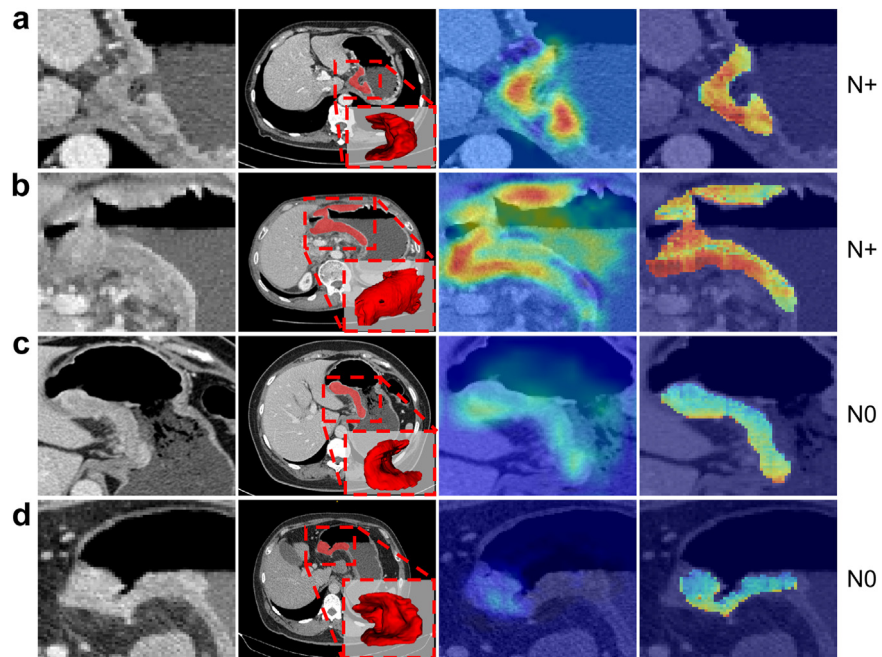
Fig. 5: **Representative images and illustration of the DLN prediction visualization**. From left to right in each row: original CT image; tumor volume of interest; resolution attention map superimposed on CT image; feature map extracted by DLN overlaid on CT image (a–b) Grad-CAM plots highlighted large, multiple regions activated within the tumor of patients with LNM after NAC, indicating that tumor heterogeneity and margin invasion contribute to the prediction of LNM by DLN; (c–d) Grad-CAM plots emphasized very few regions activated within the tumor of patients with non-LNM. Abbreviations: DLN, deep learning network; LNM, lymph node metastasis; CT, computed tomography; Grad-CAM, gradient-weighted class activation mapping; NAC, neoadjuvant chemotherapy.

across studies. In the present study, the distribution of cN stage and lesion location for constructing our clinical model significantly differed in some rather than all cohorts. Indeed, our clinical model exhibited unsatisfactory performance and low sensitivity, which aligns with findings from previous studies.[17,29] The variation in responses and prognoses among patients presenting with similar clinical signs may contribute to the poor performance of clinical models across different patient populations.

To the best of our knowledge, this large multicenter study is the first 3D analysis of entire tumors using the DLN. Previous research predominantly relied on 2D analysis of the maximum section of GC tumors to predict LNM, N-stage, and survival outcomes.[14,15] Given the high heterogeneity of LAGC, relying solely on tumor diameter may not offer comprehensive information about the entire tumor. Moreover, several studies have confirmed that tumor volume more accurately reflects dynamic changes in tumor burden and response to NAC compared to 2D analysis.[30–32] On the basis of the transformer framework, our DLN processed and integrated diverse data information simultaneously through a multi-attention mechanism, efficiently mining high-dimensional image features to comprehensively quantify tumor information. By capturing complex spatial

relationships of tumors, our DLN could mitigate subjective judgmental discrepancies among clinicians, thereby achieving sensitive identification and accurate prediction of LNM after NAC. Unsatisfactorily, our model presented poor specificity, which might be attributed to the imbalance, sophistication, or potential selection bias in real-world multicenter datasets. According to the latest evaluation criteria for LN in clinical practice,[9] it seems that higher sensitivity is more conducive to assisting clinicians in assessing LNM without missing diagnoses. Indeed, our model yielded higher sensitivity to facilitate therapeutic decision-making by clinicians, which was consistent with Tokez et al.[33]

Notably, our DLN demonstrated superior predictive efficacy compared to the clinical model. A prior study involving 523 patients revealed that DL models outperformed radiomics models in preoperatively predicting LNM for LAGC patients, achieving an optimal AUC of 0.597 in the testing cohort.[34] Our DLN showed similarly efficient performance in predicting LNM after NAC. Another prospective study involving 112 patients reported the development of a DL delta radiomics nomogram based on pre- and post-NAC CT with superb predictive performance (AUC, 0.94).[35] However, these optimistic findings were primarily attributed to
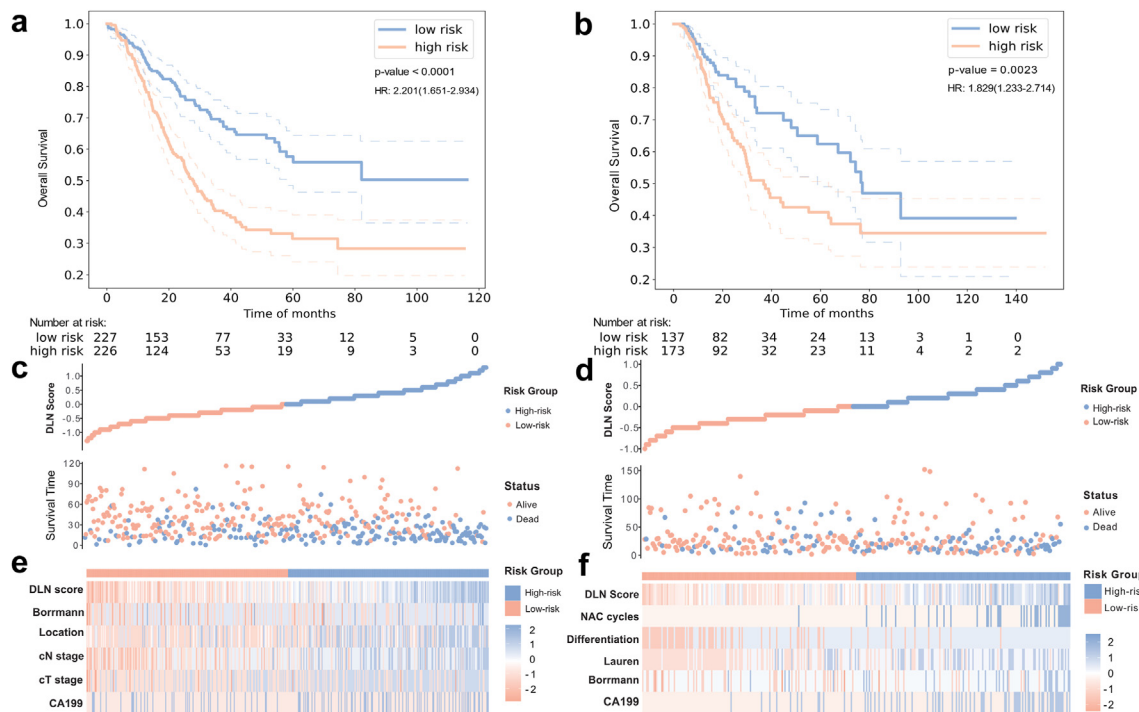
Fig. 6: Correlation of the DLN with overall survival in Center 1 and Center 2. The Kaplan–Meier survival curves of OS in Center 1 (a) and Center 2 (b). Distribution of DLN scores categorized as high-risk (blue) or low-risk (orange) and the corresponding survival status and survival time for each patient in Center 1 (c) and Center 2 (d). Heatmaps of predictors according to multivariate Cox regression analyses between high-risk and low-risk groups of OS in Center 1 (e) and Center 2 (f). Abbreviations: DLN, deep learning network; OS, overall survival; HR, hazard ratio; NAC, neoadjuvant chemotherapy.

incorporating post-NAC tumor information, which hindered early prediction to guide treatments. Additionally, these results await further substantiation due to the relatively small dataset and absence of multicenter cohort confirmation, thereby presenting opportunities for our study.

Grad-CAM serves as a foundational tool for clinicians to interpret the predictions made by the DLN by visualizing the regions within the tumor that are most influential for the model's decision. This study suggests that intra-tumoral heterogeneity and invasive margins may contribute to predicting LNM, consistent with findings by Jin C et al.[36] In contrast to non-LNM cases, the Grad-CAM in our study reveals extensive activated regions within LNM cases, which are larger and redder, particularly at the margins. Tumor heterogeneity indicates activation of multiple and widespread areas, possibly related to microvascular distribution and vigorous tumor cell proliferation.[37] Additionally, the presence of an abundant lymphatic capillary network around peri-gastric areas also increases the possibility of regional LNM.[38] These findings revealed the effective predictive ability of our DLN for LNM.

It is noteworthy that our DLN is significantly associated with OS in LAGC patients. Several studies have proved that LN status after NAC may exert a stronger effect on OS than that of primary tumors.[39–41] In this study, the DLN emerged as an independent prognostic factor for OS through multivariate Cox regression analysis, which was confirmed in two follow-up centers. The DLN score exhibits good stratification ability for OS in LAGC patients. Specifically, considering the unfavorable outcomes associated with high DLN scores, clinicians should promptly consider alternative therapies to avoid missing the optimal operative time for patients and to improve their survival outcomes.

This study still has several limitations. Firstly, being a retrospective multicenter study, there might be potential selection biases and inherent biases. Besides, the predictive capacity of the model across diverse ethnic groups remains unknown. Hence, further prospective studies are necessary to confirm the generalization of our model. Secondly, although Grad-CAM was employed to mitigate the "black box" effect of the DLN, the biological implications of these findings require further elucidation. Therefore, further studies integrating images with genetic or molecular data may unveil more microscopic information and their interrelations. Finally, this study exclusively involved patients receiving NAC regimens recommended by the

Variables	Center 1				Center 2			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
Age(y), median (IQR)	0.996 (0.981–1.011)	0.584			1.010 (0.993–1.027)	0.250		
BMI, median (IQR)	0.963 (0.923–1.004)	0.077			1.035 (0.976–1.098)	0.253		
Sex	0.840 (0.602–1.173)	0.306			1.206 (0.793–1.833)	0.381		
CEA (ng/mL)	1.334 (0.989–1.800)	0.059			1.206 (0.803–1.809)	0.367		
CA199 (U/mL)	1.384 (1.022–1.874)	0.036*	0.973 (0.706–1.340)	0.865	1.664 (1.113–2.489)	0.013*	1.712 (1.136–2.581)	0.010*
cT stage								
cT2	Ref.		Ref.		Ref.			
cT3	2.035 (0.642–6.452)	0.227	1.731 (0.516–5.806)	0.374	3.057 (0.418–22.366)	0.271		
cT4a/4b	3.664 (1.165–11.528)	0.026*	2.187 (0.632–7.575)	0.217	5.922 (0.823–42.598)	0.077		
cN stage								
cN0	Ref.				Ref.			
cN1	1.071 (0.620–1.882)	0.811	0.859 (0.481–1.535)	0.608	0.740 (0.292–1.876)	0.525		
cN2	1.648 (0.993–2.736)	0.053	1.215 (0.701–2.107)	0.488	1.055 (0.499–2.229)	0.889		
cN3a/3b	3.478 (2.053–5.892)	<0.001*	2.170 (1.207–3.904)	0.010*	1.697 (0.800–3.598)	0.168		
Location								
Cardia	Ref.		Ref.		Ref.			
Body	1.116 (0.717–1.738)	0.627	1.072 (0.681–1.689)	0.764	1.103 (0.606–2.007)	0.748		
Antrum	1.313 (0.886–1.946)	0.174	1.148 (0.765–1.722)	0.506	0.710 (0.416–1.209)	0.207		
Multiple/whole	2.123 (1.523–2.960)	<0.001*	1.465 (1.023–2.097)	0.037*	1.235 (0.789–1.934)	0.355		
Borrmann								
Type I	Ref.				Ref.		Ref.	
Type II	0.977 (0.477–1.999)	0.949	0.705 (0.337–1.476)	0.354	0.197 (0.055–0.700)	0.012*	0.207 (0.057–0.755)	0.017*
Type III	1.430 (0.729–2.807)	0.298	0.843 (0.415–1.713)	0.637	0.304 (0.095–0.973)	0.045*	0.258 (0.078–0.859)	0.027*
Type IV	2.431 (1.007–5.870)	0.048*	1.549 (0.616–3.896)	0.352	0.783 (0.231–2.652)	0.695	0.512 (0.140–1.870)	0.311
Lauren								
Intestinal	Ref.				Ref.		Ref.	
Diffuse	1.396 (0.979–1.992)	0.066			2.253 (1.436–3.535)	<0.001*	1.751 (0.960–3.195)	0.068
Mixed	1.443 (0.961–2.168)	0.077			2.191 (1.281–3.748)	0.004*	2.094 (1.084–4.044)	0.028*
Differentiation								
Well	Ref.							
Moderate	0.533 (0.262–1.086)	0.083			Ref.		Ref.	
Poor	0.782 (0.399–1.533)	0.474			2.203 (1.312–3.698)	0.003*	1.107 (0.545–2.941)	0.779
NAC regimens	0.841 (0.636–1.112)	0.224			1.640 (0.900–2.991)	0.106		
NAC cycles	1.039 (0.767–1.406)	0.806			2.172 (1.361–3.466)	0.001*	1.771 (1.067–2.956)	0.027*
DLN score (low vs. high)	2.117 (1.553–2.886)	<0.001*	1.789 (1.293–2.476)	<0.001*	1.996 (1.294–3.078)	0.002*	1.776 (1.129–2.794)	0.013*

Note: Only variables identified as significant (*Represents P < 0.05) in the univariable analyses were entered into the multivariable analysis. Abbreviations: DLN, deep learning network; HR, hazard ratio; CI, confidence interval; BMI, body mass index; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; NAC, neoadjuvant chemotherapy.

Table 4: **Univariate and multivariable Cox regression analysis of predictors of overall survival.**

guidelines, aligning more closely with current clinical practice. However, the impact of NAC regimens and prognostic analyses has not been refined in this study, warranting further exploration.

In conclusion, this study proposed and validated a transformer framework-based DLN derived from baseline CT, providing a potential approach to early prediction of LNM in patients with LAGC receiving NAC. Furthermore, the DLN exhibited strong potential for early prediction of survival outcomes and held promise for guiding individualized therapies. Future prospective analysis in larger cohorts is necessary for validation.

this study is available at https://github.com/qbingjiang/GC_N_prediction, and the pre-segmentation model ("ALIEN") can be obtained from https://github.com/ZHChen-294/ALIEN.

Appendix A. Supplementary data
Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2024.102805.

References
1 Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–263.
2 Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet*. 2020;396(10251):635–648.
3 Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. 2023;73:17–48.
4 Kang YK, Yook JH, Park YK, et al. PRODIGY: a phase III study of neoadjuvant docetaxel, oxaliplatin, and S-1 plus surgery and adjuvant S-1 versus surgery and adjuvant S-1 for resectable advanced gastric cancer. *J Clin Oncol*. 2021;39(26):2903–2913.
5 Zhang X, Liang H, Li Z, et al. Perioperative or postoperative adjuvant oxaliplatin with S-1 versus adjuvant oxaliplatin with capecitabine in patients with locally advanced gastric or gastrooesophageal junction adenocarcinoma undergoing D2 gastrectomy (RESOLVE): an open-label, superiority and non-inferiority, phase 3 randomised controlled trial. *Lancet Oncol*. 2021;22(8):1081–1092.
6 Smyth EC, Fassan M, Cunningham D, et al. Effect of pathologic tumor response and nodal status on survival in the medical research council adjuvant gastric infusional chemotherapy trial. *J Clin Oncol*. 2016;34:2721–2727.
7 Zhao B, Zhang J, Chen X, et al. The retrieval of at least 25 lymph nodes should be essential for advanced gastric cancer patients with lymph node metastasis: a retrospective analysis of single-institution database study design: cohort study. *Int J Surg*. 2017;48:291–299.
8 Ajani JA, D'Amico TA, Bentrem DJ, et al. Gastric cancer, version 2. 2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022;20(2):167–192.
9 Loch FN, Beyer K, Kreis ME, et al. Diagnostic performance of Node Reporting and Data System (Node-RADS) for regional lymph node staging of gastric cancer by CT. *Eur Radiol*. 2024;34(5):3183–3193.
10 Kim HD, Lee JS, Yook JH, et al. Radiological criteria for selecting candidates for neoadjuvant chemotherapy for gastric cancer: an exploratory analysis from the PRODIGY study. *Gastric Cancer*. 2022;25(1):170–179.
11 Fukagawa T, Katai H, Mizusawa J, et al. A prospective multi-institutional validity study to evaluate the accuracy of clinical diagnosis of pathological stage III gastric cancer (JCOG1302A). *Gastric Cancer*. 2018;21(1):68–73.
12 Zhang LX, Wei ZJ, Xu AM, Zang JH. Can the neutrophil-lymphocyte ratio and platelet-lymphocyte ratio be beneficial in predicting lymph node metastasis and promising prognostic markers of gastric cancer patients? Tumor maker retrospective study. *Int J Surg*. 2018;56:320–327.
13 Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*. 2019;25(7):1054–1056.
14 Jiang Y, Zhang Z, Yuan Q, et al. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study. *Lancet Digit Health*. 2022;4(5):e340–e350.
15 Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol*. 2020;31(7):912–920.
16 Wang X, Chen Y, Gao Y, et al. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun*. 2021;12(1):1637.
17 Cui Y, Zhang J, Li Z, et al. A CT-based deep learning radiomics nomogram for predicting the response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer: a multi-center cohort study. *eclinicalMedicine*. 2022;46:101348.
18 Wang Z, Yu L, Ding X, Liao X, Wang L. Lymph node metastasis prediction from whole slide images with transformer-guided multiinstance learning and knowledge transfer. *IEEE Trans Med Imaging*. 2022;41(10):2777–2787.
19 Ma X, Xia L, Chen J, Wan W, Zhou W. Development and validation of a deep learning signature for predicting lymph node metastasis in lung adenocarcinoma: comparison with radiomics signature and clinical-semantic model. *Eur Radiol*. 2023;33(3):1949–1962.
20 Zheng H, Fu J, Mei T, Luo J. *Learning multi-attention convolutional neural network for fine-grained image recognition*. Proceedings of the IEEE international conference on computer vision; 2017:5209–5217.
21 Lin JX, Tang YH, Lin GJ, et al. Association of adjuvant chemotherapy with overall survival among patients with locally advanced gastric cancer after neoadjuvant chemotherapy. *JAMA Netw Open*. 2022;5(4):e225557.
22 Cats A, Jansen EPM, van Grieken NCT, et al. Chemotherapy versus chemoradiotherapy after surgery and preoperative chemotherapy for resectable gastric cancer (CRITICS): an international, open-label, randomised phase 3 trial. *Lancet Oncol*. 2018;19(5):616–628.
23 Wang FH, Zhang XT, Tang L, et al. The Chinese Society of Clinical Oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer, 2023. *Cancer Commun*. 2024;44(1):127–172.
24 Amin MB, Edge SB, Greene FL, Brierley JD. *AJCC cancer staging manual*. vol. 82017;vol. 8. New York: Springer; 2017.
25 Chen Z, Yao L, Cui Y, et al. ALIEN: attention-guided cross-resolution collaborative network for 3D gastric cancer segmentation in CT images. *Biomed Signal Process Control*. 2024;96:106500.
26 Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30.
27 Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis*. 2019;128(2):336–359.
28 de Jongh C, Triemstra L, van der Veen A, et al. Pattern of lymph node metastases in gastric cancer: a side-study of the multicenter LOGICA-trial. *Gastric Cancer*. 2022;25(6):1060–1072.
29 Sun Z, Jiang Y, Chen C, et al. Radiomics signature based on computed tomography images for the preoperative prediction of lymph node metastasis at individual stations in gastric cancer: a multicenter study. *Radiother Oncol*. 2021;165:179–190.
30 Li R, Chen TW, Hu J, et al. Tumor volume of resectable adenocarcinoma of the esophagogastric junction at multidetector CT: association with regional lymph node metastasis and N stage. *Radiology*. 2013;269(1):130–138.
31 Steger S, Franco F, Sverzellati N, Chiari G, Colomer R. 3D assessment of lymph nodes vs. RECIST 1.1. *Acad Radiol*. 2011;18(3):391–394.
32 Hallinan JT, Venkatesh SK, Peter L, Makmur A, Yong WP, So JB. CT volumetry for gastric carcinoma: association with TNM stage. *Eur Radiol*. 2014;24(12):3105–3114.
33 Tokez S, Koekelkoren FHJ, Baatenburg de Jong RJ, et al. Assessment of the diagnostic accuracy of baseline clinical examination and ultrasonographic imaging for the detection of lymph node metastasis in patients with high-risk cutaneous squamous cell carcinoma of the head and neck. *JAMA Dermatol*. 2022;158(2):151–159.
34 Zhang AQ, Zhao HP, Li F, Liang P, Gao JB, Cheng M. Computed tomography-based deep-learning prediction of lymph node metastasis risk in locally advanced gastric cancer. *Front Oncol*. 2022;12:969707.
35 Zhong H, Wang T, Hou M, et al. Deep learning radiomics nomogram based on enhanced CT to predict the response of

metastatic lymph nodes to neoadjuvant chemotherapy in locally advanced gastric cancer. *Ann Surg Oncol*. 2024;31(1):421–432.

36 Jin C, Jiang Y, Yu H, et al. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. *Br J Surg*. 2021;108(5):542–549.

37 Hu C, Chen W, Li F, et al. Deep learning radio-clinical signatures for predicting neoadjuvant chemotherapy response and prognosis from pretreatment CT images of locally advanced gastric cancer patients. *Int J Surg*. 2023;109(7):1980–1992.

38 Karaman S, Detmar M. Mechanisms of lymphatic metastasis. *J Clin Invest*. 2014;124(3):922–928.

39 Moore JL, Green M, Santaolalla A, et al. Pathologic lymph node regression after neoadjuvant chemotherapy predicts recurrence and survival in esophageal adenocarcinoma: a multicenter study in the United Kingdom. *J Clin Oncol*. 2023;41(28):4522–4534.

40 Smyth EC, Nyamundanda G, Cunningham D, et al. A seven-Gene Signature assay improves prognostic risk stratification of perioperative chemotherapy treated gastroesophageal cancer patients from the MAGIC trial. *Ann Oncol*. 2018;29(12):2356–2362.

41 Fujitani K, Nakamura K, Mizusawa J, et al. Posttherapy topographical nodal status, ypN-site, predicts survival of patients who received neoadjuvant chemotherapy followed by curative surgical resection for non-type 4 locally advanced gastric cancer: supplementary analysis of JCOG1004-A. *Gastric Cancer*. 2021;24(1):197–204.