Gastric Cancer (2022) 25:1050–1059
https://doi.org/10.1007/s10120-022-01328-3

ORIGINAL ARTICLE

Deep learning predicts resistance to neoadjuvant chemotherapy for locally advanced gastric cancer: a multicenter study

Jiayi Zhang[1,3] · Yanfen Cui[2,4] · Kaikai Wei[5] · Zhenhui Li[6] · Dandan Li[2] · Ruirui Song[2] · Jialiang Ren[7] · Xin Gao[1,2,3] · Xiaotang Yang[2]

Received: 19 May 2022 / Accepted: 21 July 2022 / Published online: 6 August 2022

Abstract

Background Accurate pre-treatment prediction of neoadjuvant chemotherapy (NACT) resistance in patients with locally advanced gastric cancer (LAGC) is essential for timely surgeries and optimized treatments. We aim to evaluate the effectiveness of deep learning (DL) on computed tomography (CT) images in predicting NACT resistance in LAGC patients.

Methods A total of 633 LAGC patients receiving NACT from three hospitals were included in this retrospective study. The training and internal validation cohorts were randomly selected from center 1, comprising 242 and 104 patients, respectively. The external validation cohort 1 comprised 128 patients from center 2, and the external validation cohort 2 comprised 159 patients from center 3. First, a DL model was developed using ResNet-50 to predict NACT resistance in LAGC patients, and the gradient-weighted class activation mapping (Grad-CAM) was assessed for visualization. Then, an integrated model was constructed by combing the DL signature and clinical characteristics. Finally, the performance was tested in internal and external validation cohorts using area under the receiver operating characteristic (ROC) curves (AUC).

Results The DL model achieved AUCs of 0.808 (95% CI 0.724–0.893), 0.755 (95% CI 0.660–0.850), and 0.752 (95% CI 0.678–0.825) in validation cohorts, respectively, which were higher than those of the clinical model. Furthermore, the integrated model performed significantly better than the clinical model ($P < 0.05$).

Conclusions A CT-based model using DL showed promising performance for predicting NACT resistance in LAGC patients, which could provide valuable information in terms of individualized treatment.

Keywords Locally advanced gastric cancer · Neoadjuvant chemotherapy · Pre-treatment computed tomography · Deep learning

Jiayi Zhang, Yanfen Cui, Kaikai Wei, and Zhenhui Li have contributed equally to this work.

✉ Xiaotang Yang
yangxt210@126.com

1 Division of Life Sciences and Medicine, School of Biomedical Engineering (Suzhou), University of Science and Technology of China, Suzhou 215163, Jiangsu, China

2 Department of Radiology, Shanxi Province Cancer Hospital, Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences, Cancer Hospital, Affiliated to Shanxi Medical University, Taiyuan 030013, Shanxi, China

3 Medical Imaging Department, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, Jiangsu, China

4 Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangzhou 510080, Guangdong, China

5 Department of Radiology, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510655, Guangdong, China

6 Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming 650118, Yunnan, China

7 GE Healthcare China, Beijing 100176, China

Introduction

Gastric cancer (GC) is the second leading cause of cancer-related deaths worldwide [1], and 80–90% of GC patients are in the advanced stages at their first visits [2, 3]. Surgery is the main treatment for locally advanced gastric cancer (LAGC); however, the 5 year survival rate is only 20–30% [3]. In recent years, neoadjuvant chemotherapy (NACT) has been adopted to improve the R0 resection and prognosis of LAGC patients [4]. Nevertheless, approximate 30% of LAGC patients tend to be resistant to neoadjuvant treatment [5], which leads to an increased risk of disease progression and unnecessary toxicity for those patients. Meanwhile, histopathological examination after surgery continues to be used as the reference standard for defining the therapeutic response; yet, this approach is invasive and hysteretic [6]. Hence, identifying patients with NACT resistance before treatment is of critical importance, as it may assure that they receive timely surgeries and optimized treatments.

Clinically, computed tomography (CT) is the preferred imaging examination used to evaluate the response to NACT among LAGC patients, usually by some conventional morphologic metrics, such as the response evaluation criteria in solid tumors (RECIST) and CT volumetry, or some quantitative imaging parameters. However, these analyses tend to rely on imaging features extracted by naked eyes, where some microcosmic imaging features relevant to a patient outcome might be ignored [7]. With the development of computer-aided analysis, radiomics, which can convert medical images into mineable high-throughput quantitative features, and characterize tumors and their microenvironment [8], has been successfully applied to predict NACT response in LAGC patients [6, 9–11]. However, radiomics requires precise tumor delineation, which is the precondition for feature extraction and selection, and model building. Moreover, tumor delineations are subject to physician's experience, which conversely affects prediction accuracy and also hinders the application of radiomics in clinical practice.

Recently, deep learning (DL) has emerged as a powerful approach, which greatly reduced the difficulty of feature extraction by automatically learning the representation of key disease features directly from medical images [12–14]. In terms of GC, DL has been successfully applied in occult peritoneal metastasis identification [15] and lymph node metastases prediction [16, 17]. Still, to the best of our knowledge, DL has not yet been adopted for the prediction of resistance to NACT in LAGC patients. Therefore, we aimed to develop an end-to-end DL model based on pre-treatment CT to predict NACT resistance in patients with LAGC and externally validate the predictive ability in two independent cohorts.

Methods

Patients selection

This retrospective multicenter study was approved by the Institution Ethics Review Boards of Shanxi Cancer Hospital (No. 202223), The Sixth Affiliated Hospital of Sun Yat-sen University (No. E2021088), and Yunnan Cancer Hospital (No. KY202034), and the need for informed consent was waived due to the retrospective nature of this study.

This retrospective multicenter study recruited the patients with LAGC who received NACT in the three centers. The inclusion criteria were as follows: (1) histologically (biopsy-) confirmed gastric adenocarcinoma; (2) diagnosed with non-metastatic locally advanced stage (cT2-4N0/+M0) determined by pre-treatment CT examination or laparoscopic laparotomy based on American Joint Committee on Cancer (AJCC) TNM Staging Manual (8th Edition) [18]; (3) underwent standard baseline contrast-enhanced CT before treatment; (4) received NACT followed by gastrectomy with lymph node dissection, after which tumor response was confirmed by postoperative pathological examination; and (5) clinicopathological characteristics were available. The exclusion criteria were the following: (1) unidentified primary tumor on CT or poor CT image quality to perform measurements; (2) suffered synchronous other malignant neoplasms; (3) received anticancer therapy before the baseline CT scans; and (4) with incomplete clinicopathological data.

Baseline staging evaluations included CT, ultrasound endoscopy, and/or positron emission tomography (PET)/CT, and neoadjuvant treatment was jointly decided by the multidisciplinary discussion with radiologists, surgeons, and oncologists for the patients with LAGC (cT2-4N0/+M0) in accordance with national guidelines. Finally, 633 LAGC patients were included in this study. Among them, 346 patients from center 1 (Shanxi Cancer Hospital) were divided into a training cohort ($n=242$) and an internal validation cohort ($n=104$) at a ratio of 7:3 randomly between January 2017 and November 2020. Meanwhile, patients from center 2 (The Sixth Affiliated Hospital of Sun Yat-sen University) were grouped as external validation cohort 1 ($n=128$), and patients from center 3 (Yunnan Cancer Hospital) were grouped as external validation cohort 2 ($n=159$) between January 2016 and August 2020.

The details of the CT examination are described in the Supplementary File. All clinical characteristics before NACT were retrieved, including age, body mass index (BMI), gender, differentiation status, carcinoembryonic antigen (CEA) level, carbohydrate antigen (CA) 199 level, tumor location, clinical T stage, and clinical N stage.

Pathological evaluation

According to national guidelines, all enrolled patients with LAGC received NACT. Chemotherapy regimens included XELOX (130 mg/m^2 oxaliplatin as a 2 h infusion on day 1, followed by 1000 mg/m^2 capecitabine twice daily for 14 consecutive days), SOX (S-1 was administered orally 80 mg/m^2/day on days 1–14, while 130 mg/m^2 oxaliplatin was given intravenously 130 mg/m^2 on day 1), and FOLFOX (130 mg/m^2 oxaliplatin as a 2 h infusion, 400 mg/m^2 leucovorin, and a bolus of 400 mg/m^2 5-fluorouracil on day 1, followed by a 46-h infusion of 2400 mg/m^2 5-fluorouracil). The treatment was repeated every 3 weeks. All patients received 2–4 cycles of NACT before surgery, with adjustment to dosage or cycles based on effectiveness and patient tolerability, and gastrectomy was performed within 2 weeks after NACT. The response to NACT was assessed by the consensus of two expert gastrointestinal pathologists. Tumor regression grade (TRG) defined according to the latest National Comprehensive Cancer Network (NCCN) guideline (version 4, 2021) [19] was adopted to evaluate NACT response. TRG 0 was defined as no viable cancer cell in the primary lesion or lymph nodes, and the presence of single cells or rare small groups of cancer cells was classified as TRG 1. The evident tumor regression, but more than single cells or a rare small group of cells, was defined as TRG 2; TRG 3 was defined as no evident tumor regression with extensive residual cancer. In this study, patients with TRG 0–2 were considered as NACT responsive and patients with TRG 3 as NACT resistant [20–22].

Deep learning model building

The overall workflow for building a DL model for predicting NACT resistance is shown in Fig. 1. Portal venous-phase CT images were retrieved for further evaluation. The region of interest (ROI) was manually delineated on the center slice of CT images with the largest tumor using the Medical Imaging Interaction Toolkit (MITK) software (version 2013.12.0; http://www.mitk.org/) by a radiologist with 10 years of clinical experience in abdominal CT interpretation and was then confirmed by a radiologist with 31 years of experience in abdominal CT interpretation.

The overall preprocessing procedure was as follows: a 3D resampling process was done to regularize CT images to $1 \times 1 \times 1$ mm per voxel, and for a high contrast view of the stomach and surrounding tissues, a window of [− 350, 450] HU was truncated. To ensure that the DL network's attention focused on the most relevant part of the CT images, the slice with the largest tumor ROI and its delineation mask, as well as its ten adjacent slices (five superior and five inferior), were selected to compose a 12-layer image patch [16, 23], which means that the image patch contained 11 CT slices and one delineation mask. Meanwhile, a square covering the tumor ROI and located at the centroid of tumor ROI was used to extract the square region slice-by-slice, generating a 12-layer square image patch, which means that size of the image patch was determined by the tumor size. All patches were processed by z-score standardization, which consisted of subtracting the mean intensity and division by the standard deviation of intensity.

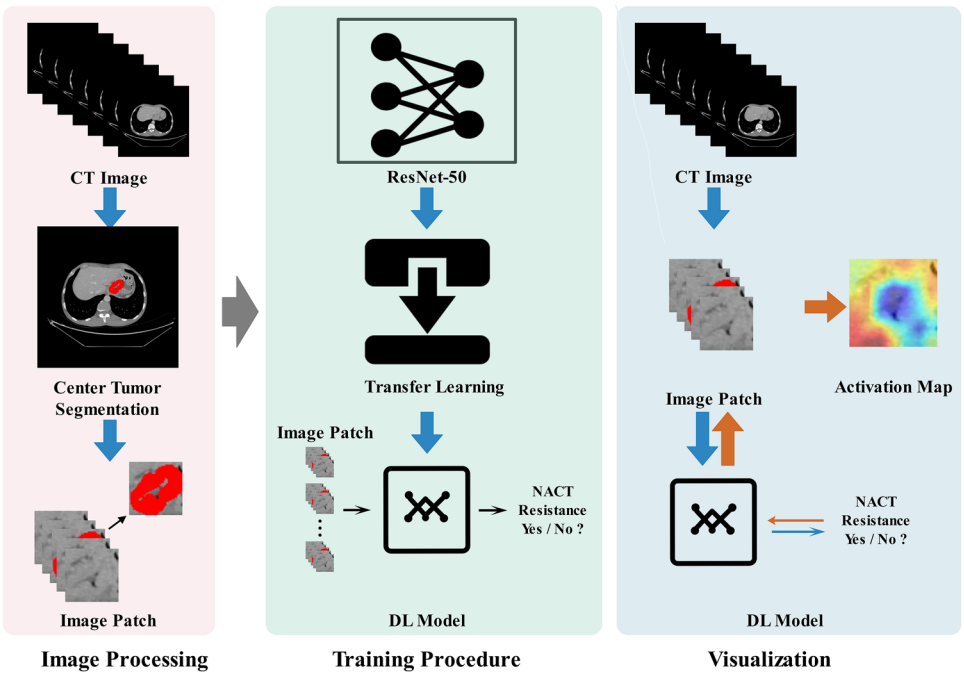

Fig. 1 Overall workflow for building a deep learning (DL) model for predicting neoadjuvant chemotherapy (NACT) resistance

ResNet-50 architecture [24], a state-of-the-art convolutional neural network (CNN), was used as the main backbone of the DL in the present study. This network demonstrates significant improvements over traditional CNNs on highly competitive object recognition benchmark tasks while requiring less computational cost and having fewer parameters, thus conferring the model a smaller size and easier accessibility for application. The training of DL models is computationally expensive and requires tremendous images because of millions of learnable parameters to be estimated. To address the lack of data, we employed data augmentation [25], including image reflection along with the patient's anterior/posterior or left/right directions, and image rotation with 90°, 180°, and 270°, and transfer learning [26]. The detailed description of data augmentation and transfer learning is documented in the Supplementary File. The DL network was implemented using PyTorch [27] in Python. The pre-trained ResNet-50 model is available online, and the training details of the DL model are described in the Supplementary File.

For an intuitive understanding of the mechanisms of the DL model, the strategy of guided gradient-weighted class activation mapping (Grad-CAM) was used to generate heat maps that could give a coarse location of the image area relevant to tumor response [28]. The technical details of heat map generation are documented in the Supplementary File.

Integrated and clinical models development

To demonstrate the incremental value of the DL model for individualized prediction of tumor resistance to NACT, an integrated model was constructed by combing the DL signature and clinical factors. The DL signature, which presented the probability of NACT resistance and whose type was float, and all mentioned clinical candidate predictors were evaluated using univariable analysis, where variables with $P < 0.05$ were kept. Then, the multivariable stepwise logistic regression was used to develop the integrated model using the remaining characteristics as input in the training cohort. The optimal combinations of the characteristics were determined using the akaike information criterion (AIC) as the stopping rule in the backward stepwise procedure [29]. A clinical model was built by including only clinical characteristics and used for comparison.

Performance evaluation

The DL, clinical, and integrated models built on the training cohort were independently validated using the internal and two external validation cohorts, while their prediction performances were evaluated using receiver operating characteristic (ROC) curve analysis and quantified by area under curve (AUC) [30]. The DeLong test [31] was used for the comparison of the AUCs. The net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were also calculated to quantify the relative improvements in prediction accuracy.

Statistical analysis

To compare the differences in the clinical characteristics between patients in different groups or cohorts, the Mann–Whitney *U* test was used for numerical variables, and the Chi-square test was used for categorical variables. The R package "pROC" was used for ROC analysis. In addition, the predictive accuracy, sensitivity, specificity, positive predict value (PPV) and negative predict value (NPV) were measured based on the optimal cutoff value, which was obtained by used the maximum Youden index method. The multivariable stepwise logistic regression was used to develop the integrated model. All statistical tests were performed using R software (4.1.0), and a two-sided $P < 0.05$ was considered statistically significant.

Results

Baseline characteristics

The baseline characteristics of all 633 patients included in the four cohorts are summarized in Table 1 and S2. The tumor resistance rates in the four cohorts were 48.8%, 49.0%, 23.4%, and 44.0%, respectively, indicating that patients were balanced for the efficacy of NACT ($P > 0.05$). Moreover, no significant difference was found between the tumor response and resistance groups in terms of the age, BMI, gender, differentiation status, CEA, and CA199 level in all four cohorts ($P > 0.05$).

Prediction performance of deep learning model

After training for 100 epochs with a loss value of 0.3 in the training cohort, which reached a plateau, we proposed an end-to-end DL model for predicting tumor resistance to NACT in patients with LAGC based on a pre-treatment CT image. The trained DL model performed well on the prediction task in all the validation cohorts. ROC curve analysis showed that the AUC was 0.808 (95% CI 0.724–0.893) in the internal validation cohort, with the sensitivity and specificity of 80.4% and 75.5%, respectively (Table 2 and S3). Similarly, the predictive performances were also confirmed in the two external validation cohorts, with AUCs of 0.755 (95% CI 0.660–0.850) and 0.752 (95% CI 0.678–0.825), respectively.

Figure 2 shows an example of CT images with superimposed heat maps. The red part of the heat map shows the key

Table 1 Clinical characteristics of patients on the training and validation cohorts

Clinical characteristics	Training cohort			Internal validation cohort			External validation cohort 1			External validation cohort 2		
	Response ($n=124$)	Resistance ($n=118$)	*P* value	Response ($n=53$)	Resistance ($n=51$)	*P* value	Response ($n=98$)	Resistance ($n=30$)	*P* value	Response ($n=89$)	Resistance ($n=70$)	*P* value
Age (years)	61 (28–89)	60 (27–75)	0.646	62 (35–75)	62 (43–75)	0.569	60 (23–75)	57 (34–77)	0.220	57 (29–77)	54 (29–76)	0.160
BMI	23.0 (14.8–31.1)	22.5 (15.9–32.5)	0.346	22.2 (17.0–29.7)	22.6 (15.4–29.7)	0.692	21.8 (13.4–28.2)	22.5 (12.4–30.0)	0.910	21.7 (15.6–30.8)	21.6 (14.5–30.1)	0.648
Gender												
Female	24 (19.4)	24 (20.3)	0.976	8 (15.1)	10 (19.6)		25 (25.5)	3 (10.0)	0.122	20 (22.5)	23 (32.9)	0.199
Male	100 (80.6)	94 (79.7)		45 (84.9)	41 (80.4)		73 (74.5)	27 (90.0)		69 (77.5)	47 (67.1)	
Differentiation status												
Well	7 (5.6)	3 (2.5)	0.456	2 (3.8)	0(0.0)	0.028	3(3.1)	3(10.0)	0.051	0(0.0)	2(2.8)	0.305
Moderately	28 (22.6)	31 (26.3)		20 (37.7)	10(19.6)		31(31.6)	4(13.3)		15(16.9)	10(14.3)	
Poorly	89 (71.8)	84 (71.2)		31(58.5)	41 (80.4)		64 (65.3)	23 (76.7)		74 (83.1)	58 (82.9)	
CEA												
≤5	91 (73.4)	80 (67.8)	0.396	39 (73.6)	30 (58.8)	0.246	72 (73.5)	21 (70.0)	0.531	62 (69.7)	48 (68.6)	0.514
>5	33 (26.6)	38 (32.2)		14 (26.4)	21 (41.2)		26 (26.5)	9 (30.0)		27 (30.3)	22 (31.4)	
CA199												
≤20	88 (71.0)	67 (56.8)	0.198	35 (66.0)	28 (54.9)	0.428	68 (69.4)	21 (70.0)	0.909	55 (61.8)	41 (58.6)	0.299
>20	36 (29.0)	51 (43.2)		18 (34.0)	23 (45.1)		30 (30.6)	9 (30.0)		34 (38.2)	29 (41.4)	
Tumor location												
Cardia	67 (54.0)	50 (42.4)	0.001*	23 (43.4)	24 (47.1)	0.212	30 (30.6)	10 (33.3)	0.259	13 (14.6)	8 (11.4)	0.742
Gastric body	23 (18.6)	40 (33.9)		11 (20.7)	15 (29.4)		25 (25.5)	10 (33.3)		27 (30.3)	23 (32.9)	
Gastric antrum	31 (25.0)	17 (14.4)		18 (34.0)	9 (17.6)		41 (41.8)	8 (26.7)		49 (55.1)	38 (54.3)	
Whole stomach	3 (2.4)	11 (9.3)		1 (1.9)	3 (5.9)		2 (2.1)	2 (6.7)		0 (0.0)	1 (1.4)	
Clinical T stage												
T2	2 (1.6)	1 (0.8)	0.012*	1 (1.9)	0 (0.0)	<0.001*	5 (5.1)	3 (10.0)	0.012*	4 (4.5)	2 (2.8)	0.009*
T3	56 (45.2)	34 (28.8)		32 (60.4)	8 (15.7)		70 (71.4)	12 (40.0)		28 (31.5)	10 (14.3)	
T4a	63 (50.8)	73 (61.9)		19 (35.8)	40 (78.4)		16 (16.3)	10 (33.3)		52 (58.4)	45 (64.3)	
T4b	3 (2.4)	10 (8.5)		1 (1.9)	3 (5.9)		7 (7.2)	5 (16.7)		5 (5.6)	13 (18.6)	
Clinical N stage												
N0	17 (13.7)	4 (3.4)	0.001*	17 (32.1)	5 (9.8)	<0.001*	9 (9.2)	1 (3.3)	0.718	15 (16.9)	7 (10.0)	0.261
N1	47 (37.9)	32 (27.1)		15 (28.3)	10 (19.6)		38 (38.8)	11 (36.7)		22 (24.7)	13 (18.6)	
N2	40 (32.3)	43 (36.4)		15 (28.3)	14 (27.5)		39 (39.8)	15 (50.0)		21 (23.6)	25 (35.7)	
N3	20 (16.1)	39 (33.1)		6 (11.3)	22 (43.1)		12 (12.2)	3 (10.0)		31 (34.8)	25 (35.7)	

Age and BMI are shown as median (range); other data are number of patients, with percentage in parentheses. *P* value is derived from the Mann–Whitney *U* test (for continuous variables) and Chi-square test (for categorical variables) analyses between clinical characteristics and NACT resistance

NACT neoadjuvant chemotherapy, *BMI* body mass index, *CEA* carcinoembryonic antigen, *CA199* carbohydrate antigen 199

*P value <0.05

Table 2 Performance of models

Models	C-index (95% CI)					
	Internal validation cohort	*P* value	External validation cohort 1	*P* value	External validation cohort 2	*P* value
Clinical	0.733 (0.650–0.816)	–	0.633 (0.532–0.733)	–	0.594 (0.527–0.661)	–
Deep learning	0.808 (0.724–0.893)	0.168	0.755 (0.660–0.850)	0.047*	0.752 (0.678–0.825)	< 0.001*
Integrated	0.835 (0.759–0.912)	0.001*	0.776 (0.680–0.871)	0.002*	0.737 (0.660–0.814)	< 0.001*
Integrated vs clinical						
NRI	0.526 (0.323–0.728)	< 0.001	0.346 (0.156–0.535)	< 0.001	0.224 (0.086–0.363)	0.002
IDI	0.093 (0.055–0.132)	< 0.001	0.138 (0.078–0.198)	< 0.001	0.072 (0.046–0.099)	< 0.001

NRI net reclassification improvement, *IDI* integrated discrimination improvement

P values for deep learning model were derived from the DeLong test between the ROCs of deep learning model and clinical model. *P* values for integrated model were derived from the DeLong test between the ROCs of integrated model and clinical model

**P* value < 0.05

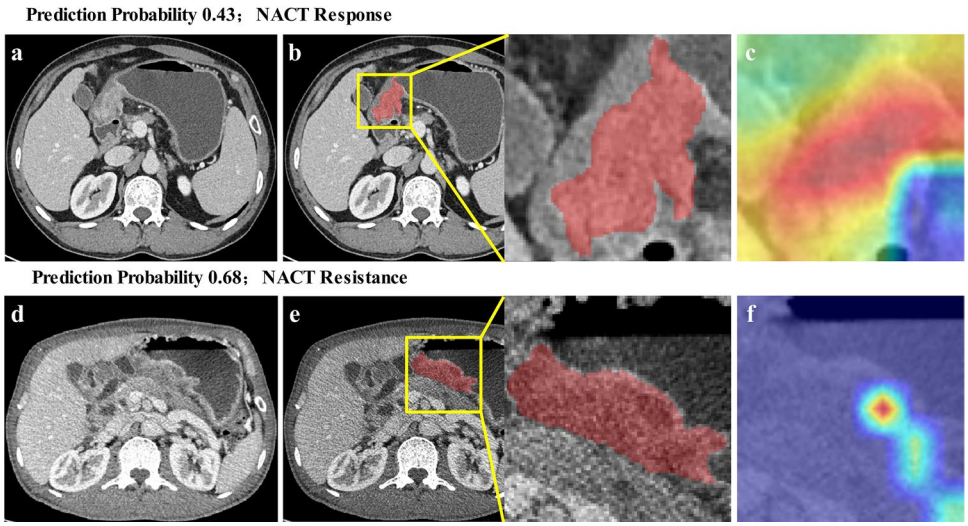

Fig. 2 Representative images for visualization of deep learning (DL) model prediction. **a**, **d** Original computed tomography (CT) images before neoadjuvant chemotherapy (NACT) in patients with response and resistance, respectively. **b**, **e** Two-dimensional (2D) region of interest (ROI) segmentation in the axial CT plane. **c**, **f** Heatmaps overlying on the original input CT images

areas of interest in the model, and conversely, the regions in blue indicate the trivial portions. It was observed that the activation map was not uniform across the entire tumor, and only certain intratumoral regions were activated, indicating that the DL model was particularly responsive to certain tumoral regions.

Integrated model development and validation

The DL signature and clinical T stage, which were significant predictors of tumor resistance to NACT (Table 3), were subsequently combined to establish the integrated model using the multivariable stepwise logistic regression method and build a nomogram (Fig. 3). For all the three validation cohorts, the C-indexes of the integrated model for tumor resistance prediction were 0.835 (95% CI 0.759–0.912), 0.776 (95% CI 0.680–0.871), and 0.737 (95% CI 0.660–0.814), respectively, yielding slightly higher accuracy than the DL model.

Table 3 Related factors for tumor resistance detection in LAGC

Intercept and variable	*β*	Odds ratio (95% CI)	*P* value
Intercept	− 1.261	–	0.949
Clinical T stages			
T2 + 3	Ref	Ref	0.017
T4a + 4b	0.658	1.930(1.128–3.333)	
Deep learning model	2.337	10.348(2.133–56.489)	< 0.005

β is the regression coefficient

LAGC locally advanced gastric cancer

**P* value < 0.05

Afterward, when the DL signature was removed from the nomogram, and only the clinical T stage was kept for building the clinical model, the C-indexes dropped to 0.733 (0.650–0.816), 0.633 (0.532–0.733), and 0.594

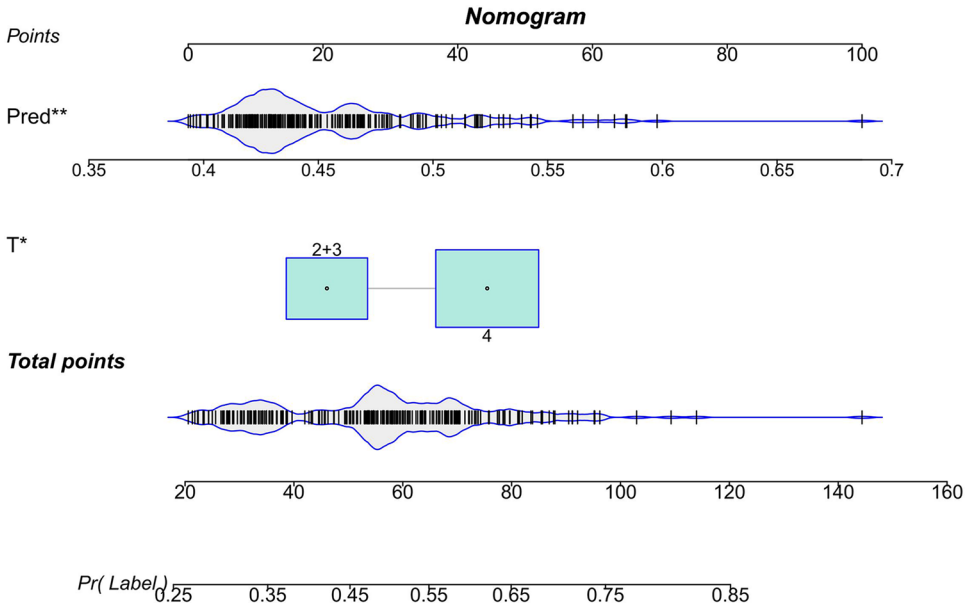

Fig. 3 Integrated model built with deep learning (DL) model output and the clinical T stage

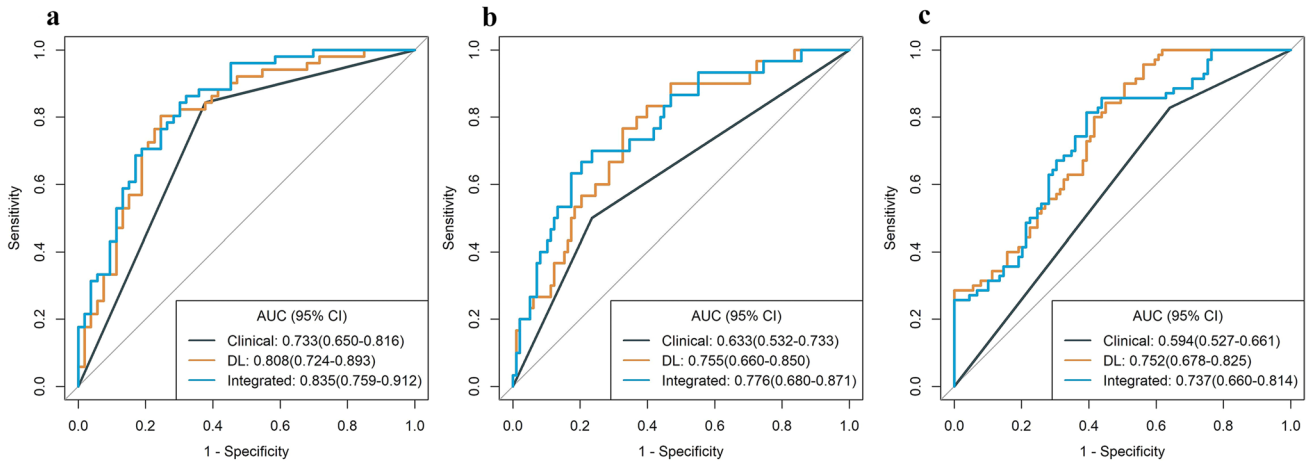

Fig. 4 Receiver operating characteristic (ROC) curves of clinical, deep learning (DL), and integrated models for neoadjuvant chemotherapy (NACT) resistance on the **a** internal validation cohort, **b** external validation cohort 1, and **c** external validation cohort 2, respectively

(0.527–0.661) in all the three validation cohorts, respectively. As shown in Fig. 4 and Table 2, the discrimination performances of the integrated model were significantly higher than the clinical model in all validation cohorts ($P < 0.05$). The calculated NRI and IDI further revealed that the integration of the DL signature into the nomogram performed satisfactorily in all validation cohorts, indicating improved classification accuracy for tumor resistance than the clinical model.

Discussion

In this study, we first developed an end-to-end DL model to predict NACT resistance in patients with LAGC using pre-treatment CT images, after which we independently validated its predictive performance in multicenter cohorts. Notably, the developed DL model showed robust and improved prediction performance in predicting NACT resistance compared to the clinical model.

In clinical practice, it is crucial to reliably identify NACT resistance in patients with LAGC for making personalized treatment decisions in the pre-treatment setting [32]. Given the wide accessibility of pre-treatment CT images, some previous studies have developed pre-treatment CT-based

models using radiomics to evaluate the NACT response for LAGC patients [6, 9–11]. The inspiring findings were mainly obtained due to the precise tumor ROI delineation and complex feature selection, making it difficult to apply these models in clinical practice. In contrast, we used DL to directly extract features from images and automatically select predictive features to build an end-to-end model based on pre-treatment CT.

Inspired by the studies mentioned above, we used the ResNet-50 architecture to build the end-to-end DL model. Then, we used the 12-layer image patch as the input data for the DL network instead of the 1-layer image patch only containing the largest tumor region, which made the DL network extracting more spatial heterogeneity information of the tumor. Moreover, in our study, the image patch size for each patient was determined by the patient's tumor size instead of a fixed value used for all patients, which made the DL network focus more on the tumor region. Accordingly, our DL model successfully predicted NACT resistance with AUCs larger than 0.75 in all internal and external validation cohorts.

Furthermore, to gain insight into how the developed DL model produced an output, we used Grad-CAM to visualize what regions were highlighted, finding that the activation map was not uniform and only certain intratumoral regions were activated (Fig. 2). We presumed that intratumoral heterogeneity might be an important factor for determining NACT resistance, which is consistent with previous studies [33, 34] arguing that imaging heterogeneity of tumor phenotypes has been associated with aggressive biology and poor prognosis in cancers.

Notably, we found that the integrated model showed far better prediction ability of NACT resistance than the clinical model in all validation cohorts ($P < 0.05$). Some previous studies have found that the clinical characteristics can be used for early diagnosis and prognostic evaluation of GC [35–37]. However, the value of clinical characteristics for predicting the resistance to NACT still remains unclear [38, 39]. A previous study showed that a high pre-treatment CA199 level was associated with a higher risk of death [39]. In our study, we found that the CA199 level was not associated with NACT resistance ($P > 0.05$), but the T stage was associated with NACT resistance ($P < 0.05$). Nevertheless, the AUCs of the clinical model in all validation cohorts were significantly lower compared to the integrated model ($P < 0.05$). This finding mainly suggests that the DL model can mine high-dimensional imaging features, which can achieve a more comprehensive quantification of intratumor heterogeneity by combining with the clinical characteristic. Moreover, the improved NRI and IDI also confirmed that the combination of DL signature and clinical characteristics could lead to much better performance in predicting NACT resistance.

The present study has some limitations that should be pointed out. First, as this was a retrospective study, inherent biases were inevitable, although a large number of LAGC patients in three centers were included. Second, in this study, the 12-layer image patch was fed into the DL network using an image with 12 channels, enabling the network to extract some spatial information about the tumor for NACT resistance prediction; however, the whole spatial information was still not fully depicted. In our subsequent studies, a 3D DL network, extracting the whole spatial information of the tumor, will be considered to achieve a better NACT resistance prediction performance. Finally, the model was developed and validated by only using data from East Asian patients. Ideally, our findings should be validated by conducting a prospective randomized trial incorporating more diverse populations in future work.

Conclusions

In conclusion, we developed and validated a CT-based model using DL for the pre-treatment prediction of resistance to NACT in patients with LAGC. The proposed model could identify LAGC patients with resistance before treatment, which provided valuable information and was of great application potential in clinical practice in terms of individual treatment.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10120-022-01328-3.

Acknowledgements Not applicable.

Author contributions YC and XY were responsible for conception and design. JZ, YC, KW, and ZL provided statistical analysis. All authors were involved in drafting and technical support in methods. JR assisted in statistical analysis. JZ and YC were involved in drafting the manuscript, and KW and ZL were involved in reviewing the manuscript. DL and RS had full access to all the data. YC and JZ verified the underlying data. All authors approved the final manuscript for submission.

Funding This study was supported by the National Natural Science Foundation of China (No. 82171923, 82001789, and 82001986), the China Postdoctoral Science Foundation (No. 2021M700897), the Key Research and Development Program of Shandong Province (No. 2021SFGC0104), the Key Research and Development Program of Jiangsu Province (No. BE2021663), the Applied Basic Research Projects of Shanxi Province, China, Outstanding Youth Foundation (No. 202103021222014), the Project of Shanxi Provincial Health Commission (No. 2021XM51, 2020064, 2020TD09, and 2019058), and the Jiangsu Province Engineering Research Center of Diagnosis and Treatment of Children's Malignant Tumor.

Data availability The data sets analyzed during the current study are not publicly available due to the metadata containing information that

could compromise the patients but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethical approval This study was approved by the ethics committee of each participating hospital.

Consent to participate The requirement for informed consent was waived.

Consent for publication Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. 2018;68(6):394–424.
2. Roukos D. Current status and future perspectives in gastric cancer management. Cancer Treat Rev. 2000;26(4):243–55.
3. Ang J, Hu L, Huang P-T, Wu J-X, Huang L-N, Cao C-H, Zheng Y-X, Chen L. Contrast-enhanced ultrasonography assessment of gastric cancer response to neoadjuvant chemotherapy. World J Gastroenterol WJG. 2012;18(47):7026.
4. Wang X-Z, Zeng Z-Y, Ye X, Sun J, Zhang Z-M, Kang W-M. Interpretation of the development of neoadjuvant therapy for gastric cancer based on the vicissitudes of the NCCN guidelines. World J Gast Oncol. 2020;12(1):37.
5. Fu J, Tang L, Li Z-Y, Li X-T, Zhu H-F, Sun Y-S, Ji J-F. Diffusion kurtosis imaging in the prediction of poor responses of locally advanced gastric cancer to neoadjuvant chemotherapy. Eur J Radiol. 2020;128: 108974.
6. Li Z, Zhang D, Dai Y, Dong J, Wu L, Li Y, Cheng Z, Ding Y, Liu Z. Computed tomography-based radiomics for prediction of neoadjuvant chemotherapy outcomes in locally advanced gastric cancer: a pilot study. Chin J Cancer Res. 2018;30(4):406.
7. Lauridsen CA, Falletin E, Hansen ML, Law I, Federspiel B, Bæksgaard L, Svendsen LB, Nielsen MB. Computed Tomography (CT) Perfusion as an early prognostic marker for treatment response to neoadjuvant chemotherapy in gastroesophageal junction cancer and gastric cancer-a prospective study. PLoS ONE. 2014;9:1–10.
8. Virendra K, Yuhua G, Satrajit B. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30(9):1234–48.
9. Xu Q, Sun Z, Li X, Ye C, Zhou C, Zhang L, Lu G. Advanced gastric cancer: CT radiomics prediction and early detection of downstaging with neoadjuvant chemotherapy. Eur Radiol. 2021. https://doi.org/10.1007/s00330-021-07962-2.
10. Wang W, Peng Y, Feng X, Zhao Y, Seeruttun SR, Zhang J, Cheng Z, Li Y, Liu Z, Zhou Z. Development and validation of a computed tomography-based radiomics signature to predict response to neoadjuvant chemotherapy for locally advanced gastric cancer. JAMA Netw Open. 2021;4(8):e2121143–e2121143.
11. Sun K-Y, Hu H-T, Chen S-L, Ye J-N, Li G-H, Chen L-D, Peng J-J, Feng S-T, Yuan Y-J, Hou X. CT-based radiomics scores predict response to neoadjuvant chemotherapy and survival in patients with gastric cancer. BMC Cancer. 2020;20:1–11.
12. Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis C-A, Gaiser T, Marx A, Valous NA, Ferber D. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med. 2019;16(1): e1002730.
13. Peng H, Dong D, Fang M-J, Li L, Tang L-L, Chen L, Li W-F, Mao Y-P, Fan W, Liu L-Z. Prognostic value of deep learning PET/CT-based radiomics: potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2019;25(14):4271–9.
14. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Vega JEV, Brat DJ, Cooper LA. Predicting cancer outcomes from histology and genomics using convolutional networks. Proc Natl Acad Sci. 2018;115(13):E2970–9.
15. Jiang Y, Liang X, Wang W, Chen C, Yuan Q, Zhang X, Li N, Chen H, Yu J, Xie Y. Noninvasive prediction of occult peritoneal metastasis in gastric cancer using deep learning. JAMA Netw Open. 2021;4(1):e2032269–e2032269.
16. Jin C, Jiang Y, Yu H, Wang W, Li B, Chen C, Yuan Q, Hu Y, Xu Y, Zhou Z. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. Br J Surg. 2021;108(5):542–9.
17. Gao Y, Zhang Z-D, Li S, Guo Y-T, Wu Q-Y, Liu S-H, Yang S-J, Ding L, Zhao B-C, Li S. Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer. Chin Med J. 2019;132(23):2804.
18. Amin MBES, Greene F, et al. AJCC Cancer Staging Manual. 8th ed. New York: Springer; 2017.
19. NCCN duideline. https://www.nccn.org/patientresources/patient-resources/guidelines-for-patients.
20. Jia F, Lei TA, Zyl C, Xtl A, Hfz B, Yss A, Jfj C. Diffusion kurtosis imaging in the prediction of poor responses of locally advanced gastric cancer to neoadjuvant chemotherapy—ScienceDirect. Eur J Radiol. 2020. https://doi.org/10.1016/j.ejrad.2020.108974.
21. Yang C, Jiang ZK, Liu LH, Zeng MS. Pre-treatment ADC image-based random forest classifier for identifying resistant rectal adenocarcinoma to neoadjuvant chemoradiotherapy. Int J Colorectal Dis. 2020. https://doi.org/10.1007/s00384-019-03455-3.
22. Zhou X, Yi Y, Liu Z, Cao W, Tian J. Radiomics-based pretherapeutic prediction of non-response to neoadjuvant therapy in locally advanced rectal cancer. Ann Surg Oncol. 2019. https://doi.org/10.1245/s10434-019-07300-3.
23. Jiang Y, Jin C, Yu H, Wu J, Li R. Development and validation of a deep learning CT signature to predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study. Ann Surg. 2020. https://doi.org/10.1097/SLA.0000000000003778.
24. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. pp 770–778
25. Dao T, Gu A, Ratner AJ, Smith V, Sa CD, Ré C. A Kernel Theory of Modern Data Augmentation. PMLR, 2019. pp 1528–1537
26. Pan SJ, Qiang Y. A Survey on Transfer Learning. IEEE Trans Knowl Data Eng. 2010;22(10):1345–59.
27. Steiner B, Devito Z, Chintala S, Gross S, Paszke A, Massa F, Lerer A, Chanan G, Lin Z, Yang E. PyTorch: an imperative style, high-performance deep learning library. Neural information processing systems; 2019. p. 8026–37.
28. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A Learning deep features for discriminative localization. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. pp 2921–2929
29. Sauerbrei W, Boulesteix A-L, Binder H. Stability investigations of multivariable regression models derived from low-and high-dimensional data. J Biopharm Stat. 2011;21(6):1206–31.
30. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.
31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating

characteristic curves: a nonparametric approach. Biometrics. 1988;44:837–45.
32. Cheong J-H, Yang H-K, Kim H, Kim WH, Kim Y-W, Kook M-C, Park Y-K, Kim H-H, Lee HS, Lee KH. Predictive test for chemotherapy response in resectable gastric cancer: a multi-cohort, retrospective analysis. Lancet Oncol. 2018;19(5):629–38.
33. Wu J, Li B, Sun X, Cao G, Rubin DL, Napel S, Ikeda DM, Kurian AW, Li R. Heterogeneous enhancement patterns of tumor-adjacent parenchyma at MR imaging are associated with dysregulated signaling pathways and poor survival in breast cancer. Radiology. 2017;285(2):401–13.
34. Wu J, Cao G, Sun X, Lee J, Rubin DL, Napel S, Kurian AW, Daniel BL, Li R. Intratumoral spatial heterogeneity at perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated with neoadjuvant chemotherapy. Radiology. 2018;288(1):26–35.
35. Feng F, Tian Y, Xu G, Liu Z, Liu S, Zheng G, Guo M, Lian X, Fan D, Zhang H. Diagnostic and prognostic value of CEA, CA19–9, AFP and CA125 for early gastric cancer. BMC Cancer. 2017;17(1):737.
36. Chen XZ, Zhang WK, Yang K, Wang LL, Liu J, Wang L, Hu JK, Zhang B, Chen ZX, Chen JP. Correlation between serum CA724 and gastric cancer: multiple analyses based on Chinese population. Mol Biol Rep. 2012;39(9):9031–9.
37. Yang AP, Liu J, Lei HY, Zhang QW, Zhao L, Yang GH. CA72–4 combined with CEA, \\{CA125\\} and CAl9–9 improves the sensitivity for the early diagnosis of gastric cancer. Clin Chim Acta. 2014. https://doi.org/10.1016/j.cca.2014.07.034.
38. Takekazu Y, Shunkichi K, Akira K, Koichi K, Takayoshi H, Norishige T, Masakazu M. Tumor markers CEA, CA19-9 and CA125 in monitoring of response to systemic chemotherapy in patients with advanced gastric cancer. Jpn J Clin Oncol. 1999;11:550–5.
39. Sun Z, Zhang N. Clinical evaluation of CEA, CA19-9, CA72-4 and CA125 in gastric cancer patients with neoadjuvant chemotherapy. World J Surg Oncol. 2014;12(1):397.