Cross-Sectional Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Dual-layer spectral detector CT-based scoring system for noninvasive TP53 expression assessment and adjuvant chemotherapy response prediction in gastric cancer: a multicenter cross-sectional study

Yinchen Wu, MD[a,b], Yu Lin, MD[a,c], Meilian Xiong, MD[a,b], Nannan Kang, MD[c], Mi Wang, MD[d,e], Yuqiong Wang, BS[a,b], Hong Zheng, BS[a,b], Chengle Ma, BS[a,b], Dairong Cao, MD[a,b,*], Dejun She, PhD[a,b,*]

Background: Noninvasive early prediction of TP53 is a key strategy for improving the prognosis of gastric cancer (GC) patients. To establish a novel, noninvasive and simple scoring system using dual-layer spectral detector computed tomography (DLCT) for preoperative prediction of TP53 expression, prognosis and response to adjuvant chemotherapy (ACT) with retrospective and prospective validation.

Methods: Between April 2021 and March 2025, 568 GC patients were retrospectively and prospectively recruited from two hospitals into a training cohort (TC), a validation cohort (VC), an internal test cohort (ITC), and an external test cohort (ETC). A nomogram prediction model was constructed based on clinical characteristics and DLCT quantitative parameters, which was further simplified to a scoring system. The performance of the system was assessed by discrimination, calibration and clinical applicability. The patients' recurrence-free survival (RFS) and benefit of ACT were evaluated by survival analysis.

Results: The nomogram outperformed clinical, conventional CT and DLCT models for TP53 prediction. The scoring system exhibited comparable predictive efficacy to the nomogram. The areas under the curve of the scoring system were 0.864 (0.815-0.914) for TC, 0.887 (0.802-0.972) for VC, 0.893 (0.844-0.942) for ITC, and 0.912 (0.838-0.985) for ETC, respectively. There were differences in RFS between patients with different TP53 types predicted by the scoring system (P = 0.006, 0.018, 0.015), in consistent with the results from truth labels. RFS differed between the patients who received the system-recommended ACT regimen and those who did not (P = 0.006, 0.027, 0.032). Furthermore, the scoring system was superior to other currently used testing techniques by using cost-effectiveness analysis (incremental cost-effectiveness ratio and incremental cost-utility ratio were 0).

Conclusion: The simple DLCT-based scoring system enables noninvasively, cost-effectively and rapidly predict TP53 expression, prognosis, and benefit from ACT in GC patient, which is expected to guide individualized treatment.

Keywords: adjuvant chemotherapy, cost-effectiveness analysis, gastric cancer, prediction, scoring system

[a]Department of Radiology, the First Affiliated Hospital, Fujian Medical University, Fuzhou, China, [b]Department of Radiology, National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, Fuzhou, China, [c]Department of Radiology, Zhongshan Hospital Affiliated to Xiamen University, Xiamen, China, [d]Department of Pathology, the First Affiliated Hospital, Fujian Medical University, Fuzhou, China and [e]Department of Pathology, National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, Fuzhou, China

Y.C.W. and Y.L. are contributed equally as the first authors

*Corresponding authors. Address: Department of Radiology, First Affiliated Hospital of Fujian Medical University, 20 Cha-Zhong Road, Fuzhou, Fujian 350005, P.R. China. Tel.: +86-591-87983333; E-mail: dejunshe@fjmu.edu.cn (D. She); E-mail: dairongcao@163.com (D. Cao).

International Journal of Surgery (2025) 111:7517–7532

Received 11 February 2025; Accepted 17 June 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 6 August 2025

http://dx.doi.org/10.1097/JS9.0000000000002912

Introduction

Gastric cancer (GC) is the fifth most common malignancy and the third leading cause of cancer death worldwide[1,2]. TP53 mutation is the most common gene mutation in GC[3,4], which is essential for drug resistance, tumor recurrence, and poor overall survival[5–8]. Molecularly targeted drugs for inhibiting TP53 mutation have been developed and proven to exhibit powerful anti-tumor capabilities[9]. The China Anti-Cancer Association Committee of GC and American Society of Clinical Oncology have also suggested that the detection of TP53 is helpful in selecting appropriate therapy protocols[10,11]. Thus, the clinical relevance of TP53 arises from its diagnostic, prognostic and therapeutic potential.

Biomarkers of TP53 are viewed as central for adjusting chemotherapy strategies timely and monitoring response to TP53-targeted therapy. TP53 mutation can be detected by immunohistochemical analysis or genomic analysis[12]. However, this requires invasive tumor biopsy or resection, which is of limited clinical relevance for longitudinal monitoring of treatment response[13–15]. In contrast, noninvasive biomarkers that can detect TP53 expression could provide useful methods for monitoring TP53 mutation status. As a novel technique, the clinical

use of liquid biopsy is limited to large-scale promotion due to long detection time, high costs, and lack of rigorous clinical guidelines[16]. The conventional computed tomography (CT) generally fails to predict TP53 expression in GC patients. Recent evidence has demonstrated that artificial intelligence (AI)-driven conventional CT could provide genetic information about tumors. However, this technique exhibited several notable drawbacks, such as algorithm complexity, advanced equipment-dependence, lack of interpretability, and so on[17–20].

Dual-layer spectral detector CT (DLCT), an advanced energy CT technology, is capable of performing high-quality scanning without increasing the patient's radiation dose and reconstructing a series of functional images[21,22]. Quantitative parameters reflecting pathological information about tumors are obtained by direct measurement from DLCT images without additional aids[21,23,24]. Previous studies have shown that DLCT is effective in raising the diagnostic accuracy of cT stage[25,26], and preoperative prediction of lymph node metastasis, perineural invasion, and expression of markers such as Ki-67[27–29]. Nevertheless, whether DLCT-based quantitative parameters can preoperatively predict TP53 expression and response to adjuvant chemotherapy (ACT) in GC patients remains unknown.

Herein, we developed a novel, noninvasive, cost-effective and simple scoring system for predicting TP53 expression using the parameters through rapid measurements on DLCT images, and evaluated its value in predicting prognosis and response to ACT as well as its economic impact. Its robustness was also assessed retrospectively and prospectively in multicenter patient validation cohorts.

HIGHLIGHTS

- Dual-layer spectral detector computed tomography (CT)-based scoring system for predicting TP53 expression consisted of clinical T stage, CT attenuation of 40 keV VMI in arterial and venous phases, NID and NZ_{eff} in venous phase.
- The efficacy of the scoring system for TP53 prediction, prognosis, and adjuvant chemotherapy response assessment was satisfactory in multicenter, retrospective, and prospective cohorts. In cost-effectiveness analyses, it was the most cost-effective.
- Physicians should focus on patients with scores greater than 21.5, recommend them for adjuvant chemotherapy, and concentrate on following up this group of patients to make timely adjustments to their treatment regimens.

Materials and methods

Ethics

This two-center study was approved by the institutional review boards of the First Affiliated Hospital of Fujian Medical University ([2022] No.334) and Zhongshan Hospital Affiliated to Xiamen University (No.2022-082). The training cohort and the validation cohort recruited patients retrospectively and the informed consent was waived. The internal test cohort and the external test cohort collected patients with informed consent prospectively.

Patients and study design

A total of 568 patients with pathohistologically confirmed GC from two hospitals were included in this study. The screening protocol is detailed in Supplementary Digital content Appendix S1, available at; http://links.lww.com/JS9/E878 and Supplementary Digital content, Fig. S1, available at; http://links.lww.com/JS9/E878. A total of 331 patients were retrospectively collected from April 2021 to July 2023 at Center 1 (the First Affiliated Hospital of Fujian Medical University), which were randomized into the training cohort (n = 232) and the validation cohort (n = 99) at a ratio of 7:3. Between August 2023 and March 2025, the internal test cohort (n = 186) and the external test cohort (n = 51) were prospectively recruited at Center 1 and Center 2 (Zhongshan Hospital Affiliated to Xiamen University), respectively. Follow-up data from patients in the training cohort, the validation cohort and the internal test cohort (from August 2023 to March 2024) were used for survival analysis and ACT response analysis. Sample estimation is described in Supplementary Digital content, Appendix S2, available at: http://links.lww.com/JS9/E878. The flowchart of the study design is shown in Figure 1. The work has been reported in line with the STROCSS criteria[30].

Pathohistological evaluation

Gastrectomy specimens were assessed for TP53 expression by immunohistochemical staining or next-generation sequencing technology to distinguish mutation from wildtype[31,32]. The procedure is provided in Supplementary Digital content, Appendix S3, available at: http://links.lww.com/JS9/E878. All patients were categorized into TP53-wildtype and TP53-mutant based on pathology.

Clinicopathology data collection

Clinicopathological data included age, sex, carbohydrate antigen 199 level, carcinoembryonic antigen level, tumor location, Lauren type, lymphovascular invasion, and perineural invasion. With reference to the American Joint Committee on Cancer 8th staging system[33] and the Chinese clinical guidelines for the diagnosis and treatment of gastric cancer[34], clinical T stage (cT stage) and clinical N stage (cN stage) were determined by CT.

Patients were routinely followed up postoperatively to record follow-up data and recurrence-free survival (RFS) time[34]. The follow-up and ACT requirements are detailed in Supplementary Digital content, Appendix S4, available at: http://links.lww.com/JS9/E878

Collection and post-processing of DLCT data

Scanning was performed on a DLCT scanner (IQon Spectral CT, Philips Healthcare, The Netherlands). An overview of the scanning protocol and information on image reconstruction is presented in Supplementary Digital content, Appendix S5, available at: http://links.lww.com/JS9/E878.

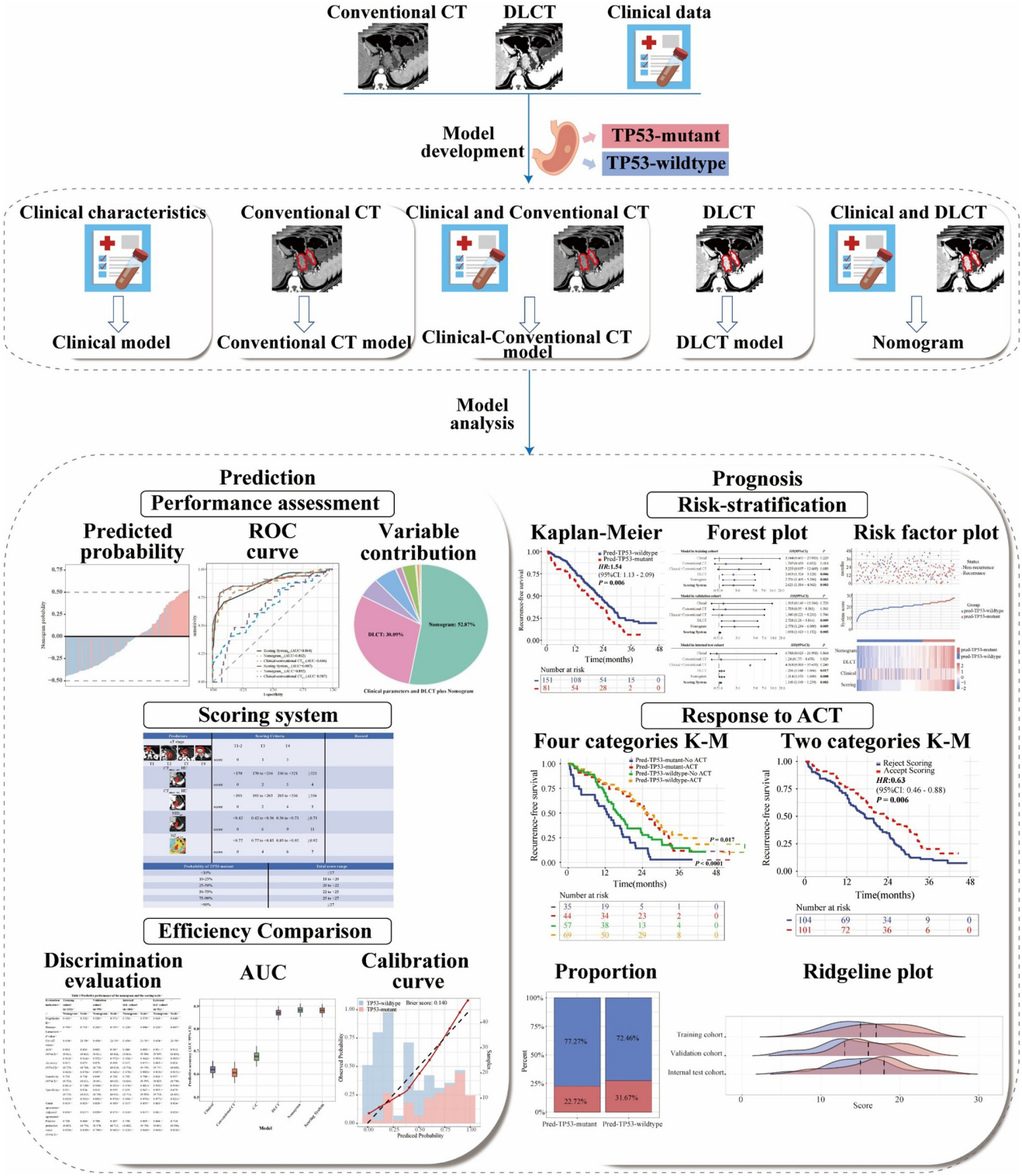

Figure 1. The overall flowchart of this study. DLCT = Dual-layer spectral detector computed tomography; ROC = Receiver-operator characteristic; ACT = Adjuvant chemotherapy; K-M = Kaplan-Meier.

Image analysis and parameter measurement

Radiomorphological parameters of the tumor were recorded, including the maximum diameter as the longest diameter of the lesion, the maximum thickness as the largest diameter perpendicular to the maximum diameter, cT stage and cN stage. CT attenuation and arterial enhancement fraction of the tumor were measured on conventional images in non-contrast scan, arterial phase (AP), and venous phase (VP). DLCT spectral

quantitative parameters were also measured in double phase, including CT attenuation of virtual monoenergetic images (VMIs) from 40 keV to 100 keV at 10 keV intervals, effective atomic number (Z_{eff}), iodine density (ID), normalized Z_{eff} (NZ_{eff}), and normalized ID (NID). VMIs mimic images generated from application of a monoenergetic beam at a single keV level, allowing a better demonstration of tissue attenuation changes[21]. Z_{eff} images are color-coded based on the effective atomic number of tissues, which describes the material composition of the tissue[21]. In ID images, iodine-containing pixels are assigned values equal to the concentration of iodine in each pixel, which allows the quantification of iodine in blood vessels and organs[21]. NZ_{eff} and NID are values after elimination of individual differences in hemodynamics and clearly reflect their uniform variation in the population[35]. Detailed overview is presented in Supplementary Digital content, Appendix S6, available at: http://links.lww.com/JS9/E878.

Model development

In the training cohort, statistically significant clinical predictors, conventional CT predictors, and DLCT predictors were identified by univariate and multivariate logistic regression analyses to construct a clinical model, a conventional CT model and a DLCT model. A clinical-conventional CT model and a clinical-DLCT model were developed by combining the relevant predictors. The latter was visualized as a nomogram. For convenience of clinical use and memorization, the model parameters were divided into hierarchical variables to create a scoring system. Details of the calculations are shown in Supplementary Digital content, Appendix S7, available at: http://links.lww.com/JS9/E878.

Statistical analysis

Data were statistically analyzed using MedCalc (v.20.0.22, Medcalc Software Ltd), and RStudio (v.4.3.2, R Foundation). Economic impact of clinical strategies for preoperative TP53 prediction was calculated through a Markov model (Supplementary Digital content, Appendix S8, available at: http://links.lww.com/JS9/E878. Data descriptive statistics, intergroup comparisons, interobserver agreement, and the comparisons of the efficacy for the models and the scoring system are provided in Supplementary Digital content, Appendix S9, available at: http://links.lww.com/JS9/E878. Bonferroni correction was used for adjusted *P*-values of multiple comparisons. A two-tailed *P*-value <0.05 was considered statistically significant.

Results

Clinicopathological characteristics

The clinicopathological characteristics of all 568 patients are demonstrated in Table 1. In the training cohort, the validation cohort, the internal test cohort, and the external test cohort, 96 (41.4%), 33 (33.3%), 72 (38.7%), and 23 (45.1%) patients with TP53-mutant were identified, respectively. The cT stage and cN stage showed significant differences between TP53-mutant and TP53-wildtype GC in all four cohorts (all $P < 0.05$). The age, tumor location, Lauren type, and perineural invasion also showed differences in some of the cohorts ($P < 0.05$). In the training cohort, after univariate and multivariate logistic regression analyses, only cT stage was retained as an independent predictor to construct the clinical model (Table 2).

Interobserver agreement

The quantitative parameters based on DLCT images, including the maximum diameter, the maximum thickness, and DLCT parameters in AP and VP, had excellent ICC values of 0. 872-0.976 (Supplementary Digital content, Table S1, available at: http://links.lww.com/JS9/E878). The ICC values for the conventional CT data, consisting of CT attenuation in non-contrast scan, AP and VP, were 0.814-0.868 (Supplementary Digital content, Table S2, available at: http://links.lww.com/JS9/E878), with poorer interobserver agreement than the results from DLCT. For subjective imaging features, cT stage and cN stage, the interobserver agreement was excellent with kappa values of 0.826-0.870 (Supplementary Digital content, Table S3, available at: http://links.lww.com/JS9/E878).

Conventional CT model and clinical-conventional CT model construction

In the training cohort, conventional CT data were analyzed by logistic regression analysis (Supplementary Digital content, Table S4, available at: http://links.lww.com/JS9/E878). Arterial enhancement fraction was the only independent predictor used to construct the conventional CT model. We combined the clinical predictor with arterial enhancement fraction to construct the clinical-conventional CT model.

DLCT model construction

In the training cohort, DLCT-based radiomorphological features and spectral quantitative parameters were analyzed by univariate and multivariate regression analyses (Table 3). The results revealed that CT_{40keV} in AP, CT_{40keV} and CT_{60keV}, NID and NZ_{eff} in VP were independent predictors ($P < 0.05$). Among them, CT_{40keV} and CT_{60keV} in VP had collinearity (Supplementary Digital content, Table S5, available at: http://links.lww.com/JS9/E878). According to previous studies, 40 keV VMI had better signal-to-noise ratio and increased diagnostic performance for GC depiction[25]. Thus, CT_{40keV} was retained and constructed as DLCT model with other predictors. The distribution of DLCT parameters in four cohorts is displayed in Supplementary Digital content, Fig. S2, available at: http://links.lww.com/JS9/E878.

Table 4 revealed that DLCT model exhibited robust predictive efficacy in four cohorts. Supplementary Digital content, Table S6, available at: http://links.lww.com/JS9/E878 showed that the predictive performance of DLCT model was significantly better than the clinical, the conventional CT and the clinical-conventional CT models (DeLong test, all $P < 0.05$).

Nomogram performance

The clinical predictor was combined with DLCT model to establish a clinical-DLCT model, which was visualized as a nomogram (Fig. 2a and Supplementary Digital content, Table S7, available at: http://links.lww.com/JS9/E878). Table 5 and Supplementary Digital content, Fig. S3, available at: http://links.lww.com/JS9/E878) demonstrated the satisfactory efficacy

Table 1

Comparison of clinicopathological variables in the training, validation, internal and external test cohorts

	Training cohort (n = 232)			Validation cohort (n = 99)			Internal test cohort (n = 186)			External test cohort (n = 51)		
Variables	**TP53-wildtype (n = 136)**	**TP53-mutant (n = 96)**	***P*-value**	**TP53-wildtype (n = 66)**	**TP53-mutant (n = 33)**	***P*-value**	**TP53-wildtype (n = 114)**	**TP53-mutant (n = 72)**	***P*-value**	**TP53-wildtype (n = 28)**	**TP53-mutant (n = 23)**	***P*-value**
Age (y)	68.00	64.50	0.031*	64.00	67.00	0.004*	67.00	67.00	0.738	70.50	66.00	0.389
median (IQR)	(62.00, 74.00)	(60.25, 70.00)		(55.75, 69.25)	(63.50, 75.50)		(58.00, 72.00)	(58.50, 72.00)		(60.25, 75.00)	(59.00, 74.00)	
Gender (%)			0.205			0.528			0.076			0.650
Male	90 (66.2%)	71 (74.0%)		46 (69.7%)	25 (75.8%)		73 (64.0%)	55 (76.4%)		22 (78.6%)	21 (91.3%)	
Female	46 (33.8%)	25 (26.0%)		20 (30.3%)	8 (24.2%)		41 (36.0%)	17 (23.6%)		6 (21.4%)	2 (8.7%)	
CA199 (n, %)			0.913			0.470			0.811			0.163
Normal	104 (76.5%)	74 (77.1%)		52 (78.8%)	28 (84.8%)		105 (92.1%)	67 (93.1%)		24 (85.7%)	16 (69.6%)	
Elevated	32 (23.5%)	22 (22.9%)		14 (21.2%)	5 (15.2%)		9 (7.9%)	5 (6.9%)		4 (14.3%)	7 (30.4%)	
CEA (n, %)			0.623			0.435			0.367			0.438
Normal	101 (74.3%)	74 (77.1%)		45 (68.2%)	25 (75.8%)		107 (93.9%)	65 (90.3%)		27 (96.4%)	21 (91.3%)	
Elevated	35 (25.7%)	22 (22.9%)		21 (31.8%)	8 (24.2%)		7 (6.1%)	7 (9.7%)		1 (3.6%)	2 (8.7%)	
cT stage (%)			0.003*			0.004*			<0.001*			0.029*
T1-2	39 (28.7%)	12 (12.5%)		19 (28.8%)	4 (12.1%)		39 (34.2%)	6 (8.3%)		11 (39.3%)	2 (8.7%)	
T3	58 (42.6%)	39 (40.6%)		16 (24.2%)	19 (57.6%)		38 (33.3%)	26 (36.1%)		9 (32.1%)	8 (34.8%)	
T4	39 (28.7%)	45 (46.9%)		31 (47.0%)	10 (30.3%)		37 (32.5%)	40 (55.6%)		8 (28.6%)	13 (56.5%)	
cN stage (%)			0.034*			0.022*			0.005*			0.020*
N0	52 (38.2%)	24 (25.0%)		34 (51.5%)	9 (27.3%)		70 (61.4%)	29 (40.3%)		20 (71.4%)	9 (39.1%)	
N1-3	84 (61.8%)	72 (75.0%)		32 (48.5%)	24 (72.7%)		44 (38.6%)	43 (59.7%)		8 (28.6%)	14 (60.9%)	
Tumor location (%)			0.188			0.501			0.025*			0.125
Cardia/fundus	37 (27.2%)	37 (38.5%)		15 (22.7%)	11 (33.3%)		21 (18.4%)	26 (36.1%)		13 (46.4%)	5 (21.7%)	
Corpus	46 (33.8%)	28 (29.2%)		23 (34.8%)	9 (27.3%)		36 (31.6%)	17 (23.6%)		9 (32.1%)	8 (34.8%)	
Antrum/pylorus	53 (39.0%)	31 (32.3%)		28 (42.4%)	13 (39.4%)		57 (50.0%)	29 (40.3%)		6 (21.4%)	10 (43.5%)	
Lauren type (%)			0.517			0.546			0.711			0.004*
Intestinal type	53 (39.0%)	39 (40.6%)		29 (43.9%)	18 (54.5%)		48 (42.1%)	26 (36.1%)		15 (53.6%)	3 (13.0%)	
Mixed type	46 (33.8%)	37 (38.5%)		22 (33.3%)	10 (30.3%)		31 (27.2%)	21 (29.2%)		4 (14.3%)	11 (47.8%)	
Diffuse type	37 (27.2%)	20 (20.8%)		15 (22.7%)	5 (15.2%)		35 (30.7%)	25 (34.7%)		9 (32.1%)	9 (39.1%)	
Lymphovascular invasion (%)			0.365			0.559			0.250			0.238
Negative	79 (58.1%)	50 (52.1%)		42 (63.6%)	19 (57.6%)		79 (69.3%)	44 (61.1%)		18 (64.3%)	11 (47.8%)	
Positive	57 (41.9%)	46 (47.9%)		24 (36.4%)	14 (42.4%)		35 (30.7%)	28 (38.9%)		10 (35.7%)	12 (52.2%)	
Perineural invasion (%)			0.327			0.663			0.832			0.022*
Negative	48 (35.3%)	28 (29.2%)		27 (40.9%)	12 (36.4%)		73 (64.0%)	45 (62.5%)		21 (75.0%)	10 (43.5%)	
Positive	88 (64.7%)	68 (70.8%)		39 (59.1%)	21 (60.6%)		41 (36.0%)	27 (37.5%)		7 (25.0%)	13 (56.5%)	

Note. The reference range for CA 199 is 0-34 U/ml and for CEA is 0-4.7 ng/ml. *P*-values were calculated by Mann-Whitney *U*-test for non-normal quantitative data and the chi-square test or Fisher's exact test for categorical variables. * *P*-value<0.05. CA 199 = Carbohydrate antigen 199, CEA = Carcinoembryonic antigen.

Table 2

Univariate and multivariate logistic regression analysis of clinicopathological characteristics in the training cohort

	Univariate analysis		Multivariate analysis	
Variables	***OR* (95% CI)**	***P*-value**	***OR* (95% CI)**	***P*-value**
Age	0.98 (0.95, 1.01)	0.113		
Sex	0.69 (0.39, 1.23)	0.206		
CA199	1.04 (0.56, 1.92)	0.913		
CEA	1.17 (0.63, 2.15)	0.623		
cT stage				
T1-2	Ref.		Ref.	
T3	2.19 (1.02, 4.69)	0.045*	1.64 (0.71, 3.81)	0.250
T4	3.75 (1.73, 8.15)	<0.001*	3.21 (1.44, 7.15)	0.004*
cN stage	1.86 (1.04, 3.31)	0.036*	1.69 (0.88, 3.25)	0.119
Tumor location				
Cardia/fundus	Ref.			
Corpus	0.61 (0.32, 1.17)	0.137		
Antrum/pylorus	0.59 (0.31, 1.10)	0.098		
Lauren type				
Intestinal type	Ref.			
Mixed type	1.09 (0.60, 1.99)	0.771		
Diffuse type	0.74 (0.37, 1.46)	0.376		
Lymphovascular invasion	1.28 (0.75, 2.16)	0.365		
Perineural invasion	1.33 (0.75, 2.33)	0.328		

Note. Only significant variables in univariate analysis (* represents P<0.05) were included in multivariate analysis. OR = Odds ratio; CI = Confidence interval; CA 199 = Carbohydrate antigen 199, CEA = Carcinoembryonic antigen.

Table 3

Univariate and multivariate logistic regression analysis of DLCT quantitative parameters in the training cohort

	Univariate analysis		Multivariate analysis	
Variables	***OR* (95% CI)**	***P*-value**	***OR* (95% CI)**	***P*-value**
Maximum diameter	0.96 (0.83, 1.11)	0.553		
Maximum thickness	1.43 (1.03, 1.99)	0.035*	1.27 (0.89, 1.91)	0.261
Arterial phase				
CT_{40keV}	1.01 (1.01, 1.02)	<0.001*	1.01 (1.00, 1.02)	0.005*
CT_{50keV}	1.01 (1.00, 1.02)	0.074		
CT_{60keV}	1.01 (1.00, 1.03)	0.073		
CT_{70keV}	1.02 (1.00, 1.04)	0.088		
CT_{80keV}	1.03 (1.00, 1.05)	0.051		
CT_{90keV}	1.02 (0.99, 1.05)	0.212		
CT_{100keV}	1.03 (0.99, 1.06)	0.107		
ID	1.48 (0.96, 2.26)	0.075		
NID	6.67 (1.09, 40.70)	0.040*	2.85 (1.00, 8.13)	0.076
Z_{eff}	1.48 (0.80, 2.73)	0.211		
NZ_{eff}	0.52 (0.01, 25.33)	0.741		
Venous phase				
CT_{40keV}	1.02 (1.01, 1.02)	<0.001*	1.05 (1.02, 1.08)	<0.001*
CT_{50keV}	1.02 (1.01, 1.03)	<0.001*	0.98 (0.94, 1.03)	0.421
CT_{60keV}	1.02 (1.00, 1.03)	0.037*	0.94 (0.90, 0.99)	0.010*
CT_{70keV}	1.01 (1.00, 1.03)	0.118		
CT_{80keV}	1.00 (0.98, 1.02)	0.949		
CT_{90keV}	1.02 (0.99, 1.05)	0.328		
CT_{100keV}	0.99 (0.96, 1.02)	0.448		
ID	1.47 (0.98, 2.20)	0.062		
NID	2.94 (0.60, 14.47)	0.044*	2.73 (1.87, 4.00)	<0.001*
Z_{eff}	1.53 (0.84, 2.81)	0.167		
NZ_{eff}	1.09 (1.01, 1.17)	0.028*	5.82 (2.28, 14.82)	<0.001*

Note. Only significant variables in univariate analysis (* represents P<0.05) were included in multivariate analysis. Since the ranges of NID and NZ_{eff} are narrow, the *OR*s (95%CI) are overlarge. Therefore, only the *OR*s (95%CI) for this two are calculated after the conversion of them with multiplication by 10. OR = Odds ratio; CI = Confidence interval; ID = Iodine density; NID = Normalized iodine density; Z_{eff} = Effective atomic number; NZ_{eff} = Normalized effective atomic number.

of the nomogram model. As shown in Figure 3a-b and Supplementary Digital content, Table S6, available at: http://links.lww.com/JS9/E878, the nomogram significantly outperformed the clinical model and conventional CT models, in terms of predictive efficacy (DeLong test, all $P < 0.05$). The nomogram, also as a combined model, showed superior discriminability against the clinical-conventional CT model (Supplementary Digital content, Fig. S4, available at: http://links.lww.com/JS9/E878). The integrated discrimination index and net reclassification index proved that the nomogram further improved the accuracy of TP53 expression classification over DLCT model (Supplementary Digital content, Table S8, available at: http://links.lww.com/JS9/E878). In all cohorts, favorable agreement between the predicted and actual TP53 mutation probabilities was observed (Brier score: 0.124-0.142), with excellent clinical net benefit (Supplementary Digital content, Fig. S5, available at: http://links.lww.com/JS9/E878).

We also evaluated the relative contribution of individual variables to TP53 prediction. Among the preoperative clinically measurable indicators and DLCT parameters, DLCT parameters were the most important factor. However, when the nomogram was added, it became the dominant variable in TP53 prediction (Fig. 3c).

Simplified scoring system for prediction

In accordance with multivariate logistic regression analysis, when the total NID point was identified as 100, the total points of CT_{40keV} in AP, CT_{40keV} and NZ_{eff} in VP were 36, 51, and 66, respectively, with cT3 stage of 7 points. The above points were divided by 7 to form a simplified scoring system (Fig. 2b).

As shown in Table 5 and Figure 3, the scoring system exhibited favorable predictive efficacy in all four cohorts, with AUCs of 0.864 (95% CI 0.815, 0.914) for the training cohort, 0.887 (95% CI 0.802, 0.972) for the validation cohort, 0.893 (95% CI 0.844, 0.942) for the internal test cohort, and 0.912 (95% CI 0.838, 0.985) for the external cohort. Multiple tests, including the non-inferiority study test, showed that the system maintained the optimized efficacy of the nomogram (Supplementary Digital content, Table S9, available at: http://links.lww.com/JS9/E878). Across the four cohorts, the system provided good agreement between predicted and actual probabilities (Supplementary Digital content, Fig. S6, available at: http://links.lww.com/JS9/E8783). The Hosmer-Lemeshow goodness-of-fit test indicated acceptable fit with all $P > 0.05$, and Brier scores ranged from 0.104 to 0.140. The system also exhibited excellent clinical utility (Supplementary Digital content, Fig. S6, available at: http://links.lww.com/JS9/E878). Figure 4 depicts TP53 expression in GC patients accurately predicted by the nomogram and the scoring system.

To further explore the predictive performance of the system for critical clinical characteristics, such as cT stage and cN stage, subgroup analyses were also performed. In four cohorts, the system was effective in predicting cT stage and cN stage with

Table 4

Predictive performance of the clinical, the conventional CT, DLCT and the clinical-conventional CT model for TP53 expression

Models	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value(95% CI)	Negative predictive value (95% CI)
Clinical						
training cohort	0.596 (0.544, 0.647)	0.547 (0.481, 0.613)	0.875 (0.794, 0.927)	0.316 (0.244, 0.399)	0.475 (0.441, 0.509)	0.782 (0.666, 0.865)
validation cohort	0.583 (0.504, 0.662)	0.485 (0.383, 0.587)	0.879 (0.727, 0.952)	0.288 (0.193, 0.406)	0.382 (0.336, 0.430)	0.826 (0.637, 0.928)
internal test cohort	0.629 (0.575, 0.684)	0.565 (0.490, 0.637)	0.917 (0.830, 0.961)	0.342 (0.261, 0.433)	0.468 (0.431, 0.505)	0.867 (0.744, 0.936)
external test cohort	0.613 (0.499, 0.728)	0.588 (0.442, 0.724)	0.870 (0.679, 0.955)	0.357 (0.207, 0.542)	0.526 (0.447, 0.604)	0.769 (0.509, 0.915)
Convention CT						
training cohort	0.592 (0.516, 0.668)	0.647 (0.581, 0.708)	0.354 (0.266, 0.454)	0.853 (0.784, 0.903)	0.630 (0.511, 0.734)	0.652 (0.614, 0.688)
validation cohort	0.625 (0.500, 0.750)	0.667 (0.565, 0.758)	0.242 (0.128, 0.410)	0.879 (0.779, 0.938)	0.500 (0.292, 0.708)	0.699 (0.652, 0.742)
internal test cohort	0.627 (0.547, 0.707)	0.565 (0.490, 0.637)	0.222 (0.142, 0.331)	0.781 (0.696, 0.847)	0.390 (0.269, 0.527)	0.614 (0.576, 0.650)
external test cohort	0.603 (0.444, 0.761)	0.588 (0.442, 0.724)	0.217 (0.097, 0.419)	0.893 (0.728, 0.963)	0.625 (0.308, 0.862)	0.581 (0.519, 0.641)
DLCT						
training cohort	0.855 (0.803, 0.908)	0.815 (0.759, 0.863)	0.688 (0.589, 0.772)	0.904 (0.843, 0.943)	0.835 (0.748, 0.897)	0.804 (0.752, 0.847)
validation cohort	0.880 (0.793, 0.968)	0.869 (0.786, 0.928)	0.788 (0.623, 0.893)	0.909 (0.816, 0.958)	0.813 (0.664, 0.905)	0.896 (0.816, 0.943)
internal test cohort	0.872 (0.819, 0.925)	0.812 (0.748, 0.865)	0.750 (0.639, 0.836)	0.851 (0.774, 0.905)	0.761 (0.668, 0.834)	0.843 (0.782, 0.890)
external test cohort	0.896 (0.813, 0.979)	0.804 (0.669, 0.902)	0.783 (0.581, 0.903)	0.821 (0.644, 0.921)	0.783 (0.613, 0.891)	0.821 (0.675, 0.911)
Clinical-conventional CT						
training cohort	0.666 (0.595, 0.737)	0.668 (0.603, 0.728)	0.385 (0.294, 0.485)	0.868 (0.801, 0.915)	0.673 (0.555, 0.772)	0.667 (0.628, 0.704)
validation cohort	0.587 (0.471, 0.702)	0.586 (0.482, 0.684)	0.152 (0.051, 0.319)	0.803 (0.692, 0.881)	0.278 (0.130, 0.497)	0.654 (0.611, 0.696)
internal test cohort	0.690 (0.615, 0.766)	0.651 (0.577, 0.719)	0.375 (0.264, 0.497)	0.825 (0.745, 0.884)	0.574 (0.451, 0.689)	0.676 (0.632, 0.718)
external test cohort	0.688 (0.540, 0.836)	0.608 (0.461, 0.742)	0.261 (0.126, 0.465)	0.893 (0.718, 0.977)	0.667 (0.359, 0.877)	0.595 (0.528, 0.659)

Note. DLCT = Dual-layer spectral detector CT; AUC = Area under the curve; CI = Confidence interval.

all $P < 0.05$ (Supplementary Digital content, Table S10-11, available at: http://links.lww.com/JS9/E878). Additionally, differences in the system score were noted in clinicopathologic factors to varying degrees, as indicated in Supplementary Digital content, Figure S7, available at: http://links.lww.com/JS9/E878.

Performance of scoring system for prognosis

We further explored the prognostic value of the scoring system in GC patients. The median follow-up was 18.3 months in the training cohort, 20.7 months in the validation cohort and 12.7 months in the internal test cohort. We categorized the patients into predicted TP53-mutant (pred-TP53-mutant) and predicted TP53-wildtype (pred-TP53-wildtype) based on the scoring system. Kaplan-Meier survival curves illustrated that the results of the predictive labels based on the scoring system were consistent with the ones based on the ground truth labels, both indicating statistically significant differences in RFS between the patients with different pred-types (all $P < 0.05$, Fig. 5a-c). The system scores, survival status, and survival time for each patient were visualized in Supplementary Digital content, Fig. S8, available at: http://links.lww.com/JS9/E878. We also performed Kaplan-Meier survival analyses for the nomogram and obtained a consistent result in Supplementary Digital content, Fig. S9, available at: http://links.lww.com/JS9/E878. Proportional hazards were determined using Schoenfeld residuals test, showing that all *P*-values were over 0.05. Figure 6a delivered the results of Cox proportional hazards regression analyses for the different models and the scoring system,

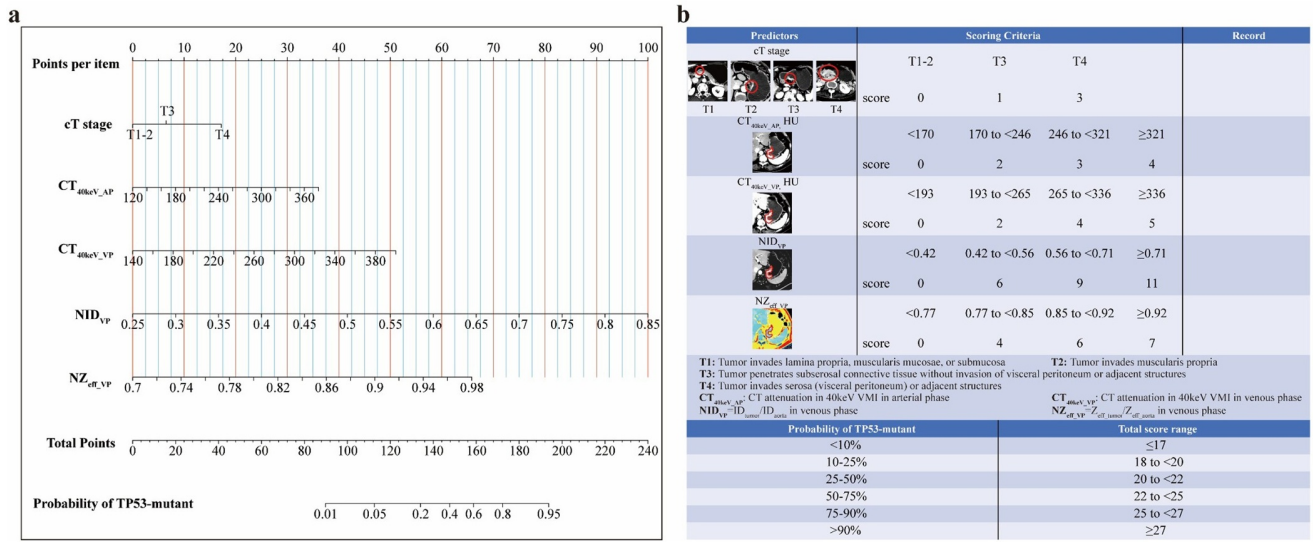

Figure 2. The nomogram and the scoring system for predicting TP53-mutant GC. (a) The nomogram shows the predicted TP53-mutant probability. The leftmost column of the scoring system (b) displays the representative images and the measurements for each variable. Readers can state the score for each variable and calculate the total score in the record box. The bottom section provides predicted probabilities based on the range of total scores. CT_{40keV_AP} = CT attenuation in 40 keV virtual monoenergetic images in arterial phase. CT_{40keV_VP} = CT attenuation in 40 keV virtual monoenergetic images in venous phase; ID = Iodine density; $NID_{VP} = ID_{tumor}/ID_{aorta}$ in venous phase; Z_{eff} = Effective atomic number; $NZ_{eff\_VP} = Z_{eff\_tumor}/Z_{eff\_aorta}$ in venous phase.

demonstrating that a low system score was a strong factor for RFS ($P < 0.05$).

Scoring system for the prediction of adjuvant chemotherapy benefit

We further assessed the predictive value of the system regarding the benefit of postoperative ACT in the follow-up patients. Propensity score matching was not performed due to the balanced baseline characteristics of patients with and without ACT. The clinicopathological characteristics of the patients grouped based on the system were shown in Supplementary Digital content, Table S12, available at: http://links.lww.com/JS9/E878. In accordance with the predicted labels and the treatment regimens, all patients were divided into four categories: pred-TP53-mutant patients receiving ACT, pred-TP53-mutant patients not receiving ACT, pred-TP53-wildtype patients receiving ACT and pred-TP53-wildtype patients not receiving ACT. As shown in Figure 5d-f, pred-TP53-mutant patients had a greater reduction in the risk of recurrence than pred-TP53-wildtype ones after ACT, implying a higher benefit therefrom. The results of statistical interaction test between TP53 types and ACT confirmed a significant interaction effect on RFS ($P < 0.05$).

Subsequently, we assessed survival differences between patients who chose to receive ACT upon the recommendation of the scoring system and those who were not advised by the system. The aforementioned four groups were further categorized into two distinct classifications: patients who complied with the system-recommended ACT (including pred-TP53-mutant-ACT and pred-TP53-wildtype-No ACT groups) and those who disregarded the system-recommended ACT (including pred-TP53-mutant-No ACT and pred-TP53-wildtype-ACT groups). Acceptance of treatment decisions suggested by the scoring system improved the patients' median survival time (the training cohort: 17.3 months vs. 23.5 months; the validation cohort: 18.9 months vs. 29.5 months; the internal test cohort: 14.7 months vs. > 18 months). Logistic regression analysis also identified the system scores as an independent predictor of ACT response (Supplementary Digital content, Table S13, available at: http://links.lww.com/JS9/E878).

For ACT patients, we categorized them into two groups, ACT-sensitive and ACT-resistant groups, depending on whether they benefited from ACT (RFS ≥ 1 or < 1 year). We found that the proportion of ACT-sensitive patients was higher or almost equal in pred-TP53-mutant group than in pred-TP53-wildtype group, and vice versa (Fig. 6b). The patients in ACT-sensitive group scored higher than those in another group ($P < 0.05$, Fig. 6c). These data suggest that the scoring system predicts the benefit from ACT in GC with different TP53 expression.

Cost-effectiveness analysis

Utilizing the accuracy calculated in this study, we performed a cost-effectiveness analysis to assess the economic impact of the scoring system when compared with other preoperative noninvasive tests. Through Markov model-based cost-effectiveness analysis, we estimated that TP53 prediction using the scoring system was significantly more cost-effective than other currently common clinical practices (incremental cost-effectiveness ratio and incremental cost-utility ratio both 0, Supplementary Digital content, Table S14, available at: http://links.lww.com/JS9/E878). Taken together, this multicenter, prospective combined with retrospective, and large sample size study demonstrated that our proposed scoring system provides a cost-efficient, robust assay for noninvasive detection of TP53 expression in GC, with potential for future clinical applications.

Table 5

Predictive performance of the nomogram and the scoring system

	Training cohort (n = 232)		Validation cohort (n = 99)		Internal test cohort (n = 186)		External test cohort (n = 51)	
Evaluation indicators	**Nomogram**	**Scoring system**	**Nomogram**	**Scoring system**	**Nomogram**	**Scoring system**	**Nomogram**	**Scoring system**
Nagelkerke R^2	0.492	0.502	0.520	0.571	0.551	0.578	0.648	0.648
Hosmer-Lemeshow *P* value	0.709	0.701	0.203	0.547	0.229	0.806	0.226	0.687
Cut-off value	0.454	21.50	0.454	21.50	0.454	21.50	0.454	21.50
AUC (95%CI)	0.862 (0.811, 0.913)	0.864 (0.815, 0.914)	0.892 (0.811, 0.972)	0.887 (0.802, 0.972)	0.886 (0.836, 0.936)	0.893 (0.844, 0.942)	0.921 (0.849, 0.993)	0.912 (0.838, 0.985)
Accuracy (95% CI)	0.815 (0.759, 0.863)	0.823 (0.768, 0.870)	0.828 (0.739, 0.897)	0.889 (0.810, 0.943)	0.817 (0.754, 0.870)	0.855 (0.796, 0.902)	0.863 (0.737, 0.943)	0.804 (0.669, 0.902)
Sensitivity (95%CI)	0.792 (0.700, 0.861)	0.708 (0.611, 0.790)	0.849 (0.691, 0.934)	0.788 (0.623, 0.893)	0.792 (0.684, 0.870)	0.708 (0.595, 0.801)	0.826 (0.629, 0.930)	0.957 (0.790, 0.998)
Specificity (95%CI)	0.831 (0.759, 0.885)	0.904 (0.843, 0.943)	0.818 (0.709, 0.893)	0.939 (0.854, 0.976)	0.833 (0.754, 0.891)	0.947 (0.890, 0.976)	0.893 (0.718, 0.977)	0.679 (0.493, 0.821)
Crude agreement	0.815	0.823	0.828	0.889	0.817	0.855	0.863	0.804
Adjusted agreement	0.810	0.817	0.820	0.873	0.810	0.847	0.861	0.824
Positive predictive value (95%CI)	0.768 (0.692, 0.829)	0.840 (0.754, 0.899)	0.700 (0.578, 0.799)	0.867 (0.712, 0.945)	0.750 (0.662, 0.821)	0.895 (0.794, 0.949)	0.864 (0.681, 0.949)	0.710 (0.586, 0.808)
Negative predictive value (95%CI)	0.850 (0.792, 0.894)	0.815 (0.762, 0.858)	0.915 (0.827, 0.961)	0.899 (0.821, 0.945)	0.864 (0.800, 0.909)	0.837 (0.782, 0.881)	0.862 (0.718, 0.939)	0.950 (0.733, 0.992)
Positive likelihood ratio (95%CI)	4.681 (3.181, 6.889)	7.410 (4.350, 12.623)	4.667 (2.742, 7.942)	13.000 (4.947, 34.163)	4.750 (3.098, 7.282)	13.458 (6.091, 29.737)	7.710 (2.604, 22.831)	2.976 (1.725, 5.133)
Negative likelihood ratio (95%CI)	0.251 (0.169, 0.373)	0.322 (0.235, 0.443)	0.185 (0.082, 0.419)	0.226 (0.117, 0.437)	0.250 (0.158, 0.395)	0.308 (0.214, 0.442)	0.195 (0.079, 0.479)	0.064 (0.009, 0.443)

Note. The cut-off value of the nomogram refers to the threshold of its occurrence probability. AUC = Area under the curve; CI = Confidence interval.

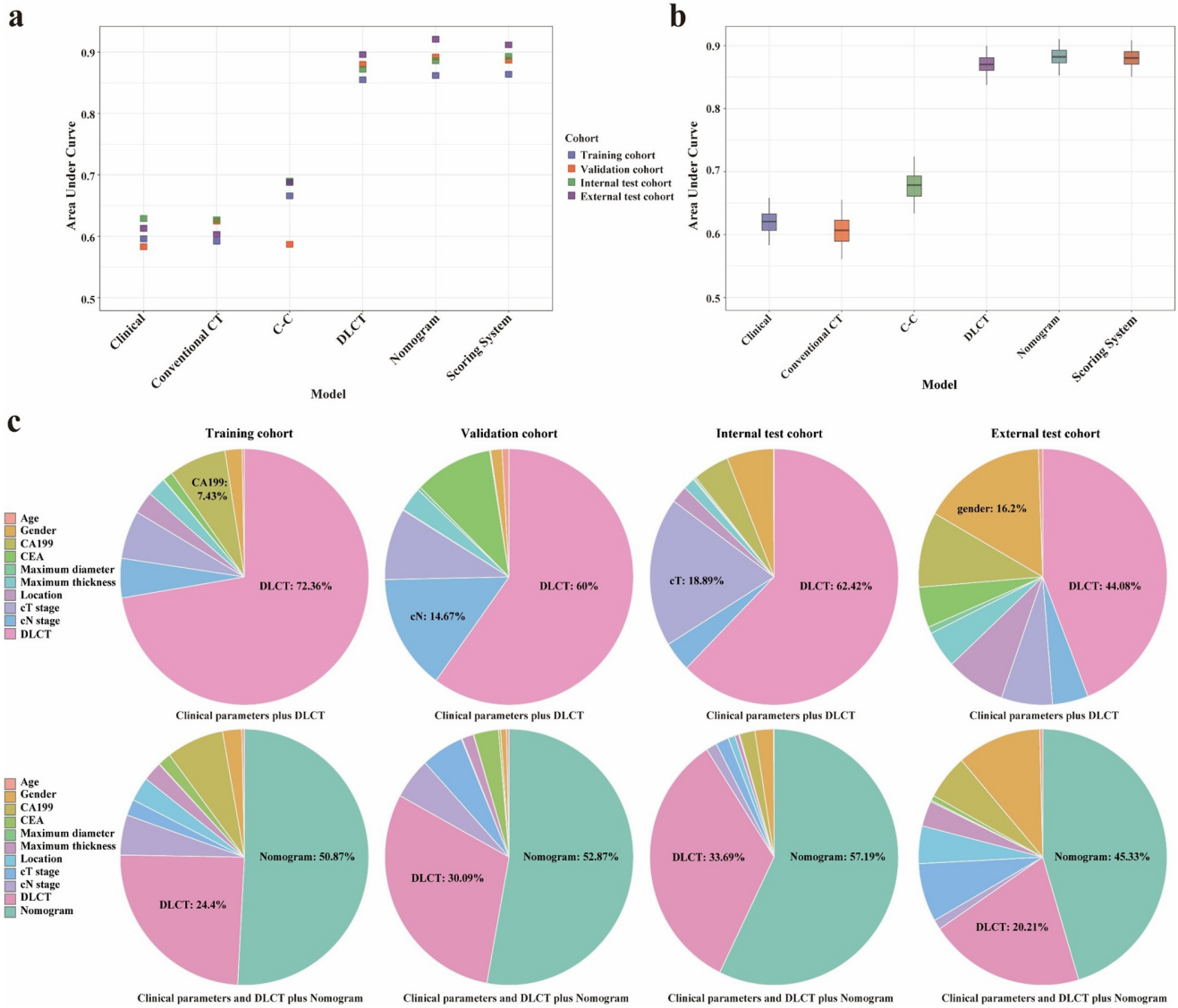

Figure 3. The predictive performance of different models. (a) The area under the receiver operating characteristic (ROC) curves (AUC) for TP53 prediction using clinical, conventional CT, DLCT and combined models in four cohorts. (b) The performance of the models in all patients (n = 568) resampled 1000 times. The center lines within the boxes indicate the mean AUC value. The box bounds represent the interquartile range (IQR) and the whiskers represent the 95% confidence intervals. (c) Relative variable contribution to the prediction by the preoperative clinically measurable indicators, DLCT and the nomogram in four cohorts using the χ^2 proportion test. AUC = Area under the curve; C-C = Clinical-Conventional CT

Discussion

In this study, we developed a simple scoring system based on measured DLCT parameters and clinical features that could allow noninvasive and rapid preoperative prediction of TP53 expression and prognosis in GC patients. We extensively validated the stability of the system in multicenter, retrospective combined prospective cohorts of patients. Critically, the system could identify the subset of patients who could derive a survival benefit from postoperative adjuvant chemotherapy (ACT), assisting the clinic in developing an individualized treatment strategy.

Our study builds on substantial evidence that TP53 expression plays an important role in disease progression and influences treatment response[8,36,37]. Although histological evaluation of specimens remains the criterion, invasive methods such as endoscopic biopsy and surgical resection are limited by sampling bias due to intratumoral spatial heterogeneity and timeliness[38,39]. Currently, abdominal CT is still the first choice for clinical evaluation of GC, with the advantage of low financial cost and high sensitivity. However, conventional CT lacks functional quantitative parameters to assess TP53 type of GC. To address this issue, we therefore utilized DLCT with wide availability, which was a quick, noninvasive and quantitative technique for assessing pathological information and predicting prognosis of tumors. Furthermore, it also could be acquired repeatedly during the treatment and was recognized as a powerful tool for longitudinal monitoring of response to treatment in GC patients.

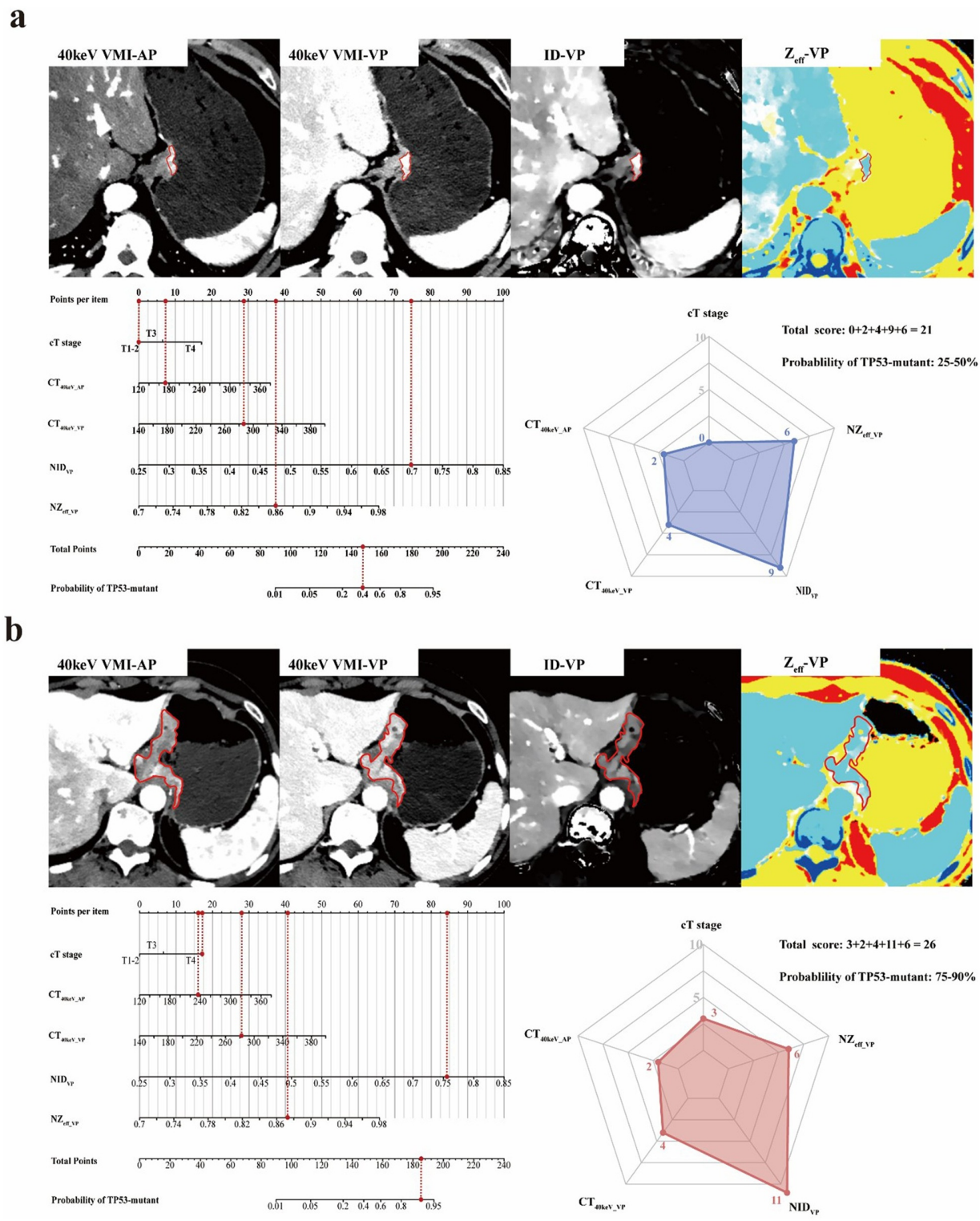

Figure 4. Representative cases and illustration of the nomogram and the scoring system prediction. 40 keV Virtual monoenergetic images (VMI) in the arterial and venous phases, iodine density (ID) and effective atomic number (Z_{eff}) maps in the venous phase in gastric cancer (GC) cases with different TP53 expression. The lesion was outlined on the images (red dashed lines), and calculations were performed by the nomogram and the scoring system (on the nomogram: shown as red circles with dotted lines, on the scoring system: shown as radar chart). (a) A 60-year-old male with GC had the nomogram point of 147.15 (probability, < 0.4539), and the system score of 21. It was predicted to be TP53-wildtype, and the postoperative pathology confirmed the prediction to be accurate. (b) A 51-year-old female with the nomogram point of 184.82 (probability, > 0.4539) and the system score of 26 predicted to experience TP53-mutant GC, and the prediction was proved by postoperative pathology. NID = ID_{tumor}/ID_{aorta}; NZ_{eff} = $Z_{eff\_tumor}/Z_{eff\_aorta}$; AP = Arterial phase; VP = Venous phase; VMI = Virtual monoenergetic images; ID = Iodine density; Z_{eff} = Effective atomic number; NID = Normalized iodine density; NZ_{eff} = Normalized effective atomic number.

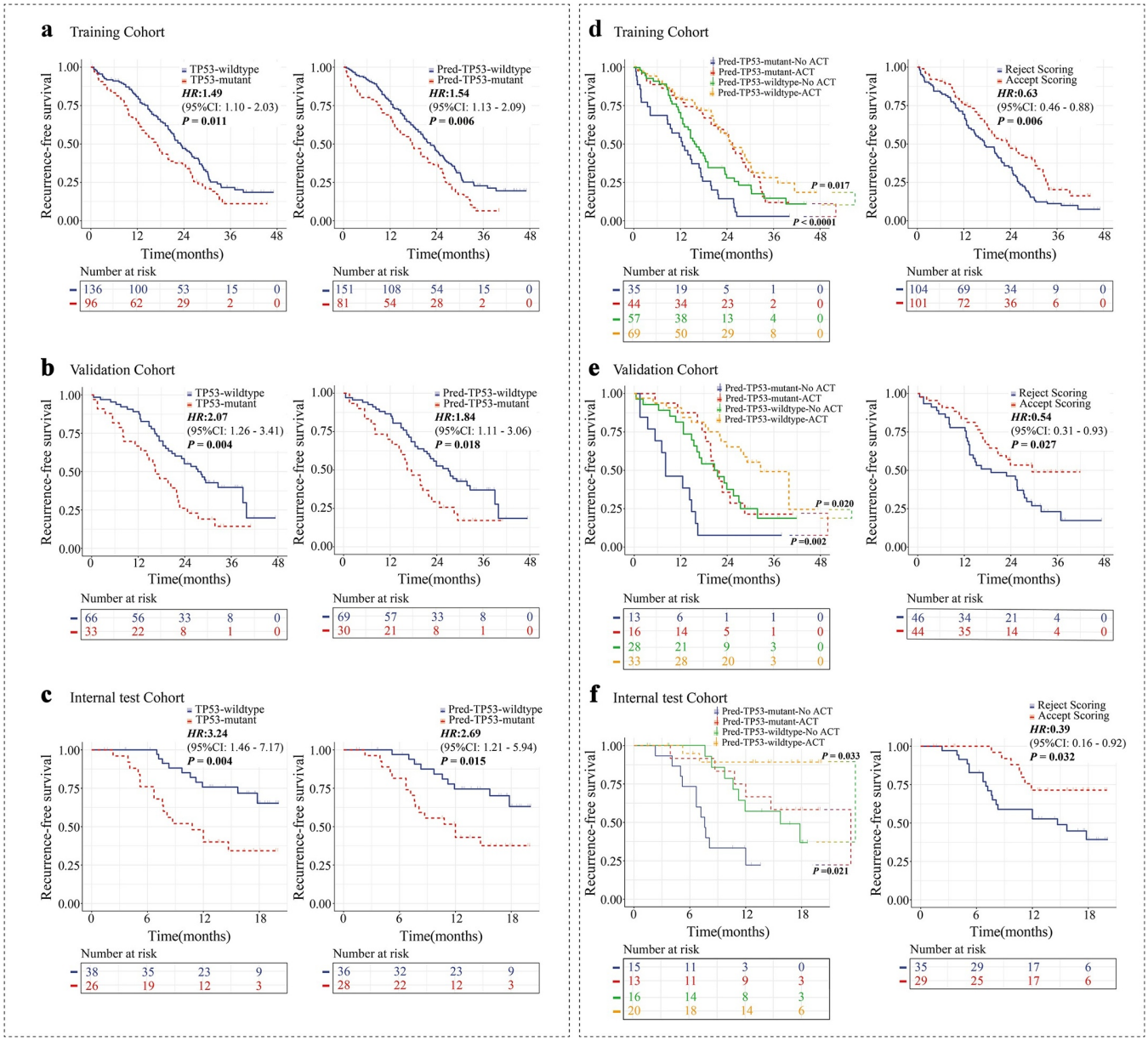

Figure 5. Kaplan-Meier survival curves related to the scoring system. Kaplan-Meier survival curves for the training cohort (a), the validation cohort (b) and the internal test cohort (c). Kaplan-Meier survival curves based on the ground truth label (left) and Kaplan-Meier survival curves based on the predicted label by the scoring system (right). In the training cohort (d), the validation cohort (e) and the internal test cohort (f), Kaplan-Meier survival curves based on the predicted results versus the adjuvant chemotherapy (ACT) in four categories (Predicted TP53-mutant patients without ACT, Predicted TP53-mutant patients with ACT, Predicted TP53-wildtype patients without ACT and Predicted TP53-wildtype patients with ACT) (left). Prognostic improvement in Pred-TP53-mutant patients over Pred-TP53-wildtype patients receiving ACT: training cohort: *HR* = 0.35, *P*<0.0001 vs. *HR* = 0.59, *P* = 0.017; validation cohort: *HR* = 0.27, *P* = 0.002 vs. *HR* = 0.45, *P* = 0.020; internal test cohort: *HR* = 0.18, *P* = 0.021 vs. *HR* = 0.25, *P* = 0.033. Kaplan-Meier survival curves based on two categories (Reject Scoring System and Accept Scoring System) (right). Reject Scoring System: patients who did not receive ACT in accordance with the scoring system recommendation (Predicted TP53-mutant patients without ACT and Predicted TP53-wildtype patients with ACT). Accept Scoring System: patients who received ACT as suggested by the scoring system (Predicted TP53-mutant patients with ACT and Predicted TP53-wildtype patients without ACT). HR = Hazard ratio; CI = Confidence interval; ACT = Adjuvant chemotherapy.

DLCT provides a series of rapidly measured and quantitative functional parameters through a single scan, and reveals subtle relationships among different gene statuses based on the unique attenuation properties of various tissues and their expression characteristics[40,41]. In this study, we interpreted its utility as an effective method to investigate the genetic characteristics of GC at the pathological level. CT_{40keV} value and NID, which intuitively reflected contrast media uptake, were higher in TP53-mutant GC than in wildtype GC, reflecting increased blood volume and elevated vascular permeability in the former. This implied that these parameters identify more active angiogenesis and immature vascular states in the tumor microenvironment of GC[42,43]. We observed that NZ_{eff}, the normalized effective atomic number, which reflected the components of

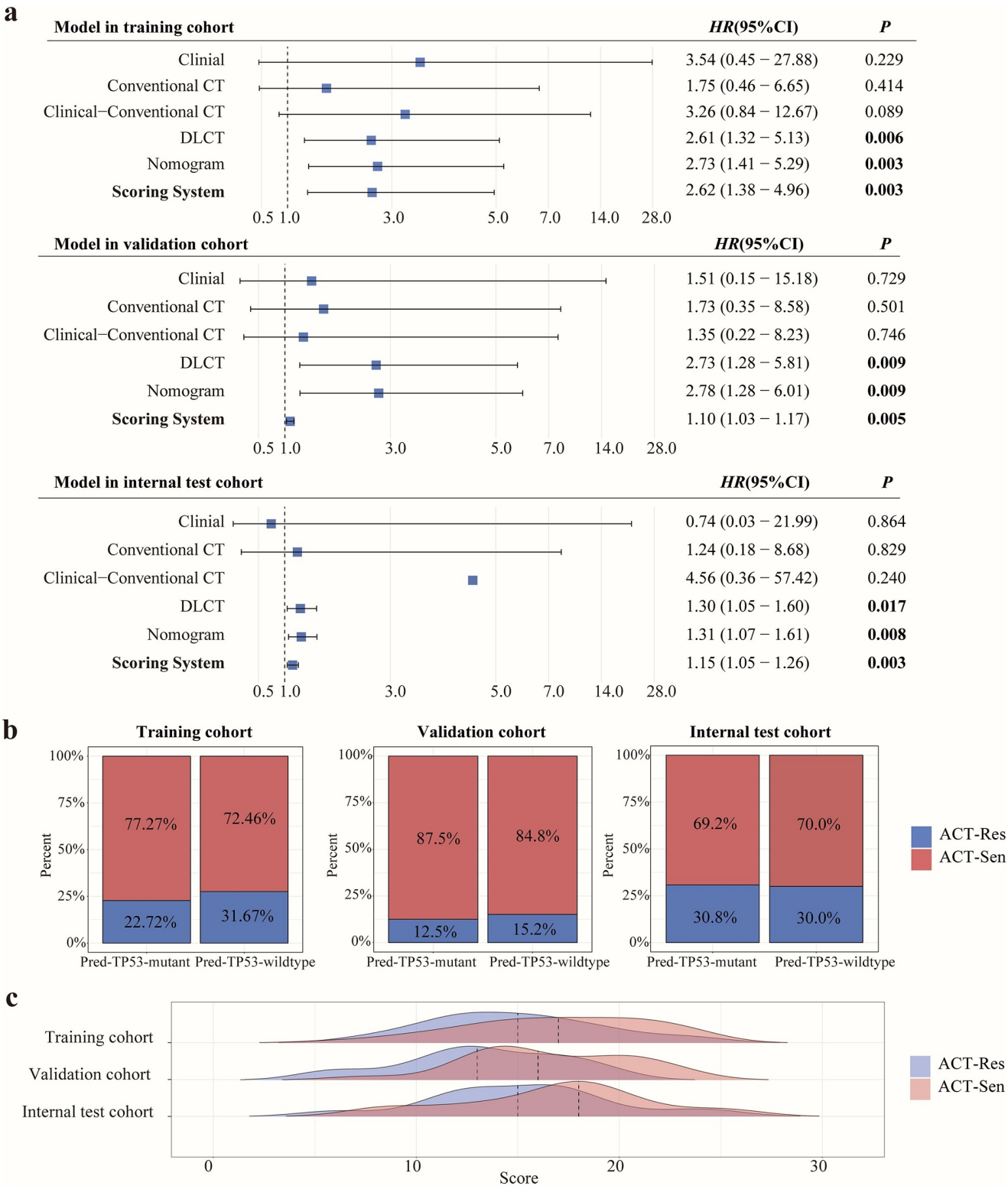

Figure 6. Association between recurrence-free survival and adjuvant chemotherapy regimen in patients with different TP53 expression predicted by the scoring system in the training cohort, the validation cohort and the internal test cohort. (a) Forest plot for impact of different predictive models and the system scores on recurrence-free survival (RFS) of the patients. *P* values are two-tailed from Cox proportional hazards regression analyses. Blue dots indicate the HR value. Error bars represent the 95% confidence intervals. (b) Proportion of the patients with different TP53 predicted by the scoring system for clinical response (resistant or sensitive) to ACT. (c) Ridgeline plot of the system scores for patients resistant or sensitive to adjuvant chemotherapy (ACT). HR = Hazard ratio; CI = Confidence interval; ACT = Adjuvant chemotherapy; ACT-Sen = Adjuvant chemotherapy sensitivity; ACT-Res = Adjuvant chemotherapy resistance.

tumor tissues at the atomic level, was higher in TP53-mutant GC. TP53 mutations promote the growth of cancer-associated fibroblasts, leading to increased cellular load and dense fibrosis[44,45]. Thus, there might be a potential close link between NZ_{eff} and cellular densities in the tumor microenvironment. Disordered neovascularization and growth of CAFs create a tumor microenvironment conductive to the survival of GC cells, which is one of the main hallmarks of the infiltrative features of TP53-mutant GC[46]. This suggests that DLCT is promising in identifying invasive status in the tumor microenvironment, and further exploration of the causal relationship between them might be a breakthrough point in the future. Our study demonstrated that DLCT had the dual advantage of using biological knowledge to guide model predictions from multiple dimensions and capturing image data more purposefully in the clinic. Indeed, as our results showed, the method improved performance over conventional CT-based predictions.

To the best of our knowledge, this is the first simple scoring system based on DLCT for predicting TP53 expression. It allowed direct estimation of TP53 mutation probability by rapid and simple measurement of the scanning images without additional complex processing. We compared the performance of this system with liquid biopsy and AI-driven conventional CT for genetic prediction of GC. For liquid biopsy, first, the sample size included in this study exceeded that of the liquid biopsy by 522, ensuring the robustness of the system[47]. Second, the accuracy of the system was 84.3%, which was superior to that of liquid biopsy (gastric wash analysis, 82.6%; plasma analysis, 80.4%)[47]. Thirdly, DLCT scanning is more operationally standardized and uniform than liquid biopsy[16]. In terms of AI-driven conventional CT, models for TP53 prediction were rare. Despite similar AUCs for GC tumor microenvironment assessment between them and the scoring system in this study, their interpretability remains a major challenge for high-risk applications such as clinical treatment decision-making[17,18]. Subsequently, AI models suffer from the potential drawbacks of relative operational complexity and workload[19]. Furthermore, the perfect deployment of advanced hardware and software is costly, both for liquid biopsy and AI-driven conventional CT[48,49]. Instead, we have not only explained the parameters of the scoring system in detail and scientifically at the pathology level but also confirmed the obvious economic benefits of this system compared with other techniques through cost-effectiveness analysis. The scoring system has been validated in a multicenter, especially prospective, broad population and is generalizable across different groups. More importantly, its nature of memorization and simple computability allows it applicable in a wider range of clinical scenarios to achieve rapid decision-making in routine practice. Especially given the relatively low cost, it may help underserved populations due to high costs or limited resources.

The significant correlation between the system scores and RFS in GC suggests that the system has a favorable stratification ability for prognosis. Given its high image quality, the system could be used to refine the current staging system and improve risk stratification for GC. Additionally, current guidelines recommend ACT for some stage Ib, stage II and stage III patients[34,50], but the optimal criteria for selecting candidates remain controversial[32,51,52]. Our results showed that the system was useful in predicting postoperative ACT benefit for GC. Pred-TP53-mutant patients based on the scoring system would benefit more significantly from postoperative ACT than wildtype patients in our study, in consistent with previous findings regarding the benefit of ACT based on molecular approaches[32,53]. In other words, more intensive therapies were needed for the Pred-TP53-wildtype patients to enhance efficacy and improve their prognosis[54]. We also found that the median time of RFS was more than 6.2 months longer in patients who accepted the treatment decisions recommended by the scoring system than those who did not. This finding has clinical implications and could complement current treatment guidelines by comprehensively screening for more appropriate ACT patient candidates. In brief, physicians should focus on patients with scores >21.5, recommend them for ACT, and concentrate on following up this subgroup to adjust treatment timely. Overall, the introduction of personalized guidance based on the scoring system has the potential to provide patients with significant survival benefits as well as more comprehensive and effective medical support.

This study is mainly limited by the method of manually outlining the regions of interest, which may involve several subjective factors. Due to the inherent nature of GC, such as irregular shape and unclear boundaries in CT images, manual outlining may be subject to human bias. Semi-automated and automated segmentation techniques have been developed. In the future, methods that combine such segmentation with manual outlining techniques may help to improve the scoring system efficacy[55]. Although we focused on the specific TP53 gene in this work, the proposed framework is general and can be applied to the study of other biological information to design interpretable prediction models. AI-driven CT could dig deeper for more information about the tumor microenvironment of GC[55]. In future work, it may be beneficial to combine the two complementary approaches to further improve the performance and scalability of the scoring system. In addition, we will make efforts to develop the system as software tools so that they could be embedded in radiology reports to form automated calculations, further facilitating the clinical utility of the scoring system.

In conclusion, we presented a noninvasive and simple scoring system for the prediction of TP53 expression using DLCT, which can predict GC prognosis and response to ACT by rapid calculation. This system provides an objective imaging basis for personalized cancer therapy.

Ethical approval

1. The institutional review boards of the First Affiliated Hospital of Fujian Medical University ([2022] No. 334).
2. Zhongshan Hospital Affiliated to Xiamen University (No. 2022-082).

Consent

The training cohort and the validation cohort recruited patients retrospectively and the informed consent was waived. The internal test cohort and the external test cohort collected patients with informed consent prospectively.

Sources of funding

Joint Funds for the Innovation of Science and Technology, Fujian Province (No. 2021Y9141). The funding sources had no role in the study design, data collection, data analysis, data

interpretation, writing of the manuscript, or decision to submit it for publication.

Author contributions

Guarantors of integrity of entire study, Y.W., Y.L., D.C., D.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, Y.W., Y.L., M.X.; clinical studies, Y.W., Y.L., M.X., N.K., M.W., Y.Q.W., H.Z., C.M.; experimental studies, M.X., N.K.; statistical analysis, Y.W., Y.L., M.X., N.K.; and manuscript editing, Y.W., Y.L., D.C., D.S.

Conflicts of interest disclosure

All authors declare that there are no conflicts of interest.

Research Registration Unique Identifying Number (UIN)

researchregistry11026.

Guarantor

Yinchen Wu, Yu Lin, Dairong Cao, Dejun She.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

Due to the privacy of patients, the data related to patients cannot be available for public access but can be obtained from the corresponding author on reasonable request approved by the institutional review board of all enrolled centers.

References

[1] Smyth EC, Nilsson M, Grabsch HI, van Grieken NCT, Lordick F. Gastric cancer. Lancet 2020;396:635–48.

[2] Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev 2014;23:700–13.

[3] Yu P, Wang Y, Yu Y, *et al*. Deep targeted sequencing and its potential implication for cancer therapy in chinese patients with gastric adenocarcinoma. Oncologist 2021;26:e756–e68.

[4] Huang S-C, Chang IY-F, Chen T-C, *et al*. Redefining aberrant P53 expression of gastric cancer and its distinct clinical significance among molecular-histologic subtypes. Asian J. Surg. 2024;47:4699–705.

[5] Sawada T, Yashiro M, Sentani K, *et al*. New molecular staging with G-factor supplements TNM classification in gastric cancer: a multicenter collaborative research by the Japan Society for gastroenterological carcinogenesis g-project committee. Gastric Cancer 2015;18:119–28.

[6] Bria E, De Manzoni G, Beghelli S, *et al*. A clinical–biological risk stratification model for resected gastric cancer: prognostic impact of Her2, Fhit, and APC expression status. Ann Oncol 2013;24:693–701.

[7] Pinto-de-sousa J, Silva F, David L, *et al*. Clinicopathological significance and survival influence of p53 protein expression in gastric carcinoma. Histopathology 2004;44:323–31.

[8] Ando K, Nakamura Y, Kitao H, *et al*. Mutational spectrum of TP53 gene correlates with nivolumab treatment efficacy in advanced gastric cancer (TP53MUT study). Br. J. Cancer 2023;129:1032–39.

[9] Wang K, Gong Z, Chen Y, *et al*. KDM4C-mediated senescence defense is a targetable vulnerability in gastric cancer harboring TP53 mutations. Clin Epigenetics 2023;15:163.

[10] Cancer CA-CACoG. Chinese expert consensus on the clinical application of next-generation sequencing for gastric cancer. Chin. J. Clin. Oncol 2023;50:309–18.

[11] Stockton S, Soares HP, Dayyani F, *et al*. A phase 2 single-arm study of berzosertib in combination with irinotecan in patients with progressive TP53 mutant gastric and gastro-esophageal junction cancer. J Clin Oncol 2023;41:4044.

[12] Yu R, Sun T, Zhang X, *et al*. TP53 co-mutational features and NGS-calibrated immunohistochemistry threshold in gastric cancer. Onco Targets Ther 2021;14:4967–78.

[13] Gullo I, Carneiro F, Oliveira C, Almeida GM. Heterogeneity in gastric cancer: from Pure morphology to molecular classifications. Pathobiology 2018;85:50–63.

[14] Kim J, Park C, Kim KH, *et al*. Single-cell analysis of gastric pre-cancerous and cancer lesions reveals cell lineage diversity and intratumoral heterogeneity. Npj Precision Oncology 2022;6:9.

[15] Chia NY, Tan P. Molecular classification of gastric cancer. Ann Oncol 2016;27:763–69.

[16] Ma S, Zhou M, Xu Y, *et al*. Clinical application and detection techniques of liquid biopsy in gastric cancer. Mol Cancer 2023;22:7.

[17] Ma T, Wang H, Ye Z. Artificial intelligence applications in computed tomography in gastric cancer: a narrative review. Transl Cancer Res 2023;12:2379–92.

[18] Ebigbo A, Messmann H, and Lee SH. Artificial intelligence applications in image-based diagnosis of early esophageal and gastric neoplasms. Gastroenterology 2025. doi:10.1053/j.gastro.2025.01.253.

[19] Chen Q, Zhang L, Liu S, *et al*. Radiomics in precision medicine for gastric cancer: opportunities and challenges. EUR RADIOL 2022;32:5852–68.

[20] Halligan S, Mallett S. Letter to the Editor: "Obsolescence of nomograms in radiomics research". EUR RADIOL 2024;34:5052–53.

[21] Rassouli N, Etesami M, Dhanantwari A, Rajiah P. Detector-based spectral CT with a novel dual-layer technology: principles and applications. Insights into Imaging 2017;8:589–98.

[22] Forghani R, De Man B, Gupta R. Dual-energy computed tomography. Neuroimaging Clin. N. Am 2017;27:385–400.

[23] Adam SZ, Rabinowich A, Kessner R, and Blachar A. Spectral CT of the abdomen: where are we now? Insights into Imaging 2021;12:138.

[24] Drljevic-Nielsen A, Mains JR, Thorup K, Andersen MB, Rasmussen F, Donskov F. Early reduction in spectral dual-layer detector CT parameters as favorable imaging biomarkers in patients with metastatic renal cell carcinoma. EUR RADIOL 2022;32:7323–34.

[25] Liu -J-J, Liu W, Jin Z-Y, *et al*. Improved visualization of gastric cancer and increased diagnostic performance in lesion depiction and depth identification using monoenergetic reconstructions from a novel dual-layer spectral detector CT. Acad. Radiol. 2020;27:e140–e7.

[26] Du M, Wang X, Zhuang S, *et al*. Quantitative parameters in novel spectral computed tomography for assessing gastric cancer and cell proliferation. Eur J Radiol 2023;167:111052.

[27] Li J, Xu S, Wang Y, *et al*. Spectral CT-based nomogram for preoperative prediction of perineural invasion in locally advanced gastric cancer: a prospective study. EUR RADIOL 2023;33:5172–83.

[28] Luo M, Chen G, Xie H, *et al*. Preoperative diagnosis of metastatic lymph nodes by CT-histopathologic matching analysis in gastric adenocarcinoma using dual-layer spectral detector CT. EUR RADIOL 2023;33:8948–56.

[29] Mao LT, Chen WC, Lu JY, *et al*. Quantitative parameters in novel spectral computed tomography: assessment of Ki-67 expression in patients with gastric adenocarcinoma. World J Gastroenterol 2023;29:1602–13.

[30] Agha RA, Mathew G, Rashid R, *et al*. Revised Strengthening the reporting of cohort, cross-sectional and case-control studies in surgery (STROCSS) guideline: an update for the age of artificial intelligence. Pjs 2025;10:100081.

[31] Pretzsch E, Bosch F, Todorova R, *et al*. Molecular subtyping of gastric cancer according to ACRG using immunohistochemistry - Correlation with clinical parameters. Pathol Res Pract 2022;231:153797.

[32] Gu Y, Sun M, Fang H, *et al*. Impact of clonal TP53 mutations with loss of heterozygosity on adjuvant chemotherapy and immunotherapy in gastric cancer. Br. J. Cancer 2024;131:1320–27.
[33] Amin MB. American Joint Committee on Cancer, American Cancer Society. AJCC Cancer Staging Manual. Eight edition/editor-in-chief, Mahul B. Amin, MD, FCAP; editors, Stephen B. EdgeMD, FACS and, 16 others; Donna M. Gress, RHIT, CTR - Technical editor; Laura R. Meyer, CAPM - Managing editor Chicago IL, American Joint Committee on Cancer, Springer;2017. Vol xvii. 1024
[34] 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;44:127–72.
[35] Zhu Y, Feng B, Cai W, *et al*. Prediction of microvascular invasion in solitary AFP-negative hepatocellular carcinoma ≤ 5 cm using a combination of imaging features and quantitative dual-layer spectral-detector CT parameters. Acad. Radiol. 2023;30:S104–S16.
[36] Bataille F, Rümmele P, Dietmaier W, Gaag D, Hartmann A. Alterations in p53 predict response to preoperative high dose chemotherapy in patients with gastric cancer. J Clin Pathol Mol Pathol 2003;56:286.
[37] Yashiro M, Inoue T, Nishioka N, Matsuoka T, Boland CR, Hirakawa K. Allelic Imbalance at p53 and microsatellite instability are predictive markers for resistance to chemotherapy in gastric carcinoma. Ann Surg Oncol 2009;16:2926–35.
[38] Yuan Y. Spatial heterogeneity in the tumor microenvironment. Cold Spring Harbor Perspectives in Medicine 2016;6:a026583.
[39] 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 multicenter cohort study. eClinicalMedicine 2022;46:101348.
[40] Segal E, Sirlin CB, Ooi C, *et al*. Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol 2007; 25:675–80.
[41] Ota T, Onishi H, Fukui H, *et al*. Prediction models for differentiating benign from malignant liver lesions based on multiparametric dual-energy non-contrast CT. EUR RADIOL 2025;35:1361–77.
[42] Sumiyoshi Y, Kakeji Y, Egashira A, Mizokami K, Orita H, Maehara Y. Overexpression of hypoxia-inducible factor 1α and p53 is a marker for an unfavorable prognosis in gastric cancer. Clin Cancer Res 2006;12: 5112–17.
[43] Miyamoto N, Yamamoto H, Taniguchi H, *et al*. Differential expression of angiogenesis-related genes in human gastric cancers with and those without high-frequency microsatellite instability. Cancer Lett 2007;254:42–53.
[44] Procopio M-G, Laszlo C, Al Labban D, *et al*. Combined CSL and p53 downregulation promotes cancer-associated fibroblast activation. Nat Cell Biol 2015;17:1193–204.
[45] Mak TK, Li X, Huang H, *et al*. The cancer-associated fibroblast-related signature predicts prognosis and indicates immune microenvironment infiltration in gastric cancer. Front Immunol 2022;13:951214.
[46] Arandkar S, Furth N, Elisha Y, *et al*. Altered p53 functionality in cancer-associated fibroblasts contributes to their cancer-supporting features. Proc Natl Acad Sci 2018;115:6410–15.
[47] Pizzi MP, Bartelli TF, Pelosof AG, *et al*. Identification of DNA mutations in gastric washes from gastric adenocarcinoma patients: possible implications for liquid biopsies and patient follow-up. Int J Cancer 2019;145:1090–98.
[48] Westwood M, Ramaekers B, Grimm S, *et al*. Software with artificial intelligence-derived algorithms for analysing CT brain scans in people with a suspected acute stroke: a systematic review and cost-effectiveness analysis. Health Technol Assess 2024;28:1–204.
[49] Kapoor R, So JBY, Zhu F, Too H-P, Yeoh K-G, Yoong JS-Y. Evaluating the use of microRNA blood tests for gastric cancer screening in a stratified population-level screening program: an early model-based cost-effectiveness analysis. Value Health 2020;23:1171–79.
[50] Ajani JA, D'Amico TA, Bentrem DJ, *et al*. Gastric Cancer, Version 2.2022, NCCN clinical practice guidelines in oncology. J National Compr Cancer Network 2022;20:167–92.
[51] Boku N, Ohtsu A, Yoshida S, *et al*. Significance of biological markers for predicting prognosis and selecting chemotherapy regimens of advanced gastric cancer patients between continuous infusion of 5-FU and a combination of 5-FU and cisplatin. Jpn J Clin Oncol 2007;37:275–81.
[52] Boku N, Chin K, Hosokawa K, *et al*. Biological markers as a predictor for response and prognosis of unresectable gastric cancer patients treated with 5-fluorouracil and cis-platinum. Clin Cancer Res 1998;4:1469–74.
[53] Nie K, Zheng Z, Wen Y, *et al*. Construction and validation of a TP53-associated immune prognostic model for gastric cancer. Genomics 2020;112:4788–95.
[54] Shitara K, Van Cutsem E, Bang Y-J, *et al*. Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer. JAMA Oncology 2020;6:1571–80.
[55] Jiang Y, Zhang Z, Wang W, *et al*. Biology-guided deep learning predicts prognosis and cancer immunotherapy response. Nat Commun 2023;14:5135.