Received: 4 April 2023 | Revised: 20 April 2023 | Accepted: 1 May 2023

DOI: 10.1002/jgm.3529

RESEARCH ARTICLE

WILEY

A combined diagnostic model based on circulating tumor cell in patients with solitary pulmonary nodules

Dong Wang[1] | Peng Li[1] | Xiang Fei[2] | Shuyu Che[3] | Jinlong Li[1] | Yunpeng Xuan[1] | Jinglong Wang[1] | Yudong Han[1] | Weiqing Gu[4] | Yongjie Wang[1]

[1]Department of Thoracic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, China

[2]Department of Thoracic Surgery, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai, China

[3]Department of Respiratory and Critical Care Medicine, The Affiliated Hospital of Qingdao University, Qingdao, China

[4]Department of Oncology, Shanghai Pulmonary Hospital Affiliated to Tongji University, Shanghai, China

Correspondence
Yongjie Wang, Department of Thoracic Surgery, The Affiliated Hospital of Qingdao University, No. 59 Haier Road, Laoshan District, Qingdao 266000, China.
Email: yjwang0408@126.com; wangyongjie@qduhospital.cn

Weiqing Gu, Department of Oncology, Shanghai Pulmonary Hospital Affiliated to Tongji University, 507 Zhengmin Road, Shanghai 200433, China.
Email: jessiegwq@163.com

Funding information
Wu Jieping Medical Foundation, Grant/Award Number: 320.6750.2021-01-4; The Affiliated Hospital of Qingdao University "Clinical Medicine+X" Scientific Research Project, Grant/Award Number: QYFY-X2021032/3731

Abstract

Background: Although many prediction models in diagnosis of solitary pulmonary nodules (SPNs) have been developed, few are widely used in clinical practice. It is therefore imperative to identify novel biomarkers and prediction models supporting early diagnosis of SPNs. This study combined folate receptor-positive circulating tumor cells (FR^+CTC) with serum tumor biomarkers, patient demographics and clinical characteristics to develop a prediction model.

Methods: A total of 898 patients with a solitary pulmonary nodule who received FR^+CTC detection were randomly assigned to a training set and a validation set in a 2:1 ratio. Multivariate logistic regression was used to establish a diagnostic model to differentiate malignant and benign nodules. The receiver operating curve (ROC) and the area under the curve (AUC) were calculated to assess the diagnostic efficiency of the model.

Results: The positive rate of FR^+CTC between patients with non-small cell lung cancer (NSCLC) and benign lung disease was significantly different in both the training and the validation dataset ($p < 0.001$). The FR^+CTC level was significantly higher in the NSCLC group compared with that of the benign group ($p < 0.001$). FR^+CTC (odds ratio, OR, 95% confidence interval, CI: 1.13, 1.07–1.19, $p < 0.0001$), age (OR, 95% CI: 1.06, 1.01–1.12, $p = 0.03$) and sex (OR, 95% CI: 1.07, 1.01–1.13, $p = 0.01$) were independent risk factors of NSCLC in patients with a solitary pulmonary nodule. The area under the curve (AUC) of FR^+CTC in diagnosing NSCLC was 0.650 (95% CI, 0.587–0.713) in the training set and 0.700 (95% CI, 0.603–0.796) in the validation set, respectively. The AUC of the combined model was 0.725 (95% CI, 0.659–0.791) in the training set and 0.828 (95% CI, 0.754–0.902) in the validation set, respectively.

Conclusions: We confirmed the value of FR^+CTC in diagnosing SPNs and developed a prediction model based on FR^+CTC, demographic characteristics, and serum biomarkers for differential diagnosis of solitary pulmonary nodules.

Abbreviations: AGR, albumin/globulin ratio; AUC, the area under the curve; CA125, carbohydrate antigen 125; CEA, carcinoembryonic antigen; CTCs, circulating tumor cells; CYFRA21-1, cytokeratin 19 fragment; ECOG, Eastern Cooperative Oncology Group; $EpCAM^+$, epithelial cell adhesion molecule positive; FRα, folate receptor alpha; FR^+CTC, folate receptor-positive circulating tumor cells; IQR, interquartile range; NSCLC, non-small cell lung cancer; NSE, neuron-specific enolase; ProGRP, pro-gastrin releasing peptide; ROC, receiver operating curve; SCCA, squamous cell carcinoma antigen; SD, standard deviation; SPNs, solitary pulmonary nodules.

Dong Wang, Peng Li and Xiang Fei contributed equally.

KEYWORDS
biomarker, circulating tumor cell, diagnose, non-small cell lung cancer (NSCLC), prediction model, solitary pulmonary nodule

1 | INTRODUCTION

Lung cancer is the most commonly diagnosed cancer (18.1%) and the leading cause of cancer death (24.1%) in China.[1] Non-small cell lung cancer (NSCLC) is the most common type of lung cancer. Early diagnosis is key to improving the long-term prognosis for NSCLC patients. Low-dose computed tomography (LDCT) has been systematically adopted for routine lung cancer screening in high-risk populations. In 2015, only about 5% of high-risk adults received LDCT screening[2] and that proportion increased to 15–20% in 2017 and 2018.[3,4] While the penetration of LDCT screening remains unsatisfactory, the increasing number of screenings has led to a substantial increase in the identification of pulmonary nodules, most of which are solitary nodules. Unfortunately, studies have shown that LDCT screening has a false-positive rate of 81–93%.[5] Additional testing or imaging is needed for the population to confirm the diagnostic result.

Solitary pulmonary nodule (SPN) is a single lung opacity less than 3 cm in diameter. Most of them are benign, such as tuberculoma and pulmonary hamartoma. Malignant nodules are relatively rare and mostly primary lung cancer. The most common pathological types of malignant SPNs are adenocarcinoma and squamous cell carcinoma.[6–8] However, both types of nodule share similar imaging features, such as lobulated and spiculated margins.[9,10] In clinical practice, differentiating malignant from benign nodules using conventional imaging methods alone has proved challenging, with false-positive and false-negative rates being 75 and 48%, respectively.[11]

Serum biomarkers have many advantages over tissue-based tests because they are non-invasive, easily repeatable and cost effective. Nevertheless, they have low sensitivities in diagnosing malignancies yet high false-positive rates in benign tumors or infections.[12] The utilities of single serum biomarkers in NSCLC diagnosis are thus limited and clinical guidelines generally recommend using combinations of serum biomarkers to improve the detection efficiency.[13] Although a large number of prediction models have been developed, few are widely used in clinical practice.[14,15] It is therefore imperative to identify novel biomarkers and prediction models to support early diagnosis of NSCLC.

Circulating tumor cells (CTCs) are tumor cells that shed into the circulatory system from primary or metastatic tumors. Studies have shown that CTCs can be used as diagnostic and prognostic markers for various types of cancer.[16–18] CellSearch is the most commonly used platform to quantify CTCs of epithelial origin by enriching epithelial cell adhesion molecule positive ($EpCAM^+$) cells in the blood.[19] Other non-$EpCAM^+$ biomarker-based approaches for CTC quantification have also been developed and investigated.[20]

Folate receptor alpha (FRα) is a glycoprotein that is anchored to the membrane of normal epithelial cells, and highly expressed in a variety of solid tumors.[21–23] For instance, the level of FRα expression is significantly upregulated in 75.7% of patients with NSCLC.[24] As such, FRα has aroused significant interest as a potential target for cancer diagnosis and treatment. A folate-integrated magnetic polymer micelle for MRI has been developed and the targeted tracer accumulates in FR-expressing tumor tissues, resulting in improved sensitivity.[25] Farletuzumab is a fully humanized IgG1 antibody and exerts its activity against FRα-positive cancer cells via multiple modes of action.[26,27]

An FR-based CTC detection has been developed, and the related FR-positive CTC (FR^+CTC) detection kit has been approved by the CFDA for clinical use. FR^+CTCs have a high sensitivity (73.2–81.8%) and specificity (84.1–93.2%) for the diagnosis of lung cancer.[28,29] The dynamic change in FR^+CTC level can also predict the outcome of NSCLC patients who have received EGFR-TKI and chemotherapy treatment.[30] Although FR^+CTC for the diagnosis of SPNs has been examined in a small prospective study,[30] the utility of FR^+CTC level combination with demographics, clinical characteristics and tumor biomarkers to build a diagnostic model in NSCLC patients with SPNs was not reported.

In this study, we aimed to explore the expression of peripheral blood FR^+CTCs and establish a diagnostic model based on FR^+CTCs, common clinical characteristics and serum biomarkers in patients with SPNs.

2 | MATERIALS AND METHODS

2.1 | Study design

This is a retrospective study conducted in The Affiliated Hospital of Qingdao University from January 2019 to December 2020. A total of 898 patients who received FR^+CTC detection were enrolled in this study. The inclusion criteria were: (1) FR^+CTC detection before anticancer therapy; (2) tests for at least three of these six serum biomarkers (carcinoembryonic antigen, CEA; cytokeratin 19 fragment, CYFRA21-1; neuron-specific enolase, NSE; carbohydrate antigen 125, CA125; pro-gastrin releasing peptide, ProGRP; squamous cell carcinoma antigen, SCCA) before anticancer therapy; and (3) definitive pathological evaluation of diseased tissue obtained through tissue biopsy or surgical resection. Exclusion criteria were: (1) non-pulmonary neoplasm; and (2) multiple nodules. The ethics committees of the Affiliated Hospital of Qingdao University approved the study (no. QYFYWZLL26937).

2.2 | Data collection

Age, gender, Eastern Cooperative Oncology Group (ECOG) score, complications, pathological diagnosis, maximum diameter of lesion, albumin/globulin ratio (AGR), values of FR^+CTC and serum biomarkers were collected. The serum biomarkers tested included CEA, SCCA, NSE, CYFRA21-1, ProGRP and CA125.

2.3 | FR^+CTC analysis

Three milliliters of peripheral blood sample were collected using an EDTA anti-coagulant vacuum tube. Samples were stored at 4°C and processed within 24 h. FR^+CTC detection was performed using a folate receptor-positive cell detection kit (Genosaber, Shanghai, China) as previously described.[29] In brief, FR^+CTCs were first enriched by the negative enrichment method, in which erythrocytes were first lysed by a lysing buffer and then leukocytes were depleted by a combination of anti-CD45 and anti-CD14 immunomagnetic beads. The enriched FR^+CTCs were then labelled by an FRα-targeting probe which contained the conjugate of folic acid and a synthesized oligonucleotide. The labelled FR^+CTCs were enumerated by quantitative PCR using the proprietary ligand-targeted PCR method.[31] A series of standards containing oligonucleotides ranging from 10^{-14} to 10^{-9} M was used for FR^+CTC quantification, representing the 2–2 × 10^5 FU/3 mL blood. "FU" was defined as the FR^+CTCs number.

2.4 | Serum biomarker analysis

A 3 mL blood sample was obtained from the patients. After centrifuging at 800–1,000 rpm for 10 min, the serum was collected. Subsequent serum biomarker testing including CEA, CYFRA21-1, NSE, CA125, ProGRP and SCCA, was performed, using the enzyme-linked immune-sorbent assay method (Roche Diagnostics, Shanghai, China). The ratio of albumin/globulin was detected using a Beckman Coulter AU5800 chemistry analyzer (Beckman Coulter, Brea, CA, USA).

2.5 | Statistical analysis

Continuous variables of normal distribution were expressed as mean ± standard deviation (SD), and continuous variables of abnormal distribution were expressed as median and interquartile range (IQR). Continuous variables were compared using the Mann–Whitney test between two groups or Kruskal–Wallis test among three or more groups. Categorical variables were expressed as counts and percentages. Categorical variables were compared using the chi-square test or Fisher's exact probability method. Binary logistic regression analysis was performed to establish a diagnostic prediction model. The receiver operating characteristic (ROC) curve was plotted and the area under the ROC curve (AUROC) was calculated to examine the diagnostic efficiency of the model. Patients were randomly assigned to the training set and validation set in a 2:1 ratio using the "caret" package of R software (https://CRAN.R-project.org/package=caret). Statistical analysis was performed using R 3.6.4 and GraphPad Prism 8. All *p*-values were based on two-sided testing. A value of $p < 0.05$ was considered to be statistically significant.

3 | RESULTS

3.1 | Characteristics of patients

A total of 898 patients with solitary pulmonary nodules were included in this study. Eight hundred and four were diagnosed with NSCLC and 94 with benign lung disease. Clinicopathologic characteristics of the training ($n = 599$) and validation ($n = 299$) cohorts are displayed in Table 1. Some 47.1% of the patients were older than 60 years old and 61.5% were females. The diameter of the lesions, benign lesions in the lung, cardiac disease, hypertension, diabetes mellitus and ECOG score were not statistically significantly different between the malignant and benign groups in either the training or the validation dataset. The positive rate (9.650 FU/3 mL is considered as a cutoff point and ≥9.650 is considered as positive, the positive rate being equal to the number of positives/the total number within the same group) of FR^+CTC between patients with NSCLC and benign disease was statistically significant in both training and validation datasets ($p < 0.001$). The AGR and NSE between patients with NSCLC and benign disease were statistically significant in the training dataset but that was not the case in the validation dataset (Table 1).

3.2 | Levels of FR^+CTCs in peripheral blood

The median and IQR were used to summarize the FR^+CTC levels in the NSCLC group and the benign group. FR^+CTC levels in the NSCLC group and the benign group were 11.60 (8.00, 15.70) FU/3 mL and 8.50 (6.85, 11.50) FU/3 mL, respectively (Figure 1). The Mann–Whitney test showed that the FR^+CTC level was significantly different between the two groups ($p < 0.001$).

3.3 | Univariate and multivariate analysis

To assess the risk factors affecting diagnosis between benign and malignant tumors, univariate and multivariate binary logistic regression analysis was performed. Age, sex, lesion diameter, ECOG score, AGR, common complications (benign lesions in the lung, cardiac disease, hypertension and diabetes mellitus), FR^+CTC and serum tumor biomarkers including CEA, CA125, SCCA, CYFRA21-1, ProGRP and NSE were used as independent variables (<cutoff value = 0, ≥cutoff value = 1) in the logistic regression model. In univariate analysis, age, sex, AGR, FR^+CTC and NSE were significantly different between the benign and the NSCLC groups (all $p<0.05$). The CEA,

TABLE 1 Characteristics of patients.

	Training set			Validation set		
	Benign ($N = 70$)	Malignant ($N = 529$)	*p*-Value	Benign ($N = 24$)	Malignant ($N = 275$)	*p*-Value
Age (years)			0.022			0.18
<60	47 (67.1%)	274 (51.8%)		16 (66.7%)	138 (50.2%)	
≥60	23 (32.9%)	255 (48.2%)		8 (33.3%)	137 (49.8%)	
Sex			0.01			<0.001
Male	38 (54.3%)	198 (37.4%)		20 (83.3%)	90 (32.7%)	
Female	32 (45.7%)	331 (62.6%)		4 (16.7%)	185 (67.3%)	
Diameter			0.19			0.95
<15 mm	36 (51.4%)	230 (43.5%)		11 (45.8%)	123 (44.7%)	
≥15 mm	32 (45.7%)	297 (56.1%)		12 (50.0%)	152 (55.3%)	
Missing	2 (2.9%)	2 (0.4%)		1 (4.2%)	0 (0%)	
Benign lesions in the lung			0.81			0.16
No	64 (91.4%)	492 (93.0%)		22 (91.7%)	268 (97.5%)	
Yes	6 (8.6%)	37 (7.0%)		2 (8.3%)	7 (2.5%)	
Cardiac disease			0.71			0.62
No	69 (98.6%)	514 (97.2%)		24 (100%)	260 (94.5%)	
Yes	1 (1.4%)	15 (2.8%)		0 (0%)	15 (5.5%)	
Hypertension			0.39			0.12
No	56 (80.0%)	394 (74.5%)		21 (87.5%)	194 (70.5%)	
Yes	14 (20.0%)	135 (25.5%)		3 (12.5%)	81 (29.5%)	
Diabetes mellitus			1			1
No	63 (90.0%)	478 (90.4%)		22 (91.7%)	246 (89.5%)	
Yes	7 (10.0%)	51 (9.6%)		2 (8.3%)	29 (10.5%)	
ECOG score			0.68			0.79
0	63 (90.0%)	465 (87.9%)		22 (91.7%)	237 (86.2%)	
1	6 (8.6%)	58 (11.0%)		2 (8.3%)	36 (13.1%)	
2	1 (1.4%)	6 (1.1%)		0 (0%)	2 (0.7%)	
$FR^{+}CTC$ (FU/3 mL)			<0.001			<0.001
<9.65	42 (60.0%)	173 (32.7%)		17 (70.8%)	85 (30.9%)	
≥9.65	28 (40.0%)	356 (67.3%)		7 (29.2%)	190 (69.1%)	
CEA (ng/mL)			0.084			0.64
<1.875	39 (55.7%)	244 (46.1%)		9 (37.5%)	116 (42.2%)	
≥1.875	25 (35.7%)	258 (48.8%)		15 (62.5%)	143 (52.0%)	
Missing	6 (8.6%)	27 (5.1%)		0 (0%)	16 (5.8%)	
AGR			0.041			0.26
<1.566	31 (44.3%)	306 (57.8%)		11 (45.8%)	165 (60.0%)	
≥1.566	39 (55.7%)	222 (42.0%)		13 (54.2%)	110 (40.0%)	
Missing	0 (0%)	1 (0.2%)		0 (0%)	0 (0%)	
SCCA (ng/mL)			0.079			1
<1.085	25 (35.7%)	256 (48.4%)		12 (50.0%)	128 (46.5%)	
≥1.085	39 (55.7%)	240 (45.4%)		12 (50.0%)	129 (46.9%)	
Missing	6 (8.6%)	33 (6.2%)		0 (0%)	18 (6.5%)	
CYFRA21-1 (ng/mL)			0.069			0.91
<1.675	27 (38.6%)	150 (28.4%)		8 (33.3%)	93 (33.8%)	
≥1.675	36 (51.4%)	340 (64.3%)		16 (66.7%)	160 (58.2%)	
Missing	7 (10.0%)	39 (7.4%)		0 (0%)	22 (8.0%)	

TABLE 1 (Continued)

	Training set			Validation set		
	Benign ($N = 70$)	Malignant ($N = 529$)	p-Value	Benign ($N = 24$)	Malignant ($N = 275$)	p-Value
ProGRP (pg/mL)			0.26			0.6
<27.81	17 (24.3%)	163 (30.8%)		8 (33.3%)	65 (23.6%)	
≥27.81	47 (67.1%)	309 (58.4%)		15 (62.5%)	174 (63.3%)	
Missing	6 (8.6%)	57 (10.8%)		1 (4.2%)	36 (13.1%)	
NSE (ng/mL)			0.032			0.74
<11.54	29 (41.4%)	155 (29.3%)		9 (37.5%)	82 (29.8%)	
≥11.54	35 (50.0%)	344 (65.0%)		15 (62.5%)	175 (63.6%)	
Missing	6 (8.6%)	30 (5.7%)		0 (0%)	18 (6.5%)	
CA125 (U/mL)			0.35			0.85
<10.14	40 (57.1%)	261 (49.3%)		13 (54.2%)	123 (44.7%)	
≥10.14	24 (34.3%)	210 (39.7%)		10 (41.7%)	113 (41.1%)	
Missing	6 (8.6%)	58 (11.0%)		1 (4.2%)	39 (14.2%)	

Abbreviations: AGR, albumin/globulin ratio; CA125, carbohydrate antigen 125; CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment; ECOG, Eastern Cooperative Oncology Group; FR$^+$CTC, folate receptor-positive circulating tumor cell; NSE, neuron-specific enolase; ProGRP, pro-gastrin releasing peptide; SCCA, squamous cell carcinoma antigen.

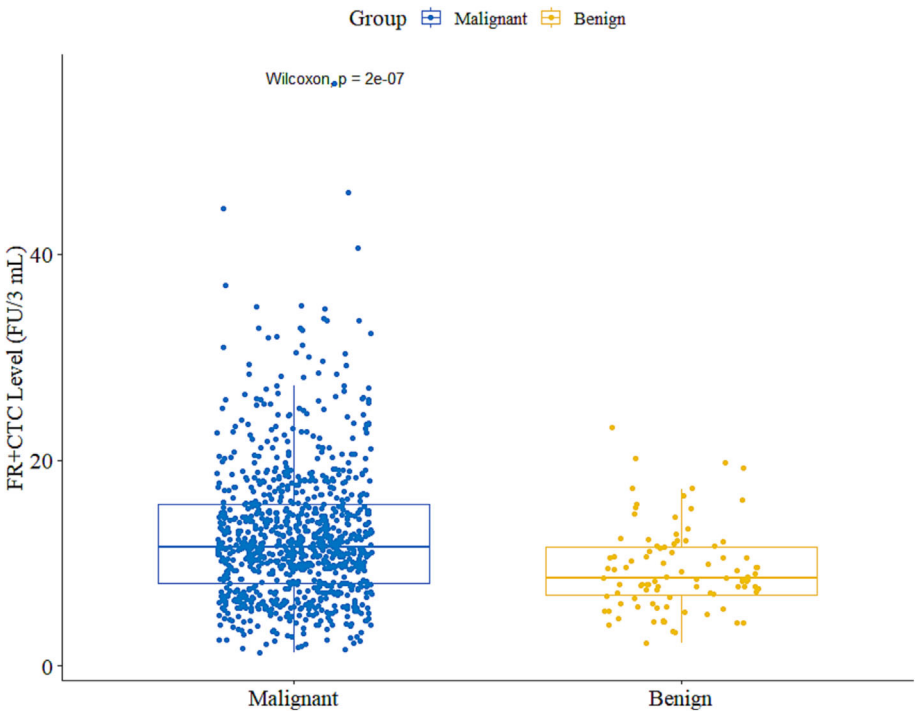

FIGURE 1 Comparison of FR$^+$CTC levels in the benign group and NSCLC group. FR$^+$CTC, folate receptor-positive circulating tumor cell; NSCLC, non-small cell lung cancer.

SCCA and CYFRA21-1 were marginally significant ($p = 0.07$, 0.06, and 0.05, respectively). In the multivariate analysis, FR$^+$CTC (odds ration, OR, 95% confidence interval, CI: 1.13, 1.07–1.19, $p < 0.0001$), age (OR, 95% CI: 1.06, 1.01–1.12, $p = 0.03$) and sex (OR, 95% CI: 1.07, 1.01–1.13, $p = 0.01$) were significant in the diagnosis of SPNs (Table 2).

3.4 | Diagnostic value of FR$^+$CTC, tumor biomarkers and the combination of other risk factors

The ROC curves were constructed to assess the diagnostic performance of FR$^+$CTC and serum biomarkers in SPNs and AUCs were calculated for each independent variable. The AUCs of FR$^+$CTC

TABLE 2 Multivariate binary logistic regression analysis for non-small cell lung cancer (NSCLC).

	Univariate			Multivariate		
Characteristics	**OR**	**95% CI**	***p***	**OR**	**95% CI**	***p***
Age (≥60, years)	1.90	1.12–3.22	0.02	1.06	1.01–1.12	0.03
Diameter (≥15 mm)	1.45	0.88–2.41	0.15			
Sex (female)	1.99	1.20–3.28	0.01	1.07	1.01–1.13	0.01
Benign lesions in the lung (yes)	0.80	0.33–1.98	0.63			
Cardiac disease (yes)	2.01	0.26–15.48	0.50			
Hypertension (yes)	1.37	0.74–2.54	0.32			
Diabetes mellitus (yes)	0.96	0.42–2.21	0.92			
ECOG score = 1	1.31	0.54–3.16	0.55			
ECOG score = 2	0.81	0.10–6.86	0.85			
FR^+CTC (≥9.65, FU/3 mL)	3.09	1.85–5.15	0.00	1.13	1.07–1.19	<0.001
CEA (≥1.875, ng/mL)	1.65	0.97–2.81	0.07			
AGR (≥1.566)	0.58	0.35–0.95	0.03	0.97	0.92–1.02	0.23
SCCA (≥1.085, ng/mL)	0.60	0.35–1.02	0.06			
CYFRA21-1 (≥1.675, ng/mL)	1.70	1.00–2.90	0.05			
ProGRP (≥27.81, pg/ml)	0.69	0.38–1.23	0.21			
NSE (≥11.54, ng/mL)	1.84	1.09–3.12	0.02	1.05	0.100–1.11	0.07
CA125 (≥10.14, U/mL)	1.34	0.78–2.30	0.28			

Abbreviations: OR, odds ratio; CI, confidence interval.

TABLE 3 Diagnostic efficiency for non-small cell lung cancer (NSCLC).

		Training set			Validation set		
Characteristic	**Cutoff**	**AUROC**	**Sensitivity**	**Specificity**	**AUROC**	**Sensitivity**	**Specificity**
FR^+CTC	9.650	0.650 (0.587–0.713)	0.673	0.600	0.700 (0.603–0.796)	0.691	0.708
CEA	1.875	0.550 (0.480–0.620)	0.514	0.609	0.464 (0.360–0.567)	0.552	0.375
AGR	1.566	0.564 (0.489–0.638)	0.580	0.557	0.571 (0.465–0.677)	0.600	0.542
SCC	1.085	0.562 (0.488–0.636)	0.516	0.609	0.501 (0.394–0.608)	0.502	0.500
CYFRA21-1	1.675	0.554 (0.429–0.694)	0.694	0.429	0.483 (0.382–0.584)	0.632	0.333
ProGRP	27.810	0.517 (0.444–0.590)	0.345	0.734	0.538 (0.435–0.641)	0.728	0.348
NSE	11.540	0.556 (0.480–0.632)	0.689	0.453	0.528 (0.425–0.631)	0.681	0.375
CA125	10.140	0.517 (0.442–0.593)	0.446	0.625	0.522 (0.414–0.630)	0.479	0.565
Age + Sex + FR^+CTC + AGR + NSE	0.883	0.725 (0.659–0.791)	0.813	0.580	0.828 (0.754–0.902)	0.767	0.792

Abbreviations: AUROC, area under the receiver operating characteristics; NSE, neuron-specific enolase.

were 0.650 (95% CI, 0.587–0.713) in the training set and 0.700 (95% CI, 0.603–0.796) in the validation set, superior to that of any other serum tumor biomarker alone (Table 3, Figure 2A,B). When FR^+CTC was combined with age, sex, AGR and NSE, the AUC in diagnosing NSCLC was 0.725 (95% CI, 0.659–0.791) in the training set and 0.828 (95% CI, 0.754–0.902) in the validation set (Table 3, Figure 2C,D).

4 | DISCUSSION

Non-small cell lung cancer accounts for about 85% of lung cancers. The 5-year survival rate of stage I/II NSCLC is 56–90%, while the 5-year survival rate of advanced stage is <24%.[32] However, owing to a lack of effective methods, only 16% of lung cancers are diagnosed at an early stage.[33] Currently, most of the commonly used models for diagnosing lung cancer, such as the Mayo model and the Veterans Affairs model, are mainly based on imaging features, which cannot meet the clinical needs.[34,35] Therefore, it is imperative to identify additional predictive factors and develop a more effective diagnostic model.

Circulating tumor cells are cells that have been detached from the primary or metastatic tumors and intravasate into the circulation system. They can cause metastases in various organs by passing through the bloodstream.[36] CTCs are rare in blood, but can be

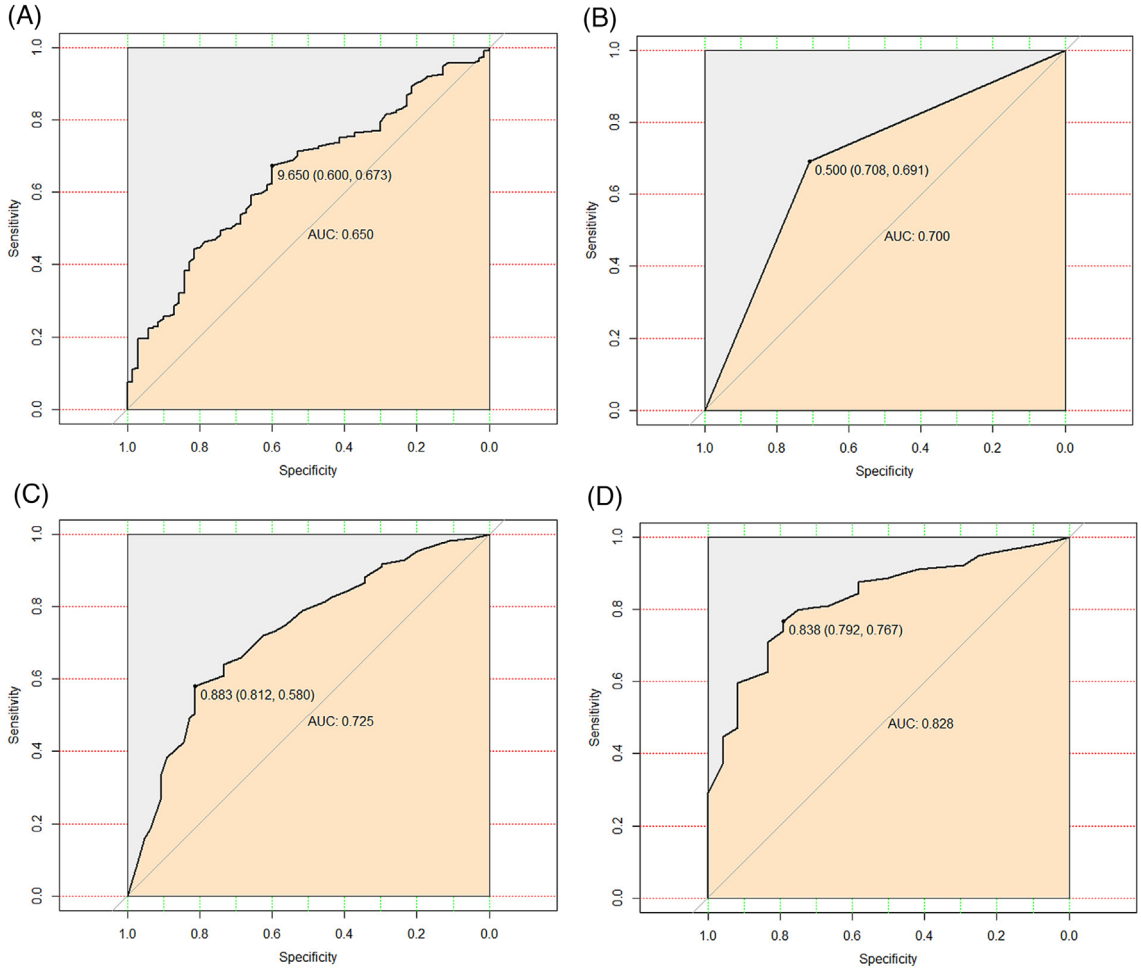

FIGURE 2 Receiver operating curve (ROC) curves of single biomarkers and different combination models in diagnosing NSCLC.

detected and counted following their separation from blood cells using various enrichment methods.[37] In previous studies, CTCs have been used as a marker for lung cancer progression, efficacy evaluation, targeted drugs development and individualized treatment.[18,30,38]

Folate receptor is a glycoprotein distributed on the cell membrane surface. It is highly expressed in cancer cells such as ovarian cancer, lung cancer and urinary cancers, but rarely expressed in normal tissues.[39] In lung cancer, it was found that the specific expression of FR varied from 72 to 83% in tumor tissues, making it a desirable target for lung cancer diagnosis and treatment.[24] FRα and its tumor specificity have been explored in cancer imaging, and a variety of methods for improving the quality of FRα-positive tumors imaging with FRα-targeted contrast-enhanced MRI have been investigated.[25,40,41] The development of FRα-targeted cancer treatment has also been actively pursued, including small molecules, monoclonal antibodies, vaccines, CAR T cells and folate–drug conjugates.[26,42–46]

Recently, FR has been adopted as a target for CTC enumeration and its applications in clinical diagnosis and prognosis in lung cancer have received much attention.[28,29,47–49] A few studies revealed that pre- and post-operative FR$^+$CTC levels are strong risk factors of recurrence and prognosis and may be used for clinical decision making.[50–53] In stage IV NSCLC patients receiving first-line chemotherapy or targeted therapy, FR$^+$CTC levels before and after treatment and the dynamic changes of FR$^+$CTC during treatment were associated with progression-free survival. Similar results were found for SCLC patients with chemotherapy.[48,54] In addition, for NSCLC patients receiving EGFR-TKI therapy, FR$^+$CTC was detected 3–6 months earlier than the clinical confirmation of PD diagnosed by traditional imaging technology, suggesting that it can be used as a biomarker for earlier prediction of tumor relapse.[30] Several studies have reported that the sensitivity of FR$^+$CTC is 3–5 times higher than that of other conventional tumor markers in the early detection of lung cancer.[28,29,55,56] Further, the sensitivity and specificity of FR$^+$CTC for preoperative prediction of tumor invasion were 73–82 and 83–88%, respectively, which could be a useful tool for determining operation strategy. Recently, a large-scale study involving more than 3,000 subjects confirmed that the overall diagnostic performance of FR$^+$CTC in patients with suspected pulmonary nodules was 87%.[57] Our results of FR$^+$CTC levels in distinguishing benign and malignant SPNs were consistent with those findings.

Owing to the insufficient sensitivity and specificity of single serum biomarkers in thediagnosis of lung cancer, several guidelines

recommend that a combination of serum biomarkers could be used to improve the sensitivity and specificity.[13] A number of prediction models utilizing different combinations of tumor biomarkers have been proposed and investigated in recent years.[14,15] However, there is still no consensus on which combination should be used for early detection of malignancy. In our study, serum tumor biomarkers and conventional clinical characteristics combined with FR^+CTC could yield satisfactory sensitivity and specificity for predicting malignancy in patients with SPNs. The results were similar to those for previous studies that included lung cancer patients at different stages.[56] Chen et al. established a predictive model for the diagnosis of lung cancer based on FR^+CTC combined with CEA, NSE and CYFRA21-1.[56] In our study, the diagnostic model included FR^+CTC, age, sex, AGR and NSE in patients with SPN. This indicates the value of the FR^+CTC combined model in the diagnosis of single lung cancer lesions.

The advantage of this study is that all patients had a single lesion, and the variables in this diagnostic model such as FR^+CTC, age, sex and AGR were easy to collect. Meanwhile the cutoff values of positive and negative used in this study were calculated based on the results obtained from the study population, rather than the clinical reference value, which can better reflect the characteristics of this cohort, but at the same time, the extrapolation is poor.

The study has several limitations. First, this was a single-center, retrospective study with a small sample size. The two groups included in the study were not balanced, with the malignant group having almost nine times as many patients as that of the benign group. The model needs to be further validated by prospective, multi-center studies with larger sample sizes. Second, in this study, the patients that received FR^+CTC detection before anticancer therapy were those with a strong suspicion of malignancy by examining physicians. In order to make this model more robust, it may be preferable to detect spontaneous SPN in patients during regular CT scans.

5 | CONCLUSIONS

In summary, our study demonstrated the value of FR^+CTC in diagnosing NSCLC, and developed a diagnostic prediction model based on FR^+CTC, demographic characteristics and serum biomarkers for screening SPNs.

AUTHOR CONTRIBUTIONS

Conception and design of the research: Yongjie Wang and Weiqing Gu. *Acquisition of data*: Dong Wang and Peng Li. *Provision of study materials or patients*: Dong Wang and Jinlong Li. *Analysis and interpretation of data*: Peng Li, Xiang Fei, and Shuyu Che. *Statistical analysis*: Yunpeng Xuan, Xiang Fei, and Yudong Han. *Drafting the manuscript*: all authors. *Revision of manuscript for important intellectual content*: Jinglong Wang. *Final approval of manuscript*: all authors.

ACKNOWLEDGEMENTS

Not applicable.

CONFLICT OF INTEREST STATEMENT

The authors have stated that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Not applicable.

REFERENCES

1. Feng R-M, Zong Y-N, Cao S-M, Xu R-H. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? *Cancer Commun*. 2019;39(1):22. doi:10.1186/s40880-019-0368-6
2. Huo J, Shen C, Volk RJ, Shih YT. Use of CT and chest radiography for lung cancer screening before and after publication of screening guidelines: intended and unintended uptake. *JAMA Intern Med*. 2017;177(3):439-441. doi:10.1001/jamainternmed.2016.9016
3. Richards TB, Soman A, Thomas CC, et al. Screening for lung cancer—10 states, 2017. *MMWR Morb Mortal Wkly Rep*. 2020;69(8):201-206. doi:10.15585/mmwr.mm6908a1
4. Narayan AK, Gupta Y, Little BP, Shepard JO, Flores EJ. Lung cancer screening eligibility and use with low-dose computed tomography: results from the 2018 Behavioral Risk Factor Surveillance System cross-sectional survey. *Cancer*. 2021;127(5):748-756. doi:10.1002/cncr.33322
5. Gartman E, Jankowich M, Baptiste J, Nici L. Providence VA lung cancer screening program: performance: comparison of local false positive and invasive procedure rates to published trial data. In: *A98. Clinical strategies to improve lung cancer early detection: who is at risk here?* American Thoracic Society; 2018:A2477.
6. Jin X, Xu L, Guan Y, Zhang Z, Li H. Bioinformatics analysis of microarray datasets to identify prognostic factors in lung adenocarcinoma. *DNA Cell Biol*. 2020;39(6):965-974. doi:10.1089/dna.2019.5203
7. Yang Y, Wang WW, Ren Y, et al. Computerized texture analysis predicts histological invasiveness within lung adenocarcinoma manifesting as pure ground-glass nodules. *Acta Radiol*. 2019;60(10):1258-1264. doi:10.1177/0284185119826536
8. Lee G, Park H, Sohn I, et al. Comprehensive computed tomography radiomics analysis of lung adenocarcinoma for prognostication. *Oncologist*. 2018;23(7):806-813. doi:10.1634/theoncologist.2017-0538
9. Shen Y, Xu F, Zhu W, Hu H, Chen T, Li Q. Multiclassifier fusion based on radiomics features for the prediction of benign and malignant primary pulmonary solid nodules. *Ann Transl Med*. 2020;8(5):171. doi:10.21037/atm.2020.01.135
10. Zhang Y, Cheng J, Hua X, et al. Can spectral CT imaging improve the differentiation between malignant and benign solitary pulmonary nodules? *PLoS ONE*. 2016;11(2):e0147537. doi:10.1371/journal.pone.0147537
11. Jia Y, Gong W, Zhang Z, et al. Comparing the diagnostic value of (18) F-FDG-PET/CT versus CT for differentiating benign and malignant solitary pulmonary nodules: a meta-analysis. *J Thorac Dis*. 2019;11(5):2082-2098. doi:10.21037/jtd.2019.05.21
12. Jiang ZF, Wang M, Xu JL. Thymidine kinase 1 combined with CEA, CYFRA21-1 and NSE improved its diagnostic value for lung cancer. *Life Sci*. 2018;194:1-6. doi:10.1016/j.lfs.2017.12.020
13. Stieber P, Hatz R, Holdenrieder S, et al. National academy of clinical biochemistry guidelines for the use of tumor markers in lung cancer. In: *NACB: practice guidelines and recommendations for use of tumor markers in the clinic lung cancer (section 3P)*. AACC Press; 2006.
14. Molina R, Marrades RM, Augé JM, et al. Assessment of a combined panel of six serum tumor markers for lung cancer. *Am J Respir Crit Care Med*. 2016;193(4):427-437. doi:10.1164/rccm.201404-0603OC
15. Xu L, Su Z, Xie B. Diagnostic value of conventional tumor markers in young patients with pulmonary nodules. *J Clin Lab Anal*. 2021;35(9):e23912. doi:10.1002/jcla.23912

16. Yang C, Chen F, Wang S, Xiong B. Circulating tumor cells in gastrointestinal cancers: current status and future perspectives. *Front Oncol*. 2019;9:1427. doi:10.3389/fonc.2019.01427
17. Cristofanilli M, Budd GT, Ellis MJ, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med*. 2004;351(8):781-791. doi:10.1056/NEJMoa040766
18. Tamminga M, de Wit S, Schuuring E, et al. Circulating tumor cells in lung cancer are prognostic and predictive for worse tumor response in both targeted- and chemotherapy. *Transl Lung Cancer Res*. 2019;8(6):854-861. doi:10.21037/tlcr.2019.11.06
19. Allard WJ, Matera J, Miller MC, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res*. 2004;10(20):6897-6904. doi:10.1158/1078-0432.ccr-04-0378
20. Rushton AJ, Nteliopoulos G, Shaw JA, Coombes RC. A review of circulating tumour cell enrichment technologies. *Cancer*. 2021;13(5):970. doi:10.3390/cancers13050970
21. Iwakiri S, Sonobe M, Nagai S, Hirata T, Wada H, Miyahara R. Expression status of folate receptor alpha is significantly correlated with prognosis in non-small-cell lung cancers. *Ann Surg Oncol*. 2008;15(3):889-899. doi:10.1245/s10434-007-9755-3
22. Shia J, Klimstra DS, Nitzkorski JR, et al. Immunohistochemical expression of folate receptor α in colorectal carcinoma: patterns and biological significance. *Hum Pathol*. 2008;39(4):498-505. doi:10.1016/j.humpath.2007.09.013
23. Basal E, Eghbali-Fatourechi GZ, Kalli KR, et al. Functional folate receptor alpha is elevated in the blood of ovarian cancer patients. *PLoS ONE*. 2009;4(7):e6292. doi:10.1371/journal.pone.0006292
24. Nunez MI, Behrens C, Woods DM, et al. High expression of folate receptor alpha in lung cancer correlates with adenocarcinoma histology and mutation. *J Thorac Oncol*. 2012;7(5):833-840. doi:10.1097/JTO.0b013e31824de09c
25. Ao L, Wang B, Liu P, et al. A folate-integrated magnetic polymer micelle for MRI and dual targeted drug delivery. *Nanoscale*. 2014;6(18):10710-10716. doi:10.1039/c4nr02484b
26. Vergote I, Armstrong D, Scambia G, et al. A randomized, double-blind, placebo-controlled, phase III study to assess efficacy and safety of weekly farletuzumab in combination with carboplatin and taxane in patients with ovarian cancer in first platinum-sensitive relapse. *J Clin Oncol*. 2016;34(19):2271-2278. doi:10.1200/JCO.2015.63.2596
27. Sato S, Itamochi H. Profile of farletuzumab and its potential in the treatment of solid tumors. *Onco Targets Ther*. 2016;9:1181-1188. doi:10.2147/OTT.S98242
28. Yu Y, Chen Z, Dong J, et al. Folate receptor-positive circulating tumor cells as a novel diagnostic biomarker in non-small cell lung cancer. *Transl Oncol*. 2013;6(6):697-702. doi:10.1593/tlo.13535
29. Lou J, Ben S, Yang G, et al. Quantification of rare circulating tumor cells in non-small cell lung cancer by ligand-targeted PCR. *PLoS ONE*. 2013;8(12):e80458. doi:10.1371/journal.pone.0080458
30. Jiang T, Zhao J, Zhao C, et al. Dynamic monitoring and predictive value of circulating tumor cells in EGFR-mutated advanced non-small-cell lung cancer patients treated with first-line EGFR tyrosine kinase inhibitors. *Clin Lung Cancer*. 2019;20(2):124-133 e2. doi:10.1016/j.cllc.2018.11.014
31. He W, Kularatne SA, Kalli KR, et al. Quantitation of circulating tumor cells in blood samples from ovarian and prostate cancer patients using tumor-specific fluorescent ligands. *Int J Cancer*. 2008;123(8):1968-1973. doi:10.1002/ijc.23717
32. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11(1):39-51. doi:10.1016/j.jtho.2015.09.009
33. Mateo J, Gerlinger M, Rodrigues DN, de Bono JS. The promise of circulating tumor cell analysis in cancer management. *Genome Biol*. 2014;15(8):448. doi:10.1186/s13059-014-0448-5
34. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med*. 1997;157(8):849-855. doi:10.1001/archinte.157.8.849
35. Gould MK, Fletcher J, Iannettoni MD, et al. Evaluation of patients with pulmonary nodules: when is it lung cancer? ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(3):108S-130S. doi:10.1378/chest.07-1353
36. Lozar T, Gersak K, Cemazar M, Kuhar CG, Jesenko T. The biology and clinical potential of circulating tumor cells. *Radiol Oncologia*. 2019;53(2):131-147. doi:10.2478/raon-2019-0024
37. Boya M, Chu CH, Liu R, Ozkaya-Ahmadov T, Sarioglu AF. Circulating tumor cell enrichment technologies. *Recent Results Cancer Res*. 2020;215:25-55. doi:10.1007/978-3-030-26439-0_2
38. Krebs MG, Sloane R, Priest L, et al. Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. *J Clin Oncol*. 2011;29(12):1556-1563. doi:10.1200/jco.2010.28.7045
39. Scaranti M, Cojocaru E, Banerjee S, Banerji U. Exploiting the folate receptor α in oncology. *Nat Rev Clin Oncol*. 2020;17(6):349-359. doi:10.1038/s41571-020-0339-5
40. Konda SD, Aref M, Brechbiel M, Wiener EC. Development of a tumor-targeting MR contrast agent using the high-affinity folate receptor: work in progress. *Invest Radiol*. 2000;35(1):50-57. doi:10.1097/00004424-200001000-00006
41. Meier R, Henning TD, Boddington S, et al. Breast cancers: MR imaging of folate-receptor expression with the folate-specific nanoparticle P1133. *Radiology*. 2010;255(2):527-535. doi:10.1148/radiol.10090050
42. Gibbs DD, Theti DS, Wood N, et al. BGC 945, a novel tumor-selective thymidylate synthase inhibitor targeted to alpha-folate receptor-overexpressing tumors. *Cancer Res*. 2005;65(24):11721-11728. doi:10.1158/0008-5472.CAN-05-2034
43. Hasegawa K, Nakamura T, Harvey M, et al. The use of a tropism-modified measles virus in folate receptor-targeted virotherapy of ovarian cancer. *Clin Cancer Res*. 2006;12(20 Pt 1):6170-6178. doi:10.1158/1078-0432.CCR-06-0992
44. Kalli KR, Block MS, Kasi PM, et al. Folate receptor alpha peptide vaccine generates immunity in breast and ovarian cancer patients. *Clin Cancer Res*. 2018;24(13):3014-3025. doi:10.1158/1078-0432.CCR-17-2499
45. Kandalaft LE, Powell DJ Jr, Coukos G. A phase I clinical trial of adoptive transfer of folate receptor-alpha redirected autologous T cells for recurrent ovarian cancer. *J Transl Med*. 2012;10:157. doi:10.1186/1479-5876-10-157
46. O'Malley DM, Matulonis UA, Birrer MJ, et al. Phase Ib study of mirvetuximab soravtansine, a folate receptor alpha (FRalpha)-targeting antibody-drug conjugate (ADC), in combination with bevacizumab in patients with platinum-resistant ovarian cancer. *Gynecol Oncol*. 2020;157(2):379-385. doi:10.1016/j.ygyno.2020.01.037
47. Qi F, Liu Y, Zhao R, et al. Quantitation of rare circulating tumor cells by folate receptor alpha ligand-targeted PCR in bladder transitional cell carcinoma and its potential diagnostic significance. *Tumor Biol*. 2014;35(7):7217-7223. doi:10.1007/s13277-014-1894-0
48. Chen X, Zhou C, Li X, Yang G. Correlation of baseline value of folate receptor-positive circulating tumor cells and efficacy of pemetrexed and dynamic monitoring study in NSCLC patients receiving first-line platinum-based chemotherapy. *J Thorac Oncol*. 2016;11(4):S125-S126. doi:10.1016/s1556-0864(16)30266-0
49. Marino FZ, Ronchi A, Accardo M, Franco R. Detection of folate receptor-positive circulating tumor cells by ligand-targeted

polymerase chain reaction in non-small cell lung cancer patients. *J Thorac Dis*. 2016;8(7):1437-1439. doi:10.21037/jtd.2016.05.38

50. He YG, Zheng L, Gao MF, et al. A preliminary study of the clinical significance of folate receptor-positive circulating tumor cell in the management of hepatobiliary-pancreatic cancers. *Transl Cancer Res*. 2020;9(11):6700-6709. doi:10.21037/tcr-20-2192
51. Wei S, Guo C, He J, et al. Effect of vein-first vs artery-first surgical technique on circulating tumor cells and survival in patients with non-small cell lung cancer: a randomized clinical trial and registry-based propensity score matching analysis. *JAMA Surg*. 2019;154(7): e190972. doi:10.1001/jamasurg.2019.0972
52. Zhou Q, Geng Q, Wang L, et al. Value of folate receptor-positive circulating tumour cells in the clinical management of indeterminate lung nodules: a non-invasive biomarker for predicting malignancy and tumour invasiveness. *EBioMedicine*. 2019;41:236-243. doi:10.1016/j.ebiom.2019.02.028
53. Zheng J, Ye X, Liu Y, Zhao Y, He M, Xiao H. The combination of CTCs and CEA can help guide the management of patients with SPNs suspected of being lung cancer. *BMC Cancer*. 2020;20(1):106. doi:10.1186/s12885-020-6524-1
54. Shen J, Zhao J, Jiang T, et al. Predictive and prognostic value of folate receptor-positive circulating tumor cells in small cell lung cancer patients treated with first-line chemotherapy. *Oncotarget*. 2017;8(30): 49044-49052. doi:10.18632/oncotarget.17039
55. Wan J-W, Gao M-Z, Hu R-J, et al. A preliminary study on the relationship between circulating tumor cells count and clinical features in patients with non-small cell lung cancer. *Ann Transl Med*. 2015;3(22): 352. doi:10.3978/j.issn.2305-5839.2015.11.18
56. Chen X, Zhou F, Li X, et al. Folate receptor-positive circulating tumor cell detected by LT-PCR-based method as a diagnostic biomarker for non-small-cell lung cancer. *J Thorac Oncol*. 2015;10(8):1163-1171. doi:10.1097/jto.0000000000000606
57. Yin W, Zhu J, Ma B, et al. Overcoming obstacles in pathological diagnosis of pulmonary nodules through circulating tumor cell enrichment. *Small*. 2020;16(25):2001695. doi:10.1002/smll.202001695

How to cite this article: Wang D, Li P, Fei X, et al. A combined diagnostic model based on circulating tumor cell in patients with solitary pulmonary nodules. *J Gene Med*. 2023;e3529. doi:10.1002/jgm.3529