sciendo

Professional paper

DOI:10.2478/rrlm-2022-0037

Diagnostic value of combination of exfoliative cytology with CA125, CEA, NSE, CYFRA21-1 and CA15-3 for lung cancer

Liang Zhang[1], Jun Lei[2], Jing Xia[3], Ang Cai[3*]

1. Laboratory Medicine, Wuhan Hospital of Traditional Chinese Medicine, China

2. Respiratory Medicine, Wuhan Hospital of Traditional Chinese Medicine, China

3. Oncology, Wuhan Hospital of Traditional Chinese Medicine, China

Abstract

***Background**: To explore the diagnostic value of combination of exfoliative cytology with detection of tumor markers carbohydrate antigen 125 (CA125), carcinoembryonic antigen (CEA), neuron specific enolase (NSE), cytokeratin 19 fragment antigen 21-1 (CYFRA21-1) and CA15-3 for lung cancer. **Methods**: A total of 256 patients were enrolled, including 164 males and 92 females aged (64.51±22.68) years old. Among them, 189 patients (100 males and 89 females) were randomly selected as Tumor group, and the remaining 67 patients were used for validation. Another 514 healthy people receiving physical examination in our hospital during the same period were selected, from which 397 cases (266 males and 131 females) were randomly selected as No Tumor group, and the remaining 117 cases were used for validation. The biochemical criteria were detected in all subjects. The diagnostic value of each index for lung cancer was analyzed using receiver operating characteristic (ROC) curves. **Results**: The results of ROC curve analysis revealed that in Tumor group, the area under curve (AUC) of exfoliative cytology, CA125, CYFRA21-1, CA15-3, CEA and NSE was ≥0.7, while that of CA72-4, CA19-9, TSGF, AFP, CA242, SCC-Ag and CA50 was <0.7. The indices in each factor were comprehensively assessed, and then exfoliative cytology, CA125, CA15-3, CYFRA21-1, CEA and NSE were screened to establish the lung cancer prediction model. The diagnostic value was comparable between the prediction model and the combined detection of 9 indices (Z=1.682, P=0.079). **Conclusions**: The lung cancer prediction model balances sensitivity and specificity without reducing the diagnostic efficiency.*

***Keywords**: pleural effusion, exfoliative cells, tumor markers, combined detection, lung cancer, diagnostic value*

Received: 19th May 2022; Accepted: 18th September 2022; Published: 28th September 2022

Introduction

Lung cancer is a malignancy frequently occurring in pulmonary alveoli and bronchi, whose mortality and fatality rates are extremely high due to indefinite early diagnosis. According to statistics, more than 80% of patients are diagnosed with mid-advanced lung cancer at the first visit. Therefore, it is of important clinical significance to develop convenient and efficient criteria for early diagnosis. Currently, tumor markers and sputum pathology are primarily used in early

* **Corresponding author**: Ang Cai, Oncology, Wuhan Hospital of Traditional Chinese Medicine, China.
E-mail: caiangwhtcm@wl-asia.com

clinical diagnosis of lung cancer, in which detection of serum tumor markers is the least invasive to patients (1). In clinical detection, however, it has been found that the levels of serum tumor markers are normal or slightly higher in some patients with lung cancer, while they significantly rise in pleural effusion. Therefore, it is more sensitive to detect tumor markers in pleural effusion (2). Exfoliative cytology characterized by high specificity and low sensitivity is often used clinically to distinguish benign and malignant pleural effusion (3). Carcinoembryonic antigen (CEA) is present in endoderm-derived digestive system cancer and normal digestive tract tissues, and also present in a trace amount in normal human serum, which can be used as a diagnostic marker for breast cancer, lung cancer, and colorectal cancer, but its specificity and sensitivity are not high (4). Carbohydrate antigen 125 (CA125), a glycoprotein that can be bound by the monoclonal antibody OC125, is most commonly found in the serum of patients with epithelial ovarian tumors, but does not exist in normal ovarian tissues, which has high diagnostic sensitivity, but low specificity (5). Neuron-specific enolase (NSE) is an acid protease unique to neuroendocrine cells and neurons, which is a specific marker for neuroendocrine tumors such as medullary thyroid carcinoma and small cell carcinoma (6). Cytokeratin 19 fragment antigen 21-1 (CYFRA21-1) is a soluble fragment of cytokeratin 19 (CYK-19) present in a large amount in malignant lung cancer tissues, and serves as a tumor marker mainly for lung cancer (7). Tumor cells derived from solid lung cancer continue to proliferate, escape from tissues, and enter the circulatory system through molecular transcription, protein modification, and phenotype changes, finally becoming circulating tumor cells (CTCs) capable of invasion and metastasis. CTC detection technology uses high-sensitivity fluorescent quantitative gene amplification technology to amplify specific target genes after enriching leukocytes with immunomagnetic beads. This technology has high sensitivity, but the cost is high, therefore promoting its application in clinical practice is difficult.

To explore convenient and efficient criteria for early diagnosis, exfoliative cells and tumor markers CEA, CA125, CYFRA21-1 and NSE in the pleural effusion of patients with lung cancer were retrospectively analyzed in this study, and the prediction model combining exfoliative cells and tumor markers was established through statistical methods, so as to provide a theoretical basis for optimizing clinical detection.

Materials and Methods

Pathological data

This study has been approved by the ethics committee of our hospital, and written informed consent has been obtained from all enrolled subjects. A total of 256 patients treated in our hospital from June 2019 to May 2021 and diagnosed with primary lung cancer by cytological or histopathological examination were enrolled, including 164 males and 92 females aged (64.51±22.68) years old. Among them, 189 patients (100 males and 89 females) were randomly selected as Tumor group, and the remaining 67 patients were used for validation. Another 514 healthy people receiving physical examination and also pleural puncture in our hospital during the same period were selected, including 318 males and 196 females aged (68.19±18.23) years old, from which 397 cases (266 males and 131 females) were randomly selected as No Tumor group, and the remaining 117 cases were used for validation.

Recording and measurement of baseline data

The patients' gender, age, body mass index (BMI) and smoking index were recorded. 5 mL of fasting venous blood was drawn in the morning. Then D-dimer (D-D) and fibrinogen (FIB) were detected using a Cobas 8000 automatic biochemical analyzer (ROCHE), and white

blood cell count (WBC), neutrophil to lymphocyte ratio (NLR) and red blood cell distribution width (RDW) were detected using a SX-500I automatic hematology analyzer (Sysmex, Japan).

Exfoliative cytology examination

The pleural effusion (30-50 mL) was routinely drawn. Then 10-20 mL of pleural effusion was placed in a centrifuge tube and centrifuged, and 5 mL of supernatant was harvested and stored at -20°C for tumor marker detection. Besides, the upper cell layer of the precipitate was harvested and prepared into two sections *via* the membrane ultra-thin cell smearing method. After drying, they were subjected to Wright-Giemsa staining and exfoliative cytology. The images were read using an OlympusBX41 microscope image analyzer.

Determination criteria of cytology

The determination criteria of Ruizhen *et al.* (8) were used for cytology (**Table 1**).

Detection of tumor markers

The patients' pleural effusion was drawn and centrifuged, and the supernatant was harvested. Then the tumor markers were detected using a Modular E170 analyzer (ROCHE) in strict accordance with the instructions.

Statistical analysis

SPSS 20.0 software was used for statistical analysis. Kolmogorov-Smirnov test of normality was conducted. Normally distributed data were expressed as ($\overline{\chi} \pm s$), intergroup comparison was performed using *t* test, and Pearson's correlation analysis was performed. Abnormally distributed data were expressed as median (M) [quartile (P25,P75)], intergroup comparison was conducted using Mann-Whitney test, and Spearman's correlation analysis was carried out. The receiver operating characteristic (ROC) curves were plotted using GraphPad Prism 5 software, and the area under curve (AUC), sensitivity, specificity, positive likelihood ratio (+LR), negative likelihood ratio (-LR), Youden index and optimal cut-off value were calculated. AUC was compared by Z test, and AUC<0.5, =0.5-0.7, =0.7-0.9 and >0.9 indicated no diagnostic value, low diagnostic value, moderate diagnostic value and high diagnostic value, respectively. Potential common factors were extracted using exploratory factor analysis (EFA), and the indices within the same factor were screened by comprehensive assessment. Multiple indices were fitted by logistic regression analysis, and principal component analysis was performed to eliminate multicollinearity. P<0.05 was considered to be statistically significant.

Results

Baseline data

According to statistics, the general data such as gender and age were not significantly different between Tumor group and No Tumor group (P>0.05), suggesting that the general data had no important influencing factors for the subsequent test results in both groups (**Table 2**).

Exfoliative cytology and expressions and differences of tumor markers

It was found by exfoliative cytology that there were 163 cases of adenocarcinoma, 11 cases of squamous cell carcinoma and 15 cases of small

Table 1. Determination criteria for cytology

Description	Criteria
Negative	No malignant tumor cells
Atypical epithelial cells	Dyskaryotic cells
Suspicious malignant tumor cells	A few dyskaryotic cells, less obvious atypia or degeneration

Table 2. Baseline data [$\overline{\chi} \pm s$, n(%)]

	Tumor (n=189)	No Tumor (n=397)	t/χ^2	P
Gender (male/female, n)	100/89	266/131	0.134	0.621
Age (Y)	64.51±22.68	68.19±18.23	0.247	0.483
BMI (kg/m^2)	22.67±2.15	23.58±2.30	0.175	0.596
Smoking index [Y·n, M(P25,P75)]	529(236~854)	535(243~867)	0.541	0.228
WBC (10^9/L)	9.41±1.67	12.36±3.27	0.801	0.173
NLR (%)	3.74±1.32	8.24±2.16	0.739	0.185
RDW (%)	13.83±1.45	15.27±1.68	0.194	0.442
FIB (g/L)	3.28±0.73	3.84±1.05	0.833	0.102
D-D (mg/L)	0.96±0.25	1.09±0.37	0.632	0.228

WBC: white blood cell count; NLR: neutrophil to lymphocyte ratio; RDW: red blood cell distribution width; FIB: fibrinogen; D-D: D-dimer.

cell carcinoma in Tumor group, while all cases were negative in No Tumor group (P<0.05).

The results of tumor marker detection revealed that the levels of all tumor markers except SCC-Ag and TSGF were significantly higher in Tumor group than those in No Tumor group (P<0.05) (**Table 3**).

ROC curve analysis results of cytology and tumor markers

The results of ROC curve analysis found that the diagnostic value of cytology, CA125, CYFRA21-1, CA15-3, CEA and NSE was moderate and above (AUC≥0.7), while CA72-4, CA19-9, TSGF, AFP, CA242, SCC-Ag and CA50 had no high diagnostic value (0.5≤AUC<0.7). NSE had the highest Youden index and sensitivity, and CEA had the highest specificity and strong ability to judge the disease status (+LR=6.58). AUC of cytology was the largest, and had significant differences from those of CA72-4, CA19-9, CA50, TSGF, CA242, SCC-Ag and AFP (P<0.01, P<0.05) (**Table 4**).

Table 3. Exfoliative cytology and expressions and differences of tumor markers [$\overline{\chi} \pm s$, n(%)]

	Tumor(n=189)	No Tumor (n=397)	$t/\chi 2$	P
Cytology	152(80.65)	0(0)	2.782	0.000
Adenocarcinoma	131(86.24)	0(0)	0.886	0.001
Squamous cell carcinoma	9(5.82)	0(0)	0.704	0.005
Small cell carcinoma	12(7.94)	0(0)	0.912	0.000
AFP (ng/mL)	2.72±0.58*	2.19±0.43	1.083	0.002
CEA (ng/mL)	8.65±1.05*	3.28±0.75	0.479	0.000
CYFRA21-1 (ng/mL)	6.24±0.33*	2.75±0.64	1.392	0.000
CA72-4 (U/mL)	4.07±1.22*	1.65±1.06	0.451	0.027
CA242 (U/mL)	5.84±0.36*	3.62±0.55	0.365	0.030
CA15-3 (U/mL)	22.06±0.82*	9.41±0.72	1.022	0.005
CA125 (ng/mL)	83.47±0.71*	18.53±0.56	0.945	0.000
CA19-9 (U/mL)	20.78±0.46*	9.72±0.32	0.609	0.018
CA50 (U/mL)	11.63±0.55*	7.23±0.18	0.592	0.011
NSE (ng/mL)	16.72±0.62*	6.02±0.58	1.066	0.000
SCC-Ag (ng/mL)	1.58±1.22	1.54±0.96	0.519	0.36
TSGF (U/mL)	57.09±15.12	60.23±15.37	0.379	0.102

*P<0.05

Table 4. ROC curve analysis results of cytology and tumor markers

	AUC	95%CI	Optimal cut-off	Sensitivity (%)	Specificity (%)	Youden index	+LR	-LR
Cytology	0.853	0.809-0.894	82.7%	81.62	93.72	0.553	3.61	0.42
CA15-3	0.822	0.782-0.874	14.90 U/mL	66.82	78.63	0.465	3.22	0.38
NSE	0.794	0.723-0.845	9.96ng/mL	71.88	75.11	0.469	3.01	0.36
CA125	0.768	0.712-0.797	40.78ng/mL	66.53	72.38	0.415	2.63	0.42
CEA	0.752	0.724-0.768	7.06ng/mL	50.80	91.41	0.439	6.58	0.51
CYFRA21-1	0.723	0.701-0.765	4.55ng/mL	60.95	83.55	0.446	3.77	0.45
CA72-4	0.698*	0.654-0.743	3.48U/mL	51.28	81.09	0.308	2.56	0.58
TSGF	0.663**	0.627-0.732	60.29U/mL	62.57	71.23	0.319	2.05	0.52
CA19-9	0.659**	0.608-0.694	18.45U/mL	52.68	72.66	0.274	2.01	0.60
CA242	0.637**	0.611-0.685	7.63U/mL	45.29	80.41	0.258	2.42	0.66
AFP	0.615**	0.578-0.647	2.53ng/mL	56.77	63.89	0.221	1.58	0.65
CA50	0.588**	0.528-0.634	33.19U/mL	28.01	93.75	0.218	4.63	0.80
SCC	0.596**	0.547-0.639	1.19ng/mL	61.18	61.22	0.173	1.39	0.72

*P<0.05, **P<0.01 vs. cytology.

Common factors extracted

The levels of TSGF and SCC-Ag had no significant differences between Tumor group and No Tumor group (P>0.05), so TSGF and SCC-Ag were not incorporated into factor analysis to improve the diagnostic accuracy. In addition, tumor markers with an AUC≥0.6 were selected. After normalization, the primary data in both groups were substituted into variables. The results of KMO and Bartlett's test of sphericity in both groups met the EFA conditions. Moreover, the results of the rotated component matrix and the rotated 3D load diagram showed that the 11 indexes in Tumor group were dimensionally reduced to 4 common factors, and then they were sorted based on the variance contribution rate. The 11 indexes in No Tumor group were dimensionally reduced to 3 common factors, and then they were sorted based on the variance contribution rate (**Table 5**).

In Tumor group, the load coefficient of cytology, CA50, CA125 and CA19-9 was higher in common factor 1; the load coefficient of CA72-4 and CYFRA21-1 was higher in common factor 2; the load coefficient of CA242 and NSE was higher in common factor 3; the load coefficient of CA15-3 and CEA was higher in common factor 4. AFP was eliminated due to its low load coefficient in common factors. The above 4 common factors were used as diagnostic indexes for lung cancer, and common factor 1 had the highest diagnostic value and was used to distinguish negative and positive results, while common factors 2/3/4 with moderate diagnostic value could assist in

Table 5. Rotated component matrix

	Common factor 1	Common factor 2	Common factor 3	Common factor 4
Tumor group	Cytology, CA50, CA125, CA19-9	CA72-4, CYFRA21-1	CA242, NSE	CA15-3, CEA
No Tumor group	AFP, CA72-4, CA125, cytology, CEA	CA19-9, CA242, CA50, CA15-3	NSE, CYFRA21-1	

Within the same common factor, each index had a higher load coefficient; the common factors were sorted based on the variance contribution rate, and the higher the variance contribution rate, the higher the diagnostic value.

the diagnosis and classification of lung cancer. Besides, in No Tumor group, AFP, CA72-4, CA125, cytology and CEA had a higher load coefficient in common factor 1; CA19-9, CA242, CA50 and CA15-3 had a higher load coefficient in common factor 2; NSE and CYFRA21-1 had a higher load coefficient in common factor 3. The above three common factors could reflect the correlation and distribution of tumor markers in normal people.

Establishment of a combined model and determination of diagnostic value

It can be seen from relevant literature that cytology and 10 tumor markers possess diagnostic value for lung cancer. However, CA242, CA72-4, CA50, CA19-9 and SCC-Ag are rarely reported in literature (9,10). On the contrary, there are many literature reports and sufficient evidence regarding cytology, CA15-3, NSE, CA125, CEA and CYFRA21-1, and NSE, CA125, CEA and CYFRA21-1 have been even listed in the guidelines of the European Group on Tumor Markers (EGTM) (11) and the National Academy of Clinical Biochemistry (NACB) (12). In this study, the results of correlation analysis on the 4 tumor markers in common factor 1 in Tumor group revealed that there were significant positive pairwise correlations among cytology, CA50, CA125 and CA19-9 (P<0.05). Cytology and CA125 had higher diagnostic efficacy, the former of which had the best diagnostic efficacy, and the diagnostic efficacy had no significant difference between CA19-9 and CA50 (P>0.05). Therefore, CA19-9 and CA50 were eliminated, and only cytology and CA125 were retained. Then the indexes in other factors were screened in the same way. There was a positive correlation between CA72-4 and CYFRA21-1 in common factor 2 (P<0.05), and the difference in their diagnostic efficacy was not significant (P>0.05). Therefore, only CYFRA21-1 was retained. There was a positive correlation between CA242 and NSE in common factor 3, and NSE had higher diagnostic efficacy, so only NSE was retained. CA15-3 and CEA were positively correlated in common factor 4 (P<0.05), and they both had higher diagnostic efficacy. To sum up, the lung cancer prediction model could be established using cytology, CA125, CYFRA21-1, NSE, CA15-3 and CEA in Tumor group (**Table 6**).

After the dependent variables were set, logistic regression analysis was conducted, and Tumor group was marked as 1, and No Tumor group as 0. Univariate analysis was first conducted to screen variables with P<0.1. It was found that except AFP (P=0.314) and CA72-4 (P=0.257), the remaining 9 indexes all met the conditions. Then the 9 tumor markers were converted to covariates to obtain the prediction probability expressed as Pre-P9, whose AUC was 0.825. Besides, the lung cancer prediction model was marked as 1, the corresponding indexes in No Tumor group as 0, and their aggregate was marked as Tumor-No Tumor. The AUC of the prediction probability expressed as Pre-P6 was 0.816, without significant difference from Pre-P9 (Z=1.682, P=0.079), indicating that the diagnostic value of the two was comparable.

Validation of combined model

The principal components of Tumor-No-Tumor were extracted, and the results showed

Table 6. Correlation analysis results and AUC values of tumor markers

	Spearman correlation coefficient			AUC (P)		
	Cytology	**CA50**	**CA19-9**	**Cytology**	**CA50**	**CA19-9**
CA125	0.219*	0.263*	0.371*	<0.01	<0.01	<0.05
Cytology		0.582*	0.609*		<0.01	<0.01
CA50			0.613*			>0.05

*P<0.05.

that KMO=0.597, Bartlett's test of sphericity P<0.001, and the cumulative variance contribution rate of 78%. Univariate analysis was performed by binary logistic regression to screen variables with P<0.1. The principal components that met the conditions were used to fit the logistic regression. The Tumor-No-Tumor statistical results showed that -2LL=190.793, Cox & Snell R^2=0.504, Nagelkerke R^2=0.618 and H-L test=0.287, and the equation was Logistic (Tumor) = -4.016 + 0.025 × cytology + -0.005 × CA125 + 0.023 × CYFRA21-1 + 0.165 × NSE + 0.071 × CA15-3 + 0.014 × CEA. The logistic (Tumor) equation was turned back into the original variable equation, and the original data were substituted to calculate the result. The ROC interaction point diagram was plotted using Medcalc v11 software, and it was found that the optimal cut-off value was -0.9476 and the doubtful value ranged from -1.305 to -0.743. The data on tumor markers of the remaining 67 patients in Tumor group and 117 patients in No Tumor group were substituted into the Tumor-No-Tumor model. The total coincidence rate, negative coincidence rate and positive coincidence rate reached 84.2%, 85.7% and 79.5%, respectively.

The AUC, standard error, sensitivity, specificity, Youden index, +LR, -LR and optimal cut-off value of the prediction model were 0.826, 0.018, 68.7%, 84.2%, 0.602, 4.28, 0.28 and -0.9617, respectively.

Table 7. Lung cancer prediction model and strategy of combined tumor marker detection

	AUC	Sensitivity (%)	Specificity (%)	+LR	-LR	Youden index
Lung cancer prediction model	0.843	71.6	84.2	4.40	0.36	0.539
CEA+NSE+CYFRA21-1 [20]	0.782	67.5	83.5	3.87	0.37	0.527
CEA+SCC-Ag+CYFRA21-1 [17]	0.786	72.8	77.6	3.36	0.33	0.518
CEA+SCC-Ag+NSE+CYFRA21-1 [22]	0.805	74.3	78.4	3.41	0.32	0.521
CEA+NSE+CA125+SCC-Ag [23]	0.796	73.8	75.8	3.29	0.33	0.497
CEA+CA125+SCC-Ag+CYFRA21-1+NSE [16]	0.808	73.9	77.6	3.27	0.33	0.521

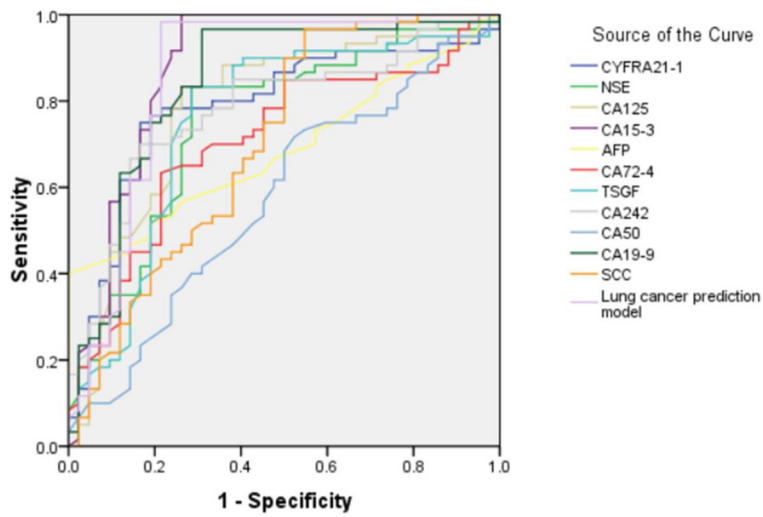

Fig. 1. Diagnostic values of prediction model and single indices (CA125, CYFRA21-1, CA15-3, CEA, NSE, CA72-4, CA19-9, TSGF, AFP, CA242, SCC, CA50) for lung cancer assessed by ROC curves.

Lung cancer prediction model and strategy of combined tumor marker detection

Compared with the strategy of combined detection, both AUC and Youden index of the lung cancer prediction model were the optimal, and its true positive detection rate, sensitivity and specificity were also greatly improved (**Table 7**). The ROC curve analysis showed that the diagnostic value of prediction model was superior to those of single indices (CA125, CYFRA21-1, CA15-3, CEA, NSE, CA72-4, CA19-9, TSGF, AFP, CA242, SCC, CA50) (**Figure 1**).

Discussion

Lung cancer frequently occurs in pulmonary alveoli and bronchi, whose mortality and fatality rates are extremely high. The prognosis of lung cancer treated with early surgical resection is good. Therefore, early detection of lung cancer and effective control of its malignant progression are of great clinical significance. Malignant pleural effusion is one of the early manifestations of lung cancer, and exfoliative cytology is primarily used in clinic to distinguish benign and malignant pleural effusion. Exfoliative cytology is a minimally-invasive and fast examination method characterized by simple operation, low costs and high reproducibility, but it has low sensitivity. For that reason, the detection amount of pleural effusion was increased in this study, the cell layer between the red blood cells and the supernatant was aspirated for uniform smearing, and the precision was enhanced in each link, thereby improving the detection sensitivity for malignant pleural effusion up to 80.65%, consistent with that (82.7%) reported by Yu *et al* (8). In addition, it was found that adenocarcinoma cells were dominant in pleural effusion, accounting for 86.24%, consistent with literature reports in China and foreign countries (13). Adenocarcinoma cells mostly from lung cancer invade the pleura, leading to pleural metastasis and ultimately producing pleural effusion, which may be responsible for the sensitivity of exfoliative cytology towards lung cancer. Thus, other markers should be combined.

Tumor markers are metabolites such as antigens, enzymes or hormones produced and released by tumor cells into tissues or body fluids, whose content is high in tumor tissues but extremely low in normal human body. In this study, it was also found that the levels of all tumor markers except SCC-Ag and TSGF were significantly higher in pleural effusion of patients with lung cancer than those in healthy people, consistent with the results of a large number of studies (14). Tumor markers are playing increasingly prominent roles in early diagnosis and prognostic monitoring of tumors, which have gradually become important clinical reference indexes. According to many studies, the content of tumor markers in pleural effusion is far higher than that in serum (15).

In China, most of the research on tumor markers in the diagnosis of lung cancer focuses on the diagnostic value of the combined detection of tumor markers, but how to obtain the combination of tumor markers has not been clarified. In this study, the lung cancer prediction model was established through an exploratory study, which contained three specific tumor markers for lung cancer (NSE, CYFRA21-1 and CEA) and two non-specific tumor markers (CA125 and CA15-3), as well as exfoliative cytology. Both NSE and CYFRA21-1 have important diagnostic value for non-small cell lung cancer (NSCLC), and CYFRA21-1 with sensitivity of 24-69% and specificity of 88% is the most sensitive tumor marker for NSCLC (16). CEA is a good marker for lung cancer, colorectal cancer and breast cancer (17). In this study, NSE, CYFRA21-1 and CEA were all contained in common factors 2/3/4, while exfoliated cytology was contained in common factor 1. The combination of detection of specific tumor markers for lung cancer

and cytology can make up for the deficiency of a single tumor marker in diagnostic efficacy. Besides, CA125 and CA15-3 are specific markers for ovarian cancer and pancreatic cancer, respectively, rather than lung cancer. However, it has been found that the immune molecule OC125 recognizing lung adenocarcinoma cells is the same as CA125 (18), and CA15-3 is also found in lung adenocarcinoma cells (19). Another study revealed that the metastasis of cancer cells in NSCLS patients can be monitored by detecting CA125 and CA15-3 (18). To sum up, CA125 and CA15-3 combined with NSE, CYFRA21-1, CEA and cytology have certain significance in auxiliary diagnosis for patients highly suspected of lung cancer.

Studies have found that the optimal combination of 3 tumor markers for early diagnosis of lung cancer is NSE+CYFRA21-1+CEA (20), and NSE+CYFRA21-1+CEA+CA125 is the optimal combination of 4 tumor markers (21). The results may be attributed to regional differences that give different optimal indices for diagnosis and prediction. To further illustrate the rationality of the lung cancer prediction model established in this study, logistic regression analysis was conducted for validation of combined detection of exfoliative cells and tumor markers in pleural effusion in lung cancer patients and healthy people. The results showed that compared with the combined detection, both AUC and Youden index of the lung cancer prediction model were the optimal, and its true positive detection rate, sensitivity and specificity were also greatly improved.

In summary, the diagnostic indexes for lung cancer were optimized and combined through clinical practice combined with statistical analysis, and 6 indexes were selected out of 11 tumor markers and cytology commonly used in the detection of pleural effusion to establish the combined model. The lung cancer prediction model balances sensitivity and specificity without reducing the diagnostic efficacy, and also eases the economic burden of patients. Nevertheless, the model should be further validated by combining the results of pathological examination.

Authors contribution

LZ: Study design
JL: Data collection
JX: Data analysis
AC: Writing
All authors read and approved the final manuscript.

Declaration of interest

The authors declare that they have no conflict of interest.

References

1. Tokito T, Azuma K, Yamada K, Naito Y, Matsuo N, Ishii H, et al. Prognostic Value of Serum Tumor Markers in Patients With Stage III NSCLC Treated With Chemoradiotherapy. In Vivo. 2019;33(3):889-95. DOI: 10.21873/invivo.11555
2. Chen Z, Wang Y, Fang M. Analysis of tumor markers in pleural effusion and serum to verify the correlations between serum tumor markers and tumor size, TNM stage of lung adenocarcinoma. Cancer Med. 2020;9(4):1392-99. DOI: 10.1002/cam4.2809
3. Tetikkurt C, Yılmaz N, Tetikkurt S, Gundogdu Ş, Disci R. The value of exfoliative cell cytology in the diagnosis of exudative pleural effusions. Monaldi Arch Chest Dis. 2018;88(3):944. DOI: 10.4081/monaldi.2018.944
4. Nithin KU, Sridhar MG, Srilatha K, Habebullah S. CA 125 is a better marker to differentiate endometrial cancer and abnormal uterine bleeding. Afr Health Sci. 2018;18(4):972-8. DOI: 10.4314/ahs.v18i4.17
5. Gomes TA, Campos EA, Yoshida A, Sarian LO, Andrade LALA, Derchain SF. Preoperative Differentiation of Benign and Malignant Non-epithelial Ovarian Tumors: Clinical Features and Tumor Markers. Rev Bras Ginecol Obstet. 2020;42(9):555-61. DOI: 10.1055/s-0040-1712993

6. Georgantzi K, Sköldenberg EG, Stridsberg M, Kogner P, Jakobson Å, Janson ET, et al. Chromogranin A and neuron-specific enolase in neuroblastoma: Correlation to stage and prognostic factors. Pediatr Hematol Oncol. 2018;35(2):156-65. DOI: 10.1080/08880018.2018.1464087
7. Qu T, Zhang J, Xu N, Liu B, Li M, Liu A, et al. Diagnostic value analysis of combined detection of Trx, CYFRA21-1 and SCCA in lung cancer. Oncol Lett. 2019;17(5):4293-8. DOI: 10.3892/ol.2019.10073
8. Yu RZ, Huang MT. [The Diagnostic Value of Tumor Markers Combined with Fiberoptic Bronchoscopic Brushing in Lung Carcinoma]. Labeled Immunoassays and Clinical Medicine. 2014;21:360-3.
9. Grunnet M, Sorensen JB. Carcinoembryonic antigen (CEA) as tumor marker in lung cancer. Lung Cancer. 2012;76(2):138-43. DOI: 10.1016/j.lungcan.2011.11.012
10. Villalobos P, Wistuba II. Lung Cancer Biomarkers. Hematol Oncol Clin North Am. 2017;31(1):13-29. DOI: 10.1016/j.hoc.2016.08.006
11. Duffy MJ, Sturgeon C, Lamerz R, Haglund C, Holubec VL, Klapdor R, et al. Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. Ann Oncol. 2010;21(3):441-7. DOI: 10.1093/annonc/mdp332
12. O'Leary PC, Feddema PH, Michelangeli VP, Leedman PJ, Chew GT, Knuiman M, et al. Investigations of thyroid hormones and antibodies based on a community health survey: the Busselton thyroid study. Clin Endocrinol (Oxf). 2006;64(1):97-104. DOI: 10.1111/j.1365-2265.2005.02424.x
13. Ferreiro L, Toubes ME, Valdés L. [Contribution of pleural fluid analysis to the diagnosis of pleural effusion]. Med Clin (Barc). 2015;145(4):171-7. DOI: 10.1016/j.medcli.2014.08.005
14. Duffy MJ, O'Byrne K. Tissue and Blood Biomarkers in Lung Cancer: A Review. Adv Clin Chem. 2018;86:1-21. DOI: 10.1016/bs.acc.2018.05.001
15. Allama AM, Abou-Elela DH, Ibrahim IM. Pleural and serum markers for diagnosis of malignant pleural effusion. Asian Cardiovasc Thorac Ann. 2020;28(9):560-5. DOI: 10.1177/0218492320948311
16. Dal Bello MG, Filiberti RA, Alama A, Orengo AM, Mussap M, Coco S, et al. The role of CEA, CYFRA21-1 and NSE in monitoring tumor response to Nivolumab in advanced non-small cell lung cancer (NSCLC) patients. J Transl Med. 2019;17(1):74. DOI: 10.1186/s12967-019-1828-0
17. Hao C, Zhang G, Zhang L. Serum CEA levels in 49 different types of cancer and noncancer diseases. Prog Mol Biol Transl Sci. 2019;162:213-27. DOI: 10.1016/bs.pmbts.2018.12.011
18. Osanyin GE, Okunade KS, Ayotunde Oluwole A. Association between serum CA125 levels in preeclampsia and its severity among women in Lagos, South-West Nigeria. Hypertens Pregnancy. 2018;37(2):93-7. DOI: 10.1080/10641955.2018.1460667
19. Chen C, Chen Q, Zhao Q, Liu M, Guo J. Value of Combined Detection of Serum CEA, CA72-4, CA19-9, CA15-3 and CA12-5 in the Diagnosis of Gastric Cancer. Ann Clin Lab Sci. 2017;47(3):260-3.
20. Wang B, He YJ, Tian YX, Yang RN, Zhu YR, Qiu H. Clinical utility of haptoglobin in combination with CEA, NSE and CYFRA21-1 for diagnosis of lung cancer. Asian Pac J Cancer Prev. 2014;15(22):9611-4. DOI: 10.7314/APJCP.2014.15.22.9611
21. Ren X, Zhang Y, Lyu Y, Jin B, Guo H, Wu J, et al. Lactate dehydrogenase and serum tumor markers for predicting metastatic status in geriatric patients with lung adenocarcinoma. Cancer Biomark. 2019;26(2):139-50. DOI: 10.3233/CBM-190201