

## Artificial neural network–based diagnostic models for lung cancer combining conventional indicators with tumor markers

Yanan Luo<sup>1</sup>, Hui Yuan<sup>2</sup>, Qin Pei<sup>1</sup>, Yiyu Chen<sup>1</sup>, Jiawen Xian<sup>1</sup>, Rongrong Du<sup>1</sup> and Ting Ye<sup>1</sup> 

<sup>1</sup>Department of Laboratory Medicine, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, P.R. China; <sup>2</sup>Department of Pathophysiology, Mudanjiang Medical University, Mudanjiang 157011, P.R. China  
Corresponding author: Ting Ye. Email: yeting1103@163.com

### Impact statement

In this study, we used artificial neural network – multilayer perceptron to establish a combined diagnosis model of lung cancer based on clinical data, including multiple laboratory indicators and tumor markers, so as to provide a more convenient, fast, and cheap detection method for early clinical screening and diagnosis of lung cancer, and improve the early diagnosis rate of lung cancer. The diagnosis model has high diagnostic performance and far-reaching significance for early auxiliary diagnosis of lung cancer.

### Abstract

This study set out to establish a lung cancer diagnosis and prediction model uses conventional laboratory indicators combined with tumor markers, so as to help early screening and auxiliary diagnosis of lung cancer through a convenient, fast, and cheap way, and improve the early diagnosis rate of lung cancer. A total of 221 patients with lung cancer, 100 patients with benign pulmonary diseases, and 184 healthy subjects were retrospectively studied. General clinical data, the results of conventional laboratory indicators, and tumor markers were collected. Statistical Product and Service Solutions 26.0 was used for data analysis. The diagnosis and prediction model of lung cancer was established by artificial neural network – multilayer perceptron. After correlation and difference analysis, five comparison groups (lung cancer-benign lung disease group, lung cancer-health group, benign lung disease-health group, early-stage lung cancer-benign lung disease group, and early-stage lung cancer-health group) obtained 5, 28, 25, 16, and 25 valuable indicators for predicting lung cancer or benign lung disease, and then established

five diagnostic prediction models, respectively. The area under the curve (AUC) of each combined diagnostic prediction model (0.848, 0.989, 0.949, 0.841, and 0.976) was higher than that of the diagnostic prediction model established only using tumor markers (0.799, 0.941, 0.830, 0.661, and 0.850), and the difference in the lung cancer-health group, the benign lung disease-health group, the early-stage lung cancer-benign lung disease group, and early-stage lung cancer-health group was statistically significant ( $P < 0.05$ ). The artificial neural network–based diagnostic models for lung cancer combining conventional indicators with tumor markers have high performance and clinical significance in assisting the diagnosis of early lung cancer.

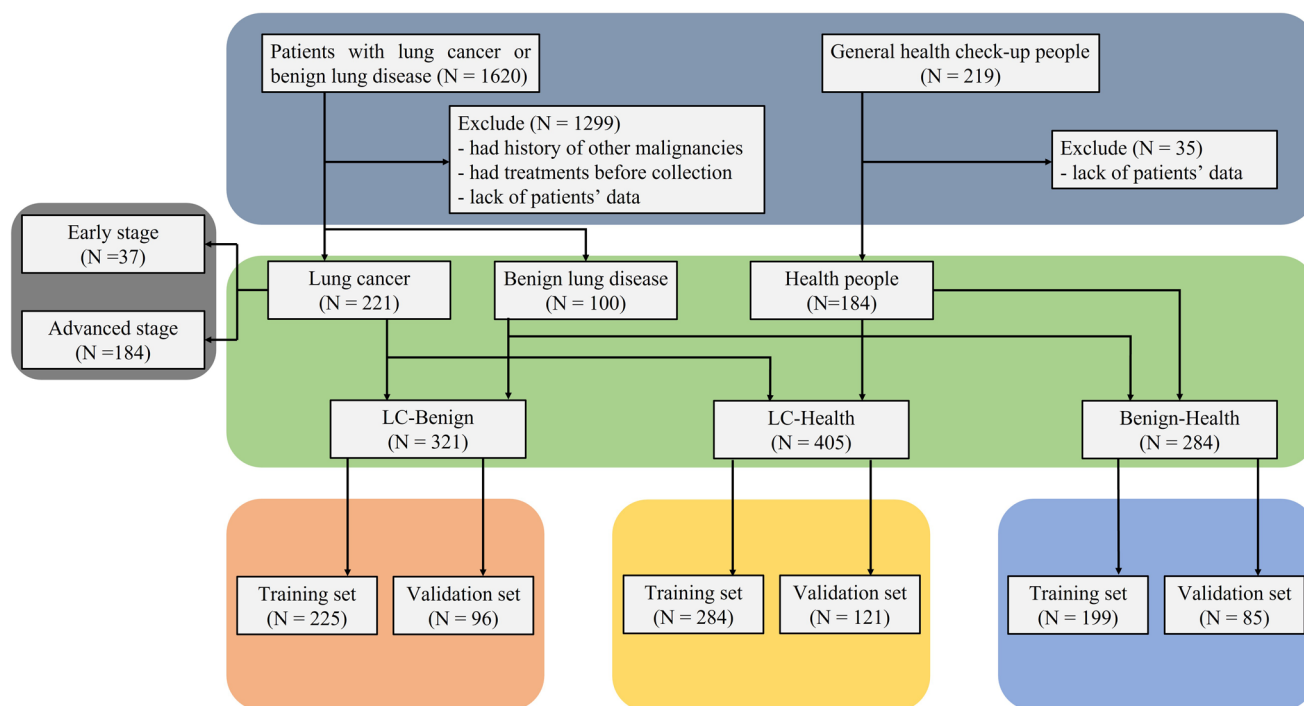
**Keywords:** Artificial neural networks, lung cancer, laboratory indicators, tumor markers, combined detection, diagnostic models

**Experimental Biology and Medicine 2023; 248: 829–838. DOI: 10.1177/15353702231177013**

### Introduction

Lung cancer is the leading cause of cancer death.<sup>1</sup> The survival rate of lung cancer is low, with a five-year survival rate of only 19.4%,<sup>2</sup> whereas the five-year overall survival rate for patients diagnosed with early-stage lung cancer is about 80%.<sup>3,4</sup> However, most patients with lung cancer do not show obvious symptoms in the early stage, so they are less likely to be diagnosed early, and most of them are in advanced stage or have metastases when they are diagnosed. Therefore, early detection is essential to improve the diagnosis rate of lung cancer and reduce lung cancer–related mortality.

Lung biopsy is the gold standard for the diagnosis of lung cancer, but lung biopsy is an invasive test and the risk of invasive procedures cannot be avoided. Existing imaging techniques such as computed tomography (CT) have proven to be effective in screening high-risk populations for lung cancer.<sup>3</sup> High-risk patients who underwent CT screening had a 20% decrease in lung cancer mortality.<sup>3,5</sup> However, CT has low specificity and high cost in the detection of lung cancer, and the radiation received by repeated scanning will cause certain damage to the human body.<sup>3,6</sup> The use of biomarkers in peripheral blood to predict cancer is a convenient, non-invasive, and widely accepted method. Tumor markers are



**Figure 1.** Flow chart of diagnosis model construction.

LC-Benign: lung cancer-benign lung disease group; LC-Health: lung cancer-health group; Benign-Health: benign lung disease-health group. Early stage: lung cancer patients in stage I and stage IIA. Advanced stage: lung cancer patients in stage IIA, stage III, and stage IV.

molecules secreted by tumor cells or produced by the interaction between tumor and host cells during carcinogenesis, and their appearance or level changes can reflect the existence of tumors,<sup>7</sup> which is of great significance in the screening, diagnosis, and therapeutic evaluation of lung cancer. However, the sensitivity of tumor markers in patients with early lung cancer is inferiority.<sup>8</sup> Moreover, no single serum tumor marker can be specifically used to detect lung cancer.<sup>9,10</sup> To improve the early diagnosis rate of lung cancer, it is a research hotspot to combine the conventional biological indicators and tumor markers in peripheral blood.

At present, artificial intelligence is widely used to assist in the diagnosis of cancer. Artificial neural network (ANN) is a non-linear and adaptive artificial intelligence method for information processing, which is composed of a large number of interconnected processing units. It is often used to assist diagnosis and monitoring of diseases in medicine. The basic ANN structure has three layers: input layer, hidden layer, and output layer. The predictive variables at the patient level are represented as nodes in the input layer, the results of patients are represented as nodes in the output layer, and the nodes in the hidden layer contain intermediate values computed by ANN, and the hidden nodes allow the ANN to model the complex relationship between the input variables and the outcome.<sup>11</sup> There is a study that constructed ANN prediction model and compared it with Lung Imaging Reporting and Data System (Lung-RADS). It was found that the sensitivity of the ANN model for lung cancer detection in Asian population was 75.0%, while the sensitivity of Lung-RADS was only 12.5%. Therefore, it is believed that ANN can provide higher sensitivity for lung cancer detection in Asian population.<sup>12</sup>

Therefore, this study uses ANN – multilayer perceptron to establish a combined diagnosis model of lung cancer based on clinical data, including multiple laboratory indicators and tumor markers, so as to provide a more convenient, fast, and cheap detection method for early clinical screening and diagnosis of lung cancer, and improve the early diagnosis rate of lung cancer.

## Materials and methods

### Clinical sample collection

The general clinical characteristics (including age and gender) and laboratory indexes (including tumor markers, liver function indexes, renal function indexes, electrolytes, blood routine indexes, coagulation indexes) of patients diagnosed with lung cancer and benign lung diseases in the Affiliated Hospital of Southwest Medical University from 2013 to 2021 were retrospectively collected. A total of 1620 cases were analyzed according to the inclusion and exclusion criteria, and 321 cases were included for follow-up analysis. Of these, 221 cases were lung cancer patients, all confirmed by pathology biopsy; 100 cases were patients with benign lung diseases. The general clinical characteristics and laboratory indexes of 219 healthy people were collected. After analyzing the data according to the inclusion and exclusion criteria, 184 cases were included for follow-up analysis. The three types of subjects were divided into three comparison groups: lung cancer-benign lung disease group (321 cases), lung cancer-health group (405 cases), and benign lung disease-health group (284 cases) (Figure 1). Non-small cell lung cancer was classified into stage I, stage IIA, stage IIB, stage III, and stage IV;<sup>13</sup> small cell lung cancer (SCLC) was classified into limited

and extensive stages. Lung cancer in stage I, stage IIA, and limited stage was classified as early-stage lung cancer (37 cases), and the rest were classified as advanced-stage lung cancer (184 cases) (Figure 1). The screening of indicators and the establishment of diagnostic prediction models were performed separately within each comparison group.

### Inclusion and exclusion criteria

**Cases of lung cancer:** cases with a confirmed diagnosis of lung cancer by pathological biopsy were included; cases with other cancers or a history of other cancers were excluded; cases that had been treated were excluded; and cases with incomplete data were excluded. **Cases of benign lung disease:** cases with confirmed benign lung disease were included; cases with a history of lung cancer or other cancers were excluded; cases with suspected lung cancer were excluded; and cases with incomplete data were excluded. **Healthy individuals:** cases that did not suggest a lung mass, nodules, and other suspected lung lesions by X-ray or CT examination and have no history of cancer were included; cases with incomplete data were excluded.

### Statistical analysis

Data analysis and graphing were performed using Statistical Product and Service Solutions 26.0. Normally distributed measures were described using  $X \pm SD$ , and non-normally distributed measures were described using median and interquartile range. Numeration data are described using frequency (percentage). The UNIFORM function was used to generate random values. The data of lung cancer patients, benign lung disease patients, and healthy subjects were randomly divided into training set and validation set with the ratio of training set / validation set = 7 / 3 using random values. Spearman correlation analysis was used for bivariate correlation testing. Independent sample *t*-test was used for comparison of normally distributed measurement data between groups. The Mann-Whitney non-parametric test was used for comparison of skewed distribution measurement data between groups. A chi-square test was used for comparison of numeration data between groups. In all analyses, a value  $P < 0.05$  was considered to be statistically significant. ANN – multilayer perceptron was used to establish lung cancer diagnosis prediction models. The receiver operating characteristic (ROC) curve was drawn and the area under the curve (AUC) was calculated to compare the diagnostic prediction models.

### Model construction

The samples are split according to the ratio of 7:3 between the training set and the test set; the architecture is taken as custom; two layers are specified for the hidden layer and a hyperbolic tangent activation function is used; the number of units is calculated automatically in the hidden layer; the sigmoid function is selected for the activation function of the output layer; the training type is processed using batches; the standard conjugate gradient optimization algorithm is used; the predicted values are saved, and the final synaptic weight estimates are exported to XML file to determine the

**Table 1.** Analysis of general clinical characteristics of serum samples.

| Study population | Gender      |            | P value | Age                        |
|------------------|-------------|------------|---------|----------------------------|
|                  | Man         | Woman      |         |                            |
| Lung cancer      | 173 (78.3%) | 48 (21.7%) | 0.000   | 62.89 ± 10.05 <sup>△</sup> |
| Benign           | 69 (69.0%)  | 31 (31.0%) | 0.000   | 58.34 ± 15.05 <sup>*</sup> |
| Health           | 94 (51.1%)  | 90 (48.9%) | 0.768   | 43.49 ± 8.91               |

Compared with healthy people, <sup>\*</sup> $P < 0.05$ ; compared with patients with benign lung disease, <sup>△</sup> $P < 0.05$ . Lung cancer: patients with lung cancer. Benign: patients with benign lung disease. Health: healthy people.

neural network model. The models' data is also exported (Supplemental Figures 1-10).

## Results

### Analysis of general clinical characteristics of subjects

Table 1 shows the general clinical characteristics of all samples. The average age of patients with lung cancer was higher than that of patients with benign lung disease and healthy people, and the difference was statistically significant ( $P < 0.05$ ). The average age of patients with benign lung diseases was higher than that of healthy people ( $P < 0.05$ ). The proportion of males in patients with lung cancer and benign lung diseases was higher than that of females ( $P < 0.05$ ), and the difference was statistically significant. There was no significant difference in gender distribution among healthy people ( $P > 0.05$ ).

### Analysis of correlation between indexes and the risk of lung cancer or benign lung disease

In the lung cancer-benign lung disease comparison group, three tumor markers (carcinoembryonic antigen [CEA], neuron-specific enolase [NSE], and carbohydrate antigen 153 [CA153]), carbon dioxide (CO<sub>2</sub>), and thrombin time (TT) were associated with an increased risk of lung cancer. The correlation coefficients were 0.314, 0.391, 0.207, 0.239, and 0.203, respectively. In the lung cancer-health comparison group, 28 indicators were associated with the risk of lung cancer. Among them, age, CEA, NSE, carbohydrate antigen 125 (CA125), CA153, globulin (GLO), r-glutamyl transferase (GGT), alkaline phosphatase (ALP), white blood cell (WBC), neutrophil (NEU), monocyte (MONO), neutrophil rate (NEU-R), monocyte rate (MONO-R), red blood cell volume distribution width standard deviation (RDW-SD), and red blood cell volume distribution width coefficient of variation (RDW-CV) were associated with the increased risk of lung cancer. There was a strong correlation between age ( $r = 0.701$ ), NEU ( $r = 0.577$ ), CA125 ( $r = 0.569$ ), and the increased risk of lung cancer. In the benign lung disease-health comparison group, 25 indicators were associated with the risk of benign lung disease, of which 11 were positively correlated (age, CEA, CA125, GLO, GGT, ALP, WBC, NEU, MONO, NEU-R, and RDW-CV). NEU-R ( $r = 0.636$ ), NEU ( $r = 0.595$ ), and CA125 ( $r = 0.550$ ) were strongly correlated with the increased risk of benign lung diseases (Table 2).

**Table 2.** Indicators associated with the risk of lung cancer or benign lung disease.

| Index                    | LC-Benign            |                       | LC-Health            |                        | Benign-Health        |                       |
|--------------------------|----------------------|-----------------------|----------------------|------------------------|----------------------|-----------------------|
|                          | Training set (N=225) | Validation set (N=96) | Training set (N=284) | Validation set (N=121) | Training set (N=199) | Validation set (N=85) |
| Age                      | 0.184*               | 0.111                 | 0.741*               | 0.701*                 | 0.492**              | 0.526**               |
| CEA (ng/mL)              | 0.299*               | 0.314*                | 0.605*               | 0.501*                 | 0.353**              | 0.226**               |
| NSE (ng/mL)              | 0.408*               | 0.391*                | 0.479*               | 0.517*                 | -0.012               | -                     |
| CA125 (U/mL)             | 0.145*               | 0.051                 | 0.614*               | 0.569*                 | 0.476**              | 0.550**               |
| CA153 (U/mL)             | 0.197*               | 0.207*                | 0.340*               | 0.278*                 | 0.112                | -                     |
| CA724 (IU/mL)            | -0.025               | -                     | -0.139*              | 0.012                  | -0.136               | -                     |
| AST/ALT                  | 0.025                | -                     | 0.144*               | 0.156                  | 0.118                | -                     |
| TP (g/L)                 | 0.092                | -                     | -0.129*              | -0.161                 | -0.227**             | -0.220**              |
| ALB (g/L)                | 0.024                | -                     | -0.641*              | -0.587*                | -0.614**             | -0.601**              |
| GLO (g/L)                | 0.099                | -                     | 0.408*               | 0.348*                 | 0.289**              | 0.397**               |
| AGR                      | -0.064               | -                     | -0.624*              | -0.571*                | -0.585**             | -0.593**              |
| TBIL ( $\mu$ mol/L)      | -0.083               | -                     | -0.457*              | -0.416*                | -0.348**             | -0.108                |
| DBIL ( $\mu$ mol/L)      | -0.006               | -                     | -0.241*              | -0.146                 | -0.201**             | 0.026                 |
| IBIL ( $\mu$ mol/L)      | -0.080               | -                     | -0.506*              | -0.479*                | -0.423**             | -0.255**              |
| GGT (U/L)                | -0.032               | -                     | 0.330*               | 0.277*                 | 0.346**              | 0.326**               |
| ALP (U/L)                | 0.184*               | 0.055                 | 0.491*               | 0.273*                 | 0.299**              | 0.236**               |
| Urea (mmol/L)            | 0.053                | -                     | 0.226*               | 0.080                  | 0.157**              | 0.071                 |
| UA ( $\mu$ mol/L)        | 0.166*               | 0.065                 | -0.141*              | -0.135                 | -0.305**             | -0.216**              |
| WBC ( $10^9$ /L)         | -0.110               | -                     | 0.389*               | 0.479*                 | 0.498**              | 0.529**               |
| NEU ( $10^9$ /L)         | -0.093               | -                     | 0.476*               | 0.577*                 | 0.547**              | 0.595**               |
| LYM ( $10^9$ /L)         | 0.012                | -                     | -0.430*              | -0.418*                | -0.364**             | -0.491**              |
| MONO ( $10^9$ /L)        | -0.073               | -                     | 0.539*               | 0.524*                 | 0.537**              | 0.505**               |
| NEU-R (%)                | -0.061               | -                     | 0.534*               | 0.549*                 | 0.515**              | 0.636**               |
| LYM-R (%)                | 0.109                | -                     | -0.626*              | -0.627*                | -0.620**             | -0.698**              |
| MONO-R (%)               | 0.012                | -                     | 0.293*               | 0.229*                 | 0.250**              | 0.181                 |
| EOS-R (%)                | 0.059                | -                     | -0.160*              | -0.035                 | -0.209**             | -0.265**              |
| BASO-R (%)               | -0.011               | -                     | -0.294*              | -0.360*                | -0.299**             | -0.332**              |
| RBC ( $10^{12}$ /L)      | -0.086               | -                     | -0.440*              | -0.361*                | -0.299**             | -0.273**              |
| HGB (g/L)                | -0.059               | -                     | -0.471*              | -0.353*                | -0.368**             | -0.260**              |
| HCT                      | -0.082               | -                     | -0.503*              | -0.385*                | -0.376**             | -0.315**              |
| MCV (fL)                 | -0.073               | -                     | -0.255*              | -0.168                 | -0.156**             | -0.130                |
| MCH (pg)                 | -0.056               | -                     | -0.241*              | -0.118                 | -0.180**             | -0.102                |
| MCHC (g/L)               | 0.088                | -                     | -0.062               | -                      | -0.154**             | 0.115                 |
| RDW-SD (fL)              | 0.193*               | 0.170                 | 0.339*               | 0.283*                 | 0.088                | -                     |
| RDW-CV (%)               | 0.003                | -                     | 0.260*               | 0.243*                 | 0.213**              | 0.268**               |
| MPV (fL)                 | 0.243*               | 0.123                 | -0.128*              | -0.208*                | -0.374**             | -0.367**              |
| PDW (%)                  | -0.090               | -                     | -0.244*              | -0.309*                | -0.175**             | -0.192                |
| P-LCR (%)                | 0.220*               | 0.102                 | -0.151*              | -0.228*                | -0.367**             | -0.374**              |
| K <sup>+</sup> (mmol/L)  | 0.206*               | 0.033                 | -                    | -                      | -                    | -                     |
| Cl <sup>-</sup> (mmol/L) | -0.190*              | -0.130                | -                    | -                      | -                    | -                     |
| CO <sub>2</sub> (mmol/L) | 0.258*               | 0.239*                | -                    | -                      | -                    | -                     |
| TT (s)                   | 0.294*               | 0.203*                | -                    | -                      | -                    | -                     |

LC-Benign: lung cancer-benign lung disease group. LC-Health: lung cancer-health group. Benign-Health: benign lung disease-health group. \*The association with the risk of lung cancer is not caused by random errors ( $P < 0.05$ ). \*\*The association with the risk of benign lung disease is not caused by random errors ( $P < 0.05$ ). CEA: carcinoembryonic antigen; NSE: neuron-specific enolase; CA125: carbohydrate antigen 125; CA153: carbohydrate antigen 153; CA724: carbohydrate antigen 724; AST: aspartate aminotransferase; ALT: alanine aminotransferase; TP: total protein; ALB: albumin; GLO: globulin; AGR: albumin/globulin ratio; TBIL: total bilirubin; DBIL: direct bilirubin; IBIL: indirect bilirubin; GGT:  $\gamma$ -glutamyl transferase; ALP: alkaline phosphatase; UA: uric acid; WBC: white blood cell; NEU: neutrophil; LYM: lymphocyte; MONO: monocyte; NEU-R: neutrophil rate; LYM-R: lymphocyte rate; MONO-R: monocyte rate; EOS-R: eosinophil rate; BASO-R: basophil rate; RBC: red blood cell; HGB: hemoglobin; HCT: red blood cell hematocrit; MCV: mean volume of red blood cells; MCH: mean hemoglobin of red blood cells; MCHC: mean hemoglobin concentration of red blood cells; RDW-SD: red blood cell volume distribution width standard deviation; RDW-CV: red blood cell volume distribution width coefficient of variation; MPV: mean platelet volume; PDW: platelet volume distribution width; P-LCR: large platelet ratio; K<sup>+</sup>: potassium ion; Cl<sup>-</sup>: chloride ion; CO<sub>2</sub>: carbon dioxide; TT: thrombin time.

### Differential analysis of laboratory indicators in lung cancer or benign lung disease

Differential analysis was performed using indicators related to the risk of lung cancer or benign lung disease to screen for indicators for identification. In the lung cancer-benign lung

disease group, the serum levels of five indicators related to the risk of lung cancer were significantly different between lung cancer patients and benign lung disease patients ( $P < 0.05$ ). In the lung cancer-health group, 28 indicators related to the risk of lung cancer were significantly different



**Table 3.** Valuable indicators for the diagnosis of lung cancer in lung cancer-benign lung disease group.

| Index                    | Lung cancer          | Benign               | <i>P</i> value |
|--------------------------|----------------------|----------------------|----------------|
| CEA (ng/mL)              | 5.39 (2.65, 30.32)   | 3.03 (2.23, 4.71)    | 0.002          |
| NSE (ng/mL)              | 5.57 (3.60, 11.17)   | 3.20 (2.66, 4.52)    | 0.000          |
| CA153 (U/mL)             | 15.88 (11.10, 23.91) | 13.67 (8.91, 18.20)  | 0.044          |
| CO <sub>2</sub> (mmol/L) | 26.89 ± 3.07         | 25.15 ± 2.95         | 0.011          |
| TT (s)                   | 16.70 (15.90, 17.83) | 16.15 (15.60, 17.10) | 0.048          |

Lung cancer: patients with lung cancer. Benign: patients with benign lung disease. CEA: carcinoembryonic antigen. NSE: neuron-specific enolase. CA153: carbohydrate antigen 153. CO<sub>2</sub>: carbon dioxide. TT: thrombin time.

**Table 4.** Valuable indicators for the diagnosis of lung cancer in lung cancer-health group.

| Index                     | Lung cancer           | Health               | <i>P</i> value |
|---------------------------|-----------------------|----------------------|----------------|
| Age                       | 62.14 ± 10.32         | 44.07 ± 8.72         | 0.000          |
| CEA (ng/mL)               | 5.39 (2.65, 30.32)    | 2.35 (1.63, 3.59)    | 0.000          |
| NSE (ng/mL)               | 5.57 (3.60, 11.17)    | 3.17 (2.66, 3.89)    | 0.000          |
| CA125 (U/mL)              | 27.58 (10.66, 61.33)  | 7.56 (5.26, 11.23)   | 0.000          |
| CA153 (U/mL)              | 15.88 (11.10, 23.91)  | 11.26 (9.8, 17.49)   | 0.002          |
| ALB (g/L)                 | 41.30 (38.10, 44.05)  | 45.96 ± 2.96         | 0.000          |
| GLO (g/L)                 | 28.75 (26.00, 31.15)  | 26.14 ± 3.61         | 0.000          |
| AGR                       | 1.44 ± 0.34           | 1.79 (1.64, 1.92)    | 0.000          |
| TBIL (μmol/L)             | 10.35 (7.60, 13.65)   | 14.30 (11.70, 18.80) | 0.000          |
| IBIL (μmol/L)             | 6.55 (4.58, 9.63)     | 10.60 (8.50, 13.80)  | 0.000          |
| GGT (U/L)                 | 38.25 (23.85, 60.85)  | 25.30 (12.70, 44.60) | 0.002          |
| ALP (U/L)                 | 89.10 (73.20, 116.73) | 77.00 (61.90, 89.10) | 0.003          |
| WBC (10 <sup>9</sup> /L)  | 7.55 (6.04, 9.88)     | 5.48 (4.90, 6.93)    | 0.000          |
| NEU (10 <sup>9</sup> /L)  | 5.51 (3.94, 7.29)     | 3.09 (2.63, 3.81)    | 0.000          |
| LYM (10 <sup>9</sup> /L)  | 1.30 (0.89, 1.66)     | 1.84 (1.47, 2.16)    | 0.000          |
| MONO (10 <sup>9</sup> /L) | 0.54 (0.40, 0.70)     | 0.34 (0.26, 0.42)    | 0.000          |
| NEU-R (%)                 | 72.5 (62.33, 79.80)   | 58.21 ± 9.49         | 0.000          |
| LYM-R (%)                 | 16.10 (11.93, 25.55)  | 33.06 ± 8.96         | 0.000          |
| MONO-R (%)                | 6.90 (5.50, 8.53)     | 5.80 (5.00, 7.20)    | 0.012          |
| BASO-R (%)                | 0.30 (0.20, 0.50)     | 0.50 (0.30, 0.60)    | 0.000          |
| RBC (10 <sup>12</sup> /L) | 4.39 ± 0.65           | 4.82 ± 0.55          | 0.000          |
| HGB (g/L)                 | 129.85 ± 19.57        | 143.60 ± 19.86       | 0.000          |
| HCT                       | 0.39 (0.37, 0.44)     | 0.44 (0.41, 0.48)    | 0.000          |
| RDW-SD (fL)               | 43.90 (41.78, 47.80)  | 42.18 ± 2.93         | 0.002          |
| RDW-CV (%)                | 13.70 (12.98, 14.55)  | 13.10 (12.70, 13.60) | 0.008          |
| MPV (fL)                  | 9.70 (9.08, 11.05)    | 10.63 ± 1.37         | 0.023          |
| PDW (%)                   | 16.00 (14.58, 16.30)  | 16.20 (15.80, 16.60) | 0.001          |
| P-LCR (%)                 | 24.10 (19.18, 34.55)  | 31.01 ± 9.37         | 0.013          |

Lung cancer: patients with lung cancer. Health: healthy people. CEA: carcinoembryonic antigen; NSE: neuron-specific enolase; CA125: carbohydrate antigen 125; CA153: carbohydrate antigen 153; ALB: albumin; GLO: globulin; AGR: albumin/globulin ratio; TBIL: total bilirubin; IBIL: indirect bilirubin; GGT: r-glutamyl transferase; ALP: alkaline phosphatase; WBC: white blood cell; NEU: neutrophil; LYM: lymphocyte; MONO: monocyte; NEU-R: neutrophil rate; LYM-R: lymphocyte rate; MONO-R: monocyte rate; BASO-R: basophil rate; RBC: red blood cell; HGB: hemoglobin; HCT: red blood cell hematocrit; RDW-SD: red blood cell volume distribution width standard deviation; RDW-CV: red blood cell volume distribution width coefficient of variation; MPV: mean platelet volume; PDW: platelet volume distribution width; P-LCR: large platelet ratio.

between lung cancer and healthy people ( $P < 0.05$ ). In the benign lung disease-health group, 25 indicators related to the risk of benign lung disease were different between patients with benign lung disease and healthy people, and the difference was statistically significant ( $P < 0.05$ ). It shows that these indicators are valuable for the diagnosis of lung cancer

**Table 5.** Valuable indicators for the diagnosis of benign lung disease in benign lung disease-health group.

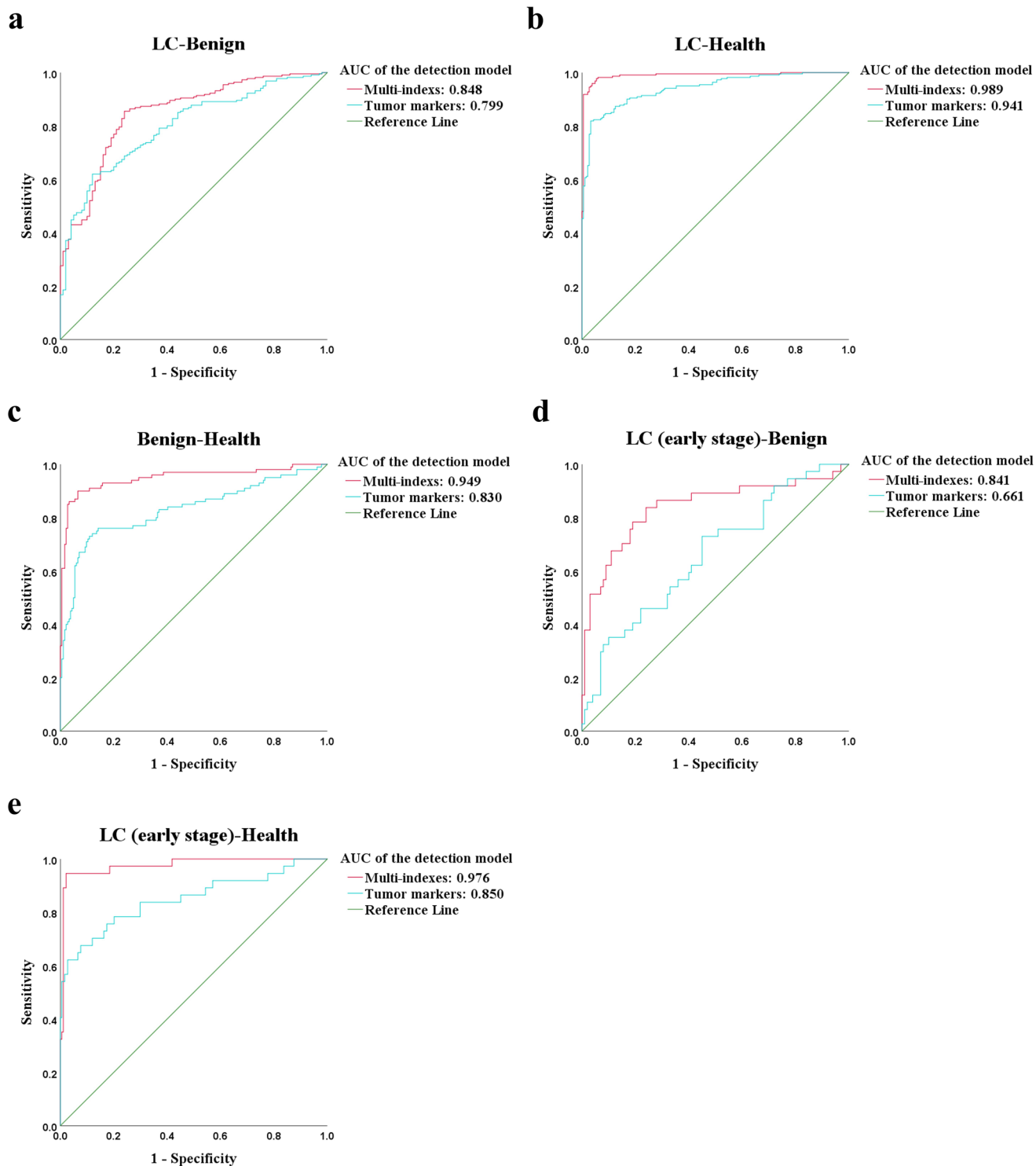
| Index                     | Benign               | Health               | <i>P</i> value |
|---------------------------|----------------------|----------------------|----------------|
| Age                       | 62.14 ± 10.32        | 44.07 ± 8.72         | 0.000          |
| CEA (ng/mL)               | 5.39 (2.65, 30.32)   | 2.35 (1.63, 3.59)    | 0.039          |
| CA125 (U/mL)              | 20.68 (11.06, 50.09) | 7.56 (5.26, 11.23)   | 0.000          |
| TP (g/L)                  | 68.98 ± 6.88         | 72.09 ± 3.82         | 0.027          |
| ALB (g/L)                 | 40.00 (35.43, 43.78) | 45.96 ± 2.96         | 0.000          |
| GLO (g/L)                 | 29.89 ± 5.27         | 26.14 ± 3.61         | 0.001          |
| AGR                       | 1.36 ± 0.38          | 1.79 (1.64, 1.92)    | 0.000          |
| IBIL (μmol/L)             | 8.40 (5.18, 11.60)   | 10.60 (8.50, 13.80)  | 0.019          |
| GGT (U/L)                 | 38.30 (24.78, 85.60) | 25.30 (12.70, 44.60) | 0.003          |
| ALP (U/L)                 | 86.95 (70.58, 98.98) | 77.00 (61.90, 89.10) | 0.030          |
| UA (μmol/L)               | 307.89 ± 101.25      | 260.50 ± 97.99       | 0.022          |
| WBC (10 <sup>9</sup> /L)  | 9.22 (7.09, 11.55)   | 5.48 (4.90, 6.93)    | 0.000          |
| NEU (10 <sup>9</sup> /L)  | 6.94 (5.03, 9.15)    | 3.09 (2.63, 3.81)    | 0.000          |
| LYM (10 <sup>9</sup> /L)  | 1.25 ± 0.57          | 1.84 (1.47, 2.16)    | 0.000          |
| MONO (10 <sup>9</sup> /L) | 0.65 ± 0.33          | 0.34 (0.26, 0.42)    | 0.000          |
| NEU-R (%)                 | 76.21 ± 12.19        | 58.21 ± 9.49         | 0.000          |
| LYM-R (%)                 | 14.74 ± 8.52         | 33.06 ± 8.96         | 0.000          |
| EOS-R (%)                 | 0.90 (0.58, 2.33)    | 1.80 (1.00, 2.70)    | 0.015          |
| BASO-R (%)                | 0.30 (0.20, 0.53)    | 0.50 (0.30, 0.60)    | 0.002          |
| RBC (10 <sup>12</sup> /L) | 4.30 ± 0.95          | 4.82 ± 0.55          | 0.008          |
| HGB (g/L)                 | 128.17 ± 29.62       | 143.60 ± 19.86       | 0.017          |
| HCT                       | 0.39 ± 0.86          | 0.44 (0.41, 0.48)    | 0.004          |
| RDW-CV (%)                | 13.60 (13.10, 14.55) | 13.10 (12.70, 13.60) | 0.014          |
| MPV (fL)                  | 9.60 ± 1.07          | 10.63 ± 1.37         | 0.001          |
| P-LCR (%)                 | 23.76 ± 7.33         | 31.01 ± 9.37         | 0.000          |

Benign: patients with benign lung disease. Health: healthy people. CEA: carcinoembryonic antigen; CA125: carbohydrate antigen 125; TP: total protein; ALB: albumin; GLO: globulin; AGR: albumin/globulin ratio; IBIL: indirect bilirubin; GGT: r-glutamyl transferase; ALP: alkaline phosphatase; UA: uric acid; WBC: white blood cell; NEU: neutrophil; LYM: lymphocyte; MONO: monocyte; NEU-R: neutrophil rate; LYM-R: lymphocyte rate; EOS-R: eosinophil rate; BASO-R: basophil rate; RBC: red blood cell; HGB: hemoglobin; HCT: red blood cell hematocrit; RDW-CV: red blood cell volume distribution width coefficient of variation; MPV: mean platelet volume; P-LCR: large platelet ratio.

or benign lung diseases and can be included in the establishment of diagnostic prediction models (Tables 3 to 5).

### Establishment and evaluation of diagnostic prediction model

In the lung cancer-benign lung disease group, five indicators were used to establish a diagnostic prediction model. The AUC of the model was 0.848; the sensitivity and specificity were 85.5% and 76.0%, respectively; and the Youden index was 0.615. In the lung cancer-health group, a diagnostic prediction model was established using 28 indicators. The AUC of the model was 0.989; the sensitivity and specificity were 97.7% and 94.6%, respectively; and the Youden index was 0.923. The benign lung disease-health group used 25 indicators to establish a diagnostic prediction model. The AUC of the model was 0.949; the sensitivity and specificity were 90.0% and 93.5%, respectively; and the Youden index was 0.835. The AUC of the three combined diagnostic prediction models was higher than that of the model using only tumor markers (0.799, 0.941, and 0.830), and the difference in the lung cancer-health group and the benign lung disease-health group was statistically significant ( $P < 0.05$ ) (Figure 2). It shows that the diagnostic prediction model established by



**Figure 2.** ROC curve of five diagnostic models. (a) ROC curve of the diagnostic model of lung cancer-benign lung disease group; (b) ROC curve of the diagnostic model of lung cancer-health group; (c) ROC curve of the diagnostic model of benign lung disease-health group; (d) ROC curve of the diagnostic model of early-stage lung cancer-benign lung disease group; (e) ROC curve of the diagnostic model of early-stage lung cancer-health group. ROC: receiver operating characteristic; LC-Benign: lung cancer-benign lung disease group; LC-Health: lung cancer-health group; Benign-Health: benign lung disease-health group; LC (early stage)-Benign: early-stage lung cancer-benign lung disease group; LC (early stage)-Health: early-stage lung cancer-health group.

combining conventional laboratory indicators and tumor markers has better performance in the diagnosis of lung cancer than the prediction model established by using tumor markers alone (Table 6).

### Diagnostic model of early-stage lung cancer

In the early-stage lung cancer-benign lung disease group, 16 indicators (CEA, NSE, uric acid [UA], CO<sub>2</sub>, WBC, NEU, NEU-R, lymphocyte rate [LYM-R], eosinophil rate [EOS-R],

**Table 6.** Discriminant analysis of diagnostic models.

| Group                   | Model         | AUC   | <i>P</i> value | Sensitivity (%) | Specificity (%) | Youden index | Accuracy (%) |
|-------------------------|---------------|-------|----------------|-----------------|-----------------|--------------|--------------|
| LC-Benign               | Multi-indexes | 0.848 | 0.149          | 85.5            | 76.0            | 0.615        | 82.8         |
|                         | Tumor markers | 0.799 |                | 62.0            | 88.0            | 0.500        | 80.0         |
| LC-Health               | Multi-indexes | 0.989 | 0.000          | 97.7            | 94.6            | 0.923        | 96.2         |
|                         | Tumor markers | 0.941 |                | 81.9            | 96.7            | 0.786        | 89.4         |
| Benign-Health           | Multi-indexes | 0.949 | 0.000          | 90.0            | 93.5            | 0.835        | 93.3         |
|                         | Tumor markers | 0.830 |                | 73.0            | 89.1            | 0.621        | 84.4         |
| LC (early stage)-Benign | Multi-indexes | 0.841 | 0.000          | 83.8            | 76.0            | 0.598        | 82.6         |
|                         | Tumor markers | 0.661 |                | 73.0            | 55.0            | 0.280        | 77.1         |
| LC (early stage)-Health | Multi-indexes | 0.976 | 0.004          | 94.6            | 97.8            | 0.924        | 96.4         |
|                         | Tumor markers | 0.850 |                | 67.6            | 92.4            | 0.600        | 92.9         |

LC-Benign: lung cancer-benign lung disease group. LC-Health: lung cancer-health group. Benign-Health: benign lung disease-health group. LC (early stage)-Benign: early-stage lung cancer-benign lung disease group. LC (early stage)-Health: early-stage lung cancer-health group. Multi-indexes: diagnostic model constructed by combining laboratory parameters with tumor markers. Tumor markers: diagnostic model constructed using only tumor markers. AUC: area under the curve.

mean hemoglobin concentration of red blood cells [MCHC], RDW-SD, mean platelet volume [MPV], large platelet ratio [P-LCR], prothrombin time [PT], prothrombin time activity [PT-%], and TT) were associated with early-stage lung cancer. Moreover, the levels of these 16 indicators were significantly different between early-stage lung cancer and benign lung disease patients. Then, use them to establish a diagnostic prediction model. The AUC of the model was 0.841 (Figure 2); the sensitivity and specificity were 83.8% and 76.0%, respectively; and the Youden index was 0.598. In the early-stage lung cancer-health group, 26 indicators were associated with early-stage lung cancer. Among them, 25 indicators (age, CEA, NSE, CA125, carbohydrate antigen 724 [CA724], albumin [ALB], GLO, albumin to globulin ratio [AGR], total bilirubin [TBIL], direct bilirubin [DBIL], indirect bilirubin [IBIL], GGT, ALP, urea, WBC, NEU, lymphocyte [LYM], MONO, NEU-R, LYM-R, MONO-R, red blood cell [RBC], hemoglobin [HGB], red blood cell hematocrit [HCT], and RDW-SD) were significantly different between early-stage lung cancer and healthy people. The AUC of the model was 0.976 (Figure 2); the sensitivity and specificity were 94.6% and 97.8%, respectively; and the Youden index was 0.924 (Table 6). Both models have good diagnostic efficacy for differentiating early-stage lung cancer from healthy individuals and patients with benign lung disease.

## Discussion

The construction of a marker model for predicting lung cancer plays a significant role in the occurrence and development of the disease.<sup>14</sup> In this study, three diagnostic prediction models were established by combining conventional laboratory indicators and serum tumor markers using ANN, and their risk prediction performance was better than that of the model using only tumor markers.

Increasing age, smoking history, occupational exposure history, and family history of lung cancer are considered to be risk factors for lung cancer.<sup>15</sup> It has been found that the risk of lung cancer increases significantly with age.<sup>16</sup> In this study, the age distribution of lung cancer cases and benign lung disease cases was significantly different in the overall and training sets ( $P < 0.05$ ), but there was no significant correlation between age and the risk of lung cancer in the

validation set, so age was not included in the establishment of the prediction model for the lung cancer-benign lung disease group, which does not mean that age cannot be used to identify benign and malignant lung diseases, and the sample size will be expanded for further verification in the future. In the comparison between healthy people and patients with lung cancer or benign lung disease, age was significantly positively correlated with the risk of lung disease and had significant differences ( $P < 0.05$ ), and the age of healthy people was significantly lower than the other two groups, so we think age is a risk factor for lung disease, but it still needs to be considered in combination with other clinical features. Globally, men are more than twice as likely as women to be diagnosed with and die from lung cancer, but this may be related to the fact that men are more likely to smoke than women.<sup>17</sup> And epidemiological studies also have identified an association between smoking and an increased risk of lung cancer.<sup>18</sup> In addition, hormonal differences between men and women may be associated with the risk of lung cancer.<sup>17</sup>

CEA is a glycoprotein complex that is expressed and released in various cancers. CEA is increased in colorectal cancer, breast cancer, lung cancer, thyroid cancer, and other cancers. It is usually absent or low in the blood of healthy adults and is considered to be associated with poor prognosis of tumors.<sup>19–24</sup> CEA may also be elevated in various benign diseases, such as obstructive pulmonary disease, hepatitis, and pancreatitis.<sup>23</sup> NSE is the main enolase isoenzyme in neurons and neuroendocrine tissues, and its serum level is often elevated in diseases with neuronal damage.<sup>25</sup> SCLC is an aggressive neuroendocrine tumor. Currently, NSE is considered a reliable tumor marker for the diagnosis of SCLC<sup>26</sup> and has certain significance for the initial screening of lung cancer. CA125 is a transmembrane glycoprotein with increased expression in a variety of pathological states, such as lung and ovarian cancer; however, the level of CA125 is also elevated during menstruation and early pregnancy.<sup>27–29</sup> CA153 is a variant glycoprotein on the surface of breast cells. It was found that the serum level of CA153 was significantly higher in patients with breast cancer than in patients with benign breast lesions.<sup>7</sup> In this study, CEA and NSE were both positively correlated with benign and malignant lung lesions. We suggest that elevated serum levels of CEA or NSE may reflect a higher risk of benign lung

disease or even lung cancer. CEA, CA125, NSE, and CA153 were differentially expressed in patients with lung cancer, patients with benign lung disease, and healthy individuals ( $P < 0.05$ ) and were therefore included in the development of the diagnostic model.

ALB and GLO are the main components of serum proteins. ALB is produced by the liver and reflects the nutritional status of the body; GLO is calculated from a formula (total serum protein minus ALB) and has been shown to be associated with immunity and chronic inflammation.<sup>30</sup> Hypoalbuminemia is an independent risk factor for survival in several cancers, and all types of GLOs produced by chronic inflammation can promote tumor development.<sup>31</sup> In this study, serum ALB was negatively correlated with the risk of lung cancer or benign lung disease, and the level of ALB was significantly higher in healthy individuals than in patients with lung cancer or benign lung disease, and the difference was significant ( $P < 0.05$ ). There was a positive correlation between GLO and the risk of lung cancer or benign lung disease, with healthy individuals having lower GLO levels than the other two groups of patients, and the difference was statistically significant ( $P < 0.05$ ). Serum levels of ALB and GLO are susceptible to interference by several factors, but a low level of AGR is a valuable predictor of long-term mortality in lung adenocarcinoma.<sup>30</sup> A negative correlation between AGR and the risk of lung cancer or benign lung disease was observed in this study, with a significant difference between healthy individuals and the other two groups of patients ( $P < 0.05$ ). GGT is distributed on the cell membrane and is most abundant in the kidney, followed by the pancreas, liver, and lung. Persistently elevated levels of GGT may reflect chronic inflammation and oxidative stress, which are risk factors for cancer.<sup>32,33</sup> Recurrent elevations in GGT have been found to be associated with a higher risk of respiratory cancers.<sup>32</sup> In this study, the levels of GGT and ALP were higher in patients with lung cancer or benign lung disease than in healthy subjects, with statistically significant differences ( $P < 0.05$ ), which is consistent with previous studies.<sup>32,34</sup> Due to its anti-inflammatory and antioxidant properties, bilirubin may have a protective effect against the development of several chronic diseases, including cancer.<sup>35</sup> In this study, we found a negative association between total serum bilirubin, IBIL, and the risk of lung cancer, with significantly lower serum levels of total and indirect bilirubin in lung cancer patients than in healthy subjects ( $P < 0.05$ ), which is consistent with the findings of existing studies.<sup>36,37</sup>

The relationship between serum UA levels and the risk of cancers is more complex. UA has both antioxidant and pro-oxidant effects, so its role in different cancers is contradictory. High UA levels may be positively associated with an increased risk of colorectal cancer in women.<sup>38</sup> Studies have also shown a negative correlation between UA and the risk of lung cancer in men.<sup>39</sup> In the results of this study, there was a negative correlation between UA and the risk of benign lung diseases in healthy people. The difference in serum UA levels between the two groups was statistically significant ( $P < 0.05$ ). The correlation between UA and the risk of lung cancer was not statistically significant in this study ( $P > 0.05$ ); the relationship between them needs further study.

Chronic inflammation is a risk factor for cancer.<sup>40</sup> The results of this study showed that total WBC and NEU were positively correlated with the risk of lung cancer or benign lung disease. The levels of WBC or NEU in the two types of patients were significantly higher than those in healthy people ( $P < 0.05$ ). LYM was negatively correlated with the risk of lung cancer or benign lung disease. LYM in patients with lung cancer or benign lung disease was lower than that in healthy people ( $P < 0.05$ ). WBC can be divided into granulocytes (including NEU, eosinophils, and basophils), MONO, and LYM according to their morphology, function, and source. When there is an inflammatory response in the body, the level of WBC in the blood is usually elevated. Studies have found that increased total WBC levels are associated with an increased risk of lung cancer, and this relationship is mainly mediated by NEU.<sup>41</sup> The antibodies secreted by LYM can regulate cancer cells and activate the complement cascade, thereby inducing tumor death.<sup>42</sup> MONO has both pro-inflammatory and anti-inflammatory properties and has an inhibitory or promotional effect on tumors, which depends on the type of cancer.<sup>43,44</sup> Studies have found that MONO in peripheral blood of lung cancer patients increased significantly.<sup>45</sup> These are consistent with the results of our findings. The RDW is a parameter that reflects the volumetric heterogeneity of RBC and indicates the variability of the size of RBC in the blood. RDW is associated with multiple inflammatory markers, and this relationship suggests that the tumor-mediated inflammatory microenvironment can lead to increased RDW expression.<sup>46</sup> RBC is an important oxygen-carrying mediator, and when its size is variable or its function is impaired, it induces a hypoxic microenvironment, which in turn leads to tumor progression.<sup>47</sup> Anemia is the most common hematologic abnormality in cancer patients and is a direct or indirect consequence of tumor progression or antitumor therapy.<sup>48,49</sup> In this study, RBC and HGB were negatively associated with the risk of lung cancer, RDW was positively associated with the risk of lung cancer, and their differences in levels between lung cancer and healthy individuals were statistically significant ( $P < 0.05$ ). In summary, we believe that RBC, HGB, and RDW have great clinical significance in the differential diagnosis of lung cancer and healthy individuals.

TT is the time required for plasma to clot after a "standardized" prothrombin solution has been added to the plasma under test. TT has been previously found to be a marker to differentiate brain metastases from other intracranial malignancies.<sup>50</sup> In this study, the prolongation of TT was positively correlated with the development of lung cancer and was significantly different between patients with lung cancer and benign lung disease ( $P < 0.05$ ), with longer TT in patients with lung cancer than in patients with benign lung disease. However, in the data collected in this study, the median values of TT in both categories of patients remained within the normal reference range. More relevant data on the role of TT in lung carcinogenesis need to be collected for further analysis.

In this study, the diagnostic efficacy of the predictive models developed by each of the three comparison groups was higher than that of the predictive models developed using



tumor markers alone, and the difference in the lung cancer-health group and the benign lung disease-health group was statistically significant ( $P < 0.05$ ). It can be assumed that the more indicators used for risk prediction of lung cancer, the higher the detection rate of lung cancer will be. The AUC of the models of early-stage lung cancer-benign lung disease group and early-stage lung cancer-health group reached 0.841 and 0.976, respectively. It indicates that the models have high diagnostic efficacy and can help in the differential diagnosis of early-stage lung cancer patients, patients with benign lung disease and healthy individuals. The five prediction models of this study can be used simultaneously for risk prediction when patients come to the hospital, so as to help early clinical detection of lung cancer and intervention to reduce lung cancer mortality. Since the data collected from healthy individuals in this study were obtained from the test results of the various physical examinations they attended, electrolyte tests as well as glomerular filtration rate were not included. Therefore, the above indicators were not included in the data analysis and the development of diagnostic models between healthy individuals and the other two populations in this study, which does not mean that these indicators do not correlate and differ between healthy individuals and those with lung cancer or benign lung lesions.

The innovation of this study is to establish a risk prediction model for lung cancer by combining several conventional laboratory indicators and tumor markers through ANNs. Next, we will further validate and improve the efficiency of the diagnostic model by improving the relevant data and expanding the sample size, and test the effect of the models in clinical practice.

## Conclusion

The diagnosis model of lung cancer based on ANN which used conventional laboratory indicators combined with tumor markers has high diagnostic performance and far-reaching significance for early auxiliary diagnosis of lung cancer.

## AUTHORS' CONTRIBUTIONS

All authors participated in the design, interpretation of the studies and analysis of the data, and review of the article; YL wrote the article.

## DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Youth Fund, China (Grant No. 82003138); the Sichuan Science and Technology Program for key Research and Development, China (Grant No. 2021YFS0226); Doctoral Research Initiation Fund of Affiliated Hospital of Southwest Medical University, China (Grant No.19077); and Medical Science and Technology

Project of Sichuan Provincial Health Commission, China (Grant No.21PJ093).

## ORCID ID

Ting Ye  <https://orcid.org/0000-0002-5704-400X>

## SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

## REFERENCES

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;**71**:209–49
- Bade BC, Dela Cruz CS. Lung cancer 2020: epidemiology, etiology, and prevention. *Clin Chest Med* 2020;**41**:1–24
- Asakura K, Kadota T, Matsuzaki J, Yoshida Y, Yamamoto Y, Nakagawa K, Takizawa S, Aoki Y, Nakamura E, Miura J, Sakamoto H, Kato K, Watanabe SI, Ochiya T. A miRNA-based diagnostic model predicts resectable lung cancer in humans with high accuracy. *Commun Biol* 2020;**3**:134
- Chansky K, Detterbeck FC, Nicholson AG, Rusch VW, Vallières E, Groome P, Kennedy C, Krasnik M, Peake M, Shemanski L, Bolejack V, Crowley JJ, Asamura H, Rami-Porta R; The IASLC Staging and Prognostic Factors Committee, Advisory Boards and Participating Institutions. The IASLC lung cancer staging project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol* 2017;**12**:1109–21
- Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD; The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;**365**:395–409
- Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, Byers T, Colditz GA, Gould MK, Jett JR, Sabichi AL, Smith-Bindman R, Wood DE, Qaseem A, Detterbeck FC. Benefits and harms of CT screening for lung cancer: a systematic review. *JAMA* 2012;**307**:2418–29
- Song X, Liang B, Wang C, Shi S. Clinical value of color Doppler ultrasound combined with serum CA153, CEA and TSGF detection in the diagnosis of breast cancer. *Exp Ther Med* 2020;**20**:1822–8
- Kim GR, Hur J, Lee HJ, Nam JE, Kim YJ, Hong YJ, Shim HS, Kim HY, Lee JW, Choi BW. Analysis of tumor markers in cytological fluid obtained from computed tomography-guided needle aspiration biopsies for the diagnosis of ground-glass opacity pulmonary lesions. *Cancer Cytopathol* 2013;**121**:214–22
- Molina R, Filella X, Augé JM, Fuentes R, Bover I, Rifa J, Moreno V, Canals E, Viñolas N, Marquez A, Barreiro E, Borrás J, Viladriu P. Tumor markers (CEA, CA 125, CYFRA 21-1, SCC and NSE) in patients with non-small cell lung cancer as an aid in histological diagnosis and prognosis. *Tumour Biol* 2003;**24**:209–18
- Bréchet JM, Chevret S, Nataf J, Le Gall C, Fréault J, Rochemaure J, Chastang C. Diagnostic and prognostic value of Cyfra 21-1 compared with other tumour markers in patients with non-small cell lung cancer: a prospective study of 116 patients. *Eur J Cancer* 1997;**33**:385–91
- Ayer T, Chhatwal J, Alagoz O, Kahn CE Jr, Woods RW, Burnside ES. Informatics in radiology: comparison of logistic regression and artificial neural network models in breast cancer risk estimation. *Radiographics* 2010;**30**:13–22
- Hsu YC, Tsai YH, Weng HH, Hsu LS, Tsai YH, Lin YC, Hung MS, Fang YH, Chen CW. Artificial neural networks improve LDCT lung cancer screening: a comparative validation study. *BMC Cancer* 2020;**20**:1023
- Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest* 2017;**151**:193–203
- Howlander N, Forjaz G, Mooradian MJ, Meza R, Kong CY, Cronin KA, Mariotto AB, Lowy DR, Feuer EJ. The effect of advances in lung-cancer treatment on population mortality. *N Engl J Med* 2020;**383**:640–9

15. Saito A, Nagase T. Hippo and TGF- $\beta$  interplay in the lung field. *Am J Physiol Lung Cell Mol Physiol* 2015;**309**:L756–67
16. MacMahon H, Austin JH, Gamsu G, Herold CJ, Jett JR, Naidich DP, Patz EF, Jr, Swensen SJ; The Fleischner Society. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology* 2005;**237**:395–400
17. Thandra KC, Barsouk A, Saginala K, Aluru JS, Barsouk A. Epidemiology of lung cancer. *Contemp Oncol* 2021;**25**:45–52
18. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. *J Am Med Assoc* 1950;**143**:329–36
19. Holdenrieder S. Biomarkers along the continuum of care in lung cancer. *Scand J Clin Lab Invest Suppl* 2016;**245**:S40–5
20. Duffy MJ, O'Byrne K. Tissue and blood biomarkers in lung cancer: a review. *Adv Clin Chem* 2018;**86**:1–21
21. Molina R, Barak V, van Dalen A, Duffy MJ, Einarsson R, Gion M, Goike H, Lamerz R, Nap M, Sölétormos G, Stieber P. Tumor markers in breast cancer – European Group on Tumor Markers recommendations. *Tumour Biol* 2005;**26**:281–93
22. Juweid M, Sharkey RM, Behr T, Swayne LC, Rubin AD, Herskovic T, Hanley D, Markowitz A, Dunn R, Siegel J, Kamal T, Goldenberg DM. Improved detection of medullary thyroid cancer with radiolabeled antibodies to carcinoembryonic antigen. *J Clin Oncol* 1996;**14**:1209–17
23. Jelski W, Mroczo B. Biochemical markers of colorectal cancer – present and future. *Cancer Manag Res* 2020;**12**:4789–97
24. Grunnet M, Sorensen JB. Carcinoembryonic antigen (CEA) as tumor marker in lung cancer. *Lung Cancer* 2012;**76**:138–43
25. Nakamura H, Nishimura T. History, molecular features, and clinical importance of conventional serum biomarkers in lung cancer. *Surg Today* 2017;**47**:1037–59
26. Xu CM, Luo YL, Li S, Li ZX, Jiang L, Zhang GX, Owusu L, Chen HL. Multifunctional neuron-specific enolase: its role in lung diseases. *Biosci Rep* 2019;**39**:BSR20192732
27. Amampai R, Suprasert P. Cancer antigen 125 during pregnancy in women without ovarian tumor is not often rising. *Obstet Gynecol Int* 2018;**2018**:8141583
28. Miralles C, Orea M, España P, Provencio M, Sánchez A, Cantos B, Cubedo R, Carcereny E, Bonilla F, Gea T. Cancer antigen 125 associated with multiple benign and malignant pathologies. *Ann Surg Oncol* 2003;**10**:150–4
29. Lahoud RM, O'Shea A, El-Mouhayyar C, Atre ID, Eurboonyanun K, Harisinghani M. Tumour markers and their utility in imaging of abdominal and pelvic malignancies. *Clin Radiol* 2021;**76**:99–107
30. Duran AO, Inanc M, Karaca H, Dogan I, Berk V, Bozkurt O, Ozaslan E, Ucar M, Eroglu C, Ozkan M. Albumin-globulin ratio for prediction of long-term mortality in lung adenocarcinoma patients. *Asian Pac J Cancer Prev* 2014;**15**:6449–53
31. Wang Y, Li S, Hu X, Wang Y, Wu Y, Li P, Che G. The prognostic value of serum albumin-globulin ratio in early-stage non-small cell lung cancer: a retrospective study. *Cancer Manag Res* 2019;**11**:3545–54
32. Lee YJ, Han KD, Kim DH, Lee CH. Determining the association between repeatedly elevated serum gamma-glutamyltransferase levels and risk of respiratory cancer: a nationwide population-based cohort study. *Cancer Med* 2021;**10**:1366–76
33. Filaire E, Dupuis C, Galvaing G, Aubret S, Laurent H, Richard R, Filaire M. Lung cancer: what are the links with oxidative stress, physical activity and nutrition. *Lung Cancer* 2013;**82**:383–9
34. Gaur P, Bhattacharya S, Kant S, S Kushwaha RA, Garg R, Pandey S, Dubey A. Hospital-based study on demographic, hematological, and biochemical profile of lung cancer patients. *J Cancer Res Ther* 2020;**16**:839–42
35. Khoei NS, Anton G, Peters A, Freisling H, Wagner KH. The association between serum bilirubin levels and colorectal cancer risk: results from the prospective cooperative health research in the region of Augsburg (KORA) study in Germany. *Antioxidants* 2020;**9**:908
36. Monroy-Iglesias MJ, Moss C, Beckmann K, Hammar N, Walldius G, Bosco C, Van Hemelrijck M, Santaolalla A. Serum total bilirubin and risk of cancer: a Swedish cohort study and meta-analysis. *Cancers* 2021;**13**:5540
37. Song Q, Wang B, Wang S, Zhang YM, Chen WH. Lower serum levels of bilirubin in the newly diagnosed lung cancer patients: a case-control study in China. *J Cancer Res Ther* 2015;**11**:C168–72
38. Mi N, Huang J, Huang C, Lin Y, He Q, Wang H, Yang M, Lu Y, Lawer AL, Yue P, Bai B, Zhang J, Zhang C, Cai T, Fu W, Gao L, Li X, Yuan J, Meng W. High serum uric acid may associate with the increased risk of colorectal cancer in females: a prospective cohort study. *Int J Cancer* 2022;**150**:263–72
39. Yiu A, Van Hemelrijck M, Garmo H, Holmberg L, Malmström H, Lambe M, Hammar N, Walldius G, Jungner I, Wulaningsih W. Circulating uric acid levels and subsequent development of cancer in 493,281 individuals: findings from the AMORIS Study. *Oncotarget* 2017;**8**:42332–42
40. Bouras E, Karhunen V, Gill D, Huang J, Haycock PC, Gunter MJ, Johansson M, Brennan P, Key T, Lewis SJ, Martin RM, Murphy N, Platz EA, Travis R, Yarmolinsky J, Zuber V, Martin P, Katsoulis M, Freisling H, Nøst TH, Schulze MB, Dossus L, Hung RJ, Amos CI, Ahola-Olli A, Palaniswamy S, Männikkö M, Auvinen J, Herzig KH, Keinänen-Kiukaanniemi S, Lehtimäki T, Salomaa V, Raitakari O, Salmi M, Jalkanen S, Jarvelin MR, Dehghan A, Tsilidis KK; The PRACTICAL consortium. Circulating inflammatory cytokines and risk of five cancers: a Mendelian randomization analysis. *BMC Med* 2022;**20**:3
41. Wong JYY, Bassig BA, Loftfield E, Hu W, Freedman ND, Ji BT, Elliott P, Silverman DT, Chanock SJ, Rothman N, Lan Q. White blood cell count and risk of incident lung cancer in the UK biobank. *JNCI Cancer Spectr* 2019;**4**:pkz102
42. Hofman P. New insights into the interaction of the immune system with non-small cell lung carcinomas. *Transl Lung Cancer Res* 2020;**9**:2199–213
43. Kwiecień I, Rutkowska E, Raniszewska A, Rzepecki P, Domagała-Kulawik J. Modulation of the immune response by heterogeneous monocytes and dendritic cells in lung cancer. *World J Clin Oncol* 2021;**12**:966–82
44. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *J Leukoc Biol* 2019;**106**:309–22
45. Yin W, Lv J, Yao Y, Zhao Y, He Z, Wang Q, Cui L, Dai H. Elevations of monocyte and neutrophils, and higher levels of granulocyte colony-stimulating factor in peripheral blood in lung cancer patients. *Thorac Cancer* 2021;**12**:2680–90
46. Zhao T, Cui L, Li A. The significance of RDW in patients with hepatocellular carcinoma after radical resection. *Cancer Biomark* 2016;**16**:507–12
47. Chen JL, Wu JN, Lv XD, Yang QC, Chen JR, Zhang DM. The value of red blood cell distribution width, neutrophil-to-lymphocyte ratio, and hemoglobin-to-red blood cell distribution width ratio in the progression of non-small cell lung cancer. *PLoS ONE* 2020;**15**:e0237947
48. Ibrahim UA, Yusuf AA, Ahmed SG. The pathophysiologic basis of anaemia in patients with malignant diseases. *Gulf J Oncolog* 2016;**1**:80–9
49. Zhang Z, Zhang F, Yuan F, Li Y, Ma J, Ou Q, Liu Z, Yang B, Wang L, Tao H, Zhang S, Li X, Zhi X, Ge X, Bao H, Wu X, Hu Y, Wang J. Pretreatment hemoglobin level as a predictor to evaluate the efficacy of immune checkpoint inhibitors in patients with advanced non-small cell lung cancer. *Ther Adv Med Oncol* 2020;**12**:0970049
50. Wang SQ, Yuan Q, Zhang GT, Qian HP, Liu ZD, Wang JW, Cai HQ, Wan JH. Preoperative blood testing for glioblastoma, brain metastases, and primary central nervous system lymphoma differentiation. *Transl Cancer Res* 2022;**11**:63–71

(Received February 14, 2022, Accepted March 16, 2023)