Risk assessment of malignancy in solitary pulmonary nodules in lung computed tomography: a multivariable predictive model study

Hai-Yang Liu[1], Xing-Ru Zhao[2], Meng Chi[3], Xiang-Song Cheng[4], Zi-Qi Wang[1], Zhi-Wei Xu[1,2], Yong-Li Li[5], Rui Yang[3], Yong-Jun Wu[6], Xiao-Ju Zhang[1,2]

[1]Department of Respiratory and Critical Care Medicine, Henan Provincial People's Hospital; Zhengzhou University People's Hospital, Zhengzhou, Henan 450000, China;
[2]Henan Joint International Research Laboratory of Diagnosis and Treatment of Pulmonary Nodules, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, Henan 450000, China;
[3]Department of Medical Imaging, Henan Provincial Chest Hospital, Zhengzhou, Henan 450000, China;
[4]Department of Respiratory and Critical Care Medicine, Fuwai Central China Cardiovascular Hospital, Zhengzhou, Henan 450000, China;
[5]Department of Radiology, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, Henan 450000, China;
[6]College of Public Health, Zhengzhou University, Zhengzhou, Henan 450000, China.

Abstract

Background: Computed tomography images are easy to misjudge because of their complexity, especially images of solitary pulmonary nodules, of which diagnosis as benign or malignant is extremely important in lung cancer treatment. Therefore, there is an urgent need for a more effective strategy in lung cancer diagnosis. In our study, we aimed to externally validate and revise the Mayo model, and a new model was established.

Methods: A total of 1450 patients from three centers with solitary pulmonary nodules who underwent surgery were included in the study and were divided into training, internal validation, and external validation sets ($n = 849$, 365, and 236, respectively). External verification and recalibration of the Mayo model and establishment of new logistic regression model were performed on the training set. Overall performance of each model was evaluated using area under receiver operating characteristic curve (AUC). Finally, the model validation was completed on the validation data set.

Results: The AUC of the Mayo model on the training set was 0.653 (95% confidence interval [CI]: 0.613–0.694). After re-estimation of the coefficients of all covariates included in the original Mayo model, the revised Mayo model achieved an AUC of 0.671 (95% CI: 0.635–0.706). We then developed a new model that achieved a higher AUC of 0.891 (95% CI: 0.865–0.917). It had an AUC of 0.888 (95% CI: 0.842–0.934) on the internal validation set, which was significantly higher than that of the revised Mayo model (AUC: 0.577, 95% CI: 0.509–0.646) and the Mayo model (AUC: 0.609, 95% CI, 0.544–0.675) ($P < 0.001$). The AUC of the new model was 0.876 (95% CI: 0.831–0.920) on the external verification set, which was higher than the corresponding value of the Mayo model (AUC: 0.705, 95% CI: 0.639–0.772) and revised Mayo model (AUC: 0.706, 95% CI: 0.640–0.772) ($P < 0.001$). Then the prediction model was presented as a nomogram, which is easier to generalize.

Conclusions: After external verification and recalibration of the Mayo model, the results show that they are not suitable for the prediction of malignant pulmonary nodules in the Chinese population. Therefore, a new model was established by a backward stepwise process. The new model was constructed to rapidly discriminate benign from malignant pulmonary nodules, which could achieve accurate diagnosis of potential patients with lung cancer.

Keywords: CT image; Lung cancer; Prediction model; Pulmonary nodules; Regression algorithm

Introduction

In 2018, there were 9.6 million cancer deaths worldwide, of which lung cancer was the leading cause (18.4% of total cancer deaths).[1] The use of low-dose computed tomography (LDCT) to screen high-risk groups for lung cancer can reduce the death rate from lung cancer by 20%.[2] However, LDCT for lung cancer screening can also have some poor results, including false positive and false negative test results, overdiagnosis, and radiation exposure.[3] Of all positive LDCT results, 95% do not result in a diagnosis of lung cancer, which affects the physical and mental health of many subjects screened for lung cancer by LDCT. To increase the effectiveness of LDCT lung cancer screening, a prediction model can be used to evaluate the

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000001507

Hai-Yang Liu and Xing-Ru Zhao contributed equally to the work.

Correspondence to: Dr. Xiao-Ju Zhang, Department of Respiratory and Critical Care Medicine, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, Henan 450000, China
E-Mail: zhangxiaoju1010@henu.edu.cn

Chinese Medical Journal 2021;134(14)
Received: 14-12-2020 **Edited by:** Pei-Fang Wei

probability of malignancy among pulmonary nodules found during lung cancer screening before CT follow-up, before surgery, or before biopsy. The goal is to achieve accurate diagnosis of patients with lung cancer, which could reduce mortality from lung cancer and limit the harm and cost of unnecessary diagnosis and treatment.

Many predictive models have been widely used in clinical practice, but whether they are applicable to local medical institutions requires verification. Complete reporting of the development and validation of clinical predictive models is critical to the modeling of external validation and clinical applications. The "Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis" (TRIPOD) guidelines formalize the reporting process and report quality evaluation of the standardized prediction model.[4] The first lung cancer risk prediction model was Swensen Mayo model, which has been recommended by organizations such as the American College of Chest Physicians and Chinese Thoracic Society.[5,6] Although several studies have attempted to externally validate prediction models for pulmonary nodules in people from China and other countries in the Asia-Pacific region, very few of them have completely adhered to the TRIPOD recommendations.[7-10] No such study from China has been conducted in multiple centers with a sample size of more than 1000.

The prediction of benign and malignant tumors is a typical binary classification problem, and the boom in statistical methods provides us with effective tools to approach such a problem: such as logistic regression, least absolute shrinkage and selection operator (Lasso) regression, and best subset selection. At present, few studies following the TRIPOD guideline have integrated the calibration of existing models and regression algorithms in risk predictive model research. In this large multicenter study in a Chinese cohort, the Mayo model was externally verified and calibrated, and a new risk predictive model was established using regression algorithms following the TRIPOD guidelines. Internal and external validations were also completed.

Methods

Ethical approval

This study was approved by the Ethics Committee and Institutional Review Board of Henan Provincial People's Hospital (No. 201986). Because of the retrospective nature of the study and desensitization data, the requirement for obtaining informed consent was waived.

Study population

We retrospectively reviewed medical records from Henan Provincial People's Hospital, Fuwai Central China Cardiovascular Hospital, and Henan Provincial Chest Hospital, three large tertiary hospitals located in China's central region. Inclusion and exclusion criteria were set according to the Swensen's study,[6] and patients were excluded if they had multiple pulmonary nodules or if the case was pathologically proven by non-surgical tissue biopsy.

From 2014 to 2020, we followed up 11,216 cases of pulmonary nodules, most of which were sub-solid at the initial CT examination. Finally, a total of 1450 subjects with pulmonary nodules on CT were enrolled for model establishment and validation [Table 1]. All nodules underwent video-assisted thoracoscopic surgery from January 2015 to May 2020. Participants from Henan Provincial People's Hospital were randomly divided into training and internal validation sets in a ratio of 0.7:0.3 using the sample function in R version 3.6.0 (www.R-project.org). The cohorts from the Fuwai Central China Cardiovascular Hospital and Henan Provincial Chest Hospital were used as an external validation dataset [Figure 1].

Clinical information and CT imaging

Information about the participants and nodules was collected from the hospital information system and picture archiving and communication systems. Smoking history was defined as a smoking index of ≥400. Determination of the tumor-vessel relationship on CT images was based on Gaeta determination method.[11] All image features were annotated by two clinicians with more than 5 years of working experience, and uncertain features were determined by a senior imaging specialist.

Mayo model and model revision

The Mayo model was as follows:

$$\frac{e^{\text{Mayo}-PI}}{(1+e^{\text{Mayo}-PI})}$$

With: Mayo-prognostic index (PI) = −6.8272 + (0.0391 × age) + (0.7917 × smoking) + (1.3388 × history of cancer) + (0.1274 × nodule diameter) + (1.0407 × spiculation) + (0.7838 × upper lobe nodule)

We denote PI as the weighted sum of the covariates $X_{1\ldots}$ X_n in the model, where the weights $\beta_{1\ldots}\beta_n$ are the regression coefficients and α is the intercept (e is the base of the natural logarithm, $e = 2.718281828$).

$$\text{PI} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots\ \ldots + \beta_n X_n$$

We generated a revised Mayo model without considering the original one. All coefficients ($\beta_{1\ldots}\beta_n$ and the intercept α) were re-estimated by fitting the original Mayo model to the training set. Because this method was used, the original PI Mayo was not preserved.

Statistical analysis

All statistical analyses were performed with R version 3.6.0 (www.R-project.org) and IBM SPSS Statistics 21 (IBM Corp, Armonk, NY, USA). In this study, continuous variables with normal distributions were expressed as mean ± standard deviation ($\overline{x} \pm s$) and those with non-normal distribution were presented as median (Q_1, Q_3). Categorical variables were compared using the Chi-square test, continuity-adjusted Chi-square test or Fisher exact

Table 1: Differences in characteristics of nodules and participants between the training, internal validation, and external validation sets.

Variables	Training set ($n=849$)				Internal validation set ($n=365$)				External validation set ($n=236$)			
	Benign ($n=218$)	Malignant ($n=631$)	Statistics	P	Benign ($n=74$)	Malignant ($n=291$)	Statistics	P	Benign ($n=123$)	Malignant ($n=113$)	Statistics	P
Nodule type			71.405*	<0.001			32.673*	<0.001			49.461*	<0.001
Non-solid	18 (8.3)	246 (39.0)			5 (6.8)	123 (42.3)			24 (19.5)	73 (64.6)		
Solid	200 (91.7)	385 (61.0)			69 (93.2)	168 (57.7)			99 (80.5)	40 (35.4)		
Gender			1.148*	0.284			4.016*	0.045			4.188*	0.041
Female	109 (50.0)	342 (54.2)			29 (39.2)	152 (52.2)			61 (49.6)	71 (62.8)		
Male	109 (50.0)	289 (45.8)			45 (60.8)	139 (47.8)			62 (50.4)	42 (37.2)		
Age (years)	53.74 ± 10.08	57.33 ± 10.09	−4.530†	<0.001	53.43 ± 10.69	57.01 ± 10.66	−2.598†	0.010	56.95 ± 9.90	57.86 ± 10.16	−1.106†	0.270
Smoking			0.176*	0.674			6.903*	0.009			4.926*	0.026
No	178 (81.7)	507 (80.3)			48 (64.9)	231 (79.4)			109 (88.6)	88 (77.9)		
Yes	40 (18.3)	124 (19.7)			26 (35.1)	60 (20.6)			14 (11.4)	25 (22.1)		
History of cancer			1.341*	0.247			-	0.214			-	0.068
No	205 (94.0)	578 (91.6)			73 (98.6)	275 (94.5)			116 (94.3)	112 (99.1)		
Yes	13 (6.0)	53 (8.4)			1 (1.4)	16 (5.5)			7 (5.7)	1 (0.9)		
Family history of lung cancer			8.184*	0.004			0.003*	0.956			4.650*	0.031
No	196 (89.9)	515 (81.6)			63 (85.1)	247 (84.9)			120 (97.6)	103 (91.2)		
Yes	22 (10.1)	116 (18.4)			11 (14.9)	44 (15.1)			3 (2.4)	10 (8.8)		
Emphysema			0.093*	0.760			0.195*	0.659			1.817*	0.178
No	184 (84.4)	527 (83.5)			61 (82.4)	246 (84.5)			115 (93.5)	100 (88.5)		
Yes	34 (15.6)	104 (16.5)			13 (17.6)	45 (15.5)			8 (6.5)	13 (11.5)		
COPD			3.075*	0.080			-	0.588			-	0.008
No	210 (96.3)	587 (93.0)			71 (95.9)	273 (93.8)			122 (99.2)	104 (92.0)		
Yes	8 (3.7)	44 (7.0)			3 (4.1)	18 (6.2)			1 (0.8)	9 (8.0)		
Tuberculosis			-	0.763			-	1.000			4.936*	0.026
No	214 (98.2)	621 (98.4)			73 (98.6)	286 (98.3)			113 (91.9)	111 (98.2)		
Yes	4 (1.8)	10 (1.6)			1 (1.4)	5 (1.7)			10 (8.1)	2 (1.8)		
ILD			-	1.000			-	1.000			-	1.000
No	217 (99.5)	629 (99.7)			74 (100.0)	290 (99.7)			122 (99.2)	112 (99.1)		
Yes	1 (0.5)	2 (0.3)			0 (0.0)	1 (0.3)			1 (0.8)	1 (0.9)		
Nodule diameter (mm)	14.55 (10.80, 19.20)	17.00 (12.20, 21.90)	−3.791‡	<0.001	15.85 (11.03, 21.65)	17.30 (11.50, 22.10)	−0.693‡	0.488	10.30 (7.10, 15.40)	14.50 (10.95, 21.30)	−5.247‡	<0.001
Upper lobe nodule			2.614*	0.106			7.240*	0.007			2.272*	0.132
No	100 (45.9)	250 (39.6)			38 (51.4)	100 (34.4)			61 (49.6)	45 (39.8)		
Yes	118 (54.1)	381 (60.4)			36 (48.6)	191 (65.6)			62 (50.4)	68 (60.2)		
Spiculation			35.580*	<0.001			8.724*	0.003			15.923*	<0.001
No	117 (53.7)	196 (31.1)			42 (56.8)	110 (37.8)			89 (72.4)	53 (46.9)		
Yes	101 (46.3)	435 (68.9)			32 (43.2)	181 (62.2)			34 (27.6)	60 (53.1)		
Lobulation			76.898*	<0.001			14.211*	<0.001			28.538*	<0.001
No	128 (58.7)	164 (26.0)			37 (50.0)	79 (27.1)			103 (83.7)	58 (51.3)		
Yes	90 (41.3)	467 (74.0)			37 (50.0)	212 (72.9)			20 (16.3)	55 (48.7)		
Vacuole sign			76.513*	<0.001			18.939*	<0.001			61.793*	<0.001
No	182 (83.5)	313 (49.6)			64 (86.5)	173 (59.5)			112 (91.1)	49 (43.4)		
Yes	36 (16.5)	318 (50.4)			10 (13.5)	118 (40.5)			11 (8.9)	64 (56.6)		
Calcification			23.748*	<0.001			-	<0.001			-	0.015
No	195 (89.4)	615 (97.5)			59 (79.7)	284 (97.6)			116 (94.3)	113 (100.0)		
Yes	23 (10.6)	16 (2.5)			15 (20.3)	7 (2.4)			7 (5.7)	0 (0)		
Pleural pull			20.947*	<0.001			10.666*	<0.001			12.002*	<0.001
No	159 (72.9)	349 (55.3)			58 (78.4)	168 (57.7)			101 (82.1)	70 (61.9)		
Yes	59 (27.1)	282 (44.7)			16 (21.6)	123 (42.3)			22 (17.9)	43 (38.1)		
With vessel			103.858*	<0.001			21.964*	<0.001			29.151*	<0.001
No	150 (68.8)	187 (29.6)			41 (55.4)	78 (26.8)			78 (63.4)	32 (28.3)		
Yes	68 (31.2)	444 (70.4)			33 (44.6)	213 (73.2)			45 (36.6)	81 (71.7)		
Satellite nodules			22.159*	<0.001			-	0.002			4.936*	0.026
No	201 (92.2)	622 (98.6)			67 (90.5)	287 (98.6)			113 (91.9)	111 (98.2)		
Yes	17 (7.8)	9 (1.4)			7 (9.5)	4 (1.4)			10 (8.1)	2 (1.8)		

Quantitative data were described as mean ± standard deviation or median (Q_1, Q_3) while qualitative variables were expressed as n (%). * χ^2 values, † t values, ‡ Z values. COPD: Chronic obstructive pulmonary disease; ILD: Interstitial lung disease; -: Not applicable.

test. Continuous variables were compared using *t*-tests for variables with normal distributions or Mann-Whitney *U* tests for variables with abnormal distributions.

Backward stepwise selection with the Akaike information criterion was used to identify variables for the multivariable logistic model. The area under the receiver operating characteristic (ROC) curve (AUC) was calculated to evaluate the model's performance. During the external validation of the model, the total score for each patient in the validation cohort was calculated according to the established model, and then logistic regression in this cohort was performed using the total score as a factor, and finally, the AUC and calibration curve were derived from the regression analysis. Furthermore, the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were used to evaluate the classification models. The model developed in this study was compared with the Mayo model using the DeLong test. $P < 0.05$ was considered statistically significant.

Results

External validation and recalibration of the Mayo model

The Mayo model fit poorly with the training data set (AUC: 0.653; 95% CI: 0.613–0.694), showing that the model poorly differentiated between benign and malignant cases. When the model was recalibrated, the coefficients for age, smoking, history of cancer, nodule diameter, spiculation, and upper lobe nodule were updated. The revised model had improved performance and achieved a better AUC value of 0.671 (95% CI: 0.635–0.706) on the training data set [Figure 2]. According to this revised Mayo model, the probability of lung cancer is $\frac{e^{PI}}{(1+e^{PI})}$, with: PI = −1.4945 + (0.0264 × age) − (0.1415 × smoking) + (0.4623 × history of cancer) + (0.0273 × nodule diameter) + (0.7864 × spiculation) + (0.3231 × upper lobe nodule).

The AUC of Mayo model and revised Mayo model was 0.609 (95% CI: 0.544–0.675) and 0.577 (95% CI: 0.509–0.646) on the internal validation data set, respectively. The

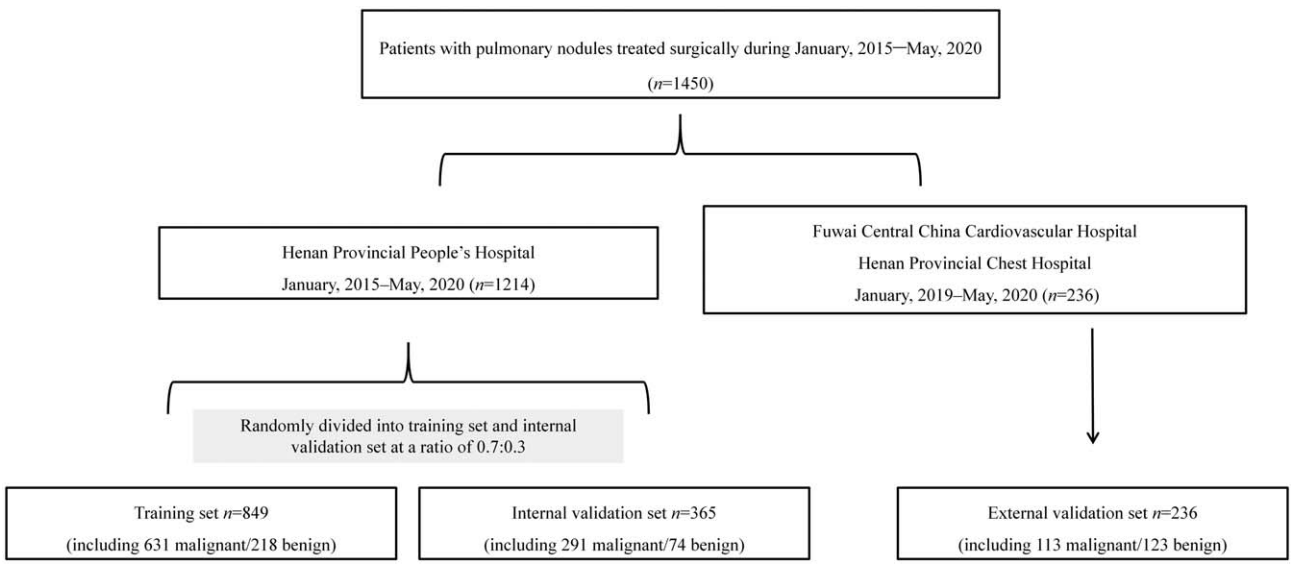

Figure 1: Distribution of datasets shown as a flowchart.

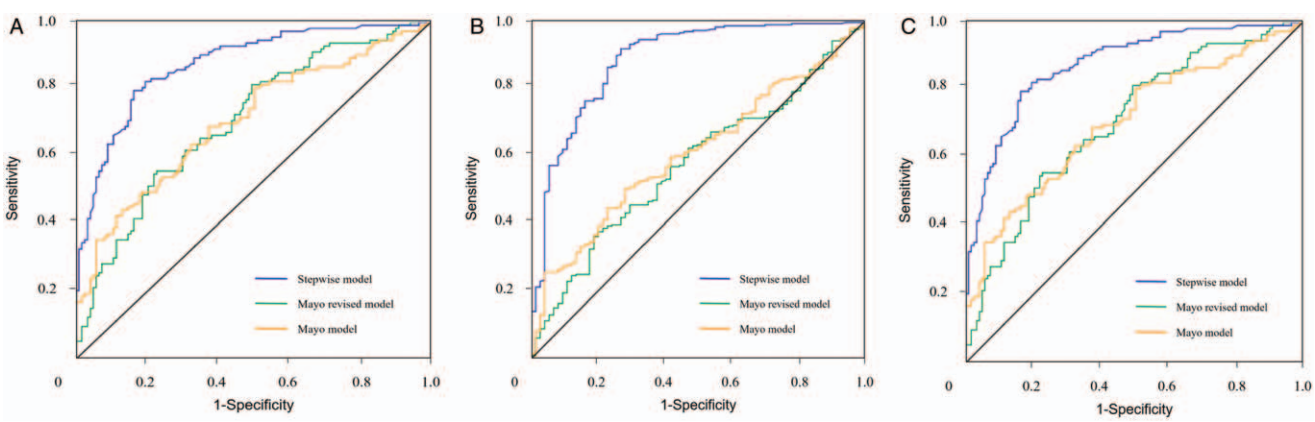

Figure 2: Comparison of ROC curves between Mayo model, revised Mayo model and new model in the data sets. (A) Training set ($n = 849$); (B) internal validation set ($n = 365$); (C) external validation set ($n = 236$). ROC: Receiver operating characteristic.

Table 2: Performance of the three models in risk assessment of malignancy in solitary pulmonary nodules in lung CT.

Models	Types	Training set ($n = 849$)		Internal validation set ($n = 365$)		External validation set ($n = 236$)	
		Malignant	Benign	Malignant	Benign	Malignant	Benign
Mayo model	Malignant	321	54	147	21	78	46
	Benign	310	164	144	53	35	77
Revised Mayo model	Malignant	361	73	110	15	63	28
	Benign	270	145	181	59	50	95
New model	Malignant	537	43	268	20	90	20
	Benign	94	175	23	54	23	103

Data are presented as *n*. CT: Computed tomography.

Table 3: Effect evaluation of the the Mayo model, revised Mayo model and new model on the training set, internal validation set, and external validation set.

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC
Mayo model						
Training set	57.13	50.87	75.23	85.60	34.60	0.653
Internal validation set	54.79	50.52	71.62	87.50	26.90	0.609
External validation set	65.68	69.03	62.60	62.90	68.75	0.705
Revised Mayo model						
Training set	59.60	57.21	66.51	83.18	34.94	0.671
Internal validation set	46.30	37.80	79.73	88.00	24.58	0.577
External validation set	66.95	55.75	77.24	69.23	65.52	0.706
New model						
Training set	83.86	85.10	80.28	92.59	65.06	0.891
Internal validation set	88.22	92.10	72.97	93.06	70.13	0.888
External validation set	81.78	79.65	83.74	81.82	81.75	0.876

AUC: Area under receiver operating characteristic curve; NPV: Negative predictive value; PPV: Positive predictive value.

AUC of Mayo model and revised Mayo model was 0.705 (95% CI: 0.639–0.772) and 0.706 (95% CI: 0.640–0.772) on the external validation data set, respectively [Figure 2]. There was a significant difference between the Mayo model and the revised Mayo model ($P < 0.001$). However, both the original Mayo model and the revised Mayo model poorly performed on our training data set and validation data set [Tables 2 and 3]. Thus, calibration of the Mayo model on the training data set did not achieve the desired results.

Establishment of the new model

The external verification and calibration of the Mayo model did not achieve good results. To achieve better prediction of benign and malignant pulmonary nodules, a new prediction model was built by the backward stepwise method on the training set, which achieved a higher AUC of 0.891 (95% CI: 0.865–0.917) than the other model. The new model contains new variables such as nodule type, history of chronic obstructive pulmonary disease (COPD), nodule margin (lobulation), nodule within the lesion (vacuole sign, calcification), and the presence of vessels and satellite nodules. For convenience, Figure 3 shows a nomogram of the new model, the regression equation of which is shown below:

$$\text{New model PI} = -1.39894 - 3.02619 \times X_1 + 0.04067 \times X_2 + 1.20883 \times X_3 + 1.52663 \times X_4 + 1.27667 \times X_5 + 0.80562 \times X_6 - 1.90613 \times X_7 + 1.52263 \times X_8 - 2.48893 \times X_9$$

Notes: X_1 represents solid nodule (1 if solid nodule, 0 if non-solid nodule);

X_2 represents age (years);

X_3 represents COPD (1 if patient has COPD, 0 if not);

X_4 represents spiculation (1 if nodule has spiculation, 0 if not);

X_5 represents lobulation (1 if nodule has lobulation, 0 if not);

X_6 represents vacuole sign (1 if nodule has vacuole, 0 if not);

X_7 represents calcification (1 if nodule has calcification, 0 if not);

X_8 represents vessels (1 if nodule has vessels, 0 if not);

X_9 represents satellite nodules (1 if nodule has satellite nodules, 0 if not).

Verification of the new model's performance

The accuracy of the new model was 83.86% on the training cohort and 88.22% and 81.78% on the internal and external validation sets, respectively, which suggests that the model has good discrimination [Table 3].

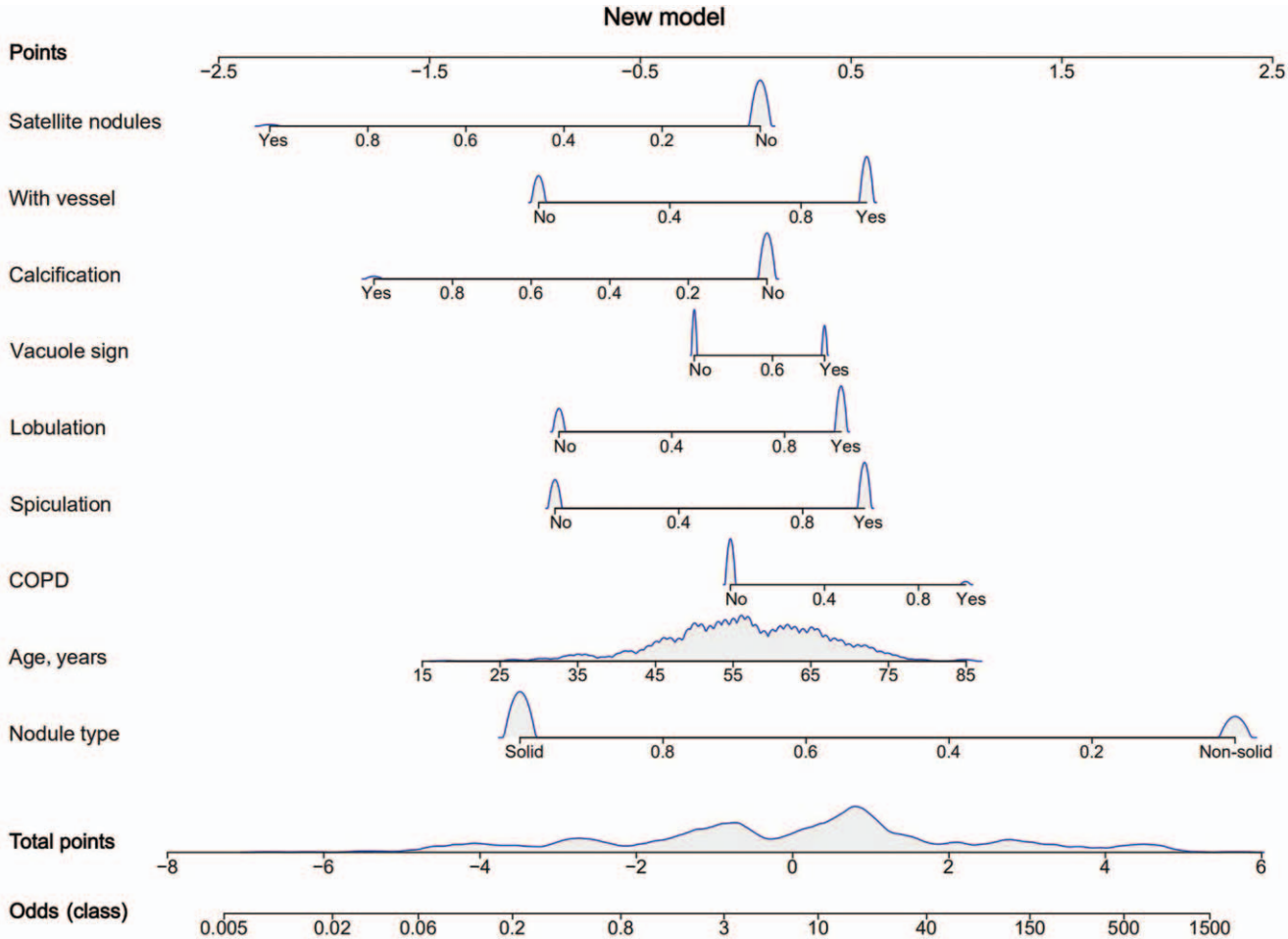

Figure 3: Development of nomogram for prediction of malignant risk. The new model is presented in the form of nomogram. After the position of each variable on the corresponding axis is determined, a line to the axis can be drawn to indicate the number of points.

The new model showed an AUC of 0.888 (95% CI: 0.842–0.934) on the internal validation set, which was higher than the corresponding value of the Mayo model (AUC: 0.609, 95% CI: 0.544–0.675) and the revised Mayo model (AUC: 0.577, 95% CI: 0.509–0.646) (both $P < 0.001$) [Figure 2]. After the new model was fit to the internal validation set, 20 and 23 benign and malignant nodules were incorrectly classified into the malignant and benign groups, respectively (sensitivity: 92.10%; specificity: 72.97%; PPV: 93.06%; NPV: 70.13%; AUC: 0.888). On the external validation set, the AUC of the new model was 0.876 (95% CI: 0.831–0.920), which was higher than the Mayo model's AUC of 0.705 (95% CI: 0.639–0.772) and the revised Mayo model's AUC of 0.706 (95% CI: 0.640–0.772) (both $P < 0.001$). The new model's accuracy on the external validation set was 81.78%, which also constitutes good performance. The model incorrectly classified 20 and 23 benign and malignant nodules as malignant and benign nodules, respectively, in the external validation set (sensitivity: 79.65%; specificity: 83.74%; PPV: 81.82%; NPV: 81.75%; AUC: 0.876) [Table 3].

Discussion

In this study, we mainly elaborated three results. First, we performed external verification and calibration of the Mayo model. Fitting our cohort to the Mayo model had similar results to those of existing studies, and the results show that the original model has weak predictive power on our cohort.[8,9,12,13] This can be attributed to the differences in living environment, living habits, and disease spectrum of the Chinese population from those of the European and American populations on which the Mayo model was developed. These differences may have caused the inadaptability of the Mayo model to the Chinese population. Following the TRIPOD guidelines, we used the training dataset to calibrate the Mayo model. The original and calibrated Mayo models performed poorly on the validation set. The final results showed that neither the Mayo model nor the calibrated Mayo model was suitable for the Chinese population cohort.

Second, we used regression algorithms to build a new model. We screened independent risk factors for lung cancer, thereby establishing a regression model, using the backward stepwise method. The results of the new model were significantly better than those of the Mayo model on all three of the datasets used in our study. Our constructed model had good predictive performance and can avoid the missed diagnosis of malignant tumors to a large extent, but the predictive performance of benign lung nodules in the training set and internal validation set was slightly

reduced. This may be because the number of malignant nodules was far greater than the number of benign nodules in the training and internal validation sets. The overestimation of malignancy risk may also be related to the fact that all pulmonary nodules admitted for treatment were considered high-risk by the clinicians.

Third, we determined nine independent risk factors for malignant pulmonary nodules through backward stepwise regression, which were similar to the risk factors (ie, age, spiculation, and calcification) included in the studies by Swensen, McWilliams, and Wu *et al*.[2,6,8] Solidity of nodules was used as a protective factor in our model, which indicates that the malignancy probability of subsolid nodules is significantly higher than that of solid nodules. A study by Henschke *et al*[14] showed that subsolid nodules, pure ground glass nodules, and solid nodules accounted for 63%, 18%, and 7% of malignant cases, respectively. In Gaeta's research,[11] nodules in which "the vessel passes through the nodule, the vessel shifts to the nodule, the vessel is cut off in the surrounding nodule, and the vessel is compressed and shifted" were all regarded as nodules with blood vessels. This is consistent with the well-known occurrence of angiogenesis and neovascularization in malignancy.[15,16] In particular, history of COPD was also an independent risk factor for lung cancer in this study. Park *et al*[17] found that the risk of lung cancer was lower in people who did not smoke and also did not have COPD than in people who smoked and did not have COPD (the risk of the latter group increased by 97%). People who did not smoke but had COPD had a 167% increase in lung cancer risk, and people who both smoked and had COPD had a 519% increase in lung cancer risk. The characteristic signs of malignant pulmonary nodules, such as the lobular and vacuole signs, were also included in the model. Calcification was included in the model as a protective factor, perhaps because of the high proportion of benign patients with tuberculosis granulomas in the cohort used to establish the model, and the presence of satellite lesions around the lesion is a characteristic manifestation of tuberculosis. In addition, in our study, 90.3% (442/489) of ground glass nodules that underwent surgical resection were eventually diagnosed as malignancy. The presence or increase in the proportion of solid components of sub-solid nodules is an important predictor in the follow-up process, and we will focus more on this aspect in the future.

Our research has the following advantages. First, all nodules included in the study were surgically removed and had clear pathological diagnoses. Second, we have verified and calibrated the existing models with external populations; established a new model; and conducted external multicenter cohort verification, which have rarely been achieved in other comparable studies, especially in China. Third, the model established by the regression algorithm is a parametric model, which can reduce the missed diagnosis of malignant nodules, and the relevant data parameters are easy to obtain. We presented the model in the form of a nomogram, which is convenient for clinical application.

This research also had several limitations. First, the external cohort of this study was small, and only two external centers were included. There may be underrepresentation of the sample in the study. In the future, a more diverse external cohort will be needed to calibrate and update our model. Second, most of the high-risk lung nodules included in the study required hospitalization for diagnosis, and the number of benign nodules in our cohort was small. The model may thus overestimate the risk that lung nodules are malignant and may be suitable for preoperative or pre-biopsy evaluation, and it may thus reduce the missed diagnosis of malignant nodules.

In conclusion, the Mayo model fit the Chinese cohort population poorly, and improvement was not obvious after calibration. When the backward stepwise regression method was employed to establish a new model to differentiate the benign from malignant lung nodules, the model had good prediction performance and was especially superior to the original and revised Mayo models. Lung cancer diagnosis systems were constructed to rapidly discriminate and diagnose benign and malignant pulmonary CT nodules.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 81670091) and the Zhongyuan Science and Technology Innovation Leading Talent Project (No. 194200510).

Conflicts of interest

None.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424. doi: 10.3322/caac.21492.
2. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, *et al*. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409. doi: 10.1056/NEJMoa1102873.
3. Patz EF Jr, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, *et al*. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014;174:269–274. doi: 10.1001/jamainternmed.2013.12738.
4. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–W73. doi: 10.7326/M14-0698.
5. Bai C, Choi CM, Chu CM, Anantham D, Chung-Man Ho J, Khan AZ, *et al*. Evaluation of pulmonary nodules: clinical practice consensus guidelines for Asia. Chest 2016;150:877–893. doi: 10.1016/j.chest.2016.02.650.
6. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997;157:849–855. doi: 10.1001/archinte.1997.00440290031002.
7. Heus P, Damen J, Pajouheshnia R, Scholten R, Reitsma JB, Collins GS, *et al*. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. BMJ Open 2019;9:e025611. doi: 10.1136/bmjopen-2018-025611.
8. Wu Z, Huang T, Zhang S, Cheng D, Li W, Chen B. A prediction model to evaluate the pretest risk of malignancy in solitary pulmonary nodules: evidence from a large Chinese southwestern population. J Cancer Res Clin Oncol 2021;147:275–285. doi: 10.1007/s00432-020-03408-2.

9. Yang B, Jhun BW, Shin SH, Jeong BH, Um SW, Zo JI, *et al*. Comparison of four models predicting the malignancy of pulmonary nodules: a single-center study of Korean adults. PLoS One 2018;13: e0201242. doi: 10.1371/journal.pone.0201242.
10. Yang D, Zhang X, Powell CA, Ni J, Wang B, Zhang J, *et al*. Probability of cancer in high-risk patients predicted by the protein-based lung cancer biomarker panel in China: LCBP study. Cancer 2018;124:262–270. doi: 10.1002/cncr.31020.
11. Gaeta M, Barone M, Russi EG, Volta S, Casablanca G, Romeo P, *et al*. Carcinomatous solitary pulmonary nodules: evaluation of the tumor-bronchi relationship with thin-section CT. Radiology 1993;187:535–539. doi: 10.1148/radiology.187.2.8475303.
12. Cui X, Heuvelmans MA, Han D, Zhao Y, Fan S, Zheng S, *et al*. Comparison of Veterans Affairs, Mayo, Brock classification models and radiologist diagnosis for classifying the malignancy of pulmonary nodules in Chinese clinical population. Transl Lung Cancer Res 2019;8:605–613. doi: 10.21037/tlcr.2019.09.17.
13. She Y, Zhao L, Dai C, Ren Y, Jiang G, Xie H, *et al*. Development and validation of a nomogram to estimate the pretest probability of cancer in Chinese patients with solid solitary pulmonary nodules: a multi-institutional study. J Surg Oncol 2017;116:756–762. doi: 10.1002/jso.24704.
14. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS, *et al*. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002;178:1053–1057. doi: 10.2214/ajr.178.5.1781053.
15. Nishida N, Yano H, Nishida T, Kamura T, Kojiro M. Angiogenesis in cancer. Vasc Health Risk Manag 2006;2:213–219. doi: 10.2147/vhrm.2006.2.3.213.
16. Raghu VK, Zhao W, Pu J, Leader JK, Wang R, Herman J, *et al*. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. Thorax 2019;74:643–649. doi: 10.1136/thoraxjnl-2018-212638.
17. Park HY, Kang D, Shin SH, Yoo KH, Rhee CK, Suh GY, *et al*. Chronic obstructive pulmonary disease and lung cancer incidence in never smokers: a cohort study. Thorax 2020;75:506–509. doi: 10.1136/thoraxjnl-2019-213732.

How to cite this article: Liu HY, Zhao XR, Chi M, Cheng XS, Wang ZQ, Xu ZW, Li YL, Yang R, Wu YJ, Zhang XJ. Risk assessment of malignancy in solitary pulmonary nodules in lung computed tomography: a multivariable predictive model study. Chin Med J 2021;134:1687–1694. doi: 10.1097/CM9.0000000000001507