### Check for updates

### OPEN ACCESS

EDITED BY Dan Liu, Wuhan University, China

### REVIEWED BY

Bing Yang, Tianjin Medical University, China Mohd Faheem Khan, M. J. P. Rohilkhand University, India

\*CORRESPONDENCE Xiaoqin Liu Iuxiaoqin@scszlyy.org.cn Qiang Li Iiqiang1907@vip.sina.com

<sup>†</sup>These authors have contributed equally to this work

RECEIVED 24 January 2025 ACCEPTED 26 February 2025 PUBLISHED 14 March 2025

### CITATION

Xiang R, Hu P, Xiao X, Li W, Liao X, Li J, Zhu W, Liu X and Li Q (2025) Development of a prognostic prediction model for nonsmoking lung adenocarcinoma based on pathological information and laboratory hematologic indicators: a multicenter study. *Front. Immunol.* 16:1566195. doi: 10.3389/fimmu.2025.1566195

#### COPYRIGHT

© 2025 Xiang, Hu, Xiao, Li, Liao, Li, Zhu, Liu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Development of a prognostic prediction model for nonsmoking lung adenocarcinoma based on pathological information and laboratory hematologic indicators: a multicenter study

Run Xiang<sup>1,2†</sup>, Peihong Hu<sup>2†</sup>, Xiaoxiong Xiao<sup>3,4</sup>, Wen Li<sup>5,6</sup>, Xiaoqing Liao<sup>7</sup>, Jun Li<sup>8</sup>, Wen Zhu<sup>1</sup>, Xiaoqin Liu<sup>2\*</sup> and Qiang Li<sup>2\*</sup>

<sup>1</sup>State Key Laboratory of Biotherapy, Sichuan University, Chengdu, China, <sup>2</sup>Department of Thoracic Surgery, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital and Institute, Sichuan Cancer Center, University of Electronic Science and Technology of China, Chengdu, China, <sup>3</sup>Department of Thoracic Surgery, Xiangya Hospital, Central South University, Changsha, Hunan, China, <sup>4</sup>Xiangya Lung Cancer Center, Xiangya Hospital, Central South University, National Clinical Research Center for Geriatric Disorders, Changsha, China, <sup>5</sup>Department of Medical Oncology, Cancer Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China, <sup>6</sup>Lung Cancer Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China, <sup>7</sup>Department of Thoracic Surgery, Dazhu County People's Hospital, Dazhou, Sichuan, China, <sup>8</sup>Department of Thoracic Surgery, Ziyang Yanjiang People's Hospital, Ziyang, Sichuan, China

**Objective:** To develop a simple and practical model to predict the prognostic survival of non-smoking patients with lung adenocarcinoma by combining general pathological information with laboratory hematologic indicators.

**Methods:** Cox univariate and multivariate analyses were used to identify the variable indicators. A Cox proportional hazards model was constructed based on the selected variables to compare survival outcomes between the high-and low-risk groups of non-smoking patients with lung adenocarcinoma and to validate the model's performance. Subsequently, a nomogram model was established to systematically evaluate the impact of selected variables on prognosis.

**Results:** Data of non-smoking patients with lung adenocarcinoma from four hospitals were retrospectively collected. We enrolled 1,172 patients, this includes 372 external validation data. Multivariate analysis identified six significant variables (P < 0.05): tumor TNM stage, tumor size, white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin level. We combined these six variables to build a model. The C-index of the training set is 0.811 (0.780–0.842), this value is 0.786 (0.737–0.835) in,test set and 0.810 (0.772–0.847) in validation set. The area under the curve (AUC) results of the predicted 3-years overall survival (OS) of the three data sets were 0.850, 0.819, and 0.860, respectively. These values for 5-years were 0.811, 0.771, and 0.849. Stratified analysis based on tumor staging showed that the model effectively distinguished

outcomes (P < 0.0001). High-risk groups demonstrated significantly poorer prognosis compared to low-risk groups (P < 0.001).

**Conclusion:** The prognostic model based on tumor TNM stage, tumor size, white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin levels effectively predicted the prognosis of non-smoking patients with lung adenocarcinoma. Compared with the more studied blood markers at present, the indicators of our model do not need conversion, Our model provides a useful reference for personalized diagnosis and treatment in clinical practice.

#### KEYWORDS

non-smoking lung adenocarcinoma, proportional hazards model, general pathological information, nomogram, prognostic prediction

### **1** Introduction

The 2022 global cancer statistics report revealed that among 36 types of cancer across 185 countries, Lung cancer has the highest incidence, at the same time, lung cancer also leads the death rate (1). Lung adenocarcinoma is the most studied type of lung cancer due to its high incidence rate in all subtypes of lung cancer. A 2023 article published in The Lancet Oncology on the global distribution of lung cancer subtypes reported that adenocarcinoma constitutes 39% of lung cancers among men and 57% among women, compared to 25% and 12%, respectively, for squamous cell carcinoma (2). Smoking is a well-established independent risk factor for lung cancer. By 2024, it is estimated that 81% of lung cancer deaths will be directly attributable to smoking (5). In some countries, a decrease in the prevalence of smoking has been accompanied by a decline in lung cancer incidence (3). However, approximately 25% of lung cancer cases globally are attributed to factors other than smoking (4). Early diagnosis and treatment strategies for nonsmoking patients with lung cancer may differ slightly from those for smokers. Therefore, early prediction of prognosis in nonsmoking patients with lung cancer is critical for designing personalized treatment plans. Previous studies have rarely distinguished patients with lung cancer according to their smoking status to compare research outcomes.

Traditionally, laboratory test data have been an essential tool for supporting clinical diagnosis and evaluating treatment efficacy. In recent years, with the rise of bioinformatics, clinical laboratory omics have been recognized for their potential to uncover previously unexplored clinical value. For instance, platelets, long known for their critical role in thrombosis, have recently garnered attention as biomarkers for monitoring tumor progression (6, 7). Similarly, serum albumin has been identified as a biomarker for small-cell lung cancer patients undergoing immunotherapy (8). Additionally, the neutrophil-to-lymphocyte ratio (NLR) has shown potential as a prognostic marker in patients with colon cancer patients with isolated liver metastases (9). Therefore, developing a prognostic prediction model for lung cancer based on common blood biomarkers is a highly promising area of investigation.

This study was intended to offer valuable perspectives for clinical decision-making and explore the potential of blood biomarkers for predicting prognosis in non-smoking lung adenocarcinoma patients. The indicators used in this study, including neutrophil percentage, lymphocyte percentage, hemoglobin level, and white blood cell count, are routinely available from complete blood count (CBC) tests. These metrics can be directly measured without additional calculations, making them the most accessible and fundamental clinical testing parameters. Their low cost, simplicity, and feasibility make them practical tools for predicting lung cancer ending.

### 2 Methods

### 2.1 Patients selection

Figure 1 illustrates the study's inclusion and exclusion criteria, as well as the study workflow. We retrospectively analyzed 2,832 cases of lung cancer diagnosed at Sichuan Cancer Hospital between November 2013 and April 2021. The ones that followed were the exclusion criteria: (1) individuals who have smoked in the past; (2) individuals with small cell carcinoma, squamous cell carcinoma, or other forms of lung cancer; (3) individuals with other concurrent malignancies; and (4) individuals lost to follow-up. Additionally, using the same inclusion and exclusion criteria, we collected samples from four hospitals-West China Hospital of Sichuan University, Xiangya Hospital of Central South University, Dazhu County People's Hospital, and Yanjiang District People's Hospital of Ziyang-for external validation of the model. The Sichuan Cancer Hospital's Ethics Committee gave its approval to this study (No. KY-2021-076). The informed consent exemption statement was carried out, and our investigation was retrospective.



### 2.2 Study design

The end point of this study was overall survival (OS). All included patients were followed up via telephone, and general patient information, such as sex, age, smoking history, tumor size, and tumor stage, was extracted from hospital integrated information system. The clinical blood test data used in the study were obtained from CBC results, which were analyzed using an automated hematology analyzer (Shenzhen Mindray, BC-5390).

### 2.3 Variable screening

Cox univariate analysis was performed to screen the CBC results of the included cases. Variables with P < 0.05 in the univariate analysis were subsequently subjected to Cox multivariate analysis to identify the key hematologic biomarkers for the study.

### 2.4 Model construction and evaluation

A training set and a test set were randomly selected from the modeling group data in a 7:3 ratio. The prediction model was established using the training set, and the model was internally verified using the test set, and the validation set data were used for external verification of the prediction performance of the model. Harrell's C-index was the main evaluation index of the model. Additionally, the Bootstrap self-sampling method was used with 500 sampling times. A time-dependent receiver operating characteristic (ROC) curve with the area under the curve (AUC) was used to evaluate the prediction performance. Additionally, we create the model's ROC curves for patients at 3- and 5- years, and we use the AUC to judge the model's accuracy. The calibration curve and decision curve were used to evaluate the calibration performance and clinical applicability of the model, respectively. According to the median OS prediction results of the model for the patients, those in the training, test, and validation sets were divided into high- and low-risk groups. Participants were split into different risk(high and low) groups for the training, test, and validation sets. The survival curves were created using the Kaplan-Meier method, and the prognostic results of each risk grouping were compared using log-rank test. Finally, we used a nomogram to visualize the established model.

### 2.5 Statistical methods

To conduct statistical studies, R software (version 4.1.0) was utilized. Continuous variables were tested for normality, when describing the ones who did not fit into a normal distribution, the median (25th to 75th percentiles; M [P25, P75]) was selected. The Kruskal–Wallis test was used for comparisons of continuous

variables, whereas the chi-squared test was applied for categorical variables. Statistical significance was set at P < 0.05.

# **3 Results**

# 3.1 Clinical pathological characteristics of non-smoking patients with lung adenocarcinoma

800 patients who did not smoke in total with lung adenocarcinoma diagnosed at Sichuan Cancer Hospital were included in this study. Additionally, 372 patients from four other hospitals—West China Hospital of Sichuan University, Xiangya Hospital of Central South University, Dazhu County People's Hospital, and Yanjiang District People's Hospital of Ziyang—were used as the external validation cohort. Lung cancer pathology staging was determined based on the eighth edition of the National Comprehensive Cancer Network (NCCN) lung cancer classification standards. All of the patients' median ages were 60 (56.4, 70.4) years. The median duration of follow-up was 42.3 (36.5, 59.7) months. Among all patients, 902 (77.0%) were female, and 322 (27.5%) died. Among all patients, 233 (19.9%) were classified as Stage III and 253 (21.6%) as Stage IV. The basic data of the patients in the training, test, and validation sets are shown in Table 1.

### 3.2 Univariate and multivariate Cox analysis

We carried out univariate and multivariate Cox analyses to investigate the association between the clinical features and OS. The results identified the following independent prognostic factors for patient outcomes: tumor TNM stage, tumor size (hazard ratio (HR): 1.01, 95% confidence interval (CI):1.01–1.02, P= 0.001), white blood cell count (HR: 1.09, 95% CI: 1.05–1.14, P< 0.001), neutrophil percentage (HR: 0.94, 95% CI: 0.91–0.98, P= 0.001), lymphocyte percentage (HR: 0.97, 95% CI: 0.87–0.94, P< 0.001), and hemoglobin level (HR: 0.98, 95% CI: 0.97–0.99, P< 0.001; Figure 2). These four laboratory markers (white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin level) significantly affect patient prognosis.

TABLE 1 Clinical features of 1172 patients with LUAD	Э.
--	----

Variables	Training set (n = 560)	Test set (n = 240)	Validation set (n = 372)	Statistic	Р
Time, Month	40.7 (36.3,58.5)	40.0 (36.5,57.8)	54.0 (40.3,64.5)	$\chi^2 = 27.91^{\#}$	<.001
Size, mm	20 (17,32)	20 (17,33)	24 (16,36)	$\chi^{2} = 2.88^{\#}$	0.237
Age, year	60 (51,66)	58 (50,67)	62 (55,68)	$\chi^2=20.05^{\#}$	<.001
WBC	5.58 (4.62,6.77)	5.65 (4.62,6.92)	5.67 (4.60,7.11)	$\chi^2=0.60^{\#}$	0.741
Percentage of Neutrophil	64.3 (57.0,70.7)	63.8 (55.8,70.0)	64.0 (56.4,69.7)	$\chi^2=0.08^{\#}$	0.960
Percentage of Lymphocyte	26.6 (20.4,33.4)	26.4 (21.2,33.9)	21.7 (1.7,31.9)	$\chi^2 = 62.79^{\#}$	<.001
Percentage of Monocyte	5.8 (4.8,6.7)	6.0 (4.8,6.9)	6.3 (5.1,7.7)	$\chi^2 = 22.51^{\#}$	<.001
Hb	130 (120,139)	129 (120,137)	127 (118,137)	$\chi^2 = 4.36^{\#}$	0.113
Status, n (%)				$\chi^2 = 0.15$	0.926
Alive	406 (72.5)	172 (71.7)	272 (73.1)		
Dead	154 (27.5)	68 (28.3)	100 (26.9)		
Sex, n (%)				$\chi^{2} = 7.26$	0.026
Female	450 (80.4)	180 (75.0)	272 (73.1)		
Male	110 (19.6)	60 (25.0)	100 (26.9)		
Stage, n (%)				$\chi^{2} = 12.23$	0.057
Ι	277 (49.5)	103 (42.9)	208 (55.9)		
II	42 (7.50)	24 (10.0)	32 (8.6)		
III	112 (20.0)	57 (23.8)	64 (17.2)		
IV	129 (23.0)	56 (23.3)	68 (18.3)		

<sup>#</sup>Kruskal–Wallis test, χ<sup>2</sup>Chi-squared test.



# 3.3 A six-variable combined model for predicting prognosis in non-smoking patients with lung adenocarcinoma

A model of Cox proportional hazards was established with the variables chosen from the Cox multivariate analysis. In the training cohort, the ROC curve analysis revealed that the AUC for the 3- and 5-years OS was 0.850 and 0.811 (Figure 3A). In the test cohort, the AUC was 0.819 and 0.771, respectively (Figure 3B). Finally, the AUC in the external validation cohort was 0.860 and 0.849, respectively (Figure 3C). Figures 3D–F show the calibration curves for the training, testing, and validation cohorts.

Decision curve analysis (DCA) was used to assess the clinical benefits of our model. Within the training dataset, the threshold probability with a practical value ranged from 0 to 0.75 (Figure 4A); in the testing cohort, it ranged from 0 to 0.7 (Figure 4B); and in the validation cohort, it ranged from 0 to 0.85 (Figure 4C). Patients were separated into different risk groups (high or low) using the median values of tumor TNM stage, tumor size, white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin level as cutoff points. Kaplan–Meier analysis was performed to evaluate the relationship between these variables and OS in non-smoking patients with lung adenocarcinoma. Survival analysis demonstrated remarkable dissimilarities in OS between the high- and low-risk sets in the training (Figure 4E), testing (Figure 4E), and validation datasets (Figure 4F; P < 0.0001).

The model's predictive ability across different clinical stages of non-smoking lung adenocarcinoma was a key focus of this study. By using the log-rank test method, Figure 5A shows the subgroup analysis results for different pathological stages in the training cohort, whereas Figure 5B presents the subgroup analysis results in the validation cohort. In both the modeling and validation datasets, the survival curves between different stages of lung adenocarcinoma were different, and the difference was statistically significant: the survival of stage I and II was better than that of stage III and IV, *P* < 0.0001.

# 3.4 Establishment and validation of the nomogram based prognostic prediction model

Based on the regression coefficients of the influencing factors from the Cox proportional hazards model, scores were assigned to each of the following factors: pathological stage, tumor size, white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin level. The scores for each factor were summed to obtain a total score. This total score was converted into an estimated survival rate, which was then used to calculate the predicted tumor progression value and create a nomogram for predicting the prognosis of non-smoking patients with lung adenocarcinoma (Figure 6). Each factor was represented by a line segment on the nomogram, with scale marks indicating the range of possible values for that factor. The length of the line segment reflects the size of the regression coefficient of the factor in relation to the outcome. The total score, including both individual and cumulative scores, was determined by drawing a vertical line to determine the corresponding 3- and 5-years survival rates. Bootstrap resampling was performed 500 times to validate the proposed model. The Cindex values for the training, testing, and validation cohorts were greater than 0.7 (Supplementary Figure S1A-C), indicating that the model had significant predictive power.

### 4 Discussion

In recent years, research on prognostic biomarkers of lung cancer has gained significant attention; however, most studies have been limited to data from approximately 200 patients. In our retrospective analysis, we used data from nearly 1,200 cases to construct a prognostic prediction model for non-smoking lung adenocarcinoma based on blood analysis indicators. Furthermore, we conducted a multicenter external validation to assess the prognostic efficacy of peripheral blood markers. Our results



Cox proportional hazard models were established and validated. (A) Receiver operating characteristic (ROC) curve of the training dataset. (B) ROC curve of the test dataset. (C) ROC curve of the validation dataset. (D) Calibration curve of the training dataset. (E) Calibration curve of the test dataset. (F) Calibration curve of the validation dataset.

demonstrated that the high-risk group, characterized by later tumor stages, larger tumor size, higher white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin levels, had a markedly worse prognosis than the low-risk group. Additionally, we developed a nomogram for prognostic prediction that serves as a valuable reference for assessing the prognosis of these patients.

White blood cells comprise neutrophils, lymphocytes, monocytes, and other types. The roles of different white blood cell components have been widely explored. Generally, white blood cells represent the body's level of inflammation, which has a critical function in the development and evolution of tumors (10-12). Some scholars have suggested that chronic inflammation is a



Meier curve of the training dataset. (E) Kaplan-Meier curve of the test dataset. (F) Kaplan-Meier curve of the validation dataset



constant feature during tumor development, leading to the idea that tumors are "non-healing wounds" (13). In previous studies, the NLR has been frequently used as a prognostic marker. The NLR has different predictive roles in various types of cancer (14–16); however, its value as a prognostic predictor remains controversial.

Some studies have suggested that the lymphocyte percentage may better reflect cancer prognosis than the absolute lymphocyte count (17). Researchers have indicated that a low lymphocyte percentage could serve as a marker for poor prognosis (18, 19), whereas others have found that high neutrophil and lymphocyte percentages are linked to a worse prognosis in patients suffering from colorectal cancer (20). Research specifically focusing on the lymphocyte percentage in lung cancer is limited. In this study, we incorporated both neutrophil and lymphocyte percentages into our model. These two markers are directly available in routine blood tests without the need to calculate a ratio, making them more feasible for clinical application.

Hemoglobin, which is primarily found in red blood cells, is closely associated with tumor development. Existing research suggests that low hemoglobin levels lead to tissue hypoxia,

Points	0	10		20	30	40	50	60		70	80	90		100
Stage	Ĺ	II		, III	IV									
Size(mm)	ō	20 40	60	80	 110									
Percentage of Neutrophil	95	90	85	80	75	70	65 6	0 55	50	45	40	35	30	
Percentage of Lymphocyte	55	50		45	40	35	30	25		20	15	10		5
Hb	180	150	120	90										
WBC	ō	5	10	1	5 2	0 25	30	35	40					
Total Points	, 0	20	40	6	io 8	30 10	0 12	0 140	160	180	200	) 22	0	 240
3-year survival prob								0.9 0.8	30.7 0	5 0.3 0.	1			
5 year survival prob														

Nomogram for predicting the prognosis of lung adenocarcinoma in non-smoking patients.

triggering a series of body responses (21, 22). A retrospective analysis based on a European population showed that high hemoglobin levels were beneficial in reducing the incidence of prostate cancer (23, 24); however, in other malignancies, hemoglobin levels may exhibit different effects. For example, high hemoglobin levels may augment the danger of cervical cancer and melanoma (25).

The influence of tumor staging and tumor size on prognosis is straightforward. Later-stage tumors generally have a poorer prognosis, and larger tumor diameters are often associated with more complications, leading to worse outcomes. On the other hand, blood test indicators have varying effects across different cancer types. Additionally, even within the same cancer type, the effects can differ at various stages or across different patient populations, possibly due to the inherent heterogeneity of tumors. Therefore, our study did not focus on a single marker but instead combined several risk factors into a model to explore their role in predicting the prognosis of non-smoking lung adenocarcinoma. Moreover, in clinical practice, these indicators are simple, readily available, and do not require complex calculations, which enhances their clinical applicability. This is a key advantage of this study. By providing individualized risk predictions and dynamic updates, our model offers actionable insights that can directly inform clinical decisionmaking, a feature that is often lacking in traditional risk scoring systems. In terms of innovation, the model performs stably in the external validation set, avoiding the overfitting problem that may exist in traditional methods. In addition, our model can more finely stratifying or predicting the risk of individual patients, thus providing support for personalized treatment.

Our study has some limitations. There are inevitable biases in sample selection, and our prediction model may perform well in specific regions and specific populations, but there may be insufficient generalization ability in different medical institutions or populations, which needs to be verified and optimized in large sample environment. In reality, the outcome of the disease may be improved by treatment, therefore, in future work, we will also incorporate propensity-score matching or difference-in-differences methods into our models to assess treatment effects on disease outcomes.

### **5** Conclusion

In conclusion, this study aimed to analyze the prognosis of nonsmoking patients with lung adenocarcinoma using a combined model including tumor stage, tumor size, white blood cell count, neutrophil percentage, lymphocyte percentage, and hemoglobin level. The nomogram developed in our research can effectively predict the prognosis of non-smoking patients with lung adenocarcinoma, helping clinicians screen high-risk patients and offering personalized treatment.

### Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

### Ethics statement

The studies involving humans were approved by Sichuan Cancer Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

### Author contributions

RX: Writing – original draft. PH: Data curation, Formal Analysis, Writing – review & editing. XX: Resources, Writing – review & editing. WL: Resources, Writing – review & editing. XLia: Resources, Writing – review & editing. JL: Resources, Writing – review & editing. WZ: Conceptualization, Writing – review & editing. XLiu: Data curation, Formal Analysis, Writing – review & editing. QL: Conceptualization, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The article was funded by Sichuan Medical Association research project (NO.2024HR13), National Natural Science Youth Foundation of China (82203838).

### Acknowledgments

We thank https://app.editage.com/ for guidance on the language of the article.

# Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

### References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2024) 74:229–63. doi: 10.3322/caac.21834

2. Zhang Y, Vaccarella S, Morgan E, Li M, Etxeberria J, Chokunonga E, et al. Global variations in lung cancer incidence by histological subtype in 2020: a population-based study. *Lancet Oncol.* (2023) 24:1206–18. doi: 10.1016/S1470-2045(23)00444-8

3. Miranda-Filho A, Piñeros M, Bray F. The descriptive epidemiology of lung cancer and tobaccocontrol:aglobaloverview2018. *SaludPublicaMex.*(2019)61:219–29. doi:10.21149/10140

4. Mu L, Liu L, Niu R, Zhao B, Shi J, Li Y, et al. Indoor air pollution and risk of lung cancer among Chinese female non-smokers. *Cancer Causes Control*. (2013) 24:439–50. doi: 10.1007/s10552-012-0130-8

5. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. (2024) 74:12-49. doi: 10.3322/caac.21820

6. Best MG, Wesseling P, Wurdinger T. Tumor-educated platelets as a noninvasive biomarker source for cancer detection and progression monitoring. *Cancer Res.* (2018) 78:3407–12. doi: 10.1158/0008-5472.CAN-18-0887

7. Zu R, Yu S, Yang G, Ge Y, Wang D, Zhang L, et al. Integration of platelet features in blood and platelet rich plasma for detection of lung cancer. *Clin Chim Acta.* (2020) 509:43–51. doi: 10.1016/j.cca.2020.05.043

8. Kuang Z, Miao J, Zhang X. Serum albumin and derived neutrophil-to-lymphocyte ratio are potential predictive biomarkers for immune checkpoint inhibitors in small cell lung cancer. *Front Immunol.* (2024) 15:1327449. doi: 10.3389/fimmu.2024.1327449

9. Polk N, Budai B, Hitre E, Patócs A, Mersich T. High neutrophil-to-lymphocyte ratio (NLR) and systemic immune-inflammation index (SII) are markers of longer survival after metastasectomy of patients with liver-only metastasis of rectal cancer. *Pathol Oncol Res.* (2022) 28:1610315. doi: 10.3389/pore.2022.1610315

10. Singh N, Baby D, Rajguru JP, Patil PB, Thakkannavar SS, Pujari VB. Inflammation and cancer. Ann Afr Med. (2019) 18:121-6. doi: 10.4103/aam.aam\_56\_18

11. Fernandes Q, Inchakalody VP, Bedhiafi T, Mestiri S, Taib N, Uddin S, et al. Chronic inflammation and cancer; the two sides of a coin. *Life Sci.* (2024) 338:122390. doi: 10.1016/j.lfs.2023.122390

12. Tezcan G, Yakar N, Hasturk H, Van Dyke TE, Kantarci A. Resolution of chronic inflammation and cancer. *Periodontol 2000.* (2024) 96:229–49. doi: 10.1111/prd.12603

13. Singel KL, Segal BH. Neutrophils in the tumor microenvironment: trying to heal the wound that cannot heal. *Immunol Rev.* (2016) 273:329–43. doi: 10.1111/imr.12459

14. Nøst TH, Alcala K, Urbarova I, Byrne KS, Guida F, Sandanger TM, et al. Systemic inflammation markers and cancer incidence in the UK Biobank. *Eur J Epidemiol.* (2021) 36:841–8. doi: 10.1007/s10654-021-00752-6

### Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1566195/ full#supplementary-material

#### SUPPLEMENTARY FIGURE 1

Self-sampling validation of the model. (A) Self-sampling validation of the training dataset. (B) Self-sampling validation of the test dataset. (C) Self-sampling validation of the validation dataset.

15. Cupp MA, Cariolou M, Tzoulaki I, Aune D, Evangelou E, Berlanga-Taylor AJ. Neutrophil to lymphocyte ratio and cancer prognosis: an umbrella review of systematic reviews and meta-analyses of observational studies. *BMC Med.* (2020) 18:360. doi: 10.1186/s12916-020-01817-1

16. Mosca M, Nigro MC, Pagani R, De Giglio A, Di Federico A. Neutrophil-tolymphocyte ratio (NLR) in NSCLC, gastrointestinal, and other solid tumors: immunotherapy and beyond. *Biomolecules*. (2023) 13:1803. doi: 10.3390/ biom13121803

17. Iseki Y, Shibutani M, Maeda K, Nagahara H, Tamura T, Ohira G, et al. The impact of the preoperative peripheral lymphocyte count and lymphocyte percentage in patients with colorectal cancer. *Surg Today.* (2017) 47:743–54. doi: 10.1007/s00595-016-1433-2

18. Ma C, Wang X, Zhao R. Associations of lymphocyte percentage and red blood cell distribution width with risk of lung cancer. *J Int Med Res.* (2019) 47:3099–108. doi: 10.1177/0300060519850417

19. Huang H, Li L, Luo W, Yang Y, Ni Y, Song T, et al. Lymphocyte percentage as a valuable predictor of prognosis in lung cancer. *J Cell Mol Med.* (2022) 26:1918–31. doi: 10.1111/jcmm.17214

20. Song X, Zhu H, Pei Q, Tan F, Li C, Zhou Z, et al. Significance of inflammationbased indices in the prognosis of patients with non-metastatic colorectal cancer. *Oncotarget.* (2017) 8:45178–89. doi: 10.18632/oncotarget.16774

21. Bellelli A, Tame JRH. Hemoglobin allostery and pharmacology. *Mol Aspects Med.* (2022) 84:101037. doi: 10.1016/j.mam.2021.101037

22. Drvenica IT, Stančić AZ, Maslovarić IS, Trivanović DI, Ilić VL. Extracellular hemoglobin: modulation of cellular functions and pathophysiological effects. *Biomolecules.* (2022) 12:1708. doi: 10.3390/biom12111708

23. Hakozaki Y, Yamada Y, Takeshima Y, Taguchi S, Kawai T, Nakamura M, et al. Low hemoglobin and PSA kinetics are prognostic factors of overall survival in metastatic castration-resistant prostate cancer patients. *Sci Rep.* (2023) 13:2672. doi: 10.1038/s41598-023-29634-5

24. Mori K, Janisch F, Mostafaei H, Lysenko I, Karakiewicz PI, Enikeev DV, et al. Prognostic value of hemoglobin in metastatic hormone-sensitive prostate cancer: A systematic review and meta-analysis. *Clin Genitourin Cancer*. (2020) 18:e402–9. doi: 10.1016/j.clgc.2019.12.002

25. Li GS, Huang T, Li JX, Liu J, Gao X, Yang N, et al. Correlation between hemoglobin and the risk of common Malignant tumors: a 1999-2020 retrospective analysis and causal association analysis. *BMC Cancer.* (2024) 24:755. doi: 10.1186/s12885-024-12495-0