Heliyon 10 (2024) e37100

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Construction of a pathomics model for predicting mRNAsi in lung adenocarcinoma and exploration of biological mechanism

Rui Chen [a], Yuzhen Liu [b], Junping Xie [a,*]

[a] Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital, Jiangxi Medical College, Nanchang University, No.1, Minde Road, Donghu District, Nanchang, Jiangxi, 330006, China
[b] Department of Oncology, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang, China

ARTICLE INFO

Keywords:
Pathomics
mRNAsi
LUAD
PyRadiomics
GBM

ABSTRACT

Objective: This study aimed to predict the level of stemness index (mRNAsi) and survival prognosis of lung adenocarcinoma (LUAD) using pathomics model.

Methods: From The Cancer Genome Atlas (TCGA) database, 327 LUAD patients were randomly assigned to a training set (n = 229) and a validation set (n = 98) for pathomics model development and evaluation. PyRadiomics was used to extract pathomics features, followed by feature selection using the mRMR-RFE algorithm. In the training set, Gradient Boosting Machine (GBM) was utilized to establish a model for predicting mRNAsi in LUAD. The model's predictive performance was evaluated using ROC curves, calibration curves, and decision curve analysis (DCA). Prognostic analysis was conducted using Kaplan-Meier curves and cox regression. Additionally, gene enrichment analysis, tumor microenvironment analysis, and tumor mutational burden (TMB) analysis were performed to explore the biological mechanisms underlying the pathomics prediction model.

Results: Multivariable cox analysis (HR = 1.488, 95 % CI 1.012–2.187, P = 0.043) identified mRNAsi as a prognostic risk factor for LUAD. A total of 465 pathomics features were extracted from TCGA-LUAD histopathological images, and ultimately, the most representative 8 features were selected to construct the predictive model. ROC curves demonstrated the significant predictive value of the model for mRNAsi in both the training set (AUC = 0.769) and the validation set (AUC = 0.757). Calibration curves and Hosmer-Lemeshow goodness-of-fit test showed good consistency between the model's prediction of mRNAsi levels and the actual values. DCA indicated a good net benefit of the model. The prediction of mRNAsi levels by the pathomics model is represented using the pathomics score (PS). PS was strongly associated with the prognosis of LUAD (HR = 1.496, 95 % CI 1.008–2.222, P = 0.046). Signaling pathways related to DNA replication and damage repair were significantly enriched in the high PS group. Prediction of immune therapy response indicated significantly reduced Dysfunction in the high PS group (P < 0.001). The high PS group exhibited higher TMB values (P < 0.001).

Conclusions: The predictive model constructed based on pathomics features can forecast the mRNAsi and survival risk of LUAD. This model holds promise to aid clinical practitioners in identifying high-risk patients and devising more optimized treatment plans for patients by jointly employing therapeutic strategies targeting cancer stem cells (CSCs).

* Corresponding author.
E-mail address: junpingxie2023@126.com (J. Xie).

https://doi.org/10.1016/j.heliyon.2024.e37100
Received 8 May 2024; Received in revised form 4 August 2024; Accepted 27 August 2024
Available online 29 August 2024

1. Introduction

Lung cancer, notorious for its high incidence, insidious onset, and poor prognosis, remains a significant health burden [1]. According to the latest Cancer Statistics data, the estimated new cases of tumors in the lung and bronchus remain substantial, with 116,310 cases in males and 118,270 cases in females, and the estimated number of deaths consistently ranks first among all categories of tumor-related deaths [2]. Lung adenocarcinoma (LUAD) is the most prevalent subtype of lung cancer [3]. Currently, the treatment modalities for LUAD mainly include surgery, targeted therapy, radiotherapy, chemotherapy, and immunotherapy, but the overall prognosis for advanced-stage patients remains poor [4]. For mid-to-late-stage lung cancer patients who have lost the opportunity for surgery, the treatment goals are to prolong survival, improve quality of life, and strive for long-term survival with the disease. LUAD prognosis is typically determined by a combination of clinicopathological factors, including tumor stage, grade, and patient age, as well as serum biomarkers such as carcinoembryonic antigen (CEA) and soluble fragment of cytokeratin 19 (CYFRA21-1) and radiological features identified by computed tomography scans [5,6]. However, as current indicators are insufficient to meet the rigorous requirements of precision medicine, there is an urgent need for the discovery of novel biomarkers to enable accurate patient stratification and personalized treatment strategies.

Cancer stem cells (CSCs), a subset of undifferentiated cells, have been identified as key drivers of tumorigenesis in various tumor types and offer significant advantages in terms of initiation, progression, and resistance to therapy [7]. These rare and elusive cells can infiltrate distant organs during early disease stages, where they may reside in specialized niches. Following primary tumor treatment, these dormant cells can be reactivated, leading to recurrent disease [8]. It has been demonstrated that poorly prognostic LUAD shows an increase in undifferentiated stem cell populations or stem cell-like characteristics, which is one of the reasons for its high invasiveness [9,10]. A study by Tathiane M Malta et al. stated that an algorithm can generate a stemness index (mRNAsi) based on mRNA expression, which is commonly used to evaluate and quantify the stemness characteristics of tumor cells, namely, whether cells possess stem cell-like properties such as self-renewal and multifunctionality differentiation potential [11]. The stemness features of tumors are intimately linked to their ability to maintain proliferation, metastasize, and evade therapy, thereby correlating with enhanced invasiveness and treatment resistance. Prognostic studies based on mRNAsi have been successfully applied to various types of cancers, such as liver cancer, squamous cell lung carcinoma, and glioblastoma, revealing its significant potential as a prognostic indicator for tumors [12–16]. Despite the plethora of studies examining the correlation between mRNAsi and various tumor prognoses, the majority of these investigations rely on indirect approaches, as mRNAsi must be inferred from gene expression data derived from samples. However, this methodology is hindered by limitations arising from the challenges associated with sample collection and high detection costs, thereby restricting the translational potential of mRNAsi in disease diagnosis and treatment.

Recent advances in artificial intelligence (AI) have revolutionized the field of pathology, enabling the conversion of pathology images into high-fidelity, high-throughput, and exploitable data. Pathomics, a subspecialty of digital pathology, leverages AI to extract quantitative features such as texture, morphology, edge gradient, and biological characteristics from images, which are then integrated for comprehensive analysis in diagnostic, molecular, and prognostic applications [17–20]. Compared to traditional visual observation, computer-assisted image analysis significantly enhances computational efficiency, accuracy, and cost-effectiveness in tumor diagnosis. Notwithstanding the complexities of tumor heterogeneity and multiple driving factors influencing cancer prognosis, recent studies have endeavored to integrate histopathological and molecular biological data to develop predictive models that provide a more comprehensive understanding of tumor behavior [21,22].

Hematoxylin and eosin (H&E) stained slides are essential for clinical diagnosis and represent the most readily available imaging

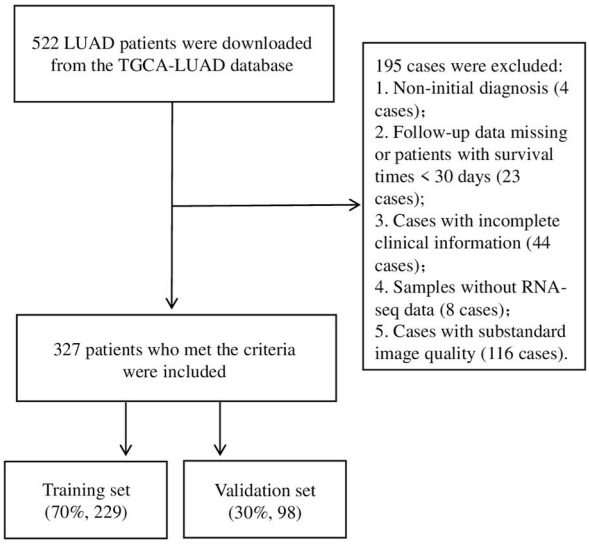

Fig. 1. Flowchart showing the inclusion and exclusion criteria for LUAD patients in this study.

data. Therefore, integrating mRNAsi with pathological image information to develop multi-omics prognostic prediction models has become a very promising research direction. Based on these factors, this study innovatively proposes predicting the mRNAsi levels of LUAD samples through pathomics technology while integrating bioinformatics analysis to explore the underlying biological mechanisms behind pathomics.

2. Materials and methods

2.1. Data source

A total of 522 LUAD patients' H&E-stained tissue pathology images were obtained from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/). A total of 327 eligible LUAD patients were selected from the initial cohort after applying strict inclusion criteria: primary diagnosis, adequate image quality, complete follow-up data, and survival time ≥30 days. A random partitioning of the patient cohort into training and validation sets was performed with a 7:3 ratio, allowing for robust evaluation of model performance (Fig. 1). We also downloaded corresponding clinical feature data, somatic mutation, and mRNA sequencing data for these patients, with an intersection sample size of 322 cases between somatic mutations and pathomics. The mRNA stemness index (mRNAsi) of tumor samples was quantified using the algorithm proposed by Tathiane M Malta et al. [11]. The R package "survMisc" was utilized to calculate the cutoff values of mRNAsi for these samples, dividing LUAD patients into high/low index groups, with the low index group as the reference. The overall workflow diagram is shown in Fig. 2.

2.2. mRNAsi as a prognostic predictor of LUAD

The Kaplan-Meier survival curve was used to illustrate changes in survival rates between high and low mRNAsi groups, and the Log-rank test was employed to assess the significance of survival rates between the two groups. Additionally, the R package "survival" was utilized for survival analysis of each variable, while the R package "survminer" was employed to summarize and visualize the analysis results. To investigate the independent risk factors for overall survival (OS) in LUAD patients, we performed univariate and multivariate Cox regression analyses on a comprehensive set of variables, including mRNAsi, age, tumor stage, primary location, and other clinicopathological factors. Furthermore, we performed subgroup analyses using univariate Cox regression to investigate the

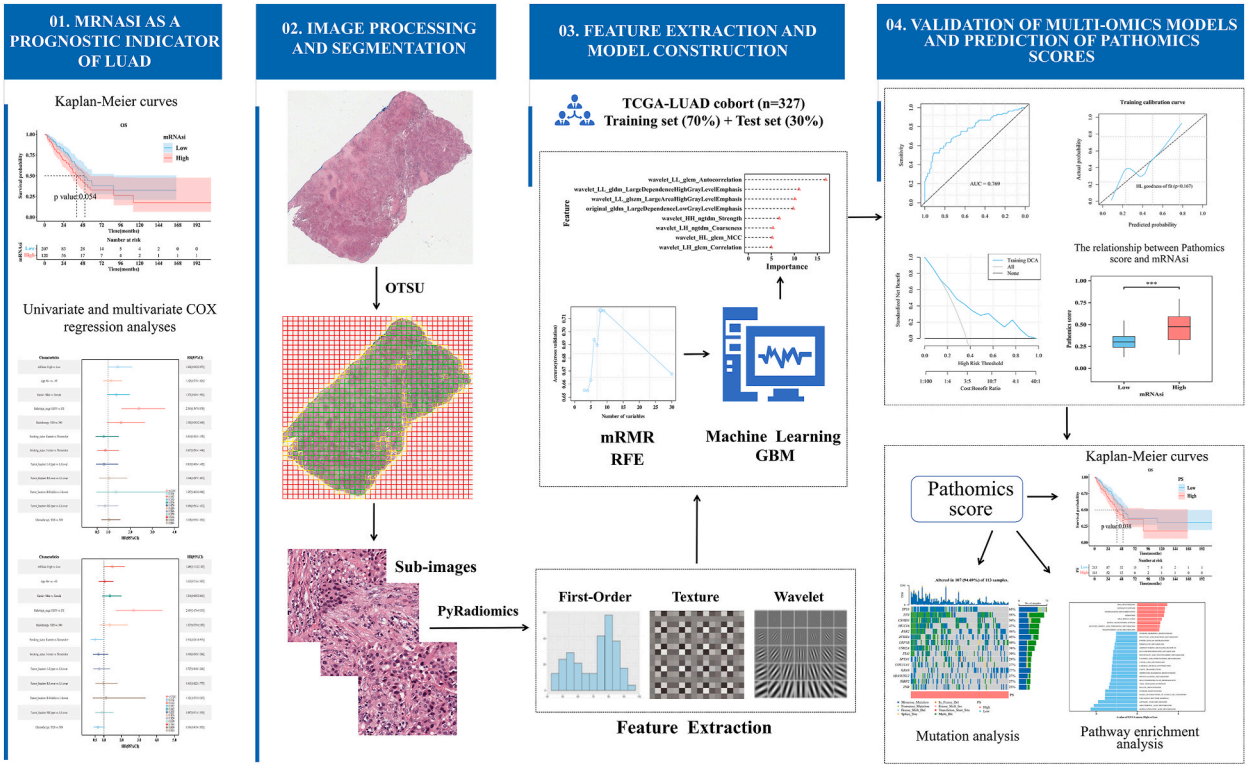

Fig. 2. The flowchart consists of four parts: (1) The potential of mRNAsi as a prognostic factor for LUAD; (2) Processing and segmentation of pathological images using the OTSU algorithm; (3) Feature extraction of pathological images using PyRadiomics, selection of important features using mRMR_RFE, and model construction using the GBM algorithm; (4) Evaluation of the model and exploration of the potential biological significance of pathomics score.

association between mRNAsi (high vs. low index group) and patient prognosis in subgroups defined by each covariate. Likelihood ratio tests were conducted to assess interactions between mRNAsi and other covariates.

2.3. Processing and segmentation of histopathological images

Qualified svs format LUAD pathological images were downloaded from the TCGA database, with a maximum magnification of 20 × or 40 × . The OTSU algorithm (https://opencv.org/) was utilized to process the pathological images and obtain the tissue regions of the slides [23]. Image segmentation was performed on the 40 × and 20 × datasets, yielding 1024 × 1024 pixel sub-images from the former and 512 × 512 pixel sub-images that were subsequently upsampled to 1024 × 1024 pixels. Subsequently, experienced pathologists carefully reviewed the images to exclude sub-images with poor quality, such as contamination, image blurring, or blank areas exceeding 50 %. Ten randomly selected sub-images from the qualified ones were then chosen for subsequent analysis.

2.4. Extraction and screening of pathomics features

The PyRadiomics open-source package (https://pyradiomics.readthedocs.io/en/latest/) [24] was used to standardize the images of the training set's sub-images and extract 93 original features (including first-order and second-order features) and 372 high-order features, namely wavelet features (LL, LH, HL, HH), consisting of intensity, gradient, and texture features extracted from wavelet-filtered images. Pathomics features were extracted from 10 sub-images of each patient's pathological image, with the mean value serving as the representative pathomics feature for each sample. To facilitate effective feature analysis, we standardized the pathomics feature values of the training set using z-scores and applied the same transformation to the validation set using the mean and standard deviation derived from the training data. Furthermore, we investigated differences in clinical variables between the datasets.

Additionally, the Maximum Relevance, Minimum Redundancy (mRMR) algorithm, and Recursive Feature Elimination (RFE) algorithm were employed to select the optimal feature subset. The mRMR algorithm selects features that balance the correlation between features, as well as the correlation between features and the target variable. On the other hand, the RFE feature selection involves ranking the predictive factors before modeling and sequentially eliminating less important factors. Its objective is to continuously train the model to find the best predictor subset that can be used to generate accurate models. In this study, the mRMR method was first used to select the top 30 features, followed by RFE feature selection to identify the optimal features.

2.5. Model construction and evaluation

In this study, the Gradient Boosting Machine (GBM) algorithm was employed to construct predictive models for mRNAsi in LUAD samples using the pathomics features selected by the mRMR and RFE algorithms in the training set. The model's effectiveness is assessed by analyzing it using the receiver operating characteristic (ROC) curve. Additionally, by generating calibration curves and applying the Hosmer-Lemeshow goodness-of-fit test, we assessed the calibration of the predictive model. We employed the Brier score to evaluate the overall predictive performance of our pathomics model, and decision curve analysis (DCA) to illustrate its clinical utility, providing a comprehensive assessment of the model's diagnostic accuracy and decision-making value.

2.6. Pathomics score as a prognostic predictor of LUAD

The prediction of mRNAsi in LUAD samples by the model we constructed is represented by the pathomics score (PS). Using the R package "ggplot2", we visualized the differences in PS between the high and low mRNAsi groups and used the Wilcoxon test to compare whether the differences between the two groups were statistically significant. Additionally, we merged the PS with the patient's clinical data and calculated the cutoff value of PS using the "survMisc" package, dividing it into high and low PS groups, and plotted baseline data tables for each clinical variable. Similar to the previous section, Kaplan-Meier survival curves, cox regression analyses, subgroup analyses, and likelihood ratio tests were used for prognostic analyses.

2.7. Functional enrichment analysis

The analysis of KEGG pathway gene sets comprised 186 pathways, while the analysis of Hallmark gene sets included 50 pathways. Differential analysis between high and low PS groups was performed using the R package "limma," with $|t| = 1$ as the threshold, and the top 30 pathways were visualized.

2.8. Analysis of the immune microenvironment and prediction of immunotherapeutic response

The differences in the expression of 37 immune-related genes between the high and low PS groups were analyzed using the Wilcoxon test [25]. We uploaded the gene expression matrix of LUAD samples to CIBERSORTx (https://cibersortx.stanford.edu/) and quantified immune cell infiltration in each sample. Subsequent analysis revealed significant differences in immune cell infiltration profiles between patients with high and low PS levels, which highlights the potential role of tumor microenvironment in patient outcomes. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was used to assess patient response to immunotherapy. Specifically, normalized transcriptomic data were uploaded to the TIDE database (http://tide.dfci.harvard.edu) to calculate the TIDE score, MSI Expr Sig (Microsatellite Instability Expression Signature), Dysfunction and Exclusion scores. Subsequently, differences in

these metrics were comparatively analyzed between the high and low PS groups.

2.9. Gene mutation analysis

The somatic mutation data in mutation annotation format (MAF) of LUAD samples were downloaded from the TCGA database (https://portal.gdc.cancer.gov/). The "maftools" package was employed to compute the tumor mutation burden (TMB) of samples and visualize the top 15 mutated genes by frequency. Furthermore, the Wilcoxon rank-sum test was utilized to analyze the differences in TMB between the high and low PS groups.

2.10. Statistical analysis

All statistical analyses and data visualizations were conducted using R language (version 4.1.0), while Python language was employed for feature extraction from pathological images. To examine the differences between groups, we used Wilcoxon's rank-sum test for continuous variables, and the Chi-square test or Fisher's exact test for categorical variables. Survival analysis was conducted using Kaplan-Meier survival curves and tested for significance using the log-rank test. All statistical tests were two-tailed, and a P-value <0.05 was deemed statistically significant.

3. Result

3.1. Prognostic value of mRNAsi in LUAD

We integrated clinical data and transcriptome profiles from 327 LUAD patients retrieved from the TCGA database. Patients were stratified into high mRNAsi group (n = 120) and low mRNAsi group (n = 207) based on mRNAsi with a cutoff value of 0.5453. Among them, there were statistically significant differences in the distribution of gender (P = 0.007) and smoking status (P = 0.002) between the high and low mRNAsi groups (Table 1). The median survival time for the low mRNAsi group was 50.93 months, while for the high mRNAsi group, it was 40.5 months. Kaplan-Meier curves revealed an association between high mRNAsi and worse overall survival (OS) in LUAD patients (Supplementary Fig. 1). In univariate cox analysis, high mRNAsi was identified as a risk factor for OS (HR = 1.434, 95 % CI 0.992–2.075, P = 0.055) (Fig. 3A). In multivariate cox analysis, high mRNAsi (HR = 1.488, 95 % CI 1.012–2.187, P = 0.043) was identified as an independent risk factor for OS (Fig. 3B). Furthermore, according to the results of the subgroup analysis, there was no interaction between mRNAsi and the included covariates regarding their effects on OS (Supplementary Fig. 2 and Supplementary Table 1).

3.2. Pathomics model

To develop a predictive pathomics model for mRNAsi in LUAD, we randomly partitioned the TCGA-LUAD dataset into a training set (n = 229) and an independent validation set (n = 98) with a 7:3 ratio, allowing for robust evaluation of model performance. Notwithstanding minor variations in tumor location (P = 0.014) and pathologic stage (P = 0.033), the baseline characteristics of patients with LUAD in both the training and validation datasets were comparable, thereby ensuring robust group equivalence (Table 2). Subsequently, we processed and segmented pathological images from the training and validation sets using the PyRadiomics

Table 1
Demographic and clinical characteristics of patients in the high and low mRNAsi groups.

Characteristics		Total	Low	High	P value
		(N = 327)N(%)	(N = 207)N(%)	(N = 120)N(%)	
Gender	Female	183 (56)	128 (62)	55 (46)	0.007
	Male	144 (44)	79 (38)	65 (54)	
Age	≤65 years old	164 (50)	98 (47)	66 (55)	0.222
	>65 years old	163 (50)	109 (53)	54 (45)	
Smoking status	Current	82 (25)	40 (19)	42 (35)	0.002
	Former	204 (62)	135 (65)	69 (57)	
	Nonsmoker	41 (13)	32 (15)	9 (8)	
Tumor location	L-Lower	56 (17)	39 (19)	17 (14)	0.682
	L-Upper	76 (23)	47 (23)	29 (24)	
	R-Lower	63 (19)	41 (20)	22 (18)	
	R-Middle	14 (4)	7 (3)	7 (6)	
	R-Upper	118 (36)	73 (35)	45 (38)	
Pathologic stage	I/II	265 (81)	173 (84)	92 (77)	0.165
	III/IV	62 (19)	34 (16)	28 (23)	
Radiotherapy	No	293 (90)	189 (91)	104 (87)	0.256
	Yes	34 (10)	18 (9)	16 (13)	
Chemotherapy	No	219 (67)	141 (68)	78 (65)	0.649
	Yes	108 (33)	66 (32)	42 (35)	

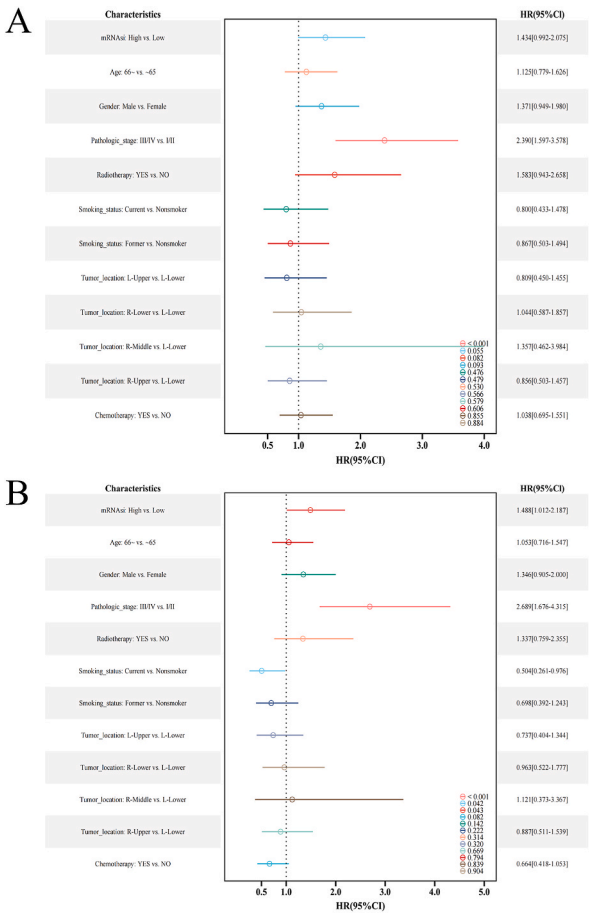

Fig. 3. Univariate cox analysis (A) and multivariate cox analysis (B) were conducted to explore whether various research factors, including mRNAsi, are independent risk factors for the overall survival (OS) of LUAD patients.

method, extracting a total of 465 pathomics features. To mitigate overfitting and optimize model performance, we employed mRMR and RFE algorithms to identify the most informative 8 features (Fig. 4A), which were deemed representative of the dataset. Additionally, the machine learning GBM algorithm was employed to construct the model based on these 8 pathomics features. Fig. 4B illustrates the ranking of the selected 8 features in terms of their importance in the GBM algorithm.

3.3. Evaluation of the model

According to the ROC curves of the model, the AUC values in the training set and validation set were 0.769 and 0.757, respectively (Fig. 5A–B). The threshold for the training set was 0.464, with an accuracy of 0.769, sensitivity of 0.524, specificity of 0.91, and Brier score of 0.185; for the validation set, the accuracy was 0.745, sensitivity was 0.75, specificity was 0.742, and Brier score was 0.198. These results indicate that the pathomics model for predicting mRNAsi has good discrimination and predictive performance. Calibration curves and the Hosmer-Lemeshow goodness-of-fit test also demonstrate good consistency between the model's predictions and the actual values of mRNAsi ($P > 0.05$) (Fig. 5C–D). The results of DCA analysis further confirm the model's good clinical utility (Fig. 5E–F).

3.4. Prognostic value of PS in LUAD

We employed our pathomics model to predict mRNAsi levels in LUAD samples, quantifying the results using the pathomics score (PS). Notably, PS exhibited significant disparities between high and low mRNAsi groups in both the training ($p < 0.001$) and validation sets ($p < 0.001$), as depicted in Fig. 6A–B. Using the R package "survMisc," a cutoff value of 0.4203 was set for PS, dividing LUAD patients into a high PS group ($n = 114$) and a low PS group ($n = 213$). Demographic and clinical characteristics were homogenous between high and low PS groups, with no significant differences observed ($p > 0.05$; Table 3). Median overall survival (OS) times were significantly different between patients with low (50.93 months) and high (39.9 months) PS groups ($P = 0.038$; Fig. 6C). High PS was associated considerably with worsened OS in LUAD patients.

Table 2
Demographic and clinical characteristics of patients.

Characteristics		Total	Train	Validation	P value
		(N = 327)N(%)	(N = 229)N(%)	(N = 98)N(%)	
mRNAsi	Low	207 (63)	145 (63)	62 (63)	1
	High	120 (37)	84 (37)	36 (37)	
OS	Alive	213 (65)	151 (66)	62 (63)	0.735
	Dead	114 (35)	78 (34)	36 (37)	
	OS.time, Median (Month) (Q1,Q3)	21.73 (14.48, 35.1)	21.73 (14.47, 36.03)	21.4 (14.54, 33.11)	0.545
Gender	Female	183 (56)	129 (56)	54 (55)	0.933
	Male	144 (44)	100 (44)	44 (45)	
Age	≤65 years old	164 (50)	122 (53)	42 (43)	0.108
	>65 years old	163 (50)	107 (47)	56 (57)	
Smoking status	Current	82 (25)	57 (25)	25 (26)	0.951
	Former	204 (62)	144 (63)	60 (61)	
	Nonsmoker	41 (13)	28 (12)	13 (13)	
Tumor location	L-Lower	56 (17)	41 (18)	15 (15)	0.014
	L-Upper	76 (23)	52 (23)	24 (24)	
	R-Lower	63 (19)	49 (21)	14 (14)	
	R-Middle	14 (4)	4 (2)	10 (10)	
	R-Upper	118 (36)	83 (36)	35 (36)	
Pathologic stage	I/II	265 (81)	193 (84)	72 (73)	0.033
	III/IV	62 (19)	36 (16)	26 (27)	
Radiotherapy	No	293 (90)	202 (88)	91 (93)	0.287
	Yes	34 (10)	27 (12)	7 (7)	
Chemotherapy	No	219 (67)	152 (66)	67 (68)	0.824
	Yes	108 (33)	77 (34)	31 (32)	

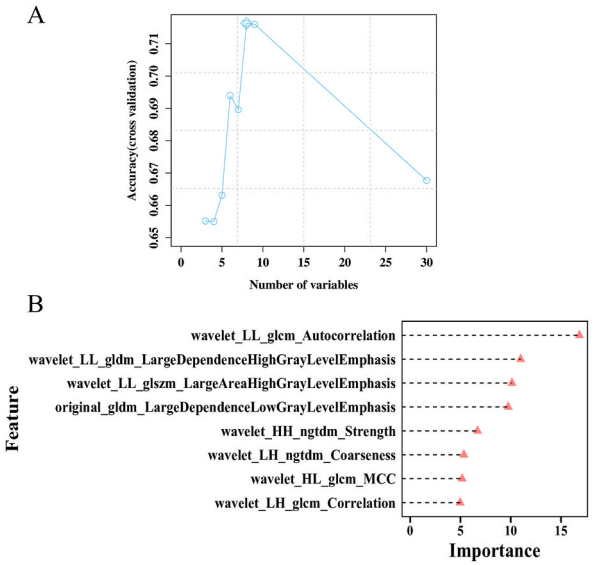

Fig. 4. Model construction: (A) The most representative eight pathomics features were selected through the joint application of the mRMR algorithm and the RFE algorithm. (B) The importance of the selected eight features in the machine learning GBM algorithm is demonstrated.

Furthermore, high PS was identified as an independent risk factor of OS in LUAD patients, with hazard ratios (HRs) of 1.479 (95 % CI 1.019–2.147, P = 0.04) in univariate cox analysis and 1.496 (95 % CI 1.008–2.222, P = 0.046) in multivariate cox analysis (Fig. 7A–B). Additionally, in subgroup analysis, it was shown that there was an interaction between PS and tumor location (p = 0.024). Analysis of the L-lower subgroup revealed a significant association between PS and OS (HR = 3.191, 95 % CI 1.304–7.804, P = 0.011), whereas, in the L-upper subgroup, no significant correlation was observed (HR = 0.814, 95 % CI 0.371–1.784, P = 0.61) (Supplementary Fig. 3 and Supplementary Table 2).

3.5. Pathway enrichment analysis

Comparative pathway enrichment analysis between high and low PS groups in LUAD identified distinct biological processes. Specifically, the high PS group exhibited significant enrichment in pathways related to RNA polymerase and mismatch repair, whereas

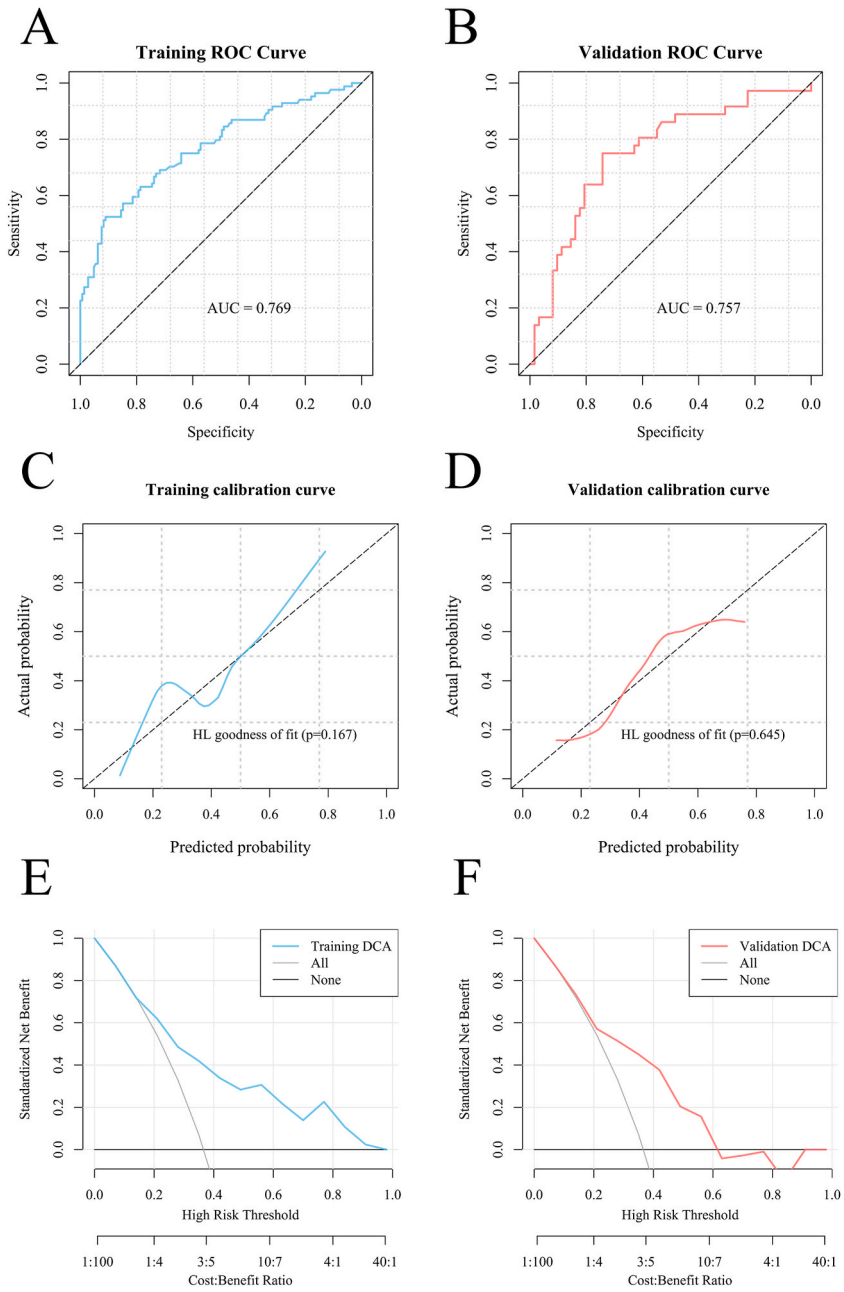

Fig. 5. Model evaluation: ROC curves (A, B), calibration curves (C, D) and DCA analyses (E, F) of the pathomics model in the training and validation sets. The x-axis of the ROC curve represents the false positive rate (1-Specificity), and the y-axis represents the true positive rate (Sensitivity).

the low PS group was enriched in pathways involved in alpha-linolenic acid metabolism and VEGF signaling (Fig. 8A). Functional analysis of the Hallmark gene set revealed that the high PS group exhibited significant enrichment in the MYC targets V2 pathway, whereas the low PS group was enriched in cholesterol homeostasis and angiogenesis pathways (Fig. 8B).

3.6. Immunological correlation analysis

Analysis of immune-related gene expression revealed significant upregulation of LAG3, TNFSF4, and CD80 in the high PS group, as well as CD40LG, HHLA2, and TNFRSF14 in the low PS group ($P < 0.05$; Fig. 9A). Immune cell infiltration analysis revealed significant differences between high and low PS groups. Specifically, T cells CD4 memory activated, NK cells activated, and Macrophages M1 exhibited higher infiltration in high PS, whereas B cell memory, T cell regulatory, and Mast cells resting showed increased infiltration in low PS ($P < 0.001$; Fig. 9B).

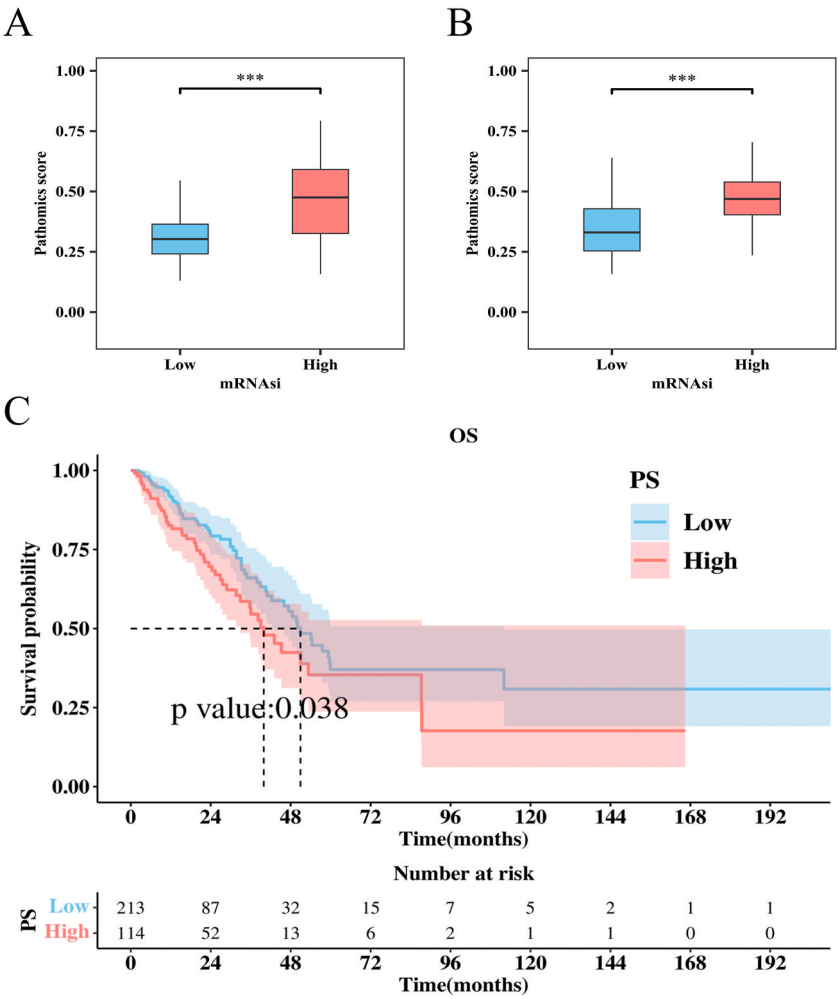

Fig. 6. The pathomics scores in the training set (A) and the validation set (B) exhibit significant differences in distribution between high and low mRNAsi groups. (C) Kaplan-Meier curves illustrate the impact of high and low PS on the prognosis of LUAD patients. (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

To assess the potential efficacy of immunotherapy, we used the TIDE software to evaluate the response of high and low PS groups. An increase in the TIDE prediction score indicates increased immune evasion, suggesting decreased responsiveness to immune therapeutic interventions. Notably, no significant intergroup differences were observed in TIDE, MSI Expr Sig, and Exclusion in the high PS group compared to the low PS group. However, a significant decrease in Dysfunction was observed in the high PS group (Fig. 10A, $P < 0.001$). This suggests that patients in the high PS group may derive greater benefit from immunotherapy.

3.7. Tumor mutational burden (TMB) and mutation analysis

The analysis results indicate a significant difference in TMB between the high and low PS groups ($P < 0.001$), with higher TMB values observed in the high PS group (Fig. 10B). Among them, Missense Mutation is the most common mutation type, followed by Nonsense Mutation and Frame Shift Deletion. Additionally, the mutation rates of TP53 and TTN genes are above 50 % in both the high and low PS groups. To compare mutant genes, we listed the top fifteen mutant genes in each of the two groups (Fig. 10C–D). Significant differences in mutation patterns and frequencies between these genes were found between the two groups, revealing potentially distinct molecular signatures.

4. Discussion

CSCs are believed to play a crucial role in the occurrence, progression, recurrence, and drug resistance of solid malignant tumors [26,27]. The heterogeneity and plasticity of CSCs contribute to tumor immune evasion and treatment resistance [28]. mRNAsi is commonly used to assess and quantify stemness characteristics of tumor cells [11]. In this study, through multivariate cox regression

Table 3
Demographic and clinical characteristics of patients in the high and low PS groups.

Characteristics		Total	Low	High	P value
		(N = 327)N(%)	(N = 213)N(%)	(N = 114)N(%)	
Gender	Female	183 (56)	127 (60)	56 (49)	0.088
	Male	144 (44)	86 (40)	58 (51)	
Age	≤65 years old	164 (50)	106 (50)	58 (51)	0.94
	>65 years old	163 (50)	107 (50)	56 (49)	
Smoking status	Current	82 (25)	46 (22)	36 (32)	0.14
	Former	204 (62)	139 (65)	65 (57)	
	Nonsmoker	41 (13)	28 (13)	13 (11)	
Tumor location	L-Lower	56 (17)	40 (19)	16 (14)	0.107
	L-Upper	76 (23)	51 (24)	25 (22)	
	R-Lower	63 (19)	47 (22)	16 (14)	
	R-Middle	14 (4)	8 (4)	6 (5)	
	R-Upper	118 (36)	67 (31)	51 (45)	
Pathologic stage	I/II	265 (81)	179 (84)	86 (75)	0.081
	III/IV	62 (19)	34 (16)	28 (25)	
Radiotherapy	No	293 (90)	192 (90)	101 (89)	0.806
	Yes	34 (10)	21 (10)	13 (11)	
Chemotherapy	No	219 (67)	141 (66)	78 (68)	0.776
	Yes	108 (33)	72 (34)	36 (32)	

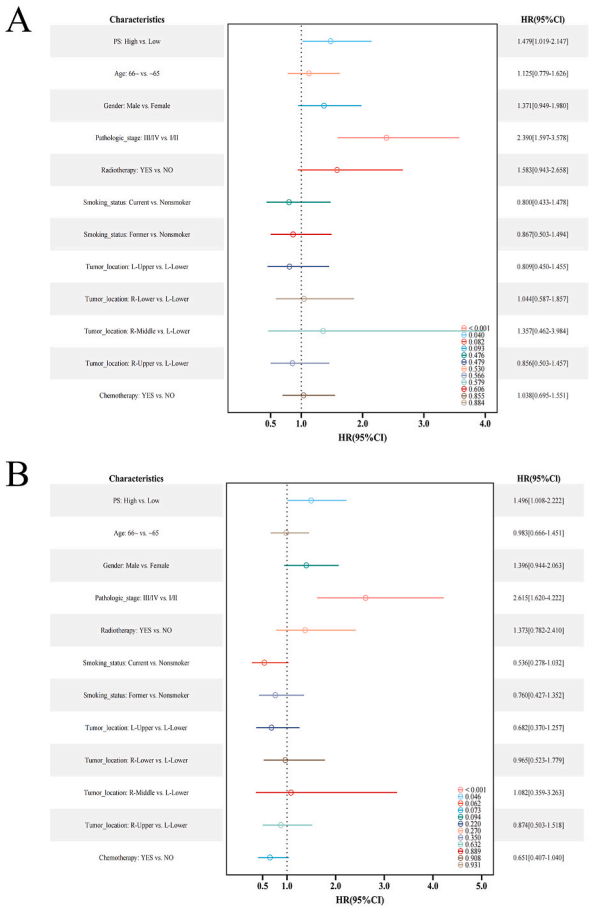

Fig. 7. Univariate cox analyses (A) and multivariate cox analyses (B) explored whether multiple study factors, including pathomics score, were independent risk factors for overall survival (OS) in patients with LUAD.

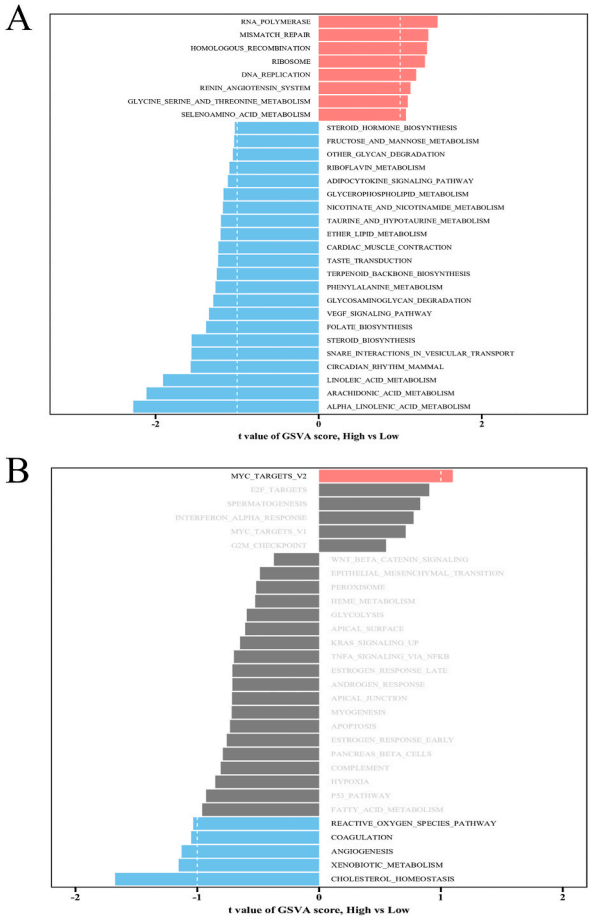

Fig. 8. Differences in pathway enrichment in KEGG pathway gene sets (A) and Hallmark gene sets (B) between high and low PS groups in LUAD.

analysis, we identified mRNAsi as an independent prognostic factor for LUAD, suggesting its potential as a prognostic indicator for LUAD patients. Leveraging the rapid development of machine learning in the medical field, this study aimed to establish a relationship with mRNAsi by extracting pathomics information from LUAD pathological slides and creating a prognostic prediction model. The results of the model's ROC curve, calibration curve, and DCA analysis demonstrate that the model we constructed has good predictive performance and clinical utility. Furthermore, according to enrichment analysis results, the high PS group showed enrichment in DNA replication and repair, as well as RNA polymerase pathways, indicating a potentially stronger proliferative ability and drug resistance in this group, which is consistent with its higher stemness characteristics. Additionally, the results of the immune-related analysis suggest that CSCs may be involved in regulating the immune microenvironment of LUAD, thereby affecting its prognosis, and patients in the high PS group may benefit more from immunotherapy. Through the constructed pathomics model, we can identify LUAD populations with higher mRNAsi, recognize patients with poorer prognoses, and tailor personalized treatment plans for them. Moreover, this approach holds promise for selecting the optimal beneficiary population for targeted CSC therapy.

In this study, multivariate cox regression analysis revealed that high mRNAsi is an independent risk factor for OS in LUAD patients. Previous studies have consistently reported a positive correlation between mRNAsi and prognosis in LUAD patients, with higher mRNAsi levels associated with poorer OS and more advanced disease stages [29]. According to the cancer stem cell theory, CSCs exhibit stem-like characteristics, promoting tumor initiation and metastasis through self-renewal and differentiation potential, and are less susceptible to anti-tumor therapies targeting primitive cells, which may contribute to treatment failure [30]. As a quantifiable indicator of tumor stemness characteristics, mRNAsi demonstrates a strong correlation with diverse tumor prognosis. Xu et al. noted that in hepatocellular carcinoma, higher mRNAsi levels are associated with higher tumor grades and poorer prognosis [31]. Similarly, Xia et al. found that in neuroblastoma, patients with high mRNAsi have a significantly worse prognosis than those with low mRNAsi [32]. Studies have also shown that in colorectal cancer and bladder cancer, patients with high mRNAsi not only have poorer prognoses but may also exhibit higher resistance to commonly used chemotherapy drugs [33,34]. These findings underscore the robust potential of mRNAsi as a prognostic predictor.

Currently, most existing pathomics models focus on predicting survival outcomes in lung cancer [35]. However, there is a lack of pathomics models specifically designed to predict mRNAsi in LUAD samples. Considering the importance of histopathological images in cancer diagnosis and prognosis prediction, we propose a workflow for image processing and data analysis to extract specific features

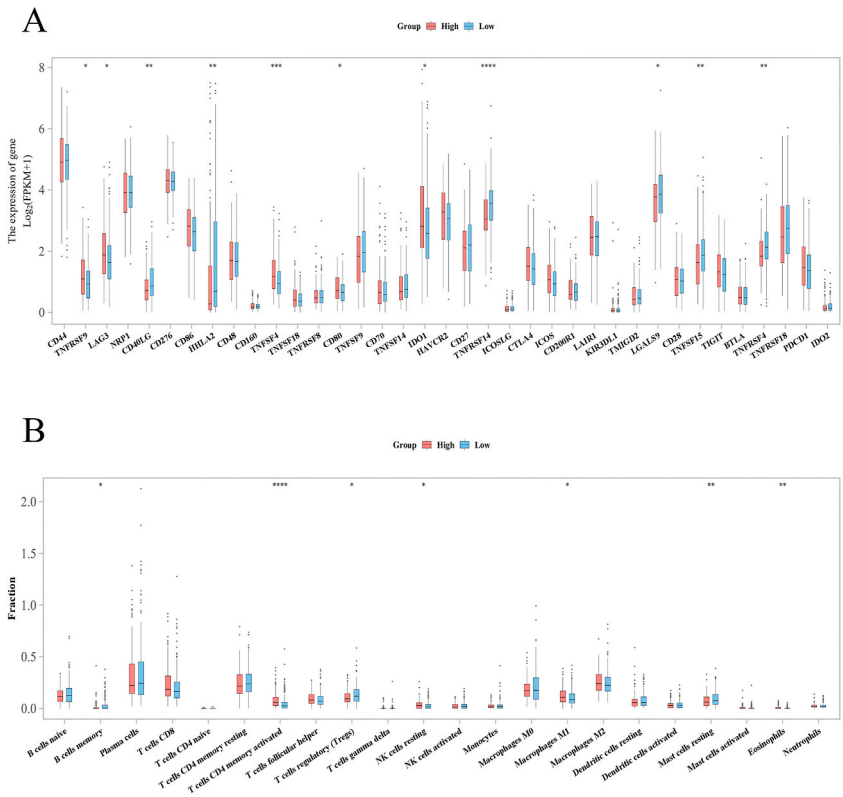

Fig. 9. (A) Differential analysis of immune-related genes between high and low PS groups; (B) Differential analysis of immune cell abundance between high and low PS groups. (*, P < 0.05; **, P < 0.01; ***, P < 0.001, ****, P < 0.0001).

from histopathological images. We utilize pathomics features and machine learning methods to develop a predictive model for mRNAsi and survival outcomes in LUAD patients. As indicated by the ROC curve of the model, the AUC value in the training set is 0.769, and in the validation set is 0.757. Combined with the calibration curve and DCA analysis results of the model, it can be concluded that the model we constructed has good clinical utility and predictive performance. Furthermore, the PS index calculated by the model shows significant differences in distribution between the high and low mRNAsi groups, with a higher PS index associated with poorer prognosis. These findings demonstrate the robustness of the model and its feasibility in prognosis assessment, providing new insights and methods for evaluating LUAD prognosis. Currently, various machine learning algorithms can be used to build reliable models based on histopathological slides to predict cancer prognosis, treatment response, gene mutations, and gene expression. For instance, Liao et al. developed a machine-learning model from histopathological slide image features to distinguish hepatocellular carcinoma (HCC) from adjacent normal tissue and predict the prognosis of HCC patients after surgical resection [36]. Machine learning algorithms based on histopathological slides of renal clear cell carcinoma can construct reliable diagnostic and prognostic prediction models [37]. Chen et al. combined histopathological images and gene expression patterns to analyze and predict the molecular characteristics and prognosis of head and neck squamous cell carcinoma [38]. Effective models integrating histopathological images and multi-omics information can predict the molecular characteristics and survival prognosis of lung adenocarcinoma [39]. This indicates that the pathomics model can not only effectively predict disease prognosis but also predict the molecular biological information contained in these samples, which is expected to promote the continuous development of personalized medicine.

The use of artificial intelligence methods has enhanced our ability to extract quantitative information from histopathological images. The potential differences in molecular expression within tumors often manifest as changes in tissue structure and nuclear morphology [40]. It has been previously discovered that machine learning can predict mutations such as STK11, EGFR, KRAS, ALK, ROS1, and TP53 in LUAD from histopathological images [39,41]. This indicates that changes in histopathological image information may be macroscopic manifestations of microscopic components such as cellular and molecular differences. CSCs are embedded in tumor tissues, where they exert crucial roles in tumor growth, metastasis, and chemoresistance. Notably, variations in mRNAsi, a quantitative indicator of CSCs, may reflect the molecular pathological basis behind the differences in the histopathological image information. This study utilized PyRadiomics to extract wavelet features from images that can characterize spatial heterogeneity within tumor cell nuclei and provide information about cell-cell interactions in the tumor microenvironment. By using this information to construct models to predict mRNAsi and prognosis, it may be possible to improve the diagnosis of LUAD and assist in guiding its treatment strategies.

In this study, we conducted a deeper exploration of the molecular biological significance behind the PS calculated by the model.

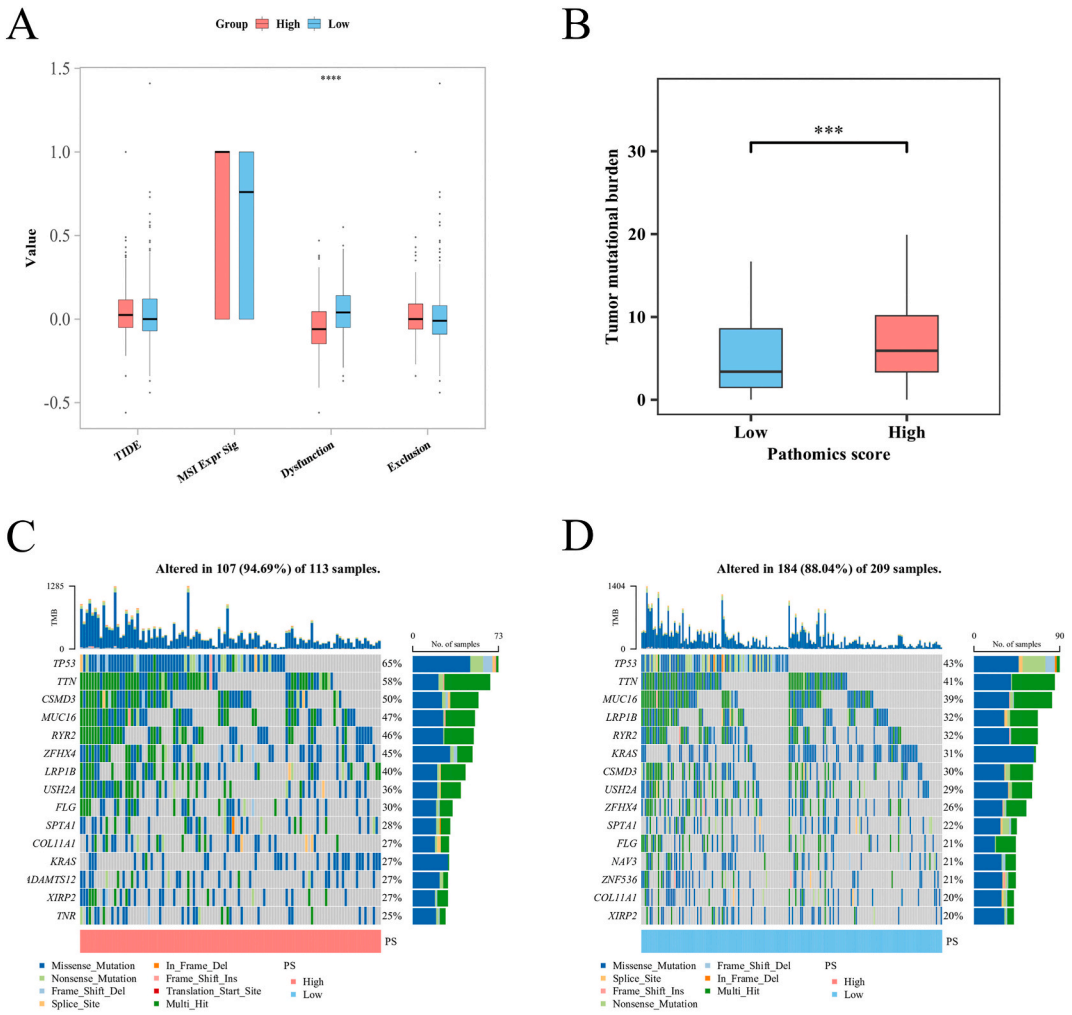

Fig. 10. (A) Tumor Immune Dysfunction and Exclusion (TIDE) algorithm to assess differential analysis of TIDE, MSI Expr Sig, Dysfunction and Exclusion between high and low PS groups. (B) Differential analysis of TMB between high and low PS groups in LUAD. Mutation analyses of the high PS group (C) and low PS group (D). (ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

According to the results of enrichment analysis, the high PS group showed significant enrichment in DNA replication and damage repair signaling pathways such as RNA polymerase, DNA replication, homologous recombination, and mismatch repair. This demonstrates the close connection between the stemness of LUAD and their proliferative capacity, which may induce tumors to develop in a more aggressive direction. Additionally, it has been reported that the main mechanisms of drug resistance in CSCs include high expression of drug transporters, strong DNA repair capability, and recruitment of a protective microenvironment [42]. CSCs exhibit a robust DNA damage repair response, akin to normal stem cells, which confers radiation and chemotherapy resistance [43,44]. The results of this study show that the high PS group is significantly enriched in pathways related to repairing DNA double-strand breaks, suggesting that the high PS group's stem cell-like features are more pronounced and may confer stronger treatment resistance.

Reportedly, CSCs engage in reciprocal interactions with tumor microenvironmental (TME) cells, fostering crosstalk that confers immune evasion and promotes their potential for recurrence [45,46]. Additionally, the TME is constantly exposed to complex nutritional, metabolic, and hypoxic environments, which can exacerbate CSCs' treatment resistance [47,48]. This study reveals significant differences in immune-related gene expression and immune cell infiltration between high- and low-PS groups, indicating that CSCs may modulate the complex immune microenvironment of LUAD to influence patient prognosis. Immunotherapy has emerged as a promising treatment approach for lung cancer, with various modalities including adoptive cell therapy, immune checkpoint inhibitors, and targeted antibodies [49,50]. Notably, numerous clinical trials have been conducted to investigate immunotherapeutic strategies for lung cancer, with several ongoing Phase III trials (e.g., NCT04489862, NCT04256421, NCT04738487) demonstrating encouraging results. Previous studies have identified mRNAsi as a crucial factor influencing tumor recurrence and immunotherapy response in LUAD [51]. Additionally, TMB has been demonstrated to predict the benefits of immune checkpoint inhibitors in non-small cell lung cancer, with the KEYNOTE 158 trial further supporting the efficacy of pembrolizumab in patients with high-TMB tumors (TMB≥10 mutations per megabase) [52]. Remarkably, our findings reveal a significantly higher TMB in the high PS group, suggesting that PS

may be a predictive marker for immunotherapy response in LUAD patients. However, the conventional detection of mRNAsi is often costly and based on local tumor tissue, which cannot accurately represent the overall tumor condition and is difficult to monitor dynamically. This study aims to establish a pathology-based predictive model by extracting information from patient tumor tissue sections, enabling objective, high-throughput, and accurate prediction of mRNAsi. Ultimately, this will facilitate the identification of potential responders to immunotherapy and provide a foundation for personalized treatment guidance.

Overall, LUAD with high mRNAsi may have poorer prognostic outcomes and are more prone to metastasis and drug resistance, which may inspire clinicians to consider adjusting follow-up frequencies and treatment strategies for these patients to maximize their benefits. Additionally, there are numerous methods for targeting CSCs, including immunotherapy, hormone therapy, (mi)siRNA delivery, gene knockout, as well as vaccines, antibodies, and CAR-T (chimeric antigen receptor T cell) cells, which can be utilized to suppress CSCs' stemness, inhibit CSCs' exogenous or endogenous signaling pathways, hypoxia, or promote cell differentiation [53–55]. Due to the common characteristics shared between CSCs and normal stem cells, targeting CSCs must be based on their unique antigens and markers [19]. Using a machine learning-based predictive model, we leveraged clinically accessible pathological tissue images to identify mRNAsi in LUAD samples, thereby prospectively identifying high-risk populations amenable to targeted interventions. Through the combined use of targeted CSC therapy measures, it is possible to develop more optimized treatment plans for patients.

However, there are some limitations to the current study. Firstly, although all eligible samples were included in the analysis, the exclusion of TCGA samples with missing data or suboptimal histopathological images may have introduced potential bias. Furthermore, the limited number of covariates included in this study may have omitted confounding factors, potentially influencing the study's conclusions. Moreover, we utilized data from LUAD patients in the TCGA database without an external validation set, which may affect the model's generalizability and practical clinical application, further validation in diverse cohorts is warranted. Finally, this study is a retrospective analysis, and further large-scale studies or experiments are needed to validate the molecular mechanisms underlying the pathomics features.

5. Conclusion

This study demonstrates the effectiveness of utilizing machine learning algorithms to construct a model for predicting mRNAsi based on LUAD pathological images. The model generates clinical impact by extracting hidden information from routinely available data, which is expected to assist pathologists and clinicians in assessing the prognosis of LUAD patients, stratifying patient risks, and making more optimized treatment decisions.

List of abbreviations

CSCs	cancer stem cells
AUC	Area under the receiver operating characteristic curve
DCA	decision curve analysis
ROC	Receiver operating characteristic curve
OS	Overall survival
GBM	Gradient Boosting Machine
mRMR	Maximum Relevance, Minimum Redundancy
RFE	algorithm and Recursive Feature Elimination
PS	pathomics score
TIDE	Tumor Immune Dysfunction and Exclusion
MSI Expr Sig	Microsatellite Instability Expression Signature
TMB	Tumor Mutational Burden

Data availability statement

The study involved the analysis of publicly accessible datasets, which can be accessed at (https://portal.gdc.cancer.gov/). Additionally, more detailed data utilized in the current study can be obtained from the corresponding author upon a reasonable request.

Funding

The present study was funded by the National Natural Science Foundation of China (grant nos. 81960425 and 81160294).

Ethics statement

The study does not require ethical approval, as TCGA participants had previously provided informed consent and obtained ethical approval in the original research.

CRediT authorship contribution statement

Rui Chen: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Formal analysis,

Data curation, Conceptualization. **Yuzhen Liu:** Writing – review & editing, Methodology, Formal analysis. **Junping Xie:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e37100.

References

[1] A. Leiter, R. Veluswamy, J. Wisnivesky, The global burden of lung cancer: current status and future trends, Nat. Rev. Clin. Oncol. 20 (9) (2023) 624–639.
[2] R. Siegel, A. Giaquinto, A. Jemal, Cancer statistics, CA A Cancer J. Clin. 74 (1) (2024) 12–49, 2024.
[3] F. Hirsch, et al., Lung cancer: current therapies and new targeted treatments, Lancet (London, England) 389 (10066) (2017) 299–311.
[4] Y. Li, B. Yan, S. He, Advances and challenges in the treatment of lung cancer, Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 169 (2023) 115891.
[5] J. Huang, et al., Baseline serum tumor markers predict the survival of patients with advanced non-small cell lung cancer receiving first-line immunotherapy: a multicenter retrospective study, BMC Cancer 23 (1) (2023) 812.
[6] X. Tang, et al., Intratumoral and peritumoral CT-based radiomics strategy reveals distinct subtypes of non-small-cell lung cancer, J. Cancer Res. Clin. Oncol. 148 (9) (2022) 2247–2260.
[7] T. Reya, et al., Stem cells, cancer, and cancer stem cells, Nature 414 (6859) (2001) 105–111.
[8] T. Phan, P. Croucher, The dormant cancer cell life cycle, Nat. Rev. Cancer 20 (7) (2020) 398–411.
[9] V. Melocchi, et al., Aggressive early-stage lung adenocarcinoma is characterized by epithelial cell plasticity with acquirement of stem-like traits and immune evasion phenotype, Oncogene 40 (31) (2021) 4980–4991.
[10] T. Ye, et al., Cdh1 functions as an oncogene by inducing self-renewal of lung cancer stem-like cells via oncogenic pathways, Int. J. Biol. Sci. 16 (3) (2020) 447–459.
[11] T. Malta, et al., Machine learning identifies stemness features associated with oncogenic dedifferentiation, Cell 173 (2) (2018) 338–354.e15.
[12] H. Lian, et al., Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma, Mol. Oncol. 13 (10) (2019) 2227–2245.
[13] W. Jiang, N. Xie, C. Xu, Characterization of a prognostic model for lung squamous cell carcinoma based on eight stemness index-related genes, BMC Pulm. Med. 22 (1) (2022) 224.
[14] D. Chen, et al., Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients, Int. J. Biol. Sci. 18 (1) (2022) 360–373.
[15] Z. Wang, et al., Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients, Briefings Bioinf. 22 (5) (2021).
[16] C. Wang, et al., mRNAsi-related genes can effectively distinguish hepatocellular carcinoma into new molecular subtypes, Comput. Struct. Biotechnol. J. 20 (2022) 2928–2941.
[17] K. Liu, J. Hu, Classification of acute myeloid leukemia M1 and M2 subtypes using machine learning, Comput. Biol. Med. 147 (2022) 105741.
[18] M. Nishio, et al., Homology-based image processing for automatic classification of histopathological images of lung tissue, Cancers 13 (6) (2021).
[19] A. Shmatko, et al., Artificial intelligence in histopathology: enhancing cancer research and clinical oncology, Nature cancer 3 (9) (2022) 1026–1038.
[20] Y. Zheng, et al., Lung cancer stem cell markers as therapeutic targets: an update on signaling pathways and therapies, Front. Oncol. 12 (2022) 873994.
[21] W. Shao, et al., Characterizing the survival-associated interactions between tumor-infiltrating lymphocytes and tumors from pathological images and multi-omics data, IEEE Trans. Med. Imag. 42 (10) (2023) 3025–3035.
[22] H. Zeng, et al., Integration of histopathological images and multi-dimensional omics analyses predicts molecular features and prognosis in high-grade serous ovarian cancer, Gynecol. Oncol. 163 (1) (2021) 171–180.
[23] Z. Tang, et al., Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline, Nat. Commun. 10 (1) (2019) 2173.
[24] J.J.M. van Griethuysen, et al., Computational radiomics system to decode the radiographic phenotype, Cancer Res. 77 (21) (2017) e104–e107.
[25] J. Xie, et al., A necroptosis-related prognostic model of uveal melanoma was constructed by single-cell sequencing analysis and weighted Co-expression network analysis based on public databases, Front. Immunol. 13 (2022) 847624.
[26] L. Walcher, et al., Cancer stem cells-origins and biomarkers: perspectives for targeted personalized therapies, Front. Immunol. 11 (2020) 1280.
[27] M.F. Clarke, Clinical and therapeutic implications of cancer stem cells, N. Engl. J. Med. 380 (23) (2019) 2237–2245.
[28] Y. Li, et al., Exploring the dynamic interplay between cancer stem cells and the tumor microenvironment: implications for novel therapeutic strategies, J. Transl. Med. 21 (1) (2023) 686.
[29] Y. Yi, et al., The tumor stemness indice mRNAsi can act as molecular typing tool for lung adenocarcinoma, Biochem. Genet. 61 (6) (2023) 2401–2424.
[30] S.Y. Yi, M.Z. Wei, L. Zhao, Targeted immunotherapy to cancer stem cells: a novel strategy of anticancer immunotherapy, Crit. Rev. Oncol. Hematol. 196 (2024) 104313.
[31] Q. Xu, et al., Immunological value of prognostic signature based on cancer stem cell characteristics in hepatocellular carcinoma, Front. Cell Dev. Biol. 9 (2021) 710207.
[32] Y. Xia, et al., Development and validation of a novel stemness-related prognostic model for neuroblastoma using integrated machine learning and bioinformatics analyses, Transl. Pediatr. 13 (1) (2024) 91–109.
[33] S. Fu, et al., Development of a stemness-related prognostic index to provide therapeutic strategies for bladder cancer, npj Precis. Oncol. 8 (1) (2024) 14.
[34] M. Weng, et al., mRNAsi-related metabolic risk score model identifies poor prognosis, immunoevasive contexture, and low chemotherapy response in colorectal cancer patients through machine learning, Front. Immunol. 13 (2022) 950782.
[35] C. Lu, R. Shiradkar, Z. Liu, Integrating pathomics with radiomics and genomics for cancer prognosis: a brief review, Chin. J. Cancer Res. 33 (5) (2021) 563–573.
[36] H. Liao, et al., Classification and prognosis prediction from histopathological images of hepatocellular carcinoma by a fully automated pipeline based on machine learning, Ann. Surg Oncol. 27 (7) (2020) 2359–2369.
[37] S. Chen, et al., Clinical use of a machine learning histopathological image signature in diagnosis and survival prediction of clear cell renal cell carcinoma, Int. J. Cancer 148 (3) (2021) 780–790.

[38] L. Chen, et al., Histopathological image and gene expression pattern analysis for predicting molecular features and prognosis of head and neck squamous cell carcinoma, Cancer Med. 10 (13) (2021) 4615–4628.
[39] L. Chen, et al., Histopathological images and multi-omics integration predict molecular characteristics and survival in lung adenocarcinoma, Front. Cell Dev. Biol. 9 (2021) 720110.
[40] A. Madabhushi, G. Lee, Image analysis and machine learning in digital pathology: challenges and opportunities, Med. Image Anal. 33 (2016) 170–175.
[41] N. Coudray, et al., Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning, Nat. Med. 24 (10) (2018) 1559–1567.
[42] C. Blanpain, et al., DNA-damage response in tissue-specific and cancer stem cells, Cell Stem Cell 8 (1) (2011) 16–29.
[43] M. Maugeri-Saccà, M. Bartucci, R. De Maria, DNA damage repair pathways in cancer stem cells, Mol. Cancer Therapeut. 11 (8) (2012) 1627–1636.
[44] A. Schulz, et al., Cancer stem cells and radioresistance: DNA repair and beyond, Cancers 11 (6) (2019).
[45] A. Guha, et al., Cancer stem cell-immune cell crosstalk in breast tumor microenvironment: a determinant of therapeutic facet, Front. Immunol. 14 (2023) 1245421.
[46] K. Chen, et al., The metabolic flexibility of quiescent CSC: implications for chemotherapy resistance, Cell Death Dis. 12 (9) (2021) 835.
[47] M. Najafi, K. Mortezaee, J. Majidpoor, Cancer stem cell (CSC) resistance drivers, Life Sci. 234 (2019) 116781.
[48] J. Behnan, et al., Differential propagation of stroma and cancer stem cells dictates tumorigenesis and multipotency, Oncogene 36 (4) (2017) 570–584.
[49] H. Mamdani, et al., Immunotherapy in lung cancer: current landscape and future directions, Front. Immunol. 13 (2022) 823618.
[50] A. Lahiri, et al., Lung cancer immunotherapy: progress, pitfalls, and promises, Mol. Cancer 22 (1) (2023) 40.
[51] H. Shi, et al., Tumor stemness and immune infiltration synergistically predict response of radiotherapy or immunotherapy and relapse in lung adenocarcinoma, Cancer Med. 10 (24) (2021) 8944–8960.
[52] A. Marabelle, et al., Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study, Lancet Oncol. 21 (10) (2020) 1353–1365.
[53] M. Najafi, B. Farhood, K. Mortezaee, Cancer stem cells (CSCs) in cancer progression and therapy, J. Cell. Physiol. 234 (6) (2019) 8381–8395.
[54] L. Yang, et al., Targeting cancer stem cell pathways for cancer therapy, Signal Transduct. Targeted Ther. 5 (1) (2020) 8.
[55] M. Liu, H. Wu, C. Xu, Targeting cancer stem cell pathways for lung cancer therapy, Curr. Opin. Oncol. 35 (1) (2023) 78–85.