Original Article

Predicting prognosis in lung adenocarcinoma by predicting TIGIT expression: a pathomics model

Peihong Hu[1#], Bo Tian[1#], Hang Gu[1,2], Haoran Liu[1,3], Qiang Li[1]

[1]Department of Thoracic Surgery, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; [2]Graduate School, Chengdu Medical College, Chengdu, China; [3]Graduate School, North Sichuan Medical College, Nanchong, China

Contributions: (I) Conception and design: P Hu; (II) Administrative support: Q Li; (III) Provision of study materials or patients: B Tian; (IV) Collection and assembly of data: H Gu, H Liu; (V) Data analysis and interpretation: P Hu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

Correspondence to: Qiang Li, BD. Department of Thoracic Surgery, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, No. 55, Section 4, South Renmin Road, Chengdu 610041, China. Email: liqiang@sichuancancer.org.

Background: Traditional diagnostic methods have limited efficacy in predicting the prognosis of lung adenocarcinoma (LUAD), T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain (TIGIT) is a new biomarker. This study aimed to evaluate TIGIT expression as a LUAD biomarker and predict patient prognosis using a pathological feature model.

Methods: Clinical data and pathological images from The Cancer Genome Atlas (TCGA) were analyzed. The prognostic value of TIGIT was verified by genetic prognostic analysis and gene set enrichment analysis (GSEA). The OTSU algorithm was used to segment LUAD pathological images, and features were extracted using the PyRadiomics package and standardized with z-scores. Feature selection was performed using min-redundancy, recursive feature elimination (RFE) and stepwise regression algorithms, and a logistic regression algorithm was used to establish the pathomics model. Receiver operating characteristics, calibration, and decision curves were used for model evaluation. The pathomics score (PS) was used to predict TIGIT gene expression and analyze prognostic value and pathological mechanisms through Spearman correlation.

Results: The study included 443 clinical samples and 327 pathological images. Prognostic analysis showed significantly higher TIGIT expression in tumor tissues (P<0.001), with TIGIT being a protective factor for overall survival (OS) in LUAD [hazard ratio (HR) =0.65; 95% confidence interval (CI): 0.44–0.95; P=0.03]. GSEA revealed significant enrichment of differentially expressed genes in the TGF-β and MAPK signaling pathways. From 465 pathological features, the four best features were selected to construct a pathomics model with good predictive performance. Higher PS values were observed in the TIGIT high-expression group, correlating with improved OS (P=0.009). PS was positively correlated with the epithelial-mesenchymal transition related (EMT-related) genes (*WIPF1*, *GLIPR1*, *IL15*) and immune checkpoints (*ICOS*, *CTLA4*, *LAG3*) (P<0.001). Increased abundance of G2/M checkpoint-related genes (*MARCKS*, *CASP8AP2*) and infiltration of $CD8^+$ T cells and M2 macrophages were noted in the high PS group (P<0.05).

Conclusions: TIGIT expression is significantly correlated with LUAD prognosis and can effectively predict patient outcomes.

Keywords: T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain (TIGIT); lung adenocarcinoma (LUAD); pathomics; prognostic model; mechanisms of pathology

Submitted Jun 17, 2024. Accepted for publication Sep 20, 2024. Published online Nov 29, 2024. This article was updated on Sep 25, 2025.

The original version was available at: https://dx.doi.org/10.21037/jtd-24-978

doi: 10.21037/jtd-24-978

Introduction

Lung adenocarcinoma (LUAD), a type of non-small cell carcinoma (1), is typically managed with surgical resection. Nevertheless, for advanced-stage patients, the prognosis is dismal, with only a 15% 5-year survival rate (2). Conventional prognostic indicators for LUAD include clinicopathological features, laboratory diagnostic indicators, such as carcinoembryonic antigen and carbohydrate antigen 125, and computed tomography (CT) imaging. Due to limited specificity and subjective interpretation, they cannot meet the precision medicine demands of today's clinical practice. Researchers are continuously exploring new prognostic markers, such as *ALG3* (3), *PZP* (4), and *TBL2* (5), aiming to refine patient prognosis stratification for personalized and precise treatment.

T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif (ITIM) domains (TIGIT) is a co-inhibitory receptor found on the surface of lymphocytes (6), including $CD4^+$ T and $CD8^+$ T cells, natural killer (NK) cells, regulatory T cells (Tregs), and follicular helper T cells (Tfh). TIGIT's overexpression on these cells allows tumors to evade the immune system by inducing and regulating immune cell activity, thus hindering antitumor immune responses at various stages of the tumor immune cycle (7). Researchers have demonstrated the potent inhibitory effect of TIGIT on tumor growth in knockout mouse models (8). TIGIT can also be used as a predictive biomarker and has been employed in efficacy detection for colorectal cancer (9) and gastric cancer (10).

At present, TIGIT expression can only be detected through peripheral blood cytokine assessment (real-time detection, expensive, lacks tumor-specific insights, and other defects) and messenger RNA (mRNA) [quantitative polymerase chain reaction (qPCR), RNA sequencing (RNA-seq)] or protein (Western blotting, flow cytometry, etc.) analysis in fresh tissue samples (challenging sample collection and operator- and antibody-dependent). Formalin fixation and paraffin embedding are commonly chosen for structural staining and immunohistochemistry. This method effectively preserves morphological details and biomarkers, offering a convenient and cost-effective approach to specimen preservation. Artificial intelligence has gradually been applied to pathology, introducing significant changes in the field. Pathomics refers to the transformation of pathological images into high-fidelity, high-throughput, and mining data using artificial intelligence. It includes quantitative features, such as texture, morphology, edge gradients, and biological characteristics, and is used for quantitative pathological diagnosis, molecular expression, and disease prognosis (11-13). Hematoxylin and eosin stained (H&E-stained) sections are commonly used in clinical pathological diagnosis. They are easily obtainable, with mature staining technology, low cost, and no dependence on operators or antibodies, making them highly suitable for pathological research.

Based on these factors, this research presents a novel method for evaluating TIGIT expression in LUAD tissue through the utilization of pathomics technology. It concurrently incorporates bioinformatics analysis to investigate the potential molecular mechanisms underlying pathomics discoveries. We present this article in accordance with the TRIPOD reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-978/rc).

Highlight box

Key findings

- The T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain (TIGIT) expression-based pathomics model can effectively predict the prognosis of lung adenocarcinoma (LUAD) patients.

What is known and what is new?

- TIGIT can be used as a predictive biomarker and has been employed in efficacy detection for cancer.
- TIGIT is combined with pathomics to predict the prognosis of LUAD.

What is the implication, and what should change now?

- More clinical samples are needed to verify the results.

Methods

Data source

Data were sourced from The Cancer Genome Atlas (TCGA)-LUAD dataset within the TCGA database (https://portal.gdc.cancer.gov/). Following the exclusion of non-primary LUAD cases, instances with missing follow-up data, cases with survival times less than 30 days, absent clinical data, and samples without RNA-seq data, the study encompassed the remaining patients for survival analysis.

Pathological images were retrieved from the TCGA-LUAD dataset, and samples with unsatisfactory image quality were removed. Samples that did not align with clinical data and RNA-seq information were also excluded,

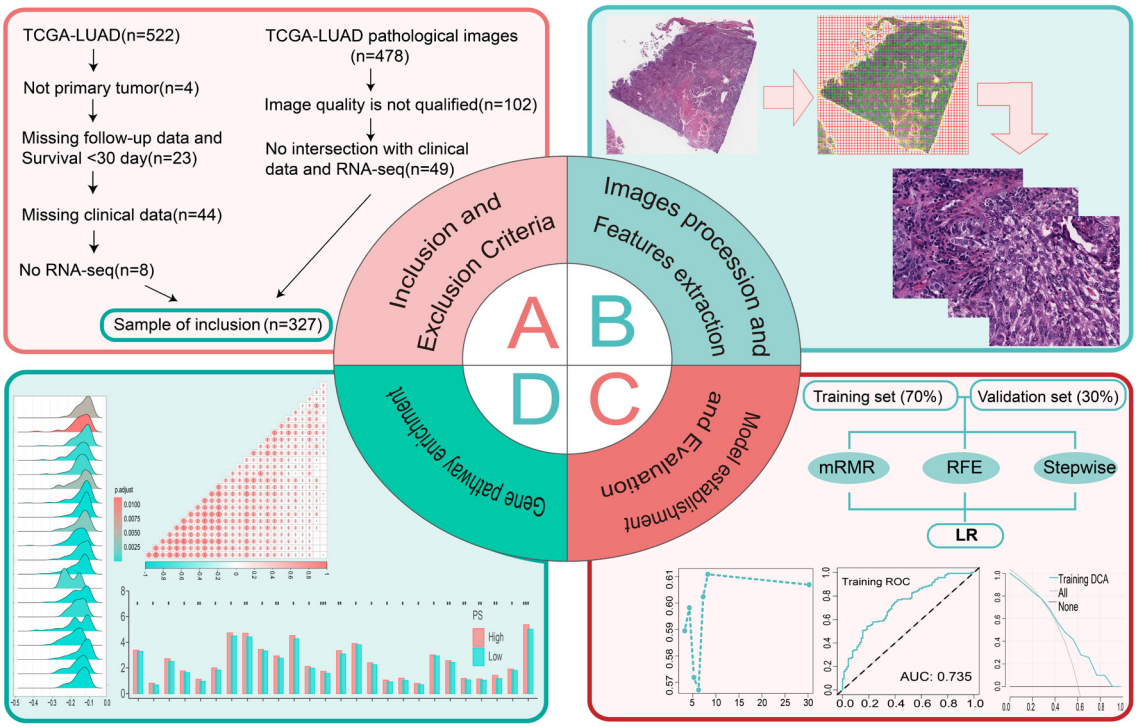

Figure 1 Data analysis workflow. (A) Inclusion criteria for the study. (B) Processing of the pathological images. (C) The process of establishing and evaluating the pathomics model. (D) The relevant mechanisms explored. TGCA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; mRMR, max-relevance and min-redundancy; RFE, recursive feature elimination; LR, logistic regression; ROC, receiver operating characteristic; AUC, area under the curve; DCA, decision curve analysis; PS, pathomics score.

resulting in the selection of final pathological image samples (refer to *Figure 1* for the screening process). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The data in this investigation were acquired from publicly available databases; therefore, the study was exempt from ethical review.

Analysis of prognostic value

RNA-seq data were downloaded and organized in transcript per million (TPM) format. The R package "ggplot2" was used for visualization. We utilized the "surminer" R package to determine the cutoff value for TIGIT gene expression levels, subsequently classifying patients into high-expression and low-expression groups. Variations in survival rates across the distinct groups were visualized using Kaplan-Meier survival curves. We conducted univariate and multivariate Cox regression analyses with the R packages "survival" and "forestplot". We performed an exploratory subgroup analysis utilizing univariate Cox regression to assess how TIGIT high expression and low expression affected patient prognosis in different subgroups for each covariate. To investigate the molecular mechanism underlying the expression disparities between the TIGIT high- and low-expression groups, we utilized the R package "clusterProfiler" to conduct gene set enrichment analysis (GSEA) on Kyoto Encyclopedia of Genes and Genomes (KEGG) (c2.cp.kegg.v7.5.1.symbols.gmt) and Hallmark (h.all.v7.5.1.symbols.gmt) gene sets, and the first 20 pathways were identified.

Screening of pathological features

We obtained H&E-stained tissue pathological images from the TCGA database (https://tcga-data.nci.nih.gov/tcga/) in ScanScope virtual slide (SVS) format, which were prepared by embedding in formalin and paraffin, with a maximum magnification of either ×20 or ×40. Pathological image processing and segmentation (14) were performed using the OTSU algorithm (15) (https://opencv.org/) to obtain

tissue areas from pathological sections. Subsequently, image segmentation was applied at both ×40 and ×20 to multiple subimages of 1,024×1,024 pixels and 512×512 pixels; ×20 images were upscaled to 1,024×1,024 pixels. The pathologist reviewed and excluded subimages with subpar image quality, including contamination, blurry images, or blank areas exceeding 50%. We randomly chose 10 subimages from each pathological image for subsequent analysis. Subsequently, we employed the PyRadiomics open-source package (https://pyradiomics.readthedocs.io/en/latest/) to standardize the subimages and extract features. Finally, the average value was considered the pathological characteristic of each sample for subsequent data analysis (16).

Construction of the pathomics model

The entire set of pathological images, gene matrices, and clinical information for the samples were randomly partitioned into training and validation sets, maintaining a ratio of 7:3. The pathological feature values from the training and validation sets were standardized using z-scores and the training set's mean and standard deviation, respectively. We then conducted an intergroup difference analysis of the clinical variables between the datasets. Features with zero variance and those exhibiting strong correlations (Pearson correlation coefficient values higher than 0.9) were removed using max-relevance and min-redundancy (mRMR), recursive feature elimination (RFE), and stepwise regression algorithms to filter out the optimal feature subset. A logistic regression (LR) algorithm was applied to fit the selected pathological features using the R "stats" package and establish a binary model for predicting gene expression.

Evaluation of the pathomics model

We assessed the calibration of pathological prediction models through the construction of calibration curves and the implementation of Hosmer-Lemeshow goodness-of-fit tests. The clinical applicability of the pathological model was established by decision curve analysis (DCA).

Clinical intersection samples and survival analysis

The pathomics score (PS) for the overlapping samples was computed using the pathomics model. PS was combined with clinical data, and a cutoff value for PS was determined, categorizing it into a binary variable (low/high). Survival analysis, Cox regression, and subgroup analysis were carried out using steps similar to those in section "*Analysis of prognostic value*".

Pathological mechanism analysis

We investigated the relationship between PS and epithelial-mesenchymal transition (EMT) gene expression using Spearman correlation analysis. We also conducted a difference analysis between the PS high and low groups and G2/M checkpoint gene expression using Wilcox. test (comes from the "*Analysis of prognostic value*" pathway enrichment results). Representative pathway-related genes were subjected to correlation analysis, and the LUAD sample gene expression matrix was uploaded to the CIBERSORTx database (https://cibersortx.stanford.edu/) to determine each sample's immune cell infiltration.

Statistical analysis

All clinical quantitative indicator results were expressed as mean and standard deviation. The differences between each indicator's results were compared using independent sample *t*-tests for normally distributed quantitative data, or the Mann-Whitney *U* test for non-normally distributed data. The chi-squared test was used for categorical variables. The log-rank test was used to determine the significance of the survival rate between groups. TIGIT expression and other factors were examined for interactions using the likelihood ratio test. By drawing the calibration curve and conducting the Hosmer-Lemeshow goodness-of-fit test, the forecast model was calibrated on the pathology evaluation set. To examine the variations between the high and low TIGIT and PS groups, the Wilcoxon test was employed. $P<0.05$ and a 95% confidence interval (CI) were used to determine statistical significance. The R program was used for all statistical studies (version 4.1.0).

Results

Baseline data

A survival analysis was conducted on a cohort of 327 LUAD patients from the TCGA database who were stratified into two groups: TIGIT high expression (n=197) and low expression (n=130), determined by applying a cutoff value of 1.0425 for TIGIT expression.

The patients' clinical information is presented in *Table 1*. No significant disparity was observed in the

Table 1 Features of The Cancer Genome Atlas cohort of patients

Variables	Total (n=327)	High (n=197)	Low (n=130)	P
Gender, n [%]				0.53
Female	183 [56]	107 [54]	76 [58]	
Male	144 [44]	90 [46]	54 [42]	
Age (years), n [%]				0.46
<66	164 [50]	95 [48]	69 [53]	
≥66	163 [50]	102 [52]	61 [47]	
Pathologic stage, n [%]				0.41
I/II	265 [81]	163 [83]	102 [78]	
III/IV	62 [19]	34 [17]	28 [22]	
Radiotherapy, n [%]				0.46
No	293 [90]	179 [91]	114 [88]	
Yes	34 [10]	18 [9]	16 [12]	
Chemotherapy, n [%]				0.71
No	219 [67]	134 [68]	85 [65]	
Yes	108 [33]	63 [32]	45 [35]	
Smoking status, n [%]				0.74
Nonsmoker	41 [13]	23 [12]	18 [14]	
Former	204 [62]	126 [64]	78 [60]	
Current	82 [25]	48 [24]	34 [26]	
Residual tumor, n [%]				0.24
R0	220 [67]	136 [69]	84 [65]	
R1/R2	13 [4]	5 [3]	8 [6]	
RX/unknown	94 [29]	56 [28]	38 [29]	
Histologic type, n [%]				0.48
NOS	204 [62]	128 [65]	76 [58]	
Others	56 [17]	32 [16]	24 [18]	
Mixed subtype	67 [20]	37 [19]	30 [23]	
Tumor location, n [%]				0.76
L-lower	56 [17]	34 [17]	22 [17]	
L-upper	76 [23]	47 [24]	29 [22]	
R-lower	63 [19]	38 [19]	25 [19]	
R-middle	14 [4]	6 [3]	8 [6]	
R-upper	118 [36]	72 [37]	46 [35]	

NOS, not otherwise specified; L, left; R, right.

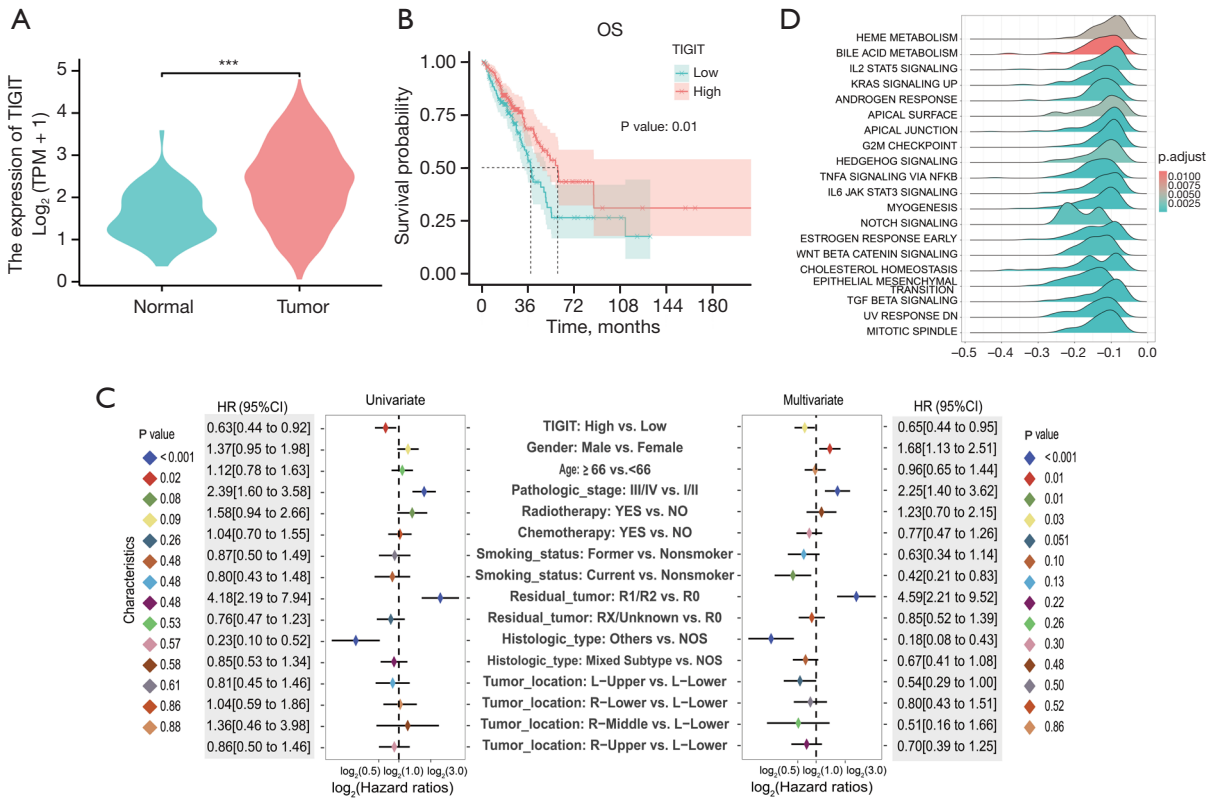

Figure 2 Relationship between TIGIT and clinical features. (A) Differential analysis of tumor versus healthy tissue. (B) Relationship between TIGIT expression and OS. (C) Univariate and multivariate Cox regression analysis. We performed log2 processing of HR. (D) GSEA in the Hallmark gene set. ***, P<0.001. TIGIT, T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain; TPM, transcript per million; OS, overall survival; HR, hazard ratio; CI, confidence interval; NOS, not otherwise specified; L, left; R, right; GSEA, gene set enrichment analysis.

distribution of clinical characteristics, including gender, age, and pathological stage, between the high and low TIGIT expression groups.

High TIGIT expression is a protective factor for overall survival (OS)

The analysis of intergroup differences between tumor and normal tissues showed that tumors expressed TIGIT at a considerably greater level (P<0.001) (*Figure 2A*). A strong relationship between increased OS and elevated TIGIT expression was highlighted by the Kaplan-Meier curve (P=0.01) (*Figure 2B*). In the univariate Cox analysis [hazard ratio (HR) =0.63; 95% CI: 0.44–0.92; P=0.02] and the multivariate analysis (HR =0.65; 95% CI: 0.44–0.95; P=0.03), TIGIT high expression was found to be a protective factor for OS. In order to better visualize all the characteristics of the data and avoid the influence of individual extreme values on the visualization, log2 was taken for HR in the process of drawing (*Figure 2C*). Exploratory subgroup analysis revealed that elevated TIGIT served as a significant protective factor against OS in subgroups younger than 66 years (HR =0.51; 95% CI: 0.30–0.86). The interaction test's P value of 0.23 shows that TIGIT's impact on OS was similar between the two age groups (Figure S1A). The top 20 pathways in the Hallmark gene set were identified by GSEA, demonstrating that signaling pathways such as the G2/M checkpoint and EMT were significantly enriched in differentially expressed genes in the low and high TIGIT groups (*Figure 2D*). Differential genes indicated considerable enrichment in the TGF-β and MAPK signaling pathways in the high and low TIGIT groups, and GSEA displayed the top 20 pathways in the KEGG gene collection (Figure S1B).

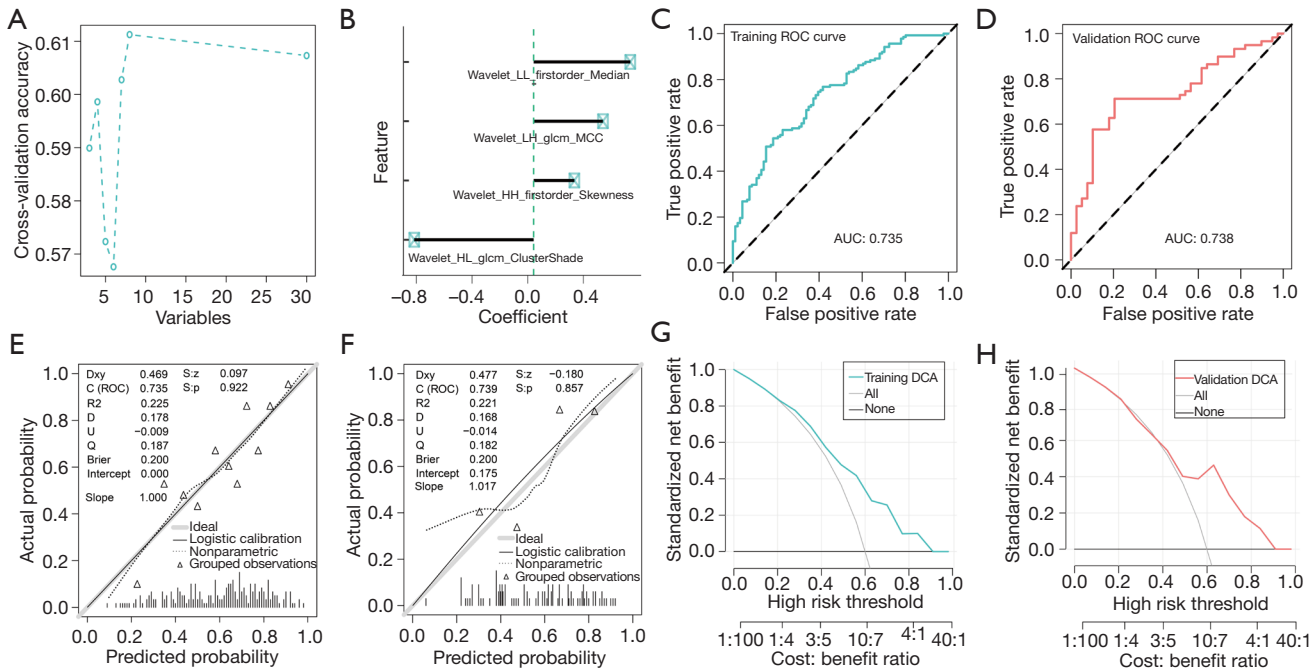

Figure 3 Pathological image feature selection, model establishment, and evaluation. (A) RFE feature screening for six characteristics. (B) The stepwise regression algorithm obtained the final four features. (C) ROC curve of the training set. (D) ROC curve of the validation set. (E) Calibration curve for the training set. (F) Calibration curve for the validation set. (G) DCA curve of the training set. (H) DCA curve of the validation set. glcm, gray-level co-occurrence matrix; MCC, maximal correlation coefficient; ROC, receiver operating characteristic; AUC, area under the curve; Dxy, the rank correlation between the predicted probability and the observed value; R2, Nagelkerke-Cox-Snell-Maddala-Magee R-squared index; D, discrimination index; U, unreliability index; Q, quality index; S:z, Z value of Spiegelhalter Z-test; S:p, P value of Spiegelhalter Z-test; DCA, decision curve analysis.

Screening of pathological features

A total of 465 features were extracted, encompassing 93 original features, including first-order and second-order features, along with higher-order features, such as wavelets [low-low (LL), low-high (LH), high-low (HL), high-high (HH)]. The P>0.05 differential analysis of each variable between groups demonstrated comparability between the groups (Table S1). The mRMR method was used to select the top 30 features, followed by RFE feature screening (*Figure 3A*), ultimately resulting in four features using a stepwise regression algorithm (*Figure 3B*).

Building and assessing the pathomics model

A model was built using the LR algorithm in the training set. *Table 2* lists the regression coefficients for the characteristics of the LR model. In the training and validation sets, the model's area under the curve (AUC) values were 0.735 (*Figure 3C*) and 0.738 (*Figure 3D*), respectively, as indicated by the receiver operating characteristic (ROC) curve. The Hosmer-Lemeshow goodness-of-fit test and calibration curve demonstrated that the pathological prediction model had a high degree of agreement between the actual value and the projected likelihood of gene overexpression (P>0.05) (*Figure 3E,3F*), and DCA confirmed the model's clinical applicability (*Figure 3G,3H*).

PS high expression is the protection factor in OS

In the training set, the PS distribution of the high and low TIGIT molecular groups differed significantly (P<0.001). The high TIGIT expression group exhibited elevated PS values, and this pattern aligned with the validation set's findings (*Figure 4A,4B*). The cutoff value of the predicted PS value from the LR model was 0.542, and patients were split into groups according to their PS expression levels: 197 in the high group and 130 in the low group (Table S2). Incorporating them into the survival analysis

Table 2 The coefficients of the features in the LR algorithm

Features	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.509848505	0.151270552	3.370441223	0.000750479
wavelet_HL_glcm_ClusterShade	−0.827517113	0.209353981	−3.952717343	7.73E−05
wavelet_HH_firstorder_Skewness	0.28227864	0.191334092	1.475318053	0.140127055
wavelet_LL_firstorder_Median	0.669757441	0.172515604	3.882300642	0.000103473
wavelet_LH_glcm_MCC	0.480485434	0.161624622	2.972847999	0.002950505

LR, logistic regression; Std. Error, standard error; Pr(>|z|), P values of the z-test were applied to the regression coefficients; glcm, gray-level co-occurrence matrix; MCC, maximal correlation coefficient.

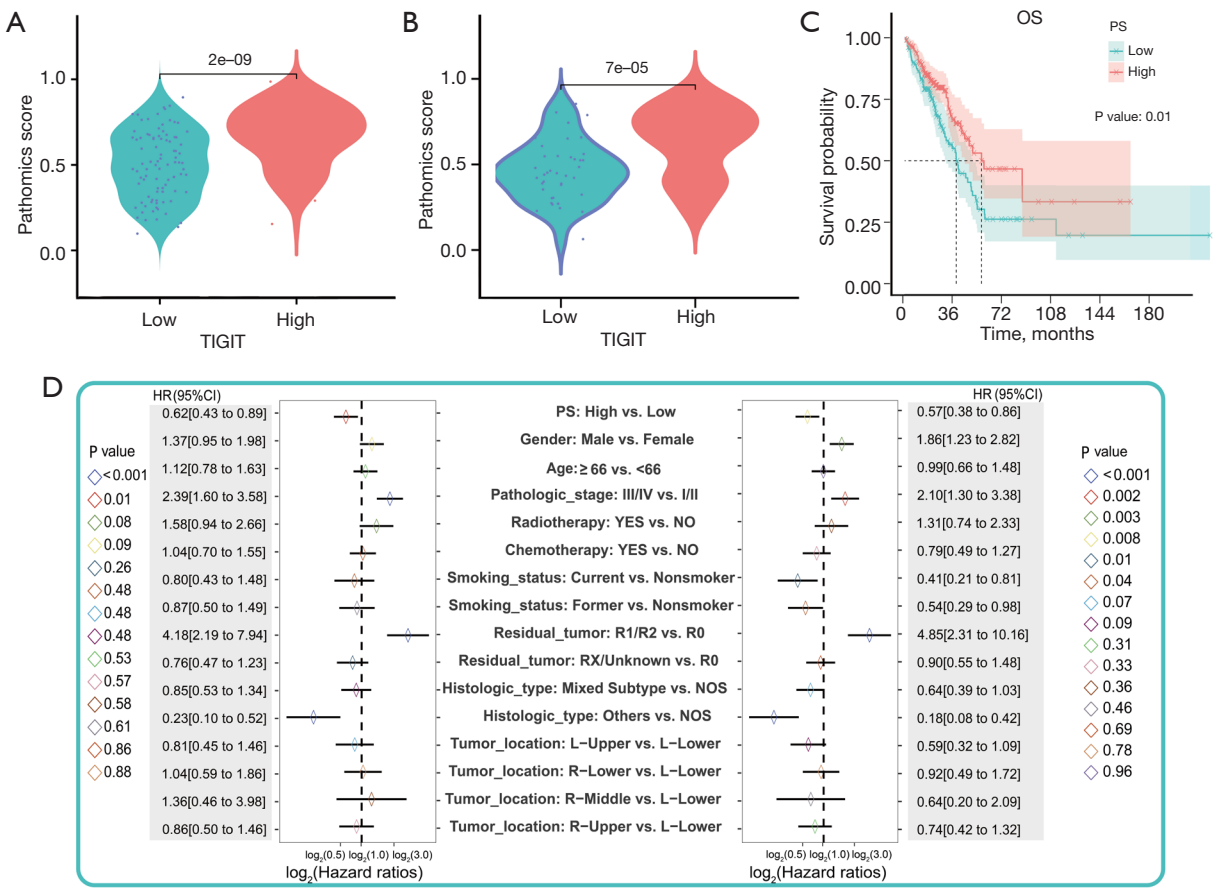

Figure 4 Clinical characteristics and PS correlation. (A) Difference analysis between tumor tissue and normal tissue in the training set. (B) Study of the variations between the validation set's normal and malignant tissues. (C) Relationship between PS high/low and OS. (D) Univariate and multivariate Cox regression analysis. OS, overall survival; PS, pathomics score; HR, hazard ratio; CI, confidence interval; NOS, not otherwise specified; L, left; R, right.

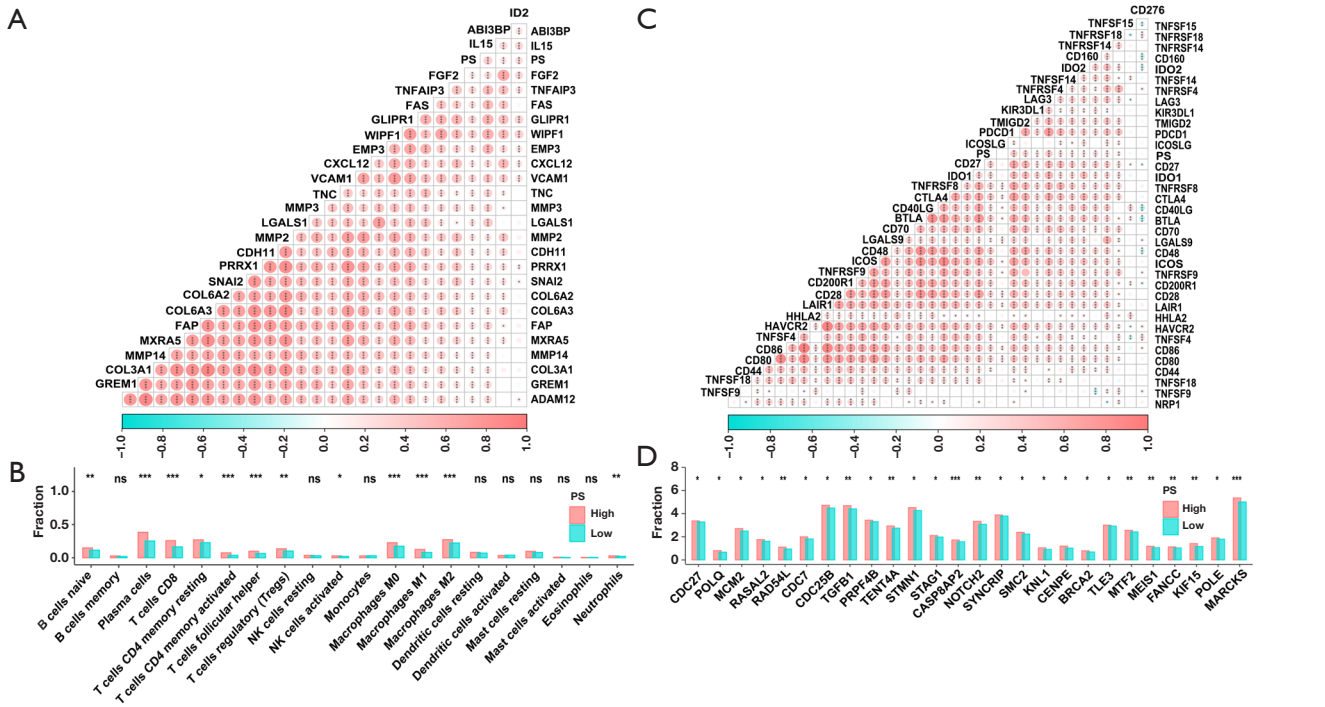

Figure 5 The correlation between PS and pathway gene analysis. (A) Investigation of the relationship between PS and EMT gene expression. (B) Immune cell abundance differences between groups with high and low PS. (C) Correlation analysis of PS high and low groups and immune checkpoint expression. (D) Difference analysis between the PS high and low groups and G2/M checkpoint. ns, not significant; *, P<0.05; **, P<0.01; ***, P<0.001. PS, pathomics score; EMT, epithelial-mesenchymal transition.

unveiled a strong correlation between increased PS and better OS (P=0.009) based on the Kaplan-Meier curve (*Figure 4C*). Elevated PS expression served as an OS preventive factor (*Figure 4D*). Subgroup analysis suggested that there was no significant interaction between PS and the different age subgroups (Figure S2).

Correlation analysis between model prediction results (PS) and pathway genes

A more thorough examination of the relationship between PS and the expression of EMT-related genes, namely *WIPF1*, *GLIPR1*, and *IL15*, showed a strong positive association (P<0.001) between PS and the expression of these genes (*Figure 5A*). Depending on the degree of immune cell infiltration in every sample, the abundance of $CD8^+$ T cells and M2 macrophage infiltration significantly increased in the high PS group (P<0.05) (*Figure 5B*). The high PS group had significantly higher (P<0.05) expression of *MARCKS* and *CASP8AP2*, as revealed by differential analysis of genes associated with the G2/M checkpoint (*Figure 5C*). A substantial positive association (P<0.001) was found in the correlation analysis between PS and immune checkpoint-related gene expression, including *ICOS*, *CTLA4*, and *LAG3* (*Figure 5D*).

Discussion

In contrast to normal tissues, tumor tissues exhibit higher levels of TIGIT expression, which this study examined and verified. Increased TIGIT expression and better OS were significantly correlated, according to the survival analysis. Both univariate and multivariate Cox regression analyses unequivocally confirmed that high TIGIT expression conferred a protective effect against OS. These findings provide substantial evidence that TIGIT may serve as a prognostic biomarker. We established a pathomics model using histopathological images to predict the expression of the main variable, TIGIT. Our results distinctly indicate that the pathological models could effectively predict TIGIT expression, thereby enabling LUAD prognosis.

Currently, the detection of TIGIT expression levels can

only be done through the following methods: detection of peripheral blood cytokines, detection of mRNA or protein levels based on fresh tissue specimens, and detection based on paraffin tissue specimens. All of these methods are subject to subjective influence of operators, antibodies, and high prices. Take immunohistochemistry (IHC) as an example. A large number of studies have found that diagnostic immunohistochemistry tests have performance differences between laboratories and cannot quantitatively and objectively evaluate the sensitivity of immunostaining (17). For example, in malignant mesothelioma, IHC specificity is not high and there is no exclusive specific antibody (18). However, H&E-stained sections are necessary for clinical diagnosis and are the most easily accessible image data. The pathomics model can digitize tissue slice information. By extracting digital features that cannot be observed by eye, it can analyze the characteristics of diseased tissues to a greater extent and has a good ability to predict the microscopic or molecular phenotype of tumors. For example, Chen *et al.* used the pathomics model to predict the prognosis of bladder cancer, and the model had good prediction performance (19). We constructed an objective batch pathomics prediction model for TIGIT expression in LUAD, with an AUC of 0.735 for the training set and 0.738 for the validation set. Based on the current criterion that AUC >0.7 is a good performance (20), the pathomics prediction model for TIGIT expression in LUAD has good prediction performance. In addition, the calibration curve shows that the model has a good calibration degree; DCA shows that the model has a high clinical net benefit. Through pathomics, objective, batch, and accurate prediction of TIGIT expression can be achieved.

While TIGIT has predominantly commanded attention for its response to immune therapy efficacy, our study was uniquely tailored to explore its role in prognosis. Notably, our findings conform to those of Li *et al.* (21), which suggested that high TIGIT expression in LUAD is indicative of better OS. Moreover, the prognosis of oral cancer patients has been shown to improve due to elevated TIGIT expression (22). According to Zhong *et al.*'s (23) research, individuals with myelodysplastic syndrome who are younger than 60 years old and have high TIGIT expression have a lower OS rate. Zou *et al.*'s (24) study also reported shorter OS in TIGIT-positive cervical cancer patients with human papilloma virus (HPV) infection. These disparities could be attributed to the difference in the immune environments of various cancers, signifying that the same molecule may play different roles in different tumors. In this study, higher TIGIT expression levels predicted a better prognosis.

In the subsequent mechanism research, we observed that the immunological infiltration of $CD8^+$ T cells and M2 macrophages was substantially linked with the PS based on TIGIT. Previous studies have demonstrated a close relationship between $CD8^+$ T cells and TIGIT expression (25). TIGIT has the ability to directly impair $CD8^+$ T cell function, thereby thwarting their tumor-killing effect (26). Additionally, M2 macrophages are immune cells, and TIGIT, being an immunosuppressive molecule, typically suppresses immune responses by negatively regulating immune cell signaling pathways (27). Therefore, we speculate that our research findings may be the reason for the upregulation of immune checkpoints in the immune microenvironment in the high-TIGIT group. The enrichment analysis conducted in the subsequent study focused on the EMT and G2/M checkpoint pathways. The expression of EMT genes, such as *WIPF1*, *GLIPR1*, and *IL15*, and the PS showed a strong positive connection. In the high PS group, the expression of the G2/M checkpoint-related genes *MARCKS* and *CASP8AP2* displayed significant augmentation. These findings have prompted us to explore the regulatory nexus between EMT genes, G2/M checkpoint-related genes, and the immune checkpoint TIGIT in our forthcoming mechanistic research.

Although EMT usually predicts a poor prognosis in the vast majority of studies, how TIGIT regulates EMT needs further mechanistic studies. We believe that it is necessary to discuss the significance of TIGIT for the prognosis of LUAD. In our study, we found that high expression of TIGIT (HR =0.65; 95% CI: 0.44–0.95; P=0.03) was a protective factor for OS. In previous studies, the effect of TIGIT on prognosis was analyzed in 33 types of cancer, and it was found that TIGIT played a completely different role in different cancers. High expression of TIGIT was associated with poor prognosis in kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), low-grade glioma (LGG), and uveal melanoma (UVM), and with good prognosis in breast cancer (BRCA), head and neck squamous cell carcinoma (HNSC), and cutaneous melanoma (28). A meta-analysis of solid tumors in East Asian populations found that no effect of TIGIT on tumor OS prognosis was found in 8 studies, and TIGIT was found to be a risk factor for tumor OS

prognosis in the remaining studies (25). Therefore, the role of TIGIT in tumors may be affected by tumor type and regional population, which still needs further research and exploration. Although the role of TIGIT in cancer is still controversial, it is undeniable that TIGIT has become an important indicator for lung cancer prognosis and immunotherapy.

It is widely acknowledged that *ICOS*, *CTLA4*, and *LAG3* are also common immune checkpoints. In our study, there was a significant positive correlation between PS and the expression of these immune checkpoints. This may explain the improved efficacy of the combination of immune checkpoints in clinical practice (29-32).

The application of pathological models to forecast clinical prognosis is a simple and convenient method. Cai *et al.* (33) studied 759 H&E-stained images from a multicenter cohort to construct a pathological model for predicting the malignant transformation of oral leukoplakia. Using pathological images from 480 patients with gastric cancer, Chen *et al.* (34) harnessed machine learning to construct a predictive model for gastric cancer prognosis. However, the predictive ability of a solitary image feature has its constraints; therefore, we employed LR algorithms to combine image features, enhancing the precision of prognosis prediction. The advantage of this study is that we not only extracted pathological features but also combined TIGIT molecules. Because of this, the model can forecast the expression of TIGIT molecules and thus the prognosis of individuals with LUAD, which is economical, fast, and highly accurate.

Nonetheless, there are some limitations in this study. For example, the data source was based on public databases, which introduced notable heterogeneity in both images and data. In addition, the sample size was small, necessitating substantial verification with large cohorts in future work. The images may have had selective offsets, and potential selective biases may have influenced the images. However, our pioneering use of pathological models for TIGIT molecule prediction represents a noteworthy contribution to the realm of clinical precision medicine.

Conclusions

As a molecular marker for LUAD prognosis research, TIGIT expression exhibits a substantial correlation with this condition's prognosis. Employing pathological models to predict LUAD prognosis is convenient, economical, and effective.

Acknowledgments

We thank Editage for language editing services.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-978/rc

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-978/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-978/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol 2015;10:1243-60.
2. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 2012;150:1107-20.
3. Yuan Y, Xie B, Guo D, et al. Identification of ALG3 as a potential prognostic biomarker in lung adenocarcinoma.

Heliyon 2023;9:e18065.
4. Chen K, Zheng T, Chen C, et al. Pregnancy Zone Protein Serves as a Prognostic Marker and Favors Immune Infiltration in Lung Adenocarcinoma. Biomedicines 2023;11:1978.
5. Kosai K, Masuda T, Kitagawa A, et al. Transducin Beta-Like 2 is a Potential Driver Gene that Adapts to Endoplasmic Reticulum Stress to Promote Tumor Growth of Lung Adenocarcinoma. Ann Surg Oncol 2023;30:7538-48.
6. Manieri NA, Chiang EY, Grogan JL. TIGIT: A Key Inhibitor of the Cancer Immunity Cycle. Trends Immunol 2017;38:20-8.
7. Harjunpää H, Guillerey C. TIGIT as an emerging immune checkpoint. Clin Exp Immunol 2020;200:108-19.
8. Zhou XM, Li WQ, Wu YH, et al. Intrinsic Expression of Immune Checkpoint Molecule TIGIT Could Help Tumor Growth in vivo by Suppressing the Function of NK and CD8(+) T Cells. Front Immunol 2018;9:2821.
9. Liang R, Zhu X, Lan T, et al. TIGIT promotes CD8(+)T cells exhaustion and predicts poor prognosis of colorectal cancer. Cancer Immunol Immunother 2021;70:2781-93.
10. Xu D, Zhao E, Zhu C, et al. TIGIT and PD-1 may serve as potential prognostic biomarkers for gastric cancer. Immunobiology 2020;225:151915.
11. Liu K, Hu J. Classification of acute myeloid leukemia M1 and M2 subtypes using machine learning. Comput Biol Med 2022;147:105741.
12. Nishio M, Nishio M, Jimbo N, et al. Homology-Based Image Processing for Automatic Classification of Histopathological Images of Lung Tissue. Cancers (Basel) 2021;13:1192.
13. Banna GL, Olivier T, Rundo F, et al. The Promise of Digital Biopsy for the Prediction of Tumor Molecular Features and Clinical Outcomes Associated With Immunotherapy. Front Med (Lausanne) 2019;6:172.
14. Chen L, Zeng H, Zhang M, et al. Histopathological image and gene expression pattern analysis for predicting molecular features and prognosis of head and neck squamous cell carcinoma. Cancer Med 2021;10:4615-28.
15. Wang X, Chen H, Gan C, et al. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Trans Cybern 2020;50:3950-62.
16. Saednia K, Lagree A, Alera MA, et al. Quantitative digital histopathology and machine learning to predict pathological complete response to chemotherapy in breast cancer patients using pre-treatment tumor biopsies. Sci Rep 2022;12:9690.
17. Sompuram SR, Vani K, Schaedle AK, et al. Quantitative Assessment of Immunohistochemistry Laboratory Performance by Measuring Analytic Response Curves and Limits of Detection. Arch Pathol Lab Med 2018;142:851-62.
18. Suster S, Moran CA. Applications and limitations of immunohistochemistry in the diagnosis of malignant mesothelioma. Adv Anat Pathol 2006;13:316-29.
19. Chen S, Jiang L, Zheng X, et al. Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer. Cancer Sci 2021;112:2905-14.
20. Wu J, Zhang H, Li L, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: A population-based analysis. Cancer Commun (Lond) 2020;40:301-12.
21. Li L, Li J. Correlation of tumor mutational burden with prognosis and immune infiltration in lung adenocarcinoma. Front Oncol 2023;13:1128785.
22. Eichberger J, Spoerl S, Spanier G, et al. TIGIT Expression on Intratumoral Lymphocytes Correlates with Improved Prognosis in Oral Squamous Cell Carcinoma. Biomedicines 2022;10:3236.
23. Zhong M, Chen C, Zhao W, et al. High Co-Expression of PDCD1/TIGIT/CD47/KIR3DL2 in Bone Marrow Is Associated with Poor Prognosis for Patients with Myelodysplastic Syndrome. J Oncol 2023;2023:1972127.
24. Zou W, Huang R, Li P, et al. Clinical significance of immune checkpoint proteins in HPV-infected cervical cancer. J Infect Public Health 2023;16:542-50.
25. Li S, Li L, Pan T, et al. Prognostic value of TIGIT in East Asian patients with solid cancers: A systematic review, meta-analysis and pancancer analysis. Front Immunol 2022;13:977016.
26. Johnston RJ, Comps-Agrar L, Hackney J, et al. The immunoreceptor TIGIT regulates antitumor and antiviral CD8(+) T cell effector function. Cancer Cell 2014;26:923-37.
27. Alemohammad H, Najafzadeh B, Asadzadeh Z, et al. The importance of immune checkpoints in immune monitoring: A future paradigm shift in the treatment of cancer. Biomed Pharmacother 2022;146:112516.
28. Wen J, Mao X, Cheng Q, et al. A pan-cancer analysis revealing the role of TIGIT in tumor microenvironment. Sci Rep 2021;11:22502.
29. Willsmore ZN, Coumbe BGT, Crescioli S, et al. Combined anti-PD-1 and anti-CTLA-4 checkpoint blockade: Treatment of melanoma and immune

mechanisms of action. Eur J Immunol 2021;51:544-56.
30. Ahn MJ, Niu J, Kim DW, et al. Vibostolimab, an anti-TIGIT antibody, as monotherapy and in combi-nation with pembrolizumab in anti-PD-1/PD-L1-refractory NSCLC. Ann Oncol 2020;31:S887.
31. Niu J, Maurice-Dror C, Lee DH, et al. First-in-human phase 1 study of the anti-TIGIT antibody vibostolimab as monotherapy or with pembrolizumab for advanced solid tumors, including non-small-cell lung cancer. Ann Oncol 2022;33:169-80.
32. Johnson ML, Patel MR, Ulahannan SV, et al. Phase I study of BI 754111 (anti-LAG-3) plus BI 754091(anti-PD-1) in patients (pts) with advanced solid cancers, followed by expansion in pts with microsatellite stable metastatic colorectal cancer (mCRC), anti-PD-(L)1-pretreated non-small cell lung cancer (NSCLC) and other solid tumors. Ann Oncol 2018;29:VIII441.
33. Cai X, Li L, Yu F, et al. Development of a Pathomics-Based Model for the Prediction of Malignant Transformation in Oral Leukoplakia. Lab Invest 2023;103:100173.
34. Chen D, Fu M, Chi L, et al. Prognostic and predictive value of a pathomics signature in gastric cancer. Nat Commun 2022;13:6903.

Cite this article as: Hu P, Tian B, Gu H, Liu H, Li Q. Predicting prognosis in lung adenocarcinoma by predicting TIGIT expression: a pathomics model. J Thorac Dis 2024;16(11):7617-7629. doi: 10.21037/jtd-24-978

Supplementary

Table S1 Baseline data of patients in the training and validation sets

Variables	Total (n=327)	Train (n=229)	Validation (n=98)	P
TIGIT				1
Low	130 (40%)	91 (40%)	39 (40%)	
High	197 (60%)	138 (60%)	59 (60%)	
Gender				0.74
Female	183 (56%)	130 (57%)	53 (54%)	
Male	144 (44%)	99 (43%)	45 (46%)	
Age (years)				0.38
<66	164 (50%)	119 (52%)	45 (46%)	
≥66	163 (50%)	110 (48%)	53 (54%)	
Pathologic_stage				0.37
I/II	265 (81%)	189 (83%)	76 (78%)	
III/IV	62 (19%)	40 (17%)	22 (22%)	
Radiotherapy				0.09
No	293 (90%)	210 (92%)	83 (85%)	
Yes	34 (10%)	19 (8%)	15 (15%)	
Chemotherapy				0.58
No	219 (67%)	156 (68%)	63 (64%)	
Yes	108 (33%)	73 (32%)	35 (36%)	
Smoking_status				0.63
Nonsmoker	41 (13%)	30 (13%)	11 (11%)	
Current	82 (25%)	60 (26%)	22 (22%)	
Former	204 (62%)	139 (61%)	65 (66%)	
Residual_tumor				0.77
R0	220 (67%)	151 (66%)	69 (70%)	
R1/R2	13 (4%)	10 (4%)	3 (3%)	
RX/unknown	94 (29%)	68 (30%)	26 (27%)	
Histologic_type				0.41
NOS	204 (62%)	144 (63%)	60 (61%)	
Mixed subtype	67 (20%)	43 (19%)	24 (24%)	
Others	56 (17%)	42 (18%)	14 (14%)	
Tumor_location				0.83
L-lower	56 (17%)	42 (18%)	14 (14%)	
L-upper	76 (23%)	55 (24%)	21 (21%)	
R-lower	63 (19%)	43 (19%)	20 (20%)	
R-middle	14 (4%)	9 (4%)	5 (5%)	
R-upper	118 (36%)	80 (35%)	38 (39%)	
OS				1
Alive	213 (65%)	149 (65%)	64 (65%)	
Dead	114 (35%)	80 (35%)	34 (35%)	
Median OS (Q1, Q3), months	21.73 (14.48, 35.1)	22.13 (14.67, 35.7)	21.07 (14.12, 33.11)	0.65

TIGIT, T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain; NOS, not otherwise specified; L, left; R, right; OS, overall survival.

https://dx.doi.org/10.21037/jtd-24-978

Table S2 Clinical characteristics of different PS groups

Variables	Total (n=327)	Low (n=130)	High (n=197)	P
Gender				0.03
Female	183 (56%%)	83 (64%)	100 (51%)	
Male	144 (44%)	47 (36%)	97 (49%)	
Age (years)				1
<66	164 (50%)	65 (50%)	99 (50%)	
≥66	163 (50%)	65 (50%)	98 (50%)	
Pathologic_stage				0.048
I/II	265 (81%)	98 (75%)	167 (85%)	
III/IV	62 (19%)	32 (25%)	30 (15%)	
Radiotherapy				1
No	293 (90%)	116 (89%)	177 (90%)	
Yes	34 (10%)	14 (11%)	20 (10%)	
Chemotherapy				0.39
No	219 (67%)	83 (64%)	136 (69%)	
Yes	108 (33%)	47 (36%)	61 (31%)	
Smoking_status				0.03
Nonsmoker	41 (13%)	14 (11%)	27 (14%)	
Current	82 (25%)	24 (18%)	58 (29%)	
Former	204 (62%)	92 (71%)	112 (57%)	
Residual_tumor				0.21
R0	220 (67%)	88 (68%)	132 (67%)	
R1/R2	13 (4%)	8 (6%)	5 (3%)	
RX/unknown	94 (29%)	34 (26%)	60 (30%)	
Histologic_type				0.11
NOS	204 (62%)	74 (57%)	130 (66%)	
Mixed subtype	67 (20%)	34 (26%)	33 (17%)	
Others	56 (17%)	22 (17%)	34 (17%)	
Tumor_location				0.81
L-lower	56 (17%)	23 (18%)	33 (17%)	
L-upper	76 (23%)	33 (25%)	43 (22%)	
R-lower	63 (19%)	21 (16%)	42 (21%)	
R-middle	14 (4%)	6 (5%)	8 (4%)	
R-upper	118 (36%)	47 (36%)	71 (36%)	

PS, pathomics score; L, left; R, right; NOS, not otherwise specified.

https://dx.doi.org/10.21037/jtd-24-978

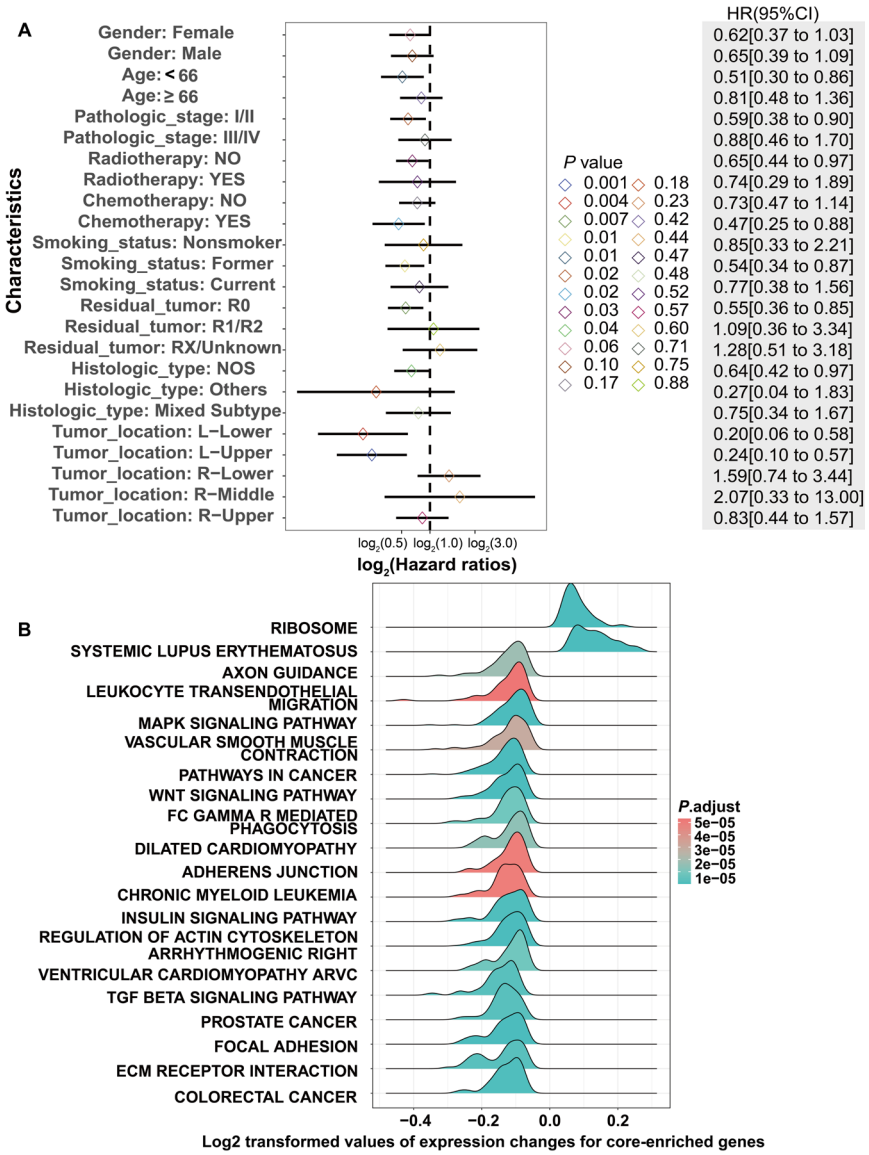

Figure S1 Univariate Cox regression subgroup analysis. (A) The effect of TIGIT high and low expression on the prognosis of patients in different subgroups of each covariate. (B) GSEA in the KEGG gene set. HR, hazard ratio; CI, confidence interval; NOS, not otherwise specified; TIGIT, T cell immunoreceptor with immunoglobulin and immunoreceptor tyrosine-based inhibitory motif domain; KEGG, Kyoto Encyclopedia of Genes and Genomes.

https://dx.doi.org/10.21037/jtd-24-978

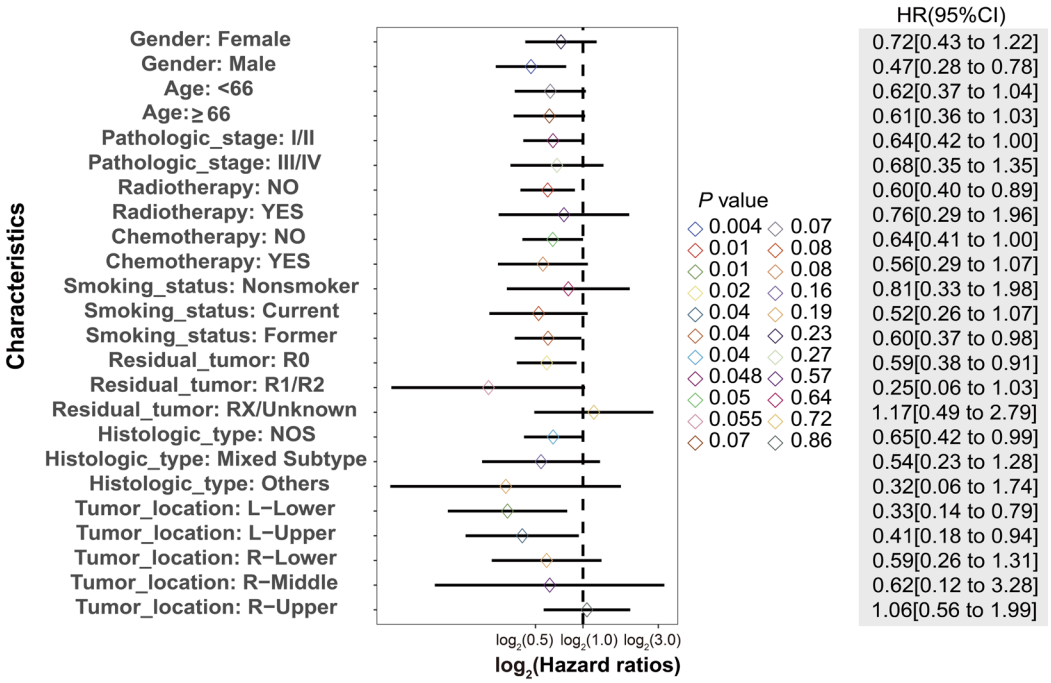

Figure S2 The effect of PS high and low expression on the prognosis of patients in different subgroups of each covariate. HR, hazard ratio; CI, confidence interval; NOS, not otherwise specified; PS, pathomics score.

https://dx.doi.org/10.21037/jtd-24-978