Heliyon 10 (2024) e31864

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

The diagnostic potential role of thioredoxin reductase and TXNRD1 in early lung adenocarcinoma: A cohort study

Guanyu Jiang [1], Xiaokun Wang [1], Yongrui Xu [1], Zhao He, Rongguo Lu, Chenghu Song, Yulin Jin, Huixing Li, Shengfei Wang [***], Mingfeng Zheng [**], Wenjun Mao [*]

Department of Thoracic Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi, 214023, China

A R T I C L E I N F O

Keywords:
TCGA
Thioredoxin reductase
Nomogram
Cohort study
Lung adenocarcinoma
Immune infiltration

A B S T R A C T

Background: Lung adenocarcinoma (LUAD) is the primary form of lung cancer, yet the reliable biomarkers for early diagnosis remain insufficient. Thioredoxin reductase (TrxR) is strongly linked to the occurrence, development, and drug resistance of lung cancer, making it a potential biomarker. However, further research is required to assess its diagnostic value in LUAD.

Methods: A retrospective analysis was performed on patients who underwent pulmonary nodule resection at our center from 2018 to 2022. Clinical data, including preoperative TrxR levels, imaging, and laboratory characteristics, were identified as study variables. Two prediction models were constructed using multiple logistic regression, and their prediction performance was evaluated comprehensively. Besides, bioinformatics analyses of TrxR coding genes including differential expression, functional enrichment, immune infiltration, drug sensitivity, and single-cell landscape were performed based on TCGA database, which were subsequently validated by Human Protein Atlas.

Results: A total of 506 eligible patients (72 benign lesions, 77 AISs, 185 MIAs and 172 IACs) were identified in the clinical cohort. Two TrxR-based models were developed, which were able to distinguish between benign and malignant pulmonary nodules, as well as pathological subtypes of LUAD, respectively. The models exhibited good predictive ability with all AUC values ranging from 0.7 to 0.9. Based on calibration curves and clinical decision analysis, the nomogram models showed high reliability. Functional analysis indicated that TXNRD1 primarily participated in cell cycle and lipid metabolism. Immune infiltration analysis showed that TXNRD1 has a strong association with immune cells and could impact immunotherapy. Then, we identified small molecular compounds that inhibit TXNRD1 and confirmed TXNRD1 expression by single-cell landscape and immunohistochemistry.

* Corresponding author. Department of Thoracic Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299 Qingyang Rd., Wuxi, 214023, China.

** Corresponding author. Department of Thoracic Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299 Qingyang Rd., Wuxi, 214023, China.

*** Corresponding author. Department of Thoracic Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299 Qingyang Rd., Wuxi, 214023, China.

E-mail addresses: ianwang0710@163.com (S. Wang), zhengmfmedical@126.com (M. Zheng), maowenjun1@njmu.edu.cn (W. Mao).

[1] These authors have contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.heliyon.2024.e31864
Received 21 September 2023; Received in revised form 22 May 2024; Accepted 22 May 2024
Available online 24 May 2024

Conclusion: This study validated the diagnostic value of TrxR and TXNRD1 in clinical cohorts and transcriptional data, respectively. TrxR and TXNRD1 could be used in the risk diagnosis of early LUAD and facilitate personalized treatment strategies.

1. Introduction

Lung cancer stands as a major cause of cancer-related mortality on a global scale, and lung adenocarcinoma (LUAD) is the most common type [1]. Early detection and treatment of LUAD could offer a favourable prognosis [2]. Despite the growing body of research on serological and histological markers for lung cancer patients, their low diagnostic sensitivity typically limits their effectiveness in clinical settings [3]. Besides, it is also urgent to determine the subtypes of LUAD before operation due to their prognostic heterogeneity [4,5]. Hence it is imperative to develop accurate and convenient biomarkers for early identification and subtype classification of LUAD.

Thioredoxin reductase (TrxR) belongs to the pyridine nucleotide-disulfide oxidoreductase family, whose primary role is to facilitate the reduction of NADPH-dependent thioredoxin (Trx), as well as other exogenous and endogenous compounds [6]. The Trx system in mammals consists of three distinct isoforms of TrxR, namely TrxR1 (cytoplasmic), TrxR2 (mitochondrial), and TrxR3 (TGR), which is primarily expressed in testicular tissue. TrxR1 is the most widely distributed and the encoding gene is TXNRD1 [7].

Compared to normal cells, the heightened metabolic activity and rapid proliferation of cancer cells result in elevated level of ROS, which is essential for maintaining tumor phenotypes. Due to the abnormal level of oxidative stress, tumor cells will enhance the activity of antioxidant system [8]. The Trx system is associated with a variety of cellular pathways, and also contribute to the progression and growth during malignant transformation [9]. After the activation of the tumor phenotype, elevated TrxR/Trx levels may enhance tumor progression and metastasis through the functions of promoting growth, anti-apoptosis, and supporting angiogenesis [10]. High TrxR expression levels have been confirmed in many tumors (especially lung cancer), which is also significantly associated with drug resistance in tumor cells and could be a critical component of the drug resistance phenotype [11]. Fernandes et al. found that TrxR expression was significantly correlated with tumor cell proliferation, patient survival and prognosis [12]. In addition, knockout of TrxR1-related genes could enhance the sensitivity of cancer cells to anticancer drug therapy (such as some selenium-containing compounds) [13]. The findings indicate that TrxR could serve as a potential and effective biomarker for the diagnosis of early LUAD. However, although previous studies have incorporated TrxR into diagnostic models for effective prediction of NSCLC, they have not accurately predicted the pathologic subtype of LUAD [14]. LUAD comprises three pathological subtypes [adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IAC)], each with distinct pathogenesis, metabolic differences, and prognosis [15–17]. Currently, there is no research to reveal whether TrxR expression varies across the three LUAD subtypes, and its potential role in subtype identification remains unexplored.

In this study, plasma TrxR levels, radiological chest CT data, laboratory characteristics, and other clinical data of patients of early LUAD from our center were collected to construct a clinical prediction model for forecasting the pathological properties of early lung nodules, which could facilitate diagnosis and further treatment. Bioinformatics analysis on TrxR-coding genes was performed to further verify the diagnostic value of TrxR, enabling the creation of more personalized and accurate treatment plans.

2. Materials and methods

2.1. Transcriptome and cohort data processing

The Cancer Genome Atlas (TCGA) database provided transcriptome data along with clinical and somatic mutation data for lung adenocarcinoma. The clinical information and CT images of 956 patients received radical surgical resection and pathologically diagnosed with benign lesion, AIS, MIA, and ICA in Wuxi People's Hospital (Wuxi, Jiangsu, China) between July 2018 and March 2022 were reviewed. The flowchart in Fig. 1 serves as a summary of this study design. Inclusion criteria were as follows [1]: serum TrxR content was detected one week before surgery [2]; age above 18 years [3] no occurrence of lung cancer or any other tumor in the five years prior to enrollment [4]; asymptomatic at diagnosis [5]; not accepted any treatment before operation. Exclusion criteria were [1]: disease-related detail is missing, such as unknown TAG, HDL, CEA, and so on [2]; presence of other pulmonary lesions [3]; a previous record of pulmonary illness [4]; metastatic tumors [5]; histopathologically diagnosed as other subtypes of lung cancer, including squamous cell carcinoma, small cell lung cancer, etc. Following the application of these inclusion and exclusion criteria, 450 patients were deemed ineligible. Finally, 506 patients (72 benign lesions, 77 AISs, 185 MIAs and 172 IACs) were analyzed (Table 1).

2.2. Gene expression analysis

Differential expression of TrxR-coding genes (TXNRD1, TXNRD2, TXNRD3) in pancancer was analyzed through TIMER2 database (Tumor Immune Estimation Resource, v2) [18]. RNA sequence data of LUAD-TCGA were applied to analyze the expression differences of TXNRD1 in paired and unpaired samples, respectively. The TISIDB database [19] was used to analyze the TXNRD1 gene expression difference in different LUAD immune subtypes.

2.3. Single cell and immunohistochemical landscape

In order to investigate TME heterogeneity across multiple datasets and cell types, the Tumor Immune Single-Cell Hub (TISCH) was utilized [20]. For further analysis of the distribution of TXNRD1 in various cell types, the dataset LUAD_EMTAB6149 of single-cell was

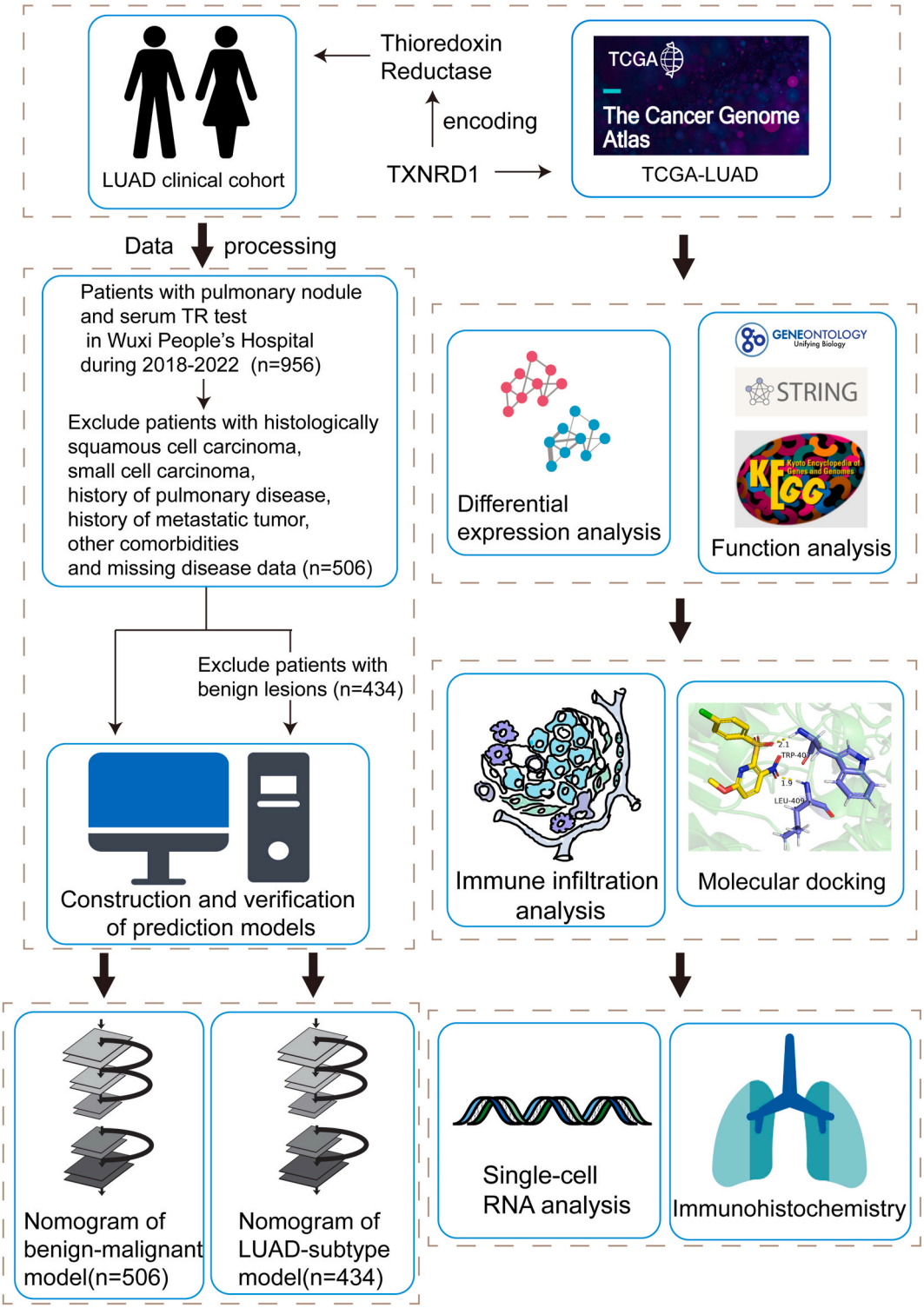

Fig. 1. Flow chart of the whole research.

chosen. The immunohistochemical study of TXNRD1 protein in LUAD and normal tissue was conducted using the Human Protein Atlas (HPA).

2.4. Study variables in the clinical cohort

In this clinical cohort, variables such as gender, age, body mass index (BMI), were collected, as well as radiological chest CT data, including pulmonary nodule components [solid, pure ground-glass opacity (pGGO) or mixed ground-glass opacity (mGGO)] and pulmonary nodule location. Laboratory characteristics, including thioredoxin reductase (TrxR), triacylglycerol (TAG), total cholesterol (Tch), low-density lipoprotein (LDL), high-density lipoprotein (HDL), Apolipoprotein A1 (ApoA1), Apolipoprotein B (ApoB), lipoprotein(a) [Lp(a)], cholinesterase (ChE), alkaline phosphatase (AKP), albumin (ALB), globulin (GLB), ALB/GLB (A/G), alpha fetal protein (AFP), carcinoembryonic antigen (CEA), carbohydrate antigen 125 (CA125), carbohydrate antigen 19-9 (CA19-9), cytokeratin 19 fragment, and neuron-specific enolase (NSE), were also recorded.

2.5. Construction and verification of models

We built two prediction models: a model for distinguishing between benign and malignant lesion (called benign-malignant model) and a model for diagnosing AIS, MIA, and IAC (called LUAD-subtype model).

506 patients (72 benign and 434 malignant lesions) were used to build benign-malignant model, a binary logistic regression. For further analysis, 70 % of screened patients (n = 354) were randomly assigned to a training cohort and 30 % (n = 152) were classified into a validation cohort to externally verify the final nomogram. After identifying significant risk factors through univariate logistic regression models, a diagnostic model was constructed using multiple logistic regression analysis. The calculation of odds ratios (ORs) and their respective 95 % confidence intervals (CIs) was performed for every risk factor. The basic equation for binary logistic regression is [21,22]:

$$y = logit(p) = \ln\frac{p}{1-p} = \partial + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m \quad (1)$$

Or, equivalently:

Table 1
Baseline characteristics of benign and malignant groups.

Characteristic	Benign	Malignant	total	p.value
	n = 72	n = 434	n = 506	
Gender, n(%)				0.12
Male	33 (45.8 %)	154 (35.5 %)	187	
Female	39 (54.2 %)	280 (64.5 %)	319	
Pulmonary nodule components, n(%)				<0.001
Solid	39 (54.2 %)	53 (12.2 %)	92	
mGGO	19 (26.4 %)	204 (47.0 %)	223	
pGGO	14 (19.4 %)	177 (40.8 %)	191	
Age, median(IQR)	54.00(46.00–64.00)	58.00(49.00–66.00)	57.00(48.00–65.00)	0.032
TrxR, median(IQR)	6.05(4.77–7.30)	7.00(5.15–8.60)	6.80(5.00–8.50)	0.003
Tumor_size, median(IQR)	11.00(8.00–16.25)	10.00(8.00–15.00)	10.50(8.00–15.00)	0.769
BMI, median(IQR)	19.07(17.77–20.07)	19.07(18.20–19.53)	19.07(18.12–19.63)	0.962
TAG, median(IQR)	1.27(0.91–1.76)	1.23(0.93–1.74)	1.24(0.92–1.74)	0.882
Tch, median(IQR)	4.85(4.03–5.35)	4.81(4.26–5.56)	4.81(4.18–5.55)	0.429
LDL, median(IQR)	2.73(2.18–3.23)	2.68(2.23–3.21)	2.69(2.23–3.22)	0.882
HDL, median(IQR)	1.12(0.91–1.31)	1.16(1.00–1.36)	1.15(0.98–1.36)	0.082
ApoA1, median(IQR)	1.40(1.23–1.54)	1.39(1.21–1.57)	1.40(1.21–1.57)	0.905
ApoB, median(IQR)	0.80(0.69–0.94)	0.80(0.69–0.95)	0.80(0.69–0.95)	0.516
LppA, median(IQR)	155.50(81.25–283.50)	129.00(62.50–279.00)	131.00(64.25–280.75)	0.235
ChE, median(IQR)	7209.00(6453.25–8539.25)	7540.50(6538.00–8539.25)	7474.00(6509.00–8539.25)	0.445
AKP, median(IQR)	74.50(63.75–87.25)	77.00(62.25–97.00)	77.00(63.00–95.00)	0.31
ALB, median(IQR)	40.55(38.68–42.62)	41.05(39.40–42.88)	41.00(39.30–42.80)	0.183
GLB, median(IQR)	26.70(25.08–30.00)	27.40(24.70–29.80)	27.30(24.70–29.90)	0.98
A/G, median(IQR)	1.48(1.40–1.61)	1.52(1.38–1.67)	1.51(1.38–1.66)	0.357
AFP, median(IQR)	2.43(1.88–3.30)	2.55(1.93–3.31)	2.54(1.92–3.31)	0.605
CEA, median(IQR)	1.14(0.77–2.20)	1.41(0.92–2.28)	1.40(0.90–2.26)	0.149
CA125, median(IQR)	8.55(6.02–11.85)	8.45(6.12–12.00)	8.45(6.10–12.00)	0.951
CA19_9, median(IQR)	6.25(4.00–9.43)	6.80(4.50–11.20)	6.70(4.40–11.15)	0.143
Cytokeratin_19_fragment, median(IQR)	2.16(1.69–2.84)	2.22(1.67–3.00)	2.22(1.69–2.95)	0.633
NSE, median(IQR)	11.37(9.89–12.86)	11.38(9.95–12.95)	11.38(9.94–12.95)	0.955

mGGO, mixed ground-glass opacity; pGGO, pure ground-glass opacity; TrxR, thioredoxin reductase; BMI, body mass index; TAG, triacylglycerol; Tch, total cholesterol; LDL, low-density lipoprotein; HDL, high-density lipoprotein; ApoA1, Apolipoprotein A1; ApoB, Apolipoprotein B; LppA, lipoprotein (a); ChE, cholinesterase; AKP, alkaline phosphatase; ALB, albumin; GLB, globulin; A/G, ALB/GLB; AFP, alpha fetal protein; CEA, carcinoembryonic antigen; CA125, carbohydrate antigen 125; CA19-9, carbohydrate antigen 19-9; NSE, neuron-specific enolase; IQR, interquartile range.

$$p = \frac{1}{1 + e^{-(\partial + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m)}} \tag{2}$$

P value represents the probability that the patient's pathology is malignant, which is between 0 and 1. The intercept (∂) represents the point where regression line intersects with the vertical Y-axis. The value of each independent variable (xi), weighted by its corresponding beta coefficient (β), is given by the expression β1x1 + β2x2 + ⋯ + βmxm. The beta coefficients indicate the extent to which the result increases with every one-unit increase in the value of the corresponding variable. We usually use the OR value to determine the influence of the independent variable change on the p value. The formula is as follows:

$$OR_{x_m} = e^{\beta_m} \tag{3}$$

OR means: For each unit change in the independent variable, p increases to eβm times the reference level. So we can get the probability of malignancy for each patient, which is used to plot the receiver operating characteristic (ROC) curve to assess the accuracy of this model. Calibration curves were conducted to establish a dependable comparison between predicted and observed outcomes within the training and verification cohorts. Decision curve analysis (DCA), an efficient algorithm assessing overall benefit of a model at various thresholds, was employed to validate the clinical value of our model.

434 patients (77 AISs, 185 MIAs and 172 IACs) were used to build LUAD-subtype model. Since three outcomes are progressive, the ordinal logistic regression was applied to examine the connection between potential predictors [23]. The patient population was separated into two cohorts, namely the training cohort (n = 303) and validation cohort (n = 131), with a division ratio of 7:3. By employing univariate and multivariate logistic regression models, we identified risk factors for pathological results in the training cohort, with concurrent recording of OR and 95 % CI. The identified risk factors were utilized to construct a nomogram for predicting AIS, MIA, and IAC. The basic equation for ordinal logistic regression [24] is:

$$y1 = logit(p1) = \ln\frac{p1}{1 - p1} = \partial_1 + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m \tag{4}$$

$$y2 = logit(p2) = \ln\frac{p2}{1 - p2} = \partial_2 + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m \tag{5}$$

Or, equivalently:

$$p1 = \frac{1}{1 + e^{-(\partial_1 + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m)}} \tag{6}$$

$$p2 = \frac{1}{1 + e^{-(\partial_2 + \beta_1 x1 + \beta_2 x2 + \ldots + \beta_m x_m)}} \tag{7}$$

in this equation, p1 indicates the probability that the pathology is no less than MIA, which means MIA or IAC. And p2 denotes the probability that the pathology is no less than IAC, which means IAC only. $\partial 1$ and $\partial 2$ are the intercepts of the two equations. Then we could get the probability of three outcomes: p (AIS) = 1-p1, p (MIA) = p1-p2, p (IAC) = p2. ROC and calibration curves for each pathological type could be performed based in the probabilities.

2.6. Interaction network and function characteristic analysis of TXNRD1

The R package "ClusterProfiler" was used to perform GO, KEGG and GSEA. The correlation between TXNRD1 and top 10 related genes was further analyzed. The STRING portal was used to obtain the top 10 TXNRD1-binding proteins.

2.7. Immune cell infiltration analysis

The application of CIBERSORT algorithm was to estimate the osmotic differences among various immune cells affected by high-and-low expression of TXNRD1. The relationship between TXNRD1 expression and the soakage of immune cell in tumor microenvironment was investigated. The calculation of immune and stromal cell scores were carried out by "ESTIMATE" package, respectively.

2.8. Immunotherapy roles of TXNRD1 in LUAD

Somatic mutation data (SNPs and small INDELs) were used to quantify the TMB of LUAD. The microsatellite instability (MSI) data was acquired from UCSC Xena (https://xena.ucsc.edu/). The analysis of relationship between TXNRD1 and TMB/MSI was conducted, respectively. To investigate the impact of TXNRD1 expression in immunotherapy, The Cancer Immunome Atlas (TCIA) [25] web portal was used for comparison.

2.9. Drug sensitivity analysis and molecular docking

By using the Gene Set Cancer Analysis (GSCA) portal, the association between TXNRD1 and small molecule drug chemosensitivity was evaluated [26]. In order to further study the interaction between small molecular substances and TXNRD1, we performed

molecular docking. Structure of TXNRD1-related small molecule compound was obtained from PubChem portal, and the best protein crystal structures were searched and filtered through Protein Data Bank (PDB) portal. The molecular docking was conducted by Autodock Vina and results were visualized with PYMOL.

2.10. Statistical analysis

The categorical variables were described as frequency and percentage, and the groups were compared by chi-square tests. Measurement data (TrxR, tumor size, etc.) were expressed using median and interquartile range (IQR) due to their non-normality, and Wilcoxon test and Kruskal-Wallis test were used to compare groups. We used chi-squared tests and Wilcoxon test to compare the baseline data of demographical and clinicopathological features in disparate pathological groups. Correlations between variables were determined by Pearson or Spelman's correlation test. All statistical analyses and graphics were generated using R software version 4.2.2. We used the R package "MASS" to perform univariable and multivariable logistic regression, which helped to screen risk factors. R package "pROC", "rms", and "rmda" were used to plot ROC curves, calibration curves and DCA curves, and R package "regplot" was utilized to plot nomograms. A P value less than 0.05, calculated using a two-sided test, was deemed to be statistically significant (*P < 0.05; **P < 0.01; ***P < 0.001).

3. Results

3.1. Expression of TXNRD1 in human cancers

Flow chart of our research was depicted in Fig. 1. In pan-cancer, comparing with TXNRD2 and TXNRD3, the expression difference of TXNRD1 in ESCA, LUAD were more prominent (Fig. 2A, sFig. 1). We further examined the result in TCGA-LUAD cohort and observed the difference of TXNRD1 expression in LUAD and paired normal samples (Fig. 2B–C). The mRNA expression pattern of TXNRD1 was evaluated in multiple immune subtypes of LUAD (Fig. 2D).

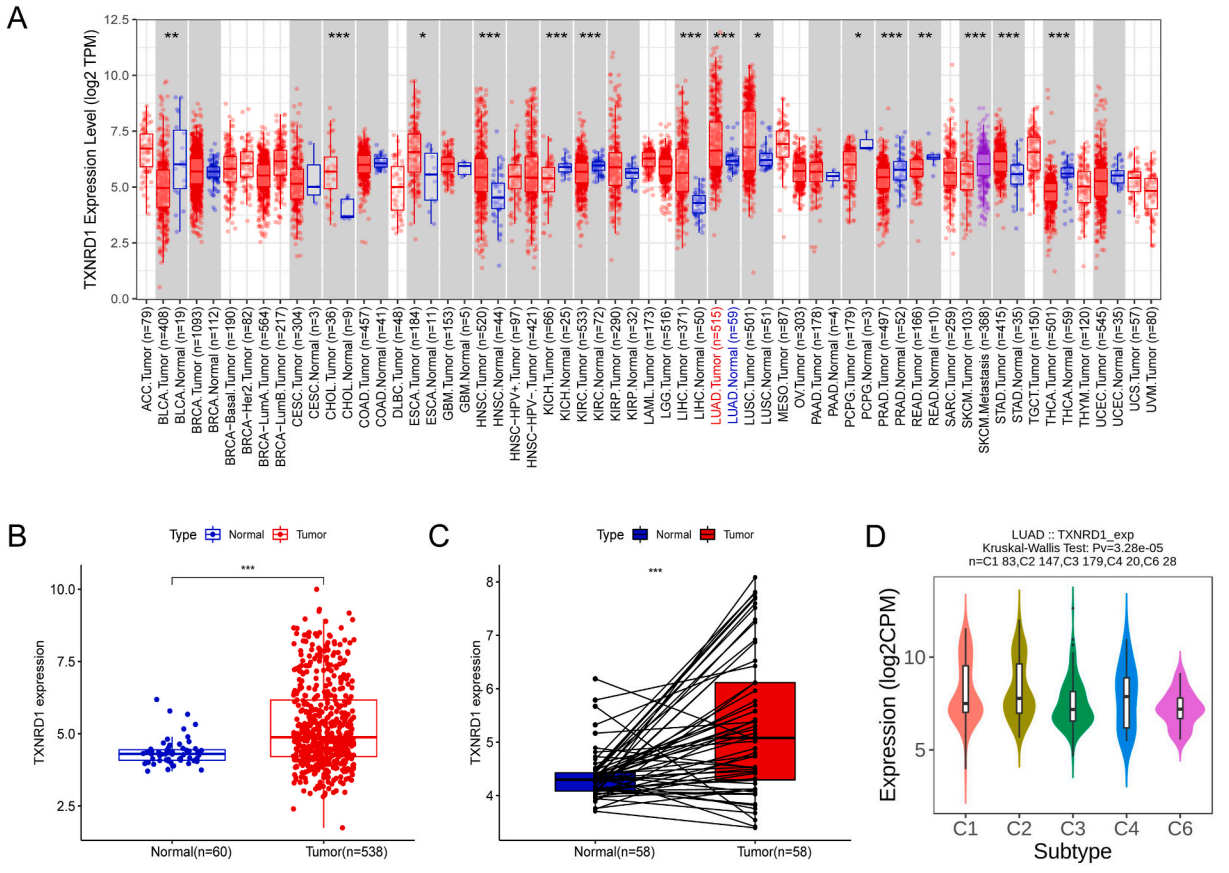

Fig. 2. Differential expression of TXNRD1. (A) Comparative TXNRD1 expression profiles in various malignancies through TIMER2 database. Differential expression levels of TXNRD1 in (B) normal and tumor tissues/(C) paired samples. (D) The discrepancies in TXNRD1 expression among immune subtypes of LUAD as determined by TISIDB. C1 (wound healing); C2 (IFN-γ dominant); C3 (inflammatory); C4 (lymphocyte depleted); C5 (immunologically quiet); C6 (TGF-β dominant).

3.2. Expression validation based on single-cell level and HPA

Based on single-cell data, the distribution of 12 cell types was identified in LUAD_EMTAB6149 dataset (Fig. 3A), as well as the expression of TXNRD1 in various cells (Fig. 3B). TXNRD1 could be found to be significantly expressed in most immune cells and cancer

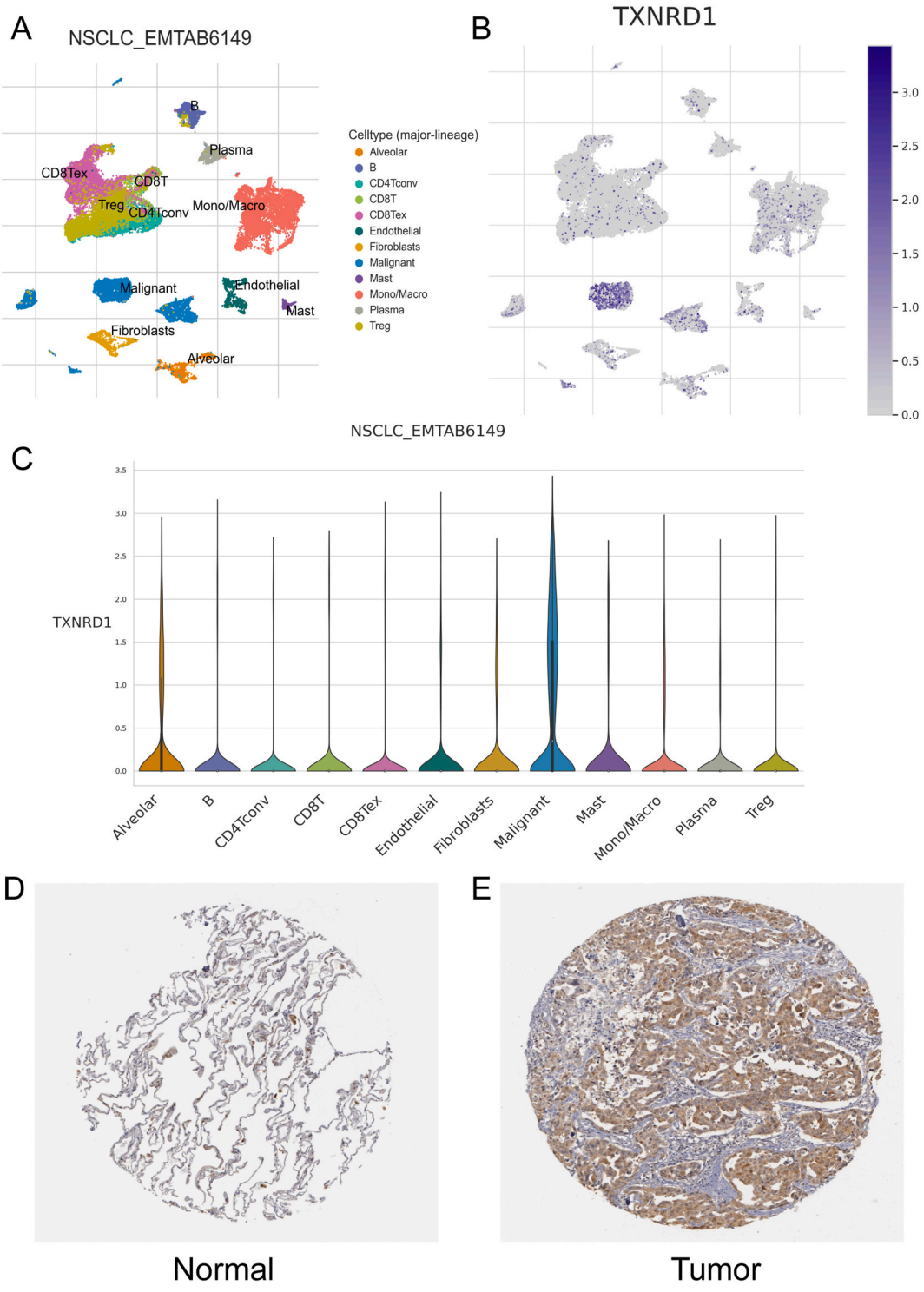

Fig. 3. Single cell and HPA landscape. (A–B) Annotation landscape of TME-related cells in EMTAB6149. (C) Expression of TXNRD1 in different cells. (D–E) The immunohistochemistry of TXNRD1 expression in normal and tumor tissues.

cells (Fig. 3C).

Finally, the expression of TXNRD1 in LUAD tissues was verified by immunohistochemical methods based on the HPA database (Fig. 3D–E).

3.3. Benign-malignant model: patients, risk analysis, construction, and validation

1 Baseline patient features.

A total of 506 eligible patients (72 benign and 434 malignant lesions) were comprised, with 354 and 152 patients assigned to the training and verification cohorts, respectively. As shown in Table 1 and sTable1, most patients were women (54.2 % in the benign cohort and 64.5 % in the malignant cohort). The benign cohort had the most solid nodules (54.2 %), while the malignant cohort had more pGGO (40.8 %) or mGGO (47.0 %). The median age of the benign and malignant cohorts were 54 years old and 58 years old, respectively ($P < 0.001$, Wilcoxon test). A statistically significant difference in TrxR was observed between benign (median = 6.05) and malignant cohorts (median = 7.00) ($P = 0.018$, Wilcoxon test). Differences could also be observed in HDL ($P = 0.082$, Wilcoxon test) in two cohorts. Both the chi-square and Wilcoxon tests demonstrated that the differences were randomized in the training and verification cohorts (sTable 2).

2 Univariate and Multivariate Logistic Regression Analysis

Results from the univariate logistic regression analysis in the training cohort revealed TrxR, pulmonary nodule components, age, HDL, Lp(a), ChE, ALB, A/G as risk factors for LUAD ($P < 0.3$), as presented in sTable 3. In the multivariate logistic regression analysis of these variables (Table 2), TrxR, pulmonary nodule components, age, HDL, and ALB were found to be independent risk factors ($P < 0.05$).

3 Construction and Validation of the Nomogram

All significant risk factors identified in the training cohort were incorporated into a model nomogram for predicting the risk of malignant lesions. A p-value of closer to 1 was considered more likely malignant lesion; the less likely a p-value was closer to 0. The predicted probabilities of malignant lesions were obtained by consolidating the scores associated with each variable and converting them to the bottom scales (Fig. 4). Outstanding forecasting ability was reflected in the AUC values for both the training cohort (0.821, 95%CI 0.752–0.889) and the verification cohort (0.7, 95%CI 0.566–0.833) (Fig. 5A–B). The precision of the nomogram was demonstrated by the calibration curves in both the training and verification cohort (Fig. 5C–D). Additionally, DCA result showed satisfactory clinical applicability in training and verification cohorts (Fig. 5E).

3.4. LUAD-subtype model: patients, risk analysis, construction, and validation

1 Role of TrxR in the diagnosis of LUAD

To delve deeper into the diagnostic role of TrxR in LUAD (AIS, MIA, and IAC), we established three models containing TrxR to distinguish them pairwise (sFigure 2, 3 and 4). AUC values of model for distinguishing between AIS and MIA were 0.677 (training cohort) and 0.547 (verification cohort). AUC values were 0.885 (training cohort) and 0.853 (verification cohort) in the model for distinguishing between MIA and IAC, and 0.945 (training cohort) and 0.928 (verification cohort) in the model for distinguishing

Table 2
Multivariate logistic regression analysis to identify risk factors for malignant transformation.

Variable	OR	95%CI	p.value
Age	1.003	1.001–1.006	0.017
Pulmonary nodule components			
Solid			reference
mGGO	1.36	1.240–1.493	<0.001
pGGO	1.469	1.334–1.617	<0.001
TrxR			
<6.8			reference
≥6.8	1.093	1.025–1.165	0.007
HDL			
<1.15			reference
≥1.15	1.089	1.022–1.160	0.009
ALB			
<41			reference
≥41	1.07	1.003–1.141	0.039

mGGO, mixed ground-glass opacity; pGGO, pure ground-glass opacity; TrxR, thioredoxin reductase; HDL, high-density lipoprotein; ALB, albumin.

between AIS and IAC. Of particular significance is that the calibration curves and DCA curves showed powerful consistency of the nomograms, which showed excellent discrimination of TrxR in three pathologies. On this basis, we established multivariate ordinal logistic regression model including AIS, MIA, and IAC (LUAD-subtype model).

2 Baseline patient features.

434 malignant cases (77 AISs, 185 MIAs and 172 IACs) were used to build LUAD-subtype model, a model for diagnosing AIS, MIA, and IAC. In accordance with a 7:3 ratio, patients were randomly assigned to a training cohort (n = 303) and a verification cohort (n = 131). The main characteristics of LUAD cases were detailed in Table 3 and sTable 4. Most patients were female (62.3 % in AIS cohort, 73.0 % in MIA cohort, and 56.4 % in IAC cohort), which was consistent with the crowd composition of benign-malignant model. Solid nodules increase with the malignancy (0 % in AIS cohort, 2.7 % in MIA cohort, and 27.9 % in IAC cohort), while pGGO showed the opposite trend (74 % in AIS cohort, 53.5 % in MIA cohort, and 12.2 % in IAC cohort). The median TrxR of three cohorts were 5.95, 6.90, and 7.49 years old, respectively ($P < 0.001$, Kruskal-Wallis test). Differences were observed in age ($P < 0.001$, Kruskal-Wallis test), tumor size ($P < 0.001$, Kruskal-Wallis test), TAG ($P = 0.023$, Kruskal-Wallis test), AKP ($P = 0.001$, Kruskal-Wallis test), GLB ($P = 0.029$, Kruskal-Wallis test), A/G ($P < 0.001$, Kruskal-Wallis test), CEA ($P < 0.001$, Kruskal-Wallis test), and cytokeratin 19 fragment ($P < 0.001$, Kruskal-Wallis test) in three cohorts. No significant disparities in demographics and clinical characteristics were detected between the training cohort and verification cohort by chi-squared test and Wilcoxon test (sTable 5).

3 Univariate and Multivariate Logistic Regression Analysis

Univariate logistic regression model was used to select risk factors, and TrxR, pulmonary nodule components, age, gender, tumor size, bmi, TAG, Tch, LDL, ApoA1, ApoB, AKP, A/G, CEA, CA125, and cytokeratin 19 fragment were potential factors in sTable 6 ($P < 0.3$). All these variables were selected for multivariate ordinal logistic regression analysis and TrxR, pulmonary nodule components, age, tumor size, Tch and LDL were identified as final factors for model construction in Table 4 ($P < 0.05$). Besides, a parallel regression analysis of the multivariate ordinal logistic regression was performed (sTable 7), and the results showed that the model satisfied the parallel regression test.

4 Construction and Validation of the Nomogram

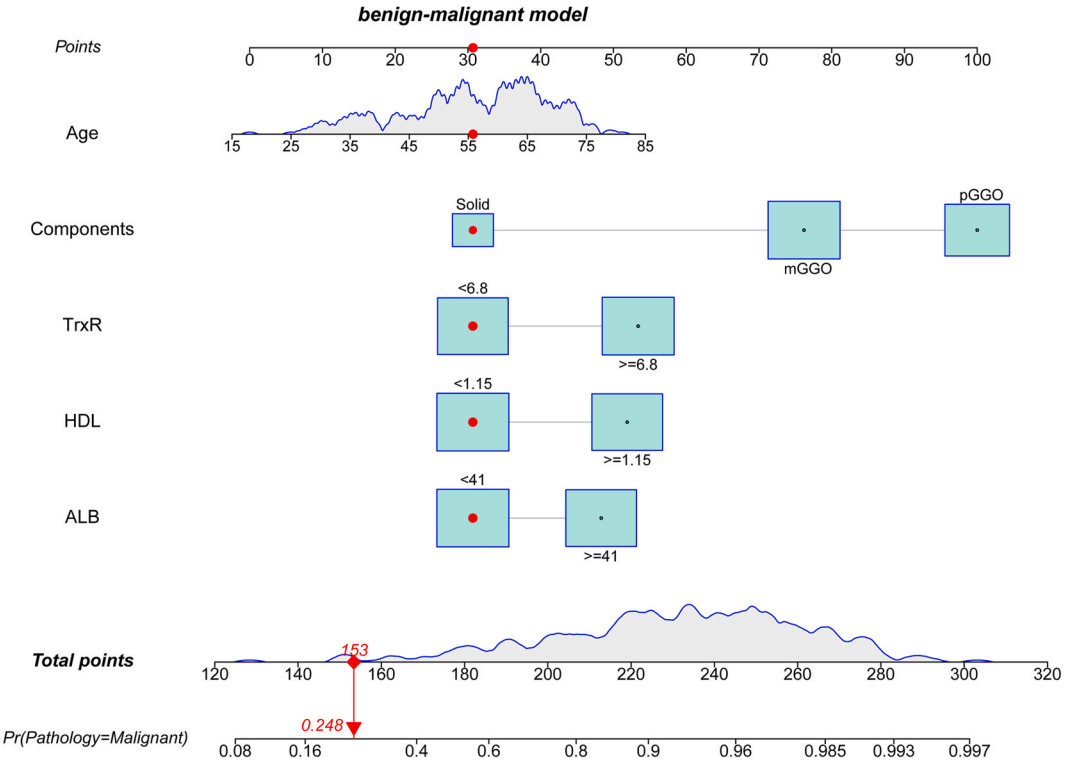

Fig. 4. Construction of benign-malignant model. The probability axis indicates the likelihood of a pulmonary nodule being LUAD. LUAD, lung adenocarcinoma.

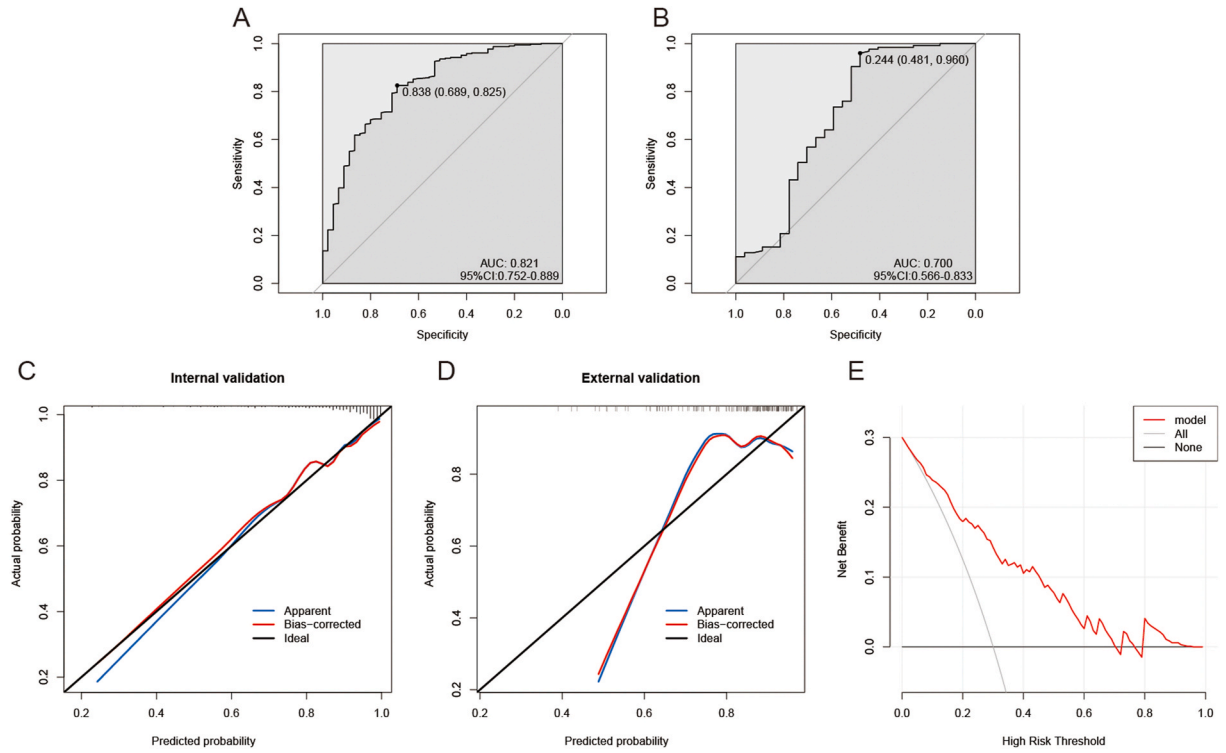

Fig. 5. Verification of benign-malignant model. ROC curves of training cohort (A) and validation cohort (B). Calibration plots of training cohort (C) and validation cohort (D). (E) DCA of benign-malignant model. The gray line represents all patients have LUAD, while the black line assumes none do. The net benefit of predicting LUAD using the nomogram model is shown by the red line. ROC, receiver operating characteristic; AUC, area under curve; DCA, decision curve analysis; LUAD, lung adenocarcinoma.

A total of six risk factors (including TrxR, pulmonary nodule components, tumor size, age, Tch, and LDL) entered the ordinal logistic regression equation. The nomogram for diagnosing AIS, MIA, and IAC is shown in Fig. 6 and the obtained score corresponds to two probabilities. The former is the probability that the pathology is MIA or IAC, and the latter is the probability that the pathology is only IAC. The overall C-index of this nomogram was 0.844, and ROC curves were applied to assess the precision for each pathological outcome in training and validation cohort. AUC values of AIS, MIA and IAC in the training cohort were 0.792 (95 % CI: 0.734–0.850), 0.716 (95 % CI: 0.658–0.773) and 0.897 (95 % CI: 0.862–0.931) (Fig. 7A,E, and I), while in the validation cohort were 0.781 (95 % CI: 0.690–0.871), 0.753 (95 % CI: 0.671–0.836) and 0.864 (95 % CI: 0.794–0.933) (Fig. 7C,G, and K). We conducted a graphical evaluation of calibration to illustrate the correlation between observed and predicted probabilities of each pathological outcome in training and validation cohort (Fig. 7B, D, F, H, J and L). The predicted and detected probabilities demonstrated excellent agreement in contrast to the ideal curve.

3.5. Functional analysis and interaction network of TXNRD1

To explore the potential biological behaviors of TXNRD1, GO showed that TXNRD1 was mainly associated with cilium movement, microtubule formation and cell cycle (sFig. 5A). KEGG analysis proved that TXNRD1 was significantly enriched in lipid metabolism, amino acid metabolism and tricarboxylic acid cycle process (sFig. 5B). The enrichment result of GSEA was shown in sFig. 5C. Besides, TXNRD1 expression was positively associated with GSR, TRIM16L, NEIL3, UGDH, NMRAL2P, GCLM in LUAD (sFig. 5D). The STRING database identified a total of 10 TXNRD1-interacting proteins and related interaction network (sFig. 5E).

3.6. Correlations between TXNRD1 and immune microenvironment

To further investigate the connection and potential role between TXNRD1 and immune microenvironment, we found that NK cells resting, $CD8^+$ T cells, Macrophages M1, $CD4^+$ memory-activated T cells were more infiltrated in TXNRD1 high expression subgroup (Fig. 8A). The relativity of TXNRD1 and the abundance of immune cell soakage in LUAD revealed that TXNRD1 expression was associated with $CD4^+$ memory-activated T cells, Macrophages M1, $CD8^+$ T cells, NK cells resting and Follicular helper T cells (Tfh) positively (Fig. 8B). The mesenchymal, immune cell and ESTIMATE scores in low TXNRD1 expression subgroup were all remarkably higher than that in high TXNRD1 expression subgroup (Fig. 8C).

Table 3
Baseline characteristics of AIS, MIA, and IAC groups.

Characteristic	AIS	MIA	IAC	total	p.value
	n = 77	n = 185	n = 172	n = 434	
Gender, n(%)					0.004
Male	29 (37.7 %)	50 (27.0 %)	75 (43.6 %)	154	
Female	48 (62.3 %)	135 (73.0 %)	97 (56.4 %)	280	
Pulmonary nodule components, n (%)					<0.001
Solid	0 (0.0 %)	5 (2.7 %)	48 (27.9 %)	53	
mGGO	20 (26.0 %)	81 (43.8 %)	103 (59.9 %)	204	
pGGO	57 (74.0 %)	99 (53.5 %)	21 (12.2 %)	177	
Age, median(IQR)	53.00(42.00–65.00)	54.00(44.00–64.00)	63.00(55.75–68.00)	58.00(49.00–66.00)	<0.001
TrxR, median(IQR)	5.95(3.91–7.53)	6.90(5.00–8.50)	7.49(5.54–9.20)	7.00(5.15–8.60)	<0.001
Tumor_size, median(IQR)	8.00(6.00–10.00)	9.00(7.50–12.00)	15.00(12.00–20.00)	10.00(8.00–15.00)	<0.001
BMI, median(IQR)	19.07(18.24–19.35)	19.07(17.83–19.46)	19.07(18.27–19.95)	19.07(18.20–19.53)	0.698
TAG, median(IQR)	1.15(0.85–1.69)	1.14(0.89–1.65)	1.33(1.01–1.89)	1.23(0.93–1.74)	0.023
Tch, median(IQR)	4.61(3.99–5.41)	4.77(4.30–5.40)	5.02(4.27–5.80)	4.81(4.26–5.56)	0.153
LDL, median(IQR)	2.70(2.23–3.37)	2.67(2.23–3.12)	2.66(2.24–3.33)	2.68(2.23–3.21)	0.533
HDL, median(IQR)	1.13(0.98–1.30)	1.19(1.03–1.36)	1.15(0.96–1.38)	1.16(1.00–1.36)	0.14
ApoA1, median(IQR)	1.35(1.19–1.47)	1.39(1.22–1.57)	1.42(1.24–1.64)	1.39(1.21–1.57)	0.099
ApoB, median(IQR)	0.79(0.67–0.89)	0.78(0.68–0.93)	0.84(0.71–0.99)	0.80(0.69–0.95)	0.052
LppA, median(IQR)	129.00(50.00–208.00)	125.00(67.00–309.00)	133.50(64.75–277.50)	129.00(62.50–279.00)	0.419
ChE, median(IQR)	7611.00 (6673.00–8541.00)	7464.00 (6553.00–8550.00)	7556.00 (6509.50–8484.50)	7540.50 (6538.00–8539.25)	0.767
AKP, median(IQR)	74.00(54.00–90.00)	75.00(60.00–94.00)	81.00(68.00–102.00)	77.00(62.25–97.00)	0.001
ALB, median(IQR)	42.00(39.60–43.60)	40.70(39.40–42.50)	41.15(39.25–43.15)	41.05(39.40–42.88)	0.114
GLB, median(IQR)	26.50(23.70–28.70)	27.70(25.00–30.20)	27.35(24.80–29.63)	27.40(24.70–29.80)	0.029
A/G, median(IQR)	1.60(1.46–1.71)	1.47(1.34–1.62)	1.51(1.39–1.68)	1.52(1.38–1.67)	0.001
AFP, median(IQR)	2.54(2.00–3.13)	2.37(1.84–3.27)	2.61(2.09–3.46)	2.55(1.93–3.31)	0.124
CEA, median(IQR)	1.31(0.86–1.93)	1.27(0.85–1.98)	1.85(1.03–2.90)	1.41(0.92–2.28)	<0.001
CA125, median(IQR)	8.80(6.60–13.40)	8.90(6.40–12.50)	8.10(5.70–10.90)	8.45(6.12–12.00)	0.081
CA19_9, median(IQR)	6.70(4.20–10.50)	6.40(4.40–9.80)	7.75(4.80–12.43)	6.80(4.50–11.20)	0.265
Cytokeratin_19_fragment, median (IQR)	2.05(1.57–2.78)	2.08(1.57–2.80)	2.44(1.92–3.23)	2.22(1.67–3.00)	0.001
NSE, median(IQR)	11.00(9.86–12.74)	11.41(9.97–13.14)	11.53(10.05–12.92)	11.38(9.95–12.95)	0.785

AIS, adenocarcinoma in situ; MIA, minimally invasive adenocarcinoma; IAC, invasive adenocarcinoma; mGGO, mixed ground-glass opacity; pGGO, pure ground-glass opacity; TrxR, thioredoxin reductase; BMI, body mass index; TAG, triacylglycerol; Tch, total cholesterol; LDL, low-density lipoprotein; HDL, high-density lipoprotein; ApoA1, Apolipoprotein A1; ApoB, Apolipoprotein B; LppA, lipoprotein(a); ChE, cholinesterase; AKP, alkaline phosphatase; ALB, albumin; GLB, globulin; A/G, ALB/GLB; AFP, alpha fetal protein; CEA, carcinoembryonic antigen; CA125, carbohydrate antigen 125; CA19-9, carbohydrate antigen 19-9; NSE, neuron-specific enolase; IQR, interquartile range.

Table 4
Multivariate logistic regression analysis to identify risk factors for subtype transformation of LUAD.

Variable	OR	95%CI	p.value
Age	1.032	1.011–1.053	0.003
Pulmonary nodule components			
Solid			reference
mGGO	0.218	0.067–0.588	0.005
pGGO	0.059	0.018–0.169	<0.001
TrxR			
<6.8			reference
≥6.8	1.748	1.076–2.854	0.025
Tumor_size			
<10.5			reference
≥10.5	6.946	3.972–12.485	<0.001
Tch			
<4.81			reference
≥4.81	2.333	1.146–4.859	0.021
LDL			
<2.69			reference
≥2.69	0.411	0.198–0.831	0.015

mGGO, mixed ground-glass opacity; pGGO, pure ground-glass opacity; TrxR, thioredoxin reductase; Tch, total cholesterol; LDL, low-density lipoprotein.

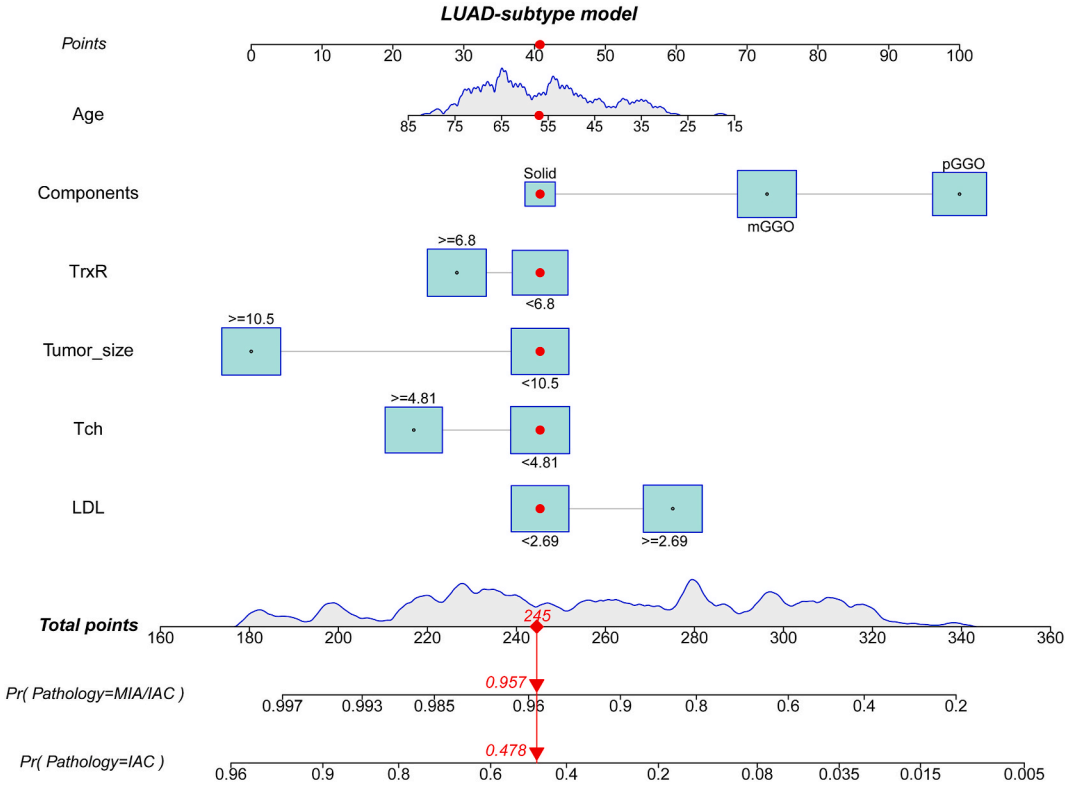

Fig. 6. LUAD-subtype model for diagnosing AIS, MIA, and IAC. The nomogram shows two probability axes, the probability of MIA or IAC, and the probability of only IAC. AIS, adenocarcinoma in situ; MIA, minimally invasive adenocarcinoma; IAC, invasive adenocarcinoma.

3.7. The impact of TXNRD1 in immunotherapy

Fig. 9A demonstrated that TXNRD1 was connected positively with TMB in LUAD. There was also a weakly negative correlation between TXNRD1 and MSI (Fig. 9B). The TCIA data showed that the immunophenotypic score (IPS) was notably higher in low TXNRD1 expression subgroup in both PD-1 positive and non-responsive patients with anti-CTLA-4 positive or negative status (Fig. 9C–F).

3.8. Drug sensitivity and molecular docking verification

The CTRP and GDSC databases also revealed a prominently positive relevance between TXNRD1 and several chemotherapeutic drug sensitivity (Fig. 10A–B). To test the potential of the molecular drug targeting TXNRD1, 2-(4-chlorophenyl) sulfonyl-6-methoxy-3-nitropyridine (TXNRD1 Inhibitor 1) was selected for molecular docking with the crystal structure of human thioredoxin reductase-thioredoxin complex (PDBID:3QFA). The docking results were shown in Fig. 10C–D.

4. Discussion

Nowadays, lung cancer remains a leading cause of cancer-related mortality, with an increasing incidence of LUAD [27,28]. It is particularly important to judge the benign and malignant of a pulmonary nodule through imaging features and laboratory examination. Different pathological types of LUAD obtained disparate follow-up treatment and showed different survival rate. Therefore, a comprehensive pathological assessment of pulmonary nodules before surgery is totally required.

While traditional biomarkers such as CEA, CA199, and NSE are still the primary choice for lung cancer screening, new biomarkers are constantly being discovered. Thioredoxin reductase (TrxR) is a selenocysteine-containing enzyme which is responsible for maintaining cellular REDOX homeostasis. It is usually upregulated in cancerous state, and works as a defense system against oxidative stress [29]. TrxR have been found to be overexpressed in tumor tissues, which can promote tumor cell growth and prevent apoptosis [30,31]. Its potential as a biomarker has been recognized across various cancers, including prostate cancer [32], breast cancer [33], and pancreatic cancer [34]. In lung cancer, elevated TrxR expression has been observed in patients with poorly differentiated and large tumors, and to be involved in the prediction of LUAD [14,35]. Our benign-malignant model, which combined blood indicators such as TrxR, HDL, and ALB with imaging features, was more comprehensive and stable than TrxR-based models in previous studies for predicting LUAD.

The 2015 World Health Organization (WHO) classification categorized lung adenocarcinoma into four histologic subtypes:

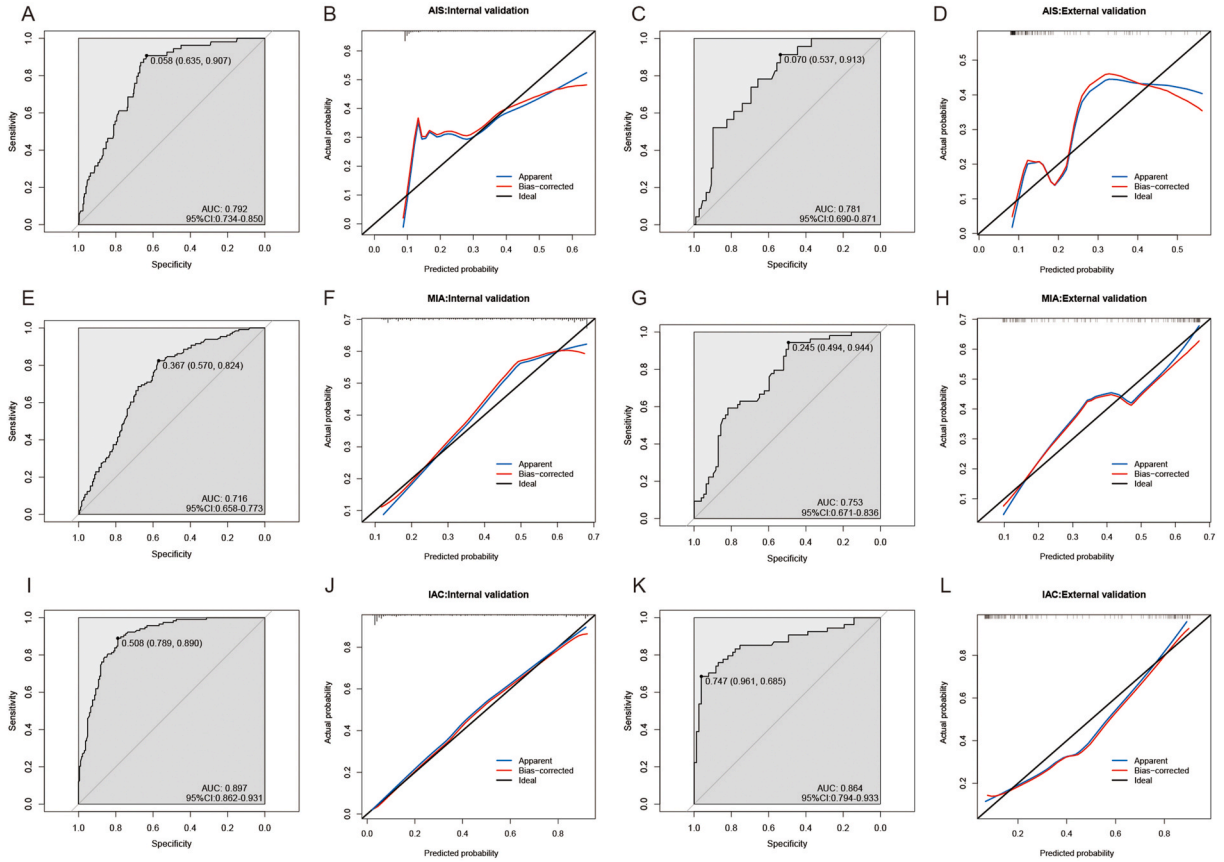

Fig. 7. Verification of LUAD-subtype model. Verification of AIS: ROC and calibration plots for the training cohort (A, B) and validation cohort (C, D). Verification of MIA: ROC and calibration plots for the training cohort (E, F) and validation cohort (G, H). Verification of IAC: ROC and calibration plots for the training cohort (I, J) and validation cohort (K, L). ROC, receiver operating characteristic; AUC, area under curve; DCA, decision curve analysis; LUAD, lung adenocarcinoma.

preinvasive lesions [including atypical adenomatous hyperplasia (AAH) and AIS], MIA, IAC, and adenocarcinoma variants [15]. The occurrence and development of LUAD is now understood to follow a continuous pattern of molecular changes that begin with normal alveolar type II cells and progress through AAH, AIS, MIA, and finally IAC. Mutations in EGFR, KRAS, and BRAF genes are widely acknowledged as pivotal factors in AAH development in normal alveolar cells [16]. AAH with EGFR mutations may progress to AIS, followed by a low propensity to advance to AIS or invasive lesions, but can do so after TP53/CDKN2A inactivation or other genomic aberrations. However, Xiang et al. reported distinct mutational features between AIS, MIA and IAC, and that specific gene mutations determine the degree of invasion of LUAD cells [17]. AIS and MIA exhibited similar mutation counts, whereas IAC had higher mutation counts when compared to AIS and MIA. The mutations in AIS and MIA were primarily concentrated in KRAS, ERBB2, BRAF, and MAP2K1, which showed an inverse correlation with the degree of invasion. The mutation rates of TP53 and EGFR were higher in IAC, and the mutation frequencies were distributed in a similar proportion along the multistep continuum. Besides, metabolic differences have been observed between AIS, MIA, and IAC in some studies [36,37]. The metabolic pathways of nicotinate and nicotinamide, β-alanine, glutathione, arginine, and proline show a progressively greater disruption as invasion advances from AIS to MIA and IAC [38]. The metabolic profiles of different LUAD subtypes are notably dissimilar. AIS mainly express spermine, spermidine, histamine, and cystine, whereas MIA has a preference for guanosine, guanine, cysteine, nicotinamide, and taurine. In summary, differences in gene mutation patterns and metabolic profiles may account for the distinct subtypes of LUAD.

Distinct molecular mechanisms could lead to different prognoses for AIS, MIA, and IAC in lung adenocarcinoma. Multiple studies have demonstrated there was no statistical disparity in 5-year survival between AIS and MIA when completely resected (both 100 %) [4,39,40]. It is widely believed that the pathological differentiation of AIS and MIA is intricate and may not hold significant clinical importance [41]. Our findings support this notion, as the ROC of the AIS-MIA model consistently underperformed the AIS-IAC and MIA-IAC models. However, accurate discrimination of IAC is paramount, as IAC exhibits a prognosis distinct from both MIA and AIS. According to a study, LPA had 68 % 5-year survival, PPA had 71 % 5-year survival, and MPA, SPA and the variants had the lowest survival (approximately 30 %–50 %) [5]. Therefore, early identification of IAC is essential to determine the appropriate treatments and improve prognosis. Precise differentiation of AIS, MIA, and IAC, particularly prior to surgery, remains a formidable task without histopathological examination. Most tumor markers used to diagnose LUAD cannot accurately distinguish AIS, MIA, and IAC. Hu et al.

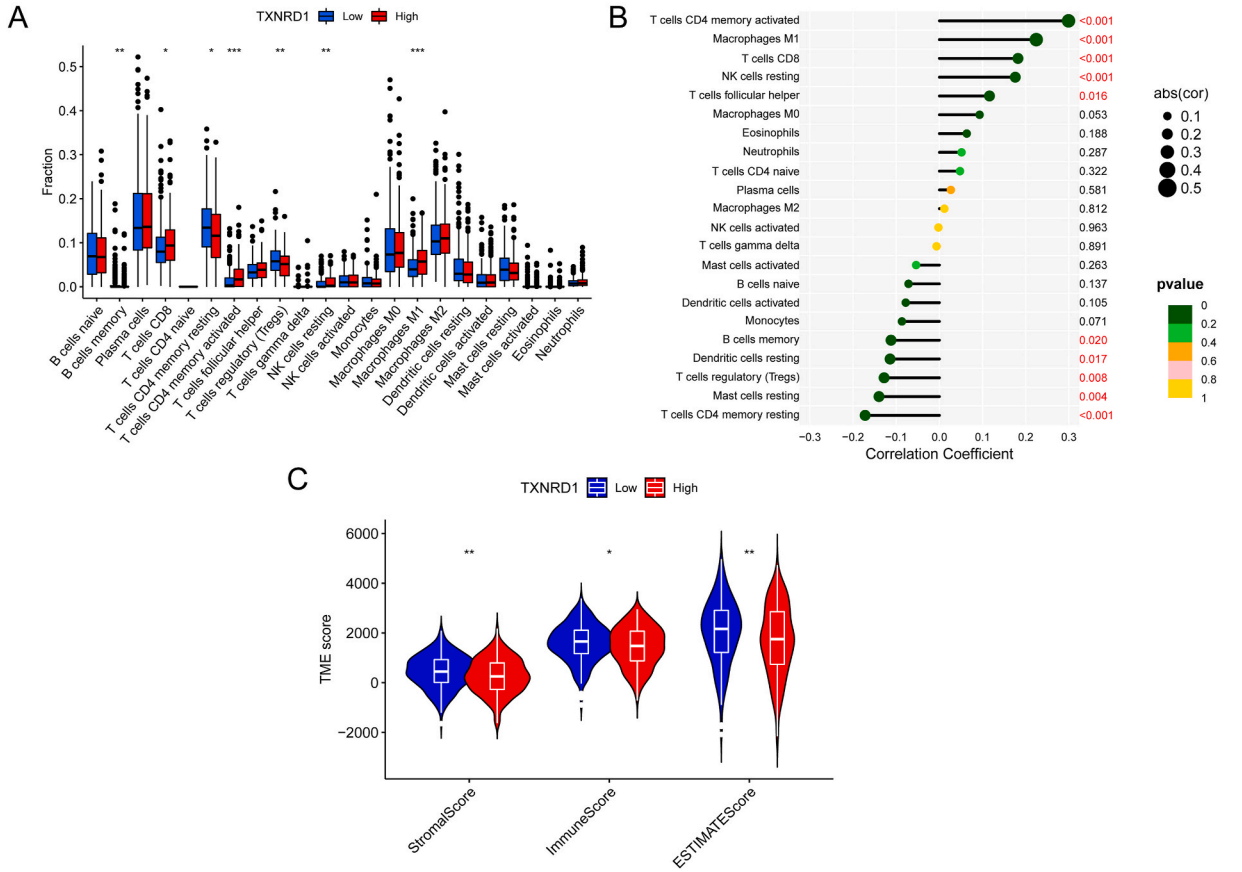

Fig. 8. Immune infiltration in tumor microenvironment. (A) Component percentage of immune cells infiltrating in groups of high and low-TXNRD1 expression. (B). Immune cell relationship with TXNRD1. (C) Subgroup differences in scores of immunity and stroma.

revealed that CEA, CYFRA21-1, SCC, NSE, and CA19-9 have limited ability to diagnose early LUAD with pGGN [42]. Besides, the five tumor biomarkers did not show significant variation in blood levels for AIS, MIA, and IAC. We speculated that early LUAD might rarely secrete tumor proteins into the bloodstream. MiR183, a tumor-associated miRNA, has been shown to be correlated with the pathological type of LUAD. It could distinguish IAC from AIS and MIA. Although it can distinguish IAC from AIS/MIA, its limited discrimination ability (AUC = 0.627) has hindered its widespread use in clinical settings [43]. Numerous novel biological markers such as insulin-like growth factor-binding protein 2, P-cadherin, and CYFRA21-1 have been demonstrated to be closely associated with micropapillary or solid components in LUAD, which showed promising predictive capabilities for poorly differentiated IAC [44,45]. In our research, we found that TrxR could serve as a biomarker for distinguishing between AIS, MIA, and IAC. The LUAD-subtype model in this study demonstrated high accuracy in identifying IAC, as evidenced by the ROC (0.897 in training cohort and 0.864 in the verification cohort) and calibration curves showing the model's stability. While the ability to distinguish between AIS and MIA is slightly insufficient, it is still considered acceptable due to the negligible difference in clinical prognosis.

As far as we know, this is the first population-based study that combining TrxR and pulmonary nodule components to predict pathology. Compared with other models used to predict the pathological properties of pulmonary nodules, our nomograms could offer distinct advantages that warrant attention. First, we have added TrxR and other blood indicators to the nomogram pioneering in order to make this model stable and reliable. Notably, the reliance solely on imaging features to predict nodule pathology, as indicated by previous models [46]. TrxR, as a potential tumor marker, could make up the absence of the model. Second, our research on TrxR was gradual. We first confirmed the predictive function of TrxR in benign and malignant aspects, and established benign-malignant model. Then the diagnostic role of TrxR in malignant LUAD (AIS, MIA, IAC) was explored, and three models including TrxR were established to distinguish them pairwise. Then a multi-classified ordered logistic model was established. Third, in the absence of external data, our study applied more robust validation method, and went back to the original data to validate the performance of the model again. The verification results of the training cohort and validation group are clear and convincing. Based on these features, we successfully explored the diagnostic efficacy of TrxR in early-stage LUAD, demonstrating that TrxR can predict pathological changes and help LUAD patients receive better treatment.

TXNRD1 mainly encodes Thioredoxin reductase 1, which is a chief component of the thioredoxin (Trx) system. Many studies have confirmed the strong correlation between TXNRD1 and tumorigenesis. Huang et al. reported that TXNRD1 knockout could

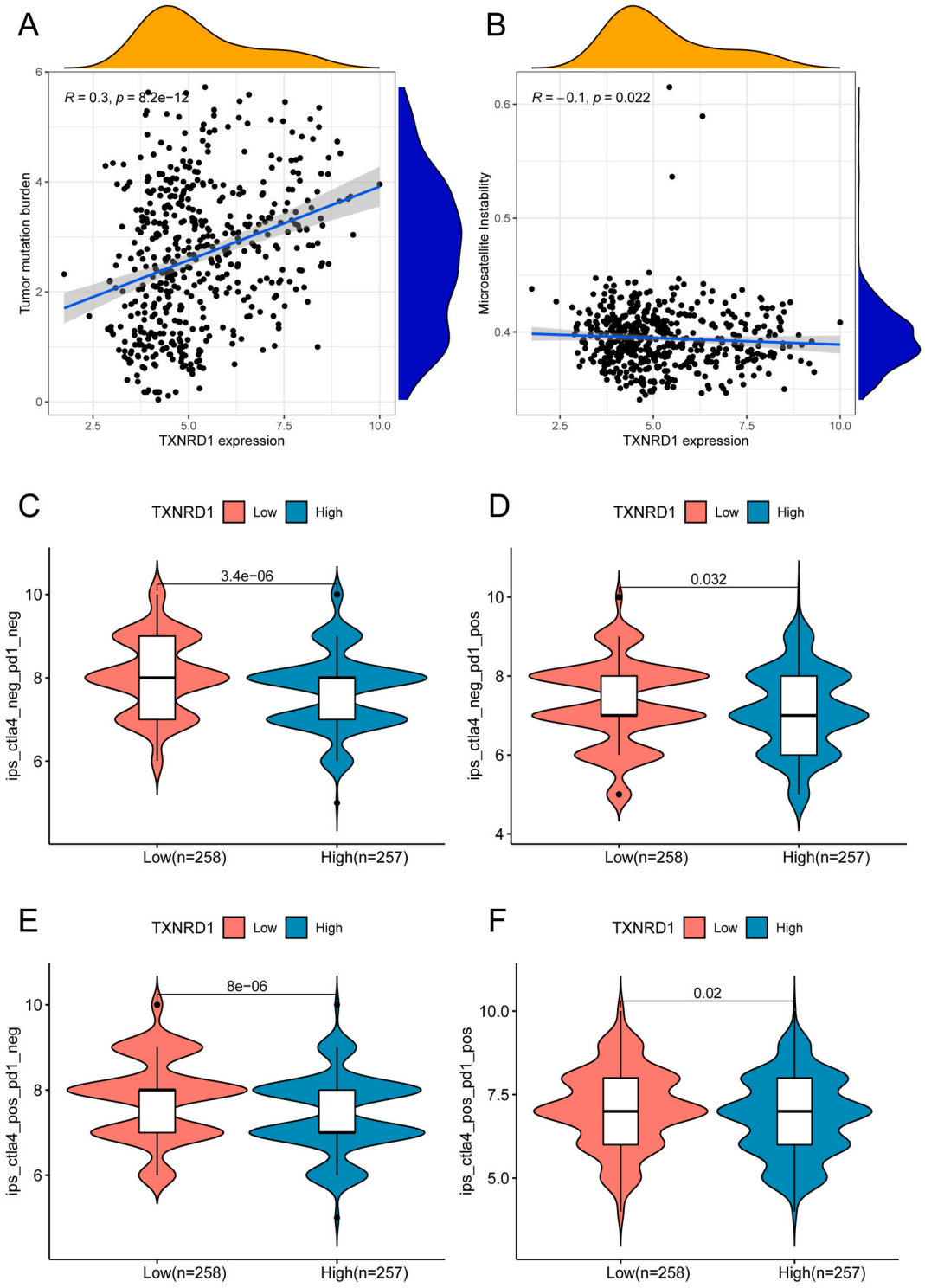

Fig. 9. Relationships between TXNRD1 expression and (A) TMB, (B) MSI. Comparison of immunophenotypic score in various groups, including (C) CTLA-4 negtive/PD-1negtive. (D) CTLA-4 negtive/PD-1 positive. (E) CTLA-4 positive/PD-1negtive. (F) CTLA-4 positive/PD-1 positive.

significantly inhibit the proliferation and metastasis of hepatocellular carcinoma cells both in vivo and in vitro, while its over-expression had the opposite effect. Furthermore, TXNRD1 dampened the interaction between Trx1 and PTEN, accelerating the degradation of PTEN, which activated Akt/mTOR signaling and target genes, ultimately promoting migration and metastasis of

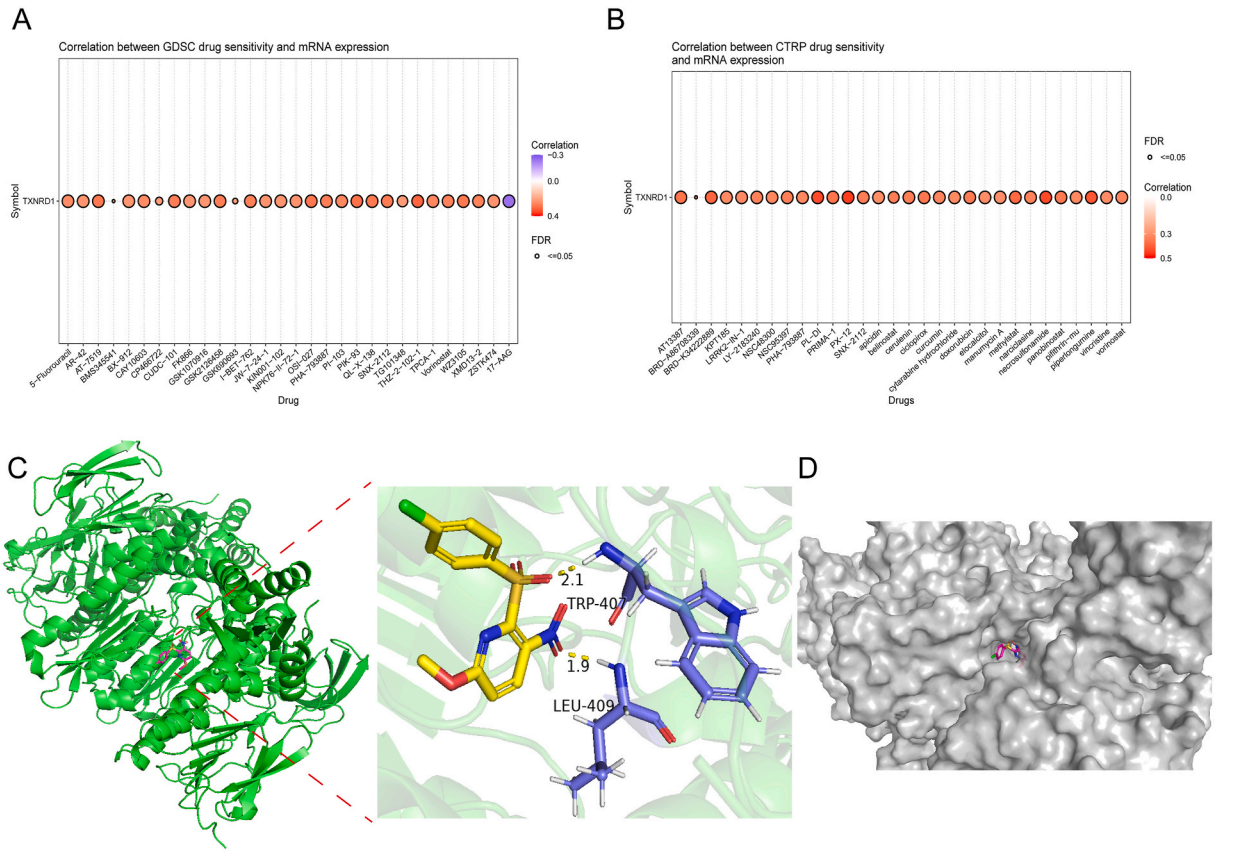

Fig. 10. Drug sensitivity and small molecule docking. (A–B) Bubble chart depicting the relationship between CTRP/GDSC drug sensitivity and TXNRD1 mRNA expression. (C–D) Molecular docking pattern diagram of TXNRD1 inhibitor 1 and thioredoxin reductase.

hepatoma cell [47].

In the TCGA cohort, comparing with normal tissues, TXNRD1 expression was significantly higher in both random and matched LUAD samples, which was consistent with prior research. Besides, TXNRD1 was expressed in disparate immune subtypes, particularly in C2(IFN-γ dominant) and C4 (lymphocyte depleted). Former studies have demonstrated that IFN-γ could play a dual role in the tumor microenvironment (TME) by both inhibiting and promoting tumor growth [48]. These findings implied a potential association between TXNRD1 expression and poor prognosis.

TME plays a critical role in cancer cell heterogeneity, thereby increasing multidrug resistance and leading to the development of tumors. Extensive evidence suggested that interactions between cancer cells and TME components promoted immune escape, proliferation, recurrence, and metastasis of tumors [49]. We found TXNRD1 expression correlated positively with infiltration of $CD8^+$ T cells, $CD4^+$ memory-activated T cells, NK resting cells, and M1 Macrophages, indicating that TXNRD1 may modulate the aggregation and regulation of tumor immune cells infiltration. In addition, transcriptome data showed that low TXNRD1 expression was related to higher TME interstitial and immune scores, suggesting a potential role of TXNRD1 in influencing immunotherapy outcomes.

In the field of tumor immunotherapy, the immune system may be more capable of recognizing and eliminating cancer cells with immune checkpoint inhibitors (ICIs) [50]. High Tumor Mutational Burden (TMB) has been established as a predictive marker for immunotherapy efficacy, indicating that T cells can easily identify tumor cells and activate an anti-tumor immune response [51]. Our study found a positive association between TXNRD1 expression and TMB, suggesting that TXNRD1 may serve as a potential predictor for immunotherapy. IPS score implied that patients with low TXNRD1 expression may exhibit improved immunotherapy sensitivity and outcomes, which required additional clinical trials for confirmation.

The identification of therapeutic biomarkers via cancer cell drug sensitivity analysis is a growing focus in precision medicine. Molecular drug reorientation is an emerging approach for disease treatment [52]. TXNRD1 contributes to drug resistance in human cancer cells against certain small molecule drugs and 2-(4-Chlorophenyl) sulfonyl-6-methoxy-3-nitropyridine could closely bind to TrxR and exert inhibitory effects.

Our findings indicate that TXNRD1 could serve as a valuable diagnostic marker. However, it is important to note the limitations. First, this is a retrospective observational study, therefore inherent selection bias is unavoidable, which suggests that further prospective comparative studies are needed. Second, this was a single-center study. As our nomograms were constructed and validated solely using data collected from our hospital, we acknowledge the potential for inherent bias due to the lack of publicly available data

from other databases. Future studies could utilize diverse and independent datasets to further validate our findings. Third, the potential interaction terms are not considered in the nomogram to improve the conciseness and interpretability. Taking all potential interaction terms in model construction may lead to better predictive performance but this will make the model more complex and difficult to use in clinical practice. Finally, additional in vivo and in vitro investigations are necessary to elucidate the underlying molecular mechanisms.

5. Conclusion

This study mainly constructed two clinical prediction models based on TrxR through clinical cohort data from our center. The former predicted benign and malignant pulmonary nodules and the latter further predicted pathological types of LUAD. Using these nomogram models, clinicians could accurately predict the pathological stage and select appropriate treatment options for LUAD patients. Finally, gene function, immune infiltration, drug sensitivity and single cell landscape were analyzed to verify the potential diagnostic value of TXNRD1, which can provide new thoughts for accurate medicine.

Funding

This work was partially supported by the Natural Science Foundation of Jiangsu Province (BK20210068), the Top Talent Support Program for Young and Middle-aged People of Wuxi Municipal Health Commission (HB2020003), the Mega-project of Wuxi Commission of Health (Z202216) and the High-end Medical Expert Team of the 2019 Taihu Talent Plan (2019-THRCTD-1).

Data availability statement

The datasets presented in this study can be found in TCGA-LUAD (The Cancer Genome Atlas; https://portal.gdc.cancer.gov/), UCSC Xena (https://xena.ucsc.edu/).

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the Ethics Committee of The Affiliated Wuxi People's Hospital of Nanjing Medical University (HS2019014).

Patient consent for publication

All patients provided signed informed consent.

CRediT authorship contribution statement

Guanyu Jiang: Writing – original draft, Methodology, Conceptualization. **Xiaokun Wang:** Writing – review & editing, Formal analysis, Data curation. **Yongrui Xu:** Data curation. **Zhao He:** Investigation. **Rongguo Lu:** Methodology. **Chenghu Song:** Data curation. **Yulin Jin:** Methodology. **Huixing Li:** Formal analysis, Data curation. **Shengfei Wang:** Formal analysis, Data curation. **Mingfeng Zheng:** Resources, Project administration. **Wenjun Mao:** Investigation, 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.

Acknowledgment

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e31864.

References

[1] K. Inamura, Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: tumor initiation and progression, Int. J. Mol. Sci. 19 (2018) E1259.
[2] A.C. Borczuk, Updates in grading and invasion assessment in lung adenocarcinoma, Mod. Pathol. 35 (2022) 28–35.

[3] M.A. Ortega, F. Navarro, L. Pekarek, et al., Exploring histopathological and serum biomarkers in lung adenocarcinoma: clinical applications and translational opportunities, Int. J. Oncol. 61 (2022) 154.
[4] M. Yotsukura, H. Asamura, N. Motoi, et al., Long-term prognosis of patients with resected adenocarcinoma in situ and minimally invasive adenocarcinoma of the lung, J. Thorac. Oncol. 16 (2021) 1312–1320.
[5] P.A. Russell, Z. Wainer, G.M. Wright, M. Daniels, M. Conron, R.A. Williams, Does lung adenocarcinoma subtype predict patient survival?: a clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification, J. Thorac. Oncol. 6 (2011) 1496–1504.
[6] F. Mohammadi, A. Soltani, A. Ghahremanloo, H. Javid, S.I. Hashemy, The thioredoxin system and cancer therapy: a review, Cancer Chemother. Pharmacol. 84 (2019) 925–935.
[7] E.S.J. Arnér, Targeting the selenoprotein thioredoxin reductase 1 for anticancer therapy, Adv. Cancer Res. 136 (2017) 139–151.
[8] J. Zhang, X. Li, X. Han, R. Liu, J. Fang, Targeting the thioredoxin system for cancer therapy, Trends Pharmacol. Sci. 38 (2017) 794–808.
[9] M. Selenius, A.-K. Rundlöf, E. Olm, A.P. Fernandes, M. Björnstedt, Selenium and the selenoprotein thioredoxin reductase in the prevention, treatment and diagnostics of cancer, Antioxidants Redox Signal. 12 (2010) 867–880.
[10] D. Mustacich, G. Powis, Thioredoxin reductase, Biochem. J. 346 (Pt 1) (2000) 1–8.
[11] M.-H. Yoo, X.-M. Xu, B.A. Carlson, Gladyshev VN and Hatfield DL: thioredoxin reductase 1 deficiency reverses tumor phenotype and tumorigenicity of lung carcinoma cells, J. Biol. Chem. 281 (2006) 13005–13008.
[12] A.P. Fernandes, A. Capitanio, M. Selenius, O. Brodin, A.-K. Rundlöf, M. Björnstedt, Expression profiles of thioredoxin family proteins in human lung cancer tissue: correlation with proliferation and differentiation, Histopathology 55 (2009) 313–320.
[13] R.L. Poerschke, P.J. Moos, Thioredoxin reductase 1 knockdown enhances selenazolidine cytotoxicity in human lung cancer cells via mitochondrial dysfunction, Biochem. Pharmacol. 81 (2011) 211–221.
[14] S. Ye, X. Chen, Y. Yao, et al., Thioredoxin reductase as a novel and efficient plasma biomarker for the detection of non-small cell lung cancer: a large-scale, multicenter study, Sci. Rep. 9 (2019) 2652.
[15] W.D. Travis, E. Brambilla, A.G. Nicholson, 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. 10 (2015) 1243–1260.
[16] K. Inamura, Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: tumor initiation and progression, Int. J. Mol. Sci. 19 (2018) 1259.
[17] C. Xiang, C. Ji, Y. Cai, et al., Distinct mutational features across preinvasive and invasive subtypes identified through comprehensive profiling of surgically resected lung adenocarcinoma, Mod. Pathol. 35 (2022) 1181–1192.
[18] T. Li, J. Fu, Z. Zeng, et al., TIMER2.0 for analysis of tumor-infiltrating immune cells, Nucleic Acids Res. 48 (2020) W509–W514.
[19] B. Ru, C.N. Wong, Y. Tong, et al., TISIDB: an integrated repository portal for tumor-immune system interactions, Bioinformatics 35 (2019) 4200–4202.
[20] D. Sun, J. Wang, Y. Han, et al., TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment, Nucleic Acids Res. 49 (2021) D1420–D1430.
[21] J.C. Stoltzfus, Logistic regression: a brief primer, Acad. Emerg. Med. 18 (2011) 1099–1104.
[22] A.S. Hess, J.R. Hess, Logistic regression, Transfusion 59 (2019) 2197–2198.
[23] L. M, Z. Hy, Z. Y, et al., A nomogram for predicting the likelihood of obstructive sleep apnea to reduce the unnecessary polysomnography examinations, Chin. Med. J. 128 (2015).
[24] S.M. Lagarde, J.B. Reitsma, A.-K.D. Maris, et al., Preoperative prediction of the occurrence and severity of complications after esophagectomy for cancer with use of a nomogram, Ann. Thorac. Surg. 85 (2008) 1938–1945.
[25] P. Charoentong, F. Finotello, M. Angelova, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (2017) 248–262.
[26] C.-J. Liu, F.-F. Hu, M.-X. Xia, L. Han, Q. Zhang, A.-Y. Guo, GSCALite: a web server for gene set cancer analysis, Bioinformatics 34 (2018) 3771–3772.
[27] Global Burden of Disease 2019 Cancer Collaboration, J.M. Kocarnik, K. Compton, et al., Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the global burden of disease study 2019, JAMA Oncol. 8 (2022) 420–444.
[28] W.D. Travis, E. Brambilla, A.G. Nicholson, 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. 10 (2015) 1243–1260.
[29] E. Chupakhin, M. Krasavin, Thioredoxin reductase inhibitors: updated patent review (2017-present), Expert Opin. Ther. Pat. 31 (2021) 745–758.
[30] X. Zhu, C. Huang, B. Peng, Overexpression of thioredoxin system proteins predicts poor prognosis in patients with squamous cell carcinoma of the tongue, Oral Oncol. 47 (2011) 609–614.
[31] K. R, N. S, P. M, et al., Immunomodulatory role of thioredoxin interacting protein in cancer's impediments: current understanding and therapeutic implications, Vaccines 10 (2022).
[32] W. Zhang, X. Zheng, X. Wang, Oxidative stress measured by thioredoxin reductase level as potential biomarker for prostate cancer, Am. J. Cancer Res. 5 (2015) 2788–2798.
[33] V. Branco, J. Pimentel, M.A. Brito, C. Carvalho, Thioredoxin, glutathione and related molecules in tumors of the nervous system, Curr. Med. Chem. 27 (2020) 1878–1900.
[34] C. Yan, B. Shieh, P. Reigan, et al., Potent activity of indolequinones against human pancreatic cancer: identification of thioredoxin reductase as a potential target, Mol. Pharmacol. 76 (2009) 163–172.
[35] R.L. Poerschke, P.J. Moos, Thioredoxin reductase 1 knockdown enhances selenazolidine cytotoxicity in human lung cancer cells via mitochondrial dysfunction, Biochem. Pharmacol. 81 (2011) 211–221.
[36] C. Zhang, J. Zhang, F.-P. Xu, et al., Genomic landscape and immune microenvironment features of preinvasive and early invasive lung adenocarcinoma, J. Thorac. Oncol. 14 (2019) 1912–1923.
[37] J. Zhu, Y. Fan, Y. Xiong, et al., Delineating the dynamic evolution from preneoplasia to invasive lung adenocarcinoma by integrating single-cell RNA sequencing and spatial transcriptomics, Exp. Mol. Med. 54 (2022) 2060–2076.
[38] M. Nie, K. Yao, X. Zhu, et al., Evolutionary metabolic landscape from preneoplasia to invasive lung adenocarcinoma, Nat. Commun. 12 (2021) 6479.
[39] S. Murakami, H. Ito, N. Tsubokawa, et al., Prognostic value of the new IASLC/ATS/ERS classification of clinical stage IA lung adenocarcinoma, Lung Cancer 90 (2015) 199–204.
[40] M. Behera, T.K. Owonikoko, A.A. Gal, et al., Lung adenocarcinoma staging using the 2011 IASLC/ATS/ERS classification: a pooled analysis of adenocarcinoma in situ and minimally invasive adenocarcinoma, Clin. Lung Cancer 17 (2016) e57–e64.
[41] K.J. Butnor, Controversies and challenges in the histologic subtyping of lung adenocarcinoma, Transl. Lung Cancer Res. 9 (2020) 839–846.
[42] F. Hu, H. Huang, Y. Jiang, et al., Discriminating invasive adenocarcinoma among lung pure ground-glass nodules: a multi-parameter prediction model, J. Thorac. Dis. 13 (2021) 5383–5394.
[43] W.-Y. Zhu, Y.-K. Zhang, Z. Chai, et al., Identification of factors for the preoperative prediction of tumour subtype and prognosis in patients with T1 lung adenocarcinoma, Dis. Markers (2016) 9354680, 2016.
[44] Z.-R. Zhao, R.W.H. Lau, H. Long, T.S.K. Mok, G.G. Chen, M.J. Underwood, Ng CSH: novel method for rapid identification of micropapillary or solid components in early-stage lung adenocarcinoma, J. Thorac. Cardiovasc. Surg. 156 (2018) 2310–2318.e2.
[45] L. Z, W. W, P. X, L. F, Z. Q, H. Z, C. L, Serum tumor markers level and their predictive values for solid and micropapillary components in lung adenocarcinoma, Cancer Med. 11 (2022).
[46] C. Jin, J. Cao, Y. Cai, L. Wang, K. Liu, W. Shen, J. Hu, A nomogram for predicting the risk of invasive pulmonary adenocarcinoma for patients with solitary peripheral subsolid nodules, J. Thorac. Cardiovasc. Surg. 153 (2017) 462–469.e1.

[47] W.-Y. Huang, Z.-B. Liao, J.-C. Zhang, et al., USF2-mediated upregulation of TXNRD1 contributes to hepatocellular carcinoma progression by activating Akt/mTOR signaling, Cell Death Dis. 13 (2022) 917.
[48] A.M. Gocher, C.J. Workman, Vignali DAA: interferon-γ: teammate or opponent in the tumour microenvironment? Nat. Rev. Immunol. 22 (2022) 158–172.
[49] P. Hernández-Camarero, E. López-Ruiz, J.A. Marchal, M. Perán, Cancer: a mirrored room between tumor bulk and tumor microenvironment, J. Exp. Clin. Cancer Res. 40 (2021) 217.
[50] M. Nishino, N.H. Ramaiya, Hatabu H and Hodi FS: monitoring immune-checkpoint blockade: response evaluation and biomarker development, Nat. Rev. Clin. Oncol. 14 (2017) 655–668.
[51] T.A. Chan, M. Yarchoan, E. Jaffee, C. Swanton, S.A. Quezada, A. Stenzinger, S. Peters, Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic, Ann. Oncol. 30 (2019) 44–56.
[52] L. Pinzi, G. Rastelli, Molecular docking: shifting paradigms in drug discovery, Int. J. Mol. Sci. 20 (2019) 4331.