



OPEN ACCESS

EDITED BY

Xiangqian Guo,
Henan University, China

REVIEWED BY

Panpan Liu,
Sun Yat-Sen University Cancer Center
(SYSUCC), China
Jiateng Zhong,
Xinxiang Medical University, China

*CORRESPONDENCE

Shupeng Zhang,
✉ tizhangsp@126.com

SPECIALTY SECTION

This article was submitted to Cancer
Genetics and Oncogenomics,
a section of the journal
Frontiers in Genetics

RECEIVED 03 December 2022

ACCEPTED 16 March 2023

PUBLISHED 06 April 2023

CITATION

Li Z, Cui Y, Zhang S, Xu J, Shao J, Chen H,
Chen J, Wang S, Zeng M, Zhang H, Lu S,
Qian ZR and Xing G (2023), Novel
hypoxia-related gene signature for
predicting prognoses that correlate with
the tumor immune microenvironment
in NSCLC.

Front. Genet. 14:1115308.

doi: 10.3389/fgene.2023.1115308

COPYRIGHT

© 2023 Li, Cui, Zhang, Xu, Shao, Chen,
Chen, Wang, Zeng, Zhang, Lu, Qian and
Xing. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Novel hypoxia-related gene signature for predicting prognoses that correlate with the tumor immune microenvironment in NSCLC

Zhaojin Li¹, Yu Cui¹, Shupeng Zhang^{1*}, Jie Xu¹, Jianping Shao¹,
Hekai Chen¹, Jingzhao Chen², Shun Wang², Meizhai Zeng²,
Hao Zhang², Siqian Lu², Zhi Rong Qian² and Guoqiang Xing¹

¹Department of General Surgery, Tianjin Fifth Central Hospital, Tianjin, China, ²Beidou Precision Medicine Institute, Guangzhou, China

Background: Intratumoral hypoxia is widely associated with the development of malignancy, treatment resistance, and worse prognoses. The global influence of hypoxia-related genes (HRGs) on prognostic significance, tumor microenvironment characteristics, and therapeutic response is unclear in patients with non-small cell lung cancer (NSCLC).

Method: RNA-seq and clinical data for NSCLC patients were derived from The Cancer Genome Atlas (TCGA) database, and a group of HRGs was obtained from the MSigDB. The differentially expressed HRGs were determined using the limma package; prognostic HRGs were identified via univariate Cox regression. Using the least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression, an optimized prognostic model consisting of nine HRGs was constructed. The prognostic model's capacity was evaluated by Kaplan–Meier survival curve analysis and receiver operating characteristic (ROC) curve analysis in the TCGA (training set) and GEO (validation set) cohorts. Moreover, a potential biological pathway and immune infiltration differences were explained.

Results: A prognostic model containing nine HRGs (STC2, ALDOA, MIF, LDHA, EXT1, PGM2, ENO3, INHA, and RORA) was developed. NSCLC patients were separated into two risk categories according to the risk score generated by the hypoxia model. The model-based risk score had better predictive power than the clinicopathological method. Patients in the high-risk category had poor recurrence-free survival in the TCGA (HR: 1.426; 95% CI: 0.997–2.042; $p = 0.046$) and GEO (HR: 2.4; 95% CI: 1.7–3.2; $p < 0.0001$) cohorts. The overall survival of the high-risk category was also inferior to that of the low-risk category in the TCGA (HR: 1.8; 95% CI: 1.5–2.2; $p < 0.0001$) and GEO (HR: 1.8; 95% CI: 1.4–2.3; $p < 0.0001$) cohorts. Additionally, we discovered a notable distinction in the enrichment of immune-related pathways, immune cell abundance, and immune checkpoint gene expression between the two subcategories.

Abbreviation: AUC, area under the curve; HRGs, hypoxia-related genes; LASSO, least absolute shrinkage and selection operator; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NSCLC, non-small cell lung cancer; ROC, receiver operating characteristic.

Conclusion: The proposed 9-HRG signature is a promising indicator for predicting NSCLC patient prognosis and may be potentially applicable in checkpoint therapy efficiency prediction.

KEYWORDS

NSCLC, hypoxia, prognosis, immune microenvironment, immunotherapy

1 Introduction

Lung cancer ranks highly among all malignancies in its prevalence and mortality. Non-small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancer diagnoses (Siegel et al., 2023). Although great advances have been made in the research and treatment of NSCLC, the majority of patients are diagnosed at an advanced stage, and the efficacy of NSCLC treatment remains unsatisfactory. NSCLC diagnosis and prognosis are currently determined by tumor stage and histopathological assessment (Cuyun Carter et al., 2014; Wood et al., 2022). Unfortunately, the accuracy of these standard diagnostic procedures is insufficient, and the resulting illness outcomes in patients with a uniform clinical diagnosis can also be inconsistent (Balachandran et al., 2015). Consequently, the development of a unique, accurate, and sensitive signature for early NSCLC detection and prognostic prediction is urgently needed.

Because the demand for oxygen in the tumor is greater than the body supplies, hypoxia has been discovered to accelerate tumor development with the induction of a hypoxic tumor context (Petrova et al., 2018). The hypoxic environment within tumors is one of their most important hallmarks. Tumor formation and incidence are frequently accompanied by a variety of adaptive changes involving transcription factor activation, cell proliferation, motility, apoptosis, and other signaling pathways through which hypoxia might enhance tumor aggressiveness and medication resistance (Gilkes et al., 2014; Muz et al., 2015; Rankin et al., 2016; Nobre et al., 2018; Jing et al., 2019).

With the popularization of transcriptome sequencing technology, an increasing number of studies have shown that hypoxic tumor microenvironments are associated with worse clinical presentation and prognosis (Magnon et al., 2007; D'Ignazio et al., 2017). Previous studies have demonstrated that hypoxia-related genes (HRGs) are linked to prognosis in LUAD patients with disease stages I–II (Sun et al., 2020). However, clinical outcomes, oncogenic pathways, and treatment responses remain unclear in NSCLC patients.

Prior studies have verified the critical roles of hypoxia in promoting immunological escape and tumor immune suppression. For example, hypoxia stimulates the recruitment of immunosuppressive cytokines and suppressive cells, consequently hampering immune effector cells and prompting immunological escape (Qian et al., 2019). Because the codependency between the immunological state and hypoxia in the tumor microenvironment could lead to alterations in immune activity, treatment response, and prognosis in NSCLC, an integrated analysis of the immunological state and hypoxia might have promising prognostic utility and provide supplementary knowledge to facilitate the development of transformational studies on and treatment strategies for NSCLC.

We thus extracted data from multiple independent NSCLC cohorts to screen candidate HRGs and established a related signature, with the aim of determining its capacity to predict

clinical outcomes, TME features, and therapeutic efficacy in NSCLC patients.

2 Methods

The overall research strategy, including the building and verification of the HRG risk model, is depicted in Figure 1.

2.1 Data preparation and processing

From the UCSC Xena database, NSCLC clinical data and transcriptome expression profiles for LUAD and LUSC patients in The Cancer Genome Atlas (TCGA) were downloaded. A total of 1,011 tumor and 108 normal samples were employed for HRG expression differential analysis and as a training set for the construction of the risk score model. All scores indicating the degree of hypoxia in the tumor, such as the Winter, Buffa, and Ragnum hypoxia scores (Winter et al., 2007; Buffa et al., 2010; Ragnum et al., 2015), were obtained from TCGA. The higher these scores, the more hypoxic is the tumor.

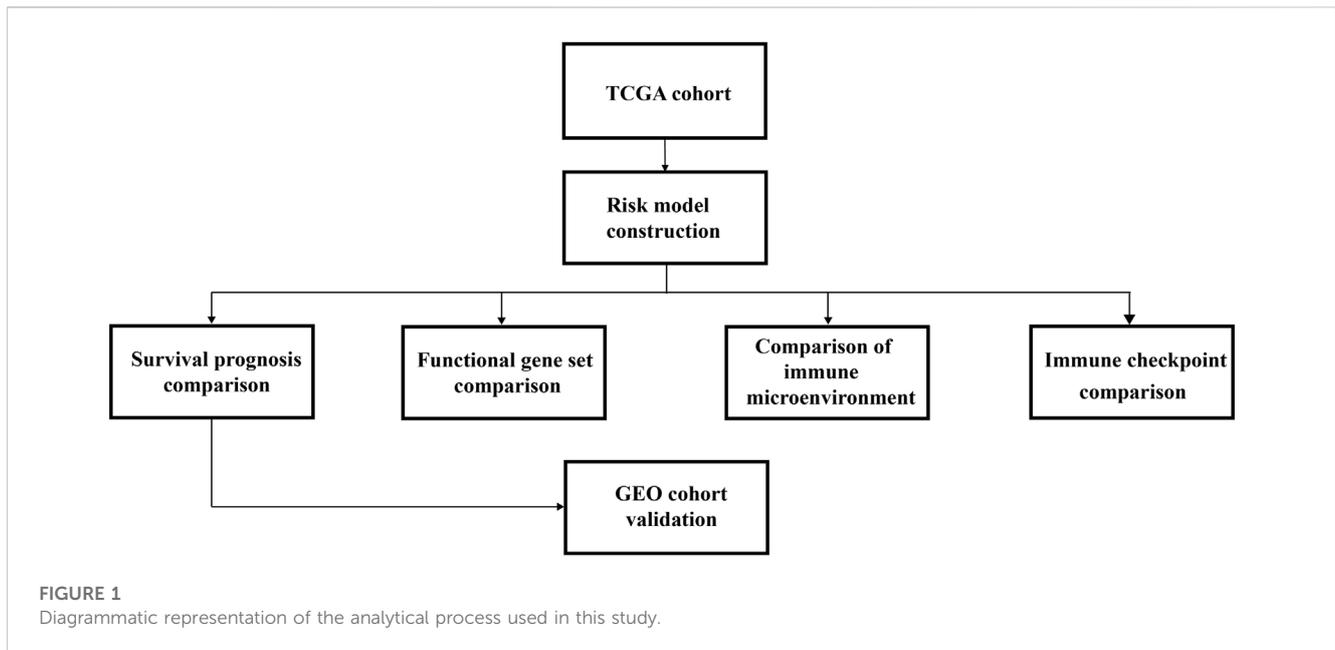
For external validation, four microarray datasets together with detailed clinical characteristics were retrieved from the Gene Expression Omnibus (GEO) (GSE30219, GSE31210, GSE37745, and GSE50081) using the GEOquery R package. A total of 628 tumor samples were acquired, and the clinical characteristics are listed in Supplementary Table S1. Batch effects were eliminated using the sva package (Leek et al., 2012), and the scale function was used to standardize the expression value before model construction and validation.

2.2 Identification of differentially expressed HRGs

A total of 190 of 200 HRGs and available data were retrieved from the MSigDB Hallmark database (Supplementary Table S2), and the expression of these genes was extracted from integrated datasets (Liberzon et al., 2015). Differential analysis between tumor and normal samples of 190 HRGs was performed using the limma package in the TCGA database. Genes with p values less than 0.05 were defined as differentially expressed HRGs (DEHRGs). Volcano mapping was then performed using the ggplot2 package.

2.3 Building and verifying an HRG-based risk model

We conducted univariate Cox regression analysis based on overall survival to determine prognostic DEHRGs ($p < 0.05$). The DEHRGs



related to OS were identified by least absolute shrinkage and selection operator (LASSO) regression analysis using the glmnet package to shrink the gene set associated with prognosis. To achieve the optimal model, we combined the remaining genes step by step and tested performance. Ultimately, nine genes were selected, and a multivariate Cox regression analysis was adopted to build a prognostic risk model. The algorithm used to determine the patient's risk level was as follows:

$$\begin{aligned}
 \text{Risk score} = & 0.10515 \times STC2_{exp} + 0.0234 \times ALDOA_{exp} \\
 & + 0.02207 \times MIF_{exp} + 0.13426 \times LDHA_{exp} \\
 & + 0.09928 \times EXT1_{exp} - 0.01151 \times PGM2_{exp} \\
 & - 0.1475 \times ENO3_{exp} - 0.16294 \times INHA_{exp} \\
 & - 0.14694 \times RORA_{exp}
 \end{aligned}$$

A risk score was then obtained for all patients. Patients were separated into two risk categories by using the median risk score of patients.

To determine the predictive ability of this model for tumor recurrence, recurrence-free survival was compared between two risk categories using stage I and II cases in the TCGA and GEO databases.

Overall survival (OS) was also compared between the two risk categories using all patients. The survivalROC R package was employed to plot the receiver operating characteristic (ROC) curves of 1-, 3-, and 5-year OS in the training cohort, and the same method was used in the validation cohort.

2.4 Independent determination of risk models and construction of clinical nomograms

Multivariate Cox regression analysis was applied to construct an independent prognostic model. A nomogram was used to predict patient survival rates, and the performance was verified by drawing

calibration curves. ROC curves and the concordance index were introduced to detect the reliability of the HRG model. Clinical features were incorporated into multivariate Cox regression to determine whether the risk model was an independent indicator of prognostic prediction.

2.5 Functional enrichment analysis

The gene sets of the Kyoto Encyclopedia of Genes and Genomes and the Hallmark gene sets were downloaded from MSigDB. The enrichment scores of individual patients were computed with the GSVA R package. Using the limma package, the enrichment score was then compared between two TCGA risk categories, and the difference in HRG expression between the two categories was also analyzed. A cluster heatmap was used to show the top 30 genes with the most significant *p* values.

2.6 Immunological features of the tumor microenvironment

Immunomodulators, a group of immunoregulatory genes containing antigen-presenting factors, ligands, and receptors, play crucial roles in cancer immunotherapy, and the expression of HRGs and immunomodulators was compared with the Spearman correlation method.

In addition, the differences in their infiltration abundance in the tumor microenvironment affected the patients' outcomes and the efficiency of immunotherapy. Therefore, the ssGSEA algorithm was employed to compute immune cell abundance in individual patients. Using the Spearman correlation method, the correlation between the hypoxia-related model and immune cells was analyzed, and, finally, the difference in immune cell abundance between the two risk categories was assessed.

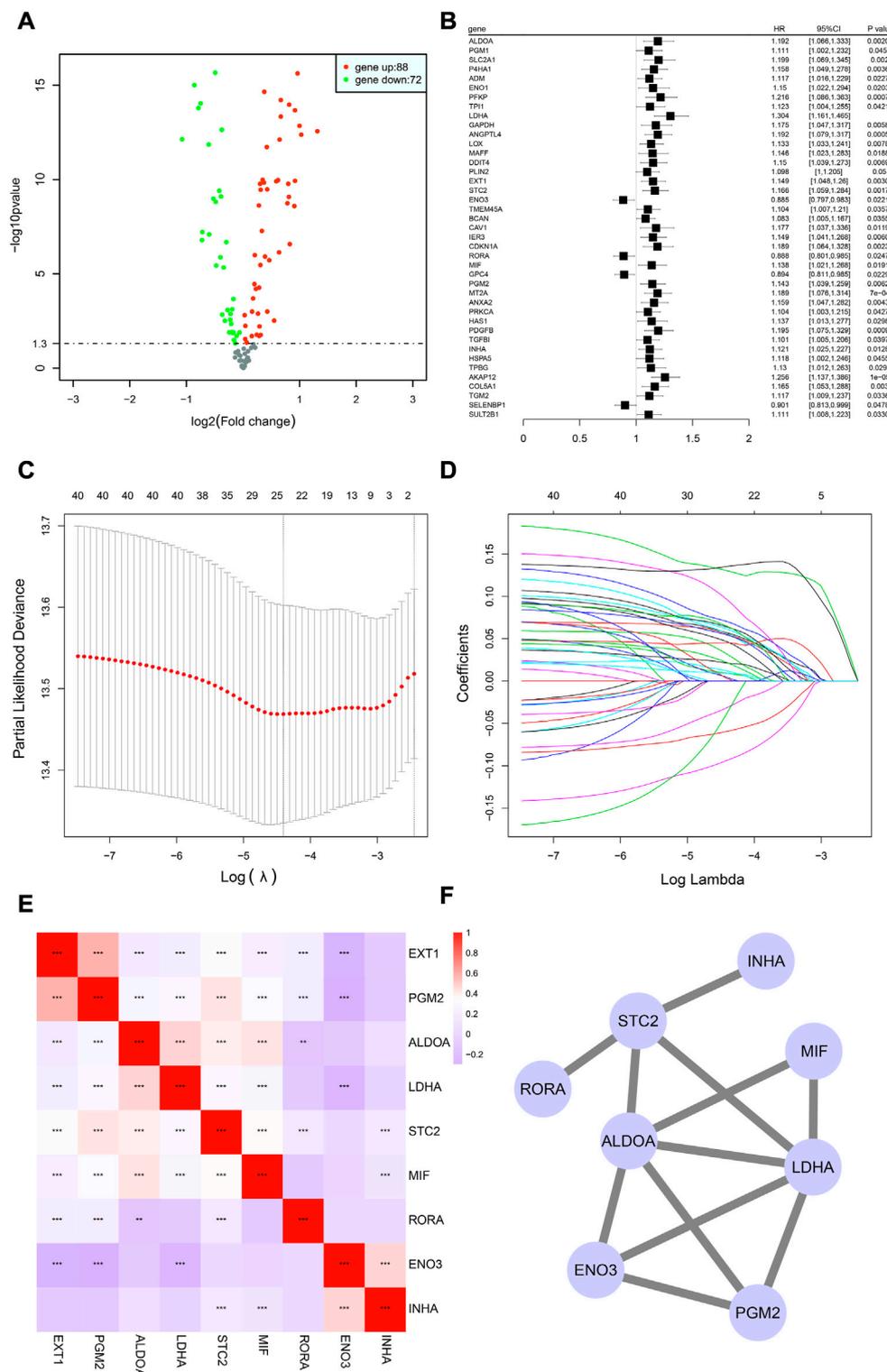


FIGURE 2 Building of the HRG model in the TCGA-NSCLC cohort. **(A)** Volcano graphic showing how HRG expression differs in tumor and normal samples based on both fold change and statistical significance. **(B)** Results of univariate Cox regression analysis presented in a forest plot, with gene name in the first column and box plot depicting the hazard ratio (HR) and its corresponding 95% confidence interval in the second column. **(C, D)** LASSO analysis for estimating the number of contributing components. **(E)** Correlation analysis of HRG expression in the risk model using Spearman's correlation. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. **(F)** Results of protein interaction analysis of the HRG model.

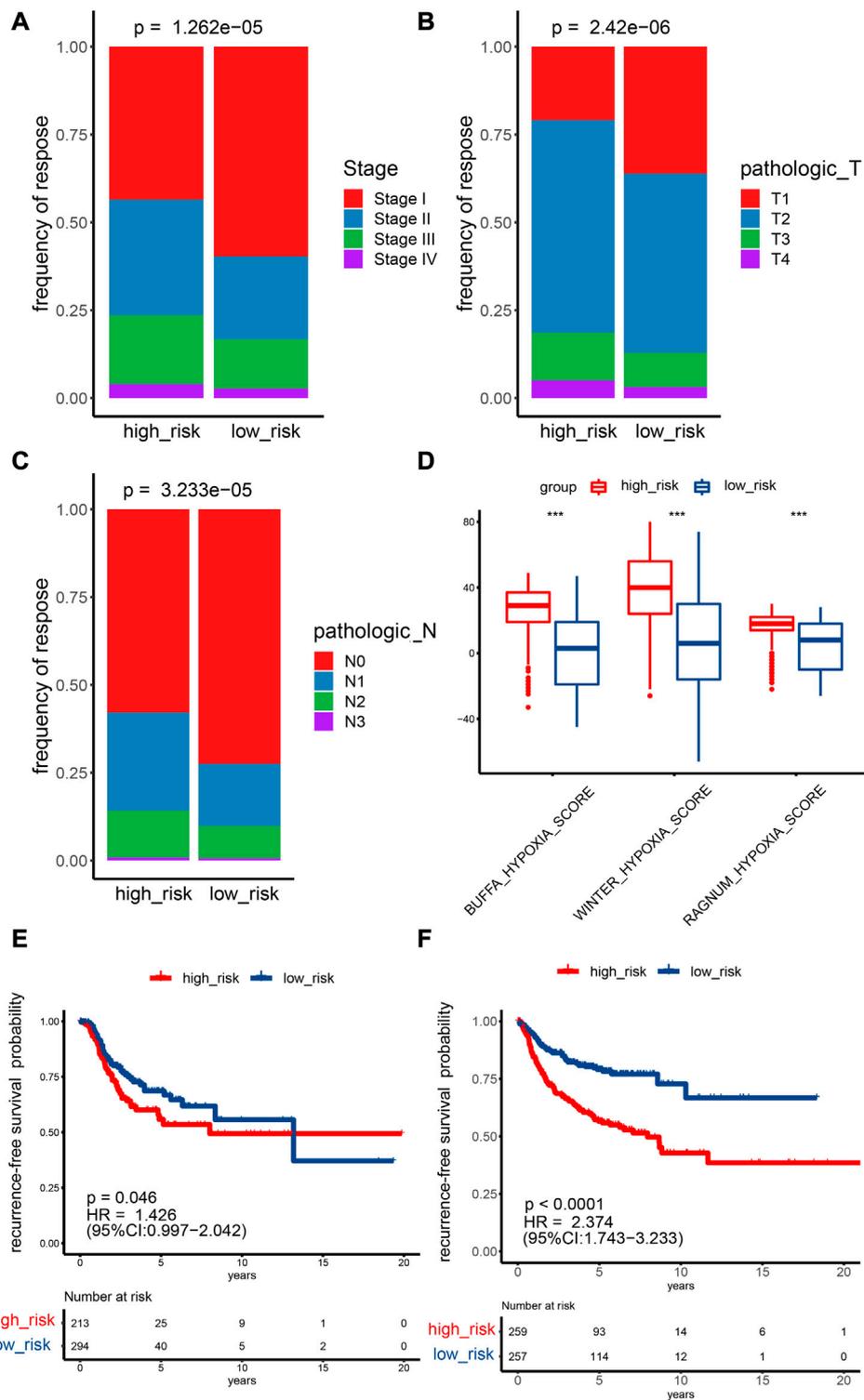


FIGURE 3

Prediction ability of the HRG model for the recurrence and stage of NSCLC patients. (A–C) TCGA cohort Fisher test outcomes for T and N stages and pathological stages comparing two risk categories. (D) Analyzing the differences between two risk categories on three known hypoxia scores. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. (E,F) Recurrence survival curves for two risk categories. E: TCGA cohort, F: GEO cohort.

Immunomodulator-related genes and immune cell-type marker genes were obtained from the TCGA immune response working group (Charoentong et al., 2017; Thorsson

et al., 2018). The immunological score was computed with the estimate package, and we compared the difference between the two risk categories.

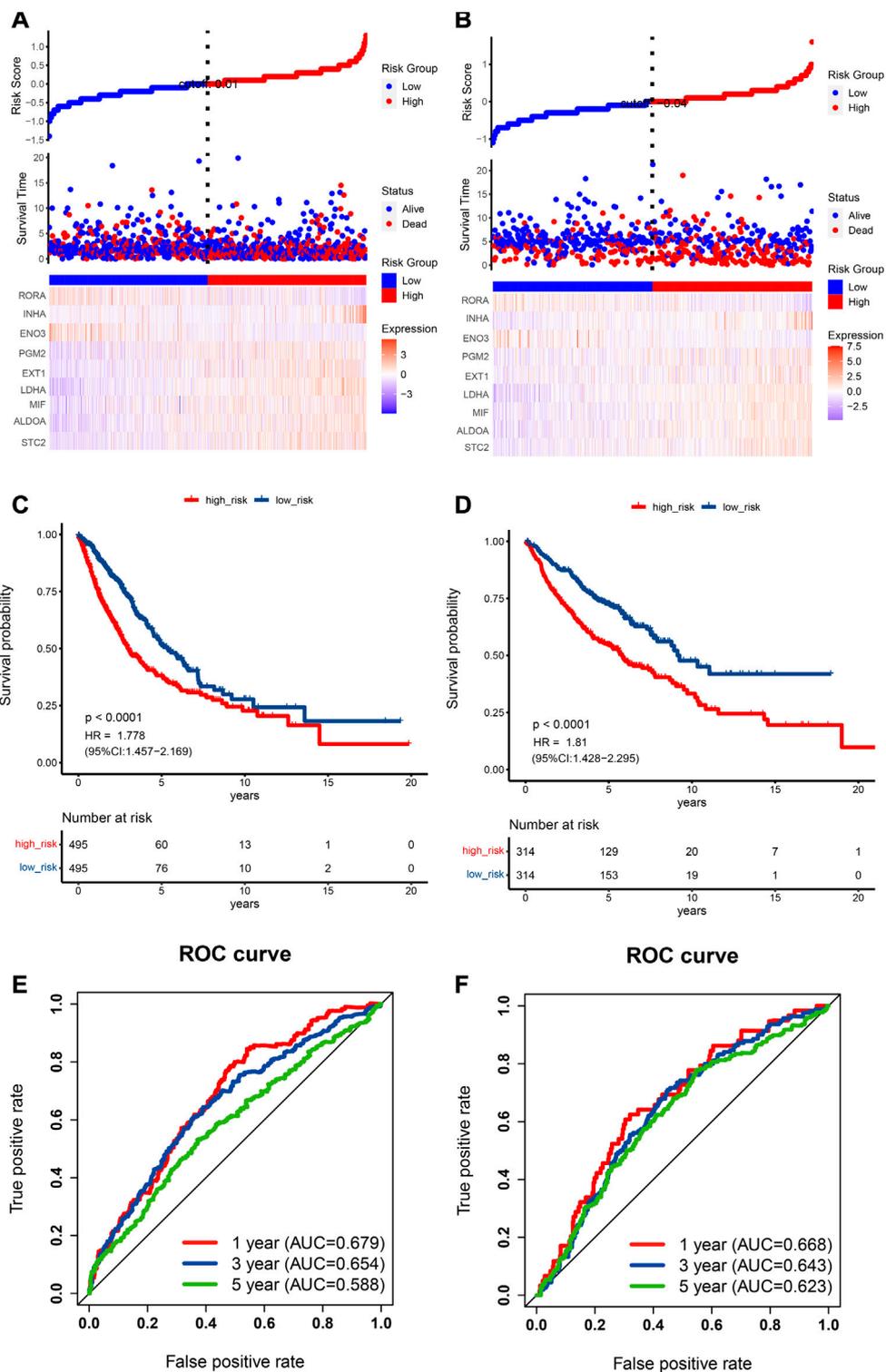


FIGURE 4

Prognostic data analysis from the training and validation cohorts as well as verification of the model's precision. (A, B) The risk score curve is displayed in the top part. The distribution of the risk score, survival duration, and patient status are presented in the middle part. A heatmap of HRGs in the classifier is shown in the bottom part. (A) TCGA cohort, (B) GEO cohort. (C, D) NSCLC patient Kaplan-Meier survival curve comparing two risk categories. (C) TCGA cohort, (D) GEO cohort. (E, F) ROC curve in the TCGA cohort (E) and GEO cohort (F).

2.7 Statistical analysis

Kaplan–Meier curves and the log-rank test were used to compare the survival rate between various subcategories. The independent prognostic significance of the clinical features in OS was detected by univariate and multivariate Cox proportional hazard regression analyses. The prognostic prediction significance of the risk models for 1/3/5-year OS was evaluated by ROC curves and area under the curve (AUC) values. Differences in TMN stage and clinicopathologic stage composition between the two risk categories of the TCGA cohort were analyzed by Fisher's test. The correlation analysis of HRGs with immune cell abundance and with immune regulator-related genes was performed using Spearman's correlation analysis. The Wilcoxon test was applied to compare immune cell infiltration. p values less than 0.05 were defined as significant. R 3.6.1 was used to conduct all the analyses.

3 Results

3.1 Differential expression analysis of HRGs in the training cohort

Genes that are expressed normally in both tumor and normal tissues do not typically play a significant role in the development and progression of tumors. Therefore, it is necessary to exclude such genes from the analysis beforehand.

Differential analysis between 1,011 tumor and 108 normal samples on 190 HRGs was performed, and 160 genes were significantly differentially expressed (Supplementary Table S3) in the TCGA cohort. Of these, 72 genes were downregulated in tumor tissues, while 88 were upregulated. The differentially expressed volcano map is shown in Figure 2A.

The analysis revealed that a majority of HRGs showed significant differences in expression, highlighting the importance of the hypoxia pathway as a mechanism underlying tumor growth and metastasis.

3.2 Building of the HRG model in the training cohort

Although there may be aberrant gene expression, it may not necessarily affect a patient's survival status or time. Thus, the identification of hypoxia-related genes that are linked to prognosis is crucial.

In total, 990 NSCLC cases in the TCGA cohort with detailed OS information were identified. To further identify prognosis-related DEHRGs, 160 DEHRGs were analyzed using univariate Cox regression analysis, and 41 genes were significantly correlated with prognosis (Supplementary Table S4). A forest map was used to display the prognosis-related HRGs' hazard ratios and the matching confidence intervals (CIs), which demonstrated that the majority of candidates were at risk (Figure 2B). To ensure the stability and feasibility of the model, LASSO regression analysis was performed with the optimum λ value to reduce the scope of candidates, and 25 DEHRGs were identified (Figures 2C, D). Finally, by performing multivariate Cox regression, nine genes (STC2, ALDOA, MIF, LDHA,

EXT1, PGM2, ENO3, INHA, and RORA) were selected to construct a risk prognostic model. In the TCGA training cohort, patients were split into two risk categories based on the median risk score. Notably, compared to patients in the low-risk category, those in the high-risk category were more likely to experience relapse (HR: 1.426; 95% CI: 0.997–2.042; $p = 0.046$; Figure 3E).

A linear regression analysis was conducted to determine the interaction of these nine genes; the results indicated that 25 candidates exhibited high correlation, such as ALDOA and LDHA, ALDOA and STC2, ALDOA and MIF, and ENO3 and EXT1 (Figure 2E). In addition, protein interaction analysis was performed on these nine genes using the STRING online database and visualized by Cytoscape. We observed that ALDOA, LDHA, ENO3, and PGR2 strongly interacted with multiple genes—in the end, these four candidates were regarded as the core genes (Figure 2F).

Fisher's test showed that the percentage of stage I samples, T1 samples, and N0 samples in the low-risk category (stage: 59.7%; T: 36.1%; N: 72.5%) was greater than the percentage in the high-risk category (stage: 43.5%; T: 20.9%; N: 57.9%) (Figures 3A–C). The difference in M0 staging in the low-risk category was not significant in comparison with the low-risk category (Supplementary Figure S1). A higher risk score indicates higher expression levels of most HRGs, which are associated with a more advanced disease stage, a larger tumor size, and increased lymph node involvement. One possible explanation for the higher proportion of stage I, T1, and N0 samples in the low-risk category compared with the high-risk category is that HRGs are expressed more strongly in later-stage and more invasive tumors, which may lead to higher risk scores and worse prognoses. In addition, the lack of a significant difference in M0 staging between the low- and high-risk categories may be due to distant metastasis being less affected by hypoxia than local tumor growth and invasion. The Buffa, Winter, and Ragnum hypoxia scores in the high-risk category were greater than those in the low-risk category (Figure 3D), implying that the established model could effectively evaluate the degree of hypoxia.

Moreover, the scatter plot of OS status showed that, in comparison with low-risk category patients, more deaths were observed in the high-risk category (Figure 4A). More importantly, high-risk category patients ($n = 495$) had a worse prognosis than low-risk patients ($n = 495$) ($p < 0.0001$; HR: 1.778; 95% CI: 1.457–2.169; Figure 4C). To accurately evaluate the prognostic performance of the nine-gene signature, ROC analysis was conducted using the 1-, 3-, and 5-year cutoff values of OS—the AUC values were 0.679, 0.654, and 0.588, respectively, indicating that our prognostic model had good predictive performance in patients with NSCLC (Figure 4E).

Early-stage cancer patients are more highly represented in the low-risk category based on HRGs than in the high-risk category. Thus, when we specifically only analyzed early-stage patients, we discovered that significant differences in overall survival still existed between high- and low-risk categories in early-stage cancer prognostic indicators (Supplementary Figure S2). This finding suggests that risk stratification based on HRGs can still be used to effectively predict the overall survival of NSCLC patients, even when analyzing only the subset of patients in the early stages of the disease.

To summarize, a model based on the expression of HRGs that are prognostically relevant can effectively assess the extent of

TABLE 1 The clinical pathological information of TCGA and GEO cohorts.

	TCGA cohort (n = 964)	GEO cohort (n = 628)
Gender		
Female	389 (40.4%)	286 (45.5%)
Male	575 (59.6%)	342 (54.5%)
Age (years)		
<45	21 (2.2%)	11 (1.8%)
≥75	181 (18.8%)	74 (11.8%)
45–60	199 (20.6%)	185 (29.5%)
60–75	563 (58.4%)	358 (57.0%)
Overall survival time		
Mean (SD)	2.52 (2.52)	4.78 (3.23)
Median [min, max]	1.78 [0.00274, 19.9]	4.61 [0.0164, 21.3]
Clinical stage		
Stage I	498 (51.7%)	446 (71.0%)
Stage II	272 (28.2%)	150 (23.9%)
Stage III	162 (16.8%)	28 (4.5%)
Stage IV	32 (3.3%)	4 (0.6%)
Recurrence-free survival time		
Mean (SD)	2.66 (2.63)	4.26 (3.29)
Median [min, max]	1.79 [0.0110,19.9]	4.31 [0.160,21.3]
Missing	391 (40.6%)	90 (14.3%)

hypoxia, exhibit a close correlation with clinical staging, and reveal significant differences in recurrence, survival, and overall survival between high- and low-risk categories.

3.3 External validation of the HRG prognostic model

External validation plays a crucial role in improving the reliability and reproducibility of scientific research. By confirming the generalizability and applicability of research findings, external validation can help to identify and correct potential problems and limitations in study design and methodology.

To evaluate the prognostic capability of this novel HRG-based risk model, external cohorts were obtained from the GEO. The risk score was computed *via* the identical formula described before. The patients were separated into two risk categories using the median score. In recurrence-free survival analysis, patients in the high-risk category had a higher tendency to relapse than low-risk patients (HR: 2.374; 95% CI: 1.743–3.233; $p < 0.0001$; Figure 3F).

Scatter plots for death events are displayed in Figure 4B. High-risk category patients showed higher mortality rates than patients in the low-risk category. Additionally, we confirmed the performance of our HRG risk model for OS in the GEO cohort. As expected, a worse OS was indicated by a higher risk score ($p < 0.0001$, HR: 1.81, 95% CI: 1.428–2.295; Figure 4D). Figure 4F shows the 1-, 3-, and 5-year cutoff OS values in the validation dataset; the AUC values were 0.668, 0.643, and 0.623, respectively. ROC analysis further verified the effectiveness and reliability of the HRG risk model for prognostic prediction in NSCLC patients.

In our external dataset validation analysis, we found that our model demonstrated strong predictive performance for both overall survival and recurrence-free survival, indicating that it is both stable and generalizable.

3.4 Analysis of independent prognostic factors in the HRG risk model

Patients with complete clinical information were identified to construct risk prognostic models; Table 1 displays the clinical information. Along with risk scores, clinical stage, age, and sex were considered in a multivariate Cox regression analysis to determine independent prognostic markers. According to the outcome, the risk score was an independent prognostic factor in the TCGA (HR: 0.58, 95% CI: 0.47–0.71; Figure 5A) and GEO cohorts (HR: 0.65, 95% CI: 0.50–0.84; Figure 5B).

We also calculated the conformance index (C-index) of these three clinical factors and risk score within 1–5 years of OS time. For the TCGA cohort, the results suggested that the C-index of the risk score was greater than that of other clinical factors (Figure 5C), and that the C-index had a similar numerical approximation between the risk score and clinicopathological stage in the GEO cohort (Figure 5D).

The ROC curve of the risk score was detected with these three clinical factors for the 3-year OS. The AUC value of the risk score was greater than that of the other three clinical characteristics in both cohorts (Figures 5E, F).

Taken together, even when other factors that may affect patient prognoses were taken into account, the risk score had

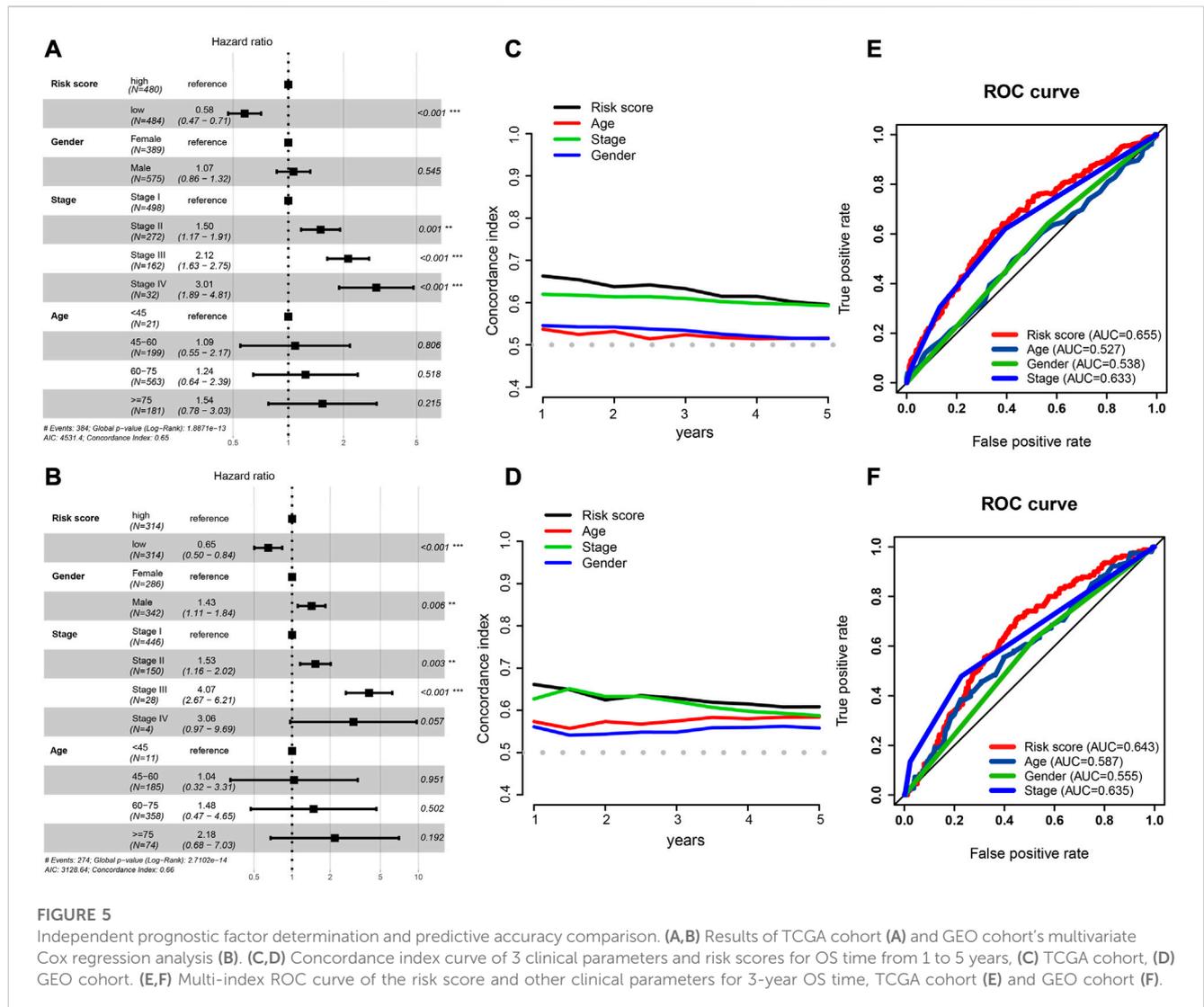


FIGURE 5

Independent prognostic factor determination and predictive accuracy comparison. (A,B) Results of TCGA cohort (A) and GEO cohort's multivariate Cox regression analysis (B). (C,D) Concordance index curve of 3 clinical parameters and risk scores for OS time from 1 to 5 years. (E,F) Multi-index ROC curve of the risk score and other clinical parameters for 3-year OS time, TCGA cohort (E) and GEO cohort (F).

a significant and distinct impact on patient outcomes. It performed better than other clinical parameters in predicting overall survival.

3.5 Building and verifying the prognostic nomogram system

To assist clinical practitioners with more accurate predictions of patient survival rates during treatment and management, we constructed a clinical nomogram that serves as an important auxiliary decision-making tool.

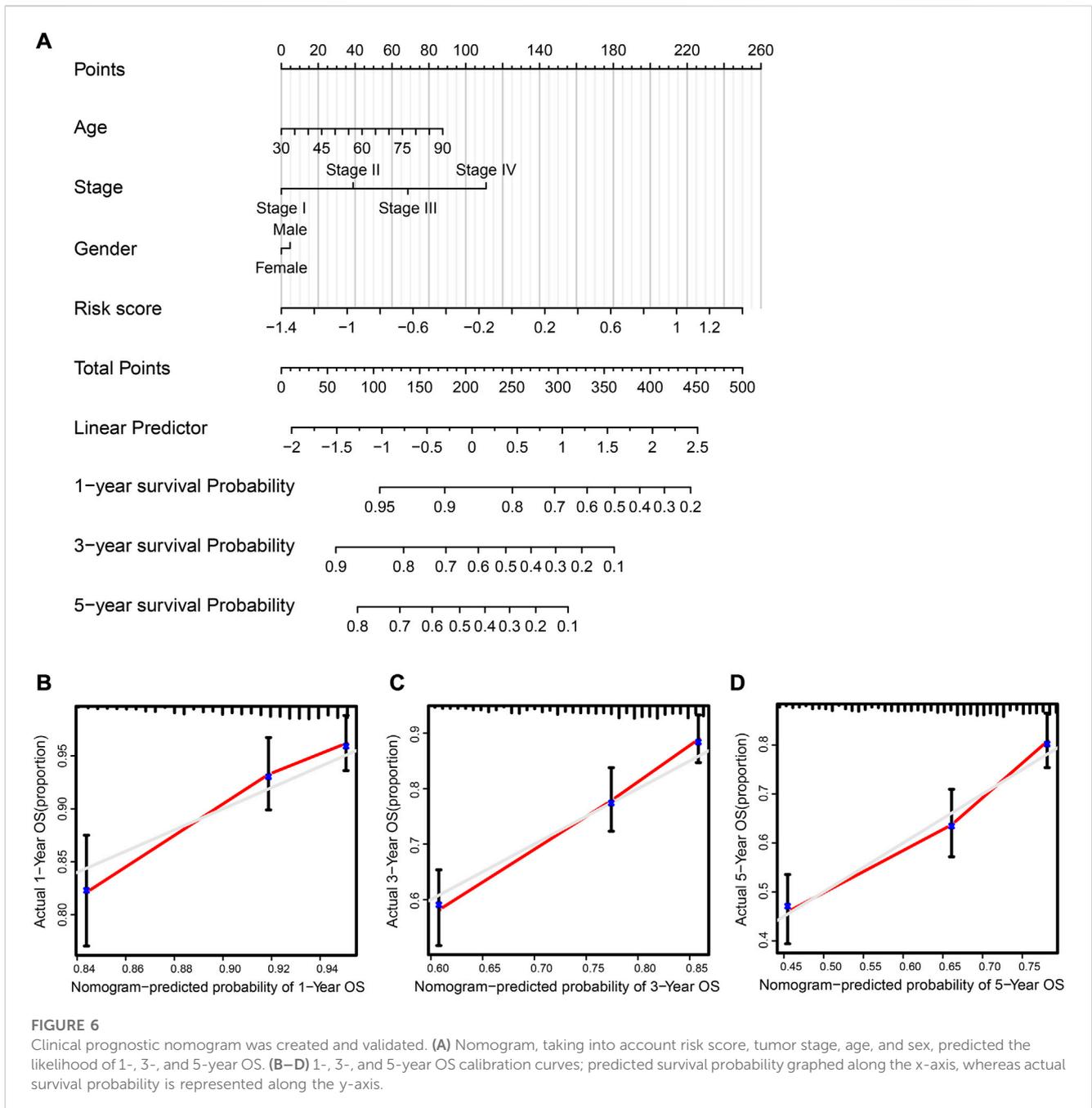
Using the “rms” and “survival” R packages, we generated a nomogram containing the risk score and valuable clinical parameters predicting 1-, 3-, and 5-year overall survival of NSCLC patients in the TCGA dataset and a calibration plot that analyzed estimated and actual overall survival to assess the effectiveness of the prognostic nomogram (Figure 6A). The calibration plot showed that the nomogram's estimated survival

and actual survival rates were highly correlated (Figures 6B–D). It is possible to evaluate the probability of a patient surviving at a certain time according to the patient's clinical information and gene expression levels using this clinical prognosis nomogram. In conclusion, this nomogram system for predicting the overall survival of patients based on gene expression and clinical information is reliable.

3.6 Biological pathway analysis related to the HRG risk model

Different biological mechanisms often lead to distinct prognostic outcomes in cancer patients. KEGG pathway and cancer-related hallmark gene set analyses were performed to clarify the biological mechanisms that were related to the risk model.

In the low-risk category, several metabolically related pathways, such as histidine, fatty acid, and ether lipid metabolism, were highly enriched (Figure 7A). The metabolism of these amino acids and



lipids was reported to be closely associated with tumors (Currie et al., 2013; Li and Zhang, 2016; Bian et al., 2021). Interestingly, immune-related pathways were observed to cluster in the low-risk category, such as the T-cell and B-cell receptor signaling pathways. Conversely, for the high-risk category, cell proliferation-related pathways such as the G2/M checkpoint, E2F targets, and MYC targets were enriched. It is worth noting that hypoxia was also identified in the high-risk category, as anticipated (Figure 7B). The expression heatmap of the 30 most significant DEHRGs between the two risk categories is shown in Supplementary Figure S3. It is evident from the heatmap that there is a significant difference in the expression levels of these genes.

In summary, there are significant mechanistic differences between the low-risk and high-risk categories, with the immune

features of the former and proliferation characteristics of the latter being consistent with their respective prognostic outcomes.

3.7 Characteristics of the tumor immune microenvironment associated with HRG models

To further explore the risk score and immunity, we performed a correlation analysis of immune cell abundance and HRG expression in the risk model. The findings revealed an extensively significant negative correlation between immune cell abundance and HRGs, such as ENO3, INHA, and ALDOA. RORA was positively correlated with most immune cell types (Figure 8A).

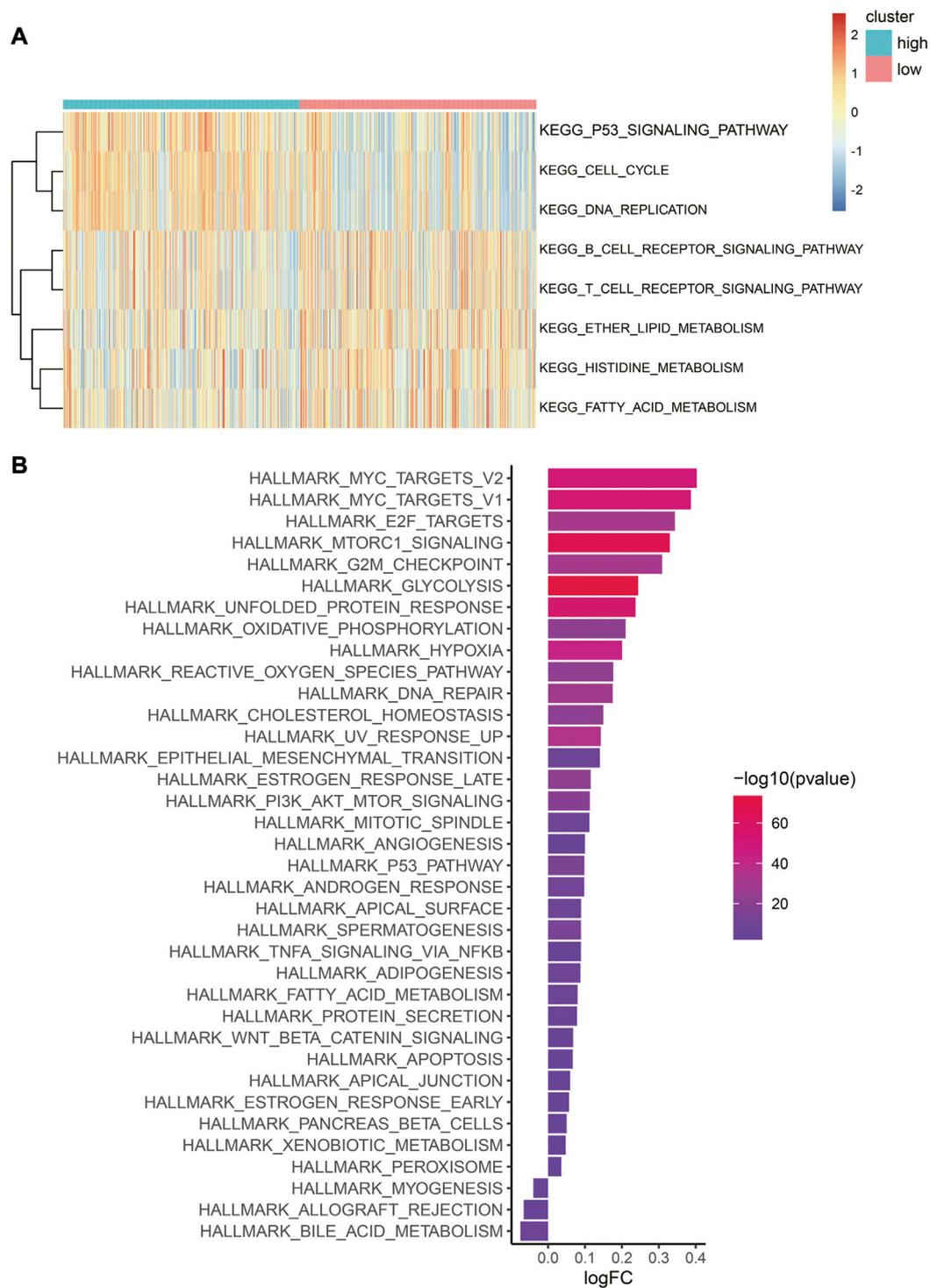


FIGURE 7

Difference in biological pathways and cancer-related gene sets between the two risk categories. **(A)** Heatmap of GSEA enrichment of the KEGG pathway. Red represents high enrichment scores, while blue represents low enrichment scores. **(B)** Bar graph of hallmark gene set enrichment score. The color indicates the significance of the difference, and the x-axis represents the enrichment score fold change in this hallmark gene set between high- and low-risk categories.

In the TCGA cohort, immune cell abundance was calculated between the two risk categories. The abundance of activated CD8 T cells, activated B cells, effector memory CD8 T cells, and

central memory CD4 T cells was higher in the low-risk category (Figure 8C). In addition, the immune score of the low-risk category was significantly greater than that of the high-risk category (Figure 8B).

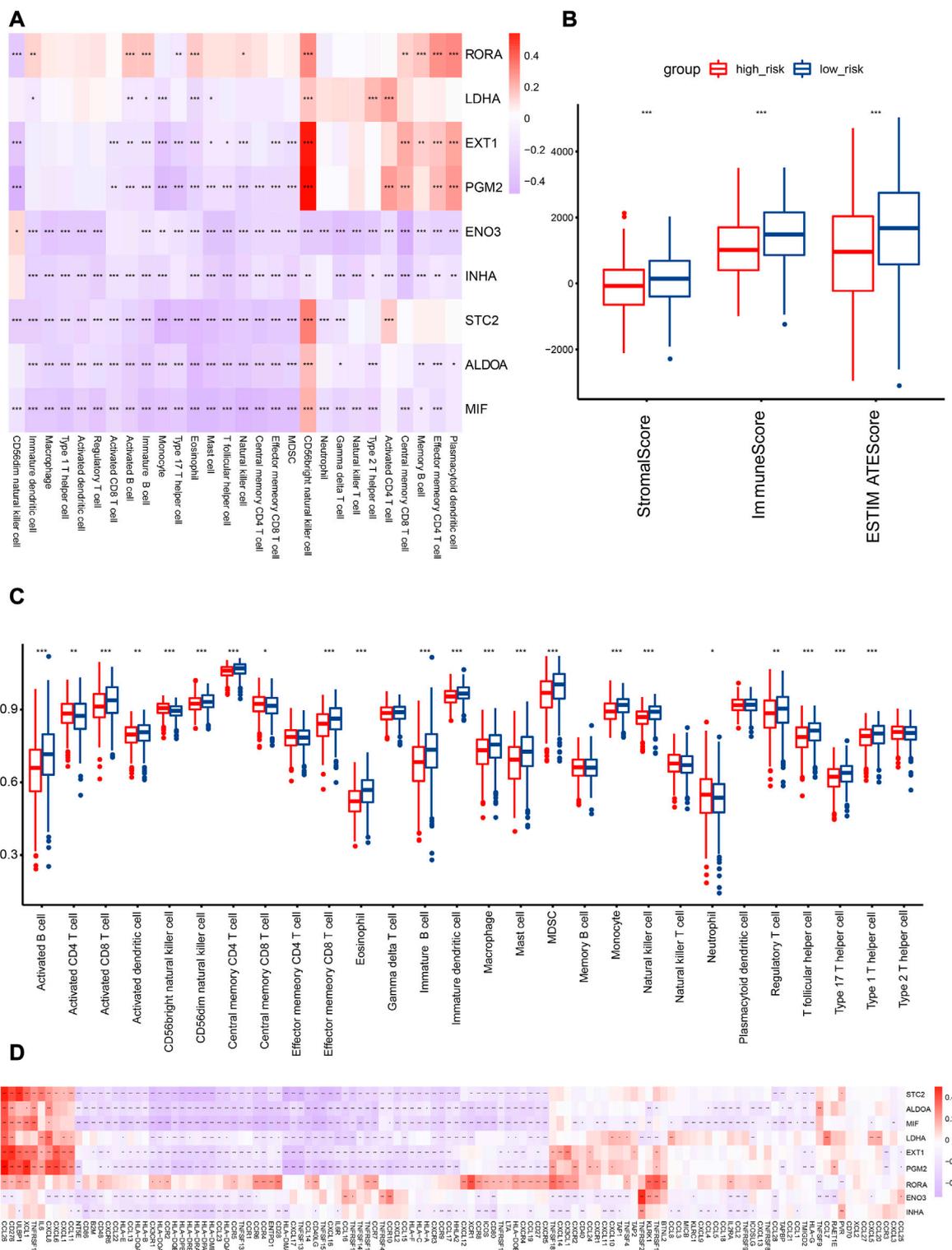
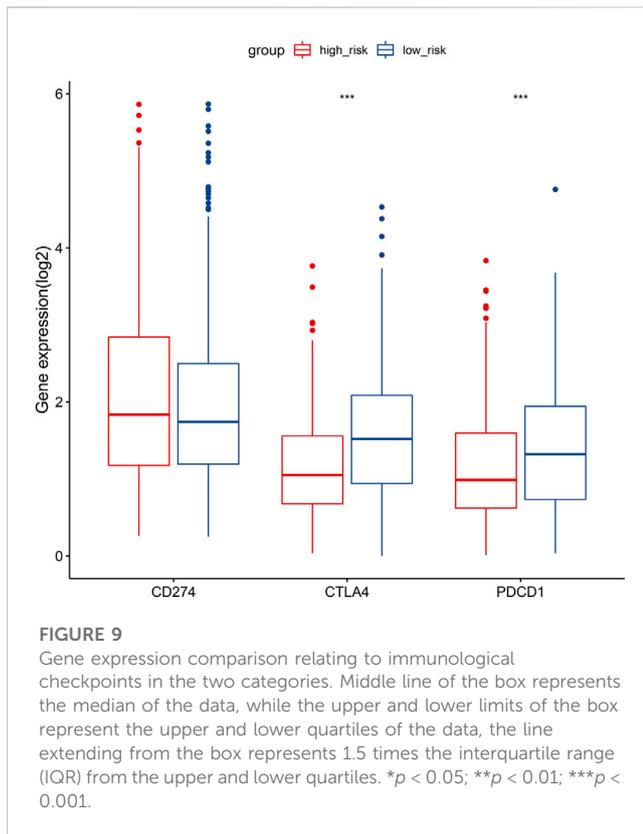


FIGURE 8 Association analysis of HRGs with TIME in the TCGA cohort. **(A)** Correlation analysis of HRGs in risk signature and immune cell abundance. **(B)** Comparison of immune score in the two subcategories. **(C)** Quantitative analysis of immune cell abundance between patients in the two risk categories. **(D)** HRGs and immunomodulator-related gene correlation analysis. Red represents a positive correlation, while blue represents a negative correlation in the heatmap. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.



To illustrate the relationship between the HRG model and the immune microenvironment, we generated the correlation pattern of the HRG model and immune-modulator-related genes. The results demonstrated that an extensively significant negative correlation existed, including STC2, ALDOA, MIF, LDHA, EXT1, PGM2, ENO3, and INHA. Only RORA was positively correlated with most immune modulators (Figure 8D).

The low-risk category was associated with high immune cell infiltration, immune-modulator-related gene expression, and immune scores, which are indicative of favorable clinical outcomes related to preexisting anticancer immunity.

3.8 Immune checkpoint expression between risk categories

In consideration of the therapeutic significance of treatment approaches related to immune checkpoint blockade in NSCLC, the correlation between the risk score and several immune checkpoints, such as PD1, PDL1, and CTLA4, were analyzed. In the low-risk category, we found a substantial increase in the expression of PD1 and CTLA4, while the expression of PDL1 was not significantly different (Figure 9).

4 Discussion

NSCLC has a significant fatality rate and is one of the most frequently diagnosed cancers worldwide. Identifying accurate and effective signatures is of great importance for NSCLC prognostic

prediction. Here, we constructed a novel HRG signature in a large NSCLC cohort, and training and validation datasets were used to confirm the signature's efficacy.

Increasing evidence suggests that NSCLC patients who have intratumoral hypoxia have a worse prognosis and a more aggressive form of the disease (Carmeliet and Jain, 2011; Krock et al., 2011; Semenza, 2012; Eales et al., 2016; Mittal, 2018; Tirpe et al., 2019). Many monogenic studies of HRGs have been related to prognosis and tumor immunity in several cancers (Lee et al., 2019; Liu et al., 2020; Zhang et al., 2020). However, there has been no systematic investigation connecting TME features with the hypoxia-related signature in NSCLC. By combining data from many separate NSCLC studies, we were able to generate and verify a unique hypoxia risk signature. The established model confirmed that this signature could be used to effectively predict clinical outcomes and TME characteristics.

The most noticeable feature of malignant tumors is hypoxia. Prior studies have demonstrated that hypoxia contributes to the aggressive development of lung cancer (Smolle et al., 2020; Ouyang et al., 2021; Lane et al., 2022). However, due to its multifaceted nature, hypoxia's precise function in NSCLC progression is unclear. In the current research, we screened nine HRGs closely associated with NSCLC. Of these, four genes had been proven to be involved in tumor malignancy. ALDOA, an oncogene, has been identified as being involved in tumor cell malignant growth and worsening prognosis in hepatoma and pancreatic cancer (Ji et al., 2016). LDHA promotes papillary thyroid carcinoma metastasis by regulating EMT gene transcription (Hou et al., 2021). STC2 could promote the metastasis of HNSCC through the PI3K/AKT/Snail signaling pathway (Yang et al., 2017). In addition, EXT1 was demonstrated to participate in the process of cancer proliferation and migration by methylation regulation (Kong et al., 2021). Our results showed that this four-gene signature affecting the prognosis of NSCLC was incorporated into the risk model, suggesting that the established signature performed well in recurrence and prognosis prediction.

Hypoxia is a very important participant in the TME through various mechanisms. HRGs affect the infiltration of various immune cell subtypes which reshape the TIME. In particular, hypoxia enhances PD-L1 expression in multiple tumor cells through HIF-1's direct binding to the HRE in the PDL1 promoter. The activity of T and NK cells is suppressed as a result of LDHA stimulation of the lactate metabolism (Brand et al., 2016). ALDOA has been shown to activate the NLRP3 inflammasome and induce the secretion of proinflammatory factors to regulate anticancer immune responses (Bai et al., 2022). RORA regulates ILC3s, macrophages, and Treg cells, which are crucial for ILC2 development (Haim-Vilmovsky et al., 2021). In addition, MIF regulates the inhibition of immune responses by enhancing harmful inflammation and eventually leads to the promotion of cancer metastasis (Sumaiya et al., 2022). In general, all the collective evidence inspired us to investigate the potential use of the hypoxia risk score to predict immunological features.

This study revealed a negative correlation between the risk score and the TIS and TIICs (such as activated B cells, activated CD8 T cells, central memory CD4 T cells, and effector memory CD8 T cells), which suggests that patients in the low-risk category had more preexisting anticancer immunity in their TME (Figure 8C). It is accepted that immune checkpoints prevent anticancer immunity in the TME (Nishino et al., 2017). Consistently, in the present study, the hypoxia risk score showed a negative correlation with the expression of several immune checkpoints, such as CTLA-4 and PD-1 (Figure 9). The TIS

and immune checkpoint expression were considerably greater in low-risk category patients, indicating increased anticancer immunity and additional therapeutic targets in the TME. In other words, low-risk patients may gain more from ICB treatment, which may reactivate dormant tumoricidal immunity in the TME (Liu et al., 2021).

The innovative aspect of this study is its development of a prognostic model for NSCLC that can be used to predict the prognosis of both lung adenocarcinoma and lung squamous cell carcinoma without the need to distinguish specific cancer types. This approach potentially simplifies the diagnostic and treatment process for clinicians. Additionally, the ability to use this model to predict recurrence in NSCLC patients enhances its clinical utility. The study's large sample size and good predictive performance in RNA-seq and microarray data platforms further highlight its strengths. Overall, the novel approach of using a prognostic model for different types of NSCLC represents a significant advancement in the field of cancer research. Nevertheless, this investigation had a few limitations. The risk score algorithm needs to be validated in real-world prospective cohort studies since the data used in the present research came from public sources. The sequencing techniques used for this research cohort varied and the accuracy of this formula might be affected to some extent. Additionally, the majority of patients were Caucasian, underscoring the need to investigate the risk score's accuracy as a predictor of outcomes in patients of other races.

5 Conclusion

In total, a unique nine-gene risk score signature was established for NSCLC. The risk score was proven to correlate with OS, functional enrichment, the immune microenvironment, and therapy responsiveness. Additionally, it accurately predicted prognosis possibility depending on age, sex, and disease stage. These observations suggest that molecular risk stratification may aid in prognostic prediction and personalized precision therapy for NSCLC.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

Ethical review and approval were not required for the study on human participants in accordance with local legislation and

institutional requirements. Written informed consent for participation was not required for this study in accordance with national legislation and institutional requirements.

Author contributions

ZL participated in the idea design, data analysis, interpretation, and drafting of the manuscript. YC, SZ, JX, JS, and HC assisted in idea design. JC and SW are responsible for data collection. MZ, HZ, SL, ZQ, and GX for literature search. All authors approved the final version and contributed to the article.

Funding

This study was funded by Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-079D) and sponsored by Tianjin Health Research Project (TJWJ2022QN101).

Acknowledgments

We greatly appreciate the information provided by the Gene Expression Omnibus and The Cancer Genome Atlas.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1115308/full#supplementary-material>

References

- Bai, D., Du, J., Bu, X., Cao, W., Sun, T., Zhao, J., et al. (2022). ALDOA maintains NLRP3 inflammasome activation by controlling AMPK activation. *Autophagy* 18 (7), 1673–1693. doi:10.1080/15548627.2021.1997051
- Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: More than meets the eye. *Lancet Oncol.* 16 (4), e173–e180. doi:10.1016/S1470-2045(14)71116-7
- Bian, X., Liu, R., Meng, Y., Xing, D., Xu, D., and Lu, Z. (2021). Lipid metabolism and cancer. *J. Exp. Med.* 218 (1), e20201606. doi:10.1084/jem.20201606
- Brand, A., Singer, K., Koehl, G. E., Kolitzus, M., Schoenhammer, G., Thiel, A., et al. (2016). LDHA-associated lactic acid production blunts tumor immunosurveillance by T and NK cells. *Cell Metab.* 24 (5), 657–671. doi:10.1016/j.cmet.2016.08.011
- Buffa, F. M., Harris, A. L., West, C. M., and Miller, C. J. (2010). Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* 102 (2), 428–435. doi:10.1038/sj.bjc.6605450
- Carmeliet, P., and Jain, R. K. (2011). Molecular mechanisms and clinical applications of angiogenesis. *Nature* 473 (7347), 298–307. doi:10.1038/nature10144

- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019
- Currie, E., Schulze, A., Zechner, R., Walther, T. C., and Farese, R. V., Jr. (2013). Cellular fatty acid metabolism and cancer. *Cell Metab.* 18 (2), 153–161. doi:10.1016/j.cmet.2013.05.017
- Cuyun Carter, G., Barrett, A. M., Kaye, J. A., Liepa, A. M., Winfree, K. B., and John, W. J. (2014). A comprehensive review of nongenetic prognostic and predictive factors influencing the heterogeneity of outcomes in advanced non-small-cell lung cancer. *Cancer Manag. Res.* 6, 437–449. doi:10.2147/CMAR.S63603
- D'Ignazio, L., Batié, M., and Rocha, S. (2017). Hypoxia and inflammation in cancer, focus on HIF and NF- κ B. *Biomedicines* 5 (2), 21. doi:10.3390/biomedicines5020021
- Eales, K. L., Hollinshead, K. E., and Tennant, D. A. (2016). Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis* 5 (1), e190. doi:10.1038/oncsis.2015.50
- Gilkes, D. M., Semenza, G. L., and Wirtz, D. (2014). Hypoxia and the extracellular matrix: Drivers of tumour metastasis. *Nat. Rev. Cancer* 14 (6), 430–439. doi:10.1038/nrc3726
- Haim-Vilmsky, L., Henriksson, J., Walker, J. A., Miao, Z., Natan, E., Kar, G., et al. (2021). Mapping Rora expression in resting and activated CD4+ T cells. *PLoS One* 16 (5), e0251233. doi:10.1371/journal.pone.0251233
- Hou, X., Shi, X., Zhang, W., Li, D., Hu, L., Yang, J., et al. (2021). LDHA induces EMT gene transcription and regulates autophagy to promote the metastasis and tumorigenesis of papillary thyroid carcinoma. *Cell Death Dis.* 12 (4), 347. doi:10.1038/s41419-021-03641-8
- Ji, S., Zhang, B., Liu, J., Qin, Y., Liang, C., Shi, S., et al. (2016). ALDOA functions as an oncogene in the highly metastatic pancreatic cancer. *Cancer Lett.* 374 (1), 127–135. doi:10.1016/j.canlet.2016.01.054
- Jing, X., Yang, F., Shao, C., Wei, K., Xie, M., Shen, H., et al. (2019). Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol. Cancer* 18 (1), 157. doi:10.1186/s12943-019-1089-9
- Kong, W., Chen, Y., Zhao, Z., Zhang, L., Lin, X., Luo, X., et al. (2021). EXT1 methylation promotes proliferation and migration and predicts the clinical outcome of non-small cell lung carcinoma via WNT signalling pathway. *J. Cell Mol. Med.* 25 (5), 2609–2620. doi:10.1111/jcmm.16277
- Krock, B. L., Skuli, N., and Simon, M. C. (2011). Hypoxia-induced angiogenesis: Good and evil. *Genes Cancer* 2 (12), 1117–1133. doi:10.1177/1947601911423654
- Lane, B., Khan, M. T., Choudhury, A., Salem, A., and West, C. M. L. (2022). Development and validation of a hypoxia-associated signature for lung adenocarcinoma. *Sci. Rep.* 12 (1), 1290. doi:10.1038/s41598-022-05385-7
- Lee, J. H., Jung, S., Park, W. S., Choe, E. K., Kim, E., Shin, R., et al. (2019). Prognostic nomogram of hypoxia-related genes predicting overall survival of colorectal cancer—Analysis of TCGA database. *Sci. Rep.* 9 (1), 1803. doi:10.1038/s41598-018-38116-y
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28 (6), 882–883. doi:10.1093/bioinformatics/bts034
- Li, Z., and Zhang, H. (2016). Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol. Life Sci.* 73 (2), 377–392. doi:10.1007/s00018-015-2070-4
- Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004
- Liu, Y., Wu, J., Huang, W., Weng, S., Wang, B., Chen, Y., et al. (2020). Development and validation of a hypoxia-immune-based microenvironment gene signature for risk stratification in gastric cancer. *J. Transl. Med.* 18 (1), 201. doi:10.1186/s12967-020-02366-0
- Liu, Z., Tang, Q., Qi, T., Othmane, B., Yang, Z., Chen, J., et al. (2021). A robust hypoxia risk score predicts the clinical outcomes and tumor microenvironment immune characters in bladder cancer. *Front. Immunol.* 12, 725223. doi:10.3389/fimmu.2021.725223
- Magnon, C., Opolon, P., Ricard, M., Connault, E., Ardouin, P., Galaup, A., et al. (2007). Radiation and inhibition of angiogenesis by conastatin synergize to induce HIF-1 α -mediated tumor apoptotic switch. *J. Clin. Invest.* 117 (7), 1844–1855. doi:10.1172/JCI30269
- Mittal, V. (2018). Epithelial mesenchymal transition in tumor metastasis. *Annu. Rev. Pathol.* 13, 395–412. doi:10.1146/annurev-pathol-020117-043854
- Muz, B., de la Puente, P., Azab, F., and Azab, A. K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia (Auckl)* 3, 83–92. doi:10.2147/HP.S93413
- Nishino, M., Ramaiya, N. H., Hatabu, H., and Hodi, F. S. (2017). Monitoring immune-checkpoint blockade: Response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* 14 (11), 655–668. doi:10.1038/nrclinonc.2017.88
- Nobre, A. R., Entenberg, D., Wang, Y., Condeelis, J., and Aguirre-Ghiso, J. A. (2018). The different routes to metastasis via hypoxia-regulated programs. *Trends Cell Biol.* 28 (11), 941–956. doi:10.1016/j.tcb.2018.06.008
- Ouyang, W., Jiang, Y., Bu, S., Tang, T., Huang, L., Chen, M., et al. (2021). A prognostic risk score based on hypoxia-immunity-and epithelial-to-mesenchymal transition-related genes for the prognosis and immunotherapy response of lung adenocarcinoma. *Front. Cell Dev. Biol.* 9, 758777. doi:10.3389/fcell.2021.758777
- Petrova, V., Annicchiarico-Petruzzelli, M., Melino, G., and Amelio, I. (2018). The hypoxic tumour microenvironment. *Oncogenesis* 7 (1), 10. doi:10.1038/s41389-017-0011-9
- Qian, X., Zhang, Q., Shao, N., Shan, Z., Cheang, T., Zhang, Z., et al. (2019). Respiratory hyperoxia reverses immunosuppression by regulating myeloid-derived suppressor cells and PD-L1 expression in a triple-negative breast cancer mouse model. *Am. J. Cancer Res.* 9 (3), 529–545.
- Ragnum, H. B., Vlatkovic, L., Lie, A. K., Axcrone, K., Julin, C. H., Frikstad, K. M., et al. (2015). The tumour hypoxia marker pimonidazole reflects a transcriptional programme associated with aggressive prostate cancer. *Br. J. Cancer* 112 (2), 382–390. doi:10.1038/bjc.2014.604
- Rankin, E. B., Nam, J. M., and Giaccia, A. J. (2016). Hypoxia: Signaling the metastatic cascade. *Trends Cancer* 2 (6), 295–304. doi:10.1016/j.trecan.2016.05.006
- Semenza, G. L. (2012). Hypoxia-inducible factors: Mediators of cancer progression and targets for cancer therapy. *Trends Pharmacol. Sci.* 33 (4), 207–214. doi:10.1016/j.tips.2012.01.005
- Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics. *CA Cancer J. Clin.* 73 (1), 17–48. doi:10.3322/caac.21763
- Smolle, E., Leko, P., Stacher-Priehse, E., Brcic, L., El-Heliebi, A., Hofmann, L., et al. (2020). Distribution and prognostic significance of gluconeogenesis and glycolysis in lung cancer. *Mol. Oncol.* 14 (11), 2853–2867. doi:10.1002/1878-0261.12780
- Sumaiya, K., Langford, D., Natarajaseenivasan, K., and Shanmughapriya, S. (2022). Macrophage migration inhibitory factor (MIF): A multifaceted cytokine regulated by genetic and physiological strategies. *Pharmacol. Ther.* 233, 108024. doi:10.1016/j.pharmthera.2021.108024
- Sun, J., Zhao, T., Zhao, D., Qi, X., Bao, X., Shi, R., et al. (2020). Development and validation of a hypoxia-related gene signature to predict overall survival in early-stage lung adenocarcinoma patients. *Ther. Adv. Med. Oncol.* 12, 1758835920937904. doi:10.1177/1758835920937904
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. doi:10.1016/j.immuni.2018.03.023
- Tirpe, A. A., Gulei, D., Ciortea, S. M., Crivii, C., and Berindan-Neagoe, I. (2019). Hypoxia: Overview on hypoxia-mediated mechanisms with a focus on the role of HIF genes. *Int. J. Mol. Sci.* 20 (24), 6140. doi:10.3390/ijms20246140
- Winter, S. C., Buffa, F. M., Silva, P., Miller, C., Valentine, H. R., Turley, H., et al. (2007). Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res.* 67 (7), 3441–3449. doi:10.1158/0008-5472.CAN-06-3322
- Wood, D. E., Kazerooni, E. A., Aberle, D., Berman, A., Brown, L. M., Eapen, G. A., et al. (2022). NCCN Guidelines[®] insights: Lung cancer screening, version 1.2022. *J. Natl. Compr. Canc. Netw.* 20 (7), 754–764. doi:10.6004/jnccn.2022.0036
- Yang, S., Ji, Q., Chang, B., Wang, Y., Zhu, Y., Li, D., et al. (2017). STC2 promotes head and neck squamous cell carcinoma metastasis through modulating the PI3K/AKT/Snail signaling. *Oncotarget* 8 (4), 5976–5991. doi:10.18632/oncotarget.13355
- Zhang, B., Tang, B., Gao, J., Li, J., Kong, L., and Qin, L. (2020). A hypoxia-related signature for clinically predicting diagnosis, prognosis and immune microenvironment of hepatocellular carcinoma patients. *J. Transl. Med.* 18 (1), 342. doi:10.1186/s12967-020-02492-9