Research article

Clinical significance and immune landscape of a novel immune cell infiltration-based prognostic model in lung adenocarcinoma

Lupeng Qiu [a,b,1], Zizhong Yang [c,1], Guhe Jia [c,1], Yanjie Liang [a], Sicheng Du [a,b], Jian Zhang [a], Minglu Liu [a], Xiao Zhao [a,**], Shunchang Jiao [a,b,*]

[a] Department of Medical Oncology, The First Medical Center, Chinese PLA General Hospital, Beijing, China
[b] Department of Graduate Administration, Chinese PLA General Hospital, Beijing, China
[c] School of Medicine, Nankai University, Tianjin, China

ARTICLE INFO

ABSTRACT

Tumor-infiltrating immune cells (TICs) play a central role in the tumor microenvironment, which can reflect the host anti-tumor immune response. However, few studies have explored TICs in predicting the prognosis of lung adenocarcinoma (LUAD). In our study, we enrolled 2470 LUAD patients from TCGA and GEO databases, and the normalized enrichment scores for 65 immune cell types were quantified for each patient. An immune-related risk score (IRRS) was built on the basis of 17 selected TICs using LASSO regression analysis, and the results showed that high-risk patients were correlated with shorter survival time for the LUAD cohorts. Correlation analyses between IRRS and clinical characteristics were also evaluated to validate the clinical use of IRRS. In addition, we analyzed the differences in the distribution of immune cell infiltration and immunoregulatory gene expression, which may facilitate individual immunotherapy. Based on the above result, we conclude that IRRS can act as a powerful predictor for risk stratification and prognosis prediction, and may facilitate the decision-making process for LUAD patients.

1. Introduction

Lung cancer is the leading cause of cancer-related mortality worldwide, with lung adenocarcinoma (LUAD) being the most prevalent subtype of the disease [1–3]. Despite an overall decline in mortality during the last few years, it remains a major threat to human health [4,5]. Nowadays, surgery, chemotherapy, targeted therapy, and immunotherapy are the main treatment options for LUAD, and the prognosis remains poor for most patients [6–10]. With rapidly evolving next-generation sequencing technologies, it is necessary to stratify LUAD patients more accurately to better guide treatment strategies and improve prognosis [11].

The tumor immune microenvironment (TIME) is an extremely complex assembly of various elements that include tumor cells, immune cells, stromal, and extracellular components [12,13]. Currently, growing evidence suggest that TIME plays a key role in LUAD development [14–16]. Tumor-infiltrating lymphocytes and tumor-associated macrophages, as easily available biomarkers in TIME,

* Corresponding author. Department of Medical Oncology, The First Medical Center, Chinese PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing, 100853, China.
** Corresponding author. Department of Medical Oncology, The First Medical Center, Chinese PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing, 100853, China.
E-mail addresses: dr.zx@163.com (X. Zhao), scjiao301@126.com (S. Jiao).
[1] These authors contributed equally to this work.
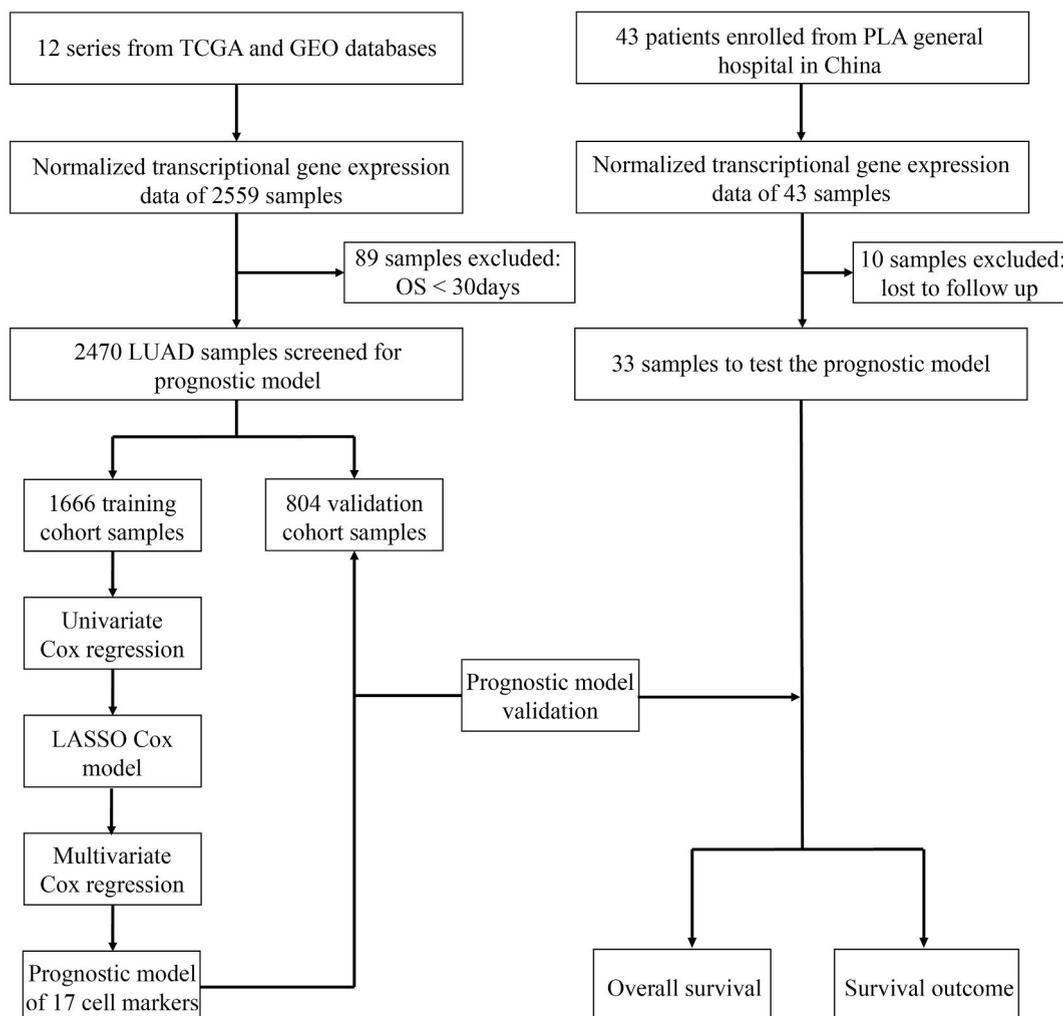
Fig. 1. Flow chart of the study.

have been shown to be associated with cancer prognosis [17–19]. Furthermore, more and more studies suggest that the presence of tumor-infiltrating immune cells (TICs) within TIME can inhibit or promote cancer growth [20–23]. Therefore, an in-depth analysis of the TIME may provide more reliable prognostic biomarkers for patients with LUAD.

Currently, many studies have examined the impact of TICs in LUAD development. Dai found that the hypoxia-related prognostic signature can predict immune cell infiltration and prognosis in LUAD patients [24]. Zuo established an immune cell characteristic score, which is a prognostic indicator for patients and can be widely applied in future clinical settings [25]. However, the small sample size and a low number of immune cell signatures may limit its statistical accuracy and overall reliability. Therefore, further studies using larger sample sizes and more comprehensive immune cell subtypes are still required.

In the present study, we focus on evaluating the features of TME in patients with LUAD, and further analyzed its immune activity and prognosis. Immune cell signatures were derived from various published literature [26–35]. 65 immune cells were included, and 17 immune cells correlated with the prognosis of LUAD were identified. Finally, a novel immune cell infiltration-based model was constructed and further examined in a validation set and an independent test set.

2. Materials and methods

2.1. Data processing

Messenger RNA (mRNA) profiling of LUAD patients were acquired from the TCGA and GEO databases. The enrollment criteria are the following: (1) the number of primary LUAD samples exceeds 50 cases; (2) OS was greater than 30 days; (3) the data type was transcriptome profiling.

The original data was downloaded from GEO database and standardized using the RMA with the "affy" R package [36]. For the

A

Cell type	HR	lower 95%CI	upper 95%CI		pvalue
Activated_CD4_T_cells	1.793	1.487	2.161		0.000
Adipocytes	0.833	0.737	0.940		0.003
B_cells	0.746	0.636	0.876		0.000
B_cells_memory	0.636	0.500	0.810		0.000
B_cells_naive	0.555	0.420	0.733		0.000
CD56dim_natural_killer_cells	1.438	1.161	1.781		0.001
Central_memory_CD4_T_cells	0.625	0.478	0.816		0.001
Cytotoxic_lymphocytes	1.247	1.132	1.374		0.000
Endothelial_cells	0.803	0.694	0.930		0.003
Eosinophils	0.553	0.377	0.812		0.003
Erythrocytes	1.544	1.364	1.748		0.000
Granulocyte_macrophage_progenitors	1.440	1.141	1.817		0.002
Keratinocytes	1.366	1.217	1.532		0.000
Mast_cells	0.552	0.446	0.684		0.000
MDSCs	1.404	1.114	1.769		0.004
Megakaryocyte_erythroid_progenitors	1.284	1.082	1.525		0.004
Multipotent_progenitors	1.307	1.103	1.550		0.002
Myeloid_dendritic_cells	0.774	0.688	0.872		0.000
Natural_killer_T_cells	2.450	1.686	3.559		0.000
Neutrophils	1.278	1.084	1.507		0.003
pDC	0.661	0.516	0.848		0.001
T_cells_CD4_memory_resting	0.694	0.538	0.894		0.005
T_cells_CD4_naive	0.637	0.490	0.827		0.001

0.5   1   1.5
Hazard ratios

B

C

Fig. 2. IRRS model Construction. (A) Forest plot of OS for 23 TICs. (B) Partial likelihood deviance of the LASSO coefficient profiles. (C) LASSO coefficient profiles of 23 TICs.

cohort from TCGA, RNA-sequencing data were transformed into the form of TPM for subsequent analysis [37]. Finally, the ComBat algorithm was used to reduce the likelihood of potential multicenter batch effects among each cohort [38]. Our workflow is depicted in Fig. 1.

Finally, a total of 2470 LUAD samples from twelve cohorts were analyzed in this study. The 1666 patients collected from the GSE31210, GSE13213, GSE50081, GSE41271, GSE68465, GSE30219 and TCGA-LUAD cohorts were used as the LUAD training cohort, while the 804 patients collected from the GSE26939, GSE14814, GSE72094, GSE37745 and GSE42127 cohorts were used as the validation cohort. Another 43 patients from the Chinese PLA general hospital were used as an external test cohort.

2.2. RNA sequencing analysis

A total of 43 fresh tumor samples were processed using the Qiagen RNeasy Mini Kit. The RNA obtained after extraction was completed for concentration measurement using Nanodrop 2000C, and Agilent 2100 Bioanalyzer was applied to RNA fragment integrity analysis. Concentration measurements were performed with Qubit4.0 Fluorometric Quantitation (Thermo Fisher Scientific, USA), and RNA library fragment analysis was performed with an Agilent 2100 Bioanalyzer. Qualified libraries were sequenced in
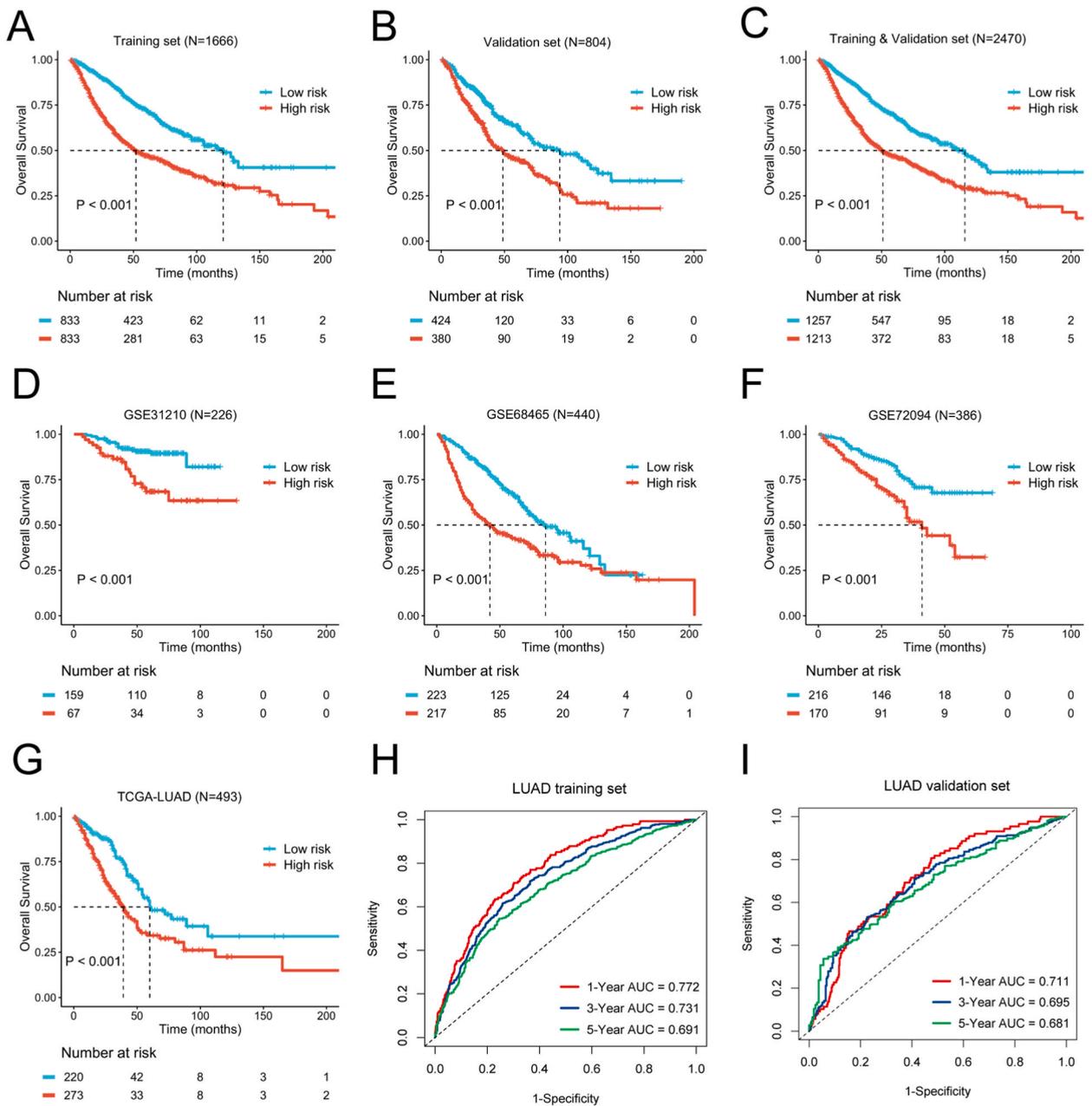
Fig. 3. Kaplan-Meier curves for OS. (A) Training cohort. (B) Validation cohort. (C) Entire cohort. Subgroup analysis of survival for the two risk groups. GSE31210 (D), GSE68465 (E), GSE72094 (F), and TCGA-LUAD (G). (H–I) ROC curves of the IRRS model for the training and validation cohorts.

PE150 mode on the Illumina HiseqX platform (Illumina US).

2.3. Gene set enrichment analysis

The 65 TICs were collected based on a comprehensive literature search, and these immune cells included T cells, NK cells, DCs et al. Therefore, we believe that the 65 TICs can truly reflect the immune status of different tumor samples. The ssGSEA implemented in the "GSVA" R package was applied to measure the per sample expression of 65 TICs [39]. The normalized enrichment score of each TIC was collected for further analysis.

Table 1

Results of Cox regression analysis in the LUAD training cohort.

Variables	Univariate analysis		Multivariate analysis	
	HR (95 % CI)	P-value	HR (95 % CI)	P-value
Age				
<65	Reference			
≥65	1.561 (1.338–1.821)	**<0.001**	1.633 (1.377–1.937)	**<0.001**
Sex				
Male	Reference			
Female	0.777 (0.666–0.907)	**0.001**	0.953 (0.797–1.138)	0.592
Smoking history				
Yes	Reference			
No	0.603 (0.485–0.750)	**<0.001**	0.730 (0.578–0.923)	**0.009**
TNM stage				
I + II	Reference			
III + IV	3.177 (2.661–3.794)	**<0.001**	2.923 (2.410–3.545)	**<0.001**
Risk score				
Low	Reference			
High	2.210 (1.884–2.593)	**<0.001**	2.091 (1.749–2.500)	**<0.001**

Table 2

Results of Cox regression analysis in the LUAD validation cohort.

Variables	Univariate analysis		Multivariate analysis	
	HR (95 % CI)	P-value	HR (95 % CI)	P-value
Age				
<65	Reference			
≥65	1.581 (1.257–1.989)	**<0.001**	1.497 (1.183–1.895)	**<0.001**
Sex				
Male	Reference			
Female	0.669 (0.537–0.832)	**<0.001**	0.726 (0.579–0.911)	**0.006**
Smoking history				
Yes	Reference			
No	0.928 (0.524–1.642)	0.797		
TNM stage				
I + II	Reference			
III + IV	1.976(1.510–2.584)	**<0.001**	1.785 (1.360–2.343)	**<0.001**
Risk score				
Low	Reference			
High	1.781 (1.428–2.222)	**<0.001**	1.689 (1.343–2.123)	**<0.001**

2.4. DEG analysis

The R package "limma" was used for DEG analysis [40]. Enrichment analysis of DEGs were calculated using GSEA, respectively. GO and KEGG analyses were implemented using the "clusterProfiler" R package [41].

2.5. Tumor purity analysis

Estimation of immune and stromal scores were calculated using the ESTIMATE algorithm from the RNA-seq data [42]. Subsequently, the association between the IRRS subgroups and the scores mentioned above was evaluated in our LUAD samples.

2.6. Immune landscape description

The relative abundance of 22 TICs was estimated using CIBERSORT with the LM22 gene set. The relationship between 65 immune-related genes and IRRS, including MHC-class family, immune checkpoints family, and several other molecules were furthered explored.

2.7. Response to immunotherapy

The TIDE web was applied to predict the immunotherapy response to immune checkpoint blockade for each patient. In addition, IMvigor210, GSE78220, GSE93157 and Check-Mate cohorts were also employed to analyze IRRS for immunotherapy.
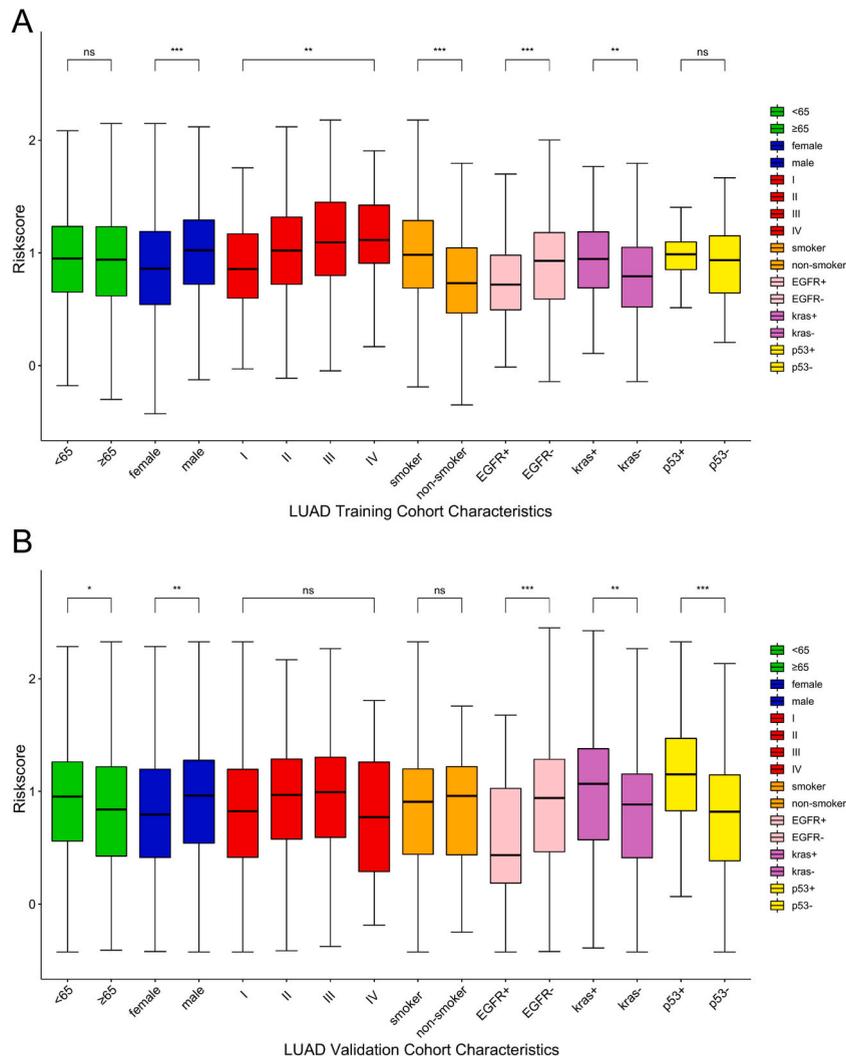
Fig. 4. Summary of the clinical characteristics stratified by IRRS. (A) LUAD training cohort. (B) LUAD validation cohort.

2.8. Statistical analysis

The IRRS model was established based on the correlation between TICs level and survival, and Lasso regression analysis was used to select highly prognosis-related TICs [43]. The formula of the IRRS was established on the basis of the value of lambda. min. The "survminer" R package was employed to draw survival curves. The nomogram model was created using the "rma" R package to predict the survival of LUAD patients. Then, the nomogram was assessed with calibration curves, which measure the relationship between the outcomes predicted by the prognostic model and the observed outcomes in the LUAD cohort. All statistical analyses were performed with R software.

3. Results

3.1. Immune cell signatures for LUAD patients

To analyze the prognostic value of 65 TICs, 12 LUAD cohorts were included. A total of 2470 LUAD patients were separated into a training and validation set. Univariate analysis demonstrated that 23 TICs were correlated with OS (P < 0.01). The rigorous cut-off point was applied to obtain a better prognostic model in LUAD samples (Fig. 2A). Subsequently, 17 out of 23 TICs highly correlated with OS were selected through LASSO regression analysis (Fig. 2B and C), and the risk score formula was established (IRRS = Activated CD4 T cells * 0.503 + B cells * (−0.309) + CTL * 0.070 + Endothelial cells * 0.033 + Eosinophils * (−0.239) + Erythrocytes * 0.041 + GMP * (−0.167) + Keratinocytes * 0.211 + Mast cells * (−0.421) + MDSCs * 0.459 + MEP * (−0.167) + Multipotent progenitors * 0.096 + Myeloid dendritic cells * (−0.076) + Natural killer T cells * 0.680 + pDC * (−0.108) + CD4 Tm resting * (−0.448) +
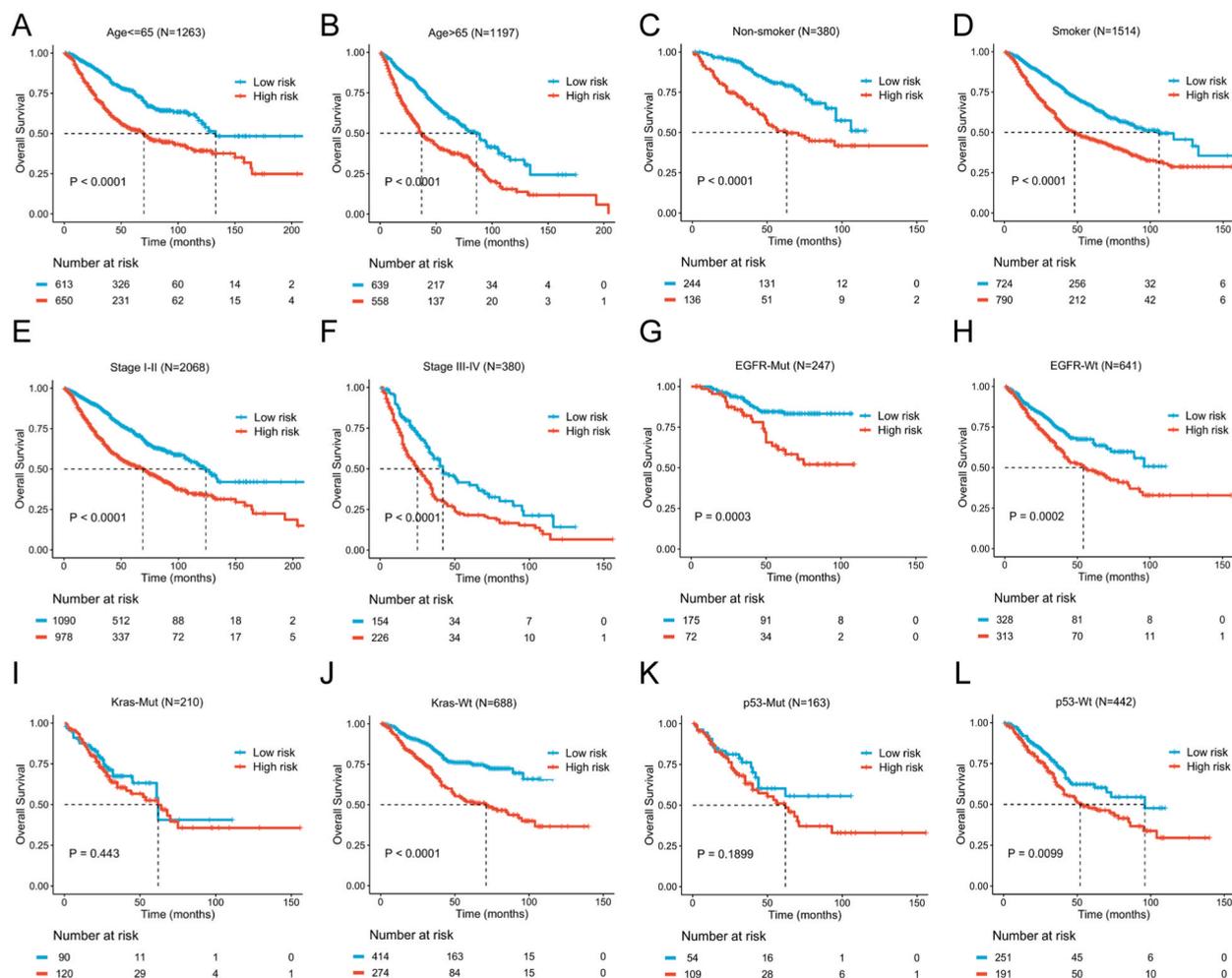
Fig. 5. Kaplan-Meier survival curves of LUAD patients in different subgroups. (A) Age ≤65. (B) Age >65. (C) Non-smokers (D) Smokers. (E) Stage I-II. (F) Stage III-IV. Positive (G) and negative (H) EGFR mutation status. Positive (I) and negative (J) KRAS mutation status. Positive (K) and negative (L) p53 mutation status.

CD4 naive T cells * (−0.093)).

Based on this formula, we calculated the IRRS for all LUAD patients. Subsequently, 1666 LUAD patients were divided into two IRRS subgroups, separated by the median value of 0.9465. The survival curves indicated that patients with low IRRS had longer OS in the LUAD training cohort (121 months vs 52 months, P < 0.001, Fig. 3A). To further test the stability of the IRRS model, we applied the same formula to the validation cohort. It was found that patients low IRRS had longer in the validation (94 months vs 49 months, P < 0.001, Fig. 3B) and the entire cohort (116 months vs 51 months, P < 0.001, Fig. 3C). In addition, among the 43 patients enrolled in the external validation cohort, 33 were eligible. The survival curves revealed that patients with low-risk scores tended to have longer OS compared with high-risk scores (Fig. S1). Additionally, multivariate analysis showed that the IRRS were significantly correlated with OS (Tables 1 and 2). We then performed the validation of IRRS in the four datasets with the largest sample size. As expected, high IRRS was correlated with a worse survival time in GSE31210, GSE68465, GSE30219, and TCGA-LUAD cohorts (Fig. 3D–G).

We also identified the prediction accuracy of our IRRS model using ROC curves. These results indicated that the 1-, 3-, and 5-year AUC were 0.772, 0.731, and 0.691 in the training cohort, and 0.711, 0.695, and 0.681 in the validation cohort (Fig. 3H and I).

3.2. Correlation analysis between IRRS and clinical features

We further investigated the association between IRRS and clinical features in the LUAD cohorts. In the training cohort, the smoking history, gender, stage, EGFR, KRAS were significantly related to IRRS (Fig. 4A). In the validation cohort, high IRRS was significantly associated with age, gender, EGFR, KRAS and p53 (Fig. 4B). Overall, from the perspective of common gene mutations in LUAD, patients with EGFR mutation, and KRAS wild-type had significantly lower IRRS.

Then, we carried out a comprehensive analysis to analyze the differences among patients with different clinical characteristics. High IRRS is related to poor OS in LUAD patients, regardless of age, smoking status, and TNM stage (Fig. 5A–F). In addition, both
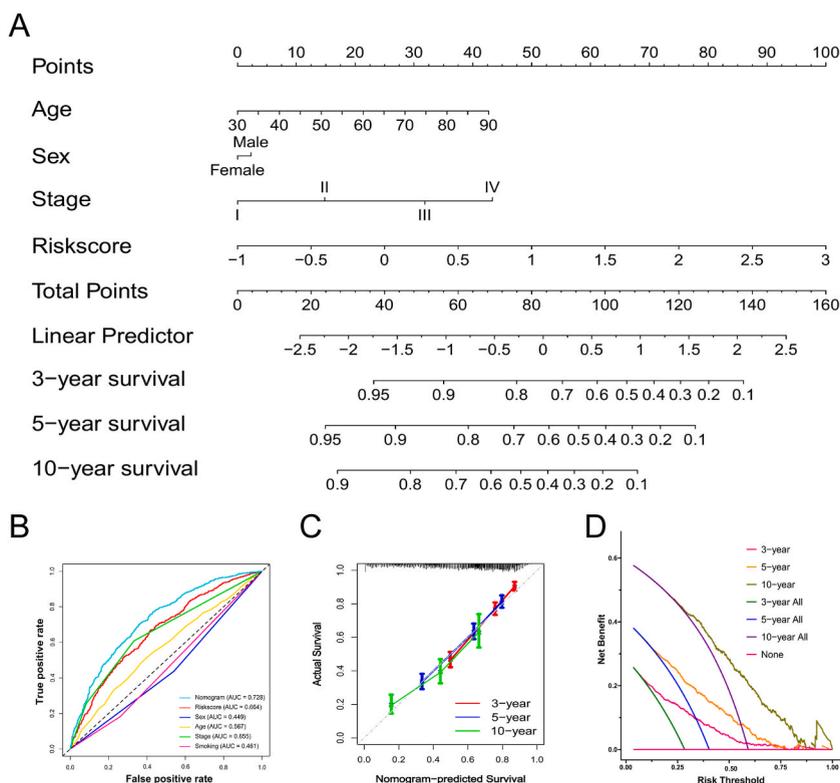
Fig. 6. Prediction model nomogram. (A) Nomogram predicting 3-, 5-, and 10-year survival. (B) The ROC curve of nomogram. (C) Calibration curves for the nomogram. (D) Decision curve analysis for the nomogram.

EGFR-mutant positive and negative patients with low IRRS had a significantly longer OS (Fig. 5G and H). For patients with negative KRAS and p53 mutation status, low IRRS patients had longer OS, while similar results were not observed in patients with positive KRAS and p53 mutation status (Fig. 5I-L).

3.3. Nomogram construction

To create a quantitative prognostic model for LUAD patients, the OS nomogram was created, and the points of LUAD patients can be used to estimate the 3-, 5- and 10-year survival (Fig. 6A). As shown, IRRS contributed the most to predicting survival of LUAD patients. ROC analysis revealed that the AUC was the largest, indicating that the IRRS model had better predictive accuracy than other clinical features (Fig. 6B). Moreover, calibration and decision curve analyses also showed good predictive accuracy for the nomogram (Fig. 6C and D).

3.4. DEG analysis

To investigate the biological significance, DEG analysis was performed between different risk groups. Among the DEGs, 2672 genes were up-regulated, and 484 genes were down-regulated (Fig. 7A). Subsequently, top 20 genes with largest statistically differences were selected, and our extended gene network revealed more interactions of genes (Fig. 7B). Kaplan-Meier curves were also constructed by dividing samples at the median expression level (Fig. S2).

We also applied three pathway analyses, GO, KEGG, and GSEA, to analyze the potential association between DEGs and immunity. GO term analysis revealed that DEGs were closely linked with lymphocyte mediated immunity, antigen presentation, cytokinesis, and T cell mediated cytotoxicity (Fig. 7C). From the results of KEGG enrichment, it was found that 48 KEGG pathways were enriched, including cell cycle, Th17 cell differentiation, and so on (Fig. 7D). We next performed GO and KEGG analyses for the down- or up-regulated genes, and displayed the network diagram of 6 immune-related pathways (Fig. S3). The results presented above showed that the differentially expressed genes can be mapped to immune, cell cycle, and metabolic pathways. Moreover, the results of the GSEA analysis indicated a close association between IRRS and several immune-related pathways (Fig. 7E).

3.5. Tumor microenvironment between two risk groups

Using the NES generated by ssGSEA, a heatmap of 17 TICs for LUAD patients was plotted. It is evident from the heatmap that the

Fig. 7. Bioinformatic analysis of LUAD subtype-specific genes and pathways. (A) Volcano plot of DEGs. (B) Gene Mania networks of the top 20 genes with the largest statistically differences. GO (C), KEGG (D), and GSEA (E) analysis of the DEGs.

two IRRS subtypes differ in the infiltration levels of 17 TICs and 8 immune-related genes (Fig. 8A and B). Additionally, we also estimated the relative abundance of 22 TICs using the CIBERSORT algorithm. The proportions of TICs for each sample were shown in Fig. 9B, and their distributions in different IRRS subtypes were illustrated in the form of boxplots (Fig. 9C). In addition, correlation analyses were carried out to measure correlations between TICs (Fig. S4).

Estimation from the RNA-seq data showed that estimate score, immune score, and stromal score were higher in the low-risk group, while the tumor purity was higher in the high-risk group (Fig. 8C–F). Subsequently, we also analyzed the expression of 65

Fig. 8. Immune landscape of LUAD patients. (A) Heatmap of the 17 immune cells. (B) The riskscore, survival status, and heatmap of 8 immune-related genes. Estimate (C), immune (D), stromal score (E), and tumor purity (F) for different risk groups.

immunomodulators in different risk groups (Fig. 9A). The results indicated that patients with low IRRS had higher immunoregulatory gene expression than patients with high IRRS. Additionally, low-risk group had increased expression of CTLA-4, CD28, CD80 and ICOS, while high-risk group had higher levels of LAG3 and PD-L2 (Fig. 9D–K). Our analysis suggests significant correlations of immune checkpoint genes in different risk groups, which has implications for the personalization of immunotherapy.

3.6. The role of IRRS in predicting immunotherapy benefits

TIDE has proven to be a useful tool to assess immunotherapy efficacy. In this study, patients with low-risk scores had lower TIDE score, implying that they may benefit more from immunotherapy (Fig. 10A–D). We then investigated whether the IRRS could predict immunotherapy response based on four independent datasets. The two IRRS subtypes differ in the expression levels of 8 immune-related genes in the IMvigor210 cohort (Fig. 10E). The survival curves indicated that patients with low-risk scores had significantly longer OS (Fig. 10I). Furthermore, patients with low-risk scores exhibited a more excellent likelihood of response to ICI therapy compared with high-risk patients (Fig. 10J and K). Similarly, the results showed a significant survival benefit for patients with low-risk scores in GSE78220, GSE93157 and CheckMate datasets (Fig. 10F–H).

4. Discussion

The tumor microenvironment is emerging as a critical component that contributes to growth and progression of malignant tumors
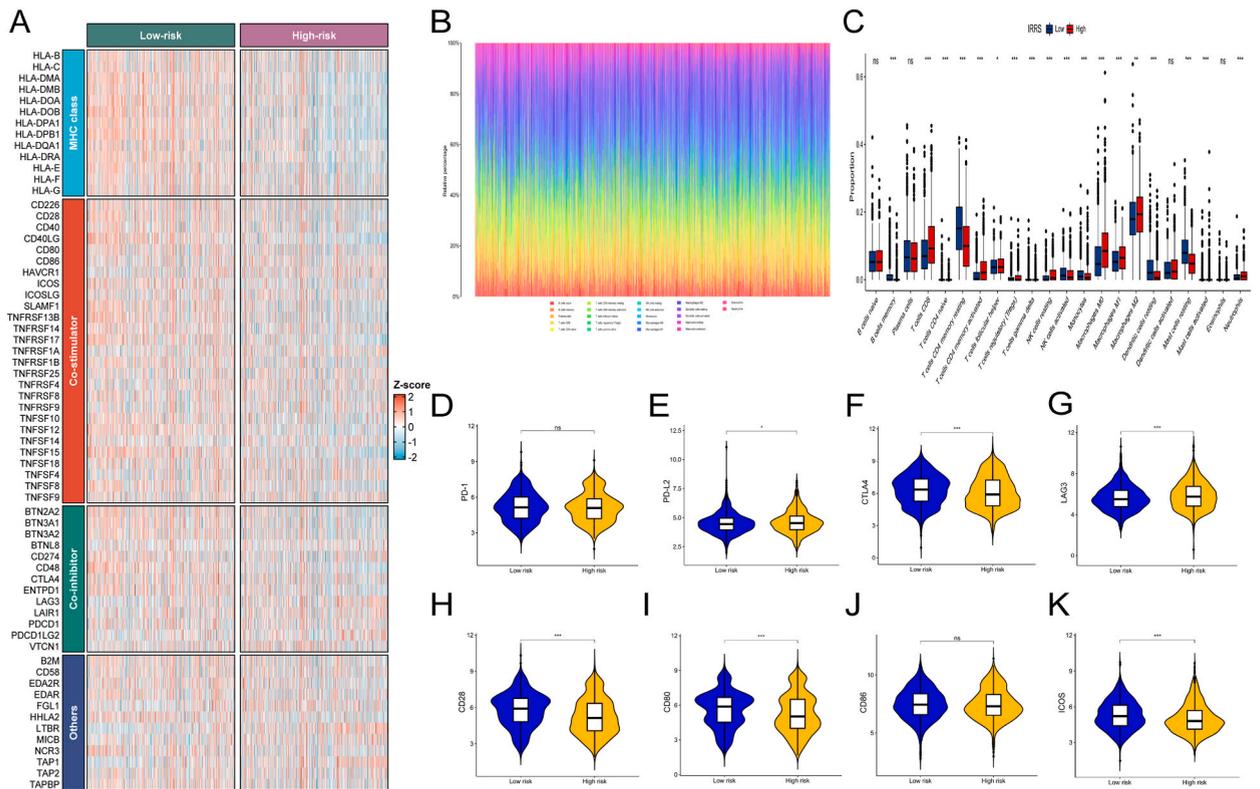
Fig. 9. Immune cell proportions and immunomodulators of the two IRRS subtypes. (A) Heatmap of 65 immune checkpoint profiles. (B) Relative proportions of 22 TICs. (C) Boxplots of the different immune cell proportions. Expression of 8 immune-related genes. (D) PD-1. (E) PD-L2. (F) CTLA4. (G) LAG3. (H) CD28. (I) CD80. (J) CD86. (K) ICOS.

[44–46]. Immune cells play vital roles in the TME, actively interacting with tumor and stromal cells [47–49]. Given the close relationship between TICs and prognosis, the exploration of TICs as a prognostic biomarker is an active research area. Previous studies have explored the role of TICs in several tumor types [50–53]. However, increasing evidence indicates that the TME is complex, and more data are needed to improve the prognostic model.

In our study, 65 TICs were included, and the IRRS model was built on the basis of 17 selected TICs. The results showed that low IRRS were correlated with longer OS in LUAD patients. Moreover, given the large number of patients with different datasets, the IRRS model was further confirmed in the subgroup analysis. As expected, IRRS was stable on independent datasets. These results indicated that the combination of TICs is a potential biomarker for LUAD patients. In addition, a nomogram based on IRRS and clinical characteristics was established to estimate the OS of LUAD patients. Our results showed that IRRS contributed the most points to the nomogram model.

We further explored the IRRS model in the subgroup of different clinical features, and the results indicated that IRRS was significantly related to gender, EGFR, and KRAS status for both the training and validation cohorts. As for other clinical characteristics, consistent results were not observed in both cohorts. Additionally, we also analyzed the association between IRRS and survival. The results showed that significant differences in OS were observed, regardless of age, TNM stage, and EGFR mutation status. However, similar results were not observed for patients with positive KRAS or p53 mutation status. In fact, there was a complex interaction between gene mutation status, environment, and prognosis of cancer patients. As genetic testing technologies rapidly evolve, a deeper and more comprehensive understanding of oncogenesis enables us to understand how cancer will behave in the future. Therefore, we need to further explore TICs in patients with different gene mutation status to better understand their interactions in the varying TIME.

We also performed a DEG analysis to analyze the biological significance between different risk groups. Among the DEGs, top 12 genes with largest statistically differences were selected, and Kaplan-Meier curves were performed by dividing patients at the median expression level. These results are consistent with previous observational studies [54–58]. Subsequently, GO term analysis implied that DEGs were prone to be statistically enhanced in immune-related pathways, including lymphocyte mediated immunity, T cell mediated cytotoxicity, and so on. Aberrant activity of genes and immune-related pathways are frequently correlated with tumorigenesis and tumor progression. In our study, differences in survival benefit between different risk groups could be explained by the DEGs in diverse cancer parameters. The results showed that several pathways, including cell cycle regulation, immune response and metabolic function, could play a major part in the TME. However, more detailed studies are required to examine the relationships between gene expression and cancer prognosis.

Multiple studies have demonstrated that immune checkpoint molecules play significant roles in tumor immune evasion, and ICIs
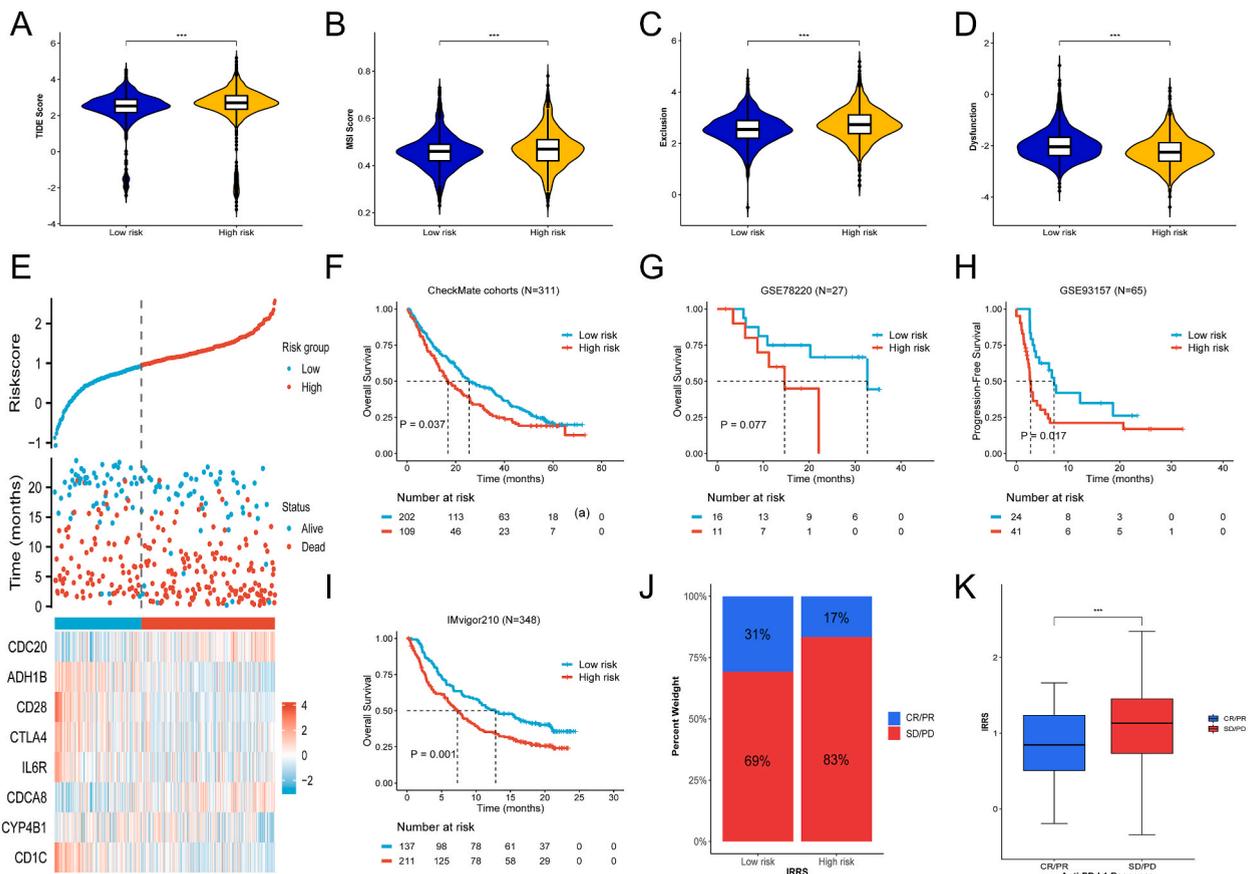
Fig. 10. The prognostic value for immunotherapy. (A) TIDE, (B) MSI, (C) T cell exclusion, and (D) T cell dysfunction score in two risk groups. (E) The riskscore, survival status, and heatmap in the IMvigor210 cohort. Kaplan-Meier curves for OS in different immunotherapy cohorts. (F) Checkmate cohort. (G) GSE78220. (H) GSE93157. (I) IMvigor210. (J) Clinical efficacy for the IMvigor210 cohort. (K) IRRS distribution of clinical efficacy in the IMvigor210 cohort.

have reshaped cancer therapy [59–61]. Therefore, an in-depth analysis between IRRS and immune checkpoint molecules was performed in this study. It was found that patients with low-risk scores had higher levels of CTLA-4, CD28, CD80 and ICOS. At present, durable survival benefit of ICIs has been observed in many cancer types. The significant of immune checkpoint molecules suggests that IRRS could be an effective indicator for patients receiving immunotherapy. Therefore, we analyzed whether the IRRS could predict immunotherapy response based on four independent datasets, and the results showed a significant survival benefit for patients with low IRRS. Additionally, the TIDE prediction scores were also lower in low-risk patients, which suggests that IRRS may be a potential indicator for response to immunotherapy.

However, there were some limitations. First, this is a retrospective observational study using public datasets, and the application of IRRS need to be further validated by future cohort studies. Second, despite the IRRS is designed to include as many TICs as possible, it is still difficult to simulate various complex aspects of the tumor immune microenvironment. Third, the optimal cut-off for IRRS is uncertain, and further clinical validation is required to determine its prognostic value. Despite the above limitations, our study indicated that IRRS is a valuable prognostic marker for patients with LUAD.

5. Conclusions

In summary, the proposed IRRS model is a powerful tool for predicting survival in LUAD. Patients with high IRRS were associated with worse survival outcomes and shorter survival for LUAD patients. In addition, the IRRS model can also help to identify patients who may benefit from immunotherapy. This study will contribute to an enhanced understanding of the differences in the distribution of immune cell infiltration and immunoregulatory gene expression between high- and low-risk patients, which may facilitate individual immunotherapy in the future.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Chinese PLA General Hospital (S2018-203-01), and written informed consent

was obtained from all included patients.

Data availability

The sequencing data used in this study can be downloaded from the TCGA and GEO databases for free. The raw data are available from the corresponding author on reasonable request.

Funding

CRediT authorship contribution statement

Lupeng Qiu: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Zizhong Yang:** Formal analysis, Methodology, Writing – review & editing. **Guhe Jia:** Conceptualization, Formal analysis. **Yanjie Liang:** Methodology, Software. **Sicheng Du:** Resources, Supervision. **Jian Zhang:** Validation, Visualization. **Minglu Liu:** Formal analysis, Project administration. **Xiao Zhao:** Funding acquisition, Visualization, Writing – review & editing. **Shunchang Jiao:** Conceptualization, Data curation, Investigation, Software, Writing – original draft, Writing – review & editing.

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.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33109.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J Clin 68 (6) (2018) 394–424.
[2] M.B. Schabath, M.L. Cote, Cancer progress and priorities: lung cancer, Cancer Epidemiol. Biomarkers Prev. 28 (10) (2019) 1563–1579.
[3] J.A. Barta, C.A. Powell, J.P. Wisnivesky, Global epidemiology of lung cancer, Ann Glob Health 85 (1) (2019).
[4] W. Hu, Z.Y. Bi, Z.L. Chen, C. Liu, L.L. Li, F. Zhang, Q. Zhou, W. Zhu, Y.Y. Song, B.T. Zhan, Q. Zhang, Y.Y. Bi, C.C. Sun, D.J. Li, Emerging landscape of circular RNAs in lung cancer, Cancer Lett. 427 (2018) 18–27.
[5] T.Y. Cheng, S.M. Cramb, P.D. Baade, D.R. Youlden, C. Nwogu, M.E. Reid, The international epidemiology of lung cancer: latest trends, disparities, and tumor characteristics, J. Thorac. Oncol. 11 (10) (2016) 1653–1671.
[6] C. Li, L. Zhang, G. Meng, Q. Wang, X. Lv, J. Zhang, J. Li, Circular RNAs: pivotal molecular regulators and novel diagnostic and prognostic biomarkers in non-small cell lung cancer, J. Cancer Res. Clin. Oncol. 145 (12) (2019) 2875–2889.
[7] N. Duma, R. Santana-Davila, J.R. Molina, Non-small cell lung cancer: epidemiology, screening, diagnosis, and treatment, Mayo Clin. Proc. 94 (8) (2019) 1623–1640.
[8] D.R. Camidge, R.C. Doebele, K.M. Kerr, Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC, Nat. Rev. Clin. Oncol. 16 (6) (2019) 341–355.
[9] D.J. Tandberg, B.C. Tong, B.G. Ackerson, C.R. Kelsey, Surgery versus stereotactic body radiation therapy for stage I non-small cell lung cancer: a comprehensive review, Cancer 124 (4) (2018) 667–678.
[10] L. Osmani, F. Askin, E. Gabrielson, Q.K. Li, Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy, Semin. Cancer Biol. 52 (Pt 1) (2018) 103–109.
[11] S. Huang, J. Yang, N. Shen, Q. Xu, Q. Zhao, Artificial intelligence in lung cancer diagnosis and prognosis: current application and future perspective, Semin. Cancer Biol. 89 (2023) 30–37.
[12] S. Maman, I.P. Witz, A history of exploring cancer in context, Nat. Rev. Cancer 18 (6) (2018) 359–376.
[13] O.M. Ozpiskin, L. Zhang, J.J. Li, Immune targets in the tumor microenvironment treated by radiotherapy, Theranostics 9 (5) (2019) 1215–1231.
[14] C. Jin, G.K. Lagoudas, C. Zhao, S. Bullman, A. Bhutkar, B. Hu, S. Ameh, D. Sandel, X.S. Liang, S. Mazzilli, M.T. Whary, M. Meyerson, R. Germain, P.C. Blainey, J. G. Fox, T. Jacks, Commensal microbiota promote lung cancer development via γδ T cells, Cell 176 (5) (2019) 998–1013.e16.
[15] Y. Zhou, X. Shi, H. Chen, B. Mao, X. Song, L. Gao, J. Zhang, Y. Yang, H. Zhang, G. Wang, W. Zhuang, Tumor immune microenvironment characterization of primary lung adenocarcinoma and lymph node metastases, BioMed Res. Int. 2021 (2021) 5557649.
[16] Y. Chen, L. Tang, W. Huang, Y. Zhang, F.H. Abisola, L. Li, Identification and validation of a novel cuproptosis-related signature as a prognostic model for lung adenocarcinoma, Front. Endocrinol. 13 (2022) 963220.
[17] J. Wu, L. Li, H. Zhang, Y. Zhao, H. Zhang, S. Wu, B. Xu, A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma, Oncogene 40 (26) (2021) 4413–4424.
[18] B. Farhood, M. Najafi, K. Mortezaee, CD8(+) cytotoxic T lymphocytes in cancer immunotherapy: a review, J. Cell. Physiol. 234 (6) (2019) 8509–8521.

[19] Y. Chen, Y. Song, W. Du, L. Gong, H. Chang, Z. Zou, Tumor-associated macrophages: an accomplice in solid tumor progression, J. Biomed. Sci. 26 (1) (2019) 78.

[20] E. Batlle, J. Massagué, Transforming growth factor-β signaling in immunity and cancer, Immunity 50 (4) (2019) 924–940.

[21] Y. Miao, H. Yang, J. Levorse, S. Yuan, L. Polak, M. Sribour, B. Singh, M.D. Rosenblum, E. Fuchs, Adaptive immune resistance emerges from tumor-initiating stem cells, Cell 177 (5) (2019) 1172–1186.e14.

[22] S.Y. Wu, T. Fu, Y.Z. Jiang, Z.M. Shao, Natural killer cells in cancer biology and therapy, Mol. Cancer 19 (1) (2020) 120.

[23] S.S. Wang, W. Liu, D. Ly, H. Xu, L. Qu, L. Zhang, Tumor-infiltrating B cells: their role and application in anti-tumor immunity in lung cancer, Cell. Mol. Immunol. 16 (1) (2019) 6–18.

[24] Z. Dai, T. Liu, G. Liu, Z. Deng, P. Yu, B. Wang, L. Cen, L. Guo, J. Zhang, Identification of clinical and tumor microenvironment characteristics of hypoxia-related risk signature in lung adenocarcinoma, Front. Mol. Biosci. 8 (2021) 757421.

[25] S. Zuo, M. Wei, S. Wang, J. Dong, J. Wei, Pan-cancer analysis of immune cell infiltration identifies a prognostic immune-cell characteristic score (ICCS) in lung adenocarcinoma, Front. Immunol. 11 (2020) 1218.

[26] D. Aran, Z. Hu, A.J. Butte, xCell: digitally portraying the tissue cellular heterogeneity landscape, Genome Biol. 18 (1) (2017) 220.

[27] E. Becht, N.A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, J. Selves, P. Laurent-Puig, C. Sautès-Fridman, W.H. Fridman, A. de Reyniès, Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression, Genome Biol. 17 (1) (2016) 218.

[28] G. Bindea, B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A.C. Obenauf, H. Angell, T. Fredriksen, L. Lafontaine, A. Berger, P. Bruneval, W.H. Fridman, C. Becker, F. Pagès, M.R. Speicher, Z. Trajanoski, J. Galon, Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer, Immunity 39 (4) (2013) 782–795.

[29] F. Finotello, Z. Trajanoski, Quantifying tumor-infiltrating immune cells from transcriptomics data, Cancer Immunol. Immunother. 67 (7) (2018) 1031–1040.

[30] W. Hendrickx, I. Simeone, S. Anjum, Y. Mokrab, F. Bertucci, P. Finetti, G. Curigliano, B. Seliger, L. Cerulo, S. Tomei, L.G. Delogu, C. Maccalli, E. Wang, L. D. Miller, F.M. Marincola, D. Bedognetti, Identification of genetic determinants of breast cancer immune phenotypes by integrative genome-scale analysis, OncoImmunology 6 (2) (2017) e1253654.

[31] M.D. Iglesia, B.G. Vincent, J.S. Parker, K.A. Hoadley, L.A. Carey, C.M. Perou, J.S. Serody, Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer, Clin. Cancer Res. 20 (14) (2014) 3818–3829.

[32] Y.R. Miao, Q. Zhang, Q. Lei, M. Luo, G.Y. Xie, H. Wang, A.Y. Guo, ImmuCellAI: a unique method for comprehensive T-cell subsets abundance prediction and its application in cancer immunotherapy, Adv. Sci. 7 (7) (2020) 1902880.

[33] G. Monaco, B. Lee, W. Xu, S. Mustafah, Y.Y. Hwang, C. Carré, N. Burdin, L. Visan, M. Ceccarelli, M. Poidinger, A. Zippelius, J. Pedro de Magalhães, A. Larbi, RNA-seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types, Cell Rep. 26 (6) (2019) 1627–1640.e7.

[34] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457.

[35] A.J. Nirmal, T. Regan, B.B. Shih, D.A. Hume, A.H. Sims, T.C. Freeman, Immune cell gene signatures for profiling the microenvironment of solid tumors, Cancer Immunol. Res. 6 (11) (2018) 1388–1400.

[36] L. Gautier, L. Cope, B.M. Bolstad, R.A. Irizarry, affy–analysis of Affymetrix GeneChip data at the probe level, Bioinformatics 20 (3) (2004) 307–315.

[37] G.P. Wagner, K. Kin, V.J. Lynch, Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples, Theor. Biosci. 131 (4) (2012) 281–285.

[38] J.T. Leek, W.E. Johnson, H.S. Parker, A.E. Jaffe, J.D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments, Bioinformatics 28 (6) (2012) 882–883.

[39] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.

[40] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.

[41] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287.

[42] K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Treviño, H. Shen, P.W. Laird, D.A. Levine, S.L. Carter, G. Getz, K. Stemke-Hale, G.B. Mills, R.G. Verhaak, Inferring tumour purity and stromal and immune cell admixture from expression data, Nat. Commun. 4 (2013) 2612.

[43] N. Simon, J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for cox's proportional hazards model via coordinate descent, J. Stat. Software 39 (5) (2011) 1–13.

[44] D.C. Hinshaw, L.A. Shevde, The tumor microenvironment innately modulates cancer progression, Cancer Res. 79 (18) (2019) 4557–4566.

[45] K. Nakamura, M.J. Smyth, Myeloid immunosuppression and immune checkpoints in the tumor microenvironment, Cell. Mol. Immunol. 17 (1) (2020) 1–12.

[46] Y. Pan, Y. Yu, X. Wang, T. Zhang, Tumor-associated macrophages in tumor immunity, Front. Immunol. 11 (2020) 583084.

[47] T.F. Gajewski, H. Schreiber, Y.X. Fu, Innate and adaptive immune cells in the tumor microenvironment, Nat. Immunol. 14 (10) (2013) 1014–1022.

[48] J. Goc, C. Germain, T.K. Vo-Bourgais, A. Lupo, C. Klein, S. Knockaert, L. de Chaisemartin, H. Ouakrim, E. Becht, M. Alifano, P. Validire, R. Remark, S. A. Hammond, I. Cremer, D. Damotte, W.H. Fridman, C. Sautès-Fridman, M.C. Dieu-Nosjean, Dendritic cells in tumor-associated tertiary lymphoid structures signal a Th1 cytotoxic immune contexture and license the positive prognostic value of infiltrating CD8+ T cells, Cancer Res. 74 (3) (2014) 705–715.

[49] Y. Huang, B.Y.S. Kim, C.K. Chan, S.M. Hahn, I.L. Weissman, W. Jiang, Improving immune-vascular crosstalk for cancer immunotherapy, Nat. Rev. Immunol. 18 (3) (2018) 195–203.

[50] R. Zhou, J. Zhang, D. Zeng, H. Sun, X. Rong, M. Shi, J. Bin, Y. Liao, W. Liao, Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer, Cancer Immunol. Immunother. 68 (3) (2019) 433–442.

[51] S. Yang, T. Liu, Y. Cheng, Y. Bai, G. Liang, Immune cell infiltration as a biomarker for the diagnosis and prognosis of digestive system cancer, Cancer Sci. 110 (12) (2019) 3639–3649.

[52] G. Mazzaschi, D. Madeddu, A. Falco, G. Bocchialini, M. Goldoni, F. Sogni, G. Armani, C.A. Lagrasta, B. Lorusso, C. Mangiaracina, R. Vilella, C. Frati, R. Alfieri, L. Ampollini, M. Veneziani, E.M. Silini, A. Ardizzoni, K. Urbanek, F. Aversa, F. Quaini, M. Tiseo, Low PD-1 expression in cytotoxic CD8(+) tumor-infiltrating lymphocytes confers an immune-privileged tissue microenvironment in NSCLC with a prognostic and predictive value, Clin. Cancer Res. 24 (2) (2018) 407–419.

[53] B. Li, E. Severson, J.C. Pignon, H. Zhao, T. Li, J. Novak, P. Jiang, H. Shen, J.C. Aster, S. Rodig, S. Signoretti, J.S. Liu, X.S. Liu, Comprehensive analyses of tumor immunity: implications for cancer immunotherapy, Genome Biol. 17 (1) (2016) 174.

[54] P. Wang, L. Zhang, C. Huang, P. Huang, J. Zhang, Distinct prognostic values of alcohol dehydrogenase family members for non-small cell lung cancer, Med Sci Monit 24 (2018) 3578–3590.

[55] J. Han, F. Wang, Y. Lan, J. Wang, C. Nie, Y. Liang, R. Song, T. Zheng, S. Pan, T. Pei, C. Xie, G. Yang, X. Liu, M. Zhu, Y. Wang, Y. Liu, F. Meng, Y. Cui, B. Zhang, Y. Liu, X. Meng, J. Zhang, L. Liu, KIFC1 regulated by miR-532-3p promotes epithelial-to-mesenchymal transition and metastasis of hepatocellular carcinoma via gankyrin/AKT signaling, Oncogene 38 (3) (2019) 406–420.

[56] X. Kang, F. Kong, K. Huang, L. Li, Z. Li, X. Wang, W. Zhang, X. Wu, LncRNA MIR210HG promotes proliferation and invasion of non-small cell lung cancer by upregulating methylation of CACNA2D2 promoter via binding to DNMT1, OncoTargets Ther. 12 (2019) 3779–3790.

[57] T. Kato, Y. Daigo, M. Aragaki, K. Ishikawa, M. Sato, M. Kaji, Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients, J. Surg. Oncol. 106 (4) (2012) 423–430.

[58] Y. Wang, F. Shi, R. Tao, J. Wu, J. Gu, R. Yang, S. Wu, The relationship between UBE2C and AGGF1 overexpression and tumor angiogenesis in non-small cell lung cancer, Cancer Manag. Res. 13 (2021) 5919–5930.

[59] X. He, C. Xu, Immune checkpoint signaling and cancer immunotherapy, Cell Res. 30 (8) (2020) 660–669.

[60] S. Qin, L. Xu, M. Yi, S. Yu, K. Wu, S. Luo, Novel immune checkpoint targets: moving beyond PD-1 and CTLA-4, Mol. Cancer 18 (1) (2019) 155.

[61] D.M. Pardoll, The blockade of immune checkpoints in cancer immunotherapy, Nat. Rev. Cancer 12 (4) (2012) 252–264.