Check for updates

OPEN ACCESS

EDITED BY Hans Raskov, Zealand University Hospital, Denmark

REVIEWED BY

Jin Zhang, I.M. Sechenov First Moscow State Medical University, Russia Chengyun Tang, I.M. Sechenov First Moscow State Medical University, Russia

*CORRESPONDENCE Yuan Tian tytytianyuan@aliyun.com

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 26 August 2022 ACCEPTED 29 September 2022 PUBLISHED 12 October 2022

CITATION

Zhu H, Zheng C, Liu H, Kong F, Kong S, Chen F and Tian Y (2022) Significance of macrophage infiltration in the prognosis of lung adenocarcinoma patients evaluated by scRNA and bulkRNA analysis. *Front. Immunol.* 13:1028440. doi: 10.3389/fimmu.2022.1028440

COPYRIGHT

© 2022 Zhu, Zheng, Liu, Kong, Kong, Chen and Tian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.

Significance of macrophage infiltration in the prognosis of lung adenocarcinoma patients evaluated by scRNA and bulkRNA analysis

Huaiyang Zhu^{1†}, Chunning Zheng^{2†}, Hongtao Liu^{3†}, Fanhua Kong^{4†}, Shuai Kong², Feng Chen⁵ and Yuan Tian^{6.7*}

¹Department of Thoracic Surgery, Shandong Second Provincial General Hospital, Jinan, China, ²Department of Gastrointestinal Surgery, Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, China, ³Department of Pathology, The First Affiliated Hospital of Shandong First Medical University and Shandong Provincial Qianfoshan Hospital, Shandong Medicine and Health Key Laboratory of Clinical Pathology, Shandong Lung Cancer Institute, Shandong Institute of Nephrology, Jinan, China, ⁴Department of Thoracic Surgery, The Affiliated Taian City Centeral Hospital of Qingdao University, Taian, China, ⁵Department of Thoracic Surgery, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, ⁶Department of Totaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Shandong University, Jinan, China, ⁷Radiotherapy Department, Shandong Second Provincial General Hospital, Shandong University, Jinan, China

Purpose: To investigate the significance of macrophage infiltration to the prognosis of lung adenocarcinoma.

Methods: R language bioinformatics analysis technology, was used to obtain macrophage infiltration-related module genes through WGCNA (Weighted Gene Co-Expression Network Analysis). Marker genes of macrophage subtypes were identified using single-cell sequencing of lung adenocarcinoma tissue. Risk score models were constructed and validated using external data cohorts and clinical samples.

Results: Analysis of cohorts TCGA-LUAD, GSE11969, GSE31210, GSE50081, GSE72094 and GSE8894, revealed a negative correlation between macrophage infiltration and survival. Immunohistochemical analyses of clinical samples were consistent with these data. Based on cell-cluster-markers and TAMs-related-genes, TOP8 genes were obtained (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RRM2, and TK1) with a significant association to prognosis. Risk score models including 9 factors (C1QTNF6, FSCN1, KPNA2, GLI2, TYMS, BIRC3, RBBP7, KRT8, GPR65) for prognosis were constructed. The efficacy, stability and generalizability of the risk score models were validated using multiple data cohorts (GSE19188, GSE26939, GSE31210, GSE50081, GSE42127, and GSE72094).

Conclusions: Macrophage infiltration negatively correlates with prognosis in patients with lung adenocarcinoma. Based on cell-cluster-markers and

TAMs-related-genes, both TOP8 genes (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RRM2, TK1) and risk score models using C1QTNF6, FSCN1, KPNA2, GL12, TYMS, BIRC3, RBBP7, KRT8, GPR65 could predict disease prognosis.

KEYWORDS

macrophages, prognosis, lung adenocarcinoma, ScRNA, bulkRNA, infilitration, marker gene, WGCNA (weighted gene co- expression network analyses)

1 Introduction

Lung cancer remains the most common malignancy worldwide and a leading cause of cancer-related death, despite advances in screening and treatment (1, 2). Whether it was for the non-small cell lung cancer (NSCLC) or small cell lung cancer (SCLC) patients, immunotherapy was the most shining one among many treatment methods, which had changed the landscape of anti-tumor therapy and brought anti-tumor therapy into a new era (3-7). However, there were still many details in the screening of immunotherapy benefit populations and related predictors needed to be further elucidated (8-16). Specific macrophage phenotypes can act as indicators of lung cancer prognosis and the efficacy of immunotherapy (17-24). Sequencing technologies and R language based bioinformatics, formerly reported (25-27), can be used for studies in this area (28–30). Based on our previous studies (28–30), we performed bioinformatics analysis and clinical sample validation to identify specific macrophage signatures that can act as indicators of therapeutic efficacy.

2 Methods

2.1 Data analysis

2.1.1 TCGA data

mRNA expression profiles, clinical information, copy number alterations and mutations of GDC TCGA Lung Adenocarcinoma (LUAD) samples were downloaded from https://xenabrowser.net/datapages/. Tumor samples were screened according to sample name. RNA-seq data for 513 tumor samples and 59 paracancerous samples were obtained.

2.1.2 GEO data

Expression data and sample survival information for GSE11969, GSE19188, GSE26939, GSE31210, GSE42127, GSE50081, GSE72094 and GSE8894 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo). Survival information of the samples were summarized as follows: (Supplementary Table 1_train_clin.tsv; Supplementary Table 1_ GSE11969_clin.txt; Supplementary Table 1_ GSE26939_clin.txt; Supplementary Table 1_ GSE31210_clin.txt; Supplementary Table 1_ GSE42127_clin.txt; Supplementary Table 1_ GSE50081_clin.txt; Supplementary Table 1_ GSE72094_clin.txt; Supplementary Table 1_ GSE3994_clin.txt; Supplementary Table 1_ GSE3994_clin.txt; Supplementary Table 1_ GSE3994_clin.txt; Supplementary Table 1_ GSE3081_clin.txt; Supplementary Table 1_ GSE30894_clin.txt). Single-cell sequencing data from GSE131907 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo). A total of 42,995 cells and 29,634 genes were obtained.

2.2 Immune infiltration analysis

Immune infiltration for each sample was calculated using IOBR of the R package for the training set TCGA expression matrix and GEO data, respectively (method = 'cibersort').

2.3 Survival analysis

For survival assessments, R packages "survminer" and "survival" were analyzed and survival curves were constructed based on survival time and status. Differences in prognosis among the groups were assessed.

2.4 Screening of modules corresponding to macrophages using WGCNA

Hierarchical clustering analysis was performed on the TCGA expression matrix using the R package "hclust", "method=average". Phenotypic information was obtained

Abbreviations: UMAP, Uniform Manifold Approximation and Projection; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; LUAD, lung adenocarcinoma; WGCNA, Weighted Gene Co-Expression Network Analysis; TAM, tumor-associated macrophage; GSVA, Gene Set Variation Analysis; FFPE, formalin fixation and paraffin embedding; IHC, immunohistochemical; DAB, diaminobenzidine; DEGs, Differential Expression Genes; LASSO, Least Absolute Shrinkage and Selection Operator.

using the infiltration ratio of macrophages. A correlation between different modules and macrophages was obtained.

2.5 Clustering analysis of samples

The R package "ConsesusClusterPlus" was used to perform consensus clustering analysis. After clustering on the TCGA and GEO data, the optimal number of categories were determined according to the change of area under the CDF curve. The k value of the cluster category ranged from 2 to 6.

2.6 Analysis of single-cell data

Single-cell data were filtered using the R package "seurat" to remove cells with $\geq 20\%$ mitochondrial expression. Data were analyzed using the "seurat" normalization pipeline. To identify tumor-associated macrophage (TAM) populations, marker genes from published studies were used to identify corresponding clusters. TAM populations were selected for standardization analysis using "Seurat".

2.7 Trajectory analysis of single-cell data

Trajectory analysis was performed on TAM subclasses using the R package "monocle" with default parameters. This resulted in differentiation trajectories and key genes determining these trajectories.

2.8 Gene set variation analysis

To investigate differences in the expression patterns of specific TAM isoforms in biological processes, GSVA enrichment analysis was performed using the R package "GSVA". GSVA is a nonparametric, unsupervised method primarily used to assess alterations in signaling pathways and biological processes in samples.

2.9 Construction of risk scoring model

Univariate cox regression analysis was performed on "cellcluster-markers" and "TAMs-related-genes", and genes significantly associated with OS survival were screened at the p<0.05 level. According to the identified prognosis-related genes, the R package 'glmnet' was used to construct a prognosis model (or classifier model) with a 10-fold cross-validation fold using the cox method. Characteristic factors were then screened. Kaplan-Meier survival analysis and ROC curves were used to evaluate the predictive power of the prognostic model.

2.10 Clinical sample validation (sample collection and immunohistochemistry)

Lung Cancer samples were collected from the First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital from June 2012 to February 2020. Written informed consent was provided by all participants. Tumor tissues were surgically resected, formalin fixed and paraffin embedded (FFPE) for histological evaluation. HE-stained and immunohistochemical (IHC) slides were examined by two independent and experienced pathologists according to the WHO criteria.

Samples were IHC stained with mouse anti-human CD68 monoclonal antibodies (MAB-0863, clone MX075) and mouse anti-human CD163 monoclonal antibodies (MAB-0869, clone MX081). CD68 was used as a general surface marker for macrophages, whilst CD163 was used as a marker for M2 macrophages (31). Double-labeled immunohistochemical staining was performed using alkaline phosphatase and horseradish peroxidase conjugated secondary antibodies. Substrates were fast red (AP-Red) and diaminobenzidine (DAB) (Roche Ltd) stained. Slides were processed using an automated Roche BenchMark XT staining system according to the manufacturer's protocol.

3 Results

3.1 Proportion of immune infiltrating cells and the prognostic efficacy of macrophages

CIBERSORT was used to evaluate the levels of immuneinfiltration from different lung adenocarcinoma datasets (TCGA-LUAD, GSE11969, GSE31210, GSE50081, GSE72094, and GSE8894). According to the median macrophage ratio, samples were divided into high- and low-levels of macrophage infiltration. Survival differences between high- and low-groups showed a significant correlation with macrophage infiltration (Figure 1; Supplementary Table 2_train_cibersort.txt; Supplementary Table 2_GSE 11969_cibersort.txt; Supplementary Table 2_GSE31 210_cibersort.txt; Supplementary Table 2_GSE31 210_cibersort.txt; Supplementary Table 2_GSE3081_cibersort.txt; Supplementary Table 2_GSE72094_cibersort.txt; Supplementary Table 2_GSE8894_cibersort.txt).

3.2 Screening of modules corresponding to macrophages

To identify macrophage-related genes related to infiltration, WGCNA module analysis was performed on the training dataset (Supplementary Figure 1, Supplementary Table 3_gene_module.txt).



Genes corresponding to red modules were named "TAMs-relatedgenes" for subsequent analysis.

3.3 TAMs-related-gene-based clustering analysis, molecular typing and prognostic assessments

We analyzed the expression profiles of TAMs-related-genes in samples from different lung adenocarcinoma datasets (TCGA-LUAD, GSE13213, GSE31210, GSE72094, and GSE8894) to construct consistent clustering profiles. Based on cumulative distribution functions and incremental area maps, we selected stable clusters of TAMs-related-genes to obtain multiple subtypes (Supplementary Figures 2A–E, Supplementary Table 4_TCGA_ consensusClass.csv; Supplementary Table 4_GSE13213_consensus Class.csv; Supplementary Table 4_GSE31210_consensusClass.csv; Supplementary Table 4_GSE72094_consensusClass.csv; Supplementary Table 4_GSE8894_consensusClass.csv).

Survival analysis was performed on cluster subtypes from different datasets, revealing significant survival differences (Supplementary Figures 2F–I). Dimensionality reduction analysis was performed on each dataset, revealing significant differences in sample characteristics between different subtypes (Supplementary Figures 2K–O).

3.4 Preprocessing of single-cell data

To further investigate the role of macrophages in lung adenocarcinoma, published single-cell sequencing data of lung adenocarcinoma patients was analyzed (PMC7210975) (32). Gene distribution and mitochondrial gene expression were screened (Supplementary Figures 3A–C). Cells with mitochondrial expression \geq 20% were identified as dead and removed.

3.5 Identification of TAMs in total cells

Markers were used to detect the presence of TAMs in the lung adenocarcinoma single-cell datasets (Supplementary Figures 3D– I). TAMs were then extracted and subtype analysis performed to obtain a TAMs subtype map (Supplementary Figure 3I).

3.6 Screening of differential expression genes among TAMs subsets

To identify marker genes amongst the different TAM subgroups, samples were screened in "Seurat". Dot and violin plots revealed the top5 marker genes for each TAM subtype (Supplementary Figure 4; Supplementary Table 5_TAM_marker_genes.txt).

3.7 Simulation of dynamic changes in macrophages

"Monocle" was used to identify dynamic changes of macrophages in the tumors and cell polarization (Supplementary Figures 5A–C). Cluster 0 could be divided into Cluster 1 and Cluster 2 amongst TAM subtypes. The identified genes were found to regulate differentiation (Supplementary Figure 5D). Gene enrichment analysis on the subtypes of TAM showed that Cluster 2 positively correlated with E2F TARGETS and G2M CHECKPOINT, whilst Cluster 4 negatively correlated with these pathways (Supplementary Figure 5E).

3.8 Screening of prognostic factors based on cell-cluster-markers and TAMS-related-genes using univariate cox regression analysis

Markers of each TAM subtype and TAMs-related-genes were used to identify genes related to the prognosis. Samples were divided into high- and low-expression groups according to the median of gene expression. Univariate Cox analysis was performed and survival curves of the top8 prognostic genes were displayed (Figures 2A-H; Supplementary Table 6_cox_significant.txt).

3.9 Construction of risk score models and evaluation of the prognostic efficacy

Based on the "GLMNET" of the R package, LASSO (Least Absolute Shrinkage and Selection Operator) regression analysis was used to construct a regression model for the expression matrix of prognosis related genes corresponding to "Cell-Cluster-Markers" and "Tams-Related-Genes". By analysis, when the value of the freedom degree was 9, the model was accurate (Figures 2I–O; Supplementary Table 7_ forest.univariate_cox.txt, Supplementary Table 7_Signature_ Coef.txt). The calculation formula of the risk score model are listed as follows:

Risk Score = 0.0354754835*C1QTNF6 (Expression Value) + 0.0023344103* FSCN1 (Expression Value) + 0.0022298189*GLI2 (Expression Value) + 0.0001616254 * KPNA2 (Expression value) + 0.0005176419*TYMS (Expression Value) + 0.0037498174 *BIRC3 (Expression Value) + 0.003257017*RBBP7 (Expression Value) + 0.0002465129 *KRT8 (Expression Value) - 0.0263442444 *GPR65 (Expression Value). Kaplan-Meier survival curves indicated a significant difference in survival between high and low risk groups. The ROC curve indicated high performance of the risk score model.

3.10 Validation of risk score prognostic models in external datasets

To further verify the stability of the risk score model, external and independent data GSE19188, GSE26939, GSE31210, GSE50081, GSE42127 and GSE72094 were used to verify predictive efficacy. Through Kaplan-Meier survival analysis, the constructed risk score model performed well for all external data predictions (Figure 3).

3.11 Robust principal component analysis of risk scoring models in clinical factors

To confirm the stability of the risk score model according to clinical characteristics, differences in survival status between high- and low-risk groups in terms of age, gender, radiotherapy, clinical characteristics and Pathlogic M were explored. Significant differences in survival between high- and low-risk groups were observed in those aged ≥ 60 and ≤ 60 years (Figures 4A, B; Supplementary Table 8_clinical_inf.txt). Similar differences were observed between gender subgroups (Figures 4C, D). In the radiotherapy group, differences between high- and low-risk groups were more pronounced (Figures 4E, F). In Pathologic M (Figures 4G, I), significant differences between high- and low-risk groups were observed for M0, indicative of higher stability.

3.12 Differences in risk score models among cancer clinical factors

To investigate the relationship between the risk score model and clinical characteristics, specific features were selected for analysis. The risk score was found to be related to radiation therapy, pathologic T and tumor stage. No significant relationship to age or gender were observed (Supplementary Figure 6).

3.13 Evaluation of risk score models through univariate and multivariate cox regression analysis

To determine whether the risk score model could act as an independent prognostic factor for cancer, the "coxph()" function in the R package "survival" was adopted for univariate and multivariate regression analysis on training and test sets, respectively. We found that in all validation and test sets, the p value of the risk score was ≤ 0.05 (Figure 5; Supplementary Table 9_TCGA_clinical.multivariate_cox.txt; Supplementary Table 9_TCGA_ clinical.univariate_cox.txt; Supplementary



Survival analysis of top8 genes is significantly associated with prognosis. Abscissa axis: survival time. Ordinate axis: survival probability. Colors: differential gene expression. Survival analysis of (A) C1QTNF6, (B) CCNB1, (C), FSCN1, (D) HMMR, (E) KPNA2, (F) PRC1, (H) RRM2, and (I) TK1. (J) Construction of the risk score model and evaluation of its prognostic efficacy. Forest plots of genes included in the risk score model. Left column: 9 genes included in the risk score model. Left column: orresponding forest plot. (K) Risk score plot for cancer samples (line graph). (L) Risk score plot for cancer samples (scatter plot graph). (M) Dynamic process diagram of variables screened by LASSO regression analysis and selection process diagram of the cross-validation parameter λ . (N) Survival analysis of the training dataset. Abscissa axis: survival time; ordinate axis: survival probability. (O) ROC curve of training datasets. Abscissa axis: specificity; Ordinate axis: sensitivity. Colors represent different years.



Validation of the risk score model with external independent data. A1: Validation (ROC curve) of the risk score model using external independent data GSE19188; A2: survival analysis. B1: Validation (ROC curve) of the risk score model using GSE26939. B2: Survival analysis using GSE26939. C1: Validation (ROC curve) of the risk score model using GSE31210. C2: Validation (survival analysis) using GSE31210. D1: Validation results (ROC curve) of the risk score model using GSE50081. D2: (survival analysis) of the risk score model using GSE50081. E1: Validation results (ROC curve) of the risk score model using GSE42127. E2: survival analysis using E42127. F1: Validation (ROC curve) of the risk score model using GSE42127. E2: survival analysis using E42127. F1: Validation (ROC curve) of the risk score model using GSE72094. Abscissa axis: survival time; Ordinate axis: survival probability. Colors: different risk groups.



Kaplan-Meier survival analysis between high- and low risk groups. Abscissa axis: survival time; Ordinate axis: survival probability. Colors: different risk groups. (A) Kaplan-Meier survival analysis between high- and low-risk groups in those aged \leq 60 years. (B) >60 years. (C) Female patients. (D) Male patients. (E) + Radiation therapy. (F) - Radiation therapy. (G) M:M0. (H) M: M1. (I) M: Mx.

Table 9_GSE19188.multivariate_cox_result.txt; Supplementary Table 9_GSE19188.univariate_cox_result.txt; Supplementary Table 9_GSE26939.multivariate_cox_result.txt; Supplementary Table 9_GSE26939.univariate_cox_result.txt; Supplementary Table 9_GSE42127.multivariate_cox_result.txt; Supplementary Table 9_GSE42127.univariate_cox_result.txt; Supplementary Table 9_GSE50081.multivariate_cox_result.txt; Supplementary Table 9_GSE50081.multivariate_cox_result.txt; Supplementary Table 9_GSE50081.univariate_cox_result.txt; Supplementary Table 9_GSE50081.univariate_cox_result.txt; Supplementary Table 9_GSE72094.multivariate_cox_result.txt). This indicated that the risk score model was an accurate independent prognostic factor for cancer.

3.14 Construction of a nomogram model of risk scores and clinical factors to predict cancer progression

We next sought to apply the risk scoring model to the prediction of cancer progression in the clinic. The R package "rms" was adopted to construct a nomogram using a variety of clinical features. Calibration curves were used to calculate 1, 2, 3, and 5-year survival times (Supplementary Figure 6, Supplementary Table 10_nomogram_patient_info_part.txt). All survival calibration curves were near the 45° slope, indicating high accuracy of the nomogram.

۰ <u> </u>	Trainin	g dataset univaria	ble Cox regression	analysis		ь	Training	dataset multivaria	able Cox regressior	n analysis	
Class	Hazard Ratio(HR)	lower 95%CI	upper 95%CI		pvalue	Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%CI		pva
Risk score	2.62	1.95	3.51		1.23e-10	Risk score	2.01	1.39	2.89		0.00
Age	1.01	0.99	1.02	•	0.35	Age	1.02	0.99	1.04	•	0.
Gender	1.04	0.78	1.39	-	0.77	Gender	0.91	0.64	1.29	-	0.
Radiation therapy	2.16	1.45	3.21		0.00014	Radiation therapy	1.47	0.96	2.26		0
pathologic T	1.53	1.27	1.85	+	5.78e-06	pathologic T	1.14	0.91	1.43	+	0
Tumor stage	1.68	1.46	1.92	+	1.84e-13	Tumor stage	1.38	1.14	1.67	+	0.0
				0 0.5 1 1.5 2 25 3 35 Coefficient						0 05 1 15 2 25 3 Coefficient	
)	GSE1	19188 univariable	Cox regression a	nalysis		D	GSE19	188 multivariate	Cox regression a	nalysis	
Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%CI	-	pvalue	Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%CI	-	p
Risk score	2.2103	1.2257	3.9858		0.0084	Risk score	2.0206	1.113	3.6671		0
Gender	2.1247	1.0607	4.2563		0.0335	Gender	1.8775	0.93	3.7902	· · · · · · · · · · · · · · · · · · ·	0
•	GSE	26939 univariable	Cox regression an	0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 Coefficient		F	GSE2	6939 multivariate	Cox regression an	0 0.5 1 1.5 2 2.5 3 3.5 4 Coefficient	
Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%Cl	aijoio	pvalue	Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%Cl	aijoio	pv
Risk score	2.5219	1.5156	4.1963		0.00036975	Risk score	2.969	1.3592	6.4853	_ •	0.00
Stage.ch1	1.2069	1.0168	1.4326	•	0.03149	Stage.ch1	1.427	1.0997	1.8516	+	0.00
Gender	1.6356	0.9639	2.7754	_	0.0682	Gender	1.6761	0.8319	3.377		0.1
Age	1.0275	1.0033	1.0523	•	0.02518	Age	1.0198	0.9786	1.0628	•	0.
smoking.status	0.9775	0.4446	2.1491		0.9549	smoking.status	0.6086	0.0668	5.5441	•	0.6
tp53.mut	0.8151	0.4597	1.4453	-	0.4843	tp53.mut	1.2412	0.5383	2.8622		0.6
kras.mut	1.3037	0.7046	2.4122		0.3981	kras.mut	0.396	0.1382	1.1347	←	0.0
<u> </u>				0 05 1 15 2 25 3 35 4 Coefficient		н				0 2 3 4 5 6 Coefficient	
	GSE4	+2127 Univariable	Cox regression a	naiysis			GSE42	127 multivariate	Cox regression a	naiysis	
Class	Hazard Ratio(HR)	lower 95%Cl	upper 95%Cl		pvalue	Class	Hazard Ratio(HR)	lower 95%CI	upper 95%CI		F
Risk score	1.4596	0.873	2.4405		0.0149	Risk score	1.3346	0.7906	2.2526		0
Gender	1.6692	0.983	2.8346		0.0578	Gender	1.2265	0.6968	2.1588		C
Age	1.0411	1.0121	1.0709	•	0.0051	Age	1.0347	1.0047	1.0657	•	C
	1.2421	1.084	1.4233		0.00179	Stage	1.183	1.0237	1.3671		(
Stage				Coefficient				0081 multivariate	Cox regression ar	Coefficient	
Stage	GSE	50081 univariable	Cox regression a	nalysis		J	GSE5				
Stage	GSE Hazard Ratio(HR)	50081 univariable	e Cox regression an upper 95%Cl	nalysis	pvalue	Class	GSE5 Hazard Ratio(HR)	lower 95%Cl	upper 95%CI		
Stage Class Risk score	GSE Hazard Ratio(HR) 1.964	50081 univariable lower 95%Cl 1.2421	e Cox regression ar upper 95%Cl 3.1054	nalysis	pvalue 0.0038	Class Risk score	GSE5 Hazard Ratio(HR) 1.7757	lower 95%Cl	upper 95%Cl 2.857		(
Class Risk score Age	GSE: Hazard Ratio(HR) 1.964 1.0187	50081 univariable lower 95%Cl 1.2421 0.9942	e Cox regression at upper 95%Cl 3.1054 1.0438	•	pvalue 0.0038 0.1348	J Class Risk score Age	GSE5/ Hazard Ratio(HR) 1.7757 1.0177	lower 95%Cl 1.1036 0.9936	upper 95%Cl 2.857 1.0424	•	(
Stage Class Risk score Age Gender	GSE: Hazard Ratio(HR) 1.964 1.0187 1.9335	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895	e Cox regression at upper 95%CI 3.1054 1.0438 3.1426	• •	pvalue 0.0038 0.1348 0.0078	Class Risk score Age Gender	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085	lower 95%Cl 1.1036 0.9936 1.264	upper 95%Cl 2.857 1.0424 3.4393	•	0
Stage Class Risk score Age Gender Stage	CSE Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881	e Cox regression au upper 95%CI 3.1054 1.0438 3.1426 1.62	• • • • • • • • • • • • • • • • • • •	pvalue 0.0038 0.1348 0.0078 0.0052	Class Risk score Age Gender Stage	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746	lower 95%Cl 1.1036 0.9936 1.264 1.0344	upper 95%Cl 2.857 1.0424 3.4393 1.5704	•	0
Stage Class Risk score Age Gender Stage Smoking.status	CSE: Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276 1.0079	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872	Cox regression and upper 95%CI 3.1054 1.0438 3.1426 1.62 1.2906	• • • • • • • • • • • • • • • • • • •	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496	J Class Risk score Age Gender Stage Smoking status	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469	lower 95%Cl 1.1036 0.9936 1.264 1.0344 0.8867	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836	• • •	
Stage Class Risk score Age Gender Stage Smoking.status	GSE Hazard Ratio(HK) 1.984 1.0187 1.9335 1.3276 1.0079 GSE	50081 univariable lower 95%C1 1.2421 0.9942 1.1895 1.0881 0.7872	Cox regression an upper 95%Cl 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression an	• • • • • • • • • • • •	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496	J Class Risk score Age Gender Stage Smoking.status	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469 GSE7.	lower 95%Cl 1.1036 0.9936 1.264 1.0344 0.8867 2094 multivariate	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836		
Stage Class Risk score Age Gender Stage Smoking.status Class	GSE: Hazard Rato(HR) 1.964 1.0167 1.9335 1.9276 1.0079 GSE Hazard Rato(HR)	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872 572094 univariable lower 95%Cl	Cox regression at upper 95%Cl 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression an upper 95%Cl	Allysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496	J Class Risk score Age Gender Stage Smoking.status L Class	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469 GSE7. Hazard Ratio(HR)	Iower 95%Cl 1.1036 0.9936 1.264 1.0344 0.8867 2094 multivariate Iower 95%Cl	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%Cl	• • • • • • • • • • • • • • • • • • • •	((((
Stage Class Risk score Age Gender Stage Smoking.status Class Risk score Risk score	GSE: Hazard Rato(HR) 1,964 1,0167 1,9335 1,0276 1,0079 GSE Hazard Rato(HR) 2,2605	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133	Cox regression at upper 95%/Cl 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression an upper 95%/Cl 3.3768	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05	J Class Risk score Age Gender Stage Smoking status L Class Risk score	GSE5 Hazard Rate(HR) 1.7757 1.0177 2.085 1.2746 1.1469 GSE7. Hazard Rate(HR) 1.4477	lower 95%Cl 1.1036 0.9936 1.264 1.0344 0.8867 2094 multivariate lower 95%Cl 0.6429	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%Cl 3.2598	• • • • • • • • • • • • • • • • • • •	((((((((((((()))))))))))))
Stage Class Risk score Age Gender Stage Smoking status Class Risk score Age	GSE Hazard Rato(HR) 1.984 1.0167 1.9335 1.3276 1.0079 GSE Hazard Rato(HR) 2.2805 1.0069	50081 univariable lower 95%Cl 1.2421 0.9842 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133 0.9676	Cox regression an upper 95%/Cl 3.1064 1.0438 3.1426 1.62 1.2906 Cox regression an upper 95%/Cl 3.3768 1.0267	alysis	pvalue 0.0038 0.1348 0.0076 0.0052 0.9496 pvalue 6.7907e-05 0.4825	J Class Risk score Age Gender Stage Smoking.status L Class Risk score Age	GSE5 Hazard Rate(HR) 1.7757 1.0177 2.085 1.2746 1.1469 GSE7 Hazard Rate(HR) 1.4477 1.0112	lower 95%Cl 1.1036 0.9936 1.264 1.0344 0.8867 2094 multivariate lower 95%Cl 0.6429 0.9725	upper 95%C1 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%C1 3.2598 1.0514	allysis	0 0 0 0 0 8 8 8 0 0 0
Stage Class Class Risk score Age Gender Stage Class Risk score Age Gender	GSE Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276 1.0079 GSE Hazard Ratio(HR) 2.2605 1.0069 1.5519	50081 univariable lower 95%Cl 1.2421 0.9942 1.0895 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133 0.9876 1.0722	Cox regression at upper 95%Cl 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression an upper 95%Cl 3.3768 1.0267 2.2462	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05 0.4825 0.0198	J Class Risk score Age Gender Stage Smoking.status L Class Risk score Age Gender	GSE5 Hazard Rate(HR) 1.7757 1.0177 2.085 1.2746 1.1469 GSE7. Hazard Rate(HR) 1.4477 1.0112 1.4613	lower 95%Cl 0.8867 2094 multivariate lower 95%Cl 0.8429 0.9725 0.6596	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%Cl 3.2598 1.0514 3.2373	• • • • • • • • • • • • • • • • • • •	0 0 0 0 7 7 7 0 0
Stage Class Class Risk score Age Gender Stage Class Risk score Age Gender Kras.mut	GSE Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276 1.0079 GSE Hazard Ratio(HR) 2.2605 1.0069 1.5519 0.0667	50081 univariable lower 95%Cl 1.2421 0.9842 1.1895 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133 0.9876 1.0722 0.4722	2 Cox regression an upper 95%Cl 3.1054 1.0438 3.1426 1.62 1.2906 2 Cox regression an upper 95%Cl 3.3768 1.0267 2.2462 0.9987	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05 0.4825 0.0198 0.0482	J Class Risk score Age Gender Stage Smoking.status Class Risk score Age Gender kras.mut	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.7469 GSE7 Hazard Ratio(HR) 1.4477 1.0112 1.4413 1.0077	lower 95%C1 0.9396 1.204 1.0344 0.8667 2094 multivariate lower 95%C1 0.6429 0.6725 0.6596 0.4532	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression at upper 95%Cl 3.2598 1.0514 3.2373 2.2406		2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
Stage Class Risk score Age Gender Stage Smoking status Class Risk score Age Gender Kras.mut Smoking status	C65E Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276 1.0079 C65E Hazard Ratio(HR) 2.2805 1.0069 1.5519 0.6867 1.0537	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872 572094 univariable fower 95%Cl 1.5133 0.9876 1.0722 0.4722 0.7815	Cox regression ai upper 95%CI 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression ai upper 95%CI 3.3768 1.0267 2.2462 0.9987 1.4208	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05 0.4825 0.0198 0.0482 0.0482 0.7311	J Class Risk score Age Gender Stage Smoking.status L Class Risk score Age Gender kras.mut Smoking.status	CSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469 CSE7. Hazard Ratio(HR) 1.4477 1.0112 1.4613 1.0077 1.0056	lower 95%C1 1.1036 0.9936 1.264 1.0344 0.8867 2094 multivariate lower 95%C1 0.6429 0.9725 0.6596 0.4532 0.4687	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%Cl 3.2598 1.0514 3.2373 2.2406 2.1575		2 C C C C C C C C C C C C C C C C C C C
Stage Class Class Risk score Age Gender Stage Smoking status Class Risk score Age Gender kras.mut Smoking status Stage	CSEE Hazard Ratio(HR) 1.964 1.0187 1.9335 1.3276 1.0079 CSEE Hazard Ratio(HR) 2.2605 1.0069 1.5519 0.6867 1.0537 1.2792	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133 0.9876 1.0722 0.4722 0.7815 1.1662	Cox regression an upper 95%/Cl 3.1054 1.0438 3.1426 1.62 1.2906 Cox regression an upper 95%/Cl 3.3768 1.0287 2.2462 0.9987 1.4208 1.4208	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05 0.4825 0.0198 0.0482 0.0482 0.7311 1.7909e-07	J Class Risk score Age Gender Stage Smoking.status L Class Risk score Age Gender kras.mut Smoking.status Stage	GSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469 0.5627 Hazard Ratio(HR) 1.4477 1.0112 1.4413 1.0077 1.0056 1.0982	lower 95%C1 1.1036 0.9936 1.284 1.0344 0.8867 2094 multivariate lower 95%C1 0.6429 0.9725 0.6556 0.4532 0.4687 0.8891	upper 95%C1 2.857 1.0424 3.4393 1.5704 1.4836 Cox regression ar upper 95%C1 3.2598 1.0514 3.2373 2.2406 2.1575 1.352		2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
Stage Class Class Risk score Age Gender Stage Smoking status Class Gender Kras.mut Smoking status Stage Stag	GSE Hazard Rato(HR) 1.964 1.0107 1.9335 1.3276 1.0079 CSE Mazard Rato(HR) 2.2605 1.0069 1.5519 0.8667 1.0537 1.0537 1.2792 1.0399	50081 univariable lower 95%Cl 1.2421 0.9942 1.1895 1.0881 0.7872 572094 univariable lower 95%Cl 1.5133 0.9876 1.0722 0.4722 0.7815 1.1662 0.989	Cox regression at upper 95%/Cl 1.0438 3.1054 1.0438 1.426 1.62 1.2206 Cox regression an upper 95%/Cl 3.3768 1.0267 2.2462 0.9987 1.4208 1.4208 1.4032 1.019	alysis	pvalue 0.0038 0.1348 0.0078 0.0052 0.9496 pvalue 6.7907e-05 0.4825 0.0198 0.0482 0.7311 1.7909e-07 0.6076	J Class Risk score Age Gender Stage Smoking.status L Class Risk score Age Gender kras.mut Smoking.status Stage tp53.mut	CSE5 Hazard Ratio(HR) 1.7757 1.0177 2.085 1.2746 1.1469 CSE7. Hazard Ratio(HR) 1.4477 1.0112 1.4473 1.0017 1.0056 1.0982 1.0028	lower 95%C1 1.1036 0.9996 1.284 1.3344 0.8867 2094 multivariate 2094 g5%C1 0.6429 0.9725 0.6596 0.4532 0.4632 0.4687 0.8821 0.3868	upper 95%Cl 2.857 1.0424 3.4393 1.5704 1.4836 • Cox regression ar upper 95%Cl 3.2598 1.0514 3.2373 2.2406 2.1575 1.352 1.019	• • • • • • • • • • • • • •)) 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

Univariate and multivariate cox regression analysis of the risk score model in training and validation datasets. (A) Univariate cox regression analysis. (B) Multivariate cox regression analysis of the fisk score model in training and validation datasets. (A) Univariate cox regression analysis. (B) Multivariate cox regression analysis. (C) GSE19188 univariate cox regression analysis. (D) GSE19188 multivariate cox regression analysis. (E) GSE26939 univariate cox regression analysis. (F) GSE26939 multivariate cox regression analysis. (G) GSE42127 univariate cox regression analysis. (I) GSE42127 multivariate cox regression analysis. (I) GSE50081 univariate cox regression analysis. (J) GSE50081 multivariate cox regression analysis. (K) GSE72094 univariate cox regression analysis. (L) GSE72094 multivariate cox regression analysis.

3.15 Prediction of immunotherapy efficacy amongst subtypes

We next investigated whether the risk score model could predict the prognosis of immunotherapy. Data were calculated using the risk score model and the K-M survival status between high- and low-risk groups evaluated (Supplementary Figure 7A, Supplementary Table 11_Immune_treatment.xlsx). Upon statistical analysis of the distribution of CR/PR and PD/SD, the proportion of treatment response rates significantly differed between high- and low-risk groups (Supplementary Figure 7B, chi-square test p= 0.004133). No significant differences in the risk scores between the different treatment response groups were observed (Supplementary Figure 7C).

3.16 Overall survival analyses of M1 and M2 macrophage subtypes in patients with lung cancer

A total of 32 patients with lung cancer were evaluated for M1 and M2 macrophage subtypes. Samples were stained using double-labeled IHC. The majority of patients were in pathological Stage II (62.5%) and the dominant histopathological type was adenocarcinoma (68.8%). The clinicopathological characteristics of the lung cancer patients are shown in (Table 1).

To identify M1 and M2 macrophage subtypes, CD68 and CD163 antibodies were used for double-labeled IHC staining. CD68 (brown/yellow) as a surface marker for all macrophages primarily localized to the cytoplasm, whilst CD163 (red)

TABLE 1 Basic characteristics of enrolled clinical samples.

Characteristic	levels	Overall
n		32
Age, n (%)	>65	12 (37.5%)
	≤65	20 (62.5%)
Gender, n (%)	Female	16 (50%)
	Male	16 (50%)
T stage, n (%)	T1	5 (15.6%)
	T2	20 (62.5%)
	Т3	7 (21.9%)
N stage, n (%)	N0	21 (65.6%)
	N1	8 (25%)
	N2	3 (9.4%)
Pathological Stage, n (%)	Ι	8 (25%)
	II	20 (62.5%)
	III	4 (12.5%)
Histologic type, n (%)	Adenomcarcinoma	22 (68.8%)
	Mucoepidermoid carcinoma	2 (6.2%)
	Squamous cell carcinoma	8 (25%)

localized to the plasma membrane. M2 macrophages were identified through double staining for CD68 and CD163. M1 macrophages were identified through staining with CD68 alone. Representative IHC images are shown in (Figures 6A, B).

The prognostic value of macrophage infiltration was next evaluated. Total macrophages, M2 to M1, and M2 macrophage infiltration were identified as detrimental to patient survival (Figures 6C, E, F), whilst M1 macrophage infiltration was beneficial to prognosis (Figure 6D). The infiltration of M1 macrophages in adenocarcinoma was significantly higher than that in squamous cell carcinoma of lung cancer. No significant differences in M2 nor total macrophage infiltration were observed between these two histological subtypes (Figure 6G).

4 Discussion

Macrophages with different phenotypes are frequently cited as indicators of the prognosis of lung cancer patients and the efficacy of immunotherapy (17–24). In our preliminary analyses, macrophage infiltration, rarely reported in lung cancer, had a significant detrimental effect on the prognosis of lung cancer patients (Figure 1). These data were consistent across cohorts (Figures 1A–F) and further verified in follow-up immunohistochemical analysis of clinical samples (Figure 6C). Collectively, these data highlight how macrophages not only act as innate immune cells to regulate immunological responses (23), but play an important role in the prognosis of lung cancer. This lays the foundation for subsequent module analysis based on macrophage infiltration (Supplementary Figure 1).

Based on the expression profiles of TAMs-related-genes, a consistent clustering profile was constructed (Supplementary Figures 2A–E). Significant differences in both survival analysis and PCA (Supplementary Figures 2F–O) were observed. These apparent differences were further identified in single-cell data (Supplementary Figure 3) confirming the importance of macrophages to the prognosis of lung cancer patients (17–24, 32). These data also highlight the need for further refinement of relevant factors to more favorably evaluate patient prognosis.

Given the advantages and progress of single-cell sequencing in lung cancer immunity (33–35), the single-cell data was further analyzed (Supplementary Figures 4, 5, and Figure 2) (32). Based on cell-cluster-markers and TAMs-related-genes, TOP 8 genes (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RRM2, and TK1) significantly associated with prognosis were obtained (Figure 2). These have obvious benefits to clinicians for the assessment of patient prognosis (36–49). The same data were used to construct a risk score model containing 9 factors (C1QTNF6, FSCN1, KPNA2, GLI2, TYMS, BIRC3, RBBP7, KRT8, and GPR65) for prognostic evaluation (Figure 2) (50– 55). The model was validated using external data cohorts (Figure 3) and identified as robust and accurate for prognostic evaluation (Figure 4). Significant differences in the risk scores



Validation analysis of clinical samples. A1: Immunohistochemical staining of M1 macrophages. A2: Enlargement of the boxed regions. B1: Immunohistochemical staining of M2 macrophages. B2: Enlargement of the boxed regions. Images were obtained at 40×10 magnification under a light microscope. (C) Kaplan-Meier survival analysis of macrophage infiltration. (D) M1 macrophage infiltration. (E) M2 macrophage infiltration. (F) M2 to M1 macrophage infiltration. (G) M1, M2 and total macrophage infiltration between adenocarcinoma and squamous cell carcinoma. Horizontal axis: survival time. Vertical axis: survival probability. Colors: macrophage infiltration.

10.3389/fimmu.2022.1028440

were observed for clinical characteristics including radiation therapy, pathologic T, and Tumor stage (Supplementary Figures 6A–E). This further highlighted the efficiency of the risk score to predict therapeutic efficacy.

Through univariate and multivariate cox regression analysis, the risk score model held utility as an independent prognostic factor for cancer, further affirming its clinical benefits (Figure 5). Furthermore, cancer progression could be more accurately predicted using nomogram models constructed based on risk scores and clinical factors (Supplementary Figures 6F–J). For prognostic assessments of immunotherapy, the risk score model could also act as an accurate evaluation tool (Supplementary Figure 7). Upon immunohistochemical analysis of clinical tissue samples to verify the correlation between the macrophage phenotype and patient prognosis, similar conclusions were obtained (Table 1; Figures 1, 6). Macrophage infiltration, particularly for the M2 phenotype, were not conducive to the prognosis and survival of patients, consistent with previous studies (20–24, 56, 57).

We used WGCNA to identify macrophage infiltrationrelated module genes and single-cell sequencing of lung adenocarcinoma tissue to identify marker genes of macrophage subtypes. This permitted the construction of a risk assessment model with high prognostic efficacy. The model performed well on external and independent datasets. Immunohistochemistry analysis of clinical samples were consistent with our data. We therefore infer that the risk score has both high clinical practicability and application.

5 Conclusion

Macrophage infiltration was negatively correlated with prognosis for patients with lung adenocarcinoma. Based on cell-cluster-markers and TAMs-related-genes, both TOP8 genes (C1QTNF6, CCNB1, FSCN1, HMMR, KPNA2, PRC1, RRM2, TK1) and the risk score model containing 9 risk factors (C1QTNF6, FSCN1, KPNA2, GL12, TYMS, BIRC3, RBBP7, KRT8, GPR65) had a high efficacy for the prediction of prognosis.

Data availability statement

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

Ethics statement

The ethics involved in the clinical samples in this study have been approved by the Ethics Committee of Qianfoshan Hospital of Shandong Province (2022-S527). The patients/participants provided their written informed consent to participate in this study.

Author contributions

Corresponding authors (YT) conceived the idea and designed this study. TY was responsible for the final submission. YT, CZ and HZ were responsible for data collection, partial data analysis and drafting the manuscript. HL performed immunohistochemical staining of the clinical tissue and related data analysis. FK, SK, and FC were reviewed the manuscript and were responsible for data corrections. All authors contributed to the article and approved the submitted version.

Funding

The study was funded by Clinical Research Fund of Shandong Medical Association Qilu Special Project (YXH2022ZX02016; YT), and Jinan Clinical Medicine Technology Innovation Program (YT).

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/ fimmu.2022.1028440/full#supplementary-material

SUPPLEMENTARY FIGURE 1

Screening of modules corresponding to macrophages using WGCNA (Weighted Gene Co-Expression Network Analysis). (A) The cluster dendrogram of modular genes associated with macrophage infiltration.

(B) Heatmap of module-trait relationships associated with macrophage infiltration. (C) Module genes relevant heatmap related to macrophage infiltration. (D) Scatter plot of correlation between modules and gene features.

SUPPLEMENTARY FIGURE 2

Sample clustering, molecular typing, and prognosis evaluation analyses based on TAMs-related-genes. (A) Subtypes of clustering profiles obtained through analysis of the expression profiles of TAMs-relatedgenes in TCGA-LUAD samples (consensus matrix k=3). (B) Subtypes of consistent clustering profiles obtained through analysis of the expression profiles of TAMs-related-genes in GSE13213 samples (consensus matrix k=3). (C) Subtypes of the consistent clustering profiles obtained from GSE31210 samples (consensus matrix k=2). (D) Clustering profiles from GSE72094 samples (consensus matrix k=2). (E) GSE8894 samples (consensus matrix k=3). (F) Survival analysis of different cluster subtypes in TCGA-LUAD samples. (G) Results of survival analysis of different cluster subtypes in GSE13213 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (H) Results of survival analysis of different cluster subtypes in GSE31210 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (I) Results of survival analysis of different cluster subtypes in GSE72094 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (J) Results of survival analysis of different cluster subtypes in GSE8894 samples; The horizontal axis represents survival time; the vertical axis represents survival probability; Curves with different colors represent different cluster subtypes. (K) Results of principal component analysis (PCA) on TCGA-LUAD samples; (L) Results of principal component analysis (PCA) on GSE13213 samples; (M) Results of principal component analysis (PCA) on GSE31210 samples; (N) Results of principal component analysis (PCA) on GSE72094 samples; (O) Results of principal component analysis (PCA) on GSE8894 samples.

SUPPLEMENTARY FIGURE 3

Single-cell data. (A) Number of genes expressed in cells; (B) Total counts. (C) Mitochondrial gene expression. (D–I) UMAP (Uniform Manifold Approximation and Projection) dimensionality reduction analysis results of TAM subgroups. (D) C1QA. (E) C1QB. (F) APOE. (G) C1QC. (H) Analysis of TAM clusters. (I) Subtypes derived from re-clustering of TAM cell populations.

SUPPLEMENTARY FIGURE 4

Screening of differentially expressed genes amongst tumor macrophage subsets. (A) Dotplot of Top5 maker genes of each subtype. Abscissa axis: marker genes. Ordinate axis: top5 of TAM subtypes. Colors: mean expression per-group; Dot sizes represent the fraction of cells in each group (%). (B–G) Violin plots of the expression of the top5 marker genes. Abscissa axis: different TAM subtypes. Ordinate axis: gene expression.

SUPPLEMENTARY FIGURE 5

Pseudo-chronological analysis of tumor macrophages for simulation of the dynamic changes of macrophages. (A–C) Differential states according to monocle trajectory analysis, distribution of TAMs in trajectories, and pseudo-sequences of differentiation. (D) Genes influencing differentiation states in the clusters. Left column: different clusters. Right column: names of genes. (E) Pathway enrichment analysis of different TAM subtypes. Color: correlation; Red: positive correlation; Blue: negative correlation. Numerical values: correlation p-value.

SUPPLEMENTARY FIGURE 6

Comparison of risk scores corresponding to the clinical characteristics of the different groups. (A) Age, (B) Gender, (C) Radiation therapy. (E) Pathologic T cells; (E) Tumor stage. Abscissa axis: Different groups. Ordinate axis: risk scores. (F) Nomogram model for risk scores and clinical factors according to the clinical characteristics of prognosis. (G) Calibration curve for 1-year survival. (H). 2-year survival. (I) 3-year survival. (J) 5-year survival. Abscissa axis: predicted probability of survival. Ordinate axis: actual survival.

SUPPLEMENTARY FIGURE 7

Assessment of immunotherapy prognosis according to the risk score. (A) Survival analysis of immunotherapy responses in the training set. Abscissa axis: survival time. Ordinate axis: survival probability. Colors represent different risk groups. (B) Comparative analysis of the proportion of treatment response states between high and low risk groups. (C) Comparative analysis of risk scores for different treatment response states.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin (2022) 72(1):7–33. doi: 10.3322/caac.21708

2. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. Lancet (2021) 398(10299):535–54. doi: 10.1016/S0140-6736(21)00312-3

3. Paz-Ares L, Dvorkin M, Chen Y, Reinmuth N, Hotta K, Trukhin D, et al. Durvalumab plus platinum-etoposide versus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): a randomised, controlled, open-label, phase 3 trial. *Lancet* (2019) 394(10212):1929–39. doi: 10.1016/S0140-6736(19)32222-6

4. Kennedy LB. Salama AKS. A Rev Cancer immunotherapy toxicity. CA Cancer J Clin (2020) 70(2):86–104. doi: 10.3322/caac.21596

5. Reck M, Remon J, Hellmann MD. First-line immunotherapy for non-Small-Cell lung cancer. J Clin Oncol (2022) 40(6):586–97. doi: 10.1200/JCO.21.01497

6. Liu SV, Giaccone G. Lung cancer: First-line immunotherapy in lung cancer - taking the first step. *Nat Rev Clin Oncol* (2016) 13(10):595–6. doi: 10.1038/ nrclinonc.2016.148

7. Rizvi NA, Hellmann MD, Brahmer JR, Juergens RA, Borghaei H, Gettinger S, et al. Nivolumab in combination with platinum-based doublet chemotherapy for first-line treatment of advanced non-Small-Cell lung cancer. *J Clin Oncol* (2016) 34 (25):2969–79. doi: 10.1200/JCO.2016.66.9861

8. The Lancet Respiratory Medicine. Lung cancer immunotherapy biomarkers: refine not reject. *Lancet Respir Med* (2018) 6(6):403. doi: 10.1016/S2213-2600(18) 30180-2

9. Rashdan S, Minna JD, Gerber DE. Diagnosis and management of pulmonary toxicity associated with cancer immunotherapy. *Lancet Respir Med* (2018) 6 (6):472-8. doi: 10.1016/S2213-2600(18)30172-3

10. Duruisseaux M, Martínez-Cardús A, Calleja-Cervantes ME, Moran S, Castro de Moura M, Davalos V, et al. Epigenetic prediction of response to anti-PD-1 treatment in non-small-cell lung cancer: a multicentre, retrospective analysis. *Lancet Respir Med* (2018) 6(10):771–81. doi: 10.1016/S2213-2600(18) 30284-4

11. Petersen I. Predictive pathology of lung cancer immunotherapy response. Lancet Respir Med (2018) 6(10):731-3. doi: 10.1016/S2213-2600(18)30333-3

12. Camidge DR, Schenk EL. Blood-based biomarkers for predicting immunotherapy benefit in lung cancer. *Cell* (2020) 183(2):303-4. doi: 10.1016/j.cell.2020.09.052

13. Hendriks L, Besse B. New windows open for immunotherapy in lung cancer. *Nature* (2018) 558(7710):376–7. doi: 10.1038/d41586-018-05312-9

14. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med (2018) 378(22):2093-104. doi: 10.1056/ NEJMoa1801946

15. Alessi JV, Awad MM. Immunotherapy in lung cancer: effective for patients with poor performance status? *Lancet Respir Med* (2020) 8(9):838–9. doi: 10.1016/S2213-2600(20)30107-7

16. Middleton G, Brock K, Savage J, Mant R, Summers Y, Connibear J, et al. Pembrolizumab in patients with non-small-cell lung cancer of performance status 2 (PePS2): a single arm, phase 2 trial. *Lancet Respir Med* (2020) 8(9):895–904. doi: 10.1016/S2213-2600(20)30033-3

17. Jeremy JW, Lin YC, Yao PL, Yuan A, Chen HY, Shun CT, et al. Tumorassociated macrophages: The double-edged sword in cancer progression. J Clin Oncol (2005) 23(5):953–64. doi: 10.1200/jco.2005.12.172

18. Aalipour A, Chuang HY, Murty S, D'Souza AL, Park SM, Gulati GS, et al. Engineered immune cells as highly sensitive cancer diagnostics. *Nat Biotechnol* (2019) 37(5):531–9. doi: 10.1038/s41587-019-0064-8

19. Gordon SR, Maute RL, Dulken BW, Hutter G, George BM, McCracken MN, et al. PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity. *Nature* (2017) 545(7655):495–9. doi: 10.1038/nature22396

20. Ohri C, Shikotra A, Waller D, Bradding P. Macrophage phenotypes and their expression in non-small cell lung cancer. *J Clin Oncol* (2008) 26(15). doi: 10.1200/jco.2008.26.15_suppl.22132

21. Hsiao YJ, Su KY, Chen HY, Yu SL. Opposite effects of M1 and M2 macrophage subtypes on lung cancer progression. *J Clin Oncol* (2015) 24 (5):14273. doi: 10.1200/jco.2015.33.15_suppl.e19148

22. Lu CS, Shiau AL, Su BH, Hsu TS, Wang CT, Su YC, et al. Oct4 promotes M2 macrophage polarization through upregulation of macrophage colony-stimulating factor in lung cancer. *J Hematol Oncol* (2020) 13(1):62. doi: 10.1186/s13045-020-00887-1

23. Conway EM, Pikor LA, Kung SH, Hamilton MJ, Lam S, Lam WL, et al. Macrophages, inflammation, and lung cancer. *Am J Respir Crit Care Med* (2016) 193(2):116–30. doi: 10.1164/rccm.201508-1545CI

24. Chen J, Sun W, Zhang H, Ma J, Xu P, Yu Y, et al. Macrophages reprogrammed by lung cancer microparticles promote tumor development via release of IL-1 β . Cell Mol Immunol (2020) 17(12):1233–44. doi: 10.1038/s41423-019-0313-2

25. Zhang Y, Zhang L, Li R, Chang DW, Ye Y, Minna JD, et al. Genetic variations in cancer-related significantly mutated genes and lung cancer susceptibility. *Ann Oncol* (2017) 28(7):1625–30. doi: 10.1093/annonc/mdx161

26. Xi Y, Shen Y, Wu D, Zhang J, Lin C, Wang L, et al. CircBCAR3 accelerates esophageal cancer tumorigenesis and metastasis *via* sponging miR-27a-3p. *Mol Cancer*. (2022) 21(1):145. doi: 10.1186/s12943-022-01615-8

27. Yang B, Zhang L, Cao Y, Chen S, Cao J, Wu D, et al. Overexpression of lncRNA IGFBP4-1 reprograms energy metabolism to promote lung cancer progression. *Mol Cancer*. (2017) 16(1):154. doi: 10.1186/s12943-017-0722-8

28. Tian Y, Zhang C, Ma W, Huang A, Tian M, Zhao J, et al. A novel classification method for NSCLC based on the background interaction network and the edge-perturbation matrix. *Aging (Albany NY).* (2022) 14(7):3155–74. doi: 10.18632/aging.204004

29. Tian Y, Liu H, Zhang C, Liu W, Wu T, Yang X, et al. Comprehensive analyses of ferroptosis-related alterations and their prognostic significance in glioblastoma. *Front Mol Biosci* (2022) 9:904098. doi: 10.3389/fmolb.2022.904098

30. Tian Y, Wang J, Wen Q, Gao A, Huang A, Li R, et al. The significance of tumor microenvironment score for breast cancer patients. *BioMed Res Int* (2022) 2022:5673810. doi: 10.1155/2022/5673810

31. Komohara Y, Jinushi M, Takeya M. Clinical significance of macrophage heterogeneity in human malignant tumors. *Cancer Sci* (2014) 105(1):1–8. doi: 10.1111/cas.12314

32. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun* (2020) 11(1):2285. doi: 10.1038/s41467-020-16164-1

33. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-induced evolution of human lung cancer revealed by singlecell RNA sequencing. *Cell* (2020) 182(5):1232–1251.e22. doi: 10.1016/ j.cell.2020.07.017

34. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* (2018) 24(7):978–85. doi: 10.1038/s41591-018-0045-3

35. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* (2018) 18(1):35–45. doi: 10.1038/nri.2017.76

36. Gu X, Chu L, Kang Y. Angiogenic factor-based signature predicts prognosis and immunotherapy response in non-Small-Cell lung cancer. *Front Genet* (2022) 13:894024. doi: 10.3389/fgene.2022.894024 37. Liu J, Wen Y, Liu Z, Liu S, Xu P, Xu Y, et al. VPS33B modulates c-Myc/p53/ miR-192-3p to target CCNB1 suppressing the growth of non-small cell lung cancer. *Mol Ther Nucleic Acids* (2020) 23:324–35. doi: 10.1016/j.omtn.2020.11.010

38. Luo A, Yin Y, Li X, Xu H, Mei Q, Feng D. The clinical significance of FSCN1 in non-small cell lung cancer. *BioMed Pharmacother*. (2015) 73:75–9. doi: 10.1016/j.biopha.2015.05.014

39. Shi J, Chen Y, Wang Z, Guo J, Tong C, Tong J, et al. Comprehensive bioinformatics analysis to identify the gene HMMR associated with lung adenocarcinoma prognosis and its mechanism of action in multiple cancers. *Front Oncol* (2021) 11:712795. doi: 10.3389/fonc.2021.712795

40. Jiang X, Tang L, Yuan Y, Wang J, Zhang D, Qian K, et al. NcRNA-mediated high expression of HMMR as a prognostic biomarker correlated with cell proliferation and cell migration in lung adenocarcinoma. *Front Oncol* (2022) 12:846536. doi: 10.3389/fonc.2022.846536

41. Stevens LE, Cheung WKC, Adua SJ, Arnal-Estapé A, Zhao M, Liu Z, et al. Extracellular matrix receptor expression in subtypes of lung adenocarcinoma potentiates outgrowth of micrometastases. *Cancer Res* (2017) 77(8):1905–17. doi: 10.1158/0008-5472.CAN-16-1978

42. Liao WC, Lin TJ, Liu YC, Wei YS, Chen GY, Feng HP, et al. Nuclear accumulation of KPNA2 impacts radioresistance through positive regulation of the PLSCR1-STAT1 loop in lung adenocarcinoma. *Cancer Sci* (2022) 113(1):205–20. doi: 10.1111/cas.15197

43. Li XL, Jia LL, Shi MM, Li X, Li ZH, Li HF, et al. Downregulation of KPNA2 in non-small-cell lung cancer is associated with Oct4 expression. *J Transl Med* (2013) 11:232. doi: 10.1186/1479-5876-11-232

44. Bu H, Li Y, Jin C, Yu H, Wang X, Chen J, et al. Overexpression of PRC1 indicates a poor prognosis in ovarian cancer. *Int J Oncol* (2020) 56(3):685–96. doi: 10.3892/ijo.2020.4959

45. Jiang X, Li Y, Zhang N, Gao Y, Han L, Li S, et al. RRM2 silencing suppresses malignant phenotype and enhances radiosensitivity *via* activating cGAS/STING signaling pathway in lung adenocarcinoma. *Cell Biosci* (2021) 11(1):74. doi: 10.1186/s13578-021-00586-5

46. Tang B, Xu W, Wang Y, Zhu J, Wang H, Tu J, et al. Identification of critical ferroptosis regulators in lung adenocarcinoma that RRM2 facilitates tumor immune infiltration by inhibiting ferroptotic death. *Clin Immunol* (2021) 232:108872. doi: 10.1016/j.clim.2021.108872

47. Cheng WC, Chang CY, Lo CC, Hsieh CY, Kuo TT, Tseng GC, et al. Identification of theranostic factors for patients developing metastasis after surgery for early-stage lung adenocarcinoma. *Theranostics* (2021) 11(8):3661–75. doi: 10.7150/thno.53176

48. Jiang ZF, Wang M, Xu JL. Thymidine kinase 1 combined with CEA, CYFRA21-1 and NSE improved its diagnostic value for lung cancer. *Life Sci* (2018) 194:1–6. doi: 10.1016/j.lfs.2017.12.020

49. Wang Z, Ren Z, Li R, Ge J, Zhang G, Xin Y, et al. Multi-omics integrative bioinformatics analyses reveal long non-coding RNA modulates genomic integrity *via* competing endogenous RNA mechanism and serves as novel biomarkers for overall survival in lung adenocarcinoma. *Front Cell Dev Biol* (2021) 9:691540. doi: 10.3389/fcell.2021.691540

50. Fan J, Zhang X, Wang S, Chen W, Li Y, Zeng X, et al. Regulating autophagy facilitated therapeutic efficacy of the sonic hedgehog pathway inhibition on lung adenocarcinoma through GLI2 suppression and ROS production. *Cell Death Dis* (2019) 10(9):626. doi: 10.1038/s41419-019-1840-6

51. Seidl C, Panzitt K, Bertsch A, Brcic L, Schein S, Mack M, et al. MicroRNA-182-5p regulates hedgehog signaling pathway and chemosensitivity of cisplatinresistant lung adenocarcinoma cells *via* targeting GLI2. *Cancer Lett* (2020) 469:266–76. doi: 10.1016/j.canlet.2019.10.044

52. Chen S, Duan Y, Wu Y, Yang D, An J. A novel integrated metabolismimmunity gene expression model predicts the prognosis of lung adenocarcinoma patients. *Front Pharmacol* (2021) 12:728368. doi: 10.3389/fphar.2021.728368

53. Frazzi R. BIRC3 and BIRC5: multi-faceted inhibitors in cancer. Cell Biosci (2021) 11(1):8. doi: 10.1186/s13578-020-00521-0

54. Chen H, Chen X, Pan B, Zheng C, Hong L, Han W. KRT8 serves as a novel biomarker for LUAD and promotes metastasis and EMT *via* NF-κB signaling. *Front Oncol* (2022) 12:875146. doi: 10.3389/fonc.2022.875146

55. Marie MA, Sanderlin EJ, Satturwar S, Hong H, Lertpiriyapong K, Donthi D, et al. GPR65 (TDAG8) inhibits intestinal inflammation and colitis-associated colorectal cancer development in experimental mouse models. *Biochim Biophys Acta Mol Basis Dis* (2022) 1868(1):166288. doi: 10.1016/j.bbadis.2021.166288

56. Ohri CM, Shikotra A, Green RH, Waller DA, Bradding P. Macrophages within NSCLC tumour islets are predominantly of a cytotoxic M1 phenotype associated with extended survival. *Eur Respir J* (2009) 33(1):118–26. doi: 10.1183/09031936.00065708

57. Cassetta L, Pollard JW. Targeting macrophages: therapeutic approaches in cancer. Nat Rev Drug Discovery (2018) 17(12):887–904. doi: 10.1038/nrd.2018.169