



## Research article

# The landscape of immunogenic cell death-related genes predicts the overall survival and immune infiltration status of non-small-cell lung carcinoma

Jian Zhang<sup>a,1</sup>, Huiying Li<sup>b,1</sup>, Xi Zhang<sup>b</sup>, Yue Yang<sup>c,\*\*</sup>, Yue Sun<sup>d,\*</sup>

<sup>a</sup> Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin, 150040, Heilongjiang, China

<sup>b</sup> Department of Pathology, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin, 150040, Heilongjiang, China

<sup>c</sup> Institute of cancer prevention and treatment, Harbin Medical University, 6 Baojian Road, Harbin, 150000, Heilongjiang, China

<sup>d</sup> Science and Technology Academic Department of Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin, 150040, Heilongjiang, China

## ARTICLE INFO

## Keywords:

NSCLC

Immunogenic cell death(ICD)

Risk model

Prognosis

Immune infiltration

## ABSTRACT

**Background:** Non-small cell lung cancer (NSCLC), which accounts for about 85 % of all lung cancers, currently exhibits insensitivity to most treatment regimens. Therefore, the identification of new and effective biomarkers for NSCLC is crucial for the development of treatment strategies. Immunogenic cell death (ICD), a form of regulated cell death capable of activating adaptive immune responses and generating long-term immune memory, holds promise for enhancing anti-tumor immunity and offering promising prospects for immunotherapy strategies in NSCLC.

**Methods:** Clinical information and expressive profiles of NSCLC genes were retrieved from the GEO and TCGA databases. By combining these databases, the researchers were able to identify the appropriate genes for use in forecasting outcomes of patients with this type of cancer. We further performed functional enrichment, gene variants and immune privilege correlation analysis to determine the underlying mechanisms. This was followed by univariate and multivariate Cox regression and LASSO regression analyses, we developed a prognostic risk model based on the TCGA cohort, which included 17 gene labels. The results of the external validation were then used to identify the appropriate genes for use in predicting the survival outcome of patients with this type of cancer. In addition, a nomogram was created to help visualise the clinical presentation of the patients. For the analyses, we performed 50 functional and immunoinfiltration assessments for two risk groups.

**Results:** Using 17 genes (AIRE, APOH, CDKN2A, CEACAM4, COL4A3, CPA, DBH, F10, FCGRB, FGFR4, MMP1, PGLYRP1, SCGB2A2, SLC9A3, UGT2B17 and VIP), The researchers then created a gene signature that could be used to identify patients with an increased risk of contracting cancer. They divided the patients into two groups based on their risk score. The low-risk group exhibited a better prognosis ( $P < 0.01$ ). The survival curve demonstrated that ICD-related models could accurately predict patient prognosis. Conversely, high-risk subgroups were closely associated with immune-related signaling pathways. The analysis of immune infiltration also showed that

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [Jian\\_cheung@hrbmu.edu.cn](mailto:Jian_cheung@hrbmu.edu.cn) (J. Zhang), [lihuiying0606@163.com](mailto:lihuiying0606@163.com) (H. Li), [zxkorea@126.com](mailto:zxkorea@126.com) (X. Zhang), [yangyue1010@sohu.com](mailto:yangyue1010@sohu.com) (Y. Yang), [moonsun95318@qq.com](mailto:moonsun95318@qq.com) (Y. Sun).

<sup>1</sup> Contributed equally to this work.

the infiltration levels of most immune cells were higher in the high risk sub-group than in the low risk sub-group. In comparison to the low-risk group, the high-risk group was more susceptible to the immune-checkpoint blockade (ICB) treatment.

**Conclusion:** Our researchers utilized a gene model to analyze the immune inflammation and prognosis of patients with non-small-cell lung cancer (NSCLC). The discovery of new ICD-related genes could lead to the development of new targeted treatments for this condition.

---

## 1. Introduction

Lung cancer is the most common cause of death from cancer in both men and women, Non-small cell lung cancer (NSCLC) accounts for approximately 85 % of all cases of lung cancer [1]. Typically, when diagnosed, the tumour has progressed to a locally advanced or metastatic stage, resulting in a poor prognosis. NSCLC includes adenocarcinoma, squamous cell carcinoma and large cell lung cancer [2]. These characteristics render NSCLC more heterogeneous compared to other types of lung cancer patients, making it challenging to identify those who are responsive to treatments. Consequently, there is an urgent need for the discovery of novel biomarkers that can be used for the prediction and improvement of the prognosis of NSCLC patients.

Malignant cells produce damage-associated patterns of molecules (DAMPs) and secrete cytokines in response to microenvironmental perturbations, which mediate chemotactic and immunostimulatory effects. Immunogenic cell death (ICD) is a unique response mode characterized by the induction of organelles and stress, ultimately leading to cell death accompanied by the exposure and release of a large number of DAMPs [3–5]. Recent evidence suggests that the mechanism of ICD varies across different cancers and is associated with patient prognosis. ICD is a ubiquitous phenomenon observed in cancer-related injuries [6], infections [7], and autoimmune diseases [8], and its occurrence is often related to patient prognosis.

For the induction of immunogenic cell death, the interaction between cancer cells and immune cells is of paramount importance (ICD). Cancer cells can induce immune cells to produce a large number of proinflammatory factors, which can promote the occurrence of ICD by activating signal transduction molecules such as Toll-like receptors (TLRs) and T cell receptors (TCRs) [9]. When immune cells come into contact with cancer cells, they may change their surface properties, shape, or interaction with cancer cells. For example, a study found that up to 95 % of healthy adults had immune cells associated with both tumor stroma and angiogenesis. However, this correlation does not always exist between cancer cells and immune cells. In some studies, tumor-associated macrophages (TAMs) [10] mediated ICD by releasing proinflammatory factors such as IL-6, TNF- $\alpha$ , and MCP-1 [11] into tissues. Furthermore, TAM play a key role in promoting ICD by increasing these pro inflammatory factors in breast cancer [4,12]. These findings provide important insights into the mechanisms of ICD and may contribute to the development of new therapeutic approaches to modulate immune responses for cancer treatment and prevention. However, the role of ICD in predicting outcomes of NSCLC patients or guiding clinical treatment is still unclear.

Therefore, this investigation focused on obtaining gene expression profiles and clinical data of NSCLC from The Cancer Genome Atlas (TCGA) [13] and Gene Expression Omnibus (GEO) [14] databases. The genes related to ICD were gathered from the Genecards database. The genes associated with overall survival (OS) prognosis in NSCLC were filtered in the TCGA cohort, and their intersection with the ICD-related gene set was determined to identify ICD-related prognostic genes. To uncover the underlying mechanisms, functional enrichment, mutagenesis and immune infiltration analyses were performed. A prognostic risk model based on nine gene signatures was developed using Cox and LASSO regression analysis in the TCGA cohort. This model was then external validated in the GEO cohort. Additionally, a nomogram was created to predict survival by integrating the clinicopathological data and prognostic gene signatures. Furthermore, A survival prediction nomogram has been developed through the integration of clinician-pathological data and prediction gene markers. The high-risk and low-risk subgroups were then analysed for differences in risk function and immunology. Based on above results, we generated an effective prognostic model for NSCLC patients, which emphasized that ICD-related genes can predict prognosis, even guide the usage of immunotherapy in NSCLC patient.

## 2. Materials and methods

### 2.1. Data collection and preprocessing

The count and transcriptome data (Count and TPM format), and clinical data of 1041 NSCLC samples and control tissues from the TCGA-LUSC and TCGA-LUAD projects were downloaded from the TCGA database. Additionally, transcriptome data and clinical data of 715 NSCLC samples, including datasets GSE30219, GSE31210, and GSE37745, were downloaded from the GEO database for validation purposes. Furthermore, 1766 ICD-related genes were retrieved from the GeneCards database (<https://www.genecards.org/>), using “immunogenicity cell death” as the keyword.

### 2.2. Establishment of ICD related models

First, Univariate Cox regression analysis was performed for ICD-related genes, and those with a p-value less than 0.05 were selected as candidate genes for the model. Subsequently, we utilized the glmnet package in R software to perform the selection of features on these candidate genes using the minimum absolute contraction and selection operator (LASSO) regression algorithm, incorporating 10-

fold cross validation. The model's formula is represented as Risk Score = (In the formula,  $n$  equal to the number of genes in the model,  $Exp_i$  is the gene expression,  $C_i$  is the corresponding coefficient). By applying the aforementioned formula, the risk score for each patient was obtained, which will be subjected to further follow-up analysis. The receiver operating characteristic (ROC) curve was used to assess the predictive ability of the model, with a larger area under the curve (AUC) indicating greater model efficiency.

### 2.3. Drawing a nomogram

In this research, to integrate the risk scores derived from the M-stage and ICD-related models, we used the R software package rms. We subsequently created a nomogram employing the Cox method to evaluate the prognostic importance of these characteristics in 993 samples obtained from the TCGA database. To assess the effectiveness of the nomogram, we calibrated the curve. The overall model exhibited a C-index of 0.626130264126878, with a 95 % confidence interval in the range of 0.600551054797142–0.651709473456614.

### 2.4. Immunoinfiltration analysis

Single sample Gene Set Enrichment Analysis (ssGSEA) involves performing GSEA on an individual sample. The sequencing approach for the gene list and the method used to compute the enrichment score (ES) are both dependent on the gene expression levels found in that specific sample. It is crucial to understand that ssGSEA does not rely on the relationship between genes and phenotypes [15]. In this study, immune cells in the tumour microenvironment of NSCLC were scored using the ssGSEA algorithm to assess immune status.

### 2.5. Drug sensitivity analysis

The Genomics of Drug Sensitivity in Cancer (GDSC), developed by the Sanger Institute in the United Kingdom, collects data on the responses of tumor cells to various treatments and their levels of sensitivity. Alterations in the cancer genome can significantly affect the effectiveness of clinical therapies, and different targets exhibit notably varied reactions to medications [16]. Therefore, this type of data is crucial for identifying potential tumor therapeutic targets. The expression of immune checkpoints can indirectly reflect a patient's sensitivity to immunotherapy. In this study, we evaluated the sensitivity to immune checkpoint blockers (ICBs) in NSCLC patients from the TCGA cohort using the Tumour Immune Dysfunction and Exclusion (TIDE) algorithm. Generally, individuals exhibiting elevated TIDE scores demonstrate reduced responsiveness to immunotherapy (30127393). Additionally, we used the GSCALite GDSC module of the database (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>) to explore the relationship between gene models and the sensitivity to chemotherapy drugs (29790900).

### 2.6. Functional enrichment analysis

Gene Set Enrichment Analysis (GSEA) was performed utilizing the R package clusterProfiler (version 3.14.3) to derive results for gene set enrichment. Parameters were set with a minimum gene set size of 5 and a maximum of 500. Statistical significance was judged by p-values of less than 0.05 and FDRs of less than 0.10.

### 2.7. Difference analysis

Limma (linear models for microarray data, <https://doi.org/10.1093/nar/gkv007>) is a method for differential expression analysis that employs a generalized linear model. In this method, the expression of each gene was fitted into a linear equation. The outcomes of the data analysis can be readily accessed and visualized via web services, significantly enhancing the use of data and research on tumors. In this study, Limma and R software (version version 3.40.6) were employed for differential expression analysis to identify the differential genes between different comparison groups and control groups. The threshold for significance was defined as the difference factor greater than 0.5 and the p-value less than 0.05.

### 2.8. Protein-protein interaction network

STRING ( ) is "https://cn.string-db.org/)is">https://cn.string-db.org/) is a protein interaction database that can be used to analyze both known and predicted interactions between proteins. Protein-protein interaction (PPI) refers to a process in which two or more protein molecules form a protein complex through non-covalent bonds [17]. Protein interactions are a key element of the biochemical response network within cells and are essential for regulating cellular functions and their signaling pathways. In this study, we applied the STRING database for PPI analysis and set the interaction score to 1.5. Subsequently, the interaction network was exported to Cytoscape software for visualization. Using the Cytohub function in Cytoscape, We examined and pinpointed the five most significant hub genes in the network. These hub genes are expected to have significant roles within the protein interaction network and may participate in essential cellular processes or signaling pathways.

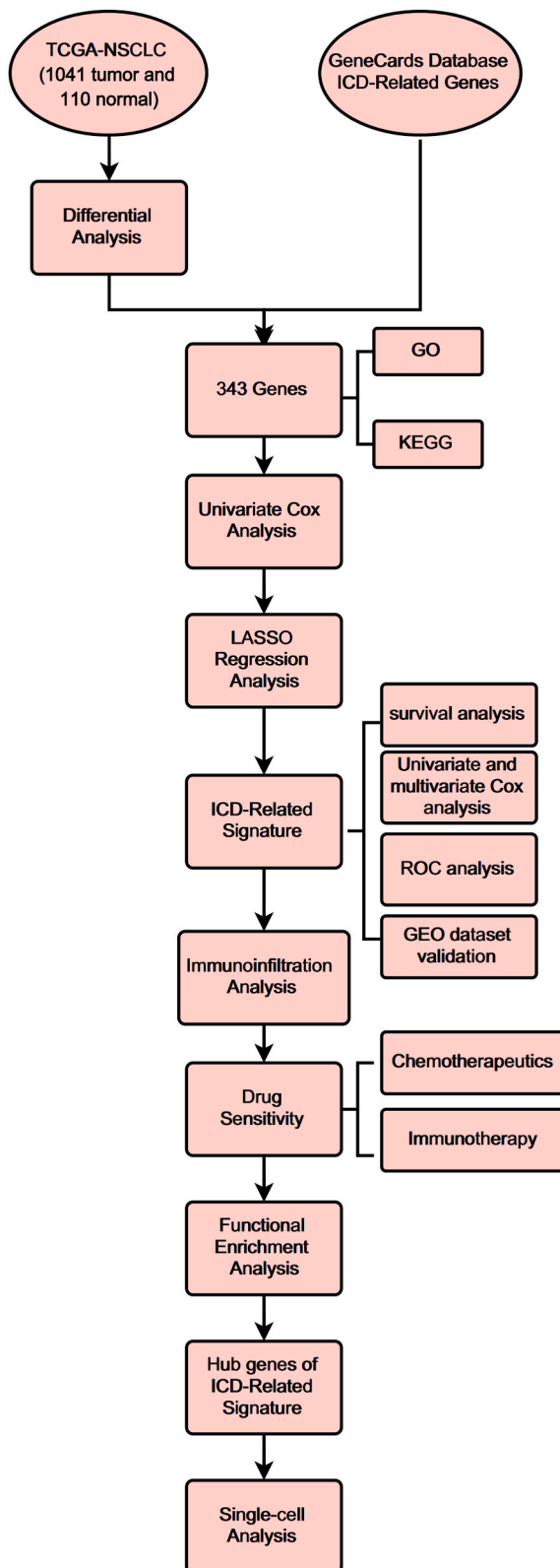
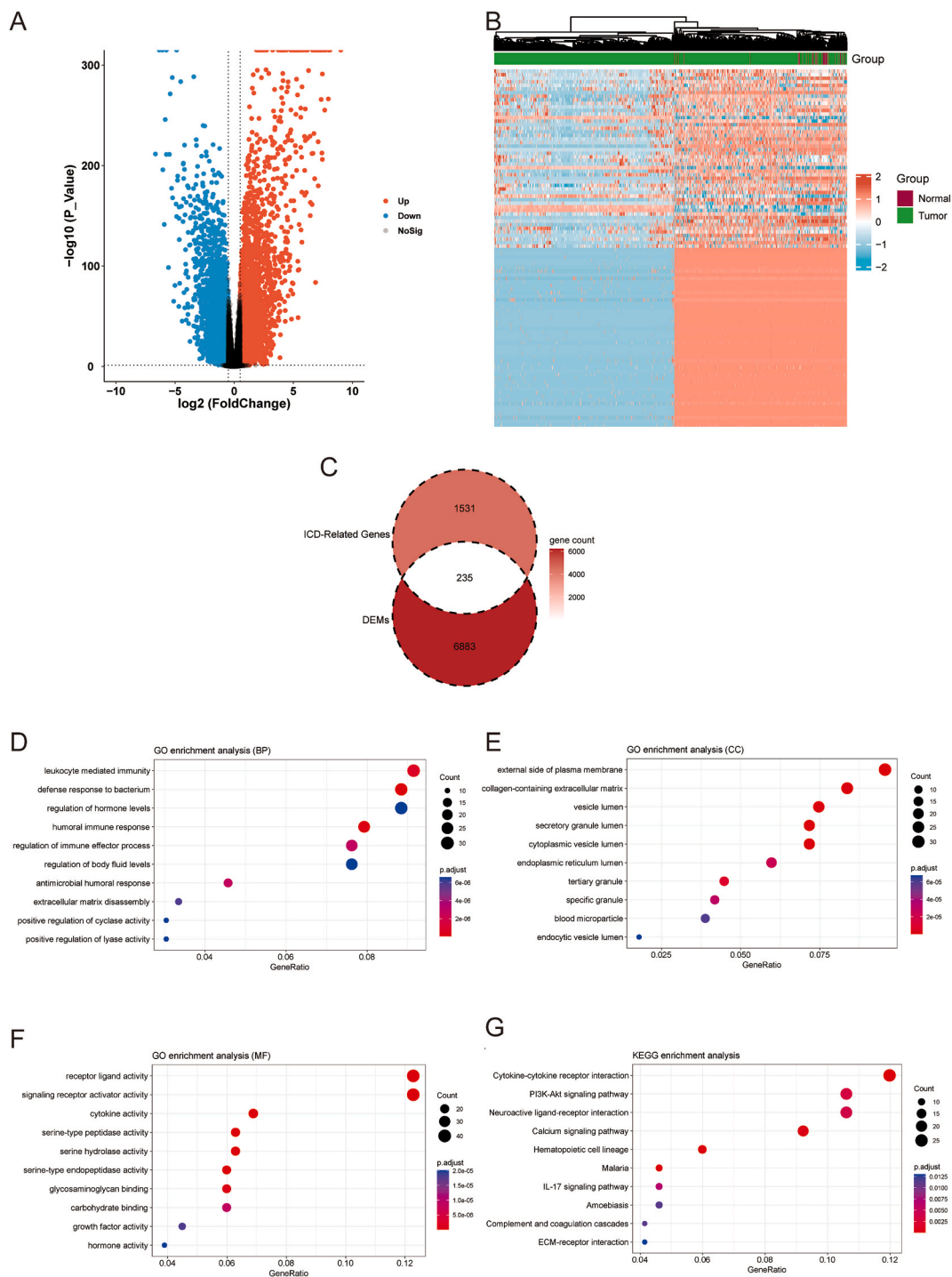


Fig. 1. Figure of the methods of this study.



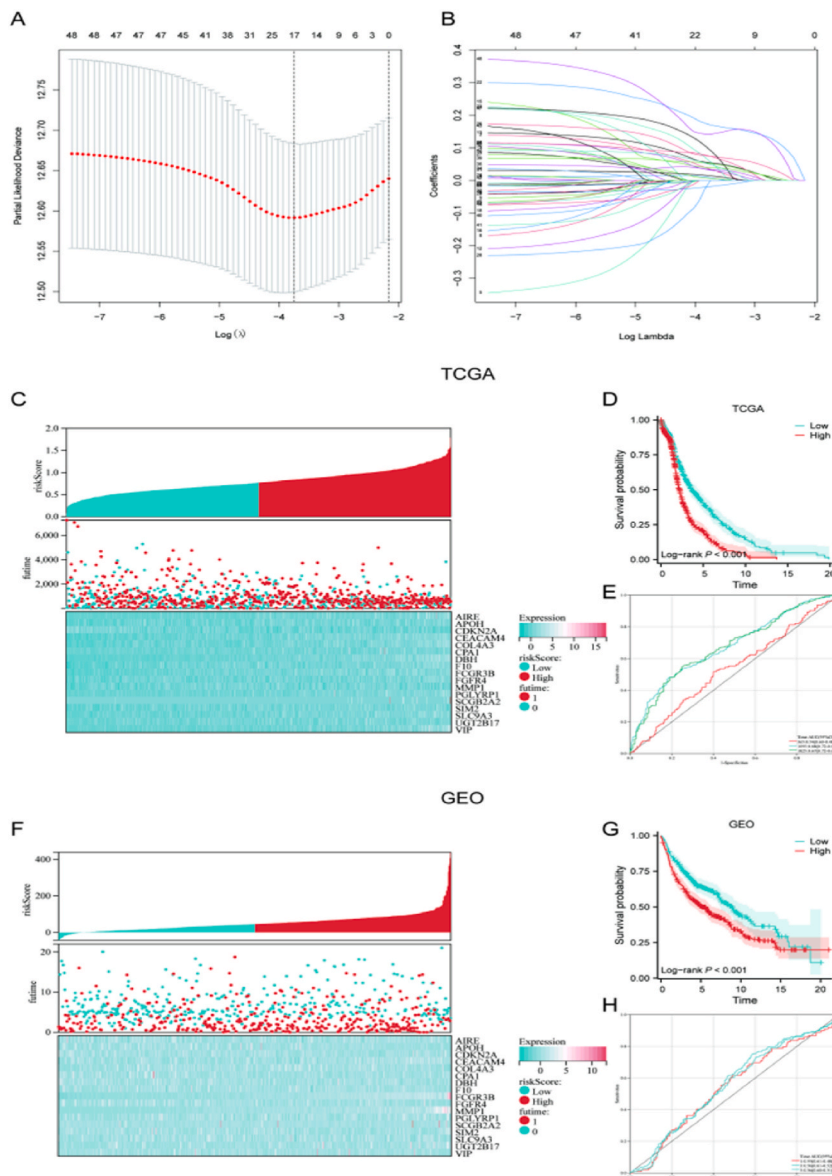
**Fig. 2.** Differential Expression of ICD Genes and Functional Enrichment Analysis in Non-Small Cell Lung Cancer (NSCLC). A. Volcano plot illustrating the differentially expressed genes in Small Cell Lung Cancer (SCLC) compared to adjacent non-cancerous tissues. B. Heatmap displaying the differentially expressed genes in Small Cell Lung Cancer (SCLC) versus adjacent non-cancerous tissues. C. Venn diagram illustrating the intersection of differentially expressed genes and ICD genes between small cell lung cancer (SCLC) and adjacent non-cancerous tissue. D. Investigation of differentially expressed ICD genes in non-small cell lung cancer (NSCLC) using Gene Ontology Biological Process (GO-BP) analysis. E. GO-MF analysis of differentially expressed ICD genes in NSCLC. F. GO-CC analysis of differentially expressed ICD genes in NSCLC. G. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of differentially expressed ICD genes in non-small cell lung cancer (NSCLC).

2.9. Single cell analysis

TIGER is an accessible web portal that integrates and analyzes gene expression data related to tumor immunology. It encompasses a substantial volume of transcriptome data, comprising bulk transcriptome information from 1508 tumor samples with associated immunotherapy clinical outcomes and 11,057 TCGA tumor/normal samples. Additionally, TIGER includes single-cell transcriptome data from 655 samples, comprising a total of 2116,945 cells. Importantly, 119,039 cells derived from 63 samples are associated with immunotherapy clinical data. In this study, we utilized NSCLC data available in the TIGER database(32042191) for single-cell analysis to confirm which cells should express the target gene.

2.10. Statistical analysis

The choice between the *t*-test and the Mann-Whitney *U* test for the comparison of the two data groups was based on the normality of



**Fig. 3.** Establishing an ICD-associated prognosis model for NSCLC. A-B. LASSO regression analysis of ICD genes associated with prognosis. C. Survivorship, hazard scores and expression levels of 17 model genes across TCGA cohorts. D. In the TCGA dataset, Kaplan-Meier curves for high-risk and low-risk groups. E. ROC curve of ICD-related prognostic model in TCGA dataset. F. Risk scores of 17 model genes in the GEO dataset, survival status and expression levels. G. Kaplan-Meier curves within the GEO dataset to illustrate the high-risk and low-risk groups. H. ROC curve analysis of ICD prediction models in GEO.

the data distribution. In the case of categorical variables, either the chi-square test or Fisher's exact test were used. For comparisons between more than two groups, the Kruskal-Wallis test was used. To assess relationships between variables, Spearman correlation analysis was used. The Kaplan-Meier method and the log-rank test were used for the analysis of survival time. To assess the independent effects of the variables, univariate and multivariate Cox regression analyses were performed. All statistical analyses were executed using R software. Statistical significance was defined as  $P < 0.05$ , with the following notation used to indicate the level of significance: \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.001$ , ns:  $P > 0.05$ .

### 3. Results

#### 3.1. Analysis of differential expression and functional enrichment of genes associated with immunogenic cell death (ICD) in non-small cell lung cancer (NSCLC)

Prior to presenting the results, we include an overall flowchart of this study in Fig. 1 to facilitate readers' understanding of our methodology. In comparison to adjacent non-cancerous tissue, we identified that 3312 genes were upregulated in NSCLC, while 3806 genes were downregulated (Fig. 2A and B). After intersecting the differential expressed genes and ICD genes, we obtained 235 genes for further analysis (Fig. 2C). GO-BP revealed that these 235 ICD genes are involved in immune response (Fig. 2D). GO-CC showed that these molecules are localized to the extracellular matrix and cytoplasm (Fig. 2E). GO-MF showed that these molecules bind to receptor ligands and are associated with cytokine activity (Fig. 2F). KEGG enrichment analysis further demonstrated that these molecules interact with cytokines and signaling pathways such as PI3K-AKT (Fig. 2G).

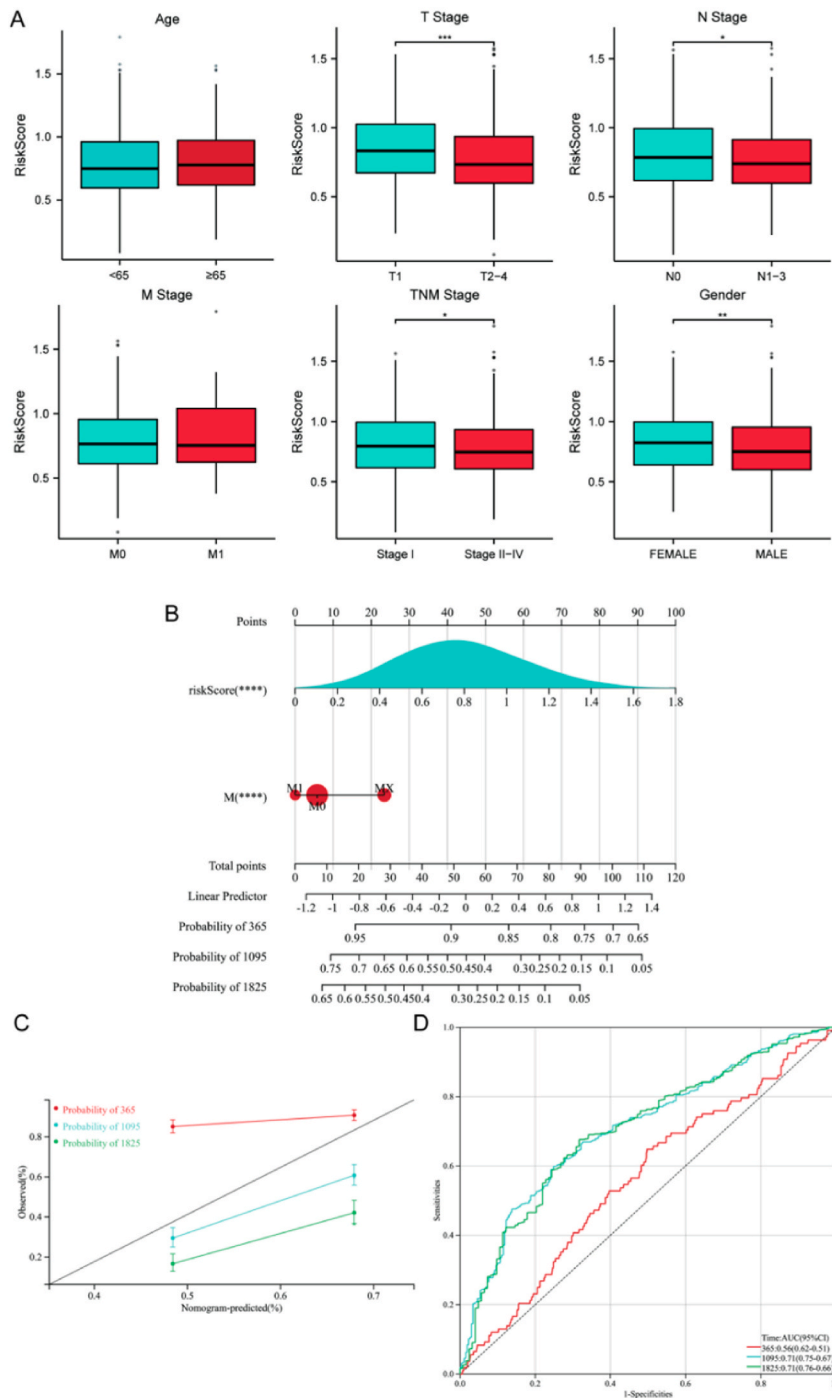
#### 3.2. Development of a prognostic model associated with immunogenic cell death (ICD)

We performed univariate Cox regression analysis of 235 ICD-associated genes using NSCLC cases from the TCGA database, and the results showed that 52 of these genes were associated with NSCLC prognosis ( $P < 0.05$ ). Subsequently, we constructed an ICD-related prognostic model based on the 52 genes (Fig. 3A and B), with the formula for this model outlined in the methodology section. For 993 NSCLC patients with available prognostic information in the TCGA database, we derived risk scores using this formula. A median risk

**Table 1**

	Levels	LUAD	LUSC
n		539	502
ajcc_stage (%)	Stage I	5 (0.9)	3 (0.6)
	Stage IA	137 (25.8)	90 (18.1)
	Stage IB	153 (28.8)	152 (30.5)
	Stage II	1 (0.2)	3 (0.6)
	Stage IIA	50 (9.4)	65 (13.1)
	Stage IIB	75 (14.1)	94 (18.9)
	Stage III	0 (0.0)	3 (0.6)
	Stage IIIA	73 (13.7)	63 (12.7)
	Stage IIIB	11 (2.1)	18 (3.6)
	Stage IV	26 (4.9)	7 (1.4)
age (median [IQR])		66.00 [60.00, 73.00]	69.00 [62.00, 74.00]
T_stage (%)	T1	71 (13.2)	50 (10.0)
	T1a	47 (8.7)	24 (4.8)
	T1b	57 (10.6)	40 (8.0)
	T2	180 (33.4)	173 (34.5)
	T2a	86 (16.0)	87 (17.3)
	T2b	27 (5.0)	34 (6.8)
	T3	49 (9.1)	71 (14.1)
	T4	19 (3.5)	23 (4.6)
	TX	3 (0.6)	0 (0.0)
	N_stage (%)	N0	349 (64.9)
N1		98 (18.2)	131 (26.1)
N2		74 (13.8)	40 (8.0)
N3		2 (0.4)	5 (1.0)
NX		15 (2.8)	6 (1.2)
M_stage (%)	M0	366 (68.7)	412 (82.7)
	M1	18 (3.4)	5 (1.0)
	M1a	2 (0.4)	1 (0.2)
	M1b	5 (0.9)	1 (0.2)
	MX	142 (26.6)	79 (15.9)
gender (%)	female	291 (54.0)	131 (26.1)
	male	248 (46.0)	371 (73.9)
smoking_status (%)	≤ 15 years	77 (32.5)	87 (50.3)
	>15 years	79 (33.3)	51 (29.5)
	Current smoker	47 (19.8)	28 (16.2)
	Lifelong Non-smoker	34 (14.3)	7 (4.0)

score was then used to divide patients into high-risk and low-risk groups. Fig. 3C shows the risk score, survival state, and expression levels of the 17 ICD-associated genes comprising the model in NSCLC patients from the TCGA cohort. Overall survival (OS) in the low-risk group was superior to the high-risk group, according to survival curve analysis ( $P < 0.001$ , Fig. 3D). The ROC curve analysis showed that the AUC value of ICD-related models was 0.54 for 1-year survival, 0.68 for 3-year survival, and 0.67 for 5-year survival in NSCLC patients (Fig. 3E). To provide further validation of the predictive performance of the ICD model, we applied it to a cohort of 715 NSCLC patients from the GEO database. Fig. 3F presents the risk scores, survival status, as well as the expression levels of the same 17



**Fig. 4.** Correlations and nomograms of ICD-Related Prognostic Models with Clinical Traits. A. Correlation of ICD-related prognostic models with clinical traits. B. Construct an ICD-related prognostic model and a nomogram for M staging. C. Nomogram consistency curve. D. Analysis of the ROC curve of the nomogram.



ICD-related genes in the GEO NSCLC patients. Consistent with our expectations, individuals in the low-risk group demonstrated better outcomes than those classified as high-risk ( $P < 0.01$ , Fig. 3G). The ROC curve analysis revealed AUC values of 0.55 for one-year survival, and both 0.56 for three- and five-year survivals among NSCLC patients (Fig. 3H). These findings suggest that the ICD-related scheme holds promise as a prognostic tool for predicting outcomes in NSCLC patients.

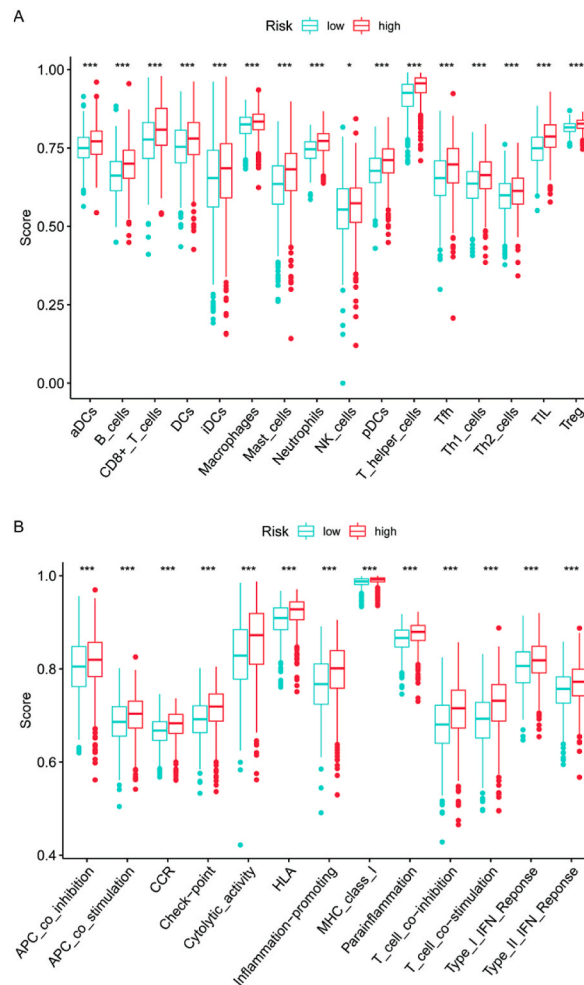
### 3.3. ICD-based models may be independently predictive of prognosis in NSCLC patients and have correlation with clinical features

Every day, approximately 200–300 billion cells undergo death and apoptosis in the human body. This process primarily occurs through Caspase-dependent apoptosis under homeostasis conditions. Despite the large quantities of apoptotic cells produced, they are difficult to observe in vivo tissues due to the efficient engulfment and removal of these cells by phagocytes through a process called efferocytosis [18]. Exocytosis, such as ICD, is essential for maintaining normal homeostasis.

To assess the impact of the ICD model as prognostic indicator for NSCLC patients, In addition to the ICD model, clinical factors including TNM stage, T stage, N stage, M stage and patient sex from the TCGA database were included. The findings indicated that both the ICD model and M stage functioned as independent prognostic markers for NSCLC patients (Table 1). Moreover, we noted that risk scores were elevated in patients with advanced T, N, and TNM stages, and male patients exhibited higher risk scores compared to their female counterparts (Fig. 4A).

### 3.4. Construction of line diagram

Since both the M staging and ICD models were identified as independent prognostic markers, we incorporated the risk scores and M stage into a nomogram to evaluate its potential utility (Fig. 4B). Good agreement between observed and predicted values was observed



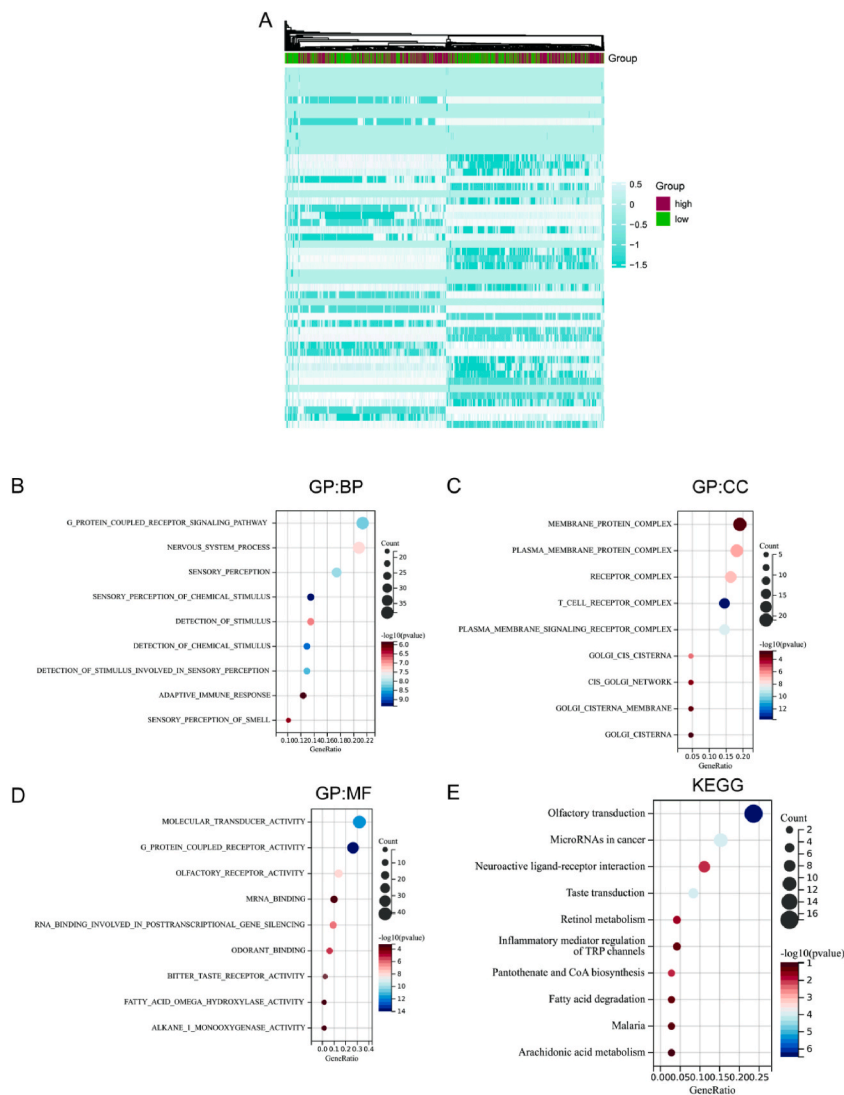
**Fig. 5.** ICD-associated Prognostic Models in relation to immune cell burden and immune status scores. A. The scores for immune cells infiltration were between the high and low risk groups. B. Differences in the immune system scores between the high and low risk groups.



and response to immune checkpoint blockade (ICB) therapy using the Tumour Immune Dysfunction and Exclusion (TIDE) algorithm. Importantly, higher TIDE score correlates with reduced sensitivity to immune checkpoint blockade (ICB) therapy. Our analysis showed that the high-risk group had a higher TIDE score compared to the low-risk group, with the following results (Fig. 6B). Furthermore, when utilizing TIDE scores to help predict how patients respond to immunotherapy, we observed that a greater proportion of patients in the high-risk cohort responded favorably to treatment than those in the low-risk cohort (Fig. 6C). This observation implies that the low-risk cohort demonstrates enhanced sensitivity to immune-checkpoint blockade (ICB) therapy in contrast to the high-risk cohort. In addition, we looked at the relationship between the levels of expression of the 17 genes that were included in our model and their sensitivity to chemotherapy drugs, which further suggested that these genes might have a crucial role in influencing responses to chemotherapy (Fig. 6D).

### 3.7. Potential biological functions of ICD-related models

We looked at the differences in the low risk and high risk groups and found 595 mRNAs were upregulated in the high risk cohort in comparison with the low-risk cohort (Fig. 7A). These variations in mRNA expression were predominantly associated with G-protein-coupled receptor signalling pathways and other biological processes, as revealed by functional enrichment analysis. Their localization

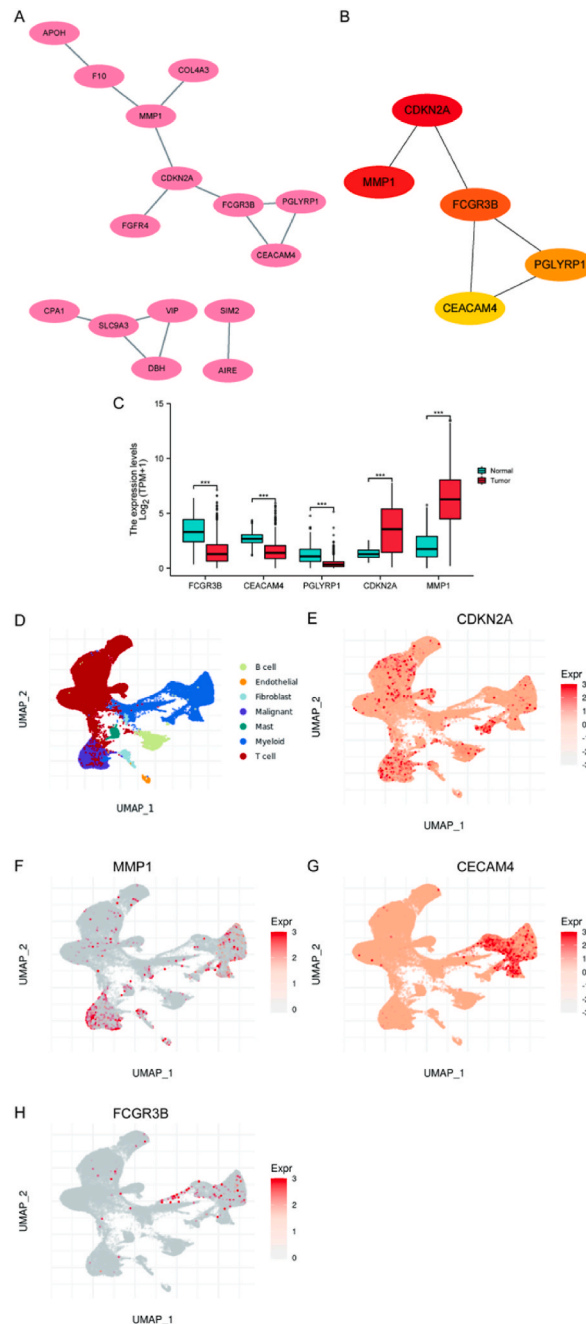


**Fig. 7.** Analysis of functional enrichment in prognostic models in the context of the ICD. A. A heat map showing genes differentially expressed between the high and low risk groups. B. GO-BP analysis for the genes that show a difference in expression when the high risk cohort is compared with the low risk cohort. E. GO-CC analysis focusing on the differential expression of genes in both high and low risk categories. F. GO-MF analysis regarding the differentially expressed genes found in both risk classifications. G. KEGG analysis highlighting the expression differences between the high-risk and low-risk groups.

was associated with the Golgi apparatus and plasma membrane, as well as G protein-coupled receptor activity (Fig. 7B–D). Additionally, KEGG analysis indicated that these genes are linked to the involvement of mRNAs in cancer (Fig. 7E).

### 3.8. Hub gene screening and single cell analysis in ICD-related models

Fig. 8A shows the PPI networks of the 17 model genes, and 8B displays the top 5 core genes. Among these top genes, FGFR3B, CEACAM4, and PGLYRP1 were down-regulated in NSCLC, while CDKN2A and MMP1 were up-regulated (Fig. 8C). Then we employed the TIGER database (see Method) to identify the specific cell types expressing these model genes. Fig. 8D shows the cluster analysis of



**Fig. 8.** Hub Gene Screening and single-cell Analysis in an ICD-Related Prognostic Model. A. PPI network of 17 genes in an ICD-related prognostic model. B. TOP5 Hub genes in the PI network. C. Differential TOP 5 hub gene expression between cancer and healthy tissue in the TCGA cohorts. D. Cluster analysis and cell type identification of SCLC single-cell dataset. Expression of E-h ub in NSCLC cell subtypes in the single cell dataset.

cell subtypes in the single-cell dataset, including B cells, endothelial cells, fibroblasts, T cells, malignant cells, mast cells, and myeloid cells. CDKN2A and MMP1 were highly expressed in malignant cells (Fig. 8E and F), while FGFR3B and CEACAM4 were highly expressed in myeloid cells (Fig. 8G and H). These results highlight that the model genes are expressed in different cell type in tumor microenvironment, and suggest that cells expressing these genes may be the potential targets for improving outcomes of NSCLC patients.

#### 4. Discussion

Lung cancer is recognized as the leading cause of cancer-related deaths worldwide. While new technologies and therapies have resulted in increased survival rates for many patients, some still succumb to relapsed metastases [19]. Over the course of the last two decades, there have been significant advances in the management of NSCLC using molecularly targeted therapies [20] and immunotherapies [21], leading to positive outcomes for patients. These novel technologies include the detection of tumor markers [22], immunomodulatory drugs [23], and innovative chemotherapy regimens [24]. However, a majority of patients with advanced NSCLC exhibit resistance to current therapies and continue to progress in the final stage of the disease. Consequently, researchers are actively seeking new approaches to improve treatment outcomes and minimize side effects. In this context, we have explored a range of recently developed breakthrough therapies for NSCLC, with a particular focus on immunotherapy and targeted therapies.

In recent years, Immune checkpoint inhibitors have proven to be pivotal agents in cancer management due to their high selectivity, low toxicity, and minimal adverse effects on the body. They have become a focal point of research within clinical antitumor immunoregulatory therapy. However, therapeutic efficacy is frequently constrained in various immunologically 'cold' tumor types, such as non-small cell lung cancer (NSCLC), owing to challenges including tumor antigen deficiency and complications related to T cell activation and infiltration. Consequently, enhancing immunity in patients with 'cold' tumors has surfaced as a significant challenge for healthcare professionals [25]. The emergence of immunogenic cell death (ICD) presents new avenues for the diagnosis and treatment of NSCLC, offering unprecedented opportunities within the medical field. Research suggests that ICD-based cancer preventive vaccines can boost immune responses against cancerous "cold" tumors, while improving responsiveness to immunotherapeutic treatment. Therefore, exploring the robust association between NSCLC and genes differentially expressed during ICD may help unlock the potential of ICD-targeted therapies for more effective treatment strategies. Furthermore, tumor cells can use different mechanisms to evade immune system attacks by inhibiting or modifying immune cell functions and evading recognition by these cells [26]. These strategies enable tumors to escape immune surveillance while continuing their proliferation and metastasis. Thus, understanding how tumors interact with the immune system—and elucidating these underlying mechanisms—is crucial for developing effective tumor immunotherapy approaches. For instance, immune checkpoint inhibitors can enhance anti-tumor efficacy by relieving inhibition on immune cells [27], while CAR-T cell therapy augments T cell cytotoxicity against malignant cells [28]. These innovative therapies have demonstrated substantial progress in clinical practice and offer novel insights into future cancer management.

In this review, we have extensively identified and analysed the prognostic features associated with immunogenic cell death (ICD) in non-small cell lung cancer (NSCLC). We have examined the expressed profiles, functional roles, and methylation alterations of 235 ICD-related genes. By leveraging 17 ICD-associated genes, We developed and calibrated a risk label for ICDs. Our validation demonstrated that this model could serve as an innovative prognostic biomarker for NSCLC patients and may facilitate predictions regarding responses to immune checkpoint inhibitors (ICBs). Furthermore, we examined the relationships between risk scores and various clinico-pathological characteristics, immune profiles, and drug sensitivity independently. KEGG pathway enrichment analysis revealed that ICD-associated genes were involved in both the PI3K-AKT signal transduction cascade and natural killer cell-mediated epithelial cell death. Earlier studies have shown that ICD-relevant genes can elicit diverse cellular phenotypes through activation of the PI3K-AKT signal transduction cascade, which is in agreement with the results of our enrichment analysis [29]. The identified risk characteristics hold significant prognostic value and can effectively estimate individual mortality risks based on patient stratification. When integrated with other clinical factors, the nomogram demonstrates enhanced predictive accuracy. Survival curve analyses showed that people categorised as low risk were more likely to survive than those categorised as high risk.

The prognostic risk model we established consists of 17 genes identified through univariate and multi-covariate Cox regression analyses, as well as LASSO regression. These genes include AIRE, APOH, CDKN2A, CEACAM4, COL4A3, CPA, DBH, F10, FCGRB, FGFR4, MMP1, PGLYRP1, SCGB2A2, SLC9A3, UGT2B17, and VIP. APECED is an autoimmune condition associated with mutations in the AIRE gene. Research suggests that this gene is instrumental in establishing thymic T-cell tolerance [30]. In molecular proteomics studies involving NSCLC patients, APOH has been recognized as a potentially significant biomarker [31]. Although CDKN2A is commonly altered across various human cancers, its association with responsiveness to immune checkpoint inhibition (ICIs) remains ambiguous [32]. Decreased expression of COL4A3 has been linked to improved survival outcomes among NSCLC patients; however the underlying mechanisms are not yet fully elucidated [33]. The long non-coding RNA DBH-AS1 plays a critical role in the progression of multiple malignant tumors and exhibits downregulation in both NSCLC cell lines and tissue samples [34]. Studies have indicated that the presence levels of FCGRs on immune cells can trigger an antibody-mediated anti-tumor response. Polymorphisms within FCGRs may influence anti-tumor efficacy against specific immunotherapies [35]. While the precise function of PGLYRP1 in NSCLC remains unclear it has been observed that prolonged administration of recombinant PGLYRP1 could result in cardiovascular complications [36]. UGT2B17 has been implicated in promoting castration-resistant prostate cancer progression; however its involvement in NSCLC is still unknown [37]. Mammaglobin-1 (SCGB2A2) serves as a breast tumour-specific marker but its correlation with NSCLC lacks clarity [38]. SLC9A3-AS1 is expressed in NSCLC patient serum and tissues and cell lines, cancer cell growth and migration have been shown to be effectively inhibited by low levels of this gene [39].

Our researchers discovered that certain genes can promote or hinder the development of cancer cells by suppressing the tumor's

immune system. However, their role in influencing the ICD and the prognosis of individuals with NSCLC is still uncertain. In this study, we have divided the data collected into two groups based on the classification criteria used in the study: the high risk group and the low risk group. The predictive value of the model was then assessed using ROC curves. The predicament of the higher-risk group was significantly poorer than that of the lower-risk group, and our risk model had good predictive value for patient survival at both three and five years. We also performed external validations using two GEO birth cohorts, which further demonstrates the performance of our predictive model. Additionally, we verified the utility of our prognostic features in multiple cancers using lung cancer data from the TCGA database, and our results show that our constructed prognostic features have some generality.

To help predict patients survival, We created a nomogram containing the various clinical factors that can affect a patient's prognosis. Our researchers then performed a genome-wide association study (GSEA) to analyze the genes that are expressed in different groups of patients. We found a large number of pathways associated with cancer in high-risk populations, which further explains the poorer prognosis in the high risk subgroup of patients. Most T-cell and B-lymphocyte markers, including CD8<sup>+</sup> T-cells, neutrophils, NK-cells, pDC-cells, helper T-cells, follicular helper T (Th1) cells and T (Tem) cells, were strongly correlated with a better prognosis. To determine the reasons for the variations in the number of immune infiltrating cells, our researchers analysed the data collected from the two groups of patients. We found that the lower-risk subgroup had more immune-infiltrating cells than the high-risk group. This suggests that, based on the presence of immune cells, the model was able to predict the progression of the disease. A poor prognosis and a negative correlation with CD8<sup>+</sup> T cells was observed in the high-risk group. It suggests that the presence of immune cells could be a potential factor that could affect the outcome of patients with this type of cancer. We concluded that the findings could help develop new targeted treatment methods for this disease.

In summary, ICD-related genes play a critical path function in shaping the immunological microenvironment. The ICD-related risk profile identified and validated in NSCLC patients in this study is linked to alterations in the immune microenvironment of NSCLC tumors and may be predictive of response to immunotherapy. This study provides a comprehensive and personalised approach to guide future tumour immunotherapy strategies and sheds new light on the underlying mechanisms of NSCLC prognosis.

## 5. Conclusion

Our researchers have analysed immune infiltration and prognosis in NSCLC patients with the help of a gene model. Our findings were that the model was predictive of disease outcome on the basis of the presence of immune cells. New targeted treatments for this disease could be developed from the discovery of novel ICD-related genes.

## The ethics statement

Approval of the research protocol by an Institutional Reviewer Board: N/A. Informed Consent: N/A. Registry and the Registration No. of the study/trial: N/A. Animal Studies: N/A.

## CRedit authorship contribution statement

**Jian Zhang:** Project administration, Funding acquisition, Formal analysis, Data curation. **Huiying Li:** Formal analysis. **Xi Zhang:** Software, Resources. **Yue Yang:** Formal analysis, Data curation. **Yue Sun:** Writing – review & editing, Writing – original draft, Project administration.

## Data availability statement

TCGA data are available at <https://portal.gdc.cancer.gov/repository>: TCGA-LUSC, TCGA-LUAD; GeneCards database (<https://www.genecards.org/>); STRING (<https://cn.string-db.org/>); GSCALite(<http://bioinfo.life.hust.edu.cn/web/GSCALite/>).

## Funding

This work was supported by National Natural Science Foundation of China (Grant Number 82303616), Postdoctoral Science Foundation of China (Grant Number 2021MD703830), Postdoctoral Science Foundation of Heilongjiang (Grant Number LBH-Z20176), Fundamental Research Fund for Provincial University, Haiyan Science Foundation of Harbin Medical University Cancer Hospital (Grant Number JJYQ2024-07), Climbing plan (PDTs2024A-05), and Individualized and precise treatment of lung cancer (Nn10py2017-04).

## 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.

## References

- [1] K. Bajbouj, A. Al-Ali, R.K. Ramakrishnan, M. Saber-Ayad, Q. Hamid, Histone modification in NSCLC: molecular mechanisms and therapeutic targets, *Int. J. Mol. Sci.* 22 (21) (2021 Oct 28) 11701, <https://doi.org/10.3390/ijms222111701>.
- [2] R. Ruiz-Cordero, W.P. Devine, Targeted therapy and checkpoint immunotherapy in lung cancer, *Surg Pathol Clin* 13 (1) (2020 Mar) 17–33, <https://doi.org/10.1016/j.path.2019.11.002>.
- [3] J. Personnaz, E. Piccolo, M. Branchereau, A. Filliol, R. Paccoud, E. Moreau, D. Calise, E. Riant, P. Gourdy, C. Heymes, R.F. Schwabe, C. Dray, P. Valet, J. P. Pradère, Macrophage-derived HMGB1 is dispensable for tissue fibrogenesis, *FASEB Bioadv* 1 (4) (2019 Feb 12) 227–245, <https://doi.org/10.1096/fba.2018-00035>.
- [4] A. Ahmed, S.W.G. Tait, Targeting immunogenic cell death in cancer, *Mol. Oncol.* 14 (12) (2020 Dec) 2994–3006, <https://doi.org/10.1002/1878-0261.12851>. Epub 2020 Dec 1.
- [5] Z. Li, X. Lai, S. Fu, L. Ren, H. Cai, H. Zhang, Z. Gu, X. Ma, K. Luo, Immunogenic cell death activates the tumor immune microenvironment to boost the immunotherapy efficiency, *Adv. Sci.* 9 (22) (2022 Aug) e2201734, <https://doi.org/10.1002/adv.202201734>. Epub 2022 Jun 2.
- [6] H. Ruan, B.J. Leibowitz, L. Zhang, J. Yu, Immunogenic cell death in colon cancer prevention and therapy, *Mol. Carcinog.* 59 (7) (2020 Jul) 783–793, <https://doi.org/10.1002/mc.23183>. Epub 2020 Mar 25.
- [7] L. Galluzzi, A. Buqué, O. Kepp, L. Zitvogel, G. Kroemer, Immunogenic cell death in cancer and infectious disease, *Nat. Rev. Immunol.* 17 (2) (2017 Feb) 97–111, <https://doi.org/10.1038/nri.2016.107>. Epub 2016 Oct 17.
- [8] C. Brieske, P. Lamprecht, A. Kerstein-Staehle, Immunogenic cell death as driver of autoimmunity in granulomatosis with polyangiitis, *Front. Immunol.* 13 (2022 Oct 6) 1007092, <https://doi.org/10.3389/fimmu.2022.1007092>.
- [9] S.M. Jin, S.N. Lee, J.E. Kim, Y.J. Yoo, C. Song, H.S. Shin, H. Phuengkham, C.H. Lee, S.H. Um, Y.T. Lim, Overcoming chemoimmunotherapy-induced immunosuppression by assemblable and depot forming immune modulating nanosuspension, *Adv. Sci.* 8 (19) (2021 Oct) e2102043, <https://doi.org/10.1002/adv.202102043>. Epub 2021 Aug 7.
- [10] X. Wang, Y. Wu, J. Gu, J. Xu, Tumor-associated macrophages in lung carcinoma: from mechanism to therapy, *Pathol. Res. Pract.* 229 (2022 Jan) 153747, <https://doi.org/10.1016/j.prp.2021.153747>. Epub 2021 Dec 18.
- [11] Q.Y. Li, H.Y. Xu, H.J. Yang, [Effect of proinflammatory factors TNF- $\alpha$ , IL-1 $\beta$ , IL-6 on neuropathic pain], *Zhongguo Zhongyao Zazhi* 42 (19) (2017 Oct) 3709–3712, <https://doi.org/10.19540/j.cnki.cjcm.20170907.004>. Chinese.
- [12] M. Xu, J.H. Lu, Y.Z. Zhong, J. Jiang, Y.Z. Shen, J.Y. Su, S.Y. Lin, Immunogenic cell death-relevant damage-associated molecular patterns and sensing receptors in triple-negative breast cancer molecular subtypes and implications for immunotherapy, *Front. Oncol.* 12 (2022 Apr 4) 870914, <https://doi.org/10.3389/fonc.2022.870914>.
- [13] A. Blum, P. Wang, J.C. Zenkunen, SnapShot: TCGA-analyzed tumors, *Clin* 173 (2) (2018 Apr 5) 530, <https://doi.org/10.1016/j.cell.2018.03.059>.
- [14] D. Toro-Domínguez, J. Martorell-Marugán, R. López-Domínguez, A. García-Moreno, V. González-Rumayor, M.E. Alarcón-Riquelme, P. Carmona-Sáez, ImaGEO: integrative gene expression meta-analysis from GEO database, *Bioinformatics* 35 (5) (2019 Mar 1) 880–882, <https://doi.org/10.1093/bioinformatics/bty721>.
- [15] B. Xiao, L. Liu, A. Li, et al., Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma, *Front. Oncol.* 10 (2020) 607622.
- [16] W. Yang, J. Soares, P. Greninger, E.J. Edelman, H. Lightfoot, S. Forbes, N. Bindal, D. Beare, J.A. Smith, I.R. Thompson, S. Ramaswamy, P.A. Futreal, D.A. Haber, M.R. Stratton, C. Benes, U. McDermott, M.J. Garnett, Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells, *Nucleic Acids Res.* 41 (Database issue) (2013 Jan) D955–D961, <https://doi.org/10.1093/nar/gks1111>. Epub 2012 Nov 23.
- [17] X. Zeng, G. Shi, Q. He, P. Zhu, Screening and predicted value of potential biomarkers for breast cancer using bioinformatics analysis, *Sci. Rep.* 11 (1) (2021 Oct 21) 20799, <https://doi.org/10.1038/s41598-021-00268-9>.
- [18] P.M. Henson, Cell removal: efferocytosis, *Annu. Rev. Cell Dev. Biol.* 33 (2017 Oct 6) 127–144, <https://doi.org/10.1146/annurev-cellbio-111315-125315>. Epub 2017 Jun 14.
- [19] J. Ko, M.M. Winslow, J. Sage, Mechanisms of small cell lung cancer metastasis, *EMBO Mol. Med.* 13 (1) (2021 Jan 11) e13122, <https://doi.org/10.15252/emmm.202013122>. Epub 2020 Dec 9.
- [20] E.N. Imyanitov, A.G. Iyevleva, E.V. Levchenko, Molecular testing and targeted therapy for non-small cell lung cancer: current status and perspectives, *Crit. Rev. Oncol. Hematol.* 157 (2021 Jan) 103194, <https://doi.org/10.1016/j.critrevonc.2020.103194>. Epub 2020 Dec 11.
- [21] T. Cascone, J. Fradette, M. Pradhan, D.L. Gibbons, Tumor immunology and immunotherapy of non-small-cell lung cancer, *Cold Spring Harb Perspect Med* 12 (5) (2022 May 27) a037895, <https://doi.org/10.1101/cshperspect.a037895>.
- [22] H. Zhang, M. He, R. Wan, L. Zhu, X. Chu, Establishment and evaluation of EGFR mutation prediction model based on tumor markers and CT features in NSCLC, *J Healthc Eng* 2022 (2022 Apr 5) 8089750, <https://doi.org/10.1155/2022/8089750>.
- [23] O. Bílek, L. Bohovicová, R. Demlová, A. Poprach, R. Lakomý, L. Zdrážilová-Dubská, Nematobuněčný karcinom plic - od imunobiologie k imunoterapii [Non-Small cell lung cancer - from immunobiology to immunotherapy], *Klin. Onkol.* 29 (4) (2016 Fall) 78–87. Czech.
- [24] E. Felip, N. Altorki, C. Zhou, T. Csőszi, I. Vynnychenko, O. Goloborodko, A. Luft, A. Akopov, A. Martinez-Marti, H. Kenmotsu, Y.M. Chen, A. Chella, S. Sugawara, D. Voong, F. Wu, J. Yi, Y. Deng, M. McClelland, E. Bennett, B. Gitlitz, H. Wakelee, IMpower010 Investigators. Adjuvant atezolizumab after adjuvant chemotherapy in resected stage IB–IIIA non-small-cell lung cancer (IMpower010): a randomised, multicentre, open-label, phase 3 trial, *Lancet* 398 (10308) (2021 Oct 9) 1344–1357, [https://doi.org/10.1016/S0140-6736\(21\)02098-5](https://doi.org/10.1016/S0140-6736(21)02098-5). Epub 2021 Sep 20. Erratum in: *Lancet*. 2021 Sep. 23.
- [25] W.M. Brueckl, J.H. Ficker, G. Zeitler, Clinically relevant prognostic and predictive markers for immune-checkpoint-inhibitor (ICI) therapy in non-small cell lung cancer (NSCLC), *BMC Cancer* 20 (1) (2020 Dec 3) 1185, <https://doi.org/10.1186/s12885-020-07690-8>.
- [26] D.S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point, *Nature* 541 (7637) (2017 Jan 18) 321–330, <https://doi.org/10.1038/nature21349>.
- [27] L.B. Kennedy, A.K.S. Salama, A review of cancer immunotherapy toxicity, *CA Cancer J Clin* 70 (2) (2020 Mar) 86–104, <https://doi.org/10.3322/caac.21596>. Epub 2020 Jan 16. PMID: 31944278.
- [28] H. Wang, X. Guo, J. Zhou, Y. Li, L. Duan, X. Si, L. Zhang, X. Liu, M. Wang, J. Shi, L. Zhang, Clinical diagnosis and treatment of immune checkpoint inhibitor-associated pneumonitis, *Thorac Cancer* 11 (1) (2020 Jan) 191–197, <https://doi.org/10.1111/1759-7714.13240>. Epub 2019 Nov 24.
- [29] S. Noorolyai, N. Shajari, E. Baghbani, S. Sadreddini, B. Baradaran, The relation between PI3K/AKT signalling pathway and cancer, *Gene* 698 (2019 May 25) 120–128, <https://doi.org/10.1016/j.gene.2019.02.076>. Epub 2019 Mar 5. PMID: 30849534.
- [30] The absence of Aire results in impaired clonal deletion of self-reactive thymocytes, which escape into the periphery and attack a variety of organs 15 Mathis D, Benoist C. *Aire*, *Annu. Rev. Immunol.* 27 (2009) 287–312.
- [31] M. Pietrowska, K. Jelonek, M. Michalak, M. Roś, P. Rodziewicz, K. Chmielewska, K. Polański, J. Polańska, A. Gdowicz-Kłosok, M. Giglok, R. Suwiński, R. Tarnawski, R. Dziadziszko, W. Rzyman, P. Widiak, Identification of serum proteome components associated with progression of non-small cell lung cancer, *Acta Biochim. Pol.* 61 (2) (2014) 325–331. Epub 2014 May 29.
- [32] E. Adib, A.H. Nassar, E.W. Akl, S. Abou Alaiwi, P.V. Nuzzo, T.H. Mouhieddine, G. Sonpavde, R.I. Haddad, K.W. Mouw, M. Giannakis, F.S. Hodi, S.A. Shukla, A. Gusev, D.A. Braun, T.K. Choueiri, D.J. Kwiatkowski, CDKN2A alterations and response to immunotherapy in solid tumors, *Clin. Cancer Res.* 27 (14) (2021 Jul 15) 4025–4035.
- [33] C.P. Jiang, B.H. Wu, S.P. Chen, M.Y. Fu, M. Yang, F. Liu, B.Q. Wang, High COL4A3 expression correlates with poor prognosis after cisplatin plus gemcitabine chemotherapy in non-small cell lung cancer, *Tumour Biol* 34 (1) (2013 Feb) 415–420.
- [34] X.Y. Shi, X.F. Tao, G.W. Wang, J.F. He, L.F. Wu, Y.Z. Sun, X.J. Sun, LncDBH-AS1 knockdown enhances proliferation of non-small cell lung cancer cells by activating the Wnt signaling pathway via the miR-155/AXIN1 axis, *Eur. Rev. Med. Pharmacol. Sci.* 25 (1) (2021 Jan) 139–144.

- [35] A.K. Erbe, W. Wang, J. Goldberg, M. Gallenberger, K. Kim, L. Carmichael, D. Hess, E.A. Mendonca, Y. Song, J.A. Hank, S.C. Cheng, S. Signoretti, M. Atkins, A. Carlson, J.W. Mier, D.J. Panka, D.F. McDermott, P.M. Sondel, FCGR polymorphisms influence response to IL2 in metastatic renal cell carcinoma, *Clin. Cancer Res.* 23 (9) (2017 May 1) 2159–2168.
- [36] Y. Han, S. Hua, Y. Chen, W. Yang, W. Zhao, F. Huang, Z. Qiu, C. Yang, J. Jiang, X. Su, K. Yang, W. Jin, Circulating PGLYRP1 levels as a potential biomarker for coronary artery disease and heart failure, *J. Cardiovasc. Pharmacol.* 77 (5) (2021 May 1) 578–585, <https://doi.org/10.1097/FJC.0000000000000996>.
- [37] H. Li, N. Xie, R. Chen, M. Verreault, L. Fazli, M.E. Gleave, O. Barbier, X. Dong, UGT2B17 expedites progression of castration-resistant prostate cancers by promoting ligand-independent AR signaling, *Cancer Res.* 76 (22) (2016 Nov 15) 6701–6711.
- [38] I.M. Talaat, M.Y. Hachim, I.Y. Hachim, R.A.E. Ibrahim, M.A.E.R. Ahmed, H.Y. Tayel, Bone marrow mammaglobin-1 (SCGB2A2) immunohistochemistry expression as a breast cancer specific marker for early detection of bone marrow micrometastases, *Sci. Rep.* 10 (1) (2020 Aug 3) 13061.
- [39] X. Huang, M. Huang, M. Chen, X. Chen, lncRNA slc9a3-AS1 promotes oncogenesis of NSCLC via sponging microRNA-760 and may serve as a prognosis predictor of NSCLC patients, *Cancer Manag. Res.* 14 (2022 Mar 9) 1087–1098.