

A novel seven-immune-gene-based model to improve prediction of prognosis of clear cell renal cell carcinoma

Qingfu Su¹, Pengsheng Chen², Haixin Guo³, Zhijiao You⁴, Wei Zhuang¹

¹Department of Urology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China

²Department of Plastic Surgery, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China

³Department of Ultrasound, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China

⁴Department of Urology, Jinjiang Municipal Hospital, No. 16 Luoshan Section, Jinguang Road, Jinjiang, Quanzhou, Fujian, China

Submitted: 23 August 2022; **Accepted:** 24 December 2022

Online publication: 29 December 2022

Arch Med Sci

DOI: <https://doi.org/10.5114/aoms/158531>

Copyright © 2022 Termedia & Banach

Abstract

Introduction: The aim of our study was to investigate the correlation of immune-related genes with clear cell renal cell carcinoma (ccRCC) prognosis and the role of immune-related genes in the tumor immune microenvironment (TIME) and to build a new prognostic model and prognostic scoring system for renal cancer.

Material and methods: We downloaded the mRNA expression data of 610 samples (538 ccRCC and 72 normal tissues) from the TCGA database and constructed an immune-related prognostic model using Cox regression analysis and LASSO analysis. Then we internally verified the scientific validity and accuracy of the model using Kaplan-Meier (KM) analysis and receiver operating characteristic (ROC) curves. Subsequently, Cytoscape was used to construct a TF-miRNA-mRNA network. The “CIBERSORT” package was used to perform the immune infiltration analysis. Finally, validation of key gene expression was performed by immunohistochemistry (IHC) and quantitative reverse transcription-PCR (qRT-PCR).

Results: The prognostic model constructed for ccRCC includes 7 genes (*KLRC2*, *PGLYRP2*, *AGER*, *CHGA*, *AVPR1B*, *IL20RB*, *LAT*). It was proven to have good prognostic performance through the K analysis and the ROC curves. We also constructed an accurate prognostic predictive scoring system by establishing a nomogram. Furthermore, the TF-miRNA-mRNA network revealed the potential mechanism of the model and the immune infiltration analysis revealed a correlation between this model and TIME.

Conclusions: The results suggest that the newly developed 7-immune-related-gene model can be a practical and reliable prognostic tool for ccRCC. It also shows T cell infiltration characteristics in TIME and can therefore be used as an immune biomarker for the diagnosis and treatment of ccRCC.

Key words: immune-related genes, prognostic model, TF-miRNA-mRNA regulatory network, tumor immune microenvironment.

Corresponding authors:

Prof. Wei Zhuang
Department of Urology
The Second Affiliated
Hospital of Fujian
Medical University
Quanzhou, Fujian
China 362000
Phone: +8613799222133
Fax: 0595-22768074
E-mail: 612032@fjmu.edu.cn

Prof. Zhijiao You
Department of Urology
Jinjiang Municipal Hospital
No. 16, Luoshan Section
Jinguang Road
Jinjiang, Quanzhou
Fujian, China 362200
Phone: +86 15159840207
Fax: 0595-85626975
E-mail: youzhijiao@fjmu.edu.cn

Introduction

Renal cell cancer (RCC) is one of the most common and deadliest malignancies of the urinary system [1]. In 2020, there were an estimated 431,288 new cases of renal cancer, 179,368 deaths caused by the disease world-

wide and 73,000 new cases diagnosed in China [2]. Clear cell renal cell carcinoma (ccRCC) is the most common type of renal cancer that is associated with different risk factors such as smoking, obesity, high blood pressure, germline, somatic mutations, etc. that predispose patients to ccRCC development [3]. Patients with localized RCC receiving treatments such as radiofrequency ablation, enucleoresection or simple nephrectomy have improved disease-free survival rates up to 90% [4, 5], while the treatment methods for advanced and metastatic ccRCC are still limited. Immunotherapy is currently used as a first-line treatment [6], but only a small proportion of patients can benefit from immunotherapy due to the complexity and highly modulated nature of the immune system [7]. Therefore, it is urgent to find more clues related to the prognosis of ccRCC from the perspective of immunity and seek new immune biomarkers to provide new ideas for the immunotherapy and prognosis of ccRCC.

The growth of research on the tumor immune microenvironment (TIME) has promoted the development of immunotherapy for renal cancer. The ccRCC is considered an immunogenic tumor that is known to mediate immune dysfunction largely by inducing infiltration of immunosuppressive cells into the TIME [8]. Existing studies suggest that TIME plays a key role in the tumor's initiation, progression, metastasis and response to treatment [9]. However, research on TIME for renal cancer remains rare. Further exploration of the immune microenvironment of ccRCC is needed for the discovery of new clues that can contribute to immunotherapy in renal cancer.

It is considered scientific to find prognostic factors using univariate and multivariate Cox regression analyses, which were also used in much relevant literature [10, 11], while little research has been done on ccRCC prognostic models, let alone immune-related prognostic models. Given that most prognostic models focus on analyzing their impact on survival while neglecting to explore the mechanism behind their impact, in our research we constructed a significant prognostic model for ccRCC, the further mechanism of which was also explored. We attempted to investigate the possible mechanisms affecting prognosis from the perspective of immune infiltration and the transcription factor (TF)-microRNA(miRNA)-mRNA regulatory network while searching for key targets or new biomarkers for renal cancer immunotherapy and prognosis.

Material and methods

Identification of differentially expressed immune genes (DEIGs) and functional enrichment

The gene expression data of ccRCC were downloaded from The Cancer Genome Atlas (TCGA) da-

tabase (<https://portal.gdc.cancer.gov/>) [12] as the training set (including 538 ccRCC samples and 72 normal samples). We further collected the clinical information of 537 ccRCC samples corresponding to the transcriptome data. Then we used the “limma” package to screen the differentially expressed genes (DEGs) of ccRCC ($\log_2 |FC| > 2$ and p -value < 0.05 were considered statistically significant). We collected the immune-related gene (IRGs) list from the Immunology Database and Analysis Portal (ImmPort) database (<https://immport.niaid.nih.gov/home>, Supplementary Table S1) and mixed the immune-related genes with DEGs. The overlap of DEGs and immune-related genes was considered as differentially expressed immune-related genes (DEIGs). Results were visualized using volcano plots and Venn diagrams. Gene ontology (GO) [13] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [14] analyses were executed using the “Cluster Profiler” R package to determine the biological processes and enrichment pathways of DEIGs. KEGG is a database that integrates genomic, chemical and systemic functional information. GO databases are used to describe cellular components (CC), biological processes (BP), and molecular functions (MF). The adjusted p -value less than 0.05 was considered statistically significant.

Development of immune risk scoring signature (IRSS) for prognosis and validation of IRSS

Univariate Cox regression analysis was performed on DEIGs to identify important immune genes associated with prognosis. Next, independent prognostic genes were obtained through least absolute shrinkage and selection operator (LASSO) regression analysis, and multivariate Cox regression analysis was used to obtain regression coefficients for the independent prognostic genes. Ultimately, 7 prognostic genes – *KLRC2*, *PGLYRP2*, *AGER*, *CHGA*, *AVPR1B*, *IL20RB*, and *LAT* – were selected. The immune risk score signature (IRSS) was established based on the 7 prognostic genes, with the formula as follows: $IRSS = EXP_{gene1} \times \beta_1 + EXP_{gene2} \times \beta_2 + \dots + EXP_{genen} \times \beta_n$ (EXP : gene expression level, β : regression coefficient). Patients were divided into a high-risk group and a low-risk group with the median risk score determined as the cut-off value. The Kaplan-Meier (KM) curves were used to compare the overall survival (OS) of the two groups. The predictive value of the model was analyzed using the ROC curve. “Survival” packages were used to plot KM curves. ROC curves were drawn by the “survival ROC” package. To assess the universality of the IRSS, 538 tumor samples were randomly divided into two groups of equal size for internal validation. Due to lost clinical information of some samples, the final

sample size of test group 1 is 257, and the size of test group 2 is 236.

Immune risk score signature combined with clinicopathological information

Univariate and multivariate Cox regression analyses combined with clinicopathological information were used to assess independent prognostic predictors for ccRCC patients. Furthermore, to comprehensively assess patient survival, independent prognostic predictors such as risk score, age, stage and grade were combined to construct a nomogram. The clinical application value of the nomogram was evaluated by decision curve analysis (DCA).

Construction of transcription factor (TF)-microRNA(miRNA)-mRNA regulatory network

MiRNA and TF that may regulate target genes were detected using the miRNAs-target interaction databases TargetScan (a web server that predicts biological targets of miRNAs) [15], TransmiR (a database for TF-miRNA regulations) [16] and miRTarBase (a database of experimentally validated miRNA targets) [17]. The TF-miRNA-mRNA network was constructed using Cytoscape (a software platform for visualizing complex networks) [18] to determine the participation of the 7 genes, miRNAs and TFs in the network. By calculating the “node degree” of the interaction, the target genes, miRNAs and TFs were found to have the highest participation in the network.

Immunohistochemistry and quantitative real-time PCR (qRT-PCR)

Immunohistochemical results were obtained from the Human Protein Atlas (HPA) database [19] (<https://www.proteinatlas.org>, clinical information of patients and the antibodies used are shown in Supplementary Tables SII and SIII). We selected 24 tissue samples of ccRCC and paired adjacent normal tissues (clinical information of patients is shown in Supplementary Table SIV) from the Department of Urology, the Second Affiliated Hospital of Fujian Medical University (2019.01.01–2020.01.01). The total RNA of ccRCC patients' tissue samples was isolated by Tissue RNA Purification Kit Plus (RN-002plus, ESscience Biotech, Shanghai, China). The total RNA quantity was measured by a NanoDrop spectrophotometer (ESscience Biotech, Shanghai, China). A Reverse Transcription Kit (CW BIO, Jiangsu, China) and UltraSYBR Mixture SYBR Green (CW BIO) were used to synthesize cDNA and for qRT-PCR detection. Finally, qRT-PCR reactions were performed using the ABI7500 Fluorescent Quantitative PCR Instrument (CW BIO). The sequences of the *AGER* primers used were as follows: 5'-CACCTTCTCTG-

TAGCTTCAGC-3' (forward), 5'-AGGAGCTACTGCTC-CACCTTCT-3' (reverse). The sequences of the *LAT* primers used were as follows: 5'-ATCCTGGAGCG-GCTAAGACTGA-3' (forward), 5'-GTTTCAGCTCCTG-CAGATTCTCG-3' (reverse).

For *GAPDH* (internal control), the primers were as follows: 5'-GGAGTCAACGGATTGGT-3' (forward), 5'-GTGATGGGATTCCATTGAT-3' (reverse).

The relative expression level in this study was calculated with the $2^{-\Delta\Delta CT}$ formula.

Analysis of immune infiltration

Based on the “CIBERSORT.R” package, the relationship between immune infiltration and the IRSS model was explored ($p < 0.05$). We used the transcriptome data from the immune infiltration file to obtain the infiltration volume of 22 immune cells in ccRCC. The difference in immune cell content between the distinct risk groups was also analyzed. In addition, the immune cells with significant differences in expression of the two groups were determined ($p < 0.05$).

Ethics statement

This study was approved by the Research Ethics Committee of Fujian Provincial Hospital and complied with the Helsinki Declaration (No. 2022469).

Statistical analysis

The statistical analysis was performed in the R language software (version 4.1.2), through which the volcano plots, Venn diagrams, forest plots, KM curves and ROC curve visualization were obtained. The univariate and multivariate Cox parameter range of prognostic mRNA is $p < 0.05$. The LASSO parameter of the prognostic mRNA is $p < 0.05$. Univariate and multivariate Cox regression analyses were used to analyze the relationship between clinical information and prognosis, and the entry criterion is $p < 0.05$.

Results

Analysis of differentially expressed immune genes (DEIGs)

By comparing 538 ccRCC samples with 72 normal samples, we performed mRNA differential expression analysis. 6,498 DEGs were extracted after the analysis (Figure 1 A) and 427 DEIGs were obtained (Figures 1 B, C), including 77 down-regulated DEIGs and 350 up-regulated DEIGs.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Analyses

The functional enrichment analysis showed that most of the 427 DEIGs in BP, CC and MF were

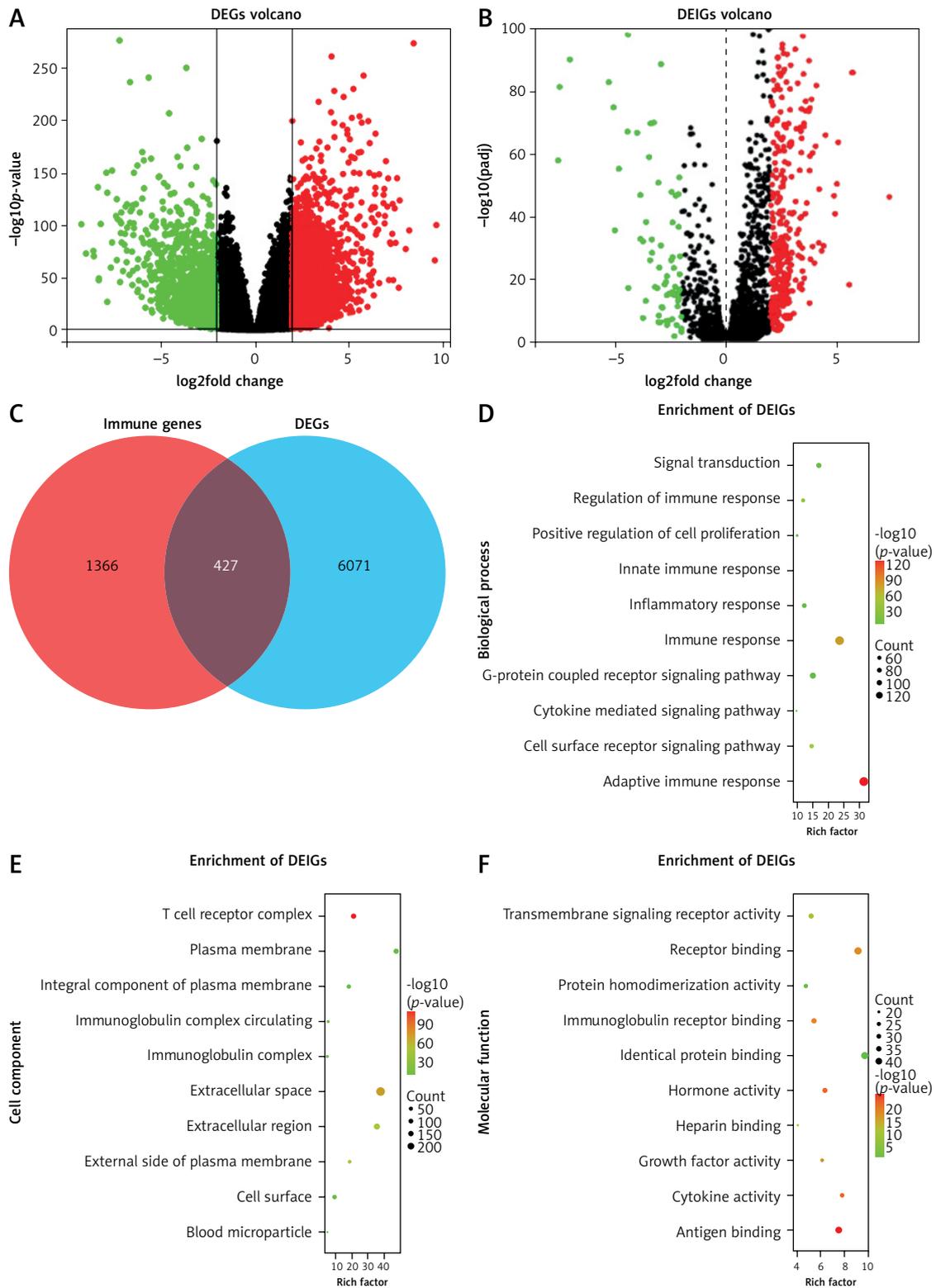


Figure 1. Screening for differentially expressed immune genes (DEIGs) and establishing the IRSS signature. **A** – Volcano plot of 6498 differentially expressed genes (DEGs). Selection criteria: $\log_2 |FC| > 2$, $p < 0.05$. **B** – Volcano plot of 427 differentially expressed immune genes (DEIGs). **C** – Venn diagram of the intersection of DEGs and immune genes. **D** – Biological process of DEIGs. **E** – Cellular component of DEIGs. **F** – Molecular function of DEIGs

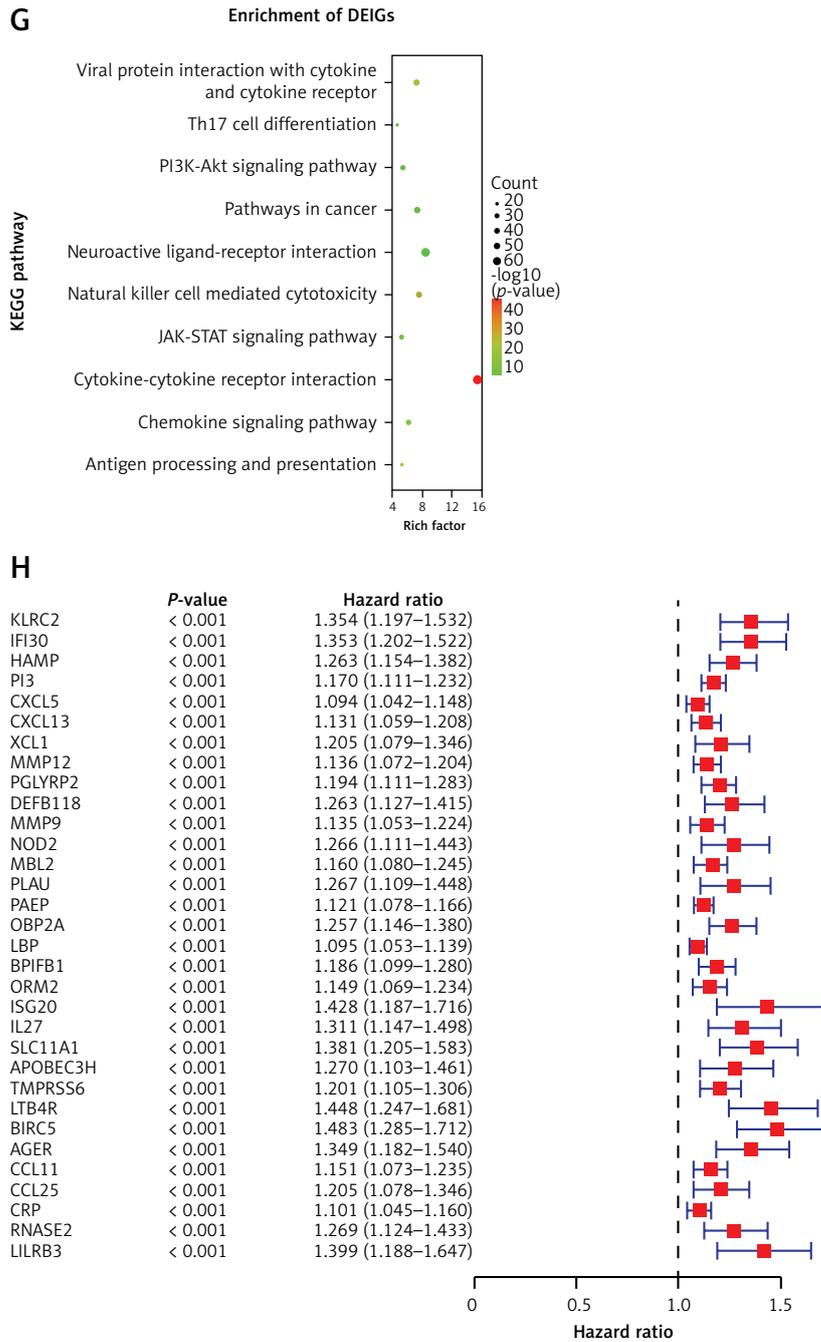
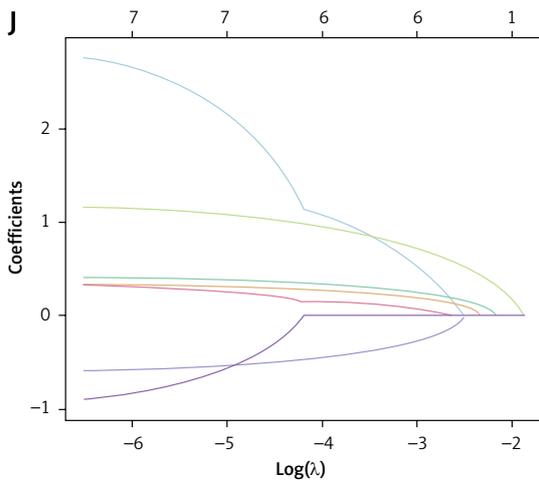
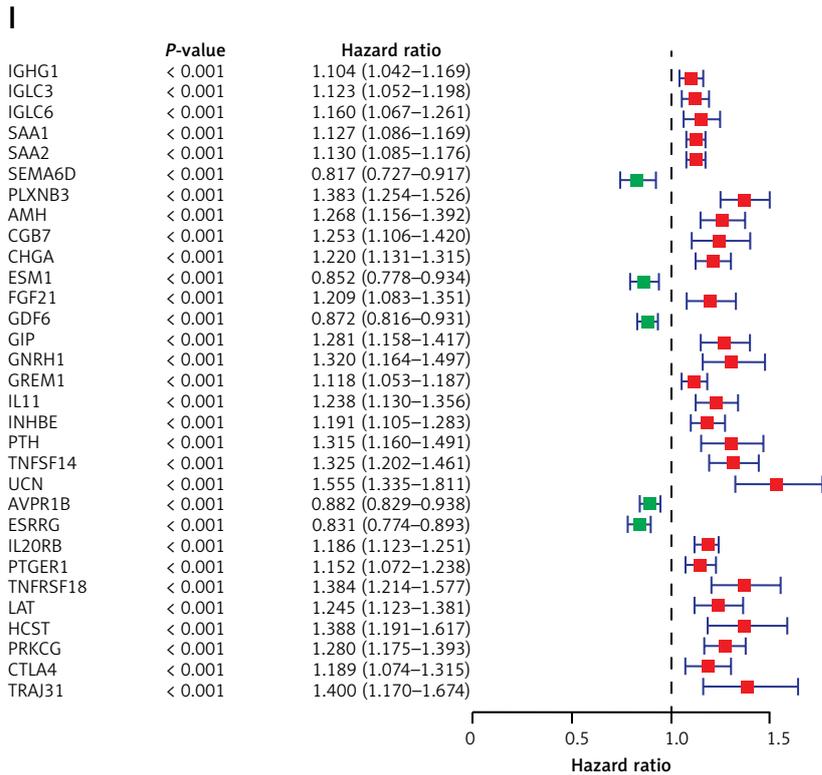


Figure 1. Cont. **G** – KEGG pathway enrichment results of DEIGs. **H** – Univariate Cox regression analysis identified 63 immune genes significantly associated with prognosis ($p < 0.05$)



K Table I. Multivariate CIX regression analysis

| Gene | Coefficient | HR | HR 95L Low | HR 95H High | P-value |
|---------|--------------|--------|------------|-------------|-------------|
| KLRC2 | 0.343855309 | 1.410 | 1.033 | 1.925 | 0.030331184 |
| PGLYRP2 | 0.408092916 | 1.504 | 1.186 | 1.907 | 0.000750895 |
| AGER | 2.92913597 | 18.711 | 4.712 | 74.304 | 3.14E-05 |
| CHGA | 0.331782282 | 1.393 | 1.108 | 1.753 | 0.004603584 |
| AVPR1B | -0.593999191 | 0.552 | 0.434 | 0.703 | 1.38E-06 |
| IL2ORB | 1.170570329 | 3.224 | 2.155 | 4.822 | 1.21E-08 |
| LAT | -0.983455296 | 0.374 | 0.197 | 0.711 | 0.002682769 |

Figure 1. Cont. I – Univariate Cox regression analysis identified 63 immune genes significantly associated with prognosis ($p < 0.05$). J – Ten-fold cross-validation for tuning parameter selection in the LASSO model ($p < 0.05$); the model fitted best with the penalty coefficient of 7. K – Multivariate Cox regression analyses revealed 7 genes as prognostic predictors ($p < 0.05$), namely *KLRC2*, *PGLYRP2*, *AGER*, *CHGA*, *AVPR1B*, *IL2ORB*, *LAT*

enriched in immune-related pathways, including adaptive immune response, T cell receptor complex, and antigen binding, which are the most common enrichment pathways (Figures 1 D–F). KEGG analysis revealed (Figure 1 G) that these DEIGs were significantly involved in cytokine-cytokine receptor interactions, natural killer cell-mediated cytotoxicity and viral protein interactions with cytokines and cytokine receptors.

Construction and validation of IRSS

Univariate Cox regression analysis identified 63 immune genes significantly associated with prognosis (Figures 1 H, I). In the following LASSO regression analysis (Figure 1 J), the model fitted best into the penalty coefficient of 7. In multivariate Cox regression analysis (Figure 1 K), the 7 immune genes (*KLRC2*, *PGLYRP2*, *AGER*, *CHGA*, *AVPR1B*, *IL20RB*, *LAT*) were determined as prognostic predictors with the corresponding regression coefficient β of 0.344, 0.408, 2.929, 0.332, -0.594, 1.171 and -0.983, respectively. According to the above-listed formula, the IRSS was finally established: $IRSS = EXP\ KLRC2 \times 0.344 + EXP\ PGLYRP2 \times 0.408 + EXP\ AGER \times 2.929 + EXP\ CHGA \times 0.332 + EXP\ AVPR1B \times -0.594 + EXP\ IL20RB \times 1.171 + EXP\ LAT \times -0.983$.

According to the formula, the risk scores of ccRCC patients were calculated. The samples were then divided into a high-risk group (245 samples) and a low-risk group (246 samples) based on the risk scores. As the risk score increases, patients' life span decreases (Figure 2 A1). The heat map in Figure 2 A1 also shows the differential expression of the 7 DEIGs in the low- and high-risk groups.

The KM curves indicated that there was a significant difference in overall survival (OS) between the low-risk group and the high-risk group ($p < 0.05$), and the prognosis of the high-risk group was worse than that of the low-risk group (Figure 2 A2). The area under the curve (AUC) of the ROC curve was 0.802 (Figure 2 A3), which showed that the model has excellent accuracy and realism. The time-dependent AUC values of 1 year, 2 years, 3 years and 5 years were 0.803, 0.745, 0.745, and 0.774, respectively (Figure 2 A4), showing that the model also has good predictive power for patients' long-term survival. These results indicated that the model is robust and accurate in predicting patient prognosis.

The test groups (test group 1, test group 2) were evaluated with the same formula to determine the reliability of the clinical prognostic model. Patients were divided into a high-risk group and a low-risk group according to the risk scores of the model. Increased risk score was associated with higher mortality (Figures 2 B1, C1), which meant that the

OS of the high-risk group was lower than that of the low-risk group (Figures 2 B2, C2). Furthermore, the accuracy of the clinical prognostic model was evaluated. The AUC of the ROC curves were 0.719 and 0.753 respectively (Figures 2 B3, C3) and the time-dependent AUC values were 0.721, 0.632, 0.680, 0.701 and 0.763, 0.748, 0.717 and 0.812, respectively (Figures 2 B4, C4). The results showed that this clinical prognostic model has great predictive potential, which can accurately predict the prognosis of ccRCC patients.

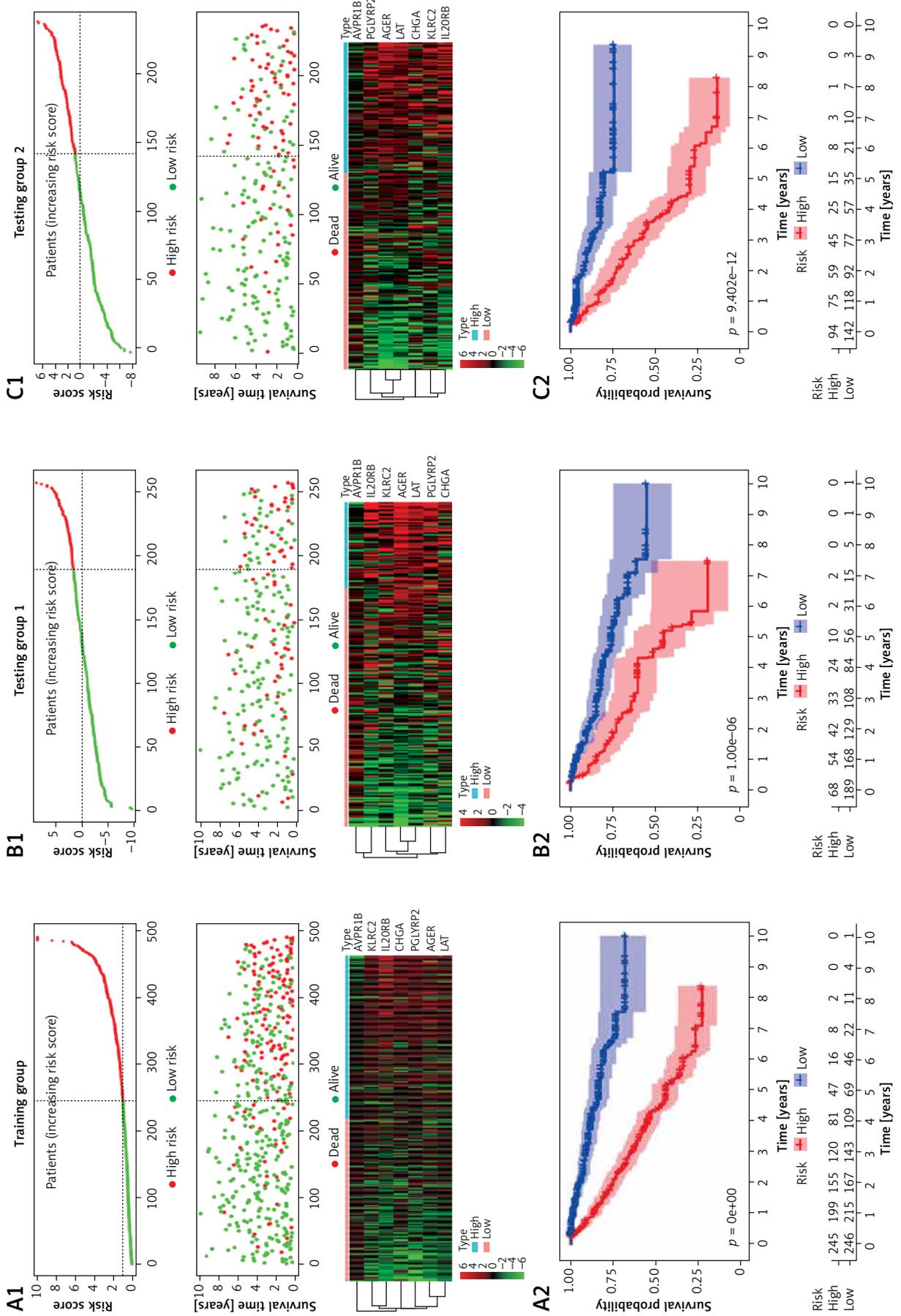
Correlations of clinical traits of the prognostic model

The univariate and multivariate Cox regression analyses showed that risk score, age, stage, and grade are independent prognostic factors (Figures 3 A, B). In addition, the association of risk scores with clinical demographic characteristics (age, sex, stage, and TNM stage) was analyzed (Figure 3 C), and the results showed that the risk scores were significantly correlated with tumor grade and stage ($p < 0.05$).

The model that incorporated the above independent predictors was developed and presented as a nomogram (Figure 3 D). The DCA results showed that the prognostic prediction ability of the nomogram in ccRCC was better than that of a single independent prognostic factor and proved that the nomogram combined with various clinical features has a better clinical application value ($p < 0.001$) (Figure 3 G). Based on clinical features and IRSS, predictive nomograms of the test groups were established to predict the prognostic survival probability of patients at 1, 2, 3 and 5 years (Figures 3 E, F). The DCA curves of the test groups suggested that there was a clear benefit from intervention following the results of the nomogram ($p < 0.001$) (Figures 3 H, I).

TF-miRNA-mRNA regulatory network

We constructed the TF-miRNA-mRNA regulatory network based on the prognostic model of the 7 DEIGs. The network was composed of 35 nodes (proteins) and 43 edges (interactions). Nodes that have the most interactions were considered hub genes. Among the 35 nodes, 2 were identified as hub genes (*LAT* and *AGER*) according to the criteria of node degree > 10 (Supplementary Table SV). Therefore, *LAT* and *AGER* are determined to be the hub genes that interact extensively with miRNAs (Figure 4 A). *MiR-584-5p* was correlated with both hub genes and the correlation coefficient was > 0.6 (Supplementary Table SVI). *E2F1* (degree = 4) is an important transcription factor in regulatory networks and interacts with hub genes *AGER* and *LAT* (Figure 4 B).



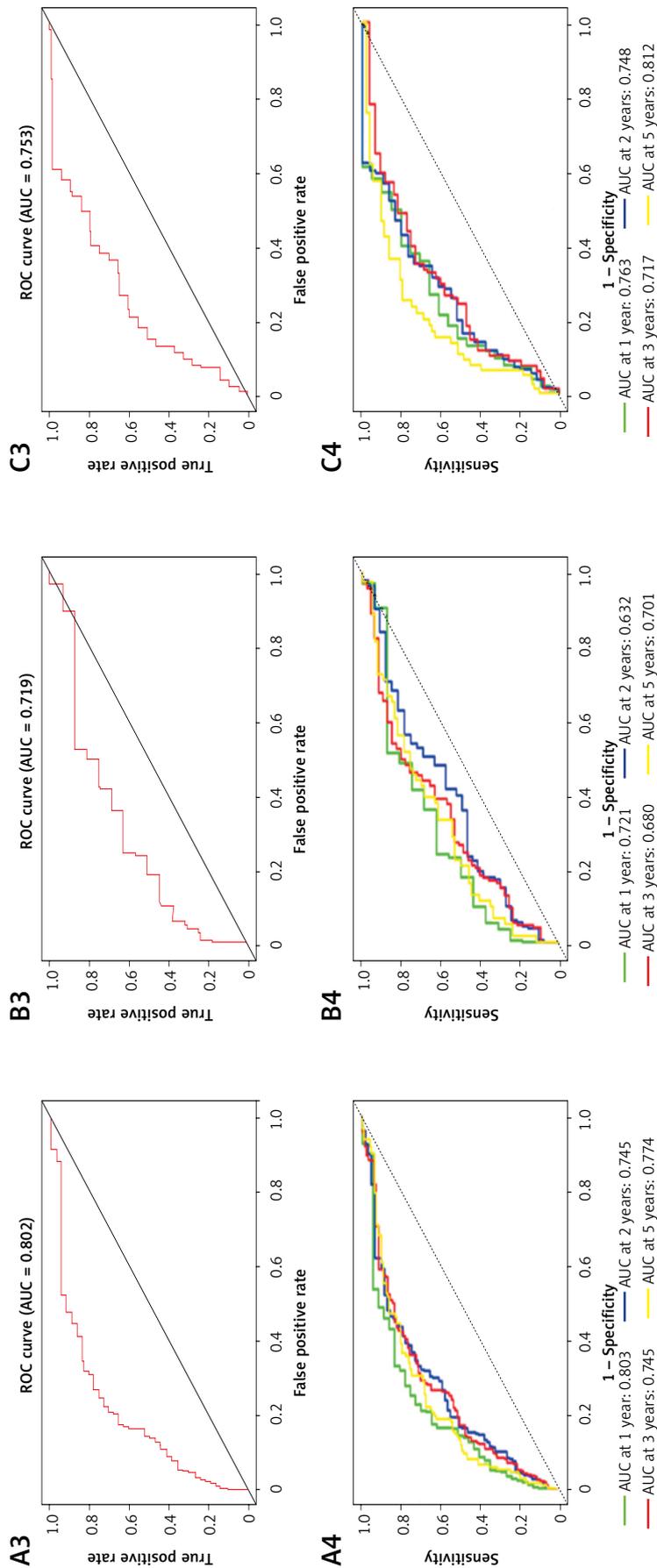


Figure 2. Evaluation and validation of clinical prognostic models. **A1, B1, C1** – The risk score, survival status and heatmap of seven immune genes in training group, testing group 1, testing group 2. **A2, B2, C2** – The risk scores of the clinical prognostic model predict survival. The KM curve suggests that the high-risk group has a worse prognosis than the low-risk group ($p < 0.05$). **A3, B3, C3** – The receiver operating characteristic curve (ROC) verifies the prognostic value of the model for the training group, testing group 1, testing group 2; AUC values are 0.802, 0.719 and 0.753 respectively. **A4, B4, C4** – The time-dependent ROC curve for the training group (AUC values were 0.803, 0.745, 0.745, and 0.774), testing group 1 (1, 2, 3, and 5 years AUC values were 0.721, 0.632, 0.680, 0.701), testing group 2 (AUC values were 0.763, 0.748, 0.717 and 0.812)

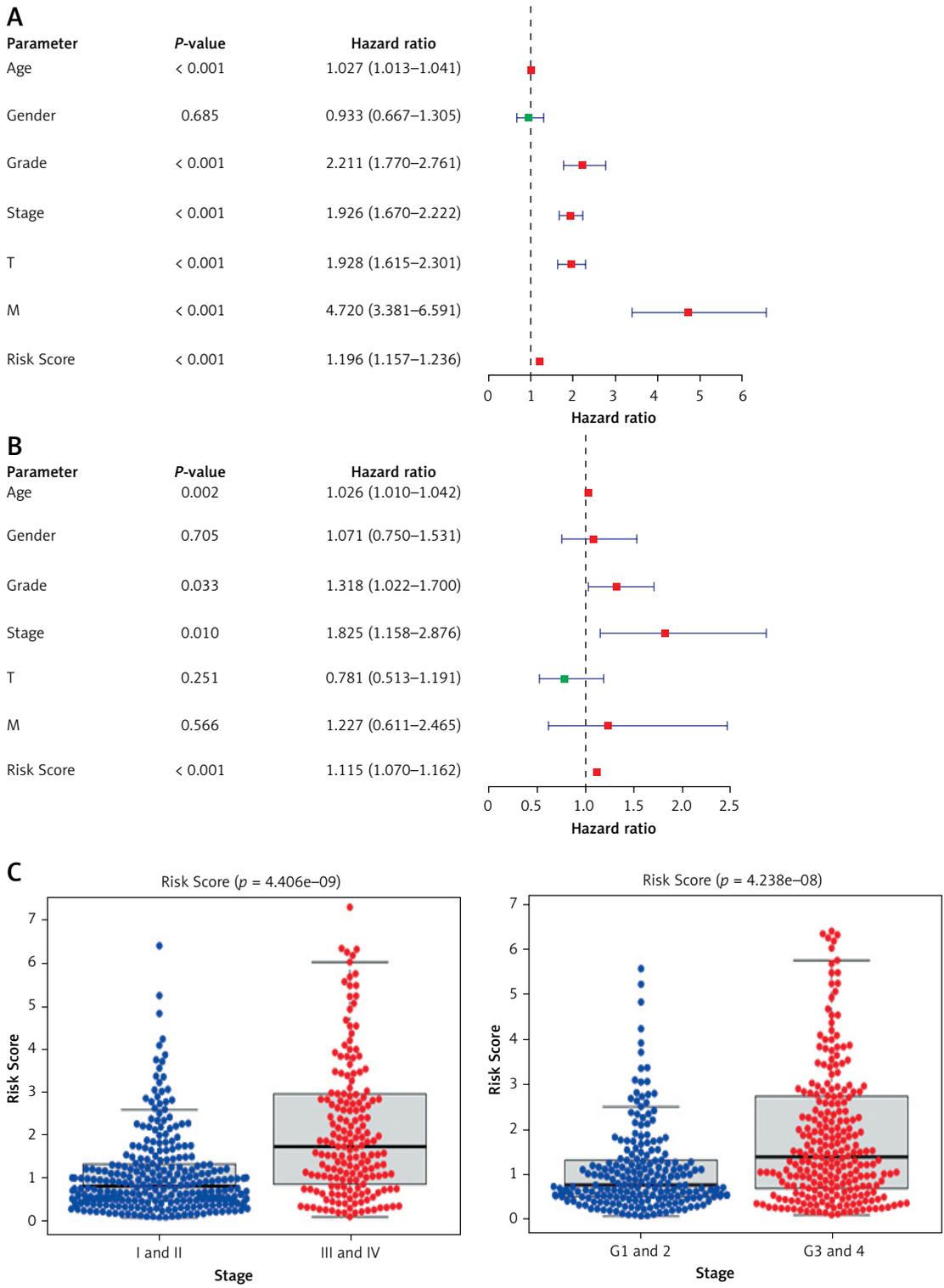


Figure 3. Independent prognostic analysis of clinical information and establishment and evaluation of nomograms for training group, testing group 1 and group 2. **A, B** – Univariate Cox regression analysis and multivariate Cox regression analysis of clinical information. Risk score, age, stage, and grade were independent prognostic factors ($p < 0.05$). **C** – Correlation between risk score and patients’ clinical and demographic characteristics

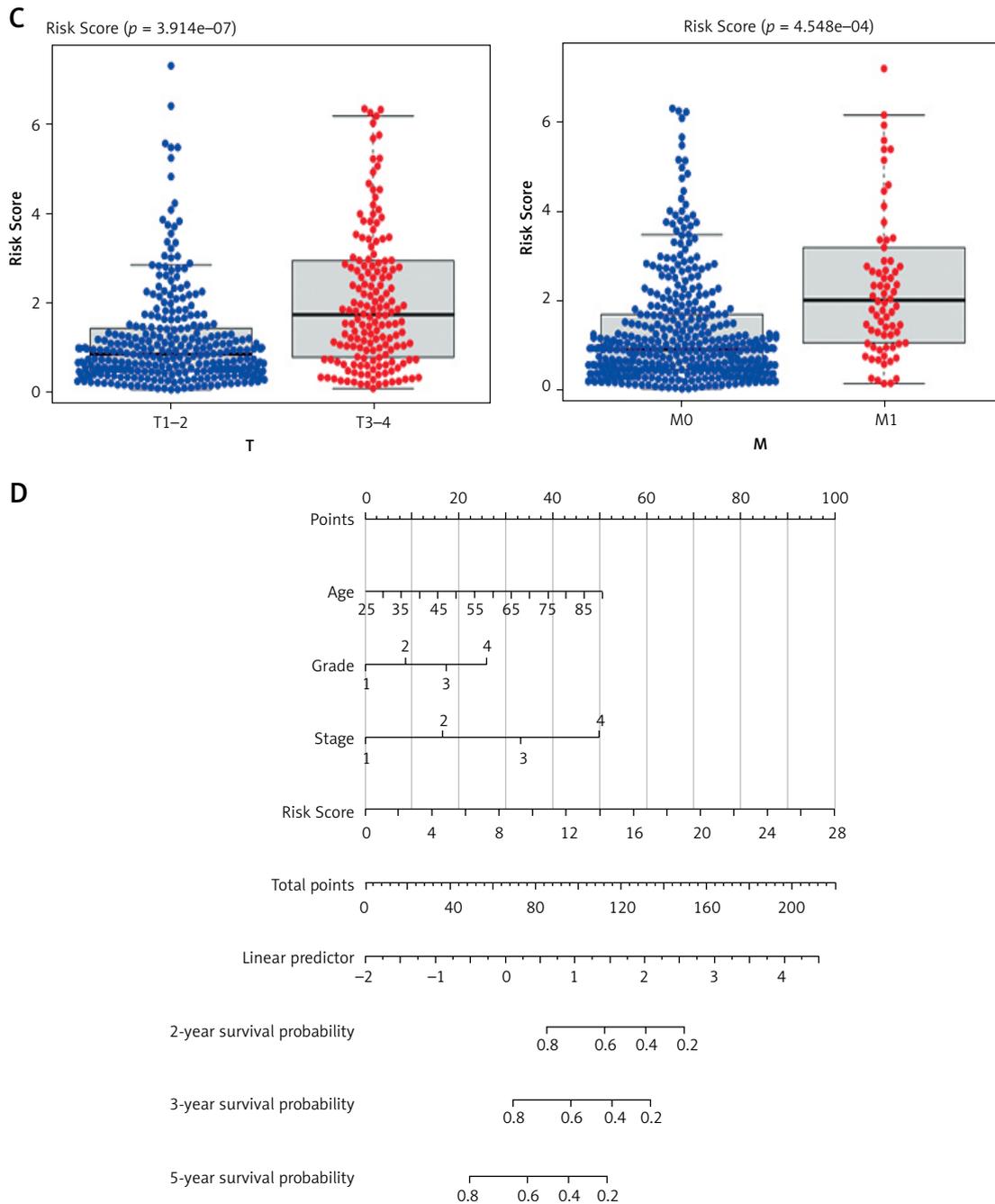
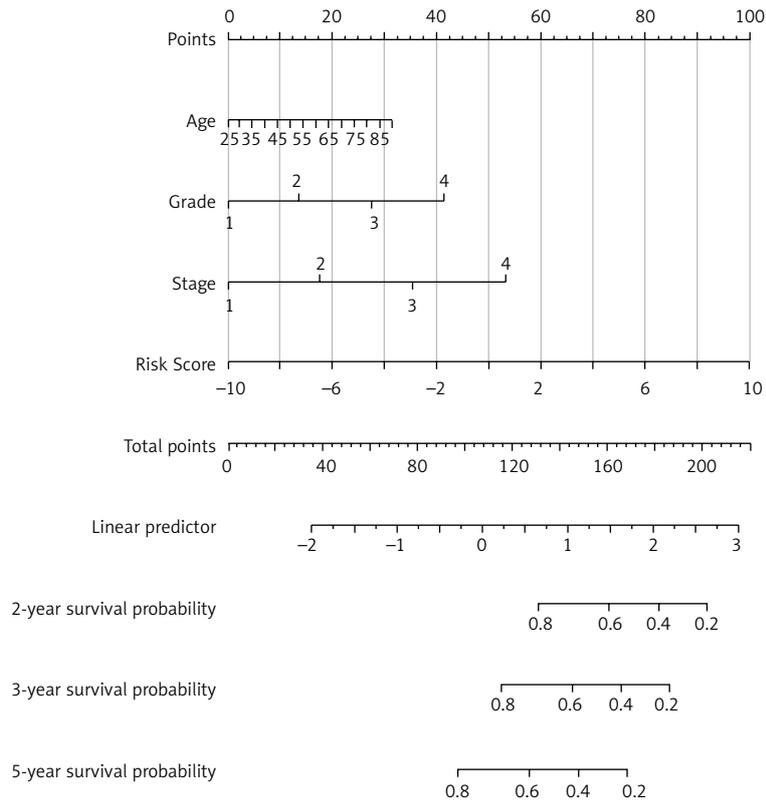


Figure 3. Cont. C – Correlation between risk score and patients' clinical and demographic characteristics. D–F - 2-, 3-, and 5-year nomogram for predicting OS of ccRCC patients. There are four components in this nomogram: age, grade, stage, risk score

E



F

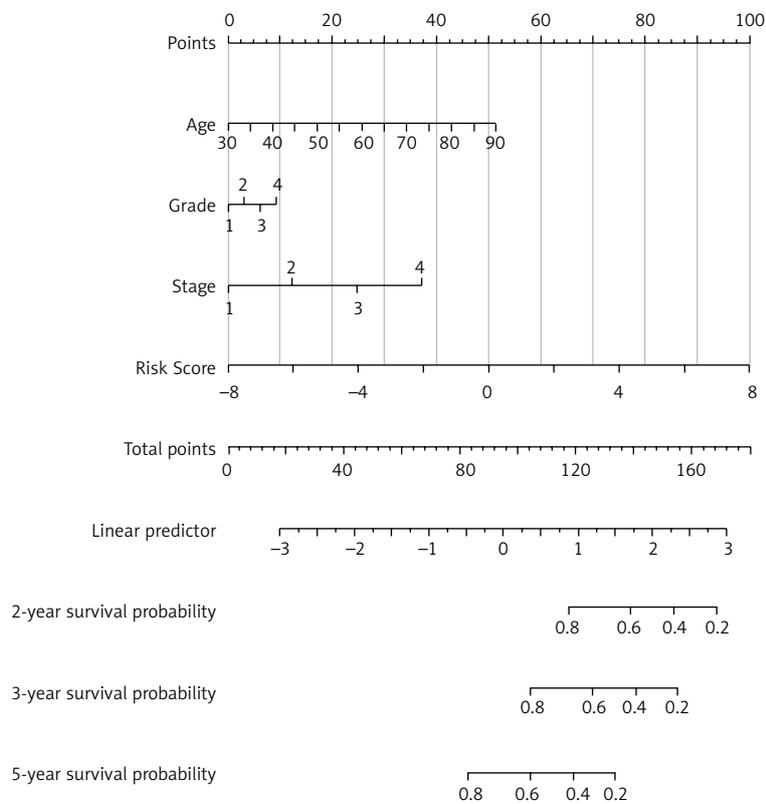


Figure 3. Cont. E-F - 2-, 3-, and 5-year nomogram for predicting OS of ccRCC patients. There are four components in this nomogram: age, grade, stage, risk score

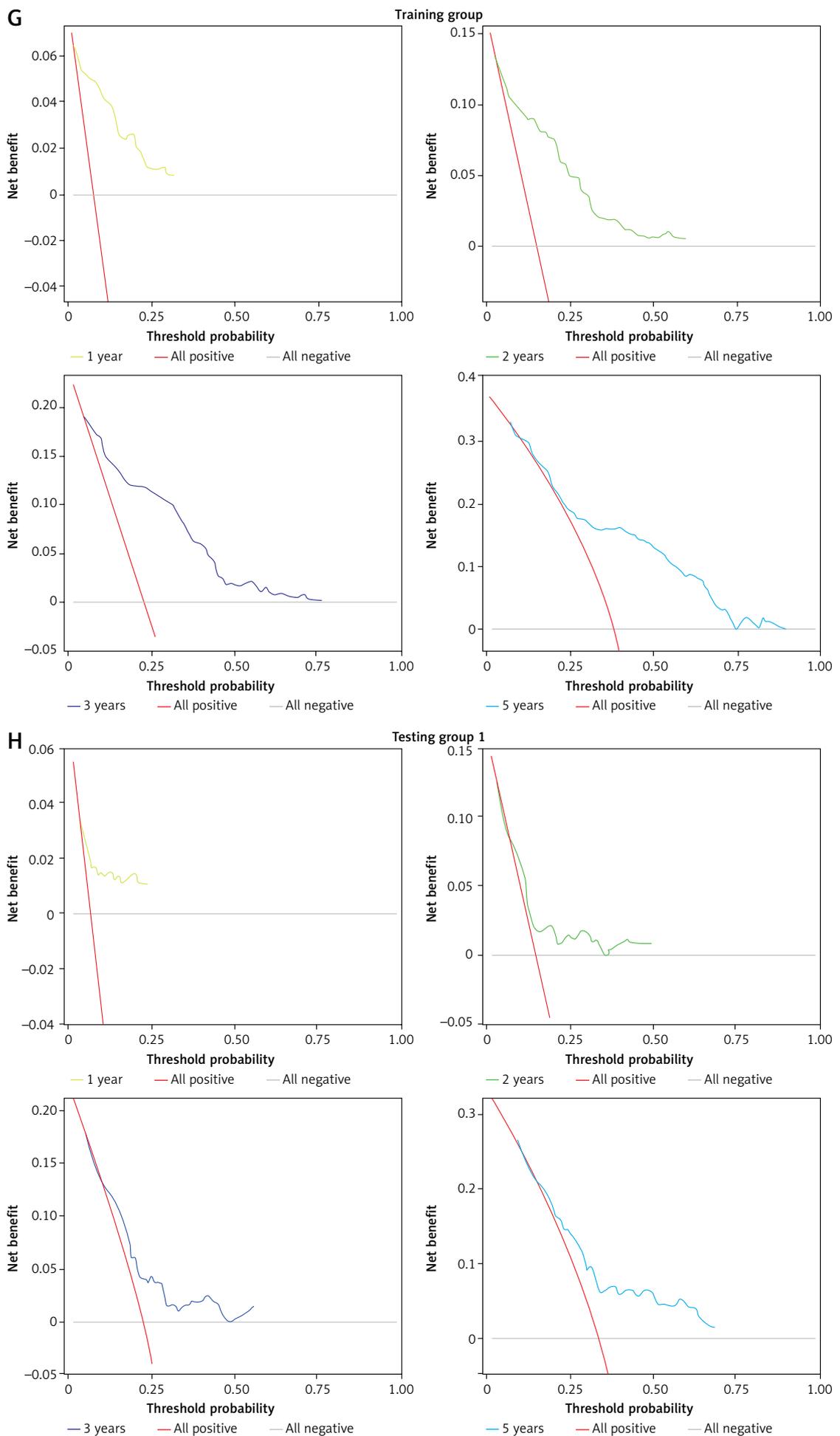


Figure 3. Cont. G–H – Decision curve analysis (DCA) for evaluation of the net benefits at 1, 2, 3, and 5 years. The x-axis represents the percentage of the threshold probability, and the y-axis represents the net benefits. Net benefits: when the threshold probability is exceeded, the use of column line plots to predict survival adds more benefit than all other prediction scenarios

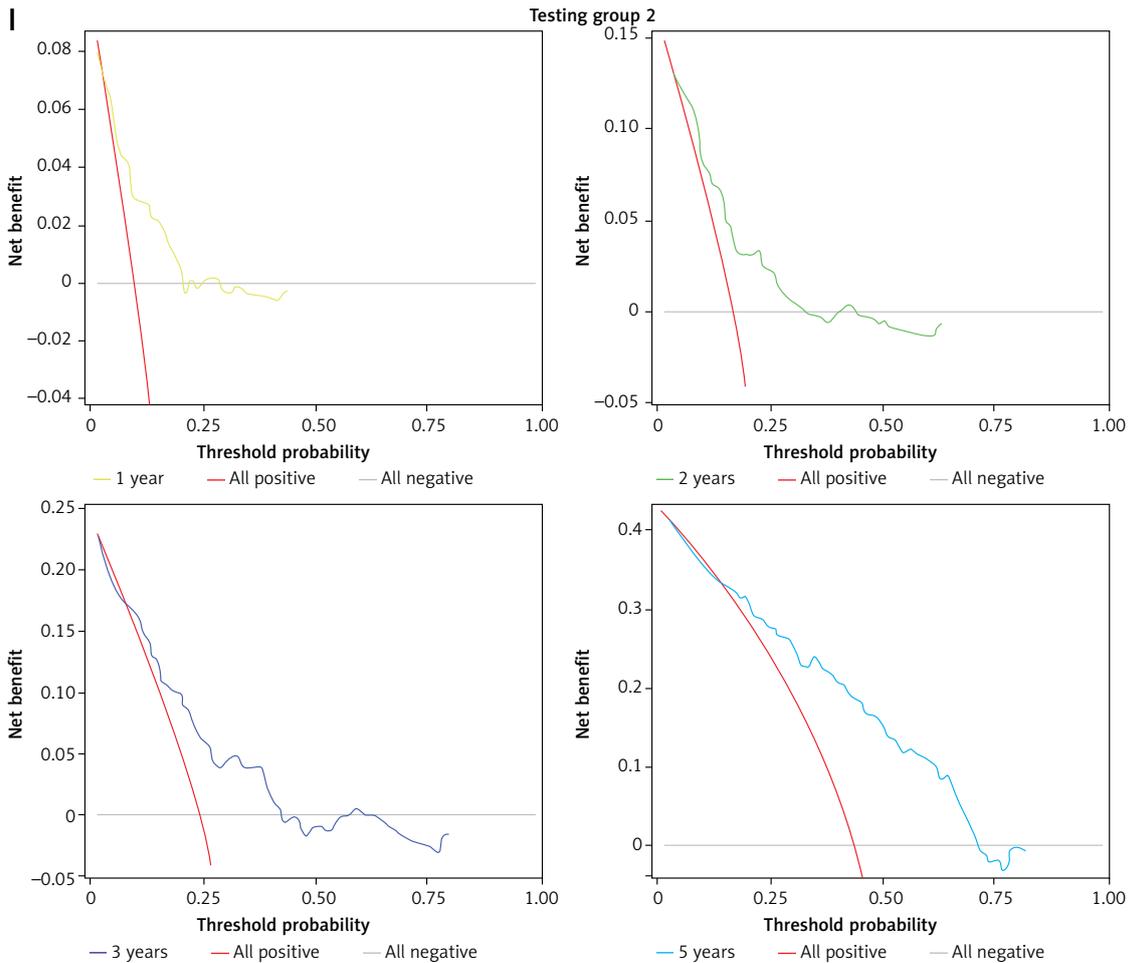


Figure 3. Cont. I – Decision curve analysis (DCA) for evaluation of the net benefits at 1, 2, 3, and 5 years. The x-axis represents the percentage of the threshold probability, and the y-axis represents the net benefits. Net benefits: when the threshold probability is exceeded, the use of column line plots to predict survival adds more benefit than all other prediction scenarios

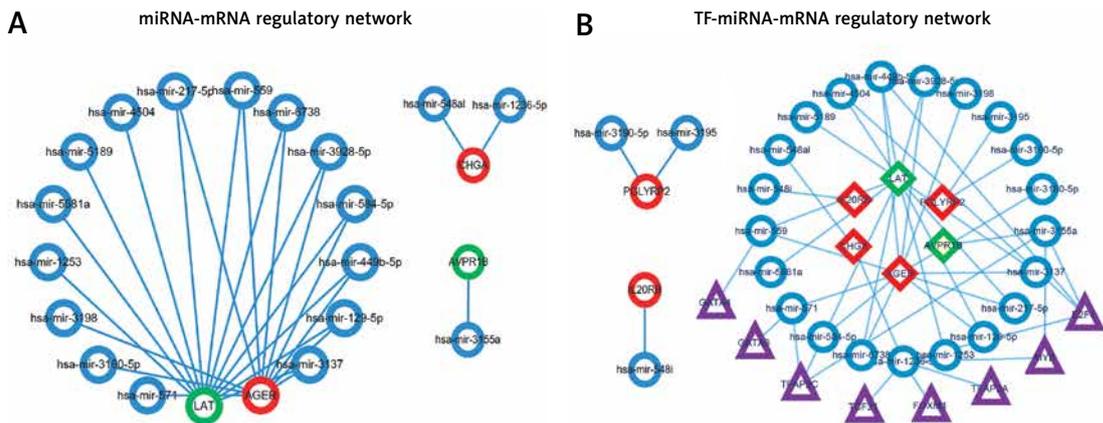


Figure 4. Interaction network of seven mRNA and reverse-predict miRNA and TF. **A** – Construct miRNA-mRNA regulatory network. Red circles represent highly expressed mRNA, green circles represent lowly expressed mRNA, and blue circles represent miRNA. **B** – Construct TF-miRNA-mRNA regulatory network. Red cubes represent highly expressed mRNAs, green cubes represent lowly expressed mRNAs, blue circles represent miRNAs, and purple triangle represents transcription factor (TF)

Immunohistochemistry (IHC) and qRT-PCR results of the hub genes

The results of immunohistochemistry showed that *AGER* was highly expressed and *LAT* was lowly expressed in ccRCC (Figure 5 A). This is consistent with the results we obtained from the TCGA database. The qRT-PCR results showed that the expression of *AGER* in ccRCC samples was significantly higher and the expression of *LAT* was significantly lower ($p < 0.01$) compared with non-cancerous samples (Figure 5 B, detailed results are provided in Supplementary Table SVII). This means that the dysregulated expression of *AGER* and *LAT* may play a role in ccRCC development.

Immune infiltration characteristics of the prognostic model

To determine whether the immune prognostic model accurately reflects the state of the tumor immune microenvironment, we analyzed the relationship between the prognostic model and immune cell infiltration. The proportions of follicular helper T cells ($p < 0.001$), CD8+ T cells ($p < 0.001$), regulatory T cells (Tregs) ($p < 0.05$), and CD4 memory activated T cells ($p < 0.01$) in the high-risk group were significantly higher than those in

the low-risk group. However, the proportions of monocytes ($p < 0.05$), activated resting mast cells ($p < 0.001$), M0 macrophages ($p < 0.001$), M2 macrophages ($p < 0.001$), neutrophils ($p < 0.01$), and activated dendritic cells ($p = 0.01$) in the high-risk group were significantly lower than those in the low-risk group (Figure 6 A). The results showed that the risk scores were significantly associated with T-cell infiltration, which was associated with poor prognostic factors (Figure 6 B). *AGER* was related to follicular helper T cells (cor = 0.22) and CD8 T cells (cor = 0.13). *LAT* was related to follicular helper T cells (cor = 0.33), CD8 T cells (cor = 0.25) and Tregs (cor = 0.12). *IL20RB* was related to Tregs (cor = 0.39), CD4 memory activated T cells (cor = 0.19) and follicular helper T cells (cor = 0.18). *PGLYRP2* was related to Tregs (cor = 0.19), T cells CD8 (cor = 0.10) and follicular helper T cells (cor = 0.13). *KLRC2* was related to CD8 T cells (cor = 0.39), follicular helper T cells (cor = 0.26) and Tregs (cor = 0.11). *AVPR1B* was related to resting dendritic cells, and *CHGA* was related to plasma cells (Figure 6 C–I).

Discussion

Several studies have constructed prognostic models for kidney cancer [20]; however, we at-

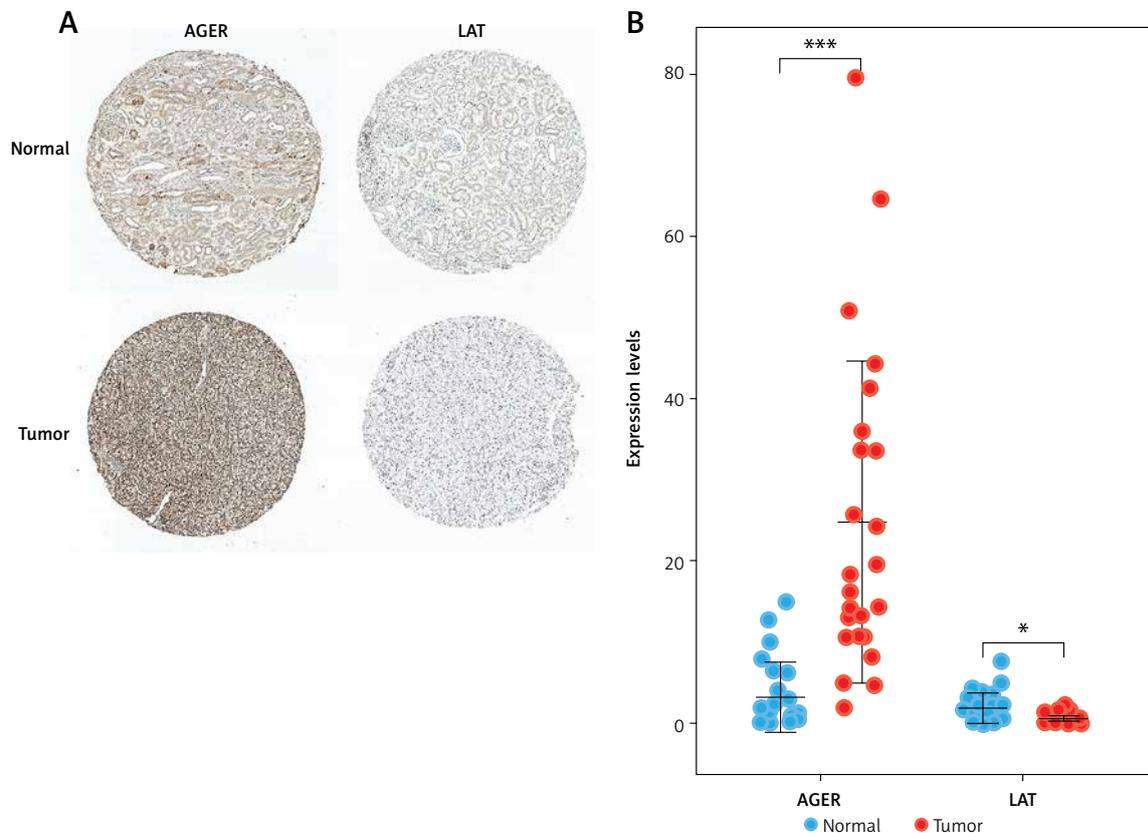


Figure 5. IHC and qRT-PCR results. **A** – IHC showed the *AGER* and *LAT* protein levels in ccRCC and adjacent nontumor tissues based on HPA database. **B** – *AGER* and *LAT* were analyzed by qRT-PCR assays in 24 ccRCC tissues and adjacent nontumor specimens from 24 patients

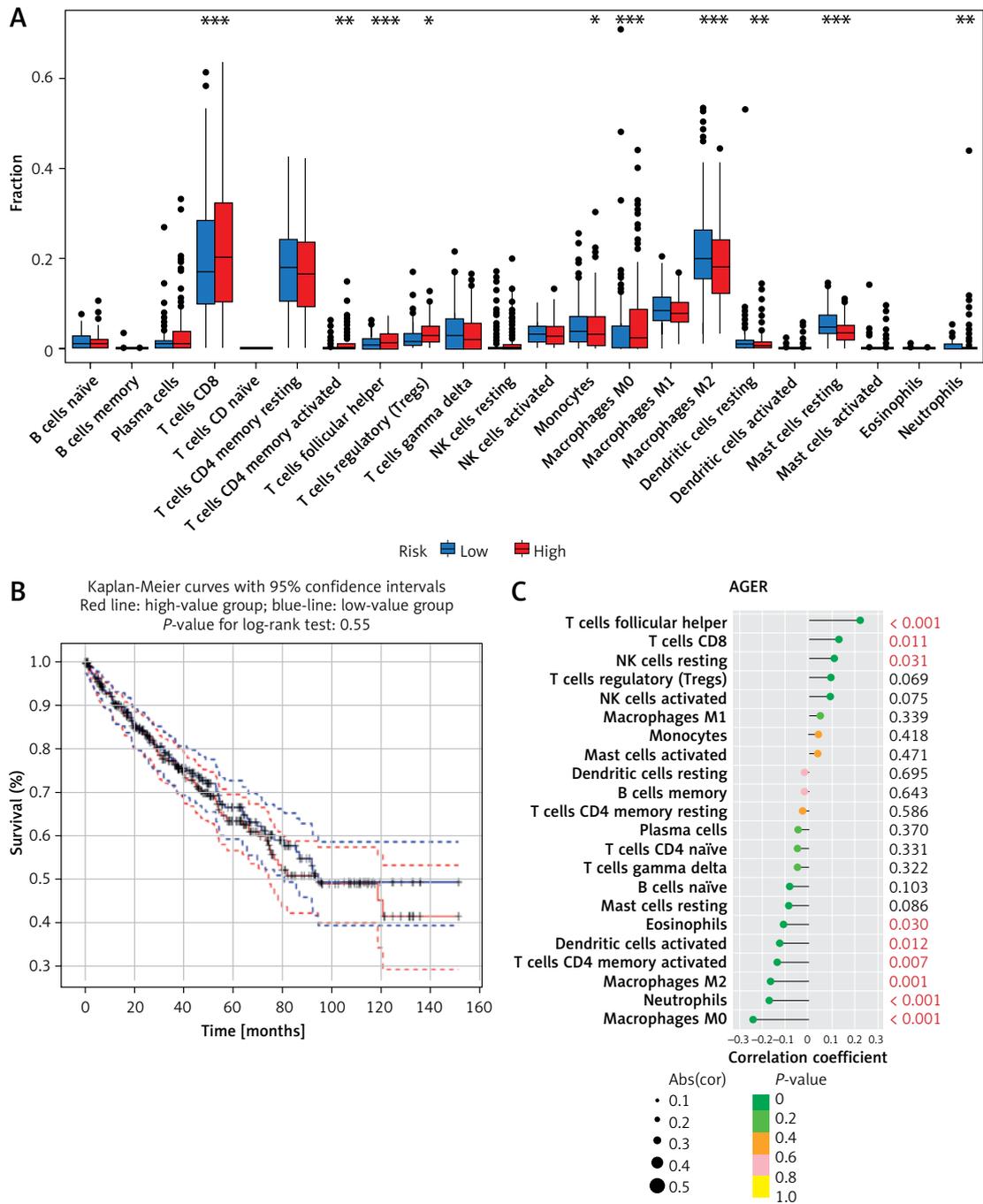


Figure 6. Relationship between immune prognosis model and immune cell infiltration. **A** – Differences in immune cell infiltration between high-risk and low-risk groups in prognostic models. **B** – T cell infiltration is highly correlated with overall survival of ccRCC patients. **C** – Relationship between seven immune-related genes (IRGs) and immune cell infiltration

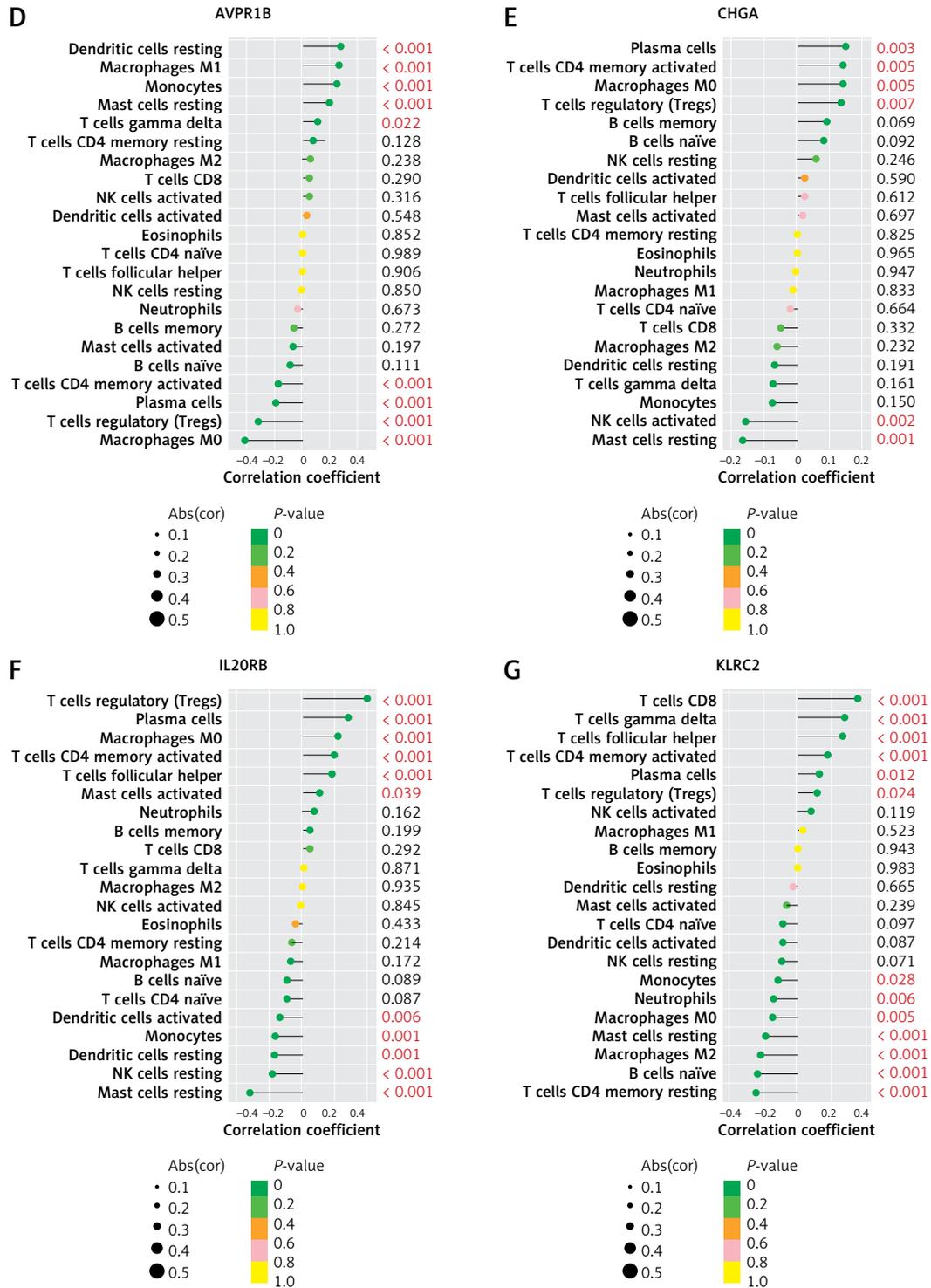


Figure 6. Cont. D–G Relationship between seven immune-related genes (IRGs) and immune cell infiltration

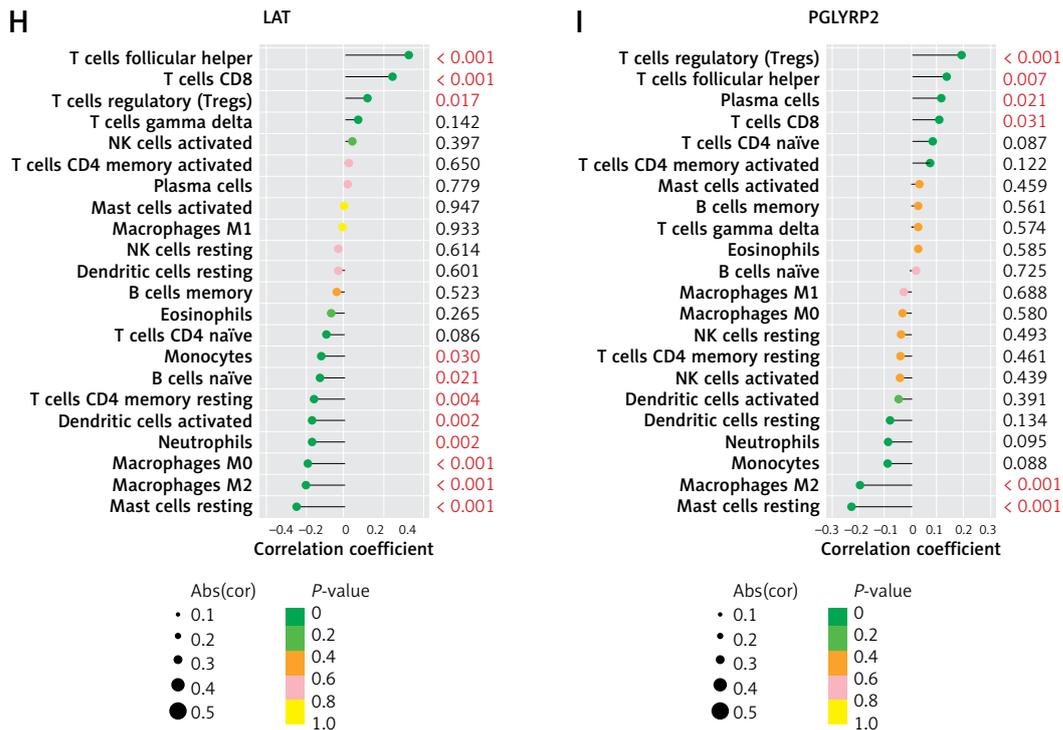


Figure 6. Cont. H-I Relationship between seven immune-related genes (IRGs) and immune cell infiltration

tempted to develop a more accurate prognostic model through using univariate and multivariate Cox regression analyses and LASSO analysis. Different immune-related genes were obtained using various screening criteria, and the overlap of the two genes (*AGER* and *LAT*) [21] again confirmed the reliability of our established model. Additionally, the prognostic model developed in this work appeared to be more accurate than previous research based on the AUC value of the ROC curves. Most of the models in previous literature were not validated, but the model described in this paper showed significant scientific validity after rigorous internal validation. As the independent prognostic analysis suggested that the model remained a strong independent prognostic predictor, it can be used as an independent prognostic indicator in the treatment of patients with ccRCC. The nomogram was plotted on the basis of the model to construct a new prognostic scoring system for renal cancer. It should be emphasized that the DCA decision curves showed that the predictive ability of the nomogram was better than that of a single independent prognostic factor, and could more sensitively predict the prognosis of patients with ccRCC.

The prognostic risk model we constructed contains 7 genes. Among these genes, *AGER* and *LAT* have been reported to act as immune prognostic factors for renal cancer [21]. This is similar to our findings. *AVPR1B* has rarely been studied in cancer and has only been found to be possibly associ-

ated with the development of Cushing's syndrome [22]. Our study suggests that *AVPR1B* is associated with a good prognosis of kidney cancer and its high expression is associated with low grades and early stages of tumors, suggesting that it may play an anti-tumor role in kidney cancer and is a reliable biomarker. High expression of *CHGA* is negatively correlated with OS in prostate cancer [23] and is an important regulator of tumor angiogenesis and tumor cell migration [24]. *IL20RB* has been reported as a prognosis-related immune gene for pancreatic cancer [25], lung cancer [26] and colorectal cancer [27]; *KLRC2* has also been predicted to play a role among many immune cell subpopulations [28, 29]. The current study also demonstrates that *IL20RB* and *KLRC2* affect the infiltration of immune T-cell subsets in renal cancer, which is expected to further explain the mechanism of immune action of *IL20RB* and *KLRC2* in renal cancer. *PGLYRP2* is thought to play an anti-tumor role in immune surveillance and immunotherapy of liver cancer [30]. However, our research shows that it may be a risk factor for ccRCC prognosis. These two contrasting mechanisms need to be further investigated.

To further explore the mechanism of this model in renal cancer, we constructed a TF-miRNA-mRNA regulatory network. Based on the node degrees, *AGER* and *LAT* are considered hub genes in the network, which suggests that they may play a key regulatory role in ccRCC. We further verified the differential expression of the key genes by IHC

and RT-PCR, and the results showed high expression of *AGER* and low expression of *LAT* in ccRCC. This is consistent with our results from the TCGA database analysis, which indicates the reliability of our study and shows the potential of *AGER* and *LAT* as immune markers for ccRCC. The results of multiple verification by TargetScan, TransmiR and miRTarBase showed that *AGER* and *LAT* may be the target genes of *miR-584-5p*. *MiR-584-5p* is a tumor suppressor that can limit the malignant progression of lung [31] and gastric cancers [32]. Our study also revealed that *miR-584-5p* improved the prognosis of patients with ccRCC (Supplementary Figure S1), suggesting that it may be able to act as silencers as immunotherapeutic targets for renal cancer. Thus, *AGER* and *LAT* may play an important role with *miR-584-5p* as a mediator, which deserves further exploration. In addition, among TFs, the oncogenic transcript *E2F1* has the highest degree of interaction with genes such as *AGER*, *CHGA*, *LAT* and *AVPR1B*. In previous research the association of significantly up-regulated *E2F1* expression in various cancers with poor prognoses was reported several times [33], which is consistent with our results (Supplementary Figures S2 A, B). This suggests that *E2F1* is expected to become an immune checkpoint for ccRCC treatment.

The results of GO and KEGG enrichment analyses showed that DEIGs were significantly enriched in immune and tumor-related signaling pathways. This is consistent with previous studies which showed that inducing immune inflammation is one of the important mechanisms of tumorigenesis [34]. At the same time, it also suggested that the prognostic model is closely related to immunity. Therefore, we further performed an immune infiltration analysis, whose results showed that T-cell infiltration (CD8 T cells, CD4 memory activated T cells, follicular helper T cells, and Tregs) was diagnostic and predictive of ccRCC. Previous studies have reported that Tfr and Tregs play an important role in tumor cells evading immune detection and establishing immune tolerance [35, 36]. Our results further demonstrate that T-cell infiltration may be an unfavorable prognostic factor for renal cancer. It may be an immune factor influencing the prognosis of ccRCC in this model. An in-depth immune infiltration analysis of the seven genes showed that *AGER* and *LAT* were related to follicular helper T cells, *IL2ORB* and *PGLYRP2* were related to Tregs, and *KLRC2* was related to T cells CD8. Therefore, *AGER*, *IL2ORB*, *KLRC2*, *LAT*, and *PGLYRP2* all have obvious T-cell infiltration characteristics, which may be the immune mechanism of their poor prognosis.

To sum up, the innovation of this study is that a new and more accurate ccRCC prognostic scoring system based on immune-related genes has

been constructed. After full internal validation, excellent consistency evaluation results were obtained. This study also attempts to explore the possible regulatory mechanisms of these possible targets for new immunotherapy molecules in the future and preliminarily verify them through molecular experiments. The results of our study could help bring new ideas to the new immunotherapy for ccRCC.

In conclusion, in this study, we established a prognostic model with good predictability and robustness. The chosen immune genes may become immune biomarkers for ccRCC. The immune characterization of the model suggests that T-cell infiltration may be highly correlated with the clinical prognosis of ccRCC. Our results may help reveal the clinical and biological implications of the ccRCC immune microenvironment.

Although this study established an immune-related prognostic model based on TCGA-KIRC mRNA expression profiles, there are still some limitations. First, differences in race, gender, age, and tumor stage of patients with ccRCC can lead to heterogeneity. Second, although our results were validated on the test set of the TCGA dataset, these results were not validated by abundant experiments *in vitro*, which implies that further experiments on a larger number of samples and clinical patients are needed.

Acknowledgments

Wei Zhuang and Zhijiao You have contributed equally to this work.

This research was funded by the Natural Science Foundation of Fujian Province (2022J01273) and Young and Middle-Aged Talent Training Project in Fujian Provincial Health System (2019-ZQNB-9).

Conflict of interest

The authors declare no conflict of interest.

References

1. Prattichizzo C, Gigante M, Pontrelli P, et al. Establishment and characterization of a highly immunogenic human renal carcinoma cell line. *Int J Oncol* 2016; 49: 457-70.
2. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; 71: 209-49.
3. Bukavina L, Bensalah K, Bray F, et al. Epidemiology of renal cell carcinoma: 2022 update. *Eur Urol* 2022; 82: 529-42.
4. Shulyak A, Banyra O. Radical or simple nephrectomy in localized renal cell carcinoma: what is a choice. *Cent European J Urol* 2011; 64: 152-5.
5. Lesnyak O, Stroy O, Banyra O, et al. Assessment of the effectiveness of radiofrequency ablation as a technique

- for destroying small renal tumors in patients older than 70. *Cent European J Urol* 2020; 73: 416-22.
6. Yao C, Zhang T, Wu T, Brugarolas J. Facts and hopes for immunotherapy in renal cell carcinoma. *Clin Cancer Res* 2022; 28: 5013-20.
 7. Manneh R, Lema M, Carril-Ajuria L, et al. Immune checkpoint inhibitor combination therapy versus sunitinib as first-line treatment for favorable-IMDC-risk advanced renal cell carcinoma patients: a meta-analysis of randomized clinical trials. *Biomedicines* 2022; 10: 577.
 8. Pi YN, Qi WC, Xia BR, Lou G, Jin WL. Long non-coding RNAs in the tumor immune microenvironment: biological properties and therapeutic potential. *Front Immunol* 2021; 12: 697083.
 9. Díaz-Montero CM, Rini BI, Finke JH. The immunology of renal cell carcinoma. *Nature reviews. Nephrology* 2020; 16: 721-35.
 10. Kunc M, Gabrych A, Dulak D, et al. Systemic inflammatory markers and serum lactate dehydrogenase predict survival in patients with Wilms tumour. *Arch Med Sci* 2022; 18: 1253-61.
 11. Yi-Wen W, Long-Long L, Ming L, Hao L, Kong-Wang H. Stem cell-like circulating tumor cells indicate poor prognosis in gastric cancer. *Arch Med Sci* 2022; 18: 1297-307.
 12. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-analyzed tumors. *Cell* 2018; 173: 530.
 13. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; 25: 25-9.
 14. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; 28: 27-30.
 15. McGeary SE, Lin KS, Shi CY, et al. The biochemical basis of microRNA targeting efficacy. *Science* 2019; 366: DOI: 10.1126/science.aav174.
 16. Tong Z, Cui Q, Wang J, Zhou Y. TransmiR v2.0: an updated transcription factor-microRNA regulation database. *Nucleic Acids Res* 2019; 47: D253-8.
 17. Chou CH, Shrestha S, Yang CD, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018; 46: D296-302.
 18. Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape automation: empowering workflow-based network analysis. *Genome Biol* 2019; 20: 185.
 19. Uhlén M, Fagerberg L, Hallström BM, et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015; 347: 1260419.
 20. Wu G, Xu Y, Zhang H, et al. A new prognostic risk model based on autophagy-related genes in kidney renal clear cell carcinoma. *Bioengineered* 2021; 12: 7805-19.
 21. Li W, Meng X, Yuan H, Xiao W, Zhang X. A novel immune-related ceRNA network and relative potential therapeutic drug prediction in ccRCC. *Front Genetics* 2021; 12: 755706.
 22. Menini S, Iacobini C, de Latouliere L, et al. The advanced glycation end-product N(ϵ)-carboxymethyllysine promotes progression of pancreatic cancer: implications for diabetes-associated risk and its prevention. *J Pathol* 2018; 245: 197-208.
 23. Giridhar KV, Sanhueza C, Hillman DW, et al. Serum chromogranin-A-based prognosis in metastatic castration-resistant prostate cancer. *Prostate Cancer Prostatic Dis* 2018; 21: 431-7.
 24. Mahata SK, Corti A. Chromogranin A and its fragments in cardiovascular, immunometabolic, and cancer regulation. *Ann N Y Acad Sci* 2019; 1455: 34-58.
 25. Haider S, Wang J, Nagano A, et al. A multi-gene signature predicts outcome in patients with pancreatic ductal adenocarcinoma. *Genome Med* 2014; 6: 105.
 26. Fan T, Pan S, Yang S, et al. Clinical significance and immunologic landscape of a five-IL(R)-based signature in lung adenocarcinoma. *Front Immunol* 2021; 12: 693062.
 27. Dai S, Xu S, Ye Y, Ding K. Identification of an immune-related gene signature to improve prognosis prediction in colorectal cancer patients. *Front Genet* 2020; 11: 607009.
 28. Ferez M, Knudson CJ, Lev A, et al. Viral infection modulates Qa-1b in infected and bystander cells to properly direct NK cell killing. *J Exp Med* 2021; 218: e20201782.
 29. Rölle A, Pollmann J, Ewen EM, et al. IL-12-producing monocytes and HLA-E control HCMV-driven NKG2C+ NK cell expansion. *J Clin Investig* 2014; 124: 5305-16.
 30. Yang Z, Feng J, Xiao L, et al. Tumor-derived peptidoglycan recognition protein 2 predicts survival and antitumor immune responses in hepatocellular carcinoma. *Hepatology* 2020; 71: 1626-42.
 31. Lee SB, Park YS, Sung JS, Lee JW, Kim B, Kim YH. Tumor suppressor miR-584-5p inhibits migration and invasion in smoking related non-small cell lung cancer cells by targeting YKT6. *Cancers* 2021; 13: 1159.
 32. Li Q, Li Z, Wei S, et al. Overexpression of miR-584-5p inhibits proliferation and induces apoptosis by targeting WW domain-containing E3 ubiquitin protein ligase 1 in gastric cancer. *J Exp Clin Cancer Res* 2017; 36: 59.
 33. Shats I, Deng M, Davidovich A, et al. Expression level is a key determinant of E2F1-mediated cell fate. *Cell Death Differentiation* 2017; 24: 626-37.
 34. Kiely M, Lord B, Ambs S. Immune response and inflammation in cancer health disparities. *Trends Cancer* 2022; 8: 316-27.
 35. Eschweiler S, Clarke J, Ramírez-Suástegui C, et al. Intratumoral follicular regulatory T cells curtail anti-PD-1 treatment efficacy. *Nature Immunol* 2021; 22: 1052-63.
 36. Watson MJ, Vignali P, Mullett SJ, et al. Metabolic support of tumour-infiltrating regulatory T cells by lactic acid. *Nature* 2021; 591: 645-51.