Am J Transl Res 2022;14(5):2825-2843
www.ajtr.org /ISSN:1943-8141/AJTR0140294

Original Article

An RNA-binding protein-related risk signature can predict the prognosis and tumor immunity of patients with testicular germ cell tumors

Yang Fu[1], Shanshan Sun[2], Jianbin Bi[1], Chuize Kong[1], Du Shi[1]

[1]*Department of Urology, The First Hospital of China Medical University, Shenyang 110001, Liaoning, China;*
[2]*Department of Pharmacy, People's Hospital Affiliated of China Medical University, Shenyang 110015, Liaoning, China*

Received November 3, 2021; Accepted March 31, 2022; Epub May 15, 2022; Published May 30, 2022

Abstract: Background: The functions of RNA-binding proteins (RBPs) in the occurrence and development of tumors remain largely unexplored. We established a risk signature based on RBPs to predict the prognosis, tumor-related immunity, and treatment benefits of patients with testicular germ cell tumors (TGCTs). Methods: A risk signature was built based on RBPs closely related to survival obtained from TGCT data in The Cancer Genome Atlas (TCGA) database. The ability of the signature to predict prognosis was analyzed by survival curves and Cox regression. The risk signature was validated using the Gene Expression Omnibus (GEO) database. The connection between tumor immunity and the risk score was evaluated. Risk score-related drug sensitivity and biofunctions were also explored. Results: A risk signature including four selected RBP genes (PARP12, USB1, POLR2E and EED) was established. The prognosis of high-risk TGCT patients was worse than that of low-risk TGCT patients. The risk score was considered a critical factor closely related to prognosis, as determined via Cox regression, and was also closely associated with multiple characteristics of tumor immunity, chemotherapy drugs and biofunctions. Conclusion: The established risk signature including four selected RBPs in TGCTs could predict the prognosis, tumor-related immunity and treatment benefits of patients with TGCTs. Utilization of this signature could help clinicians make personalized treatment decisions.

Keywords: RNA-binding proteins, TCGA, prognosis, testicular germ cell tumors

Introduction

Testicular germ cell tumors (TGCTs) are relatively rare (1% of solid tumors in men) but are considered the most common malignances in young adult men [1]. Patients with TGCTs comprised more than 95% of patients with testicular origin cancers, and TGCTs can be further divided into seminomas and nonseminomas [2, 3]. Although the cure rate of TGCTs through conventional surgical resection, radiotherapy and chemotherapy can reach over 90%, approximately 15% of patients with TGCTs are not sensitive to chemotherapy and have a poor prognosis [4-6]. Therefore, identification of other sensitive TGCT biomarkers to better predict the prognosis and treatment benefits of patients with TGCTs would be valuable.

RNA-binding proteins (RBPs) are essential proteins closely related to various RNAs [7, 8]. The main functions of RBPs are to coordinate the stability, splicing, modification, and positioning of various RNAs and to maintain cell homeostasis by participating in posttranscriptional gene regulation [9]. To date, 1542 RBPs have been identified, and recent studies have also demonstrated that the dysregulation of RBPs is closely related to tumors [10]. The overexpression of RNA-binding motif protein 3 (RBM3) can activate hepatocellular carcinoma (HCC) cell proliferation and predict a poor patient prognosis [11]. Quaking (QKI) inhibits tumor progression by regulating the alternative splicing process in lung cancer [12]. In ovarian cancer, epithelial splicing regulatory protein 1 (ESRP1) promotes tumor epithelial-mesenchymal transition (EMT)

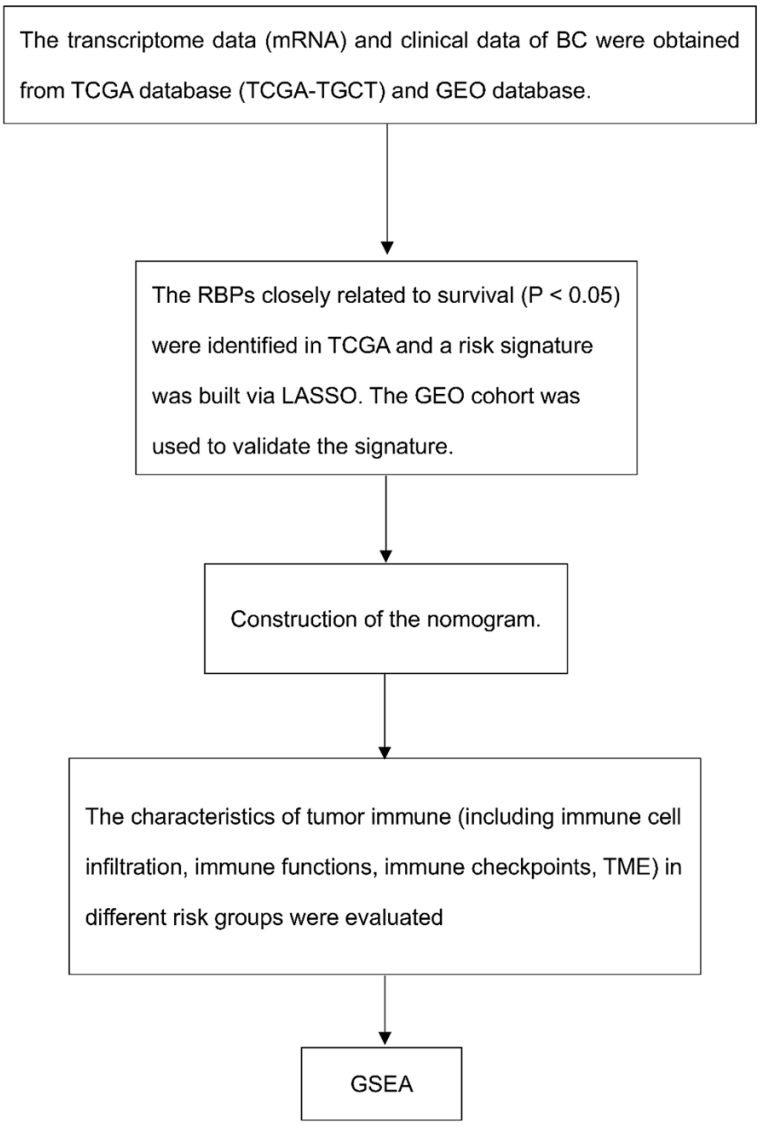


Figure 1. Flowchart of this research. TCGA, The Cancer Genome Atlas; RBPs, RNA-binding proteins; TME, tumor microenvironment; TGCT, testicular germ cell tumor; GEO, Gene Expression Omnibus; LASSO, least absolute shrinkage and selection operator; GSEA, gene set enrichment analysis.

and is related to a poor 5-year survival rate [13]. However, the functions of RBPs in the occurrence and development of tumors remain largely unexplored.

In our study, we established a risk signature based on RBPs to predict the survival and treatment benefits of patients with TGCTs from data in The Cancer Genome Atlas (TCGA) database (TCGA-TGCT) and validated this signature with the Gene Expression Omnibus (GEO) database. Furthermore, tumor-related immunity characteristics [including immune cell infiltration, immune functions, immune checkpoints, and the tumor microenvironment (TME)] in different risk groups were evaluated, and risk score-related biological functions were also explored (**Figure 1**).

Materials and methods

Patient samples included in the study

The normalized RNA expression data and clinical information data of patients with TGCTs were obtained from the official website of the TCGA database (TCGA-TGCT; http://cancergenome.nih.gov/). TCGA-TGCT included transcriptome information with survival data [the survival index was progression-free survival (PFS)] for 134 patients, and complete clinical information was available for 103 of these patients. The basic information of the cohort from TCGA is shown in **Table 1**. The GEO cohorts (GSE3218 and GSE10783) from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) were utilized for validation, and the data from 108 TGCT patients with complete clinical information [including only overall survival (OS)] and RNA sequencing information were extracted for further research. A list of RBPs was obtained from the literature [10]. To establish the signature, the TCGA cohort was used as the training set, and the GEO cohort was used as the validation set. The ethical approval was unnecessary because the data were obtained from public databases.

Establishment of a risk signature

Before establishing the risk signature, the RBPs closely related to survival ($P < 0.05$) were identified from the TCGA dataset by univariate Cox analysis. Then, the risk signature was established via least absolute shrinkage and selec-

Table 1. Characteristics of the TGCT patients obtained from the TCGA database

Basic information		TCGA (n = 103)
Age		31 (median)
Stage	I	72
	II & III	31
T classification	T1	58
	T2 & T3	45
N classification	N0	73
	N1 & N2 & N3	30
M classification	M0	95
	M1	8
Type	Seminoma	45
	Nonseminoma	58
Postoperative therapy	None	51
	Pharmaceutical	36
	Radiation	16

TGCT, testicular germ cell tumor; TCGA, the Cancer Genome Atlas.

tion operator (LASSO) Cox regression in the "glmnet" R package based on the expression data of the selected genes in TCGA. LASSO Cox regression is a regression method for high-dimensional predictive variables that can retain valuable variables, estimate parameters simultaneously and avoid overfitting [14]. This method has been widely used in survival analysis of high-dimensional data. We then calculated the risk score according to the coefficients obtained from LASSO Cox regression as follows: risk score = $\sum_{i=1}^{n}(\text{coefficient}i \times \text{expression of signature gene}i)$. According to the median risk score, patients with TGCTs were divided into high- and low-risk groups. The accuracy of the risk signature was evaluated through receiver operating characteristic (ROC) curves using the "ROC" R package and the C-index. The distribution patterns of the different risk groups were then estimated by principal component analysis (PCA). Survival curves (log-rank test) were used to compare differences in prognosis between the two risk groups. Cox regression was then performed to assess the ability of the risk score to independently predict the prognosis of patients with TGCTs. Factors that were significant in both univariable and multivariable Cox regressions ($P < 0.05$) were considered to affect the outcome of patients independently. The effect of each included gene on survival was also evaluated using Kaplan-Meier curves. A nomogram was constructed to predict the 1-, 3- and 5-year survival probabilities using the "rms" R package. Calibration curves and decision curve analysis (DCA) were performed to assess the effectiveness of the nomogram using the "rmda" R package. The GEO cohort was then used to verify the risk signature. All Cox regression analyses and the log-rank test were completed using the "survival" R package.

Risk score and tumor immunity

Twenty-nine gene markers of immune-related characteristics were identified in a previous study [15]. Single-sample gene set enrichment analysis (ssGSEA) using the "GSVA" R package was performed to calculate the enrichment level of each sample based on these gene markers and to quantify the infiltration of immune cells and immune function scores. The differences in tumor immunity (including the infiltration of immune cells, immune functions and expression of 47 common immune checkpoints) between the different risk groups were then studied. The stromal score (level of stromal cells), immune score (level of immune cells), estimation of stromal and immune cells in malignant tumor tissues using expression data (ESTIMATE) score (the stromal score plus the immune score) and tumor purity were obtained using ESTIMATE in the "estimate" R package [16, 17]. The ESTIMATE algorithm could infer the infiltration levels of stromal cells and immune cells in the tissue based on the gene expression profile of the sample (the sum of the calculation results of the two cells was defined as the tumor purity) [18]. The relationships between the risk score and the ESTIMATE results were further assessed. Pearson's test was used for the correlation analysis, and the effect of the TME on survival in the two risk groups was also evaluated.

Risk score and drug sensitivity

The relationships between the half-maximal inhibitory concentration (IC50) of six common chemotherapy drugs (bleomycin, docetaxel, cisplatin, doxorubicin, gemcitabine and paclitaxel) and the risk score were investigated using the "pRRophetic" R package. The algorithm constructed a regression model based on gene expression and drug sensitivity data in cancer

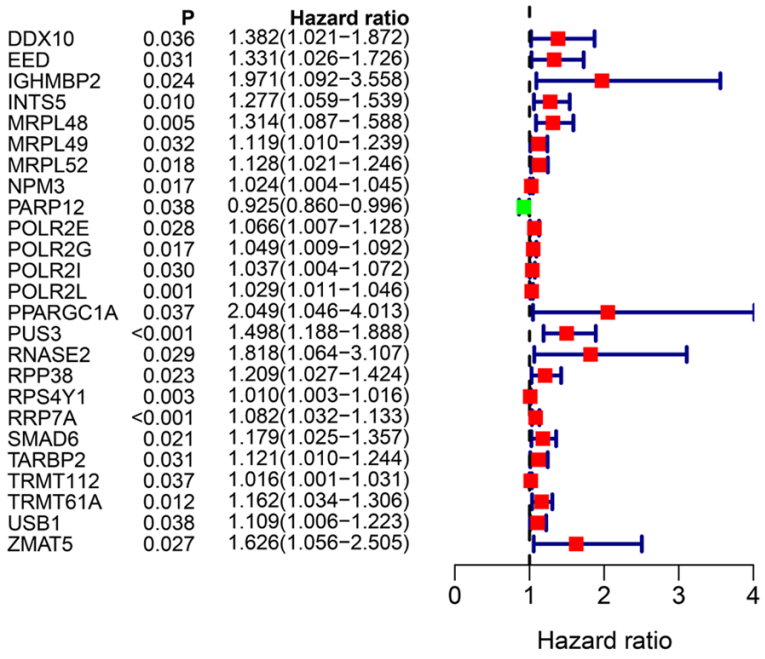


Figure 2. Univariate Cox regression. RBPs closely related to survival (P < 0.05) were identified from the TCGA dataset by univariate Cox regression. RBPs, RNA-binding proteins; TCGA, The Cancer Genome Atlas.

cell lines obtained from Genomics of Drug Sensitivity in Cancer (GDSC) (www.cancerrxgene.org/) and then applied the model to gene expression data from TCGA to evaluate drug sensitivity in vivo [19, 20].

Biofunctions associated with the risk score

Gene set enrichment analysis (GSEA) was performed to analyze the biological functions of the genes in the risk score. GSEA is one of the most commonly used methods for biological function analysis. The results were based on gene sets rather than individual genes and were thus more reliable and flexible than those obtained using traditional methods [21]. The "c5.all.v7.4.symbols.gmt" gene set for Gene Ontology (GO) analysis and the "c2.cp.kegg.v7.4.symbols.gmt" gene set for Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were downloaded from the Molecular Signatures Database (MSigDB). A nominal (NOM) *P* value < 0.05 and false discovery rate (FDR) Q value < 0.25 were considered to indicate significance.

Statistical analysis

The differences in measurement data between the patients with TGCTs in the two risk groups were detected using the Wilcoxon rank sum test. All statistical analyses were completed using R 4.03 software. We confirmed that all methods were performed in accordance with relevant guidelines and regulations.

Results

Construction of the risk signature

RBPs closely related to survival (P < 0.05) were identified from the TCGA dataset by univariate Cox regression (**Figure 2**). A risk signature was then established using LASSO Cox regression according to the selected genes; ultimately, four genes, namely, poly (ADP-ribose) polymerase family member 12 (PARP12), U6 snRNA biogenesis phosphodiesterase 1 (USB1), RNA polymerase II, I and III subunit E (POLR2E) and embryonic ectoderm development (EED), were included in the signature, and the risk score was calculated using the coefficients obtained by LASSO Cox regression (**Table 2**). We then divided the patients into high- and low-risk patient groups based on the median risk score (the cutoff value was 1.276) (**Figure 3A-C**). The areas under the curve (AUCs) were 0.768 at 1 year, 0.708 at 3 years and 0.669 at 5 years (C-index = 0.695), which indicated that the credibility of the risk signature was low to medium (**Figure 3D**). The PCA results suggested that the two groups exhibited different distribution patterns and could be clearly distinguished (**Figure 3E**). The survival curve revealed that the PFS rate of high-risk TGCT patients was lower than that of low-risk TGCT patients (P = 0.002) (**Figure 3F**). The univariable Cox regression suggested that the risk score, clinical stage and N stage were associated with PFS in TGCTs (**Figure 3G**) (all *P* values < 0.05). Multivariable Cox regression further proved that the risk score independently predicted the prognosis of patients with TGCT (P < 0.001) (**Figure 3H**). The external GEO cohort was used to validate the signature (**Figure 4A-C**), and the results yielded AUCs of 0.703 at 1 year, 0.817 at 3 years and 0.783 at

Table 2. The coefficients of included genes obtained by LASSO Cox regression

Gene	Coefficients
POLR2E	0.030
PARP12	-0.045
USB1	0.037
EED	0.199

LASSO, least absolute shrinkage and selection operator; POLR2E, RNA polymerase II, I and III subunit E; PARP12, poly (ADP-ribose) polymerase family member 12; USB1, U6 snRNA biogenesis phosphodiesterase 1; EED, embryonic ectoderm development.

5 years (C-index = 0.729) (**Figure 4D**). The PCA suggested that the two groups exhibited different distribution patterns and could be clearly distinguished (**Figure 4E**). The survival curve revealed that the OS of high-risk TGCT patients was lower than that of low-risk TGCT patients (P < 0.001) (**Figure 4F**). The results of the impact of each included gene on survival showed the same trend in the GEO and TCGA cohorts, decreased expression of PARP12 suggested a poor prognosis, and increased expression of USB1, POLR2E and EED suggested a poor prognosis (**Figure 5A-H**). The nomogram was built (**Figure 6A**), the calibration curve (**Figure 6B-D**) and the DCA (**Figure 6E**) confirmed that the nomogram could appropriately predict the survival probability of patients with TGCTs.

The risk score could indicate the characteristics of tumor immunity

ssGSEA was conducted to calculate the scores of immune cells and elucidate the immune-related functions for each sample. The results of the TCGA cohort indicated that there were no differences in dendritic cells (DCs), macrophages or immature dendritic cells (iDCs) among the different risk groups. The infiltration level of mast cells in high-risk patients was significantly increased, and the infiltration levels of other immune cells were significantly increased in low-risk patients (all *P* values < 0.05) (**Figure 7A**). The immune function scores in low-risk patients were significantly higher than those in high-risk patients (all *P* values < 0.05) (**Figure 7B**). Interestingly, the expression levels of most immune checkpoints in high-risk patients were lower than those in low-risk patients (**Figure 7C**). In the analysis of the GEO cohort, the results were roughly the same as those obtained with the TCGA dataset (**Figure 7D-F**).

The relationships between the TME (including the immune score, stromal score, ESTIMATE score and tumor purity) and risk score were assessed via ESTIMATE. The results obtained with the cohort from TCGA revealed that the risk score negatively correlated with the immune score (**Figure 8A**) and ESTIMATE score (**Figure 8B**) (all *P* values < 0.05), whereas the risk score positively correlated with tumor purity (P < 0.001) (**Figure 8C**). No significant difference was found between the stromal score and the risk score (P = 0.14) (**Figure 8D**). We found similar results with the GEO cohort (**Figure 8E-H**). The analyses of the cohorts from TCGA (**Figure 9A-D**) and GEO (**Figure 9E-H**) revealed that lower values of the immune score and ESTIMATE score suggested a poor OS, that increased tumor purity led to worse prognosis and that the stromal score was not significantly associated with survival.

Risk score and drug sensitivity

The IC50 values of six common chemotherapy drugs were predicted in the different groups. The results obtained with the cohort from TCGA revealed that bleomycin (**Figure 10A**), cisplatin (**Figure 10B**), docetaxel (**Figure 10C**), doxorubicin (**Figure 10D**), gemcitabine (**Figure 10E**) and paclitaxel (**Figure 10F**) all had higher IC50 values in low-risk patients (all *P* values < 0.05), which could indicate that high-risk patients were more sensitive to these chemotherapy drugs. We obtained similar results with the GEO cohort (**Figure 10G-L**).

Biological functions

The biological functions of the risk score were evaluated via GSEA. The most significant biofunctions enriched in high-risk patients based on GO and KEGG analyses are listed in **Tables 3** and **4**, respectively [22]. The most significant biofunctions enriched in low-risk patients based on a KEGG analysis are listed in **Table 5**, and the GO analysis identified no enriched pathways in low-risk patients.

Discussion

It is well known that an imbalance of RBPs is significantly related to the occurrence and development of tumors and can further affect patient's survival [23-25]. Unfortunately, the current research on RBPs in tumors is not com-

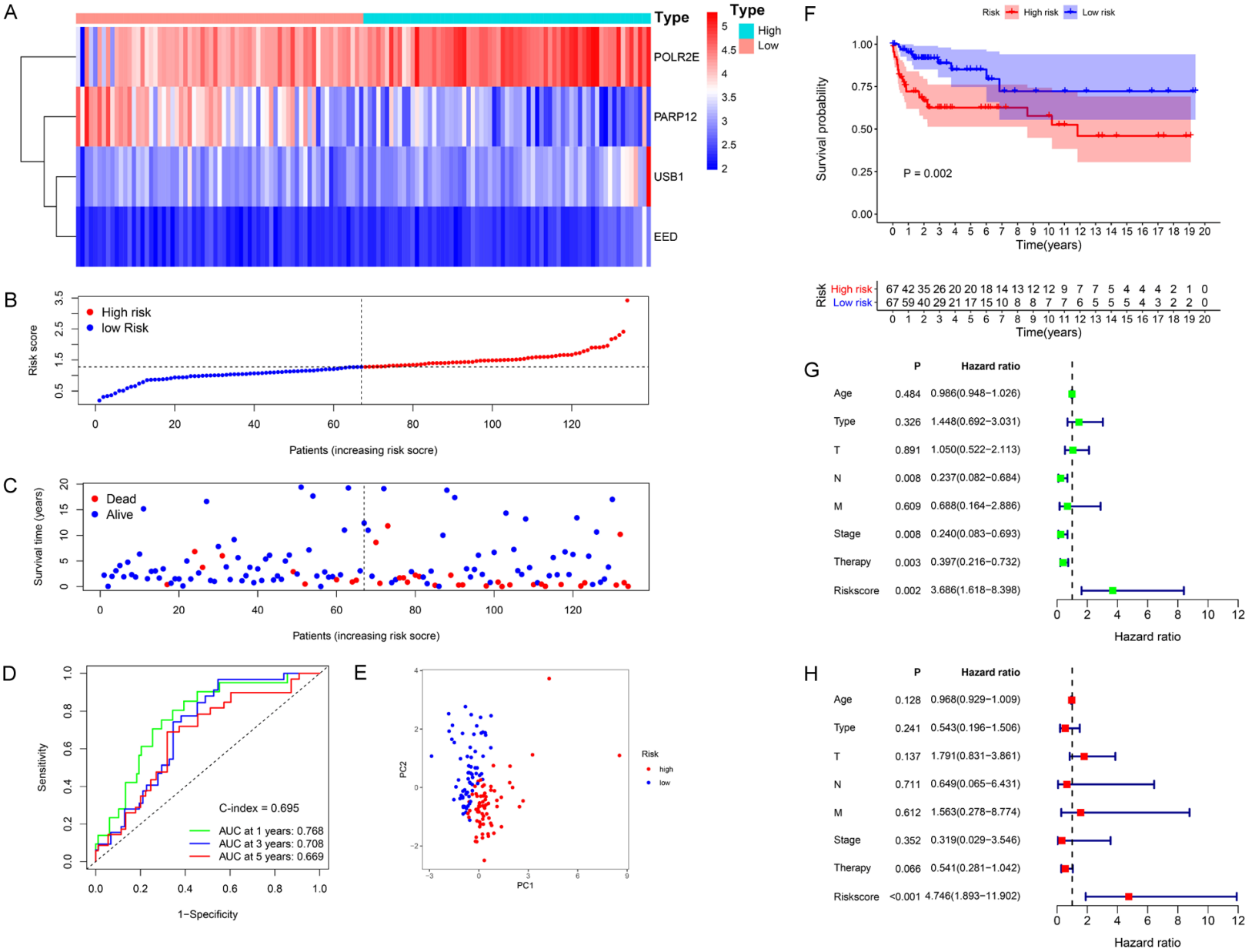

A
Type
POLR2E
PARP12
USB1
EED
Type
High
Low
B
High risk
low Risk
Risk score
Patients (increasing risk socre)
C
Dead
Alive
Survival time (years)
Patients (increasing risk socre)
D
Sensitivity
1−Specificity
C-index = 0.695
AUC at 1 years: 0.768
AUC at 3 years: 0.708
AUC at 5 years: 0.669
E
PC1
PC2
Risk
high
low
F
Risk
High risk
Low risk
Survival probability
P = 0.002
Time(years)
Risk
High risk 67 42 35 26 20 20 18 14 13 12 12 9 7 7 5 4 4 4 2 1 0
Low risk 67 59 40 29 21 17 15 10 8 8 7 7 6 5 5 5 4 3 2 2 0
Time(years)
G
P
Hazard ratio
Age 0.484 0.986(0.948−1.026)
Type 0.326 1.448(0.692−3.031)
T 0.891 1.050(0.522−2.113)
N 0.008 0.237(0.082−0.684)
M 0.609 0.688(0.164−2.886)
Stage 0.008 0.240(0.083−0.693)
Therapy 0.003 0.397(0.216−0.732)
Riskscore 0.002 3.686(1.618−8.398)
Hazard ratio
H
P
Hazard ratio
Age 0.128 0.968(0.929−1.009)
Type 0.241 0.543(0.196−1.506)
T 0.137 1.791(0.831−3.861)
N 0.711 0.649(0.065−6.431)
M 0.612 1.563(0.278−8.774)
Stage 0.352 0.319(0.029−3.546)
Therapy 0.066 0.541(0.281−1.042)
Riskscore <0.001 4.746(1.893−11.902)
Hazard ratio

Figure 3. The risk signature could independently predict a poor PFS of patients with TGCTs. Expression of selected genes in different risk groups of patients with TGCTs (A). Distribution of patients with TGCTs into different risk groups (B). Survival status of patients in different risk groups of patients with TGCTs (C). AUC based on the ROC curve (D). A PCA suggested that the two groups exhibited different distribution patterns and could be clearly distinguished (E). The survival curve suggested that the PFS of high-risk TGCT patients was lower than that of low-risk TGCT patients (F). Cox regression confirmed that the risk score was a factor that independently predicted the prognosis of patients with TGCTs. (G, H) AUC, area under the curve; TGCT, testicular germ cell tumor; ROC, receiver operating characteristic; PFS, progression-free survival; PCA, principal component analysis.

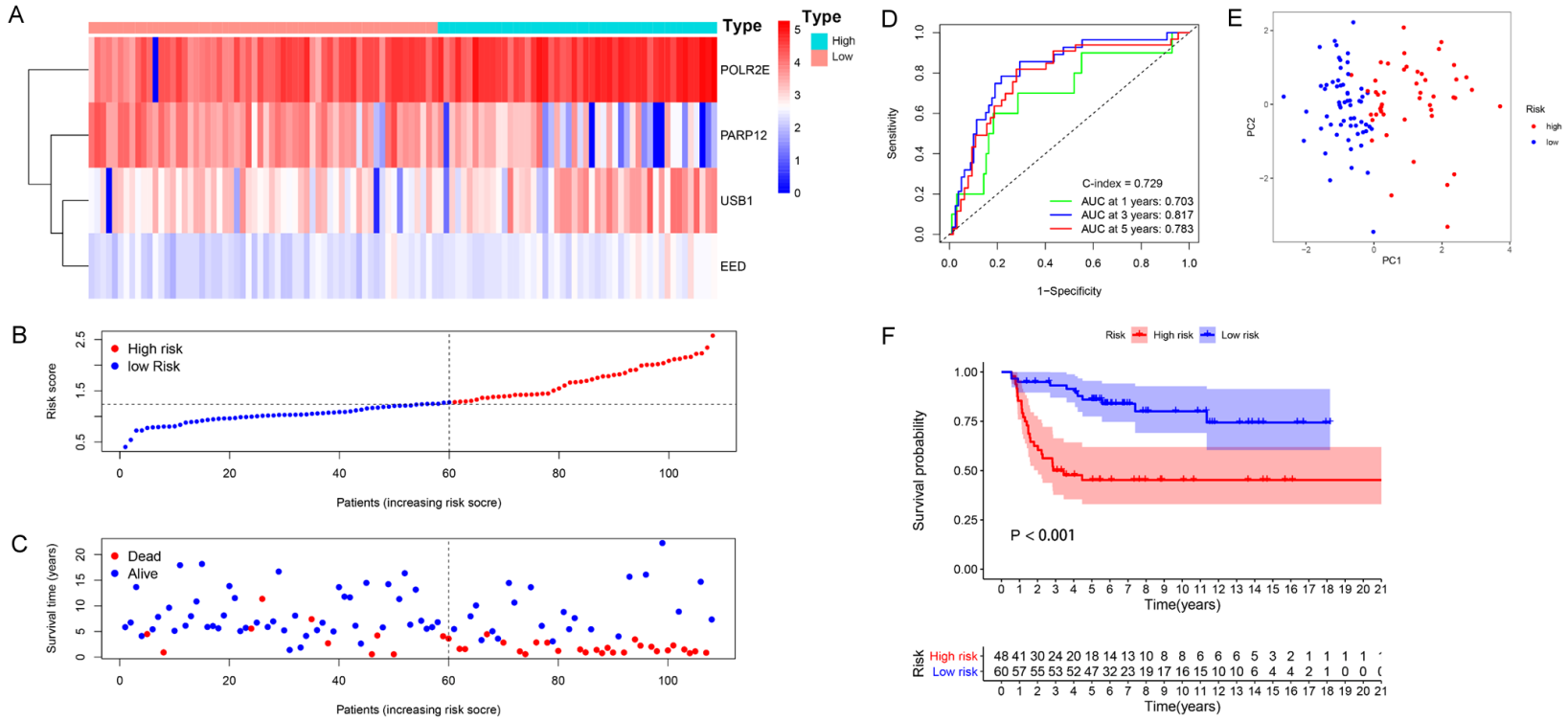


Figure 4. The risk signature was validated with the GEO database. Expression of selected genes in different risk groups of patients with TGCTs (A). Distribution of patients with TGCTs into different risk groups (B). Survival status of patients in different risk groups of patients with TGCTs (C). AUC based on the ROC curve (D). A PCA suggested that the two groups exhibited different distribution patterns and could be clearly distinguished (E). Survival curves revealed that high-risk TGCTs were significantly related to poor OS (F). AUC, area under the curve; TGCT, testicular germ cell tumor; GEO, Gene Expression Omnibus; ROC, receiver operating characteristic; PCA, principal component analysis; OS, overall survival.

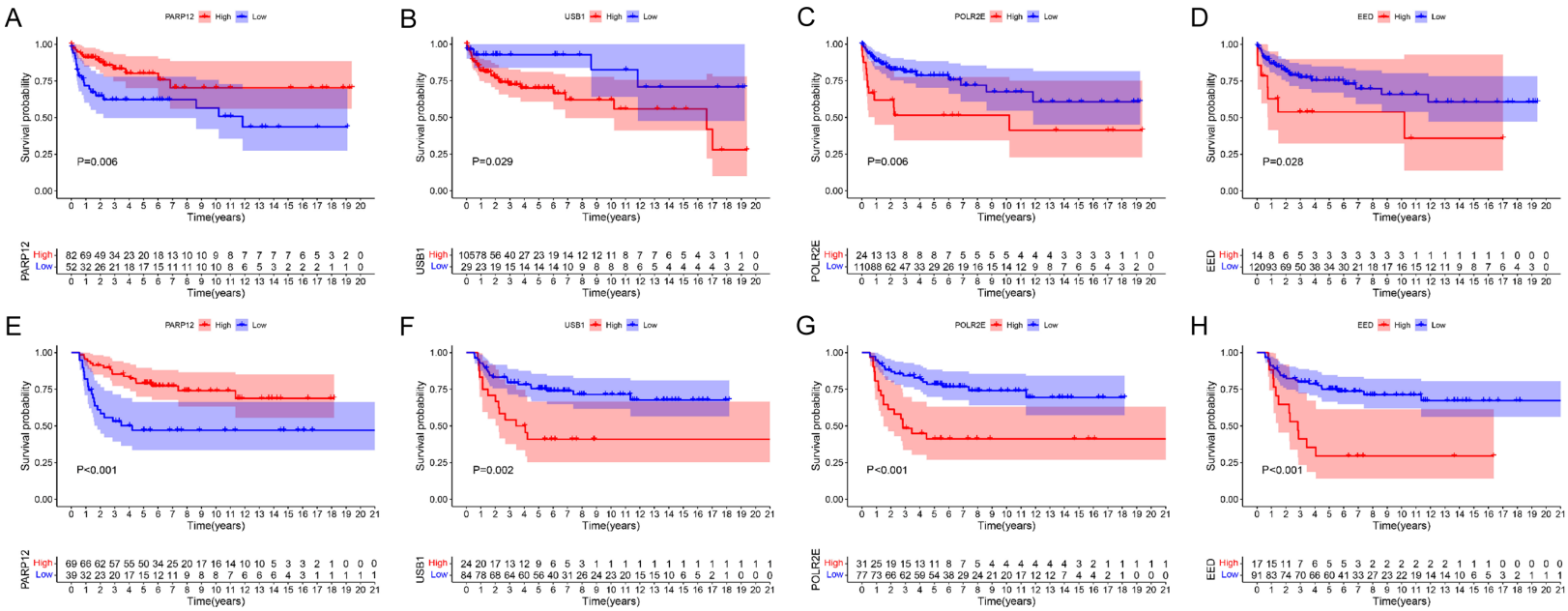


Figure 5. Impacts of each included gene on survival. In the TCGA cohort, decreased PARP12 expression suggested a poor prognosis (A). Increased expression of USB1, POLR2E and EED suggested a poor prognosis (B-D). The analysis of the GEO cohort yielded results that were roughly the same as those obtained with the TCGA cohort (E-H). TCGA, The Cancer Genome Atlas; PARP12, poly (ADP-ribose) polymerase family member 12; USB1, U6 snRNA biogenesis phosphodiesterase 1; POLR2E, RNA polymerase II, I and III subunit E; EED, embryonic ectoderm development; GEO, Gene Expression Omnibus.

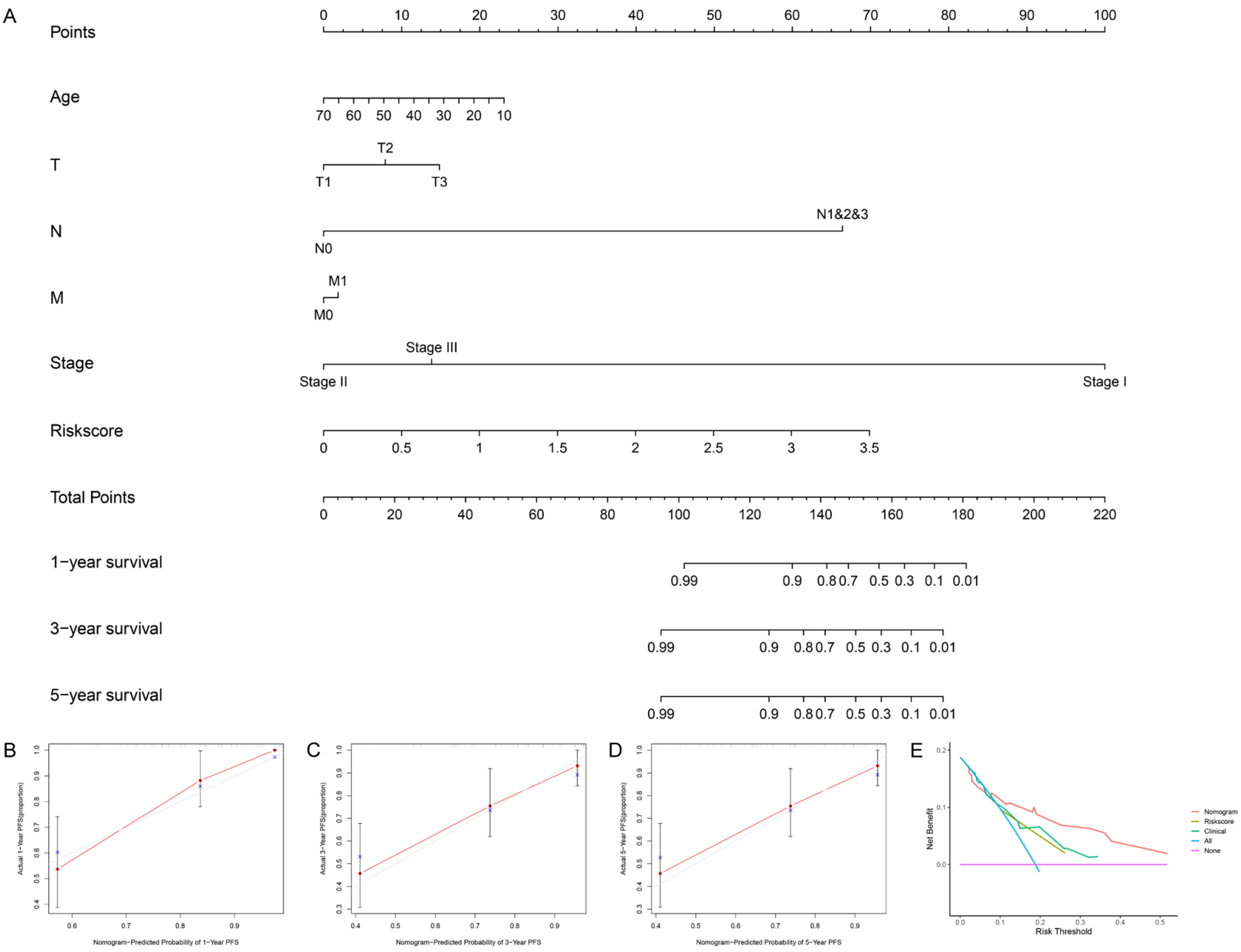

A
Points
0 10 20 30 40 50 60 70 80 90 100
Age
70 60 50 40 30 20 10
T
T2
T1 T3
N
N1&2&3
N0
M
M1
M0
Stage
Stage III
Stage II
Stage I
Riskscore
0 0.5 1 1.5 2 2.5 3 3.5
Total Points
0 20 40 60 80 100 120 140 160 180 200 220
1−year survival
0.99 0.9 0.8 0.7 0.5 0.3 0.1 0.01
3−year survival
0.99 0.9 0.8 0.7 0.5 0.3 0.1 0.01
5−year survival
0.99 0.9 0.8 0.7 0.5 0.3 0.1 0.01
B
Actual 1−Year PFS(proportion)
Nomogram−Predicted Probability of 1−Year PFS
C
Actual 3−Year PFS(proportion)
Nomogram−Predicted Probability of 3−Year PFS
D
Actual 5−Year PFS(proportion)
Nomogram−Predicted Probability of 5−Year PFS
E
Net Benefit
Risk Threshold
Nomogram
Riskscore
Clinical
All
None

Figure 6. Nomogram. A nomogram was constructed (A), and the results of the calibration curve (B-D) and DCA (E) showed that the nomogram could appropriately predict the survival probability of patients with TGCTs. DCA, decision curve analysis; TGCT, testicular germ cell tumor.

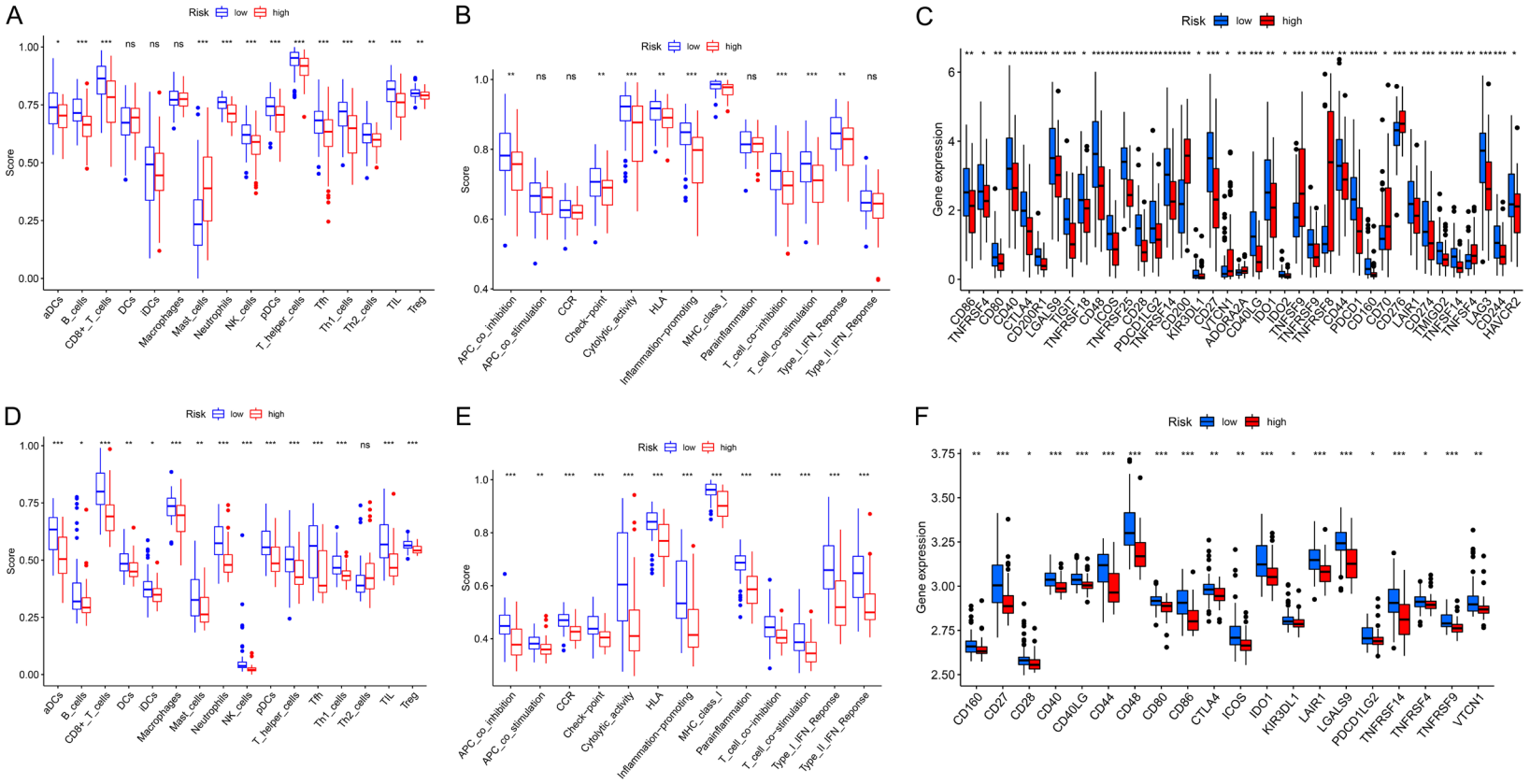


Figure 7. Risk score and tumor immunity. The results obtained with the cohort from TCGA indicated that the infiltration levels of immune cells were closely related to the risk score (A). The immune function scores of low-risk TGCT patients were significantly higher than those of high-risk TGCT patients (B). The expression levels of most immune checkpoints in high-risk TGCT patients were lower than those in low-risk TGCT patients (C). The analysis of the GEO cohort yielded results that were roughly the same as those obtained with the cohort from TCGA (D-F). GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas.

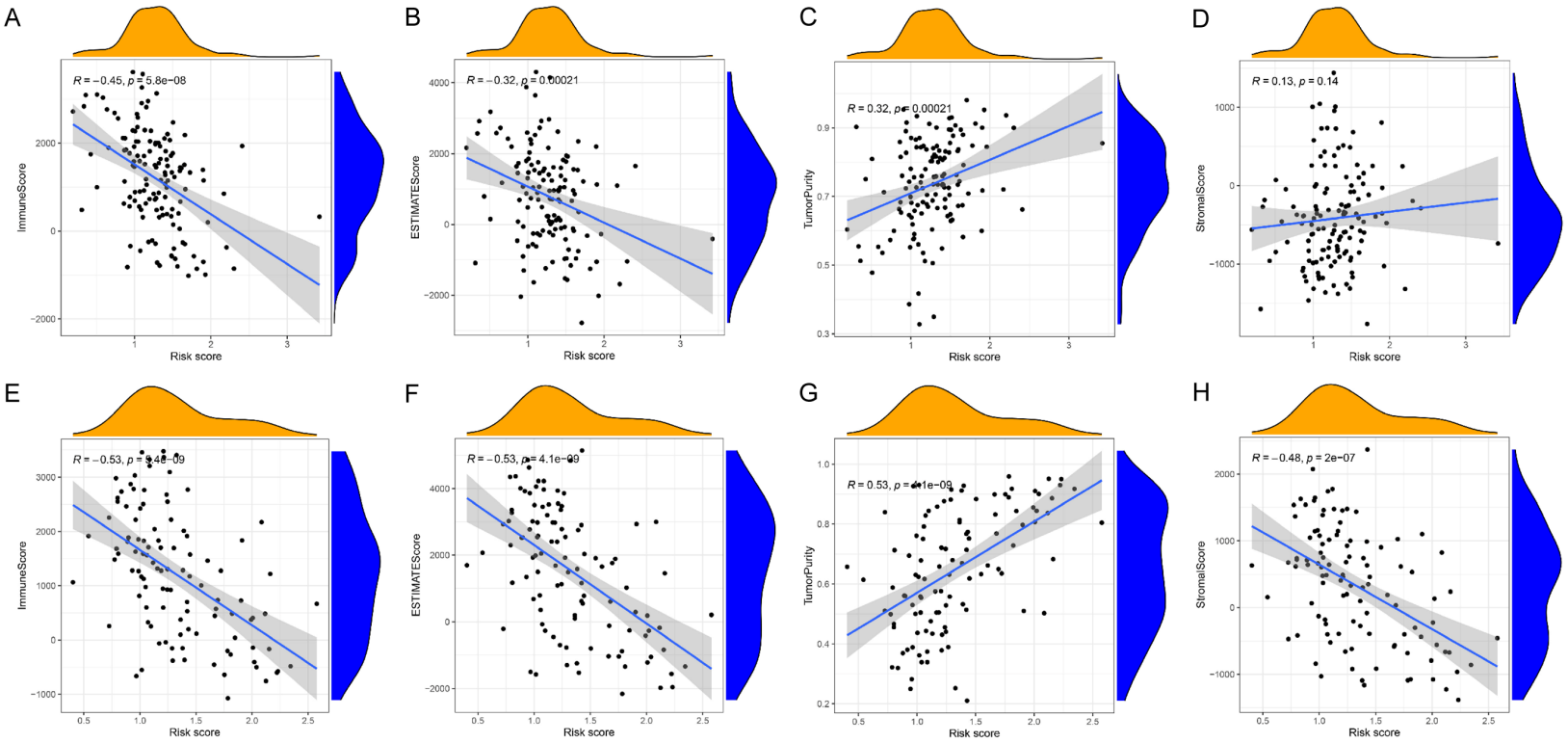


Figure 8. Risk score and TME. The results obtained with the cohort from TCGA demonstrated that the risk score negatively correlated with the immune score and ESTIMATE score, whereas the risk score positively correlated with tumor purity. No significant difference was found between the stromal score and the risk score. Similar results were obtained with the GEO cohort. TME, tumor microenvironment; TCGA, The Cancer Genome Atlas; ESTIMATE, estimation of stromal and immune cells in malignant tumor tissues using expression data; GEO, Gene Expression Omnibus.

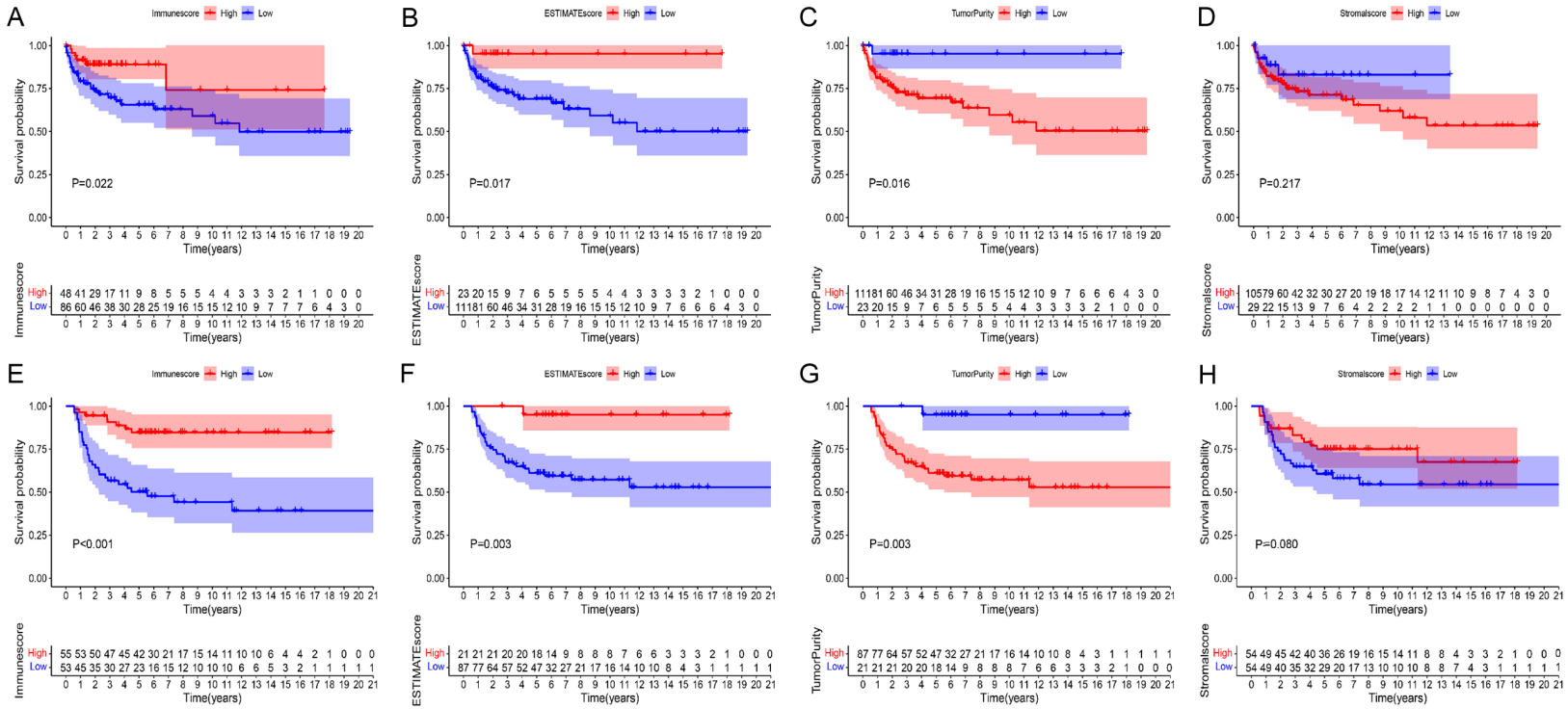
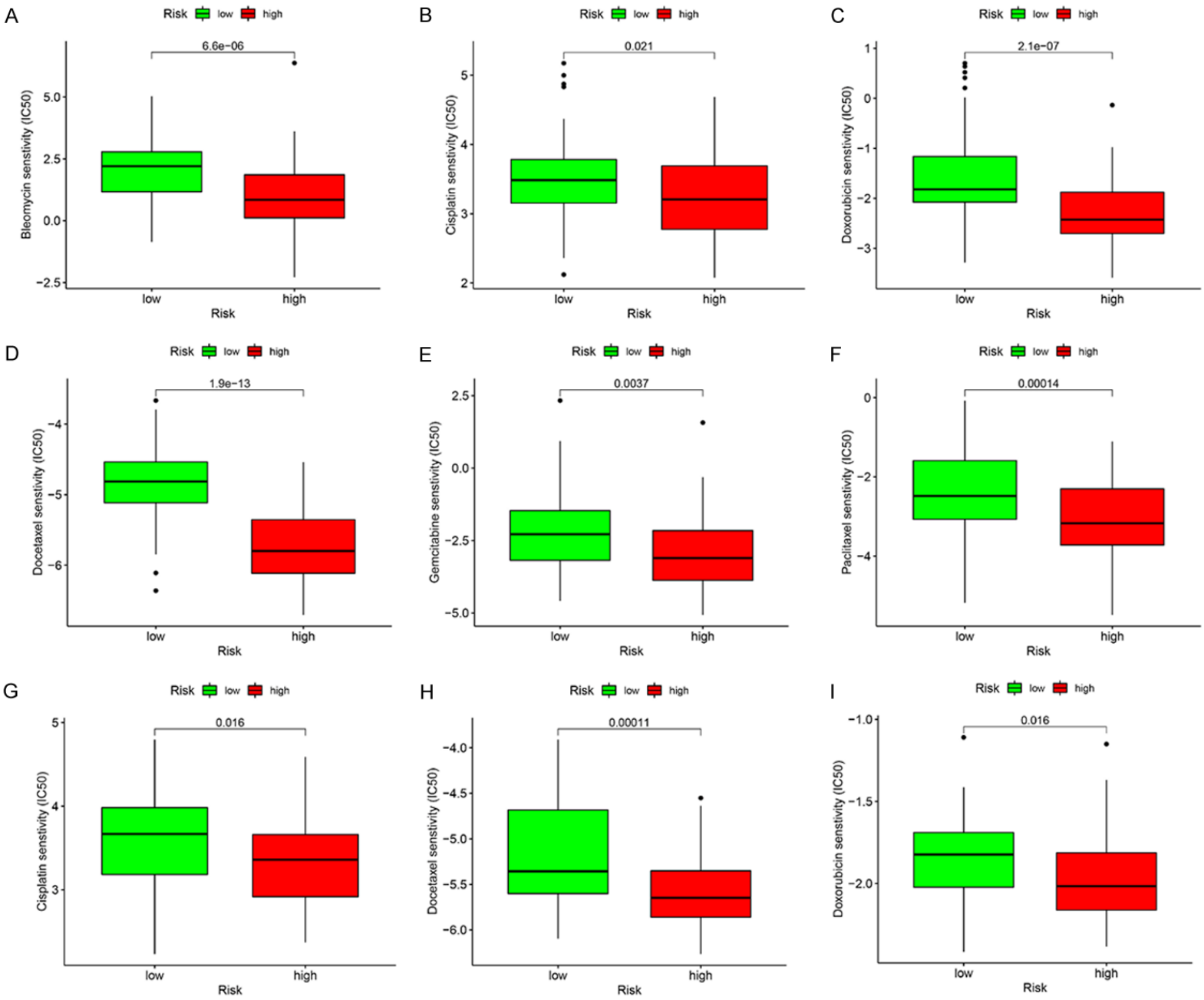


Figure 9. Impact of the TME on survival. In the cohort from TCGA, decreased values of the immune score (A) and ESTIMATE score (B) suggested poor PFS, an increased tumor purity (C) led to a worse prognosis, and the stromal score was not significantly related to survival (D). The analysis of the GEO cohort yielded results that were roughly the same as those obtained with the cohort from TCGA (E-H). TME, tumor microenvironment; TCGA, The Cancer Genome Atlas; PFS, progression-free survival; ESTIMATE, estimation of stromal and immune cells in malignant tumor tissues using expression data; GEO, Gene Expression Omnibus.


A
Risk low high
6.6e−06
Bleomycin senstivity (IC50)
5.0
2.5
0.0
−2.5
low
high
Risk
B
Risk low high
0.021
Cisplatin senstivity (IC50)
5
4
3
2
low
high
Risk
C
Risk low high
2.1e−07
Doxorubicin senstivity (IC50)
1
0
−1
−2
−3
low
high
Risk
D
Risk low high
1.9e−13
Docetaxel senstivity (IC50)
−4
−5
−6
low
high
Risk
E
Risk low high
0.0037
Gemcitabine senstivity (IC50)
2.5
0.0
−2.5
−5.0
low
high
Risk
F
Risk low high
0.00014
Paclitaxel senstivity (IC50)
0
−2
−4
low
high
Risk
G
Risk low high
0.016
Cisplatin senstivity (IC50)
5
4
3
low
high
Risk
H
Risk low high
0.00011
Docetaxel senstivity (IC50)
−4.0
−4.5
−5.0
−5.5
−6.0
low
high
Risk
I
Risk low high
0.016
Doxorubicin senstivity (IC50)
−1.0
−1.5
−2.0
low
high
Risk

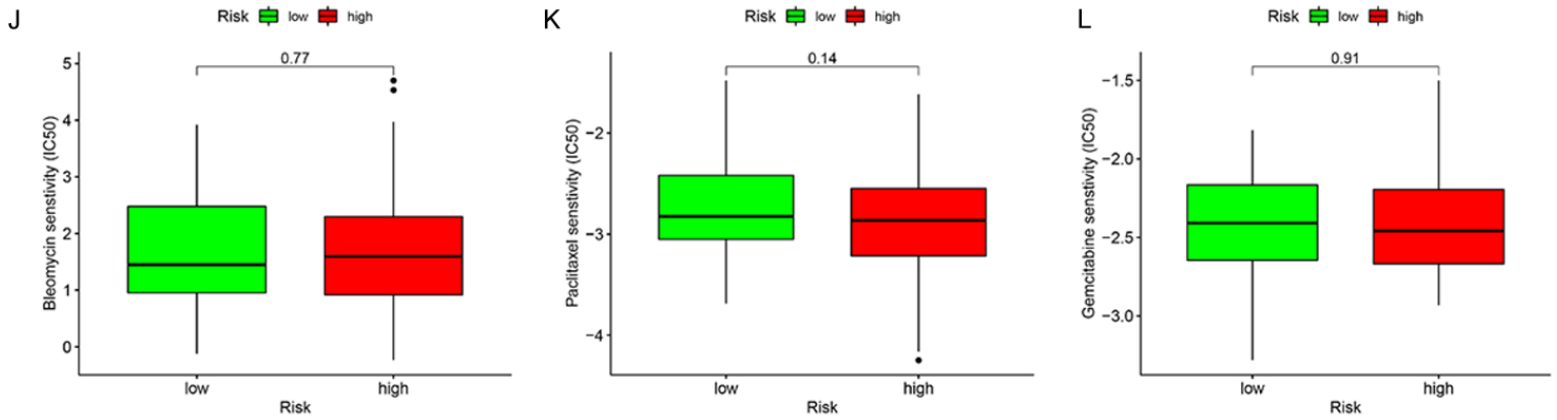


Figure 10. Risk score and drug sensitivity. The results obtained with the cohort from TCGA revealed that bleomycin (A), cisplatin (B), docetaxel (C), doxorubicin (D), gemcitabine (E) and paclitaxel (F) all had higher IC50 values in low-risk TGCT patients, which suggested that high-risk patients were more sensitive to these chemotherapy drugs. We found similar results with the GEO cohort (G-L). TCGA, The Cancer Genome Atlas; IC50, half-maximal inhibitory concentration; GEO, Gene Expression Omnibus.

Table 3. Gene sets enriched in the high risk phenotype via GO

Gene set name	NES	NOM *p*-val	FDR *q*-val
GOBP_FIBROBLAST_GROWTH_FACTOR_RECEPTOR_SIGNALING_PATHWAY	1.957	0.000	0.151
GOBP_SOMATIC_STEM_CELL_POPULATION_MAINTENANCE	1.901	0.000	0.109
GOMF_CELL_CELL_ADHESION_MEDIATOR_ACTIVITY	1.810	0.000	0.196
GOBP_GLUCOSE_6_PHOSPHATE_METABOLIC_PROCESS	1.804	0.010	0.198
GOBP_CHONDROCYTE_PROLIFERATION	1.783	0.004	0.220
GOMF_CELL_ADHESION_MEDIATOR_ACTIVITY	1.726	0.004	0.242
GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION	1.701	0.008	0.239
GOBP_GLUCOSE_CATABOLIC_PROCESS	1.699	0.010	0.238
GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION	1.687	0.012	0.244
GOBP_POSITIVE_REGULATION_OF_WNT_SIGNALING_PATHWAY	1.649	0.004	0.249
GOBP_GLUCOSE_METABOLIC_PROCESS	1.647	0.000	0.248
GOBP_FIBROBLAST_GROWTH_FACTOR_RECEPTOR_SIGNALING_PATHWAY	1.957	0.000	0.151

GO, Gene Ontology; NES, Normalized enrichment score; NOM, Nominal; FDR, False discovery rate. Gene sets with NOM *p*-val < 0.05 and FDR *q*-val < 0.25 were considered significant.

Table 4. Gene sets enriched in the high risk phenotype via KEGG

Gene set name	NES	NOM *p*-val	FDR *q*-val
KEGG_GALACTOSE_METABOLISM	1.702	0.008	0.132
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM	1.697	0.011	0.124
KEGG_GLYCOLYSIS_GLUCONEOGENESIS	1.677	0.016	0.131
KEGG_CELL_CYCLE	1.616	0.038	0.146
KEGG_TGF_BETA_SIGNALING_PATHWAY	1.609	0.036	0.145
KEGG_HUNTINGTONS_DISEASE	1.573	0.036	0.149
KEGG_PYRIMIDINE_METABOLISM	1.549	0.038	0.166
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_LACTO_AND_NEOLACTO_SERIES	1.507	0.042	0.197
KEGG_INSULIN_SIGNALING_PATHWAY	1.396	0.046	0.256
KEGG_GALACTOSE_METABOLISM	1.702	0.008	0.132
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM	1.697	0.011	0.124
KEGG_GLYCOLYSIS_GLUCONEOGENESIS	1.677	0.016	0.131

KEGG, Kyoto Encyclopedia of Genes and Genomes; NES, Normalized enrichment score; NOM, Nominal; FDR, False discovery rate. Gene sets with NOM *p*-val < 0.05 and FDR *q*-val < 0.25 were considered significant.

Table 5. Gene sets enriched in the low risk phenotype via KEGG

Gene set name	NES	NOM *p*-val	FDR *q*-val
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	-1.607	0.066	0.272
KEGG_JAK_STAT_SIGNALING_PATHWAY	-1.602	0.045	0.224
KEGG_PRIMARY_IMMUNODEFICIENCY	-1.584	0.068	0.213
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY	-1.552	0.053	0.230
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION	-1.465	0.140	0.237
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY	-1.440	0.116	0.234
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	-1.382	0.171	0.247
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	-1.378	0.158	0.240
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	-1.607	0.066	0.272

KEGG, Kyoto Encyclopedia of Genes and Genomes; NES, Normalized enrichment score; NOM, Nominal; FDR, False discovery rate. Gene sets with NOM *p*-val < 0.05 and FDR *q*-val < 0.25 were considered significant.

prehensive. In this study, we established a risk signature based on RBPs to predict the survival and treatment benefits of patients with TGCTs. A total of 4 genes (PARP12, USB1, POLR2E and EED) were included in the construction of the risk signature. The log-rank test and Cox analysis confirmed that the risk score could be used as a factor to independently predict the prognosis of patients with TGCTs. Previous studies have revealed the connections between these 4 genes and tumors. PARP12 deficiency accelerated HCC cell migration and invasion via regulation of the EMT process [26]. USB1 inhibits thyroid cancer cell proliferation by inducing cell cycle arrest [27]. The overexpression of POLR2E significantly reduced the survival rate of patients with acute myeloid leukemia [28]. Increased expression of EED was associated with advanced clinical characteristics and worse disease-free survival (DFS) in patients with colorectal cancer [29]. These genes included in the gene signature deserve further in-depth study as potential targets in TGCT therapy. We also validated the signature using the GEO cohort. In addition, a nomogram composed of various clinical features and the risk score was built to predict the PFS of patients with TGCTs. Calibration curves were used to estimate the effectiveness of the nomogram. The DCA suggested that the net benefit of the nomogram was greater than that of the risk score and the clinical characteristics alone. Therefore, the clinical applicability and robustness of the signature were both satisfactory.

The relationships between various tumor immune-related parameters and the risk score were analyzed in our study. We found that tumor-related immunity features (including various effector immune cells and immune functions) were significantly activated in low-risk patients. Disorder of the immune system was recently confirmed as a vital process of tumorigenesis, and immunotherapy has also become an emerging treatment method for tumors, including TGCTs [30, 31]. The infiltration of T cells could improve the prognosis of patients with TGCTs [32]. Interestingly, the expression of immune checkpoints, which cause immune escape to suppress the immune response, was significantly increased in low-risk patients in both the GEO and TCGA cohorts [33]. According to literature, the immune checkpoint inhibitors exhibit significant effects in some tumors, and the frequent expression of PD-1 can be observed in TGCT tissues [34, 35]. Anti-PD-1 therapy has been administered to patients with TGCTs who were not sensitive to radiotherapy and chemotherapy [36]. However, we should also acknowledge that the effect of immunotherapy varies greatly among patients, and some patients do not respond to this type of treatment [37, 38]. Our study revealed that the immune response and immune checkpoint levels were increased in low-risk TGCT patients with a better survival rate, which indicated that the immune response enhancement effect in low-risk patients was greater than the effect of immune checkpoints on immune response inhibition; thus, the application of immune checkpoint inhibitors to low-risk TGCT patients could further activate the immune response and exert better anticancer effects.

The stromal score (level of stromal cells), immune score (level of immune cells), ESTIMATE score (stromal score plus immune score) and tumor purity were calculated via ESTIMATE. In the cohort from TCGA, increases in the risk score were associated with a decrease in the level of immune cell infiltration and a shorter PFS, but no significant difference in the stromal infiltration level was found. Moreover, the increase in tumor purity caused by a decrease in the level of immune cell infiltration also led to poor survival. These results were consistent with the ssGSEA results. Similar results were obtained with the GEO cohort.

TGCTs characteristically show sensitivity to chemotherapy drugs, we therefore assessed whether the risk score reflected drug sensitivity [39]. The results demonstrated that high-risk patients were more sensitive to 6 common chemotherapy drugs. This suggested that the administration of adjuvant chemotherapy to high-risk TGCT patients and that of immune checkpoints to low-risk TGCT patients could achieve more significant clinical effects. The results from the KEGG analysis of the low-risk group indicated that a variety of immune-related pathways were enriched, which suggested that the immune functions of low-risk TGCT patients were significantly enhanced, and this finding was consistent with the results from the analysis of immune parameters. The results from the KEGG and GO analyses revealed that some pathways related to glucose metabolism were significantly enriched in high-risk TGCT patients. The enhancement of glucose metabo-

lism could provide the energy needed for the biological behavior of tumors (cell division and metastasis); thus, glucose metabolism is considered to be closely related to the pernicious phenotype [40]. The product of glucose metabolism, lactic acid, can strongly inhibit the function of natural killer (NK) cells and T cells and thereby suppress the immune response [41]. These results showed that targeting glucose metabolism might also serve as a new direction for the treatment of patients with TGCTs.

Nevertheless, our study had some limitations. First, although we performed a systematic bioinformatics analysis of the RBP-related signature of TGCTs, these results still need to be confirmed by further basic experiments and clinical analyses in the future. Second, the clinical information from TCGA and GEO data was not specific, TCGA data did not include details of patients receiving systemic treatment, and the GEO data only contained survival data, which might affect the effect of the signature. Third, the histology of our research was not strictly differentiated, and TGCT seminomas and nonseminomas were pooled; we thus look forward to improving this analysis in the future.

Conclusion

A risk signature including four selected RBPs in TGCTs was constructed and could predict the prognosis, tumor-related immunity characteristics and treatment benefits of patients with TGCTs. The use of this signature could help clinicians make personalized treatment decisions.

Acknowledgements

We thank TCGA and GEO for sharing large amounts of data. This work was supported by the Project of Liaoning Distinguished Professor [Grant No. [2012]145], the Shenyang Plan Project of Science and Technology [Grant No. F17-230-9-08], China Medical University's 2017 Discipline Promotion Program [Grant No. 3110117040], China Medical University's 2018 Discipline Promotion Program, and the 2017 National Key R&D Program Key Projects of Precision Medical Research [No. 2017YFC-0908000].

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Du Shi, Department of Urology, The First Hospital of China Medical University, No. 155 Nanjing North Street, Heping District, Shenyang 110001, Liaoning, China. E-mail: shidu_cmu@163.com

References

[1] Tsili AC, Sofikitis N, Stiliara E and Argyropoulou MI. MRI of testicular malignancies. Abdom Radiol (NY) 2019; 44: 1070-1082.
[2] Bosl GJ and Motzer RJ. Testicular germ-cell cancer. N Engl J Med 1997; 337: 242-253.
[3] Moch H, Cubilla AL, Humphrey PA, Reuter VE and Ulbright TM. The 2016 WHO classification of tumours of the urinary system and male genital organs - Part A: renal, penile, and testicular tumours. Eur Urol 2016; 70: 93-105.
[4] Einhorn LH. Curing metastatic testicular cancer. Proc Natl Acad Sci U S A 2002; 99: 4592-4595.
[5] Diamantopoulos N and Kortsaris A. Testicular germ cell tumors. J Buon 2010; 15: 421-434.
[6] Cong R, Ji C, Zhang J, Zhang Q, Zhou X, Yao L, Luan J, Meng X and Song N. m6A RNA methylation regulators play an important role in the prognosis of patients with testicular germ cell tumor. Transl Androl Urol 2021; 10: 662-679.
[7] Hentze MW, Castello A, Schwarzl T and Preiss T. A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol 2018; 19: 327-341.
[8] Perron G, Jandaghi P, Solanki S, Safisamghabadi M, Storoz C, Karimzadeh M, Papadakis AI, Arseneault M, Scelo G, Banks RE, Tost J, Lathrop M, Tanguay S, Brazma A, Huang S, Brimo F, Najafabadi HS and Riazalhosseini Y. A general framework for interrogation of mRNA stability programs identifies RNA-Binding proteins that govern cancer transcriptomes. Cell Rep 2018; 23: 1639-1650.
[9] Pereira B, Billaud M and Almeida R. RNA-binding proteins in cancer: old players and new actors. Trends Cancer 2017; 3: 506-528.
[10] Gerstberger S, Hafner M and Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet 2014; 15: 829-845.
[11] Dong W, Dai ZH, Liu FC, Guo XG, Ge CM, Ding J, Liu H and Yang F. The RNA-binding protein RBM3 promotes cell proliferation in hepatocellular carcinoma by regulating circular RNA SCD-circRNA 2 production. EBioMedicine 2019; 45: 155-167.
[12] Zong FY, Fu X, Wei WJ, Luo YG, Heiner M, Cao LJ, Fang Z, Fang R, Lu D, Ji H and Hui J. The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. PLoS Genet 2014; 10: e1004289.
[13] Jeong HM, Han J, Lee SH, Park HJ, Lee HJ, Choi JS, Lee YM, Choi YL, Shin YK and Kwon

MJ. ESRP1 is overexpressed in ovarian cancer and promotes switching from mesenchymal to epithelial phenotype in ovarian cancer cells. Oncogenesis 2017; 6: e389.
[14] Chen W, Ou M, Tang D, Dai Y and Du W. Identification and validation of immune-related gene prognostic signature for hepatocellular carcinoma. J Immunol Res 2020; 2020: 5494858.
[15] Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, Fröhling S, Chan EM, Sos ML, Michel K, Mermel C, Silver SJ, Weir BA, Reiling JH, Sheng Q, Gupta PB, Wadlow RC, Le H, Hoersch S, Wittner BS, Ramaswamy S, Livingston DM, Sabatini DM, Meyerson M, Thomas RK, Lander ES, Mesirov JP, Root DE, Gilliland DG, Jacks T and Hahn WC. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 2009; 462: 108-112.
[16] Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612.
[17] Li B, Geng R, Wu Q, Yang Q, Sun S, Zhu S, Xu Z and Sun S. Alterations in immune-related genes as potential marker of prognosis in breast cancer. Front Oncol 2020; 10: 333.
[18] Liu Y, Wu J, Huang W, Weng S, Wang B, Chen Y and Wang H. Development and validation of a hypoxia-immune-based microenvironment gene signature for risk stratification in gastric cancer. J Transl Med 2020; 18: 201.
[19] Song C, Guo Z, Yu D, Wang Y, Wang Q, Dong Z and Hu W. A prognostic nomogram combining immune-related gene signature and clinical factors predicts survival in patients with lung adenocarcinoma. Front Oncol 2020; 10: 1300.
[20] Geeleher P, Cox N and Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014; 9: e107468.
[21] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[22] Kanehisa M and Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.
[23] Neelamraju Y, Gonzalez-Perez A, Bhat-Nakshatri P, Nakshatri H and Janga SC. Mutational landscape of RNA-binding proteins in human cancers. RNA Biol 2018; 15: 115-129.
[24] Wurth L. Versatility of RNA-binding proteins in cancer. Comp Funct Genomics 2012; 2012: 178525.
[25] Conrad T, Albrecht AS, de Melo Costa VR, Sauer S, Meierhofer D and Ørom UA. Serial interactome capture of the human cell nucleus. Nat Commun 2016; 7: 11212.
[26] Shao C, Qiu Y, Liu J, Feng H, Shen S, Saiyin H, Yu W, Wei Y, Yu L, Su W and Wu J. PARP12 (ARTD12) suppresses hepatocellular carcinoma metastasis through interacting with FHL2 and regulating its stability. Cell Death Dis 2018; 9: 856.
[27] Ma Y, Yin S, Liu XF, Hu J, Cai N, Zhang XB, Fu L, Cao XC and Yu Y. Comprehensive analysis of the functions and prognostic value of rna-binding proteins in thyroid cancer. Front Oncol 2021; 11: 625007.
[28] Wei J, Xie Q, Liu X, Wan C, Wu W, Fang K, Yao Y, Cheng P, Deng D and Liu Z. Identification the prognostic value of glutathione peroxidases expression levels in acute myeloid leukemia. Ann Transl Med 2020; 8: 678.
[29] Liu YL, Gao X, Jiang Y, Zhang G, Sun ZC, Cui BB and Yang YM. Expression and clinicopathological significance of EED, SUZ12 and EZH2 mRNA in colorectal cancer. J Cancer Res Clin Oncol 2015; 141: 661-669.
[30] Sharma P and Allison JP. The future of immune checkpoint therapy. Science 2015; 348: 56-61.
[31] Chovanec M, Mardiak J and Mego M. Immune mechanisms and possible immune therapy in testicular germ cell tumours. Andrology 2019; 7: 479-486.
[32] Siska PJ, Johnpulle RAN, Zhou A, Bordeaux J, Kim JY, Dabbas B, Dakappagari N, Rathmell JC, Rathmell WK, Morgans AK, Balko JM and Johnson DB. Deep exploration of the immune infiltrate and outcome prediction in testicular cancer by quantitative multiplexed immunohistochemistry and gene expression profiling. Oncoimmunology 2017; 6: e1305535.
[33] Roufas C, Chasiotis D, Makris A, Efstathiades C, Dimopoulos C and Zaravinos A. The expression and prognostic impact of immune cytolytic activity-related markers in human malignancies: a comprehensive meta-analysis. Front Oncol 2018; 8: 27.
[34] Wolchok JD. PD-1 blockers. Cell 2015; 162: 937.
[35] Fankhauser CD, Curioni-Fontecedro A, Allmann V, Beyer J, Tischler V, Sulser T, Moch H and Bode PK. Frequent PD-L1 expression in testicular germ cell tumors. Br J Cancer 2015; 113: 411-413.
[36] Zschäbitz S, Lasitschka F, Hadaschik B, Hofheinz RD, Jentsch-Ullrich K, Grüner M, Jäger D and Grüllich C. Response to anti-programmed cell death protein-1 antibodies in

men treated for platinum refractory germ cell cancer relapsed after high-dose chemotherapy and stem cell transplantation. Eur J Cancer 2017; 76: 1-7.

[37] Shah S, Ward JE, Bao R, Hall CR, Brockstein BE and Luke JJ. Clinical response of a patient to anti-PD-1 immunotherapy and the immune landscape of testicular germ cell tumors. Cancer Immunol Res 2016; 4: 903-909.

[38] Adra N, Abonour R, Althouse SK, Albany C, Hanna NH and Einhorn LH. High-dose chemotherapy and autologous peripheral-blood stem-cell transplantation for relapsed metastatic germ cell tumors: the Indiana University experience. J Clin Oncol 2017; 35: 1096-1102.

[39] Rossini E, Bosatta V, Abate A, Fragni M, Salvi V, Basnet RM, Zizioli D, Bosisio D, Piovani G, Valcamonico F, Mirabella G, Berruti A, Memo M and Sigala S. Cisplatin cytotoxicity in human testicular germ cell tumor cell lines is enhanced by the CDK4/6 inhibitor palbociclib. Clin Genitourin Cancer 2021; 19: 316-324.

[40] Tuo Z, Zheng X, Zong Y, Li J, Zou C, Lv Y and Liu J. HK3 is correlated with immune infiltrates and predicts response to immunotherapy in non-small cell lung cancer. Clin Transl Med 2020; 10: 319-330.

[41] San-Millán I and Brooks GA. Reexamining cancer metabolism: lactate production for carcinogenesis could be the purpose and explanation of the Warburg Effect. Carcinogenesis 2017; 38: 119-133.