

81670\_Auto\_Edited.docx

# Machine learning algorithm to construct cuproptosis- and immune-related prognosis prediction model for colon cancer

Huang YY *et al.* Machine learning algorithm for colon cancer

Yuan-Yi Huang, Ting-Yu Bao, Xu-Qi Huang, Qi-Wen Lan, Ze-Min Huang, Yu-Han Chen, Zhi-De Hu, Xu-Guang Guo

## Abstract

### BACKGROUND

Over the past few years, research into the pathogenesis of colon cancer has progressed rapidly, and cuproptosis is an emerging mode of cellular apoptosis. Exploring the relationship between colon cancer and cuproptosis benefits in identifying novel biomarkers and even improving the outcome of the disease.

### AIM

To look at the prognostic relationship between colon cancer and the genes associated with cuproptosis and immune system in patients. The main purpose was to assess whether reasonable induction of these biomarkers reduces mortality among patients with colon tumors.

### METHOD

Data obtained from <sup>19</sup>The Cancer Genome Atlas and Gene Expression Omnibus were used in differential analysis to explore differential expression genes associated with cuproptosis and immune activation. The <sup>8</sup>Least Absolute Shrinkage and Selection Operator Cox regression algorithm was applied to build a cuproptosis- and immune-related combination model, and the model was utilized for principal component analysis and survival analysis to observe the survival and prognosis of the patients. A series of statistically meaningful transcriptional analysis results demonstrated an intrinsic relationship between cuproptosis and the micro-environment of colon cancer.

## RESULTS

Once prognostic characteristics were obtained, [the](#) CDKN2A and DLAT genes related to cuproptosis were strongly linked to colon cancer: The first was a risk factor, whereas the second was a protective factor. The finding of the validation analysis showed that the comprehensive model associated with cuproptosis and immune was statistically significant. Within the component expressions, the expressions of HSPA1A, CDKN2A, and UCN3 differed markedly. Transcription analysis primarily reflects the differential activation of related immune cells and pathways. Furthermore, genes linked to immune checkpoints that were expressed differently between the subgroups revealed the mechanism of tumor immune evasion.

**Key Words:** Cuproptosis; Immune; Colon cancer; Prognosis models; Immune infiltration analysis

Huang YY, Bao TY, Huang XQ, Lan QW, Huang ZM, Chen YH, Hu ZD, Guo XG. Machine learning algorithm to construct cuproptosis- and immune-related prognosis prediction model for colon cancer. *World J Gastrointest Oncol* 2023; In press

**Core Tip:** We comprehensively <sup>30</sup> analyzed the effect of cuproptosis and immunity on the prognosis of colon cancer based on the close association between the three. Scrupulously, a variety of algorithms were applied to analyze and construct a three-gene prognostic model whose efficacy was statistically significant <sup>1</sup> in both the training set and the external validation set. Moreover, in order to provide help for prognosis assessment and personalized treatment of colon cancer, we performed immunoinfiltration analysis and immunocheckpoint inhibitor-related genes expression analysis in different risk groups according to the cuproptosis- and immune-related combination model.

## INTRODUCTION

Colon cancer is among the most serious cancers worldwide and is genetically complex<sup>[1,2]</sup>. According to the World Health Organization (WHO), over 940000 cases are reported worldwide, and almost 500000 people die annually of colon cancer<sup>[3]</sup>. As with other solid cancers, colon cancer is infested with various types of immune cells. Among these, factors favouring tumors are the inflammatory reaction of B cells, innate immune cells and numerous T cell subtypes. Unlike other epithelial cancers, many inflammatory T cells that infiltrate colon cancer are specific for the commensal microbiota, rather than specific tumor antigens or autoantigens. Therefore, they do not directly kill cancer cells. However, natural killer (NK) cells, CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells are still the main immune monitors<sup>[4,5]</sup>. At the same time, CD11b<sup>+</sup>Gr1<sup>+</sup> cells recruited in the late stage contribute to the suppression of the antitumor immune response and tumor angiogenesis, which share characteristics with monocytes, macrophages, neutrophils and dendritic cells<sup>[5]</sup>. The tumor microenvironment (TME) formed in the tumor development process consists mostly of stromas and immune cells<sup>[6,7]</sup>. Changes in their composition and proportion will affect the immunosuppressive effect of tumors and subsequently affect pathological state, treatment and prognosis. At present, further research on immune biomarkers is required to provide evidence for immune checkpoint blockade therapy in colon cancer, which is more appropriate for the majority of colon cancers<sup>[6-8]</sup>.

Copper, the micronutrient essential to human health, is an essential cofactor of biological function. However, it becomes toxic when its concentration exceeds the threshold maintained by evolutionary conservative homeostatic mechanisms<sup>[9]</sup>. Cuproptosis occurs when copper accumulates. Too much copper binds to the lipid components of the tricarboxylic acid (TCA) cycle, resulting in the accumulation of lipoprotein and the loss of iron-sulfur tuft, eventually leading to protein-toxic stress and cell death<sup>[10]</sup>. Studies have shown that cuproptosis may have a high cytotoxic effect on tumor cells. Abnormal copper levels can affect tumor growth by causing irreversible damage to organelles and consequently inducing tumor cell death through multiple

mechanisms, which could be used as new targets for anticancer therapy<sup>[11-13]</sup>. Meanwhile, to adapt to adverse microenvironments, several copper chaperones are upregulated in cancer cells, which bind to cytoplasmic copper and transfer to SOD1, supporting antioxidant function and indirectly mediating cancer cell metastasis<sup>[14]</sup>. In the relationship between excessive copper and immunity, excessive copper can induce splenomegaly, disorder of the white medulla in the spleen, expansion of the red medulla and accumulation of dead cells, decrease the number of thymocytes, inhibit the proliferation of lymphocytes in other tissues and induce their necrosis and apoptosis<sup>[15]</sup>. At the same time, serum Cu and ferroxidase ceruloplasmin were positively correlated with inflammation<sup>[16]</sup>.

To further study the effect of cuproptosis and immunity on colon cancer, we built a prognostic model based on the currently known genes, which is likely to provide guidance for the treatment of colon cancer.

## **MATERIALS AND METHODS**

*The idea of this study is shown in Figure 1 in the form of a flow chart*

**Data collection and handling:** Clinical information and RNA-seq data of 443 patients with colon cancer and 41 normal controls were downloaded from The Cancer Genome Atlas (TCGA) (<https://portal.gdc.com>), and we obtained differentially expressed genes (DEGs) with the "Limma" package of R software. For external validation purpose, we downloaded the GSE39582 dataset containing 568 colon cancer samples from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). Information about the training set and validation set is shown in Table 1.

*Screening cuproptosis- and immune-related differentially expressed genes independently associated with prognosis*

We collected ten cuproptosis-related genes (CRGs) from existing reports<sup>[10]</sup>, and the differentially expressed cuproptosis-related genes (CR-DEGs) were obtained by intersecting them with the DEGs of colon cancer. Through the GeneCards database

(<https://www.genecards.org/>), 1793 immune-related genes (IRGs) with  $\text{score} \geq 7$  were obtained<sup>[17]</sup> and intersected with the DEGs with  $|\log_2\text{FC}| > 1$  and  $\text{p.adjust} < 0.05$  to obtain immune-related differentially expressed genes (IR-DEGs). The "Survival with Forestplot" package in R was used to perform univariate and multivariate Cox regression analyses on CR-DEGs and IR-DEGs, respectively, to screen out candidate signatures with significant independent predictive values for the prognosis of colon cancer ( $P < 0.05$ ).

### *Construction and external validation of prognostic risk score models*

We utilized the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression algorithm to construct three risk scoring models with candidate signatures using 10-fold cross-validation. This method was implemented with the "glmnet" package<sup>[18,19]</sup>. On the basis of the median risk score, the samples were divided into two groups: The high-risk group was greater than the median, and the low-risk group was less than the median. Then, we applied the "survminer" package to determine the optimal cutoff value for survival analysis and to compare overall survival (OS) in different subgroups using the Kaplan-Meier method and the log-rank test. The "timeROC" package was applied for Receiver Operating Characteristic (ROC) analysis to evaluate the performance of the model. Finally, we used the GSE39582 dataset for external validation of the integrated cuproptosis- and immune-related model.

### *Principal Component Analysis and Construction of Nomograms*

We performed Principal Component Analysis (PCA) on colon cancer patients based on the cuproptosis- and immune-related model and used the "ggplot2" package for visualization. According to the results of multivariate Cox proportional hazards analysis, nomograms including risk score and other clinical predictors were built using the "rms" package to predict 1-, 3-, and 5-year OS<sup>[20]</sup>.

### *Functional enrichment and immune profiling*

Based on the cuproptosis- and immune-related model, we carried out Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA) analyses of DEGs between the high- and low-risk groups. To explore the immune profiles of the high- and low-risk groups, we applied the CIBERSORT and EPIC algorithms to calculate the abundance of infiltrating immune cells in separate subgroups. We compared the expression of genes associated with immune checkpoints between the subgroups and displayed the results with great significance ( $P < 0.05$ ).

## RESULTS

### *Acquisition of cuproptosis-related and immune-related differentially expressed genes*

A heatmap and volcano map present the results of gene differential expression analysis (Figure 2A and B). Ten CR-DEGs were obtained by intersecting the CRGs and DEGs (Figure 2C), suggesting that these genes may be involved in the development of colon cancer. Similarly, 299 IR-DEGs were obtained (Figure 2D).

### *Identification of prognostic features and construction of risk scoring models*

Univariate analysis of ten CR-DEGs showed that two genes were significant ( $P < 0.05$ ). Among them, CDKN2A was a risk factor ( $HR > 1$ ), while DLAT was a protective factor ( $HR < 1$ ) (Figure 2E). Multivariate analysis showed that both were still significant (Figure 2F), indicating that they may be independent factors affecting the OS of colon cancer. LASSO-Cox regression analysis screened the coefficients of the predictors (Figure 3A), and a colon cancer prognostic risk score model related to copper death was constructed. The model is a formula, and the optimal model is taken when  $\lambda_{\min} = 0.002$ : Riskscore =  $(0.1582) \times CDKN2A + (-0.3203) \times DLAT$ .

Analogous results were obtained in the analysis of 299 IR-DEGs. We showed 20 IR-DEGs associated with prognosis at  $P < 0.05$ , 4 of which had independent effects, except for UCN3 as a protective factor ( $HR < 1$ ). HSPA1A, ULBP2 and FABP4 were all risk factors ( $HR > 1$ ). (Figure 3B and C) When  $\lambda_{\min} = 0.0017$  (Figure 3D), the best immune-

related colon cancer prognostic risk score model was as follows: Riskscore =  $(0.2507)*HSPA1A + (0.1563)*ULBP2 + (0.1065)*FABP4 + (-0.1704)$ .

After multivariate analysis of CDKN2A, DLAT, UCN3, HSPA1A, ULBP2 and FABP4, the results showed that CDKN2A and HSPA1A may be independent risk factors for colon cancer prognosis ( $HR > 1$ ), while UCN3 is an independent protective factor ( $HR < 1$ ) (Figure 3E and F). After coefficient screening, when  $\lambda_{min} = 0.0044$ , the optimal cuproptosis- and immune-related prognostic model was constructed, namely: Riskscore =  $(0.2391)*HSPA1A + (0.1849)*CDKN2A + (-0.1784)*UCN3$ .

### *Model evaluation and external validation*

Figure 4A shows the risk score and survival time of 443 patients in the model based on two CR-DEGs. The KM curve in Figure 4B shows that patients with lower risk scores had a better prognosis. The ROC curves demonstrated that the area under the curve (AUC) of the model was greater than 0.6 at 1, 3, and 5 years (Figure 4C). This indicates that the model has better predictive performance for prognosis. We performed the same analysis on two other models and obtained similar results (Figures 4D-F and 5A-C). In parallel, the cuproptosis- and immune-related risk scoring model was validated in an external dataset. Both the KM curve and the ROC curve drawn using the 568 patients with colon cancer in GSE39582 were meaningful (Figure 5D-F). Therefore, we concentrated on the analysis of this model at a later stage.

### *PCA and the independent predictive value of the risk score*

The PCA showed that different subgroups could be clearly distinguished in the cuproptosis- and immune-related models. (Figure 6A and B) Univariate and multivariate Cox regression analyses of the risk score and other clinical features, such as age and pTNM stage, were carried out. Cuproptosis- and immune-related risk scores were independently related to the OS of patients with colon cancer, indicating that they can serve as independent prognostic predictors. (Figure 6C and D) Afterwards, nomograms were drawn to make it possible to derive a score for each factor from

information about a specific patient and finally to estimate the probability of survival for that patient based on the total score (Figure 6E).

### *Functional enrichment and immune profiling*

The differential expression of the three genes is presented in Figure 6F. Among them, CDKN2A and HSPA1A were upregulated, while UCN3 was downregulated in the high-risk group. Functional enrichment revealed that the DEGs between the two distinct groups were enhanced in the tumor immune microenvironment (Figures 6G and 7A-C). Therefore, we performed immune cell infiltration analysis. The heatmap in Figure 7D shows the infiltration abundance of different immune cells in different subgroups, among which memory B cells, endothelial cells, NK cells and so on were significantly different (Figure 7D and E). The differential expression results of immune checkpoint inhibitors (ICIs) showed that the expression of mRNA expression of perforin (PRF1) and TBX2 in the high-risk group was higher than that in the low-risk group, which may be associated with stronger tumor immune escape in the high-risk group (Figure 7F and G).

## **DISCUSSION**

Colon cancer is the third most common cancer and the second most common cause of global cancer mortality<sup>[21-23]</sup>. During the past decades, the incidence and mortality of this cancer have progressively increased<sup>[21]</sup>, making it imperative and urgent to identify potential biomarkers to evaluate its prognosis and treatment.

A recent study reported that cuproptosis, as a novel type of cell death, is relevant to the development and progression of malignancy<sup>[10]</sup>. Cuproptosis may have a high cytotoxic effect on tumor cells, and abnormal copper levels can affect tumor growth by causing irreversible damage to organelles and consequently inducing tumor cell death through multiple mechanisms, which could be used as new targets for anticancer therapy<sup>[11-13]</sup>. Another study found that immune cell infiltration was associated with a variety of reactions to the prognosis and treatment of colon cancer<sup>[24]</sup>.

In our research, we constructed a cuproptosis- and immune-related prognostic risk model, combined with HSPA1A, CDKN2A and UCN3, which significantly differed between the low-risk and high-risk groups for colon cancer. HSPA1A and CDKN2A exhibited carcinogenesis, and their overexpression was associated with poorer prognosis and survival in colon cancer patients, whereas high expression of the UCN3 gene showed a tumor suppressor profile, which was associated with better prognosis and survival. As one of the HSP70 family members, extracellular HSPA1A is thought to elicit antitumor immune responses<sup>[25]</sup>. Previous studies have shown that the extracellular environment Hsp70 acts as a risk-associated molecular pattern (DAMP) and increases cytokine production, which can form an inflammatory microenvironment conducive to tumor progression<sup>[26]</sup>. In addition, Hsp70 promotes tumor invasion by binding to Toll - like receptors 2 and 4<sup>[27]</sup>. CDKN2A has been declared to be associated with gene mutation and cancer progression, such as colon and breast cancers<sup>[28]</sup>. CDKN2A is upregulated in colon cancer through two potential mechanisms: Immunosuppression and progression of epithelial-mesenchymal transition (EMT), leading to poor prognosis<sup>[29]</sup>. UCN3 might participate in local inflammatory stress response pathways to assist in the maintenance of tissue integrity and homeostasis<sup>[30]</sup>. In undifferentiated epithelial cells of colonic crypts and colon cancer cell lines, corticotropin-releasing factor 2 (CRF2) protein is expressed preferentially. UCN3-induced CRF2 signaling decreases the expression of transcription factors involved in intestinal epithelial cell differentiation and modifies the cytoplasmic and transcellular permeability of related cells<sup>[31]</sup>.

Functional enrichment analysis between the high- and low-risk groups were enriched in oxygen transport, cellular response to monosaccharide stimulus, intermediate filament cytoskeleton, G protein-coupled receptor binding, receptor-ligand activity, and structural constituent of cytoskeleton. It seems that the cellular response may be altered in different risk groups<sup>[32]</sup>. In a previous study, the basal epithelial intermediate filament cytoskeleton was confirmed to be necessary for mass invasion of cancer<sup>[33]</sup>. G protein-coupled receptor binding has been implicated in the tumorigenesis,

progression, and metastasis of cancers<sup>[34]</sup>. Of the many factors, mediating disease ligands and receptor activity may be implicated in cancers<sup>[35]</sup>. Through research on multiple cancer datasets, the regulation of actin cytoskeleton signaling pathways has been identified as a cancer-related factor<sup>[36]</sup>. KEGG enrichment analysis showed that bile secretion, neuroactive ligand-receptor interaction and the PPAR signaling pathway were significantly enriched. Several studies have investigated whether the secretion of bile might be linked to the pathogenesis of colon cancers<sup>[37]</sup>. Recent studies have reported that neuroactive ligand-receptor interaction pathways are relevant to colorectal cancers<sup>[38]</sup>. GSEA showed that immature to mature B lymphocytes, dendritic cell maturation, and silencing by the TME were mainly enriched. However, as antigen-presenting cells (APCs), both dendritic cells and B-lymphocytes can deliver antigens to T-lymphocytes, thus activating immature T-lymphocytes to induce an antitumor immunization effect, which can consequently affect the prognosis of tumors<sup>[39-41]</sup>. Consequently, the prognosis for colon cancer may be regarded as being related to changes in immune cells and the TME. Meanwhile, a growing number of studies on the TME have shown that tumor-infiltrating immune cells play a vital role in cancer progression and aggressiveness, and their molecular quantitative traits are increasingly considered to have predictive value and correlate with the prognosis of colon cancer, so we followed up with an immune infiltration analysis for high-risk and low-risk groups<sup>[42,43]</sup>.

B-cell memory, endothelial cell, and macrophage infiltration were higher in the high-risk group, and since it has been shown that tumor tissue infiltrating B cells are mainly B-cell memory and that macrophages and endothelial cells can contribute to cancer development and malignant progression by stimulating angiogenesis, tumor cell invasion, and intravascular infiltration, such infiltration is related to poor prognosis and reduced survival in most cancers<sup>[44-47]</sup>. The low-risk group, on the other hand, had more activated CD4<sup>+</sup> memory T cells, CD8<sup>+</sup> T cells, and NK cells. CD4<sup>+</sup> T cells can recognize tumor antigens presented by APCs and secrete specific cytokines and costimulatory signaling molecules to stimulate CD8<sup>+</sup> T cells to proliferate and differentiate into tumor

effector T cells or memory T cells, which have strong tumor cytotoxic and immunosurveillance functions<sup>[48,49]</sup>. In addition, NK cells can produce a large number of cytotoxic granule proteins and cytokines without prestimulation to induce apoptosis, which can rapidly exert a tumor-cytotoxicity effect at the early stage of lesions and can complement and coordinate with the immunosurveillance function of cytotoxic T lymphocytes to inhibit the development of tumor cells<sup>[50,51]</sup>, so their infiltration is mostly associated with a better prognosis. Immune checkpoint inhibitors, such as PRF1 and TBX2, suggest that patients in the high-risk group are the most likely to experience immune escape and have a relatively poor prognosis. It has been shown that immune cytolytic activity (CYT) is an immune indicator of cancer that is based on the PRF1 and granzyme A<sup>[52]</sup>. Tumors with high CYT have a high mutational load, which predisposes them to microsatellite instability. Unstable microsatellite colorectal tumors can overexpress several immune checkpoint molecules triggering immune evasion and immunosuppression, which are associated with poorer prognosis<sup>[53]</sup>. There is a significant correlation between tumor TBX2 overexpression and distant tumor metastasis as well as late-stage and recurrence<sup>[53,54]</sup>. However, the specific mechanisms by which TBX2 promotes tumorigenesis and progression are currently unclear. According to the available studies, TBX2 may inhibit the cellular senescence program by participating in the regulation of cyclin kinase inhibitor p19 ARF and p21, as well as providing cells with the ability to bypass the cytokinesis block, which leads to increased genomic instability<sup>[55-57]</sup>. Therefore, TBX2 may act as an immortalizing gene that avoids cellular senescence to participate in the regulation of cyclin and thus influence tumor progression and prognosis<sup>[53]</sup>.

In this study, based on cuproptosis- and immune-related genes, we identified prognostic signatures for colon cancer and validated them by using an external dataset. Nevertheless, additional confirmation foresight and experimental validation is necessary.

## CONCLUSION

The prognosis of <sup>17</sup>the high-risk group evaluated in the combined model was poorer, and cuproptosis was highly correlated with the prognosis of colon cancer. It is possible that we may be able to improve patients' prognosis using the *in vitro* intervention risk score.

<sup>22</sup>**Figure 1 A brief flowchart of the study.** TCGA: The Cancer Genome Atlas; DEG: Differentially expressed gene; ROC: Receiver Operating Characteristic; CR: Cuproptosis-related; IR: Immune-related.

**Figure 2 Acquisition of cuproptosis-related and immune-related differentially expressed genes.** A: <sup>14</sup>Heatmap of differentially expressed genes in The Cancer Genome Atlas colon cancer patients compared with normal controls; B: <sup>14</sup>Volcano plot of differentially expressed genes (DEGs), where blue represents downregulated genes and red represents upregulated genes; C: <sup>24</sup>Ten cuproptosis-related genes DEGs (CR-DEGs) were calculated by taking the intersection of DEGs and cuproptosis-related genes; D: <sup>1</sup>The DEGs with  $|\log_2FC| > 1$  and  $p.adjust < 0.05$  were intersected with immune-related genes to obtain IR-DEGs; E: The result of univariate Cox regression analysis of CR-DEGs; F: <sup>2</sup>The result of multivariate Cox regression analysis of CR-DEGs. CRG: Cuproptosis-related gene; IRG: Immune-related genes.

**Figure 3 Identification of prognostic features and construction of risk scoring models.** A: <sup>7</sup>Coefficient screening plot generated when the model was constructed using Least Absolute Shrinkage and Selection Operator (LASSO)\_Cox regression analysis with CDKN2A and DLAT as variables. When  $\lambda$  takes the minimum value, it is the best model; B: <sup>2</sup>Univariate Cox regression analysis of immune-related differentially expressed genes (IR-DEGs) showed significant results; C: <sup>2</sup>Multivariate Cox regression analysis of IR-DEGs identified four genes independently associated with prognosis; D: Coefficient screening plots generated when the model was constructed using LASSO\_Cox regression analysis with UCN3, HSPA1A, ULBP2, and FABP4 as variables. When  $\lambda$

takes the minimum value, it is the best model; E: Multivariate Cox regression analysis of two cuproptosis-related (CR)-DEGs and four IR-DEGs. CIR-DEGs with  $P < 0.05$  were independently associated with prognosis; F: Coefficient screening plots when the model was constructed with the selected CDKN2A, HSPA1A, and UCN3 variables.

**Figure 4 Model evaluation.** A: The expression differences of DLAT and CDKN2A in high- and low-risk groups and the risk scores and survival time of different subgroups are shown; B: KM curve of the cuproptosis-related model. The higher the cuproptosis-related risk score was, the worse the prognosis; C: The Receiver Operating Characteristic (ROC) curve of the cuproptosis-related model, and the size of the area under the curve can reveal the quality of the model. In the figure, the area under the curve (AUC) values in years 1, 3, and 5 are all greater than 0.6, indicating that the model has strong predictive ability; D: The expression differences of the four immune-related differentially expressed genes used to construct the model in the high- and low-risk groups and the risk scores and survival times of different subgroups are shown; E: KM curve of the immune-related model; the higher the risk score was, the worse the prognosis; F: The ROC curve of the immune-related model, and its AUC values at 1, 3, and 5 years were all greater than 0.6. TCGA: The Cancer Genome Atlas database; CR: Cuproptosis-related; IR: Immune-related.

**Figure 5 External validation.** Based on CDKN2A, HSPA1A, and UCN3, a cuproptosis- and immune-related prognostic model was constructed. In samples of the The Cancer Genome Atlas database. A: The expression differences of these three genes in high- and low-risk groups and the risk score and survival time of different subgroups; B: The KM curve indicates that the higher the score is, the worse the prognosis; C: The Receiver Operating Characteristic curve shows that the model has area under the curve values

greater than 0.6 in years 1, 3, and 5; D-F: In the Gene Expression Omnibus database, the model obtained similar results, and  $P < 0.05$ . TCGA: The Cancer Genome Atlas database.

**Figure 6 Principal Component Analysis and the independent predictive value of the risk score.** A and B: The Principal Component Analysis plots in The Cancer Genome Atlas and Gene Expression Omnibus, respectively, and both high- and low-risk groups can be obviously distinguished; C and D: Univariate and multivariate Cox regression analyses of the cuproptosis- and immune-related risk score and other clinical features to screen for factors independently associated with prognosis; E: A nomogram helps predict a patient's prognosis based on the patient's information; F: Differential expression of CDKN2A, HSPA1A, and UCN3 in the high- and low-risk groups; G: Functional enrichment consequences of differentially expressed genes in the high- and low-risk groups. TCGA: The Cancer Genome Atlas database.  $^cP < 0.001$ .

**Figure 7 Functional enrichment and immune profiling.** A-C: Gene Set Enrichment Analysis functional enrichment results of differentially expressed genes in the high- and low-risk groups; D: Heatmap showing the infiltrating abundance of different immune cells in different groups analyzed by CIBERSORT and EPIC algorithms; E: Significant ( $P < 0.05$ ) results presented in immune infiltration analysis; F and G: Differential expression of immune checkpoint inhibitors in the high- and low-risk groups. Both were highly expressed in the high-risk group.  $^aP < 0.05$ ;  $^bP < 0.01$ ;  $^cP < 0.001$ .

**Table 1** Information on **The** Cancer Genome Atlas and Gene Expression Omnibus samples

	Charar	TCGA ( <i>n</i> = 443)	GSE39582 ( <i>n</i> = 568)
Status	Alive	346	382
	Dead	97	186
Age	Mean (SD)	66.9 (13.1)	66.8 (13.2)
	Median	[MIN, 68 [31,90]	68 [22,97]
	MAX]		
Gender	FEMALE	211	257
	MALE	232	311
pTNM_stage	I	75	40
	II	176	265
	III	128	203
	IV	64	60

TCGA: The Cancer Genome Atlas.

# 22%

SIMILARITY INDEX

### PRIMARY SOURCES

1	<a href="http://www.frontiersin.org">www.frontiersin.org</a> Internet	307 words — 7%
2	<a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a> Internet	55 words — 1%
3	<a href="http://www.researchgate.net">www.researchgate.net</a> Internet	52 words — 1%
4	<a href="http://www.spandidos-publications.com">www.spandidos-publications.com</a> Internet	44 words — 1%
5	<a href="http://www.hindawi.com">www.hindawi.com</a> Internet	38 words — 1%
6	<a href="http://bmccancer.biomedcentral.com">bmccancer.biomedcentral.com</a> Internet	37 words — 1%
7	<a href="http://www.dovepress.com">www.dovepress.com</a> Internet	35 words — 1%
8	<a href="http://academic-accelerator.com">academic-accelerator.com</a> Internet	34 words — 1%
9	<a href="http://downloads.hindawi.com">downloads.hindawi.com</a> Internet	34 words — 1%
10	<a href="http://www.researchsquare.com">www.researchsquare.com</a> Internet	

32 words — 1%

11 [www.wjgnet.com](http://www.wjgnet.com)  
Internet

22 words — 1%

12 Benjamin Ducarouge, Marjolaine Pelissier-Rota, Rebecca Powell, Alain Buisson, Bruno Bonaz, Muriel Jacquier-Sarlin. "Involvement of CRF2 signaling in enterocyte differentiation", World Journal of Gastroenterology, 2017  
Crossref

18 words — < 1%

13 Tianhao Cong, Yingen Luo, Yu Liu, Chao Yang, Hongcai Yang, Yujie Li, Jingui Li, Xiao Li. "Cuproptosis-related immune checkpoint gene signature: Prediction of prognosis and immune response for hepatocellular carcinoma", Frontiers in Genetics, 2022  
Crossref

18 words — < 1%

14 [pubmed.ncbi.nlm.nih.gov](http://pubmed.ncbi.nlm.nih.gov)  
Internet

18 words — < 1%

15 [www.lion-oncology.com](http://www.lion-oncology.com)  
Internet

18 words — < 1%

16 [www.mdpi.com](http://www.mdpi.com)  
Internet

18 words — < 1%

17 Zhenzhen Mao, Ye Nie, Weili Jia, Yanfang Wang, Jianhui Li, Tianchen Zhang, Xinjun Lei, Wen Shi, Wenjie Song, Xiao Zhang. "Revealing Prognostic and Immunotherapy-Sensitive Characteristics of a Novel Cuproptosis-Related LncRNA Model in Hepatocellular Carcinoma Patients by Genomic Analysis", Cancers, 2023  
Crossref

14 words — < 1%

18 Cao-Jie Chen, Hiroki Kajita, Noriko Aramaki-Hattori, Shigeki Sakai, Kazuo Kishi. "Screening of Autophagy-Related Prognostic Genes in Metastatic Skin Melanoma", Disease Markers, 2022  
Crossref 12 words — < 1%

19 doaj.org  
Internet 12 words — < 1%

20 e-dmj.org  
Internet 12 words — < 1%

21 euroblight.net  
Internet 12 words — < 1%

22 www.aging-us.com  
Internet 12 words — < 1%

23 hal.archives-ouvertes.fr  
Internet 11 words — < 1%

24 Jiahui Wei, Baoyi Huang, Yingqi Nong, Qianyu Zhang et al. "Identification of a Novel Cuproptosis-Related Gene Signature in Eutopic Endometrium of Women with Endometriosis", Reproductive Sciences, 2022  
Crossref 9 words — < 1%

25 Yajie Zhang, Mingxin Zhu, Limin Xie, Haowei Zhang, Tuo Deng. "Identification and validation of key immune-related genes with promising diagnostic and predictive value in systemic sclerosis", Life Sciences, 2023  
Crossref 9 words — < 1%

26 Zhuo Lin, Zhuoyan Chen, Dan Miao, Qian Xu, Liuwei Zeng, Luying Zhao, Jun Xu, Xian Song, Fujun Yu. "Pyroptosis-related genes IRAK1, NOD2, POP1 and YWHAB 9 words — < 1%

impact the prognosis through immune functional pathways in hepatocellular carcinoma", Research Square Platform LLC, 2021

Crossref Posted Content

27 Zihao Zhang, Bingcheng Wang, Xiaoqin Xu, Tao Xin. "Cuproptosis-related gene signature stratifies lower-grade glioma patients and predicts immune characteristics", Frontiers in Genetics, 2022

Crossref

28 [www.medicinescience.org](http://www.medicinescience.org)

Internet

29 Jiarong He, Zhongzhong Jiang, Jiawei Lei, Wen Zhou, Yan Cui, Biao Luo, Mingming Zhang. "Prognostic Value and Therapeutic Perspectives of CXCR Members in the Glioma Microenvironment", Frontiers in Genetics, 2022

Crossref

30 Liukai Ma, Ruoling Jia, Mengyao Wang, Cheng Yan. "A new prognostic model for pancreatic cancer based on pyroptosis-related genes", Cold Spring Harbor Laboratory, 2021

Crossref Posted Content

31 Yuwei Liao, Simon X Liang, Dong He, Zhikai Chen, Jinxiu Yao, Hongfeng Liang, Liye Yang. "m6A-related lncRNAs are Potential Biomarkers for Predicting Prognoses in Patients with Breast Cancer", Research Square Platform LLC, 2021

Crossref Posted Content

32 Zheng Li, Hui Wang, Xia Deng, Jing Zhang, Wanyan Tang, Ling Wang, Wenxin You, Weiqi Nian. "Construction and Validation of an Immune-related lncRNA

---

33 [bmcgastroenterol.biomedcentral.com](https://bmcgastroenterol.biomedcentral.com) 8 words — < 1%  
Internet

---

34 [worldwidescience.org](https://worldwidescience.org) 8 words — < 1%  
Internet

---

35 Aibin Liu, Yanyan Li, Lin Shen, Na Li, Liangfang Shen, Zhanzhan Li. "Comprehensive analysis of Histone deacetylases genes in the prognosis and immune infiltration of glioma patients", Cold Spring Harbor Laboratory, 2022 7 words — < 1%  
Crossref Posted Content

---

36 Xiangyu Li, Hao Cai, Weicheng Zheng, Mengsha Tong et al. "An individualized prognostic signature for gastric cancer patients treated with 5-Fluorouracil-based chemotherapy and distinct multi-omics characteristics of prognostic groups", Oncotarget, 2016 7 words — < 1%  
Crossref

---

37 Xingyu Peng, Jinfeng Zhu, Sicheng Liu, Chen Luo, Xun Wu, Zitao Liu, Yanzhen Li, Rongfa Yuan. "Signature construction and molecular subtype identification based on cuproptosis-related genes to predict the prognosis and immune activity of patients with hepatocellular carcinoma", Frontiers in Immunology, 2022 7 words — < 1%  
Crossref

---

38 Hong-kai Xu, Shao-liang Han, Hao-feng Lu, Rui-shuai Dai, Sai-yi Han, Wei-dong Xie. "Construction and Validation of a Coagulation Factor-Related Prognostic Model for Colorectal Cancer Based on the Public Database", Research Square Platform LLC, 2022 6 words — < 1%

---

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	OFF	EXCLUDE MATCHES	OFF