Identification of four Ferroptosis Gene signatures and Establishment of a predictive model for the overall survival rate of gastric cancer

Jiang Liu
Department of Gastrointestinal Surgery, the Second Affiliated Hospital of Nanchang University, Nanchang

Hua-kai Tian
Department of General Surgery, the First Affiliated Hospital of Nanchang University, Nanchang

Nan-tao Fu
Department of day ward, the first Affiliated Hospital of Nanchang University, Nanchang

Ce-gui Hu
Department of Gastrointestinal Surgery, the Second Affiliated Hospital of Nanchang University, Nanchang

Fei-long Zou
Department of Gastrointestinal Surgery, the Second Affiliated Hospital of Nanchang University, Nanchang

Zhen Zong
Department of Gastrointestinal Surgery, the Second Affiliated Hospital of Nanchang University, Nanchang

Zhi-kun Ning (huihanmo594@163.com)
Department of day ward, the first Affiliated Hospital of Nanchang University, Nanchang

Research Article

Keywords: ferroptosis, gastric cancer, nomogram, prognosis, the cancer genome atlas, gene expression omnibus

Posted Date: April 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1535844/v1

Abstract

Background: Gastric cancer (GC) is highly aggressive and recurrent. Ferroptosis is believed to be closely related to the occurrence and development of tumors, but the specific mechanism is still unclear. This study aims to construct a new prognostic model based on ferroptosis-related gene scores to assess the prognosis of GC patients and guide clinical treatment.

Materials and Methods: The gene expression information of GC patients with clinical data were obtained from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO). Differentially expressed genes (DEGs) between tumor and normal tissue were screened from TCGA dataset, and ferroptosis-related genes were downloaded from the ferroptosis database. The ferroptosis-related DEGs were identified by intersecting the ferroptosis-related genes with DEGs. Univariate Cox and LASSO regression analysis was applied to identify survival-related intersection genes (SRIGs) and 4 hub genes. The TCGA dataset was randomly divided into 2 cohorts in a ratio of 0.7:0.3, the training cohort for construction of signature and the testing cohort for internal validation. The GSE26901 was used for external validation. Receiver operating characteristic (ROC), Kaplan-Meier curve and risk curve of 3 cohorts were plotted to evaluate performance of the signature. Multivariate Cox regression analysis was used to determine independent prognostic factors for gastric cancer, and then a prognostic nomogram was established based on risk score and clinicopathological parameters.

Results: A four-gene signature comprising MYB, CHAC1, NOX4 and AFT3 was constructed to predict overall survival of GC. Using the best cut-off value to divide patients into high-risk and low-risk groups, we found that the survival rate of the low-risk group in the training group, test group and external validation group was significantly higher than that of the high-risk group ($P<0.05$). The area under the ROC curve for 1 year, 3 year, and 5 year were (0.635, 0.666, 0.713), (0.63, 0.681, 0.602) and (0.758, 0.679, 0.677). Both univariate and multivariate Cox regression analysis showed that 4 gene markers are independent prognostic factors for GC. The 1-year, 3-year and 5-year overall survival consistency tests show that the nomogram has good predictability. Immune correlation analysis of high-risk population indicated that tumor progression may be related to the high expression of Macrophages.

Conclusion: Our study identified a signature with 4 ferroptosis-related genes and established a reliable prognostic nomogram for predicting the overall survival rate of GC patients. The results may help medical decision-making, and provide a novel reference for predicting the prognosis of GC and even potential molecular targets.

1. Introduction

GC is the fifth most common malignancy diagnosed worldwide and the third highest death toll of cancer, with more than 1 million cases diagnosed every year[1–3]. Patients have almost no typical clinical symptoms before advanced stage, which leads to the early diagnosis is very difficult[4]. Many patients with GC are diagnosed at an advanced stage, by which time the majority of GC have distantly

metastasized and the surgery is too late[5, 6]. The treatment options for GC are mainly radical resection, supplemented by postoperative radiotherapy, chemotherapy, targeted therapy and immunotherapy, etc[7]. The prognosis of GC is not ideal, and the five-year survival rate does not exceed 30%[8]. At present, most of the main factors predicting the prognosis of GC are clinicopathological factors, but there are huge differences between individualization. Only relying on clinical information parameters cannot accurately predict the prognosis of GC patients and guide individualized treatment[9]. Therefore, more and more scholars use genes related to tumor prognosis in combination with clinical information parameters to evaluate patient prognosis and better guide individualized clinical treatment, providing possible targets for individualized targeted therapy[10, 11].

Ferroptosis coined in 2012 to describe an iron-dependent regulated form of cell death is driven by loss of activity of the lipid repair enzyme glutathione peroxidase 4 (GPX4) and subsequent accumulation of lipid-based reactive oxygen species (ROS), particularly lipid hydroperoxides[12, 13]. Features of ferroptosis have been observed periodically over the last several decades, and have been shown recently to play significant roles in numerous biological contexts[14, 15]. Ferroptosis plays an important role in regulating the progression of multiple tumors and may provide a new opportunity to therapies of tumor[16–18]. Ma,et al (2016)[19] found ferroptosis was induced in breast cancer cells following siramesine and lapatinib treatment. Sun,et al (2016)[20] found sorafenib-resistant cells are more sensitive to ferroptosis in hepatocellular carcinoma. Liu,et al (2020)[21] proposed that ferroptosis and its regulatory factors may be related to the survival rate of cancer patients and other clinical characteristics. These novel Insights were provided in the mechanism of ferroptosis and the development of cancer. However, the relationship between ferroptosis and prognosis in GC is still unclear.

In the present study, we explored ferroptosis-related genes and screened prognosis-related genes. Finally, a signature is established and verified reliability by bioinformatic methods to predict prognosis of patients with GC. Altogether these data, we believe, will contribute to help clinicians in clinical decision-making, provide new reference for patient prognosis assessment, and may become a new target for gene individualization therapy.

2. Materials And Methods

2.1 Download and pre-processing of datasets

The gene expression data and corresponding clinical information of GC patients, including age, gender, TNM stage, Lauren parting and survival information, were downloaded from the TCGA database (https://portal.gdc.cancer.gov/projects), for a total of 365 GC samples.. Ferroptosis-related genes were retrieved in the ferroptosis database (http://www.zhounan.org/ferrdb/index.html), including total ferroptosis maker genes, ferroptosis suppressor genes and ferroptosis driver genes. Another chip expression data GSE26901 with 109 cases was obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) for the following external validation. The samples for

which the gene expression was “zero” or clinical information was “none” were excluded from the analysis. Data were collected from September 22 to October 8, 2020.

2.2 Screening of DEGs

The TCGA gene counts were used to screen differentially expressed genes(DEGs) in the “DESeq2” R package and transformed to normalized data using variance stabilization transformation (VST) in order to the further analysis. Through calculations, we screened out the DEGs between tumor samples and normal samples, and the criteria was the absolute fold change (FC) > 1 and the adjusted P value < 0.05. A volcano plot was made using the “ggplot2” R package, and a heat map was generated using the “pheatmap” R package to visualize DEGs.

2.3 Selection and bioinformatic analysis of ferroptosis-related DEGs and survival-related intersection genes (SRIGs)

We crossed the ferroptosis-related genes and the DEGs to select the ferroptosis-related DEGs. Venn diagrams were plotted in the “VennDiagram” R package to display the intersection genes, and the difference of expression of these genes were showed in a heat map between tumor samples and normal samples. In order to further understanding of ferroptosis-related DEGs, Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed in the “clusterProfiler” R package[22]. GO enrichment contained biological processes (BP), cellular components (CC), and molecular functions (MF), and the top 10 terms of each part was showed in the generated figure.

To guarantee that different databases can be analyzed together, the expression levels of TCGA and GEO are standardized to the range of 0 ~ 1. TCGA dataset was used to identify overall survival-related intersection genes (SRIGs) using Univariate Cox regression analysis; the genes with $p < 0.05$ was considered statistically significant and used in the subsequent analysis. Then, a forest plot was made in the “forestplot” R package to display these genes and their hazard radio. GO enrichment analysis were applied to reveal potential function of SRIGs. Protein-protein interaction (PPI) network analysis was constructed in the STRING database (https://string-db.org) to identify the potential interaction of SRIGs. And the network visualization of these genes was conducted with the “igraph” R package.

2.4 Construction and validation of prognostic gene signature

we performed least absolute shrinkage and selection operator (LASSO) regression analysis with the “glmnet” R package to screen genes with the best predictive performance using 10-fold cross validation, and finally, the 4 most relevant genes were selected to define the later model[23]. LASSO is a statistical regularization method that imposes a penalty term on regression-based parameter estimation to prevent overfitting of the model[24]. All TCGA patients were divided into 2 groups, the training cohort and the

testing cohort, in a ratio of 0.7:0.3 using the stratified randomization method. The testing cohort was used for internal validation and GSE26901 for external validation. Based on the linear combination of gene expression level and its regression coefficients from Lasso Cox regression, we calculated the risk score of each patient in the training cohort. The prognostic signature was showed as risk score = Σ expression level of gene Xi × Cox coefficient of gene Xi. The optimal cut-off value was calculated by means of maximally selected rank statistics in the training cohort for overall survival (OS) using the "survminer" R package, and patients were divided into high risk and low risk groups from each cohort according to the cut-off. Next, time-dependent receiver operating characteristic (ROC) curve of training cohort, testing cohort and external validation cohort were plotted and the area under the curve (AUC) was employed to evaluate accuracy of the prognostic signature using "timeROC" R package. The Kaplan-Meier survival curve of 3 cohorts were used to compare OS difference between high and low groups using the "survival" R package. Risk curve, scatter plot and risk heatmap also showed the difference among two groups.

2.5 Difference between high-/low-risk groups and immune score analysis

The method of dimensionality reduction, like PCA and t-NSE, could be further demonstrated the separation of high and low risk groups. Principal component analysis (PCA) which used orthogonal transformation linearly transformed the observations of a series of possibly related variables to the values of a series of linear unrelated variables and projected it. Subsequently, the principal components were used for t-Distributed Stochastic Neighbor Embedding (t-SNE) clustering, and the tSNE plot was then generated using the "Rtsne" and the "ggplot2" R package. We also conducted single sample gene set enrichment analysis (ssGSEA) independent for each sample were scored, using the results of immune function and infiltration of immune cells to show differences in the level of each group. The method was performed using the "GSVA" R package.

Furthermore, in order to analyze the difference in immune infiltration of high- and low-risk groups, we performed 7 evaluation methods to improve the credibility of the results[25]. These methods contained XCELL, QUANTISEQ, MCCOUNTER, EPIC, TIMER, CIBERSORT-CBS and CIBERSORT-CBA.

2.6 Independent prognostic factors and establishment and validation of nomogram

Univariate and multivariate Cox proportional-hazards analysis were performed through the "rms" R package to determine independent prognostic factors in TCGA; only a factor with pvalue < 0.05 was considered statistically significant and included to participate in the establishment of the nomogram to predict 1-, 3-, and 5-year overall survival of GC patients in the TCGA dataset. Predictive accuracy of nomogram was measured using Harrell's concordance index (C-index). We also established calibration curves to evaluate the nomogram, and the closer to the dotted line, the more accurate it was, 1, 3, and 5 years respectively.

2.7 Analysis of hub genes in the prognostic model

Gene Expression Profiling Interactive Analysis (GEPIA) was then used to analyse expression difference of the 4 hub genes between normal and tumor tissue in the web survey (http://gepia.cancer-pku.cn/). The expression of 4 hub genes in Stomach adenocarcinoma (STAD) were read and downloaded to demonstrate. Similarly, 4 KM curves were made from the Kaplan-Meier plotter (http://kmplot.com/analysis/) to display differential expression.

2.8 Statistical Analysis

All statistical analysis was performed in R v. 3.6.1, and each package we used was listed in the proceeding. In all the tests, p-value less than 0.05 was considered statistically significant. Statistical significance is shown as *p < .05, **p < .01, ***p < .001 and **** p < .0001.

3. Results

3.1 Screening and function enrichment analysis of ferroptosis of DEGs in TCGA

Total 4751 DEGs screened between tumor and normal samples were displayed in a volcano plot (Fig. 1A) and a heatmap (Fig. 1B), and the expression difference could be visualized in the heatmap. By exploring the ferroptosis database, we downloaded total 111 ferroptosis maker genes, 70 ferroptosis suppressor genes and 259 ferroptosis driver genes. The specific quantity could be browsed in the Venn plot (Fig. 1C), which was used to display 59 intersection genes between ferroptosis-related genes and DEGs. Similarly, we made a heat map (Fig. 1D) to show that the expression of ferroptosis-related DEGs is significantly different in normal and tumor tissues.

GO and KEGG enrichment analysis were used to discover function of ferroptosis-related genes (Fig. 1E). KEGG analysis showed the main enrichment pathways of genes were AGE-RAGE signaling pathway in diabetic complications□HIF-1 signaling pathway□African trypanosomiasis□Bladder cancer. GO analysis showed the main enrichment in BP: response to oxidative stress□cellular response to chemical stress□ response to nutrient; CC: oxidoreductase complex□NADPH oxidase complex□invadopodium; MF: oxidoreductase activity,acting on NAD(P)H□iron ion binding□oxidoreductase activity, acting on single donors with incorporation of molecular oxygen.

3.2 survival-related intersection genes (SRIGs) and function enrichment analysis

Subsequently, 5 genes without prognostic information in TCGA from 59 intersection genes were identified and excluded, so the remaining 54 genes were used Univariate Cox regression analysis to select survival-related intersection genes (SRIGs). As a result, 14 SRIGs (CAV1□DUSP1□ZEB1□NOX4□RGS4□MYB□ZFP36□

ATF3、TSC22D3、TXNIP、CDO1、HELLS、FANCD2、CHAC1) were screened through univariate COX analysis and drawed a forest map. (Fig. 2A and Supplementary Table 1).

GO enrichment analysis indicated 14 SRIGs were mainly enriched in BP : response to antibiotic、response to steroid hormone、response to oxidative stress、negative regulation of MAPK cascade、negative regulation of transferase activity、regulation of T cell activation、regulation of MAP kinase activity (Fig. 2B); while CC and MF enrichment were not statistically significant. Visualization of the PPI network and relevance network clarified the link between SRIGs (Fig. 2C, D).

3.3 Construction and validation of prognostic gene signature

The 14 genes obtained by the above univariate Cox regression analysis were subjected to lasso regression analysis (Fig. 3A, B), and 4 hub genes (MYB, CHAC1, NOX4 and ATF3) were distinguished and used to construct prognostic gene signature (Supplementary Table 2). The down-regulation of MYB and CHAC1 indicated that these genes were tumor suppressor genes (HR < 1), while the up-regulation of NOX4 and ATF3 indicated that these genes were oncogenes (HR > 1). According to the optimal cut-off value calculated from “survminer” R package, patients were divided into high risk and low risk groups (Fig. 3C). The risk score = 1.057289 × NOX4 expression + (-1.007024) × MYB expression + (-1.273476) × CHAC1 expression + 1.239175 × ATF3 expression. The TCGA database is divided into training group and test group according to 0.7:0.3, and GSM26901 is used as an external verification group for verification. The expression levels of TCGA and GEO were standardized to the range of 0–1. The KM curves from 3 cohorts revealed visible difference in OS between high risk and low risk groups (P < 0.05). The area under the ROC curve (AUC) of 1 year, 3 years and 5 years respectively arrived at 0.635,0.666,0.713 in training cohort, 0.63,0.681,0.602 in testing cohort and 0.758,0.679,0.677 in external validation cohort. The risk curves of the three cohorts all showed that there was higher mortality in the high-risk group then the low. Heat maps further demonstrated that MYB and CHAC1 were down-regulated and NOX4 and ATF3 were up-regulated in the tumor group (Fig. 3D, E and F). All these results indicated that the 4 hub genes had high sensitivity and accuracy, and could well predict the survival rate of gastric cancer patients.

3.4 Difference between high-/low-risk groups and immune score analysis

Through dimensionality reduction processing, gene expression can be visualized by PCA analysis and t-SNE analysis. The results clearly showed that the high- and low-risk groups could be separated in both datasets (Fig. 4A, B, E and F). Independent immune scores were performed on each sample through ssGSEA analysis to evaluate immune function and immune cell infiltration. In TCGA dataset, immune cells expressed differences between high and low risk groups were: B_cells, CD8 + _T_cells, Macrophages, Mast_cells, Neutrophils, T_helper_cells, TIL, and all immune cells were highly expressed in the high-risk group; immune function differentially expressed were: APC_co_stimulation, CCR, HLA, MHC_class_I, Parainflammation, Type_II_IFN_Reponse; except MHC_class_I, all the others were highly expressed in the high-risk group. Yet, different results of ssGSEA were presented in GSE26901. Only Macrophage was

expressed differences between high- and low-risk groups, and different expression of immune function were: APC_co_inhibition, MHC_class_I and Type_II_IFN_Reponse (Fig. 4G, H). Meanwhile, analyzing the differences between TCGA high and low risk groups by using different immunoinfiltrative assessment methods, we found Macrophages in most methods showed that there were differences in expression between high- and low-risk groups, and it was highly expressed in high-risk groups (Figure S1).

3.5 Establishment and validation of nomogram

The TCGA database provided a series of complete clinical information, including gender, age, staging, TNM staging, grading, tumor location, Lauren classification, etc(Table 1). Univariate and multivariate Cox regression analysis were conducted to identify prognostic factors of GC. The univariate analysis indicated that age ($p = 0.037$), grade ($p < 0.001$), T ($p = 0.013$), N ($p < 0.001$), M ($p = 0.001$), and risk score ($p < 0.001$) were prognostic factors (Fig. 5A), and the following multivariate analysis further suggested age ($p = 0.023$) and risk score ($p < 0.001$) were independent prognostic factors for OS (Fig. 5B). Based on TCGA, we then constructed a nomogram to predict 1-year, 3years, 5years OS using independent prognostic factors – age and risk score (Fig. 5C). The C-index of risk score was 0.733, and 1-year, 3-years, and 5-years OS consistency test showed that the model had good predictability and the predicted value is close to the actual value (Fig. 5D).

Table 1
Baseline

	GEO			TCGA		
	high	low	p.overall	high	low	p.overall
	N = 76	N = 33		N = 242	N = 123	
Gender:			0.932			0.675
female	18 (35.3%)	22 (37.9%)		79 (35.1%)	53 (37.9%)	
male	33 (64.7%)	36 (62.1%)		146 (64.9%)	87 (62.1%)	
Age:			0.730			0.169
>55	22 (43.1%)	28 (48.3%)		181 (80.8%)	120 (87.0%)	
≤55	29 (56.9%)	30 (51.7%)		43 (19.2%)	18 (13.0%)	
Stage:			0.086			0.018
Stage I	12 (23.5%)	26 (44.8%)		20 (9.71%)	28 (20.6%)	
Stage II	11 (21.6%)	9 (15.5%)		65 (31.6%)	45 (33.1%)	
Stage III	18 (35.3%)	18 (31.0%)		99 (48.1%)	48 (35.3%)	
Stage IV	10 (19.6%)	5 (8.62%)		22 (10.7%)	15 (11.0%)	
Lauren:			0.582			0.115
diffuse	7 (14.9%)	4 (7.84%)		44 (51.2%)	17 (35.4%)	
intestinal	38 (80.9%)	44 (86.3%)		42 (48.8%)	31 (64.6%)	
mixed	2 (4.26%)	3 (5.88%)		-	-	
Grade:			-			0.001
G1	-	-		5 (2.22%)	5 (3.57%)	
G2	-	-		64 (28.4%)	66 (47.1%)	
G3	-	-		151 (67.1%)	65 (46.4%)	
GX	-	-		5 (2.22%)	4 (2.86%)	
T_Stage:			-			0.001
T1	-	-		4 (1.78%)	12 (8.57%)	
T2	-	-		43 (19.1%)	35 (25.0%)	
T3	-	-		103 (45.8%)	62 (44.3%)	

	GEO			TCGA		
T4	-	-		67 (29.8%)	31 (22.1%)	
TX	-	-		8 (3.56%)	0 (0.00%)	
N_Stage:			-			0.012
N0	-	-		54 (24.1%)	53 (38.1%)	
N1	-	-		66 (29.5%)	28 (20.1%)	
N2	-	-		41 (18.3%)	32 (23.0%)	
N3	-	-		51 (22.8%)	22 (15.8%)	
NX	-	-		12 (5.36%)	4 (2.88%)	
M_Stage:			-			0.629
M0	-	-		197 (87.6%)	127 (90.7%)	
M1	-	-		16 (7.11%)	8 (5.71%)	
MX	-	-		12 (5.33%)	5 (3.57%)	
Location:			0.192			0.180
antrum	25 (49.0%)	31 (53.4%)		78 (36.1%)	53 (39.8%)	
body	15 (29.4%)	20 (34.5%)		49 (22.7%)	40 (30.1%)	
cardia	1 (1.96%)	0 (0.00%)		60 (27.8%)	26 (19.5%)	
fundus	4 (7.84%)	0 (0.00%)		29 (13.4%)	14 (10.5%)	
entire	6 (11.8%)	7 (12.1%)		-	-	

3.6 Expression feature of four hub genes

Through analysis in GEPIA (http://gepia.cancer-pku.cn/), we knew that the expression of NOX4 and MYB in STAD tumor tissues had increased significantly, however, CHAC1 and ATF3 showed no significant difference between tumor and normal tissues (Figure S2E). Figure S2ABCD also showed that NOX4 and MYB highly expressed in tumor tissues from the Cho Gastric, DErrico Gastric, and Cui Gastric datasets. After that, we analyzed the probability of the expression of the four genes and found higher probability of NOX4 with low expression, higher probability of MYB and ATF3 with high expression and no significant difference in expression of CHAC1 (Figure S2F). Finally, the typical immunohistochemistry (IHC) of 3 genes (except NOX4) were showed in tumor and normal gastric tissue (Figure S3).

4. Discussion

GC is a relatively common gastrointestinal tumor, and its incidence is gradually increasing. It is characterized by a high recurrence rate and a large difference in prognosis[26, 27]. Currently, treatment options for GC is mainly radical excision, supplemented simultaneously by postoperative radiotherapy, chemotherapy, targeted therapy and immunotherapy[28]. Accurate prediction of patient prognosis can benefit patients in subsequent treatment. Therefore, it is possible to develop an appropriate treatment plan for the individual differences in order to improve the prognosis[29]. Traditional clinicopathological factors have been widely used to predict the prognosis and progression of tumors at present[30, 31]. While, individual difference in reaction is more genetic difference. With the standardized quantification of gene testing procedures, gene expression can be used to predict the prognosis of tumors and even served as a target for new treatments[32, 33]. As a newly discovered form of regulatory cell death different from apoptosis, necrosis and autophagy in terms of morphologic, biochemical and genetic aspects, ferroptosis is induced by abnormal accumulation of iron-dependent lipid reactive oxygen species. Studies have found that ferroptosis was closely related to tumor development, and it could promote apoptosis in tumor cell by inducing ferroptosis[34]. With the study of the mechanism of ferroptosis, increasing ferroptosis-related genes have been found and the genes associated with prognosis may well predict prognosis of clinical patients[35, 36]. Compared with traditional prediction models, this study combines the ferroptosis-related prognostic gene risk score and clinical tumor pathological information to predict the prognosis of GC, which will be more accurate and be beneficial for clinical decision-making and individualized treatment.

In this study, analysing in TCGA dataset and ferroptosis-related genes, we screened 59 ferroptosis-related DEGs in gastric cancer. Functional enrichment analysis showed that the DEGs were strongly linked biological processes of tumor cell oxidation, and AGE-RAGE signaling pathway and HIF-1 signaling pathway were the most abundant signaling pathways. The univariate survival prognostic analysis showed that 14 genes were associated with prognosis and the further lasso-COX regression analysis revealed 4 hub genes associated with ferroptosis that influenced the prognosis of gastric cancer. The down-regulation of MYB and CHAC1 expression was identified as tumor protection genes, and the up-regulation of NOX4 and ATF3 expression was related to the poor prognosis of tumors. In the results, the risk scores of the four genes were independent prognostic factors for GC, and the prognosis of patients in the low-risk group was significantly higher than that of the high-risk group. The training cohort in TCGA and GSE26901 in GEO were respectively used to be internal and external validation, which confirmed the predictive property of four hub genes. At last, the nomogram was constructed by a combination of the risk score with clinicopathological information, and performed to accurately predict overall survival (OS).

Four hub genes that predict the prognosis of GC in this study have been previously reported to be related to tumor development. MYB has been proved to be associated with tumor-associated macrophages in breast cancer, and high expression of MYB can be regarded as evidence of macrophage aggregation[37]. MYB is also related with ferroptosis, and some study have shown that the ferroptosis produced by a classic inducer erastin could reduce through inhibition of MYB. The specific mechanism of the action may be related to cysteine dioxygenase 1 (CDO1)[38]. Studies have shown that artesunate can enhance

the ferroptosis in DAUDI and CA-46 cells by activating ATF4-CHOP-CHAC1 pathway, and reducing expression of CHAC1 can enhance resistance of tumor cell to ferroptosis[39]. NOX4 knock-down could significantly improve remodeling of left ventricular and reduce cardiomyocyte death, meanwhile, knocking down NOX4 could inhibit ferroptosis due to NOX4 mainly involved in hydrogen peroxide and lipid peroxidation in biological processes[40–42]. Wang et al. found that ATF3 could accelerate ferroptosis induced by Erastin by inhibiting the XC-system, which was essential for glutamate synthesis[43]. In previous studies, the four hub genes mentioned above (MYB, CHAC1, NOX4, AFT3) were all promoters for ferroptosis, while MYB and CHAC1 with down-regulation were considered to be tumor protective genes and NOX4 and ATF3 with up-regulation were associated with poor prognosis. Whether these genes play a role in the prognosis of GC patients by influencing the process of iron death remains to be further investigated.

Immunotherapy is a conventional method for tumor treatment. Ferroptosis has been shown to play a significant role in tumor development, but further studies are remained to be explored to uncover the potential mechanisms between ferroptosis and tumor immunity. Therefore, we also analyzed the correlation of tumor immunity between high and low risk groups. In this study, there are significant differences in the content of the antigen presentation process between the high- and low-risk groups. The mechanism may be that ferroptosis cells attract antigen-presenting cells (APCs) to accumulate inside by releasing different signals, such as lipid signals. Furthermore, we found both in TCGA dataset and GSE26901, the content of Macrophages reached a high level in the high-risk groups. Previous studies have also shown that the aggregation of tumor macrophages is closely related to the occurrence and development of tumors, as it may play a role in immune invasion[44, 45]. In addition, the low expression of MHC_class_I in high-risk patients was found in both databases, which may be related to tumor immune impairment.

Combining the characteristics of the four hub genes and clinicopathological parameters, our nomogram can be used by clinicians to determine the prognosis of each patient. Its scoring system is easy to understand and facilitates customized treatments and medical decisions. Our research can provide new insights into the molecular mechanism and prognosis prediction of GC. In addition, the DEGs obtained in this study comes from a comprehensive analysis of multiple data sets and has high reliability. Prior to this study, four genes signature has not yet seen the report related to GC. These genes may be potential molecular targets against GC.

However, we must acknowledge that the study has some limitations. First of all, our main source of clinical information is from the data sets of TCGA and GEO databases, in which most of patients are white, African or Latino. So, we must be cautious when extending our findings to patients of other races. Secondly, the establishment and validation of the nomogram were based on existing database. Therefore, it is necessary to use prospective data with complete clinical information and gene expression information to verify its clinical effectiveness in the future. In addition, it should be emphasized that the link between ferroptosis-related DEGs risk score and immune activity has not been resolved in experiments, which is also one of our regrets.

5. Conclusion

In summary, we have studied a new prognostic signature consisting of four hub genes. This model has been proved to be related to overall survival both in internal and external verifications, and is basically in line with reality, providing a reference for predicting the prognosis of GC and even potential molecular targets. However, the underlying molecular mechanism between ferroptosis-related genes and tumor immunity in gastric cancer is still unclear, so the further experimental research is needed.

Declarations

Funding

This study was supported by the National Natural Science Foundation of China (Grant Number: 81860433 and 82103645), the Natural Science Youth Foundation of Jiangxi Province (Grant Number: 20192BAB215036), the Key Technology Research and Development Program of Jiangxi Province (Grant Number: 20202BBG73024), the Foundation for Fostering Young Scholar of Nanchang University (Grant Number: PY201822).

Conflicts of interest

All authors have completed the ICMJE uniform disclosure form. The authors have no conflicts of interest to declare.

Authors' contributions

Z.Z. and Z.K.N. conceived and designed this study. Z.K.N. and N.T.F collected and assembled the data. J.L. and H.K.T. analysed and interpreted the data. J.L. and H.K.T. drafted the manuscript. C.G.H. and F.L.Z. prepared the figures and tables. All authors read and approved the final manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet. 2020;396(10251):635–648.
3. Venerito M, Link A, Rokkas T, et al. Review: Gastric cancer-Clinical aspects. Helicobacter. 2019;24 Suppl 1:e12643.
4. Gao J, Cao R, Mu H. Long non-coding RNA UCA1 may be a novel diagnostic and predictive biomarker in plasma for early gastric cancer. Int J Clin Exp Pathol. 2015;8(10):12936–12942.
5. Li P, Chen S, Chen H, et al. Using circular RNA as a novel type of biomarker in the screening of gastric cancer. Clin Chim Acta. 2015;444:132–136.

6. Li T, Mo X, Fu L, et al. Molecular mechanisms of long noncoding RNAs on gastric cancer. Oncotarget. 2016;7(8):8601–8612.
7. Yang L, Zou S, Shu C, et al. CYP2A6 Polymorphisms Associate with Outcomes of S-1 Plus Oxaliplatin Chemotherapy in Chinese Gastric Cancer Patients. Genomics Proteomics Bioinformatics. 2017;15(4):255–262.
8. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin. 2015;65(1):5–29.
9. Tang S, Lin L, Cheng J, et al. The prognostic value of preoperative fibrinogen-to-prealbumin ratio and a novel FFC score in patients with resectable gastric cancer. BMC Cancer. 2020;20(1):382.
10. Lim C, Xu JC, Chen TY, et al. Ubiquitin-specific peptide 22 acts as an oncogene in gastric cancer in a son of sevenless 1-dependent manner. Cancer Cell Int. 2020;20:45.
11. Sakaki A, Kanamori J, Sato A, et al. Case report: Gastric tube cancer after esophagectomy-Retrograde perfusion after proximal resection of right gastroepiploic artery. Int J Surg Case Rep. 2019;59:97–100.
12. Dixon SJ, Lemberg KM, Lamprecht MR, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. Cell. 2012;149(5):1060–1072.
13. Yang WS, Stockwell BR. Ferroptosis: Death by Lipid Peroxidation. Trends Cell Biol. 2016;26(3):165–176.
14. Galluzzi L, Bravo-San Pedro JM, Vitale I, et al. Essential versus accessory aspects of cell death: recommendations of the NCCD 2015. Cell Death Differ. 2015;22(1):58–73.
15. Linkermann A, Stockwell BR, Krautwald S, et al. Regulated cell death and inflammation: an auto-amplification loop causes organ failure. Nat Rev Immunol. 2014;14(11):759–767.
16. Fearnhead HO, Vandenabeele P, Vanden Berghe T. How do we fit ferroptosis in the family of regulated cell death? Cell Death Differ. 2017;24(12):1991–1998.
17. Mou Y, Wang J, Wu J, et al. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. J Hematol Oncol. 2019;12(1):34.
18. Niu Y, Zhang J, Tong Y, et al. Physcion 8-O-β-glucopyranoside induced ferroptosis via regulating miR-103a-3p/GLS2 axis in gastric cancer. Life Sci. 2019;237:116893.
19. Ma S, Henson ES, Chen Y, et al. Ferroptosis is induced following siramesine and lapatinib treatment of breast cancer cells. Cell Death Dis. 2016;7(7):e2307.
20. Sun X, Ou Z, Chen R, et al. Activation of the p62-Keap1-NRF2 pathway protects against ferroptosis in hepatocellular carcinoma cells. Hepatology. 2016;63(1):173–184.
21. Liu Z, Zhao Q, Zuo ZX, et al. Systematic Analysis of the Aberrances and Functional Implications of Ferroptosis in Cancer. iScience. 2020;23(7):101302.
22. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284–287.
23. Simon N, Friedman J, Hastie T, et al. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. J Stat Softw. 2011;39(5):1–13.

24. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2011;73(3):273–282.
25. Sturm G, Finotello F, Petitprez F, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics. 2019;35(14):i436-i445.
26. Okines A, Verheij M, Allum W, et al. Gastric cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2010;21 Suppl 5:v50-54.
27. Su X, Zhang J, Yang W, et al. Identification of the Prognosis-Related lncRNAs and Genes in Gastric Cancer. Front Genet. 2020;11:27.
28. Nardone G. Review article: molecular basis of gastric carcinogenesis. Aliment Pharmacol Ther. 2003;17 Suppl 2:75–81.
29. Katz SJ, Morrow M. Addressing overtreatment in breast cancer: The doctors' dilemma. Cancer. 2013;119(20):3584–3588.
30. Fan B, Dachrut S, Coral H, et al. Integration of DNA copy number alterations and transcriptional expression analysis in human gastric cancer. PLoS One. 2012;7(4):e29824.
31. Liu Z, Gao P, Liu S, et al. Tumor volume increases the predictive accuracy of prognosis for gastric cancer: A retrospective cohort study of 3409 patients. Oncotarget. 2017;8(12):18968–18978.
32. Robichaud N, Sonenberg N, Ruggero D, et al. Translational Control in Cancer. Cold Spring Harb Perspect Biol. 2019;11(7).
33. Truitt ML, Ruggero D. New frontiers in translational control of the cancer genome. Nat Rev Cancer. 2017;17(5):332.
34. Li J, Cao F, Yin HL, et al. Ferroptosis: past, present and future. Cell Death Dis. 2020;11(2):88.
35. Tang B, Zhu J, Li J, et al. The ferroptosis and iron-metabolism signature robustly predicts clinical diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. Cell Commun Signal. 2020;18(1):174.
36. Wu G, Wang Q, Xu Y, et al. A new survival model based on ferroptosis-related genes for prognostic prediction in clear cell renal cell carcinoma. Aging (Albany NY). 2020;12(14):14933–14948.
37. Volodko N, Gutor T, Petronchak O, et al. Low infiltration of tumor-associated macrophages in high c-Myb-expressing breast tumors. Sci Rep. 2019;9(1):11634.
38. Hao S, Yu J, He W, et al. Cysteine Dioxygenase 1 Mediates Erastin-Induced Ferroptosis in Human Gastric Cancer Cells. Neoplasia. 2017;19(12):1022–1032.
39. Wang N, Zeng GZ, Yin JL, et al. Artesunate activates the ATF4-CHOP-CHAC1 pathway and affects ferroptosis in Burkitt's Lymphoma. Biochem Biophys Res Commun. 2019;519(3):533–539.
40. Chen X, Xu S, Zhao C, et al. Role of TLR4/NADPH oxidase 4 pathway in promoting cell death through autophagy and ferroptosis during heart failure. Biochem Biophys Res Commun. 2019;516(1):37–43.
41. Poursaitidis I, Wang X, Crighton T, et al. Oncogene-Selective Sensitivity to Synchronous Cell Death following Modulation of the Amino Acid Nutrient Cystine. Cell Rep. 2017;18(11):2547–2556.

42. Wang Z, Ding Y, Wang X, et al. Pseudolaric acid B triggers ferroptosis in glioma cells via activation of Nox4 and inhibition of xCT. Cancer Lett. 2018;428:21–33.
43. Wang L, Liu Y, Du T, et al. ATF3 promotes erastin-induced ferroptosis by suppressing system Xc(). Cell Death Differ. 2020;27(2):662–675.
44. Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. Cell. 2010;141(1):39–51.
45. Tan QL, Zhou CY, Cheng L, et al. Immunotherapy of Bacillus Calmette–Guérin by targeting macrophages against bladder cancer in a NOD/scid IL2Rg–/– mouse model. Mol Med Rep. 2020;22(1):362–370.

Figures

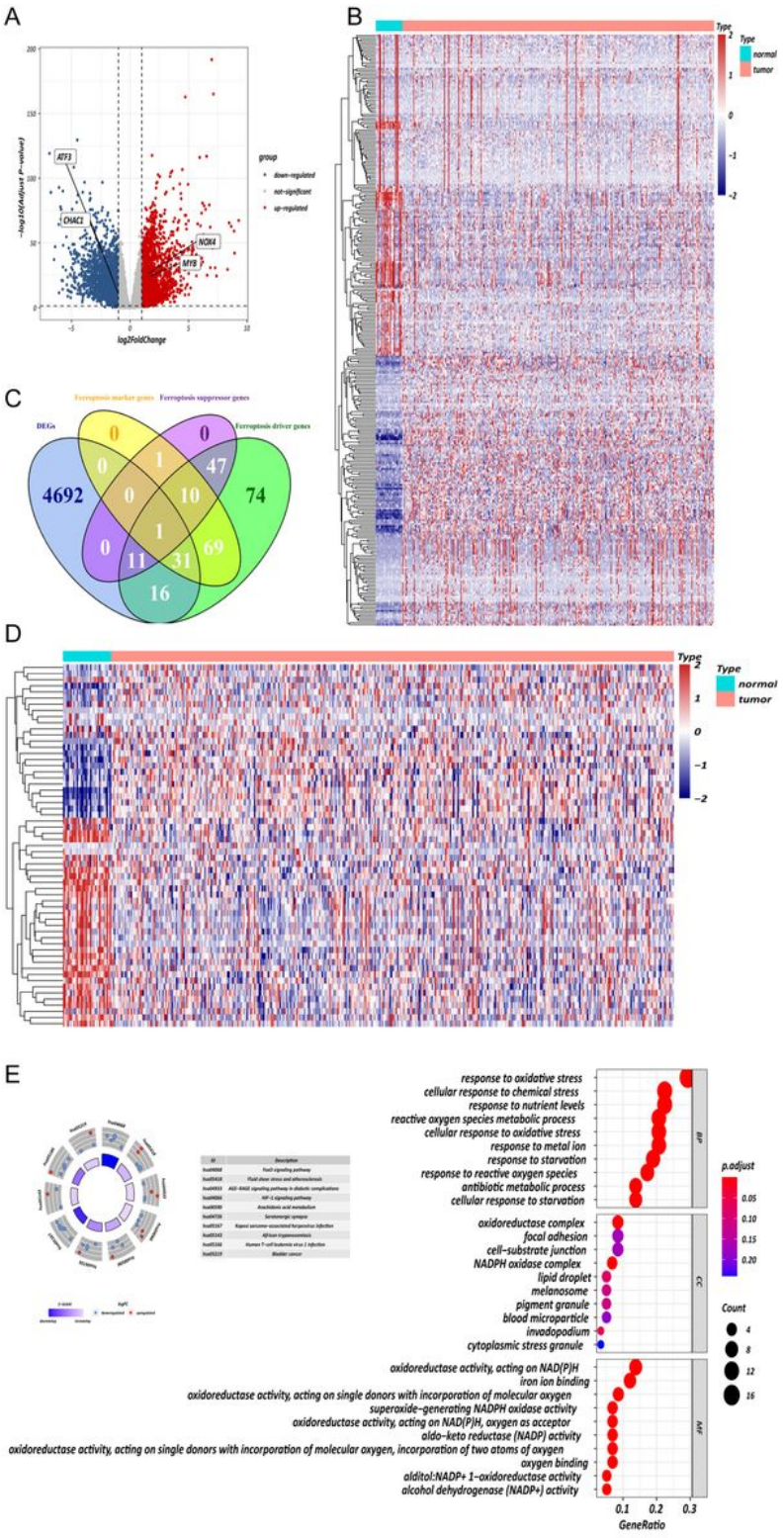

Figure 1

Screening and bioinformatic analysis of ferroptosis-related differentially expressed genes (DEGs). (A) Volcano plot of DEGs. (B) Heatmap of DEGs between normal and tumor tissue. (C) Venn diagram: making intersection of DEGs and ferroptosis-related genes to screen ferroptosis-related DEGs. (D) Heatmap of ferroptosis-related DEGs. (E) GO and KEGG enrichment analysis of ferroptosis-related DEGs.

Figure 2

Survival-related intersection genes (SRIGs) and function enrichment analysis. (A) Forest plot of SRIGs with univariate Cox regression analysis. (B) GO enrichment analysis of 14 SRIGs. (C-D) Visualization of the PPI network and relevance network of SRIGs.

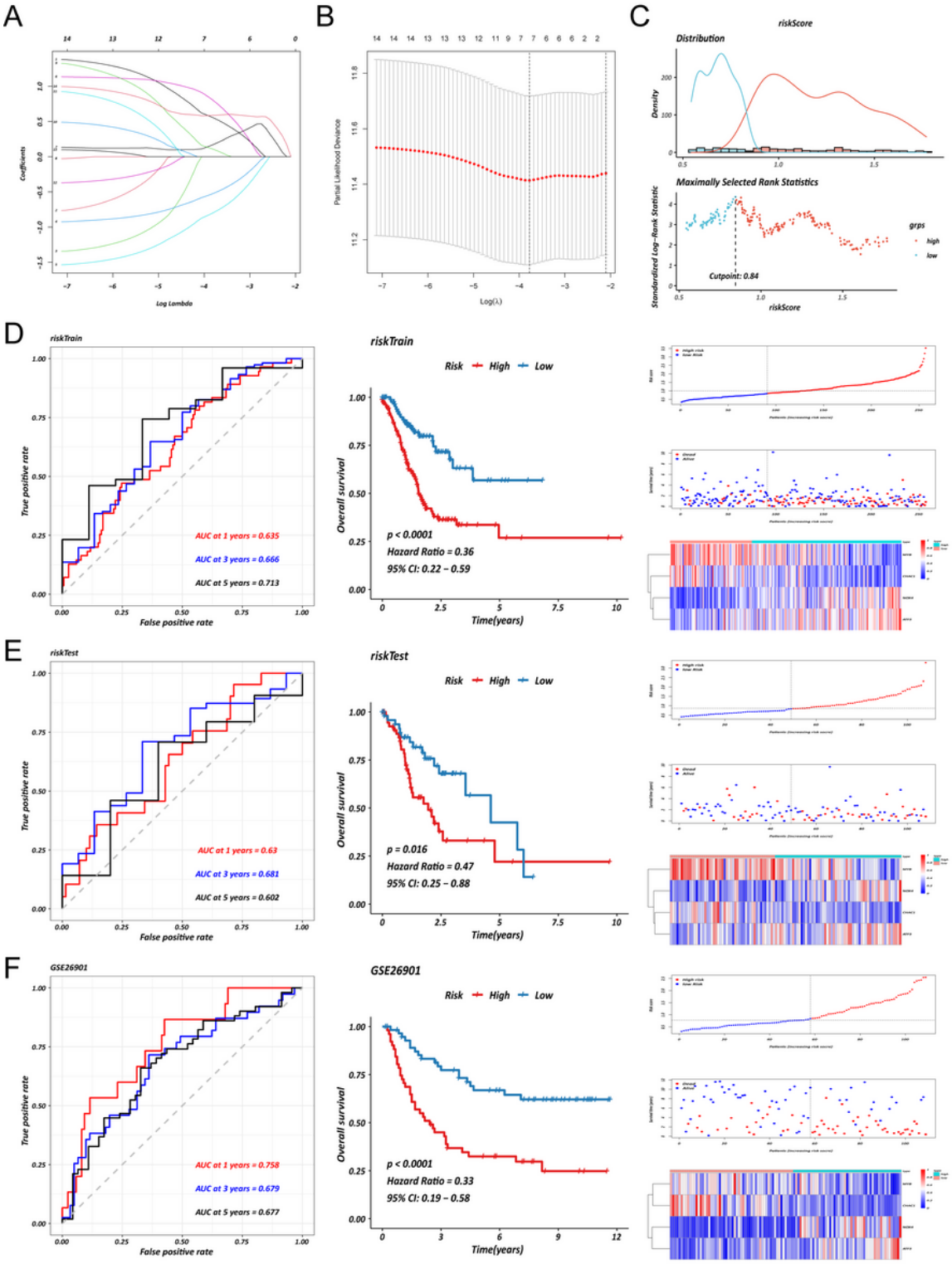

Figure 3

Establishment and verification of the signature. (A-B) Results of lasso regression analysis. (C) Selection of the optimal cut-off value. (D-F) analysis of 3 cohorts (training cohort, testing cohort and GSE26901): time-dependent ROC curves of overall survival (OS) of 1year, 3year and 5year; Kaplan-Meier curves, risk curves and heatmaps of high- and low-risk groups according the optimal cut-off.

Figure 4

Analysis of expression differences including principal component analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE), and single sample gene set enrichment analysis (ssGSEA) of high- and low-risk groups. (A-B) PCA plot and t-SNE plot in TCGA. (C-D) Box plot of ssGSEA in TCGA. (E-F) PCA plot and t-SNE plot in GSE26901. (G-H) Box plot of ssGSEA in GSE26901.

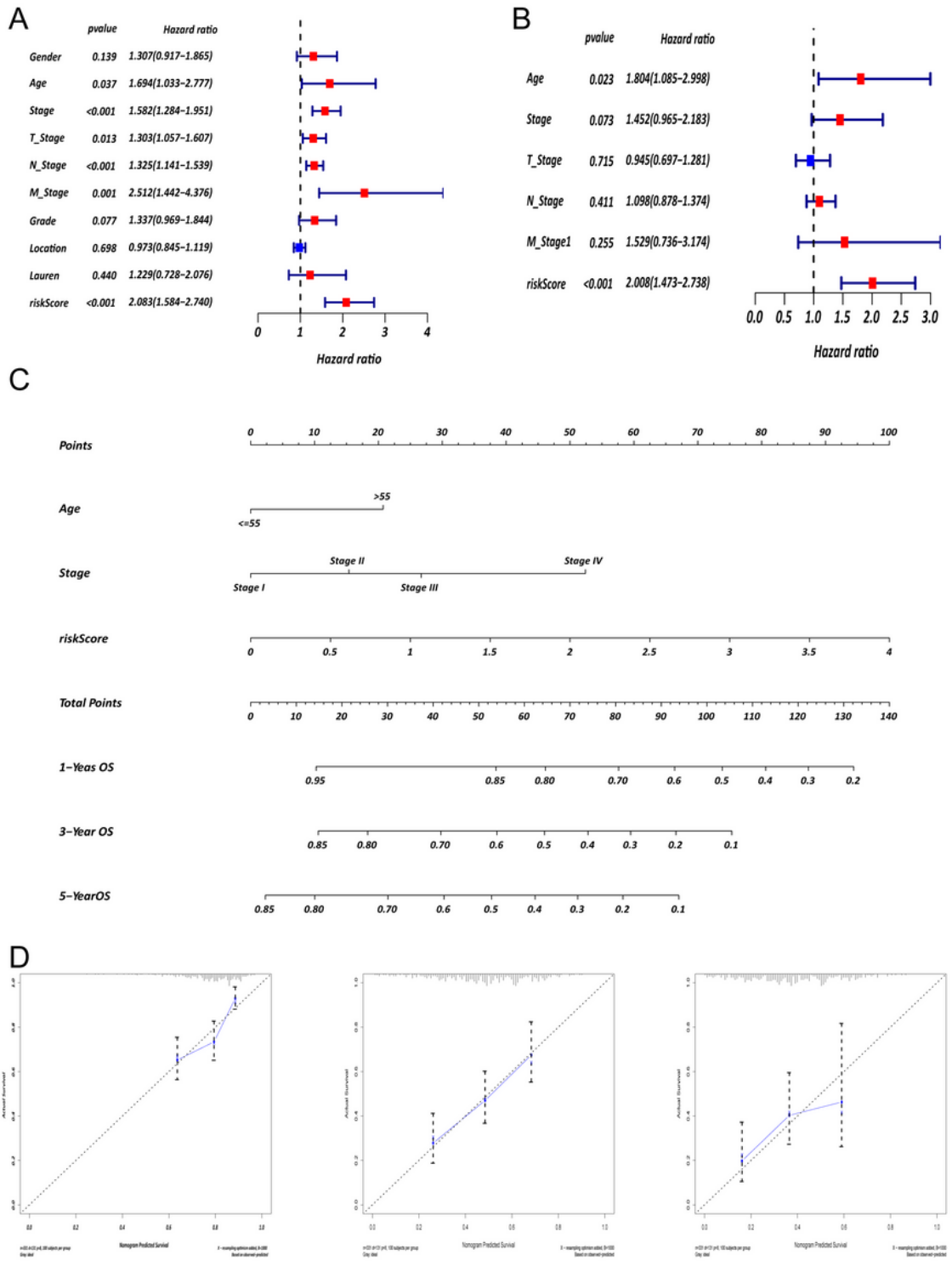

Figure 5

Establishment and validation of nomogram. (A-B) The forest plots of univariate and multivariate Cox regression analysis to identify prognostic factors of gastric cancer. (C) Nomogram of GC patient OS combining the risk score and clinicopathological variables. (D) Calibration curves of the nomogram at 1, 3, and 5 years.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.xls
- SupplementaryTable2.xls
- FIGS1.tif
- FIGS2.tif
- FIGS3.tif