Heliyon 9 (2023) e16157

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Construction and validation of a prognostic prediction model for gastric cancer using a series of genes related to lactate metabolism

Si-yu Wang [a,1], Yu-xin Wang [b,1], Ao Shen [c], Rui Jian [a], Nan An [a], Shu-qiang Yuan [a,*]

[a] Department of Gastric Surgery, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, 510060, China
[b] The First Hospital of Jilin University, Changchun, 130000, China
[c] State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, 510060, PR China

ARTICLE INFO

Keywords:
Gastric cancer
Lactate metabolism
Prognosis
Prediction model
Model construction
Model validation
Model application

ABSTRACT

Background: Gastric cancer (GC) is one of the most common clinical malignant tumors worldwide, with high morbidity and mortality. The commonly used tumor-node-metastasis (TNM) staging and some common biomarkers have a certain value in predicting the prognosis of GC patients, but they gradually fail to meet the clinical demands. Therefore, we aim to construct a prognostic prediction model for GC patients.

Methods: A total of 350 cases were included in the STAD (Stomach adenocarcinoma) entire cohort of TCGA (The Cancer Genome Atlas), including the STAD training cohort of TCGA (n = 176) and the STAD testing cohort of TCGA (n = 174). GSE15459 (n = 191), and GSE62254 (n = 300) were for external validation.

Results: Through differential expression analysis and univariate Cox regression analysis in the STAD training cohort of TCGA, we screened out five genes among 600 genes related to lactate metabolism for the construction of our prognostic prediction model. The internal and external validations showed the same result, that is, patients with higher risk score were associated with poor prognosis (all $p < 0.05$), and our model works well without regard of patients' age, gender, tumor grade, clinical stage or TNM stage, which supports the availability, validity and stability of our model. Gene function analysis, tumor-infiltrating immune cells analysis, tumor microenvironment analysis and clinical treatment exploration were performed to improve the practicability of the model, and hope to provide a new basis for more in-depth study of the molecular mechanism for GC and for clinicians to formulate more reasonable and individualized treatment plans.

Conclusions: We screened out and used five genes related to lactate metabolism to develop a prognostic prediction model for GC patients. The prediction performance of the model is confirmed by a series of bioinformatics and statistical analysis.

1. Introduction

As a serious global health problem, the global morbidity of gastric cancer (GC) ranks fifth and it is the fourth leading cause of

* Corresponding author.
E-mail address: yuanshq@sysucc.org.cn (S.-q. Yuan).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.heliyon.2023.e16157
Received 16 February 2023; Received in revised form 6 May 2023; Accepted 8 May 2023
Available online 10 May 2023

cancer-related death [1], with more than one million people newly diagnosed with it worldwide [2]. The majority of cases of gastric cancer are usually not diagnosed until an advanced stage, usually leading to a poor prognosis [3,4]. Although overall survival (OS) for GC patients has improved greatly due to combined treatment [4,5], it is still difficult to evaluate the prognosis of GC patients due to the great differences in individual prognosis [6].

At present, tumor-node-metastasis (TNM) staging system has been generally applied to predicting the prognosis of GC and more and more prognostic prediction models have also been constructed based on the general data (such as gender and age) and clinicopathological data of GC patients [7]. However, there is a certain lag in time of these data, resulting in its less accuracy.

In addition, a series of traditional biomarkers (such as CEA, CA 12-5, CA 19-9 and CA 72-4) and novel biomarkers (such as microRNAs, long non-coding RNAs and circular RNAs) have been exploited to assist in the diagnosis and prognostic prediction of GC and achieve good clinical effects [8]. Unfortunately, most of novel biomarkers show shortcomings due to limitation by high-throughput technology and some of them are still at the organ level [9]. Some researchers have found several sorts of genes related to GC and explored their prognostic value, but the prognostic value of an individual gene is still not ideal and perfect [10,11]. With the rapid development of precision medicine and gene sequencing technology, multiple genes signature is more and more applied to the construction of prognostic prediction models to achieve more prospective, accurate and individualized prediction performance [12]. Perhaps it will be a promising direction to build related prediction models with the help of this kind of multiple genes signature, as energy metabolism is closely related to tumor cells.

Reprogramming of energy metabolism and evading immune are two emerging hallmarks of cancers [13], and the former is a common feature of most tumor cells [14]. Tumor cells were found to upmake a large amount of glucose and metabolize into lactate under oxygen-rich conditions, known as “Warburg effect” [15]. The same phenomenon was also found in GC cells [16]. Lactate, as a signaling molecule, is discovered to play important roles in the regulation of the metabolic pathways and the immune response within the tumor microenvironment [17], which is closely related to the occurrence, invasion, metastasis, drug resistance and poor prognosis of GC [16]. Sonveaux found that lactate promotes endothelial cell division, thus promoting angiogenesis [18]. Some researchers have successfully constructed corresponding metabolism-related gene signatures for prognostic prediction of gliomas [19,20], GC [21] and lung adenocarcinomas (LUAD) [22]. All the above indicates the availability and feasibility of applying lactate metabolic genes to predicting the prognosis of GC, which is theoretical basis for us to construct a prognostic prediction model for GC patients.

As lactate metabolism is a complex biological process involving multiple genes, the predictive value of those models based on multiple metabolism related genes will be better than those based on a single gene. Therefore, we aim and manage to construct a prognostic prediction model by screening a series of genes related to lactate metabolism to better predict the prognosis of GC patients.

2. Methods

2.1. Acquisition of genes related to lactate metabolism

Genes related to lactate metabolism, known as LM (lactate metabolism) genes, were collected in GeneCards database (https://www.genecards.org/), which provides a wide range of information about human genes. The term “lactate metabolism” was used as a keyword for the search, and sorted by relevance scores, the top 600 genes were considered to be LM genes.

2.2. Collection of datasets

The RNA-seq data and clinical characteristics of TCGA cohorts (ACC: Adrenocortical carcinoma; BLCA: Bladder urothelial carcinoma; BRCA: Breast invasive carcinoma; CESC: Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL: Cholangiocarcinoma; COAD: Colon adenocarcinoma; DLBC: Lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: Esophageal carcinoma; GBM: Glioblastoma multiforme; HNSC: Head and neck squamous cell carcinoma; KICH: Kidney chromophobe; KIRC: Kidney renal clear cell carcinoma; KIRP: Kidney renal papillary cell carcinoma; LAML: Acute myeloid leukemia; LGG: Brain lower grade glioma; LIHC: Liver hepatocellular carcinoma; LUAD: Lung adenocarcinoma; LUSC: Lung squamous cell carcinoma; MESO: Mesothelioma; OV: Ovarian serous cystadenocarcinoma; PAAD: Pancreatic adenocarcinoma; PCPG: Pheochromocytoma and paraganglioma; PRAD: Prostate adenocarcinoma; READ: Rectum adenocarcinoma; SARC: Sarcoma; SKCM: Skin cutaneous melanoma; STAD: Stomach adenocarcinoma; TGCT: Testicular germ cell tumors; THCA: Thyroid carcinoma; THYM: Thymoma; UCEC: Uterine corpus endometrial carcinoma; UCS: Uterine carcinosarcoma; UVM: Uveal melanoma) were downloaded from Xena website (https://xenabrowser.net/datapages/). At the same time, we also downloaded the TCGA Pan-Cancer Datasets of Immune Subtype, Immune Signature Scores (Denise Wolf et al.), Stemness Score (DNA methylation based) and Stemness Score (RNA based) from Xena website. The gene-expression profile matrices of the STAD cohort GEO (Gene Expression Omnibus): GSE15459 and GSE62254 were collected from GEO database (https://www.ncbi.nlm.nih.gov/geo/) for external validation. The “limma” package was used to do the normalization for the expression profiles. The average expression level was retained for repeated genes.

2.3. Differential expression genes analysis and gene function analysis

By using the “limma” package, those LM genes mentioned above were investigated with differential expression analysis in the STAD cohort of TCGA and visualized as heatmap plots. |Fold change (FC)| more than 0 and *p*-value (adj. *p*) less than 0.05 were considered to be statistically apparent for distinguishing the differential expression genes. We selected the intersection of the differential expression genes and LM genes, known as differential expression LM genes, for further analysis.

The analysis of gene function is an important procedure in translating molecular findings from high-throughput data into biological significance. By using the "clusterProfiler" package in R, we perform statistical analysis and visualize functional profiles of LM genes, including Genetic Ontology (GO) annotation analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis. A *p*-value less than 0.05 was considered to be statistically significant.

2.4. The construction of prognostic gene signature

By performing the univariate Cox proportional hazards regression analysis [23] on each differential expression LM gene, we selected genes evidently associated with OS in the STAD training cohort of TCGA. Then, to confirm independence of prognostic factors, a *p*-value less than 0.05 was considered to be statistically significant. Multivariate Cox regression was then conducted to develop the LM-associated predictive model. Utilizing these remaining genes as variables, from these independent prognostic genes and their corresponding coefficients, a signature was established as follow:

$$\text{Risk score} = \sum_{i=1}^{q} (\beta_i \times Z_i)$$

β: coefficients; Z: gene expression level; q: number of genes with nonzero coefficients used in our prediction model.

The "surv_cutpoint" function in the "survminer" package was used to obtain the best cut-off value according to the risk score to divide the STAD cohort of TCGA into high- and low-risk groups. The identical formula and identical cut-off value were selected to two GEO datasets and our own sequencing data for validation.

Performing univariate and multivariate Cox proportional hazards regression analyses to investigate whether the differential expression LM genes-based prognostic model was an independent prognostic factor. To assess the survival differences between two groups, we constructed the Kaplan-Meier (KM) survival curve and used the log-rank test. At the same time, to examine the specificity and sensitivity of the prognostic performance, we explored the receiver operating characteristic (ROC) curve analysis. The area under curve (AUC) values indicated discrimination.

2.5. Gene set enrichment analysis (GSEA)

To find the possible pathways and molecular mechanisms in the high- and low-risk groups in the STAD cohorts of TCGA, we use the analysis of GSEA (https://www.gsea-msigdb.org/gsea/index.jsp). Gene sets were considered significantly enriched with the *p*-value less than 0.05 after 1000 permutation.

2.6. Investigation of tumor-infiltrating immune cells

To explore the relationship between the immune-cell characteristics and the model and genes, we considered the CIBERSORT method to calculate the immune infiltration statues in the samples from the datasets of the STAD cohorts of TCGA. By using the Wilcoxon signed-rank test, we explored the differences in immune infiltrating cell content between high- and low-risk groups of the constructed model.

2.7. Tumor microenvironment analysis

To explore the immune microenvironment of GC patients, we applied the ESTIMATE algorithm in "estimate" and "limma" package of R Software to calculate the Immune Score and Stromal Score. The Estimate Score was used to describe tumor purity. This analysis was based on the gene expression data retrieved from the datasets of the STAD cohorts of TCGA.

2.8. Exploration of the significance of the model in clinical treatment

To explore the model for gastric cancer treatment in the clinic, we calculated the IC_{50} of common approved chemotherapeutic drugs from the datasets of the STAD cohorts of TCGA by using "pRRophetic" package of R Software. The difference between the high- and low-risk groups in the IC_{50} was compared by using Wilcoxon signed-rank test.

2.9. NCI-60 analysis

We downloaded the drug-related data from the NCI-60 database (NCI: National Cancer Institute), which contains data on 60 different cancer cell lines in nine different types of cancer. *F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3* mRNA expression levels and z scores for cell sensitivity data (GI_{50}) were collected from 59 cell lines and were used to explore the relationship between gene expression and drug sensitivity by using Pearson correlation. In the correlation analysis, we analyzed the drug response of 262 drugs approved by FDA (Food and Drug Administration) or on clinical trials.

2.10. Statistical analysis

All statistical analyses were done with R software (version 4.0.3) (https://www.r-project.org/). In our study, the endpoint used in all survival analysis was overall survival (OS). KM estimates for survival functions and log-rank test were used to compare the survival differences between the two groups. The survival curves were drawn by the "survminer" package of R [24]. The Cox proportional hazard model was used to evaluate the association between risk scores and OS [23]. Hazard ratios (HR) and 95% confidence intervals (CI) were evaluated by univariate and multivariate Cox regression models. We used the default Wilcoxon's test when comparing numeric variables in terms of differences between the two groups. The correlation was calculated with Spearman's rank correlation in gene expression levels (*F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*), immune related index (PD1, PDL1, immune score, etc.) and IC_{50}. And a *p*-value less than 0.05 was considered to be statistically significant. The R packages used in our study included "tidyverse" [25], "clusterprofiler" [26,27], "limma" [28], "pheatmap" [29], "glmnet" [30], "survival" [31], "timeROC" [32] and "pRRophetic" [33].

3. Result

3.1. Calculation of the risk score and construction of the prognostic prediction model

A total of 350 cases in the STAD cohort of TCGA were downloaded from Xena as the entire cohort, which were further randomly divided into the STAD training cohort of TCGA (n = 176) and the STAD testing cohort of TCGA (n = 174). The former was for the construction of our prognostic prediction model and the latter was for internal validation (Fig. 1). The details for "Data availability" and "Data split" were provided in the Supplementary Materials.

After search and screening on GeneCards database, we collected 600 LM genes. Among them, 376 differential expression LM genes were screened out after differential expression analysis between cancer and normal tissues in the STAD training cohort of TCGA. The gene-expression profiles of the top 60 of them are displayed in Fig. 2A. After univariate Cox regression analysis, finally, we obtained five genes (*F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*) with nonzero coefficients. The expressions of the five genes were all significantly different between GC and normal tissues in the STAD training cohort of TCGA (all $p < 0.01$, Fig. 2B–F), verifying the rationality and feasibility of using these five genes to develop our prediction model:

$$\text{Risk score} = 0.1496 \times F5 + 0.1838 \times MTTP + 0.1581 \times SERPINE1 + 0.4849 \times CYP19A1 + (-0.3032) \times SLC52 \quad \text{A3}$$

The risk scores of each patient were calculated separately using the above formula, and all these patients were divided into high-

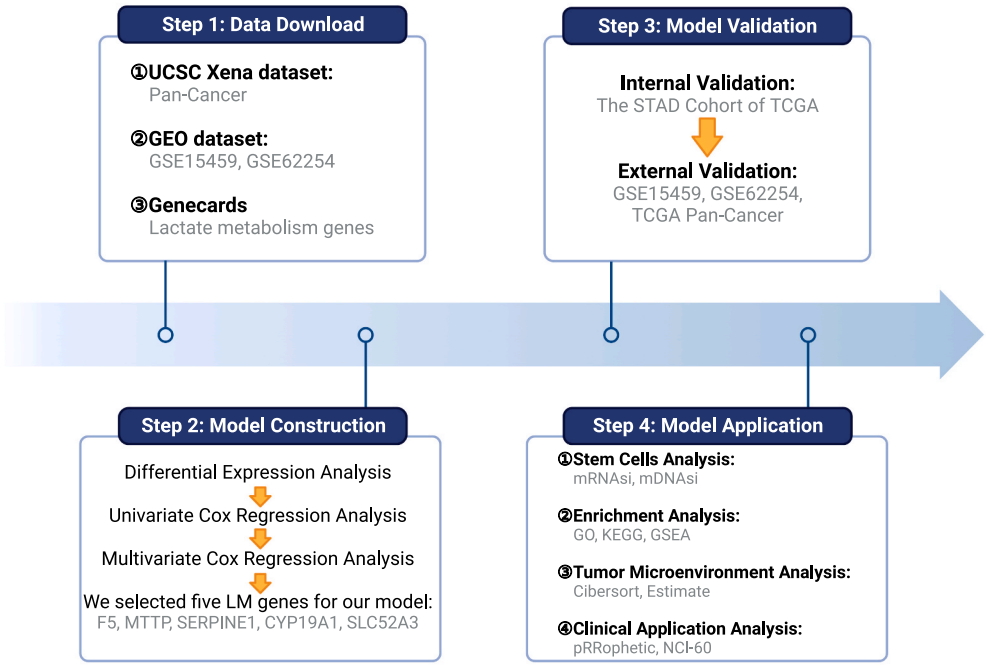


Fig. 1. Overview of the whole study. First, all the data needed was downloaded. Second, the five LM genes for our model were selected. Third, internal validation and external validation were performed. Finally, potential application and clinical significance of our model were analyzed.

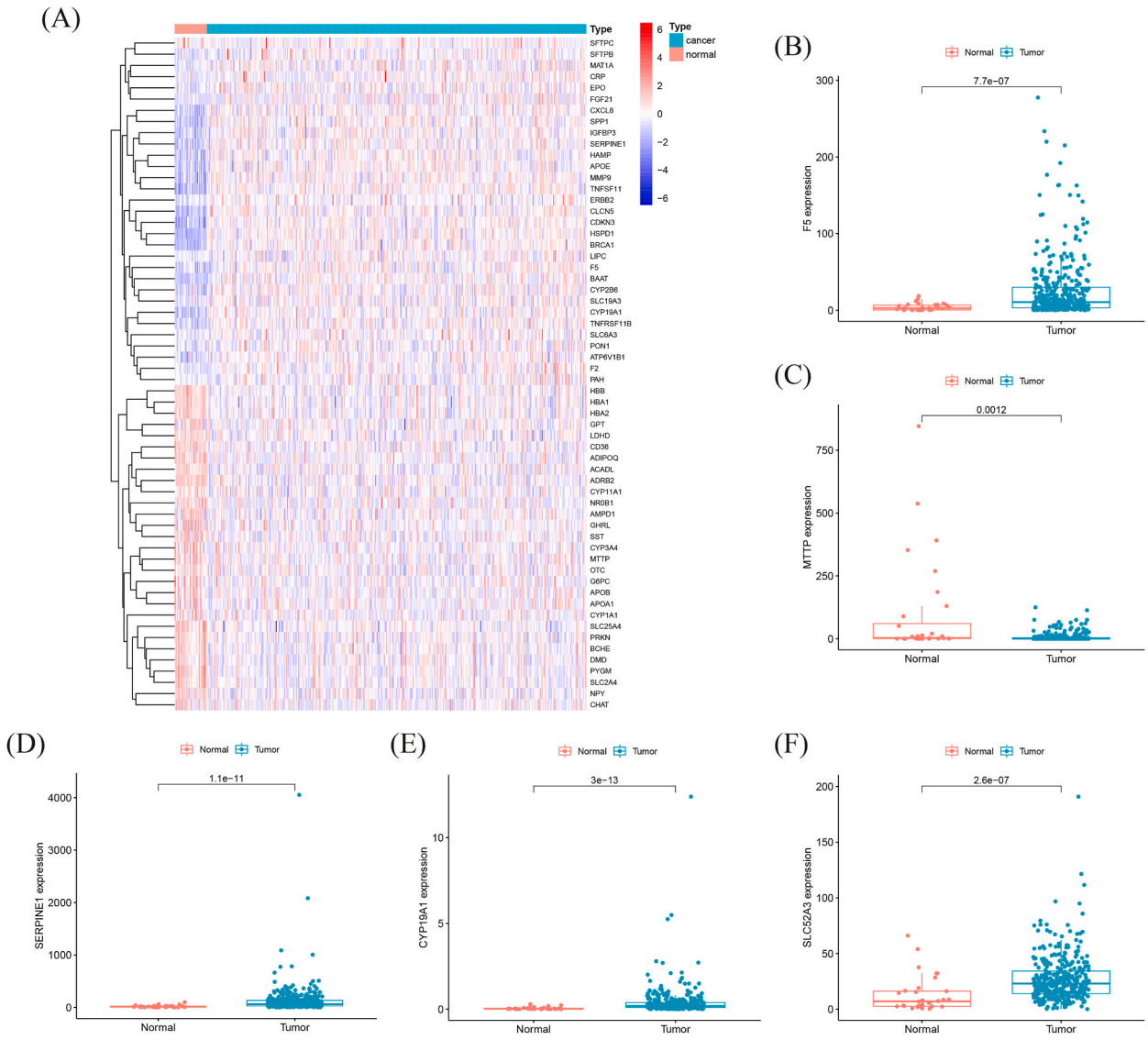


Fig. 2. Screening of the five LM genes. (**A**) The gene-expression profiles of the top 60 differential expression genes. Red indicates genes with higher expression and blue indicates genes with lower expression. (B–F) Gene expressions of *F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3* between GC and normal tissues, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and low-risk group according to the optimal cut-off value we calculated.

3.2. Evaluation and internal validation of the prognostic prediction model

In training cohort, KM estimates for survival functions revealed that patients of high-risk group obviously had worse prognosis compared to those of low-risk group, with significant difference ($p < 0.001$, Fig. 3A). And the corresponding receiver operating characteristic (ROC) curve of 5-year survival was shown in Fig. 3D, with the AUC of 0.763. Univariate and multivariate Cox regression analyses in the STAD training cohort of TCGA showed that the risk score was confirmed to be an independent prognostic factor for patients' survival (both $p < 0.05$, Fig. 4A,D), with the same results in the testing cohort (both $p < 0.05$, Fig. 4B,E) and the entire cohort (both $p < 0.05$, Fig. 4C,F).

For internal validation, the patients in the STAD testing cohort of TCGA were also divided into high- and low-risk group according to the same cut-off value, showing significant difference in KM estimates for survival functions ($p < 0.001$, Fig. 3B) and with 5-year AUC of 0.781 (Fig. 3E). In entire cohort, our prognostic prediction model also had a good performance (Fig. 3C,F). The distributions of the expressions of the five LM genes (*F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*) in the training cohort, the testing cohort and the entire cohort were shown in Fig. 3G–I, respectively.

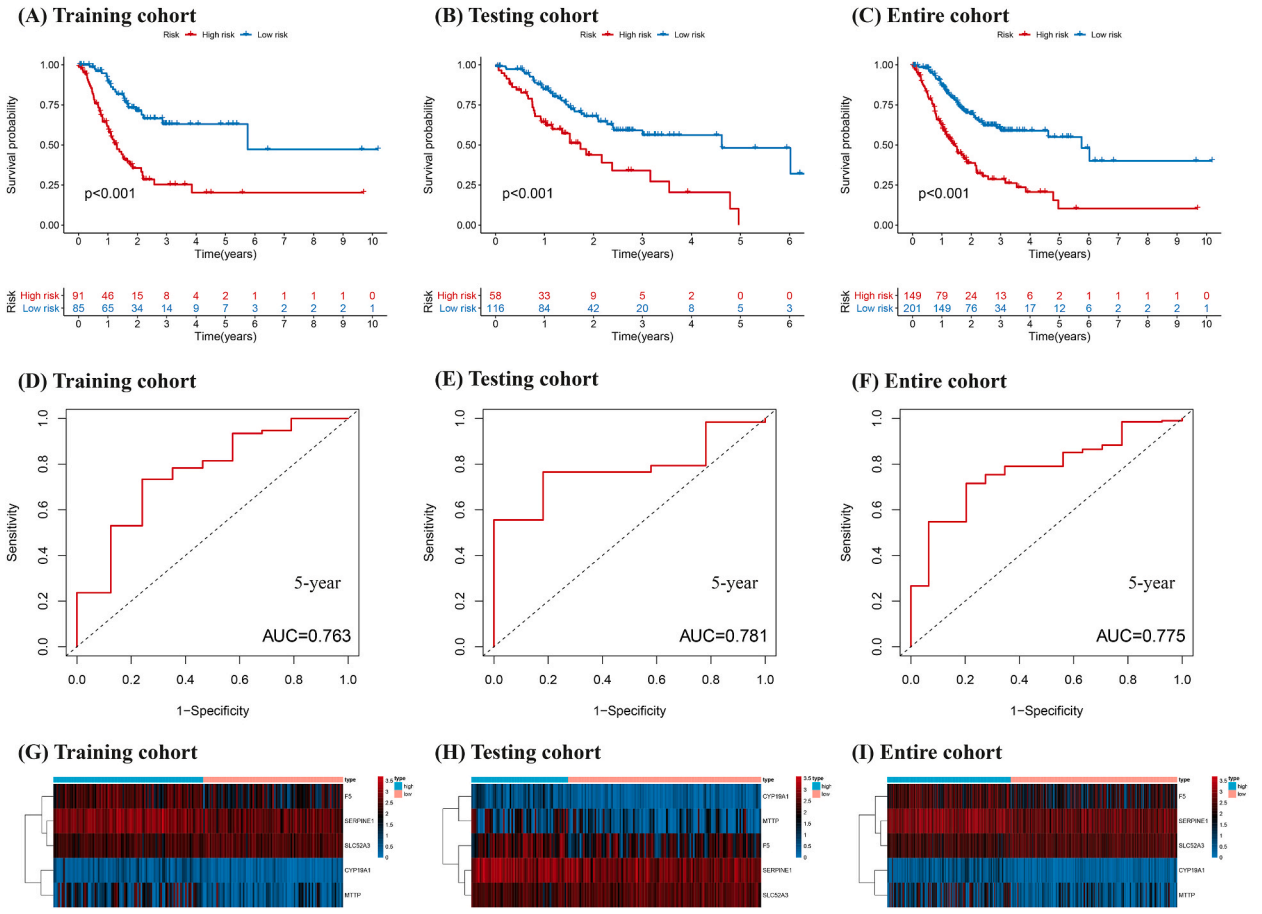


Fig. 3. Kaplan-Meier (KM) estimates for survival functions, receiver operating characteristic (ROC) analysis and heatmaps of prognostic prediction model in (A, D) training cohort, (B, E) testing cohort and (C, F) entire cohort, respectively. The distributions of the expressions of the five LM genes (*F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*) in (G) training cohort, (H) testing cohort and (I) entire cohort, respectively.

3.3. Performance of the prognostic prediction model on different subgroups

All the patients in the STAD entire cohort of TCGA were further divided into two subgroups according to their age (>65 or ≤65 years old), gender (male or female), tumor grade (G1-2 or G3), clinical stage (stage I-II or stage III-IV), T stage (T1-2 or T3-4), N stage (N0 or N1-3) and M stage (M0 or M1). All the corresponding KM curves were shown in Fig. 5A1-G2. It should be specially explained that between high- and low-risk patients with M1, the result of log-rank test did not reach the level of statistical significance ($p = 0.056$), but with obviously separated survival curve. We think it might be caused by lack of numbers of cases. While in other subgroups, there were all significant differences between high- and low-risk groups (all $p < 0.05$), revealing that our prognostic prediction model works well without regard of patients' age, gender, tumor grade, clinical stage or TNM stage, that is, the patients in the high-risk group were usually associated with a poor prognosis.

3.4. External validation of the prognostic prediction model

According to the same cut-off value in the STAD training cohort of TCGA, we did KM estimates for survival functions on two independent datasets: GSE15459 (n = 191), and GSE62254 (n = 300). The results showed that the patients in the high-risk group were significantly associated with lower OS (all $p < 0.05$, Fig. 6A and B), which was consistent with findings in the STAD training cohort of TCGA.

To evaluate the sensitivity and specificity of the prognostic prediction model, external validation in the independent dataset of 32 types of TCGA tumors were performed. According to the results of KM estimates for survival functions, it was found that our signature could significantly distinguish high- and low-risk patients with one of the following 20 types of TCGA tumors respectively: adrenocortical carcinoma (ACC), bladder urothelial carcinoma (BLCA), colon adenocarcinoma (COAD), glioblastoma multiforme (GBM), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), acute myeloid leukemia (LAML), lower grade glioma (LGG), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), mesothelioma (MESO), pancreatic

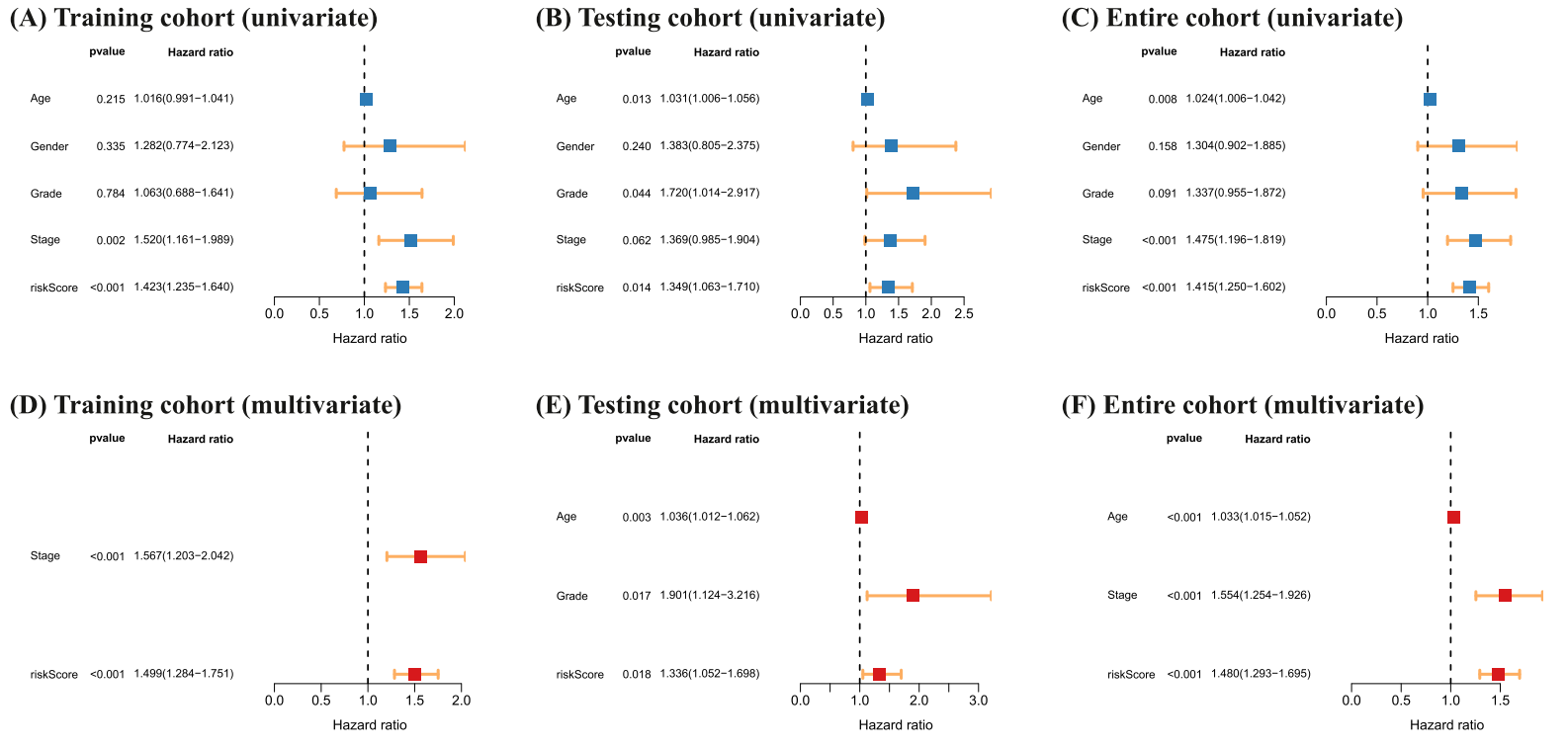


Fig. 4. Univariate and multivariate Cox regression analyses in training, testing and entire cohorts. (A) Training cohort (univariate); (B) Testing cohort (univariate); (C) Entire cohort (univariate); (D) Training cohort (multivariate); (E) Testing cohort (multivariate); (F) Entire cohort (multivariate).

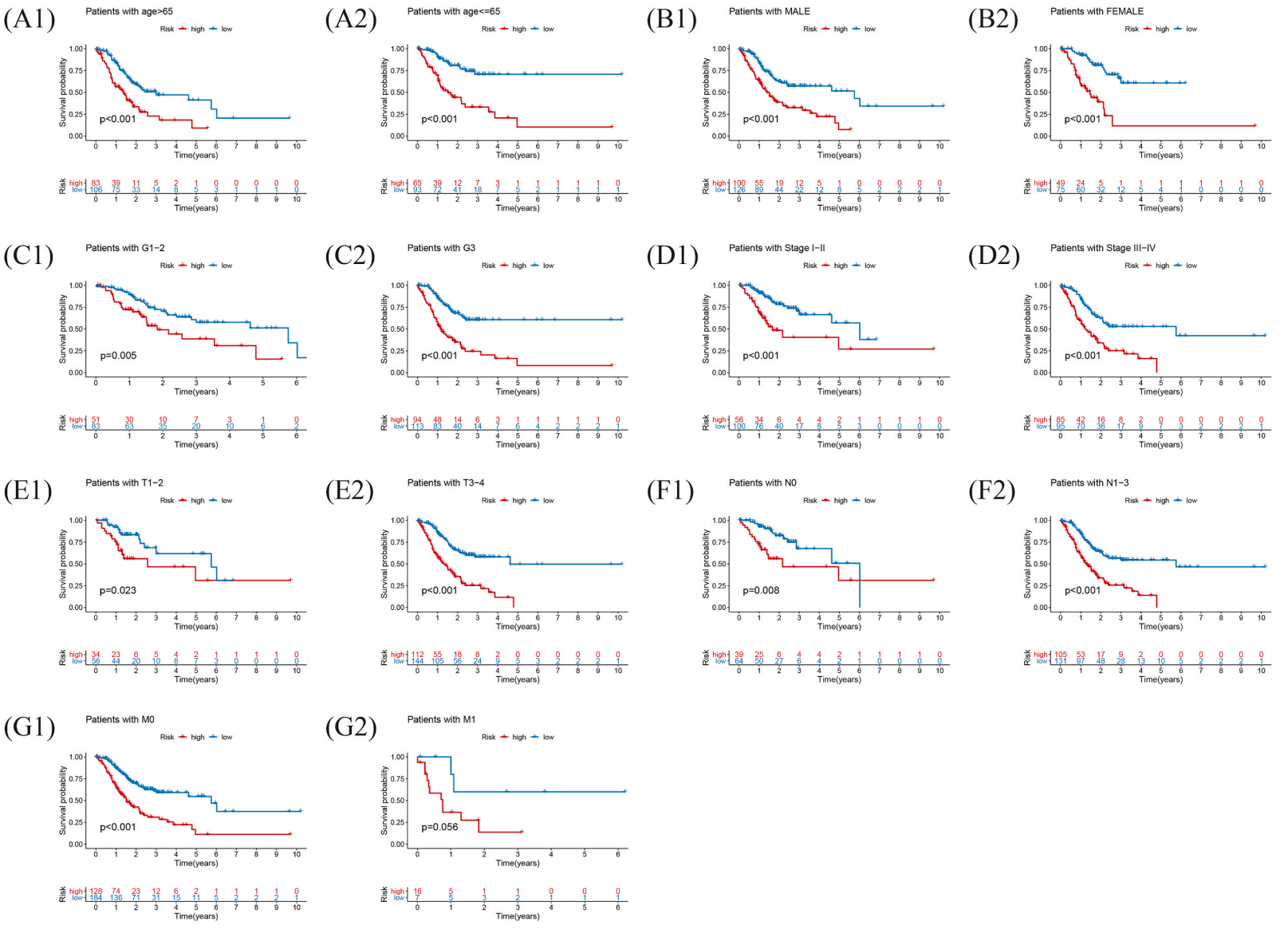


Fig. 5. Kaplan-Meier (KM) estimates for survival functions in different subgroups. (A1, A2) Patients with age >65 or ≤65 years old. (B1, B2) Male or female patients. (C1, C2) Patients with tumor grade of G1-2 or G3. (D1, D2) Patients with stage I-II or stage III-IV. (E1, E2) Patients with T1-2 or T3-4. (F1, F2) Patients or N0 and N1-3. (G1, G2) Patients or M0 and M1.

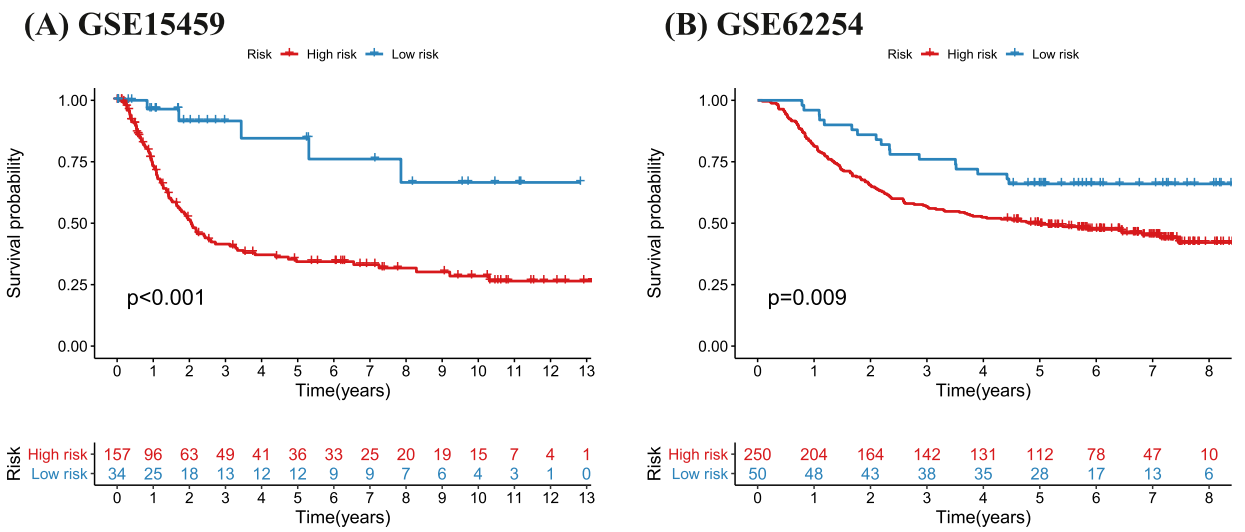


Fig. 6. Kaplan-Meier (KM) estimates for survival functions in independent datasets: (**A**) GSE15459, n = 191; (**B**) GSE62254, n = 300.

adenocarcinoma (PAAD), pheochromocytoma and paraganglioma (PCPG), sarcoma (SARC), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), thymoma (THYM), uterine corpus endometrial carcinoma (UCEC) and uveal melanoma (UVM) (all $p < 0.05$; Fig. S1A-T). Interestingly, the high-risk group has a better prognosis than the low-risk group in some specific types of TCGA tumors. There may be some unknown factors, and further researchers may be needed if necessary.

3.5. Gene function analysis and gene set enrichment analysis (GSEA)

To further explore the potential biological functions of the differential expression LM genes, we performed GO annotation analysis and KEGG enrichment pathway analysis to visualize their functional profiles. The GO analysis showed that the top five enriched terms in BP (biological process) group were "energy derivation by oxidation of organic compounds", "small molecule catabolic process", "ATP metabolic process", "cellular respiration" and "electron transport chain"; the top three enriched terms in CC group were "mitochondrial matrix", "mitochondrial inner membrane" and "mitochondrial protein complex"; the top two enriched terms in MF (molecular function) group were "electron transfer activity" and "tetrapyrrole binding" (Fig. 7A and B). Meanwhile, KEGG analysis showed that the differential expression LM genes were mainly enriched in "thermogenesis", "carbon metabolism" and "oxidative phosphorylation" (Fig. 7D and E). The above 13 enriched terms are all associated with cellular energy metabolism.

Furthermore, the transcript message of patients stratified by risk score into high- and low-risk groups were analyzed by GSEA (Fig. 7C,F). "Collagen containing extracellular matrix", "extracellular matrix structural constituent", "calcium signaling pathway", "focal adhesion" and "neuroactive ligand receptor interaction" were more enriched in the high-risk group, while "mitochondrial membrane organization", "mitochondrial protein complex", "mitochondrial respiratory chain complex assembly" and "mitochondrial transport" were more enriched in the low-risk group, suggesting that mitochondrial function of GC cells in the high-risk group was inhibited to a certain extent.

3.6. Analysis of the five LM genes

To evaluate the impact of the five LM genes on patients' OS individually, the optimal cut-off values of the expression levels of the five LM genes were calculated according to their expression levels separately, and patients in the STAD entire cohort of TCGA were divided into high- and low-expression groups to explore the impact of the five LM genes on patients' survival. The results of KM estimates for survival functions showed that all these five LM genes exerted significant effects on patients' survival individually (all $p < 0.05$, Fig. 8A–E). High expressions of *F5*, *MTTP*, *SERPINE1* and *CYP19A1* may be risk factors for GC, while high expression of *SLC52A3* may be a protective factor, which were consistent with our risk score formula.

3.7. Investigation of tumor-infiltrating immune cells

We also analyzed immune infiltration levels of several specific types of immune cells. Macrophages M1, NK cells activated, T cells CD8 and T cells follicular helper were more enriched in the low-risk group, while Mast cells activated and NK cells resting were more enriched in the high-risk group (all $p < 0.05$, Fig. 9A–F). The KM estimates for survival functions showed that high level of Mast cells activated was associated with a poor prognosis, and the high levels of NK cells activated, T cells CD8 and T cells follicular helper were associated with a better prognosis (both $p < 0.05$, Fig. 9G–J). Unfortunately, Macrophages M2 was also analyzed, but resulted in no significantly difference.

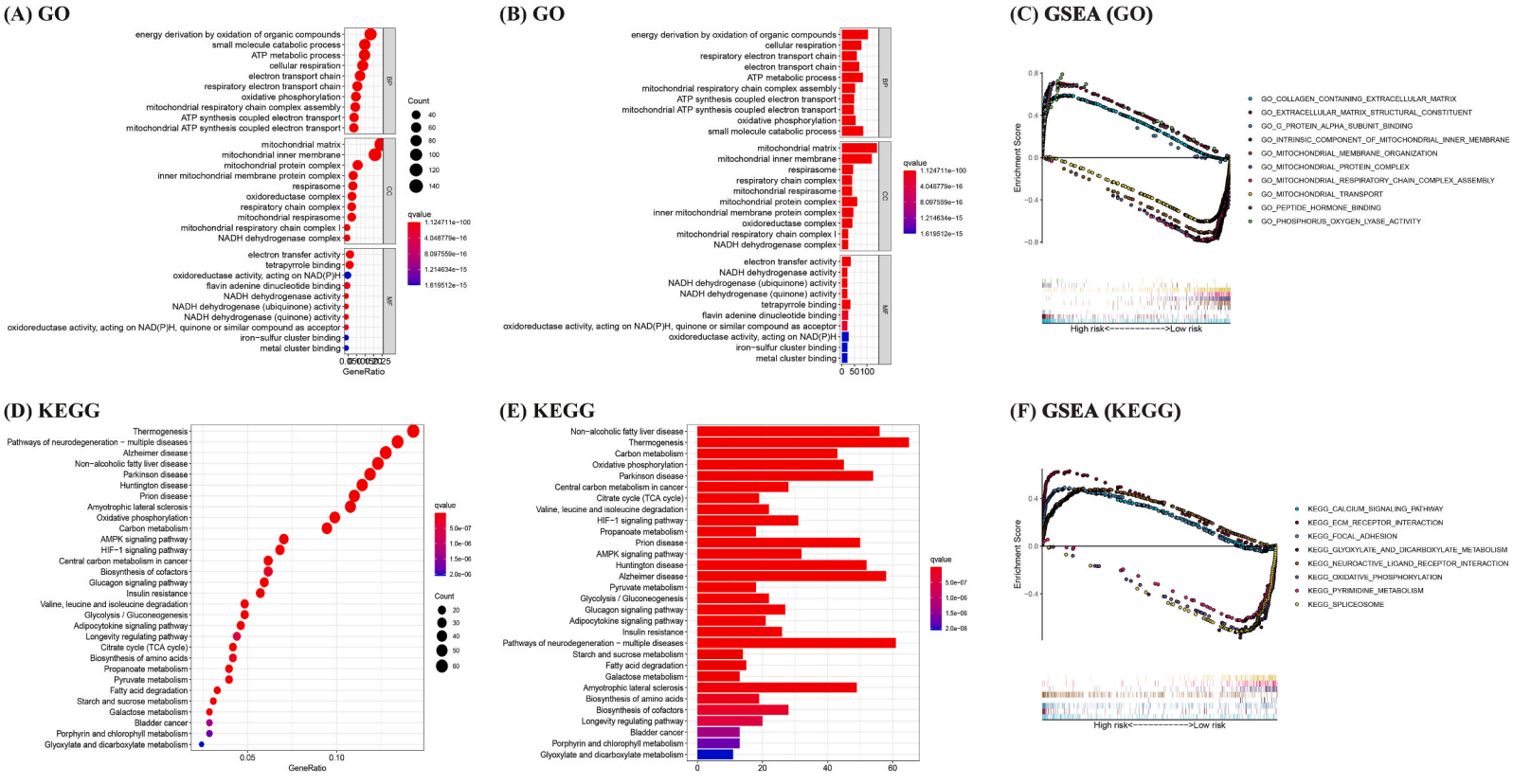


Fig. 7. Gene function analysis and gene set enrichment analysis. (A, B) GO annotation analysis and (D, E) KEGG enrichment pathway analysis of the differential expression LM genes. (C, F) GSEA analysis between high- and low-risk groups. GO: Gene Ontology; KEGG: Kyoto encyclopedia of genes and genomes; GSEA: gene set enrichment analysis; BP: biological process; CC: cellular component; MF: molecular function.

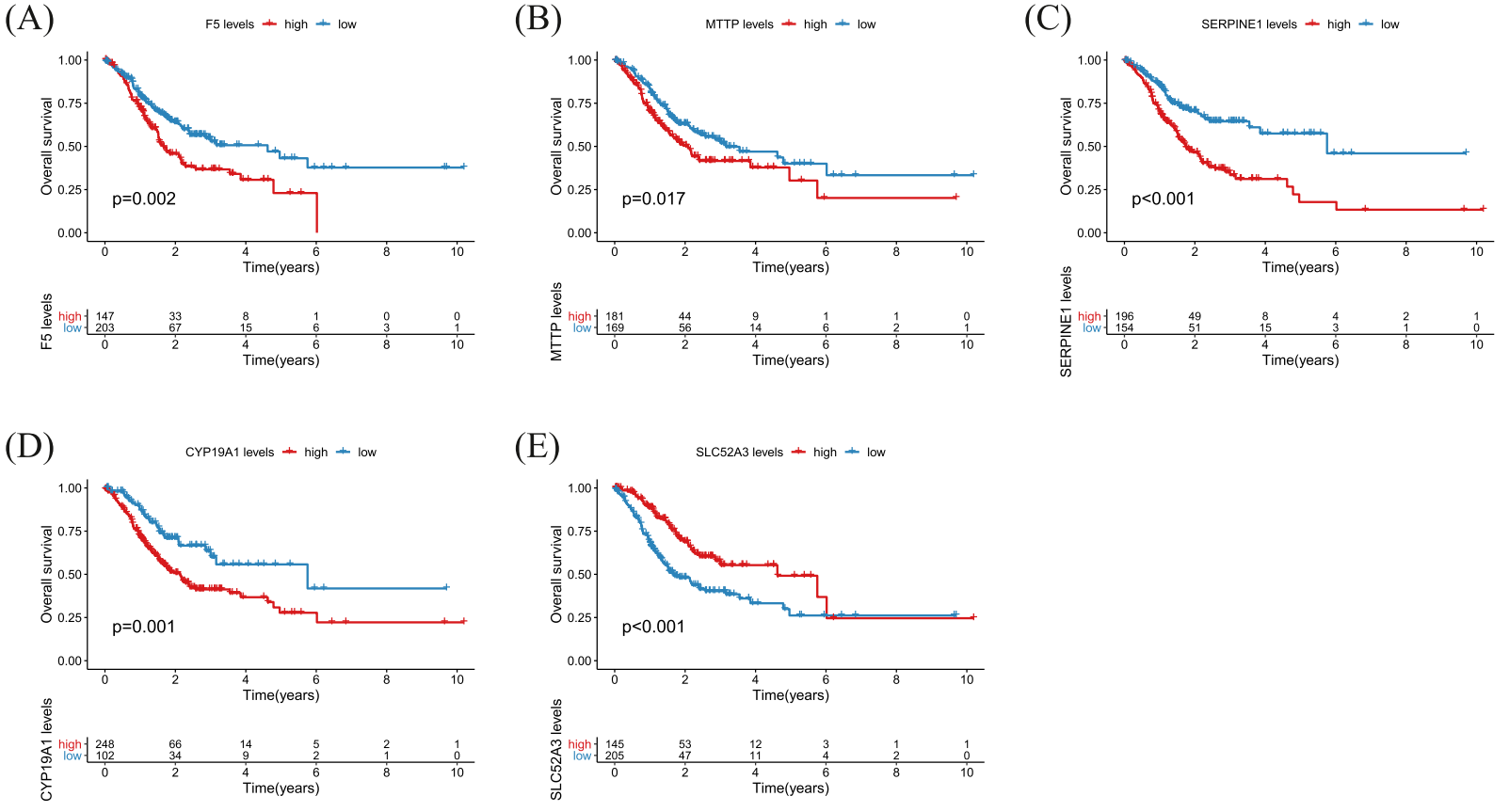


Fig. 8. Impact of the five LM genes on patients' OS. (A–E) Kaplan-Meier estimates for survival functions of *F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*, respectively.

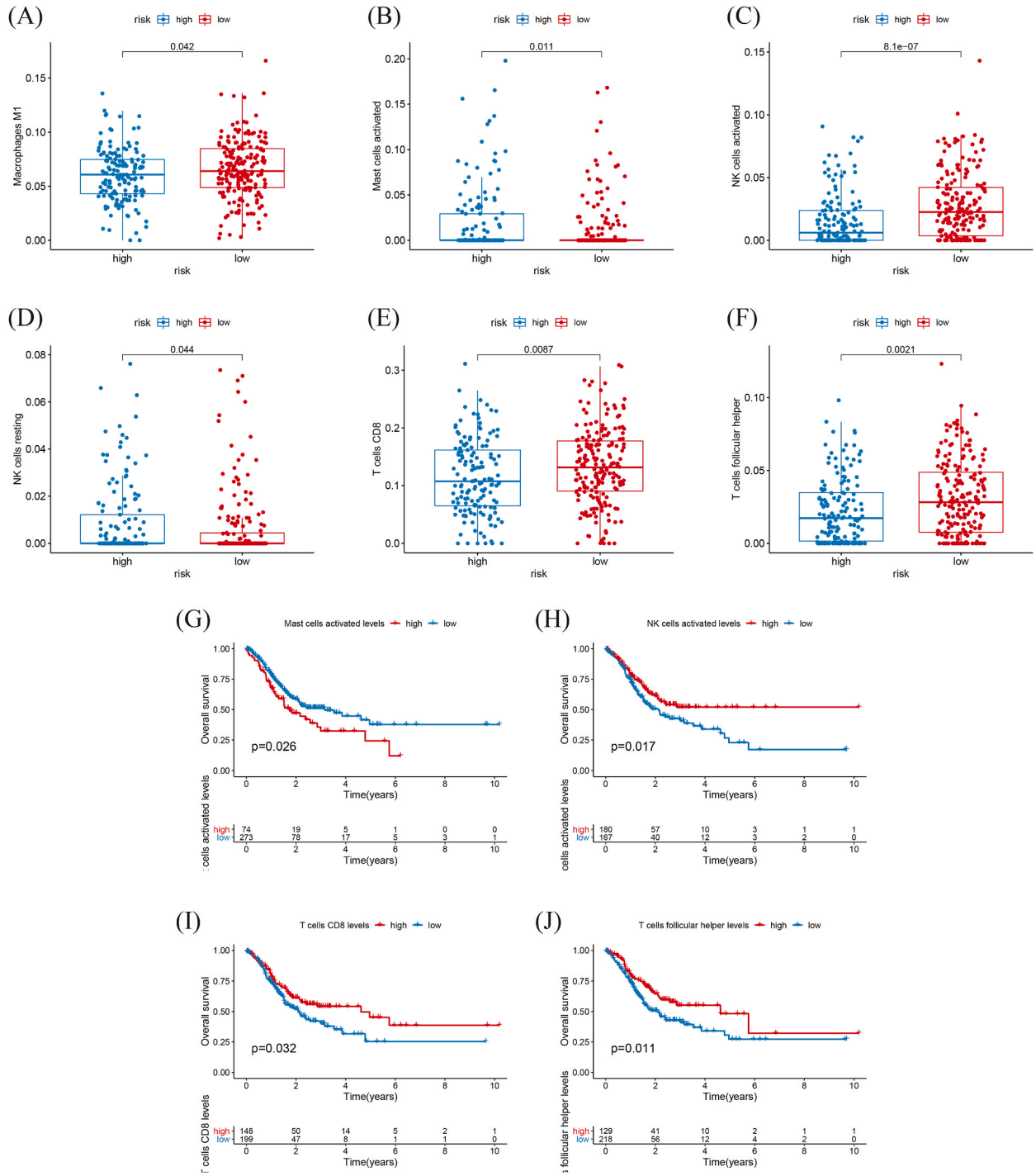
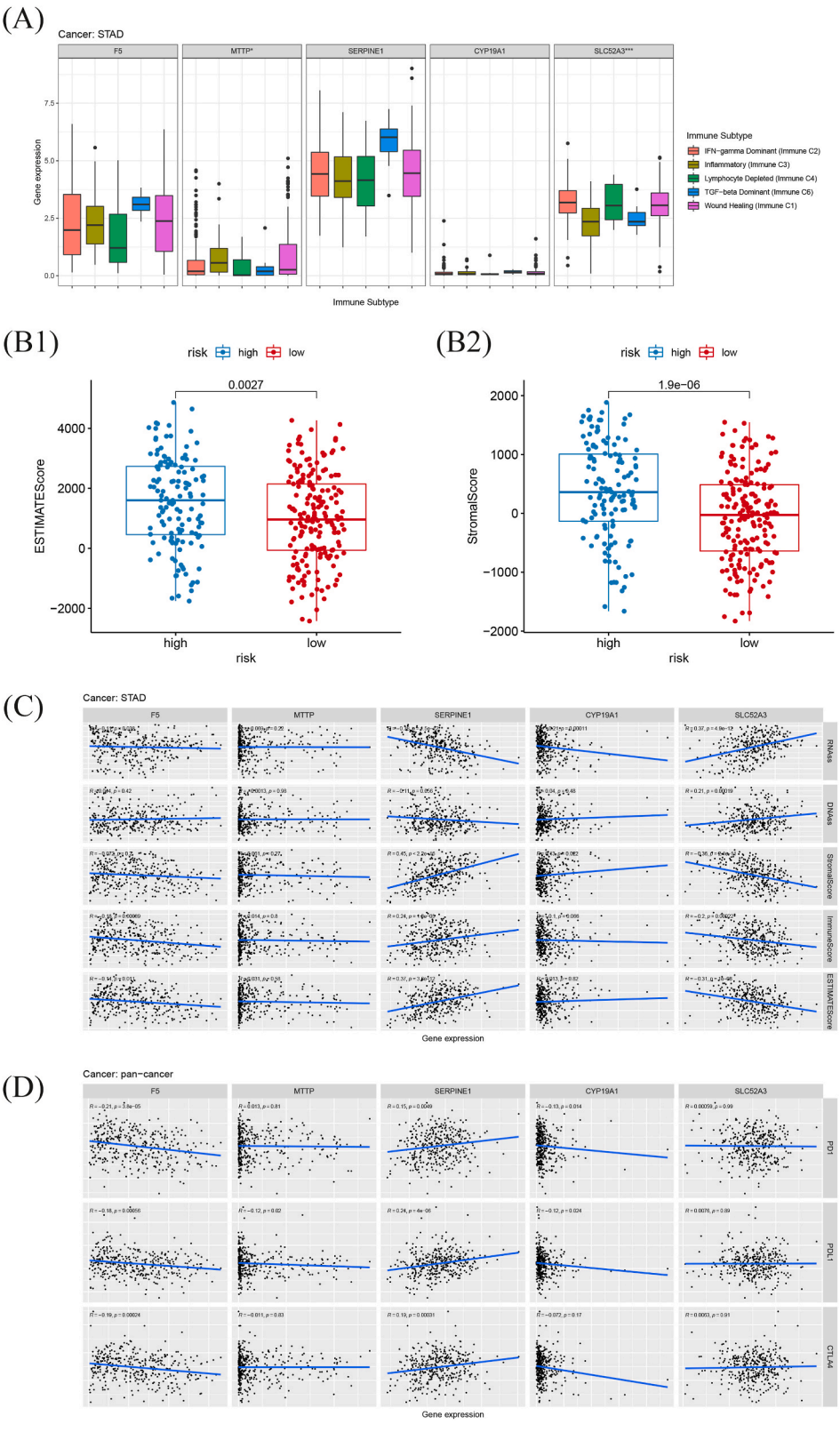


Fig. 9. Investigation of tumor-infiltrating immune cells. (A–F) Immune infiltration levels of Macrophages M1, Mast cells activated, NK cells activated, NK cells resting, T cells CD8 and T cells follicular helper between high and risk groups, respectively. (G–J) Correlation between patients' OS and the levels of Mast cells activated, NK cells activated, T cells CD8 and T cells follicular helper, respectively.

3.8. Tumor microenvironment analysis

To further explore the association of the five LM genes' expressions with immune microenvironment, we further compared their expressions among the five immune subtypes in the STAD entire cohort of TCGA. As shown in Fig. 10A, high expressions of *F5*, *SERPINE1* and *CYP19A1* were associated with Immune C6 (TGF-β dominant), and high expression of *MTTP* was associated with



(caption on next page)

Fig. 10. Tumor microenvironment analysis. (A) Association of the five LM gene expressions with immune infiltrate subtypes. (B) The Estimate Score and Stromal Score between high- and low-risk groups. (C) The correlation between the expressions of the five LM genes and RNAss, DNAss, Stromal Score, Immune Score or Estimate Score. (D) The correlation between the expressions of the five LM genes and immune checkpoint molecules (PD-1, PD-L1 and CTLA4).

Immune C3 (Inflammatory). However, high expression of *SLC52A3* was usually associated with Immune C2 (IFN-γ dominant).

The Estimate Score and Stromal Score calculated by "estimate" package showed the patients in the high-risk group were associated with significantly higher Estimate Score and Stromal Score, indicating their poor prognosis (both $p < 0.001$, Fig. 10B1 and B2). The results of the correlation between the expressions of the five LM genes and RNAss, DNAss, Stromal Score, Immune Score or Estimate Score in the STAD cohort of TCGA showed that higher expression of *SLC52A3* was associated with lower Estimate Score ($p < 0.001$), indicating *SLC52A3* correlated with reduced risk of GC (Fig. 10C).

We also analyzed the correlation between the expressions of the five LM genes and immune checkpoint molecules among TCGA pan-cancer, as shown in the scatter plots of Fig. 10D. The results showed that the higher expression of *SERPINE1* was associated with higher levels of immune checkpoint molecules (PD-1, PD-L1 and CTLA4), indicating that *SERPINE1* might promote tumor development by down-regulating immune response ($p < 0.01$).

3.9. Exploration of the significance of the model in clinical treatment

We analyzed the estimated IC_{50} of the 9 chemotherapeutic drugs (Camptothecin, Cisplatin, Mitomycin C, Vinblastine, Docetaxel, Doxorubicin, Erlotinib, Paclitaxel and Sorafenib), the results showed that the patients in the high-risk group had significantly higher estimated IC_{50} levels of the 4 chemotherapeutic drugs (Camptothecin, Cisplatin, Mitomycin C and Vinblastine, all $p < 0.05$, Figs. S2A–D), revealing their lower drug sensitivity for high-risk patients. However, there were no significant declines of drug sensitivity among Docetaxel, Doxorubicin, Erlotinib, Paclitaxel and Sorafenib (Figs. S3A–E).

Scatter plots were also drawn to study the correlation between the expressions of the five LM genes and drug sensitivity. We found that the expression of the five LM genes were related to increased or decreased drug sensitivity of a number of chemotherapeutic drugs, see Fig. S2E for more details.

4. Discussion

Gastric cancer (GC) is one of the most common clinical malignant tumors worldwide, with high morbidity and mortality. The commonly used TNM staging and some common biomarkers, such as CEA, CA 12-5, CA 19-9 and CA 72-4, have a certain value in predicting the prognosis of GC patients, but they gradually fail to meet the clinical demands at present because of their lag in time and there are great differences in individual prognosis. Therefore, the construction of a prospective, accurate and individualized prognostic prediction model for GC is one of the most urgent clinical tasks at the present stage.

After differential expression analysis and univariate/multivariate Cox regression analysis, we screen out the five LM genes for the construction of our model (*F5*, *MTTP*, *SERPINE1*, *CYP19A1* and *SLC52A3*). In previous studies, *F5* gene was found to be significantly upregulated in GC tissues, and may be a potential prognostic biomarker for GC [34]. *SERPINE1*, as a tumor-promoting gene of gastric cancer [35], its high expression is significantly related to the poor prognosis of GC patients [36], which can be used as a biomarker for the diagnosis and prognosis of GC alone [37]. *CYP19A1* was found to be correlated with the poor prognosis of lung cancer [38] and colon/rectal cancer [39], unfortunately there is no research on GC.

After calculating the risk score and successfully constructing the prognostic prediction model, to ensure its effectiveness and stability, the study first used the STAD testing cohort of TCGA for internal verification, and used three independent datasets for external verification. These results all suggest that patients in the high-risk group were negatively correlated with their OS. We also evaluated the impact of the five LM genes on patients' OS individually, all the results were consistent with our risk score formula.

To study whether the differences in clinical information of patients have an impact on the predictive performance of the model, we further divided the STAD entire cohort of TCGA into several subgroups according to their age, gender, tumor grade, clinical stage and TNM stage. The results of the KM estimates of survival probabilities were all in line with the expected results, showing that our model is not easily disturbed by the above clinical information of patients, suggesting its wide range of clinical predictive value.

For further analysis of the specificity and sensitivity of the model, 32 types of TCGA tumors were selected for external verification of the model. The results of the KM analysis showed that a higher risk score was associated with a lower OS among 22 TCGA tumors. Our explanation for this result is that these tumors may have similar metabolic characteristics and share some same nodes in their metabolic pathways. The identification of these nodes may contribute to a more in-depth study of the mechanism of these tumors and update the existing prognostic prediction models.

In addition, GO analysis and KEGG analysis were performed to further explore the potential biological functions of the differential expression LM genes, and it was found that all these enriched terms are associated with cellular energy metabolism. For deeper study of the biological functions, we performed GSEA to explore the different enrichment pathways between high- and low-risk groups in the STAD entire cohort of TCGA. Interestingly, it was found that all the "mitochondrial-related" pathways were more enriched in the low-risk group, suggesting that mitochondrial function of GC cells in the high-risk group was inhibited to a certain extent, which is consistent with the fact that even under the condition of sufficient oxygen, GC cells still absorb a large amount of glucose and metabolize it into lactate [40].

Importantly, tumor-infiltrating immune cells seems to have more and more significant prognostic value of tumors [41,42], including GC [43]. In our study, Mast cells activated were more enriched in the high-risk group and led to poor prognosis. It is found in previous studies that Mast cells enriched in tumor tissue can produce a large amount of pro-angiogenic factors, which promote angiogenesis and lead to poor prognosis [44,45]. Sammarco found the same phenomenon in GC [46]. Macrophages M1, NK cells activated, T cells CD8 and T cells follicular helper were found to be more abundant in patients of low-risk group, and the latter three significantly affected patients' OS. Husain found tumor-derived lactate inhibits NK cell function [47]. Fischer found tumor cells can increase surrounding lactate concentration and further inhibit the function of cytotoxic T cells [48]. Interestingly, M1 and M2 macrophages were found to have almost opposite effects on tumor patients' prognosis, and the former has been proved to improve patients' OS [49].

As for tumor microenvironment analysis, Baumann found lactate induces TGF-β expression [50], and some other researchers' results showed TGF-β promotes GC malignancy [51] and metastasis [52], which can be suppressed by attenuating TGF-β [53]. It is worth noting in our study that *SERPINE1* and *CYP19A1*, may act as oncogenic genes, were found to associated with Immune C6 (TGF-β dominant), which is consistent with the above researches. *SLC52A3* was found to be associated with Immune C2 (IFN-γ dominant) in our study. Intriguingly, IFN-γ was once recognized as a tumor-killing cytokine, but there is controversial in recent years [54] and its mechanism is not clear. Based on the ESTIMATE algorithm, patients in the high-risk group were associated with significantly higher Estimate Score and Stromal Score, indicating a correlation with poor prognosis. And higher expression of *SLC52A3* was associated with lower Estimate Score, indicating better correlation. Part of the above results, as well as the results of immune checkpoint molecules, are not in line with expectations, which we think may be related to the complexity and polyvalence of metabolic pathways. On the whole, however, these results accord with our expectations.

According to our drug sensitivity analysis, Camptothecin, Cisplatin, Mitomycin C and Vinblastine showed lower drug sensitivity among GC patients in the high-risk group compared with those of low-risk, while there are no significant reduced drug sensitivity among Docetaxel, Doxorubicin, Erlotinib, Paclitaxel and Sorafenib. Meanwhile, the expression of the five LM genes were found to be related to increased or decreased drug sensitivity of a number of chemotherapeutic drugs. These results can provide a new and fresh basis for clinicians to formulate individualized chemotherapy regimens for GC patients in the high-risk group based on our prognostic prediction model.

However, some limitations exist in our study. Most of our data was downloaded from the public database, with the clinical information of some patients incomplete, and there is only a small number of cases from Sun Yat-sen University Cancer Center. At the same time, as a retrospective study, information bias is inevitable [55], so it is necessary to prove the clinical predictive value of the model through more convincing prospective studies in future clinical practice.

5. Conclusion

In this study, we excavated and screened a series of the differential expression LM genes and developed a prognostic prediction model for GC patients based on these genes. Through a series of bioinformatics and statistical analysis, our model has been validated to have a good performance to predict the prognosis of GC patients, and has a great potential to guide clinicians to provide more effective and individualized therapeutic schemes for GC patients. Further prospective studies and ideally laboratory experiments are needed to validate and determine its accuracy and practicability.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Yes.

Data availability

The data could be download at (https://portal.gdc.cancer.gov/, https://xenabrowser.net/, and https://www.ncbi.nlm.nih.gov/geo/; GSE15459 and GSE62254) and the code used during the current study are available from the corresponding author on reasonable request.

Funding

This research was not solicited and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

Si-yu Wang: Investigation, Methodology, Visualization, Software, Roles/Writing - original draft. **Yu-xin Wang:** Data curation, Roles/Writing - original draft, Writing - review & editing. **Ao Shen:** Resources. **Rui Jian, Nan An:** Data curation. **Shu-qiang Yuan:** Conceptualization, Resources, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e16157.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (2021) 209–249, https://doi.org/10.3322/caac.21660. PubMed PMID: 33538338.
[2] A.P. Thrift, H.B. El-Serag, Burden of gastric cancer, Clin. Gastroenterol. Hepatol. 18 (2020) 534–542, https://doi.org/10.1016/j.cgh.2019.07.045. PubMed PMID: 31362118; PubMed Central PMCID: PMC8859863.
[3] H. Katai, T. Ishikawa, K. Akazawa, Y. Isobe, I. Miyashiro, I. Oda, et al., Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001–2007), Gastric Cancer 21 (2018) 144–154, https://doi.org/10.1007/s10120-017-0716-7. PubMed PMID: 28417260.
[4] Z. Tan, Recent advances in the surgical treatment of advanced gastric cancer: a review, Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. 25 (2019) 3537–3541, https://doi.org/10.12659/MSM.916475. PubMed PMID: 31080234; PubMed Central PMCID: PMC6528544.
[5] T. Kiyokawa, T. Fukagawa, Recent trends from the results of clinical trials on gastric cancer surgery, Cancer Commun. 39 (2019) 11, https://doi.org/10.1186/s40880-019-0360-1. PubMed PMID: 30917873; PubMed Central PMCID: PMC6437915.
[6] N.R. Costa, R.M. Gil da Costa, R. Medeiros, A viral map of gastrointestinal cancers, Life Sci. 199 (2018) 188–200, https://doi.org/10.1016/j.lfs.2018.02.025. PubMed PMID: 29476768.
[7] Y. Muneoka, K. Akazawa, T. Ishikawa, H. Ichikawa, A. Nashimoto, H. Yabusaki, et al., Nomogram for 5-year relapse-free survival of a patient with advanced gastric cancer after surgery, Int. J. Surg. 35 (2016) 153–159, https://doi.org/10.1016/j.ijsu.2016.09.080. PubMed PMID: 27664559.
[8] L. Necula, L. Matei, D. Dragu, A.I. Neagu, C. Mambet, S. Nedeianu, et al., Recent advances in gastric cancer early diagnosis, WJG 25 (2019) 2029–2044, https://doi.org/10.3748/wjg.v25.i17.2029. PubMed PMID: 31114131; PubMed Central PMCID: PMC6506585.
[9] D. Ye, G. Xu, W. Ma, Y. Li, W. Luo, Y. Xiao, et al., Significant function and research progress of biomarkers in gastric cancer (Review), Oncol. Lett. (2019), https://doi.org/10.3892/ol.2019.11078. PubMed PMID: 31897111; PubMed Central PMCID: PMC6924079.
[10] K. Guan, X. Liu, J. Li, Y. Ding, J. Li, G. Cui, et al., Expression status and prognostic value of M6A-associated genes in gastric cancer, J. Cancer 11 (2020) 3027–3040, https://doi.org/10.7150/jca.40866. PubMed PMID: 32226518; PubMed Central PMCID: PMC7086255.
[11] J. Pan, H. Zhou, L. Cooper, J. Huang, S. Zhu, X. Zhao, et al., LAYN is a prognostic biomarker and correlated with immune infiltrates in gastric and colon cancers, Front. Immunol. 10 (2019), https://doi.org/10.3389/fimmu.2019.00006. PubMed PMID: 30761122; PubMed Central PMCID: PMC6362421.
[12] C.-T. Yeh, G.-Y. Liao, T. Emura, Sensitivity analysis for survival prognostic prediction with gene selection: a copula method for dependent censoring, Biomedicines 11 (2023) 797, https://doi.org/10.3390/biomedicines11030797. PubMed PMID: 36979776; PubMed Central PMCID: PMC10045003.
[13] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646–674, https://doi.org/10.1016/j.cell.2011.02.013. PubMed PMID: 21376230.
[14] U.E. Martinez-Outschoorn, M. Peiris-Pagés, R.G. Pestell, F. Sotgia, M.P. Lisanti, Cancer metabolism: a therapeutic perspective, Nat. Rev. Clin. Oncol. 14 (2017) 113, https://doi.org/10.1038/nrclinonc.2017.1. PubMed PMID: 28094266.
[15] O. Warburg, F. Wind, E. Negelein, The metabolism of tumors in the body, J. Gen. Physiol. 8 (1927) 519–530, https://doi.org/10.1085/jgp.8.6.519. PubMed PMID: 19872213; PubMed Central PMCID: PMC2140820.
[16] Y. Liu, Z. Zhang, J. Wang, C. Chen, X. Tang, J. Zhu, et al., Metabolic reprogramming results in abnormal glycolysis in gastric cancer: a review, OTT 12 (2019) 1195–1204, https://doi.org/10.2147/OTT.S189687. PubMed PMID: 30863087; PubMed Central PMCID: PMC6389007.
[17] L. Ippolito, A. Morandi, E. Giannoni, P. Chiarugi, Lactate: a metabolic driver in the tumour landscape, Trends Biochem. Sci. 44 (2019) 153–166, https://doi.org/10.1016/j.tibs.2018.10.011. PubMed PMID: 30473428.
[18] P. Sonveaux, T. Copetti, C.J. De Saedeleer, F. Végran, J. Verrax, K.M. Kennedy, et al., Targeting the lactate transporter MCT1 in endothelial cells inhibits lactate-induced HIF-1 activation and tumor angiogenesis, PLoS One 7 (2012), e33418, https://doi.org/10.1371/journal.pone.0033418. PubMed PMID: 22428047; PubMed Central PMCID: PMC3302812.
[19] Y.-Q. Liu, R.-C. Chai, Y.-Z. Wang, Z. Wang, X. Liu, F. Wu, et al., Amino acid metabolism-related gene expression-based risk signature can better predict overall survival for glioma, Cancer Sci. 110 (2019) 321–333, https://doi.org/10.1111/cas.13878. PubMed PMID: 30431206; PubMed Central PMCID: PMC6317920.
[20] F. Wu, Z. Zhao, R. Chai, Y. Liu, G. Li, H. Jiang, et al., Prognostic power of a lipid metabolism gene panel for diffuse gliomas, J. Cell Mol. Med. 23 (2019) 7741–7748, https://doi.org/10.1111/jcmm.14647. PubMed PMID: 31475440; PubMed Central PMCID: PMC6815778.
[21] T. Luo, Y. Li, R. Nie, C. Liang, Z. Liu, Z. Xue, et al., Development and validation of metabolism-related gene signature in prognostic prediction of gastric cancer, Comput. Struct. Biotechnol. J. 18 (2020) 3217–3229, https://doi.org/10.1016/j.csbj.2020.09.037. PubMed PMID: 33209209; PubMed Central PMCID: PMC7649605.
[22] L. He, J. Chen, F. Xu, J. Li, J. Li, Prognostic implication of a metabolism-associated gene signature in lung adenocarcinoma, Mol. Ther.- Oncolytics 19 (2020) 265–277, https://doi.org/10.1016/j.omto.2020.09.011. PubMed PMID: 33209981; PubMed Central PMCID: PMC7658576.
[23] T. Emura, S. Matsui, H.-Y. Chen, compound.Cox: Univariate feature selection and compound covariate for predicting survival, Comput. Methods Progr. Biomed. 168 (2019) 21–37, https://doi.org/10.1016/j.cmpb.2018.10.020. PubMed PMID: 30527130.

[24] A. Kassambara, M. Kosinski, P. Biecek, survminer: Drawing survival curves using 'ggplot2', R package version 0.4.8, https://CRAN.R-project.org/package=survminer, 2020.
[25] H. Wickham, M. Averick, J. Bryan, Welcome to the {tidyverse}, J. Open Source Softw. 4 (2019) 1686. https://CRAN.R-project.org/package=tidyverse.
[26] G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R Package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (2012) 284–287, https://doi.org/10.1089/omi.2011.0118. PubMed PMID: 22455463; PubMed Central PMCID: PMC3339379.
[27] T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, et al., clusterProfiler 4.0: a universal enrichment tool for interpreting omics data, Innovation 2 (2021), 100141, https://doi.org/10.1016/j.xinn.2021.100141. PubMed PMID: 34557778; PubMed Central PMCID: PMC8454663.
[28] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47, https://doi.org/10.1093/nar/gkv007. PubMed PMID: 25605792; PubMed Central PMCID: PMC4402510.
[29] R. Kolde, pheatmap, Pretty Heatmaps, R package version 1.0.12. https://CRAN.R-project.org/package=pheatmap, 2019.
[30] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Software 33 (2010) 1–22. https://CRAN.R-project.org/package=glmnet.
[31] T.M. Therneau, A Package for Survival Analysis in R, 2023. R package version 3.5-5. https://CRAN.R-project.org/package=survival.
[32] P. Blanche, J.F. Dartigues, H. Jacqmin-Gadda, Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks, Stat. Med. 32 (2013) 5381–5397. https://CRAN.R-project.org/package=timeROC.
[33] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, PLoS One 9 (2014), e107468, https://doi.org/10.1371/journal.pone.0107468. PubMed PMID: 25229481; PubMed Central PMCID: PMC4167990.
[34] Y. Liu, X.-W. Liao, Y.-Z. Qin, X.-W. Mo, S.-S. Luo, Identification of F5 as a prognostic biomarker in patients with gastric cancer, BioMed Res. Int. 2020 (2020) 1–13, https://doi.org/10.1155/2020/9280841. PubMed PMID: 32190689; PubMed Central PMCID: PMC7064826.
[35] J.-D. Yang, L. Ma, Z. Zhu, SERPINE1 as a cancer-promoting gene in gastric adenocarcinoma: facilitates tumour cell proliferation, migration, and invasion by regulating EMT, J. Chemother. 31 (2019) 408–418, https://doi.org/10.1080/1120009X.2019.1687996. PubMed PMID: 31724495.
[36] L. Li, Z. Zhu, Y. Zhao, Q. Zhang, X. Wu, B. Miao, et al., FN1, SPARC, and SERPINE1 are highly expressed and significantly related to a poor prognosis of gastric adenocarcinoma revealed by microarray and bioinformatics, Sci. Rep. 9 (2019), https://doi.org/10.1038/s41598-019-43924-x. PubMed PMID: 31127138; PubMed Central PMCID: PMC6534579.
[37] P. Liao, W. Li, R. Liu, J.K. Teer, B. Xu, W. Zhang, et al., Genome-scale analysis identifies SERPINE1 and SPARC as diagnostic and prognostic biomarkers in gastric cancer, OTT 11 (2018) 6969–6980, https://doi.org/10.2147/OTT.S173934. PubMed PMID: 30410354; PubMed Central PMCID: PMC6199229.
[38] J. Zhang, Y. Yin, X.-M. Niu, Y. Liu, D. Garfield, S.-F. Chen, et al., CYP19A1 gene polymorphisms and risk of lung cancer, J. Int. Med. Res. 41 (2013) 735–742, https://doi.org/10.1177/0300060513477291. PubMed PMID: 23669293.
[39] M.L. Slattery, A. Lundgreen, J.S. Herrick, S. Kadlubar, B.J. Caan, J.D. Potter, et al., Variation in the CYP19A1 gene and risk of colon and rectal cancer, Cancer Causes Control 22 (2011) 955–963, https://doi.org/10.1007/s10552-011-9768-x. PubMed PMID: 21479914; PubMed Central PMCID: PMC3225228.
[40] O. Warburg, The chemical constitution of respiration ferment, Science 68 (1928) 437–443, https://doi.org/10.1126/science.68.1767.437. PubMed PMID: 17782077.
[41] H. Angell, J. Galon, From the immune contexture to the Immunoscore: the role of prognostic and predictive immune markers in cancer, Curr. Opin. Immunol. 25 (2013) 261–267, https://doi.org/10.1016/j.coi.2013.03.004. PubMed PMID: 23579076.
[42] W.H. Fridman, F. Pagès, C. Sautès-Fridman, J. Galon, The immune contexture in human tumours: impact on clinical outcome, Nat. Rev. Cancer 12 (2012) 298–306, https://doi.org/10.1038/nrc3245. PubMed PMID: 22419253.
[43] D. Zeng, R. Zhou, Y. Yu, Y. Luo, J. Zhang, H. Sun, et al., Gene expression profiles for a prognostic immunoscore in gastric cancer, Br. J. Surg. 105 (2018) 1338–1348, https://doi.org/10.1002/bjs.10871. PubMed PMID: 29691839; PubMed Central PMCID: PMC6099214.
[44] K. Mukai, M. Tsai, H. Saito, S.J. Galli, Mast cells as sources of cytokines, chemokines, and growth factors, Immunol. Rev. 282 (2018) 121–150, https://doi.org/10.1111/imr.12634. PubMed PMID: 29431212; PubMed Central PMCID: PMC5813811.
[45] G. Paolino, P. Corsetti, E. Moliterni, S. Corsetti, D. Didona, M. Albanesi, et al., Mast cells and cancer, G. Ital. Dermatol. Venereol. 154 (2019), https://doi.org/10.23736/S0392-0488.17.05818-7. PubMed PMID: 29192477.
[46] G. Sammarco, G. Varricchi, V. Ferraro, M. Ammendola, M. De Fazio, D.F. Altomare, et al., Mast cells, angiogenesis and lymphangiogenesis in human gastric cancer, IJMS 20 (2019) 2106, https://doi.org/10.3390/ijms20092106. PubMed PMID: 31035644; PubMed Central PMCID: PMC6540185.
[47] Z. Husain, Y. Huang, P. Seth, V.P. Sukhatme, Tumor-derived lactate modifies antitumor immune response: effect on myeloid-derived suppressor cells and NK cells, J. Immunol. 191 (2013) 1486–1495, https://doi.org/10.4049/jimmunol.1202702. PubMed PMID: 23817426.
[48] K. Fischer, P. Hoffmann, S. Voelkl, N. Meidenbauer, J. Ammer, M. Edinger, et al., Inhibitory effect of tumor cell–derived lactic acid on human T cells, Blood 109 (2007) 3812–3819, https://doi.org/10.1182/blood-2006-07-035972. PubMed PMID: 17255361.
[49] J. Jackute, M. Zemaitis, D. Pranys, B. Sitkauskiene, S. Miliauskas, S. Vaitkiene, et al., Distribution of M1 and M2 macrophages in tumor islets and stroma in relation to prognosis of non-small cell lung cancer, BMC Immunol. 19 (2018), https://doi.org/10.1186/s12865-018-0241-4. PubMed PMID: 29361917; PubMed Central PMCID: PMC5781310.
[50] F. Baumann, P. Leukel, A. Doerfelt, C.P. Beier, K. Dettmer, P.J. Oefner, et al., Lactate promotes glioma migration by TGF-β2–dependent regulation of matrix metalloproteinase-2, Neuro Oncol. 11 (2009) 368–380, https://doi.org/10.1215/15228517-2008-106. PubMed PMID:19033423; PubMed Central PMCID: PMC2743217.
[51] J.-Y. Im, D.-M. Kim, H. Park, M.-J. Kang, D.-Y. Kim, K.Y. Chang, et al., VGLL1 phosphorylation and activation promotes gastric cancer malignancy via TGF-β/ERK/RSK2 signaling, Biochim. Biophys. Acta Mol. Cell Res. 1868 (2021), 118892, https://doi.org/10.1016/j.bbamcr.2020.118892. PubMed PMID: 33069758.
[52] X. Zhang, P. Zhang, M. Shao, X. Zang, J. Zhang, F. Mao, et al., SALL4 activates TGF-β/SMAD signaling pathway to induce EMT and promote gastric cancer metastasis, CMAR 10 (2018) 4459–4470, https://doi.org/10.2147/CMAR.S177373. PubMed PMID: 30349378; PubMed Central PMCID: PMC6188178.
[53] J. Liu, X. Dai, X. Guo, A. Cheng, S.M. Mac, Z. Wang, Circ-OXCT1 suppresses gastric cancer EMT and metastasis by attenuating TGF-β pathway through the circ-OXCT1/miR-136/SMAD4 Axis, OTT 13 (2020) 3987–3998, https://doi.org/10.2147/OTT.S239789. PubMed PMID: 32523351; PubMed Central PMCID: PMC7236241.
[54] M. Song, Y. Ping, K. Zhang, L. Yang, F. Li, C. Zhang, et al., Low-dose IFNγ induces tumor cell stemness in tumor microenvironment of non, Small Cell Lung Cancer 79 (2019) 3737–3748, https://doi.org/10.1158/0008-5472.CAN-19-0596. PubMed PMID: 31085700.
[55] T. Tamura, M. Ikegami, Y. Kanemasa, M. Yomota, A. Furusawa, R. Otani, et al., Selection bias due to delayed comprehensive genomic profiling in Japan, Cancer Sci. 114 (2023) 1015–1025, https://doi.org/10.1111/cas.15651. PubMed PMID: 36369895; PubMed Central PMCID: PMC9986065.