

Ferroptosis-related Genes Predicting Survival Status in Gliomas with Machine Learning Algorithms

Yanliang Tang¹, Xiaoli Zhang¹, Xiaofei Tang¹, Ye Yuan¹ and Wenwen Wang^{2*}

¹Department of Neurology, Hangzhou Fuyang Hospital of Traditional Chinese Medicine, Hangzhou, China

²Department of Internal Medicine, Hangzhou Fuyang Hospital of Traditional Chinese Medicine, Hangzhou, China

***Corresponding Author:** Wenwen Wang, Department of Internal Medicine, Hangzhou Fuyang Hospital of Traditional Chinese Medicine, No.2-4 Guihua Road, Fuchun Street, Hangzhou, 311400, China, E-mail: 15906667745@163.com

Received Date: October 02, 2024 **Accepted Date:** November 02, 2024 **Published Date:** November 05, 2024

Citation: Yanliang Tang, Xiaoli Zhang, Xiaofei Tang, Ye Yuan, Wenwen Wang (2024) Ferroptosis-related Genes Predicting Survival Status in Gliomas with Machine Learning Algorithms. J Cancer Res Therap Oncol 12: 1-25

Abstract

Introduction: Glioma is one of the most prevalent tumors in the central nervous system and has been classified into low-grade glioma (LGG) and glioblastoma (GBM). Ferroptosis is a form of iron-dependent programmed cell death. The objective of this study is to construct a survival status prediction model using ferroptosis-related genes for patients with LGG and GBM.

Methods: RNA-seq data and clinical information pertaining to patients with gliomas were collected from TCGA. Predictive models were constructed based on selected specifically expressed lncRNAs and mRNAs in GBM and LGG that demonstrated a high degree of correlation with ferroptosis genes. To assess the performance of the prediction models, we examined the areas under the curve (AUC). In order to achieve a balanced dataset, we applied Random Oversampling Examples (ROSE).

Results: A comparison of the expression of 59 ferroptosis-related genes in LGG and GBM revealed a predominance of highly expressed gene transcripts in LGG, as compared to GBM. The results of our enrichment and pathway analyses revealed notable differences in the functions and pathways of exclusively expressed ferroptosis-related lncRNAs and mRNAs in LGG and GBM. XGBoost and random forest were employed to forecast the survival status of glioma patients based on the top 20 lncRNAs and mRNAs selected from elastic net. Upon evaluating model performance using AUC, we observed that XGBoost exhibited superior performance in predicting survival outcomes for patients with LGG and GBM. Furthermore, we observed a notable improvement in model performance for XGBoost following ROSE in the GBM cohort, while this was less pronounced in the LGG subgroup.

Conclusion: We constructed a survival status prediction model for patients with LGG and GBM. Furthermore, with the aid of elastic net and Cox regression, we were able to identify several survival ferroptosis-related mRNAs in GBM and LGG, respectively. These potential biomarkers warrant further research for the purpose of validating their prognostic value.

Keywords: Ferroptosis; Gliomas; Prediction Model; Machine Learning; Survival Status

Abbreviations

CDF: cumulative density function; CNS: central nervous system; GBM: glioblastoma; GTEX: Genotype-Tissue Expression; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; lncRNA: long non-coding RNA; LGG: low-grade glioma; mRNA: messenger RNA; OS: Overall Survival; PCA: Principal Component Analysis; ROSE: Random Over Sampling Examples; TCGA: The Cancer Genome Atlas Program; TERT: Telomerase Reverse Transcriptase; WHO: World Health Organization.

Introduction

Glioma is one of the most common tumors in the central nervous system (CNS) [1]. According to the fifth edition of the WHO classification of tumors of the CNS in 2021, several main subtypes of gliomas are as follows: astrocytoma is a type of lower grade glioma (LGG) with isocitrate dehydrogenase (IDH) mutation, which has the key diagnostic genes, including IDH1, IDH2, ATRX, TP53 and CDKN2A/B. Oligodendroglioma is another type of LGG with IDH mutation and 1p/19q-codeleted, which has the key diagnostic genes of IDH1, IDH2, 1p/19q, telomerase reverse transcriptase (TERT) promoter, capicua transcriptional repressor (CIC), far upstream element binding protein 1 (FUBP1), and NOTCH1. The conventional treatment for LGG is surgical resection. Patients who underwent surgery exhibited a superior survival outcome compared to those who did not [2-4]. However, the recurrence rate of LGG is relatively high, particularly in cases where the patient is at high risk. Consequently, a considerable number of LGG patients are required to undergo radiotherapy or chemotherapy [4,5]. A number of studies have demonstrated that the IDH mutation is more prevalent in grade II and III than in grade IV, indicating that IDH mutations are involved in the early stages of tumor progression [6,7].

Glioblastoma (GBM) is grade IV glioma, with IDH-wildtype, and the key genes and molecular characteristics are IDH-wildtype, TERT promoter, chromosomes 7/10 and epidermal growth factor receptor (EGFR) [8,9]. GBM is the

highest grade of glioma, which studies confirmed that over 90% of GBM are IDH-wildtype tumors [10]. Similar to the treatment in LGG, the standard care of GBM is the combination of surgery and radiotherapy and temozolomide chemotherapy, but some studies suggest targeted therapy for patients that show aberrant CpG methylation of a particular gene, O⁶-methylguanine DNA methyltransferase (MGMT) gene [11-14]. The targeted therapy based largely on the target pathways that are common in GBM, including the phosphoinositide 3-kinase (PI3K), protein kinase B (AKT), mammalian target of rapamycin (mTOR), and the p53 and the retinoblastoma (RB) pathways [10]. In order to have a better understanding of the treatment and prognosis of patients with gliomas, numerous studies have focused on certain lncRNAs or mRNAs to further investigate the underlying mechanisms and potential prognostic values [15-19]. The interest in long non-coding RNA (lncRNA) and messenger RNA (mRNA) has been outburst over the years, for their diverse functions in gene regulation and potential therapeutic treatment [20-25].

Ferroptosis is a form of regulated cell death but different from the other programmed death such as apoptosis [26,27] (Dixon et al., 2012, Liu et al., 2020). Ferroptosis is an iron-dependent cell death that with the process of iron accumulation, and thus increased lipid peroxidation would lead to the loss of lipid repair enzyme glutathione peroxidase (GPX4) and results in the increase of lipid-based reactive oxygen species (ROS) [28-31]. Furthermore, mounting evidence supported the idea of the therapeutic value of fer-

roptosis, through constructing prognostic gene models or targeted therapy with ferroptosis specific pathways in various cancers [29,32-36]. It was demonstrated that the induction of iron death in glioma cells disrupts their antioxidant homeostasis, leading to the inhibition of tumor growth [37,38]. Glioblastoma (GBM) cells require substantial quantities of iron to promote tumour growth and progression, thereby rendering these cells susceptible to destruction through iron death induction [39]. A study has indicated that neutrophils facilitate tumor necrosis through the induction of ferroptosis in glioblastoma progression [40].

Furthermore, the relationship between iron death and the immune microenvironment of gliomas is also a topic of interest. Modulation of iron death-related signaling pathways by glioma cells has the potential to influence the infiltration and function of immune cells in the tumor microenvironment, thereby evading immune surveillance and attack [41,42]. In GBM (glioblastoma), there is a notable increase in mutations in P53, and the status of P53 may serve as an additional prognostic indicator for the treatment of glioma with iron death-inducing agents [43,44].

A few studies have investigated the prognostic model constructed by ferroptosis-related lncRNAs or mRNAs in gliomas, and these models could be used to identify potential biomarkers and further treatment applications [45-47]. The aim of this study is to differentiate the biomarkers in LGG and GBM and to make a contribution to the potential target treatment for glioma patients, with the objective of achieving a better prognosis. In particular, our objective was to distinguish the exclusively expressed genes in LGG and GBM and to explore the corresponding pathways in GBM and LGG, respectively. Moreover, based on the exclusively expressed genes, we constructed several survival prediction models using machine learning algorithms on the training set and validated their performance on the test set. To enhance the model performance, we employed Random Over-Sampling Examples (ROSE), a data balancing method, to address the issue of an unbalanced dataset and observed improved results, particularly in patients with GBM.

Method

Data Acquisition

In this study, data were collected from a number of common sources, including TCGA, GTEx and multiple peer-reviewed literature sources. The initial step involved the retrieval of gene expression data and clinical information pertaining to patients diagnosed with GBM and LGG, respectively, from TCGA. The database was queried to obtain data on 169 patients with GBM and 532 patients with LGG. As TCGA lacks normal glioma tissues, gene expression data of normal glioma tissues were downloaded from GTEx (n = 1132). The normal tissues will be employed to identify differentially expressed genes in conjunction with tumor tissues. All data from TCGA and GTEx were transformed to $\log_2(x+1)$ form for further analysis, and genes exhibiting consistent expression of 0 were also removed. The 59 ferroptosis-related genes were extracted from a previous study [47].

Correlation Heatmap & Venn Plot

The correlation between each ferroptosis-related gene was quantified using Pearson correlation, with only those pairs exhibiting a statistically significant correlation included in a heatmap ($p < 0.05$). In order to explore the relationship between ferroptosis-related genes and long non-coding RNAs (lncRNAs) and messenger RNAs (mRNAs), we measured the Pearson correlation of lncRNAs and ferroptosis-related genes, with the limitation of an absolute value of correlation greater than 0.7 and a p-value less than 0.05. Similarly, a Pearson correlation was conducted to examine the correlation between mRNAs and ferroptosis-related genes. However, the criteria for inclusion were more stringent, with only those with an absolute correlation value greater than 0.7 and a p-value less than 0.001 being retained.

The principal rationale for the elevation of the mRNA selection threshold was the abundance of mRNAs that exhibited a high degree of correlation with ferroptosis-related genes. Consequently, it was necessary to implement a further elimination process through the adjustment of the p-value. The ferroptosis genes were employed to facilitate a comparison of the expression differences between GBM

and LGG through the utilization of a box plot. An analysis of the expression differences of ferroptosis genes between GBM and LGG may facilitate an understanding of the levels of gene expression and glioma progression at different levels. Furthermore, a heatmap was constructed to facilitate a comparative analysis of the expression differences between the two cohorts, incorporating the corresponding clinical in-

formation (age, living status, and gender). With the clinical information obtained from TCGA, it is reasonable to have a summary of how the patients were distributed with regard to age, gender, and living status. The statistical summary of patients with GBM and LGG is presented in Table 1 and Table 2. To further refine the scope of our investigation, we excluded the genes that were commonly identified in ferroptosis-related mRNAs and lncRNAs in GBM and LGG.

Table 1: The statistics summary table of patients with GBM

	Alive (N=30)	Dead (N=130)	Overall (N=160)
Futime			
Mean (SD)	0.923 (0.713)	1.21 (1.10)	1.16 (1.05)
Median [min, max]	0.671 [0.036, 2.62]	1.05[0.0137, 7.35]	0.986 [0.0137,7.35]
Missing	0 (0%)	1 (0.8%)	1 (0.6%)
Age			
20-40	4(13.3%)	9 (6.9%)	13 (8.1%)
41-60	13 (43.3%)	50 (38.5%)	63 (39.4%)
61-80	12 (40.0%)	64 (49.2%)	76 (47.5%)
81-90	1 (3.3%)	7 (5.4%)	8 (5.0%)
Gender			
Female	12 (40.0%)	44 (33.8%)	56 (35.0%)
Male	18 (60.0%)	86 (66.2%)	104 (65.0%)

Table 2: The statistics summary table of patients with LGG

	Alive (N=388)	Dead (N=125)	Overall (N=513)
Futime			
Mean (SD)	2.41 (2.40)	3.34 (3.14)	2.64 (2.63)
Median [min, max]	1.72 [-0.00274, 17.6]	2.23 [0.0192, 14.2]	1.85 [-0.0024, 17.6]
Missing	1 (0.3%)	0 (0%)	1 (0.2%)
Age			
20-40	194 (13.3%)	43 (6.9%)	237 (8.1%)
41-60	153 (39.4%)	50 (40.0%)	203 (39.6%)
61-80	37 (9.5%)	31 (24.8%)	68 (13.3%)
81-90	0 (0%)	1 (0.8%)	1 (0.2%)
Missing	4 (1.0%)	0 (0%)	4 (0.8%)
Gender			

Female	174 (44.8%)	54 (43.2%)	228 (44.4%)
Male	214 (55.2%)	71 (56.8%)	258 (55.6%)
Grade			
G2	212 (54.6%)	37 (29.6%)	249 (48.5%)
G3	175 (45.1%)	88 (70.4%)	263 (51.3%)
Unknown	1 (0.3%)	0 (0%)	1 (0.2%)

Functional Enrichment Analysis

By extracting exclusively expressed ferroptosis-related long non-coding RNAs (lncRNAs) or messenger RNAs (mRNAs), we sought to elucidate the gene functions and pathways in GBM and LGG. The "clusterProfiler" package was employed for the functional enrichment analysis, and our genes were mapped with the "org.Hs.eg.db" package. Gene Ontology (GO) enrichment was used to identify gene functions in biological process, molecular function, and cellular component in GBM and LGG, respectively. Similarly, the Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to identify pathways in which the selected genes were enriched.

Consensus Clustering

To further examine the influence of ferroptosis genes on patient survival in GBM and LGG subgroups, we conducted a consensus clustering analysis using the R package "Consensus ClusterPlus." In this analysis, we sought to group the data into $k = 2$ to $k = 8$ clusters, employing the cumulative density function (CDF) and survival rate, to identify the optimal number of subgroups.

Feature Selection (elastic net)

A considerable body of research employs Lasso regression in conjunction with Cox regression to identify and select genes that are associated with survival (46). However, there are a few shortcomings associated with Lasso regression. It is not effective when several genes with potential survival-related characteristics are highly correlated. In such instances, Lasso will automatically select one or two genes and shrink the others to zero, thereby achieving the objective of variable selection. Another limitation is that when the number of patients is significantly smaller than the number of genes, Lasso can only select as many genes as there are pa-

tients. In other words, the number of genes selected will be constrained by the number of patients. One potential solution is to employ random Lasso in place of Lasso regression (48). However, the computational power required by random lasso is considerable, with the running time for processing all the genes in the dataset taking several hours. Therefore, it is necessary to identify an alternative resolution.

The aforementioned limitations could be partially alleviated by employing the elastic net, a comparable algorithm that could potentially relax the constraints imposed by the lasso. In the case of variables that are highly correlated, elastic net performs better than lasso, which tends to select only one variable, whereas elastic net maintains both variables. Lasso regression is a general regression with an L-1 penalty, which serves to reduce the dimensionality of a given dataset. In contrast, the elastic net algorithm incorporates an L-2 penalty in addition to the Lasso regression. The most advantageous aspect of elastic net is that it is capable of consistently identifying a solution to the presented problem, and the computational time required for optimization is relatively brief.

A predetermined random seed was used to split the data into training and test sets in a ratio of 6:4, with the same ratio of survival outcome in both datasets. Furthermore, the bootstrap technique was employed to eliminate any bias that might have arisen from randomness. The distribution of survival outcome was found to be unbalanced in patients with GBM and LGG, and thus ROSE was applied to balance the two classes in the training sets. In the application, the "caret" package in R was employed to configure the parameters for the training set. Specifically, within the training set, 10-fold cross-validation was performed five times. With regard to the regularization parameter in elastic net, the search grid for alpha ranged from 0 to 1, with a length

of 15, and the search grid for lambda ranged from 0.0001 to 0.2, with a length of 5. Those variables with non-zero coefficients were retained for the purpose of model construction.

Prediction Model (XGBoost, random forest)

After the feature selection, we were able to employ machine learning techniques to predict the survival status of patients with gliomas. We chose random forest as our first prediction model. Random forest is an algorithm that can be classified as bagging, which is distinct from boosting method. Similar to XGBoost, in random forest, we constructed multiple decision trees as well. In boosting methods, the later learner depends on the result of the previous learner, whereas in random forest, each decision tree is independent from each other. Each decision tree from random forest will vote for a result, and the final output of the random forest is the majority of votes from individual trees. Generally, random forest is capable to process high dimension data without dimension reduction and feature selection. Moreover, with Gini purity, it can also output the importance of each variable. The second algorithm applied in this study is XGBoost. XGBoost is an efficient gradient boosting decision tree (GBDT), the idea of boosting is to compile multiple weak learners into a strong one to improve the prediction ability. Comparing to the general GBDT, XGBoost improves the performance by optimizing the loss function from first order Taylor expansion to the second order Taylor expansion and making use of the L2 regularization to simplify and avoid overfitting. The evaluation of the model performance was conducted using the receiver operating characteristic (ROC) and area under the ROC curve (AUC) metrics. Additionally, ROSE was employed in the training set, comprising both LGG and GBM patients, to eliminate the influence of dataset imbalance and improve model performance.

Statistical Analysis

All figure plotting, machine learning, and statistical analysis were conducted using the R statistical computing platform (version 4.2.2). The correlations evaluated in the heatmap were all quantified using the Pearson correlation coefficient, with a significance level of $p < 0.05$. A boxplot was constructed to compare the gene expression difference in GBM and LGG, with a Wilcoxon test ($p < 0.05$) em-

ployed to this end. The GO and KEGG enrichment analysis were performed with the "clusterProfiler" package in R, while survival analysis was conducted using the "survival" package in R.

Results

Correlation of Ferroptosis-related Genes

The methodology employed in this study is illustrated in Figure 1. Figure 2A illustrates the heatmap of 59 ferroptosis-related genes in GBM. All pairs with correlation and a p-value of less than 0.05 were presented. In general, the majority of the pairs presented in the heatmap exhibited a positive correlation. Figure 2B illustrates that the correlation of each pair of ferroptosis genes in LGG exhibits a relatively distinct pattern compared to that observed in GBM. The number of pairs with correlation and a p-value less than 0.05 in Figure 2B was less than in Figure 2A. Furthermore, the number of pairs with a negative correlation was similar to the number of pairs with a positive correlation in LGG. The Venn diagrams presented the exclusively expressed mRNAs and lncRNAs, as well as the common mRNAs and lncRNAs shared in GBM and LGG, respectively (Figure 2C-2D). Even with a more rigorous approach to filtering ferroptosis-related mRNAs, the number of highly correlated mRNAs with ferroptosis genes remains substantial. A total of 4704 mRNAs exhibited an absolute Pearson correlation value of greater than 0.7 in GBM, while 719 mRNAs demonstrated a similar correlation in LGG. The number of shared mRNAs was significantly higher than the number exclusively expressed in GBM (Figure 2C). Conversely, the number of highly correlated lncRNAs was relatively low in both GBM and LGG. Notably, the common lncRNAs no longer comprised the majority of the ferroptosis-related lncRNAs in LGG compared to mRNAs (Figure 2C-2D).

The expression level discrepancy in each ferroptosis-related gene between GBM and LGG was illustrated in a boxplot (Figure 3A). Among the 59 genes, five exhibited no differential expression between GBM and LGG. Conversely, 32 of the 59 ferroptosis-related genes exhibited significantly elevated expression levels in LGG compared to GBM. A considerable number of genes exhibited high expression levels in LGG, while 22 genes demonstrated higher expression lev-

els in GBM than in LGG. Furthermore, we combined the clinical information of the patients with the differentially expressed genes in GBM and LGG (Figure 3B). In general, the

expression of ferroptosis-related genes in the two cohorts was different, particularly when they were split into smaller subgroups.

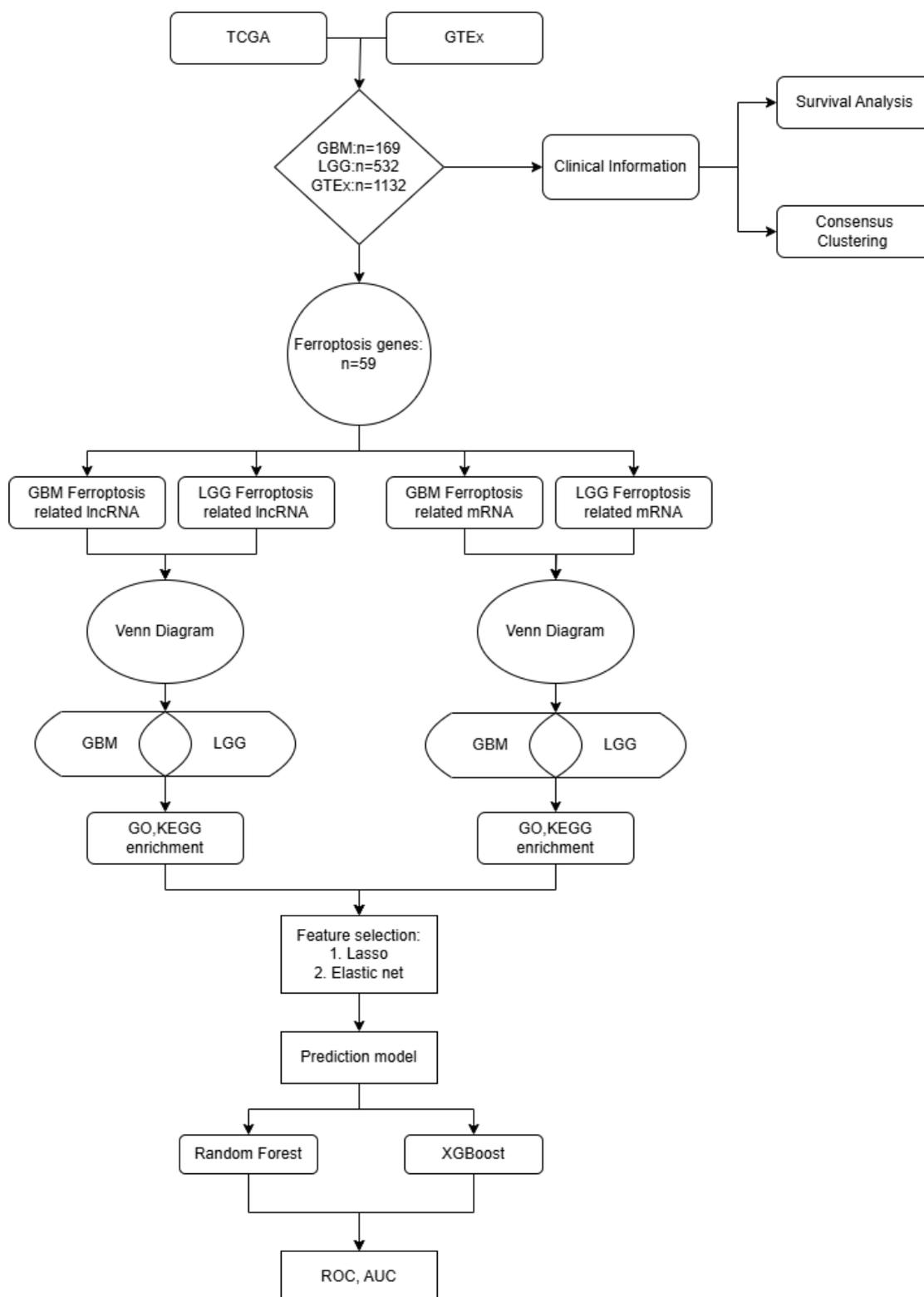


Figure 1: Workflow of this study

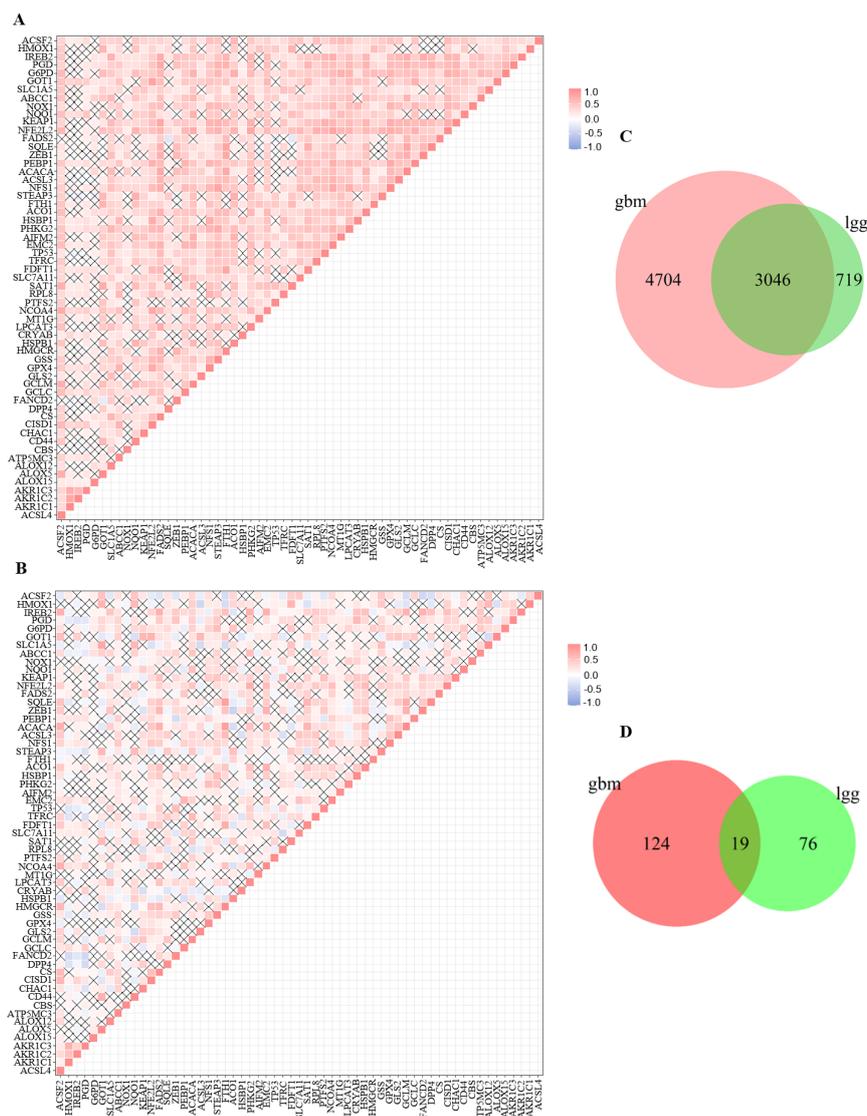


Figure 2: Correlation of ferroptosis-related genes in pairs and Venn diagrams for ferroptosis-related genes. (A-B) Heatmap of the correlation of ferroptosis-related genes in GBM (A) and LGG (B), respectively. Only pairs with correlation and t-test with $p < 0.05$ were shown; (C) Venn diagram of mRNAs related to ferroptosis genes that the filtering condition is $|\text{cor}| > 0.7$, $p < 0.001$; (D) Venn Diagram of lncRNAs related to ferroptosis genes that the filtering condition is $|\text{cor}| > 0.7$, $p < 0.01$.

Functional Enrichment Analysis

By investigating exclusively expressed genes in two distinct phases of glioma development, we are able to ascertain their functional roles and associated biological pathways. In the GO analysis, a significant proportion of the genes were found to be involved in mitochondrial-related functions in mRNAs in GBM (Figure 4A). In the molecular function section, some genes were found to be involved in oxidoreduction and ubiquitin-related functions. While the majority of the ferroptosis-related mRNAs in LGG were shared with GBM, the functions of the exclusively expressed mRNAs in LGG were distinct from those of the genes in

GBM (Figure 4B). The functions in the biological process category were predominantly related to synaptic and neurotransmitter transport, a pattern that was also observed in the cellular component and molecular function sections, where channel activities and synaptic and neuronal activities were similarly prominent. With regard to the enrichment pathways in mRNAs in GBM, the majority of the pathways were found to be enriched in various neural diseases (Figure 5A). In contrast, in LGG, the pathways were predominantly enriched in multiple signaling pathways and a range of substance addictions, such as morphine addiction (Figure 5B).

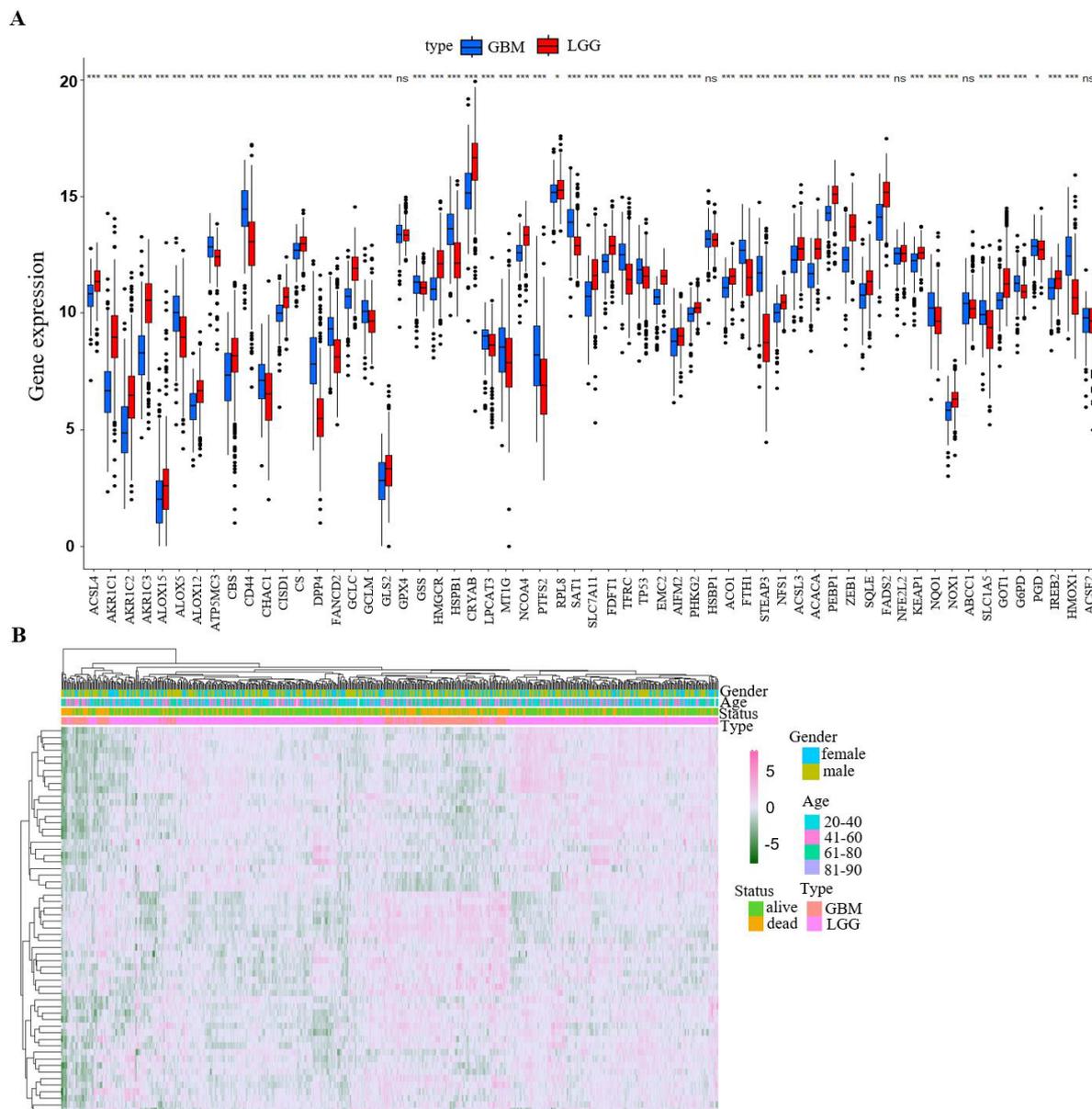


Figure 3: Gene expression difference of ferroptosis-related genes in GBM and LGG. (A) A boxplot demonstrates the gene expression level between GBM and LGG in each ferroptosis-related gene; (B). Heatmap of expression difference in GBM and LGG combined with clinical information of patients.

Consensus Clustering of Glioma Patients and Clinical Subtypes

It would be beneficial to determine whether there are any subgroups with differing survival rates, allowing for a more in-depth examination of the condition and more accurate decision-making when managing gliomas. The "ConsensusClusterPlus" package revealed that the optimal clustering of patients with GBM was achieved with $k=4$ (see Supplementary Figure 1A). However, there was a discrepancy between the consensus CDF and the area under the curve

and the effect of subgroup splitting (Supplementary Figure 1B-C), despite the evidence of a significant difference among subgroups (Supplementary Figure 1D). Meanwhile, the optimal number of subgroups in patients with LGG was 3 (Supplementary Figure 1E). The effect of subgroup splitting was similar to that observed in GBM, whereby the subgroups did not yield favorable outcomes (Supplementary Figure 1F-G). Furthermore, the subgroups in LGG did not demonstrate differences in survival rates (Supplementary Figure 1H).

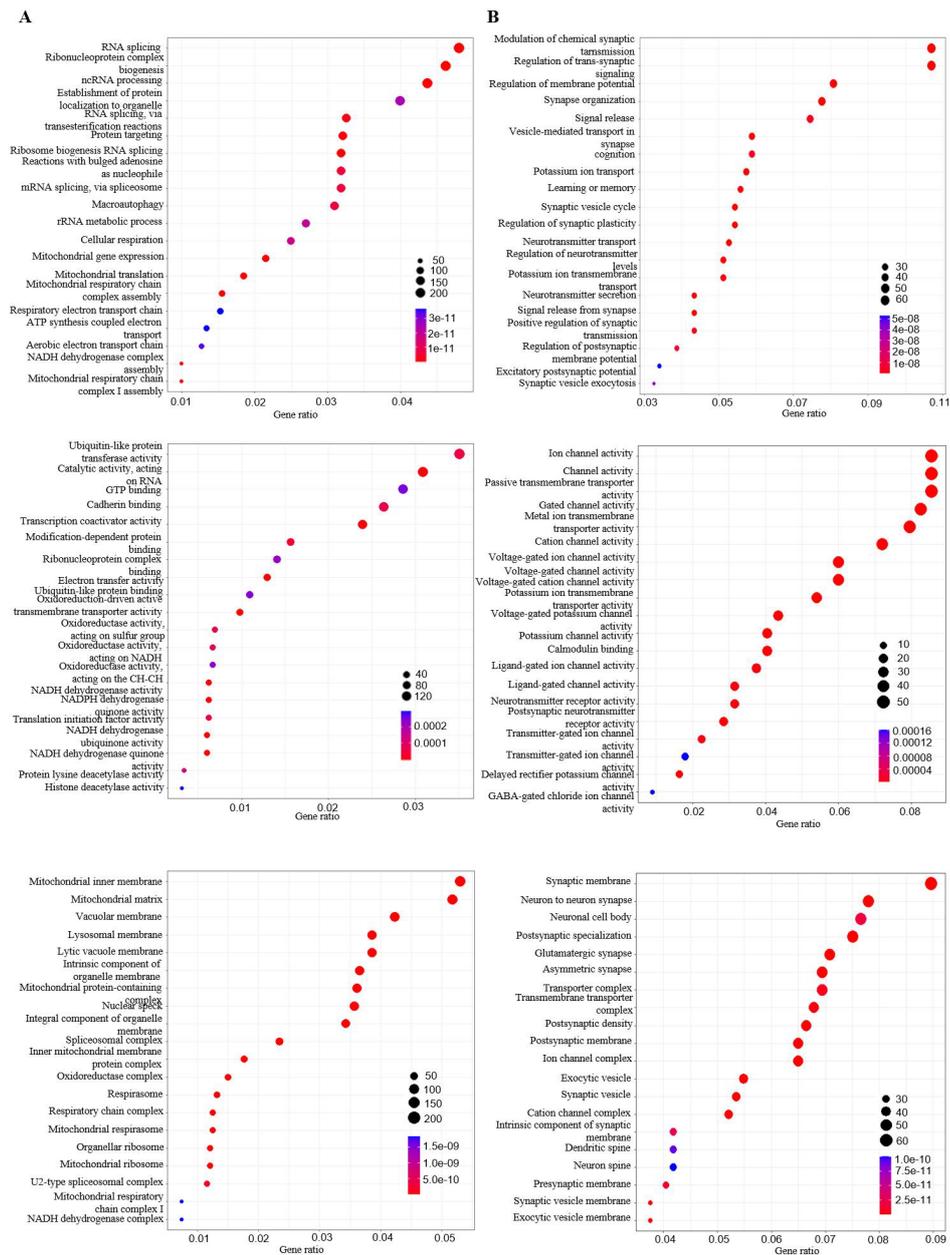


Figure 4: Performed GO functional analysis on mRNAs from GBM and LGG. (A) Top: GO analysis on biological process on ferroptosis-related mRNAs in GBM. A few of the biological processes were focused on mitochondrial functions. Middle: GO analysis on molecular functions, and a great amount of functions were focused on oxidoreductase activities. Bottom: GO analysis on cellular component functions, the main functions in this plot were involved with mitochondrial again; (B) Top: GO analysis on ferroptosis-related mRNAs in LGG. GO analysis on biological process that most functions were about chemical synaptic transmission. Middle: GO analysis on molecular functions, and a lot were chemical channel activities. Bottom: GO analysis on cellular component functions, and many functions were about synaptic activities.

Despite the consensus clustering didn't show the survival difference among subgroups, we inspected the survival differences in clinical subtypes in both GBM and LGG patients. Survival analysis was performed to assess the effect of gender and age (divided into over 45 and under 45 categories) in survival on training and test sets. In GBM pa-

tients, the analysis showed no statistically significant differences between the two characteristics (Supplementary Figure 2A). In addition, we did the same analysis for LGG patients and found that there was no difference in survival between men and women, but that patients in the two age groups had significantly different survival rates (Supplemen-

tary Figure 2B).

Feature Selection and Prediction Model

The process of feature selection was conducted using elastic net as the fundamental method, prior to conducting the survival analysis. The progression of the search grid, which was run using elastic net, and its results for mRNAs in GBM are presented in Figure 6A. The mixing percentage was plotted against the accuracy for different values of the

regularization parameters in the training set. In general, when the regularization parameter was in the range of 0.15-0.2, the highest accuracy was observed; conversely, when the parameter was between 0.0001 and 0.05, the accuracy was observed to decrease as the parameter increased. With the optimized parameter selected using elastic net, a histogram of the variable importance in the model was acquired, in which the importance is in descending order (Figure 6B).

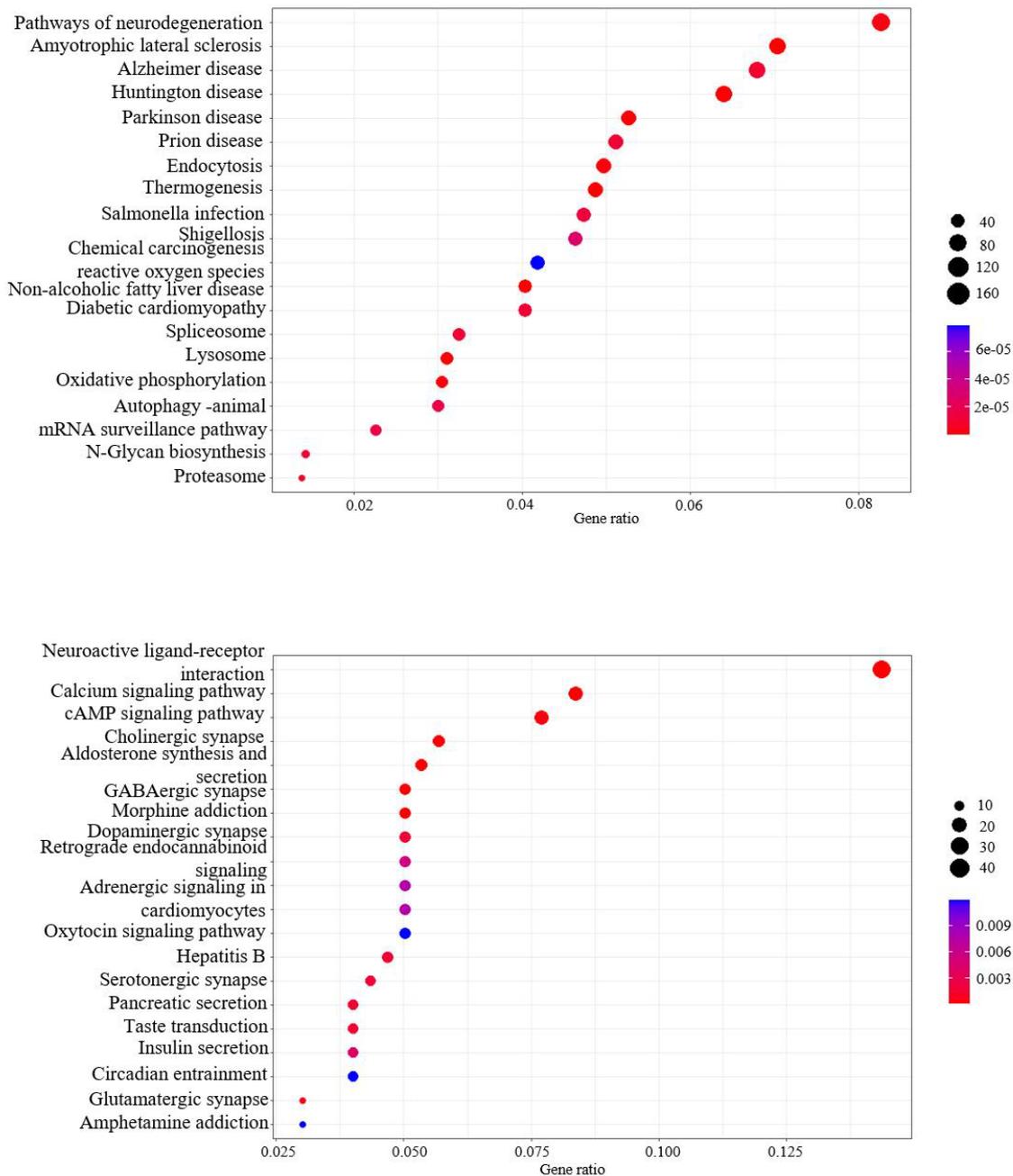


Figure 5: KEGG analysis in GBM and LGG. (A) The KEGG pathway enrichment analysis for exclusively expressed mRNAs in GBM; (B) The KEGG pathway enrichment analysis for exclusively expressed mRNAs in LGG.

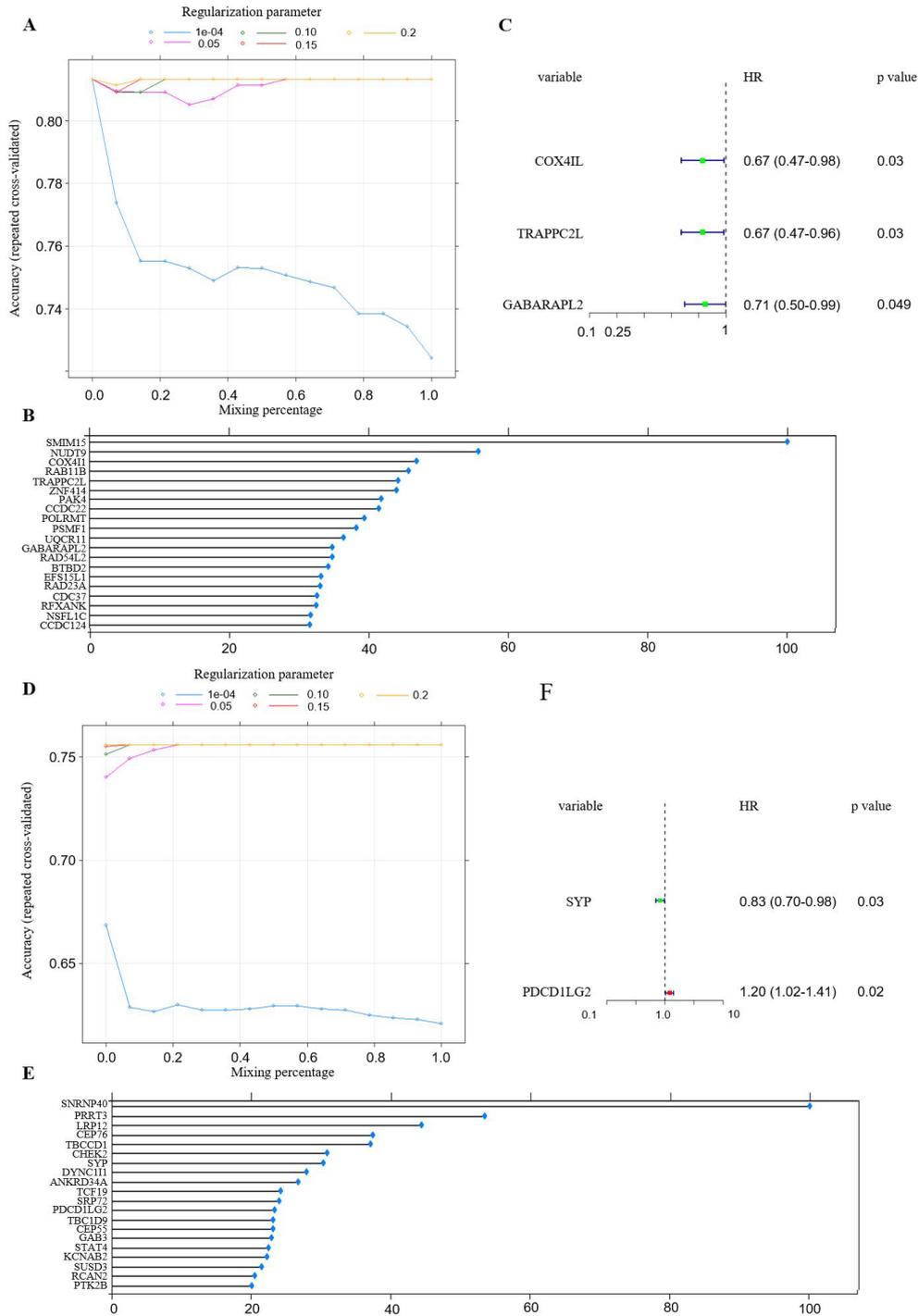


Figure 6: Parameters of feature selection using elastic net. (A) The plot showed the changes in accuracy when the algorithm took in different values of regularization parameter where the blue dotted line represented a series of parameter that was equally cut between 1e-04 to 0.05, the same to all the other lines; (B) The variable importance of the features selected by elastic net after we obtained the optimal value from previous plot. (C) The forest plot of the 20 selected ferroptosis-related mRNAs and ran them into survival analysis, only three genes showed survival related and both lncRNAs had HR less than 1; (D) The plot of the changes in accuracy when the mixing percentage changed at different value of regularization parameters; (E) The top 20 ferroptosis-related mRNAs in LGG that were selected from elastic net, and ranked with variable importance; (F) The two mRNAs among the 20 selected variables that were survival related, and SYP had the HR less than 1 whereas the HR of PDCD1LG2 was greater than 1.

Following the reduction of the number of genes, survival analysis was conducted using the 20 selected genes, with the results displayed in a forest plot (Figure 6C). Mean-while, the plot demonstrating the optimized parameter indicated that the algorithm reached its highest level of accuracy when the regularization parameter was set between 0.15 and 0.2 (Figure 6D). Of the 20 selected genes with the highest variable importance, a survival analysis was performed on them. Among these genes, only two were found to be significantly related to survival (Figure 6E-6F). In addition to investigating the mRNAs in gliomas, we have also selected ferroptosis-related lncRNAs with elastic net (Supplementary Figure 3). However, with the selected lncRNAs in elastic net, none of the genes were found to be survival-related in both GBM and LGG (Supplementary Figure 3C and 3F).

Conversely, we employed a random forest approach to forecast the algorithm's performance in predicting survival outcomes. The variables that had been filtered using elastic net were then extracted and incorporated into the random forest model. The variable importance and Gini impurity were subsequently displayed (see Figure 7A). Following the fitting of the genes into the random forest, the accuracy of the training set and test sets were presented (Figure 7B). The AUC from the random forest model was 1.0 for the training set and 0.5 for the test set, which was an exceptionally high value for the training set and a surprisingly low value for the test set. During the analysis, it was observed that the incidence of death among patients diagnosed with glioblastoma (GBM) (77%) was significantly higher than that of patients who survived gliomas (23%). The discrepancy between the accuracy of the training set and test set was primarily attributable to the imbalanced dataset inherent to the disease. To address this, we applied ROSE to obtain a relatively balanced data set. Utilizing the algorithm, the AUC obtained from the random forest analysis of the balanced dataset was 0.84, although the AUC in the test set remained 0.5 (Figure 7C).

In the mRNAs in LGG, the same procedures were

followed as in GBM, resulting in two plots demonstrating the variable importance from random forest (Figure 7D-E). Subsequently, the performance of the random forest was evaluated based on the AUCs obtained from the training and test sets. As with the model performance in GBM, the AUC was 1.0 and 0.5 for the training and test sets, respectively. In the case of the balanced dataset produced by ROSE, the AUC in the training set was reduced to 0.83 due to the decrease in the proportion of deaths, whereas the AUC in the test set was only increased by 0.01 in comparison to the original dataset (Figure 7G).

The results of the model performance in XGBoost were significantly superior to those of random forest. Upon completion of 70 iterations, it was observed that the AUC value exhibited a gradual increase with an increase in the number of iterations in the training set for GBM (Figure 8A). A comparison of the training sets before and after data balancing in GBM revealed that the AUC was 0.82 before ROSE at the beginning of the iteration. However, after ROSE, the AUC was already close to 0.9, and it reached its peak at a faster rate in the dataset with ROSE than in the one without. Nevertheless, in the AUC test set of XGBoost with ROSE, the AUC decreased as the number of iterations increased.

The AUC in the test set of XGBoost without the ROSE demonstrated an initial increase, followed by a decrease, and a subsequent period of minor fluctuation, ultimately reaching a stable AUC after 40 iterations. The results for LGG exhibited a somewhat disparate pattern. The area under the curve (AUC) of the XGBoost training process commenced at a relatively low value of 0.6 and reached its peak around 20 iterations. In contrast, the AUC in the training set with ROSE commenced at 0.83 and reached 1 at a faster rate. Nevertheless, the AUCs in the test sets exhibited a high degree of intertwining before the first 20 iterations, and the AUCs in the test set with ROSE demonstrated consistently lower values than those in the test set without ROSE (Figure 8B).

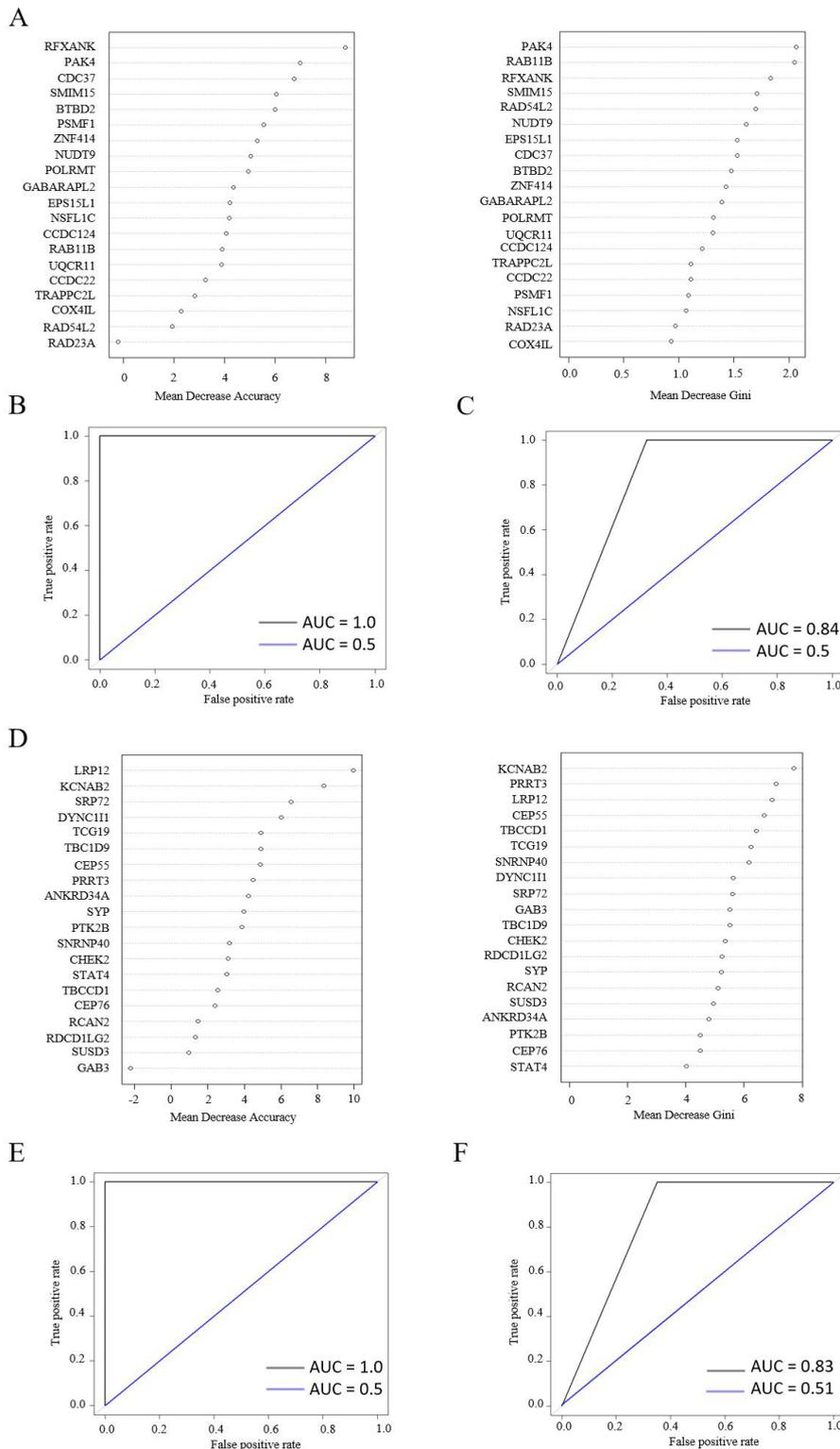


Figure 7: Performance of random forest in GBM and LGG for both training and test sets. (A) The variable importance of the selected genes from GBM, and we evaluated the importance on the metric of mean decrease accuracy and mean decrease gini; (B) ROC of the random forest on training and test sets for GBM, the ROC for training set was in black, and ROC for the test set was in blue; (C) ROC of the random forest after ROSE on training (black) and test (blue) sets; (D) The variable importance of the selected genes from LGG, and evaluated the importance on the metric of mean decrease accuracy and mean decrease gini; (E) ROC of the random forest on training (black) and test (blue) sets, respectively; (F) ROC of the random forest after ROSE on training (black) and test (blue) sets.

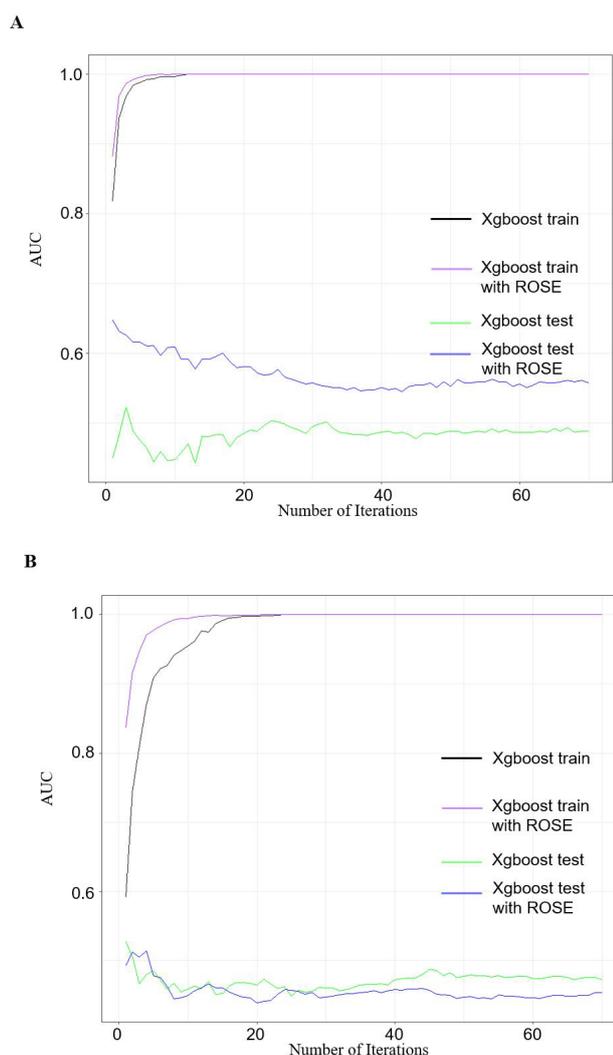


Figure 8: AUCs on Xgboost algorithm in training and test sets before and after ROSE. (A) Using 20 genes in GBM that were selected from elastic net, and put them into Xgboost. The x-axis is the number of iterations and y-axis is the value of AUCs. Each line represents the model performance on that dataset, the black and green curves represent the training and test sets before ROSE, and the purple and blue curves represent the training and test sets after ROSE; (B) In LGG, we applied the 20 genes in LGG selected from elastic net and ran the Xgboost, the performances of each group were shown in the plot.

Discussion

Ferroptosis is relatively new form of programmed death that has been gradually studied since 2012 [48]. A study demonstrated that in GBM, LGG and non-tumor groups, a substantial number of ferroptosis-related genes were differentially expressed. Furthermore, a prognostic risk model was established to predict the survival of patients with gliomas, and the risk score was also correlated with IDH mutation status. This risk score proved to be an effective estimator for predicting immunotherapy [48]. This study explored the differential expression of 59 ferroptosis-related genes between GBM, LGG, and non-tumor tis-

sues. Our investigation involved the identification of the top 100 differentially expressed lncRNAs and mRNAs, exclusive to GBM and LGG, and correlated with ferroptosis. Our findings revealed that the majority of the 59 ferroptosis-related genes exhibited higher expression levels in LGG than in GBM. This observation provides a promising foundation for further investigation into the expression differences between these two levels of gliomas and the prospect of developing targeted therapies against LGG and GBM, with the aim of achieving more favorable treatment outcomes.

A comparison of the expression of ferroptosis genes in GBM and LGG revealed that over 50% of genes in-

volved in ferroptosis demonstrated higher expression in LGG compared to GBM, as previously indicated in the results section. The most highly expressed gene in both GBM and LGG was CRYAB. Prior studies have indicated that the overexpression of CRYAB correlates with tumor progression and a poor prognosis in patients diagnosed with ovarian cancer [49], lung cancer [50], and breast cancer [51]. Another study reported that in oligodendroglioma, CRYAB was more highly expressed in astrocyte-like cells than in oligodendrocyte-like cells [52]. It seems plausible that the elevated expression of CRYAB in LGG is attributable to the inherent heterogeneity of LGG and the likelihood of its advancement to a more malignant grade. On the other hand, among the ferroptosis genes that have higher expression in GBM than in LGG, CD44 expressed the highest level in GBM. Moreover, several studies have demonstrated that the overexpression of CD44 correlates with poorer prognosis in patients with gliomas [53,54].

A growing body of evidence indicates that mitochondria play a pivotal role in ferroptosis through the deprivation of cysteine [55]. While some studies have indicated a controversial and context-based role for defense mechanisms in ferroptosis mitochondria, other studies have shown that a mitochondria-targeted antioxidant, MitoTEMPO, induces a protective effect [55-57]. These findings align with our GO analysis of mRNAs in GBM, which revealed a significant enrichment of biological processes within mitochondrial activities. In the gene ontology (GO) enrichment analysis of GBM, oxidoreductase activity was identified as a recurring theme. Additionally, research has demonstrated that oxidoreductases, such as glutaredoxin 2 (Grx2c), play a role in facilitating glioma cell migration and invasion [58]. However, the GO analysis of LGG did not exhibit a comparable pattern to that observed in GBM. Furthermore, the role of mitochondria in ferroptosis has been predominantly investigated in GBM. It can be postulated that the regulation of mitochondria is more closely associated with the mechanism of ferroptosis in more severe glioma cells [59,60].

Additionally, amongst the three genes that exhibited survival-related characteristics in GBM, COX4IL belongs to the family of cytochrome c oxidase (COX), which is located within the respiratory chain of mitochondria. However,

the gene itself has seldom undergone extensive investigation. TRAPPC2L is a component of the transport protein particle complex and has been identified in human substantia neurons. However, there is a paucity of research on this gene in gliomas [61,62]. The two survival-related genes in LGG were SYP and PDCD1LG2. The hazard ratio for SYP was less than 1 in the survival analysis, whereas the ratio for PDCD1LG2 was greater than 1. The result for SYP, with a hazard ratio less than 1, indicated that patients with higher SYP expression had a lower risk of death than those with lower SYP expression. In addition to our own findings, another study has also identified PDCD1LG2 as a potential biomarker for predicting the prognosis of IDH mutation status in gliomas [63].

The genes selected via elastic net were subjected to analysis using random forest and XGBoost. The resulting ROC is illustrated in Figure 7. Two algorithms were used to predict the survival status of glioma patients; however, even with data balancing using ROSE, the outcomes remain sub-optimal, particularly in LGG. In general, XGBoost demonstrated superior performance compared to random forest, both with and without ROSE, particularly in the context of GBM. It would be beneficial to investigate further why the performance in the test set with ROSE was less favorable than without data balance. A potential enhancement in the predictive efficacy of the model could be accomplished by taking the intersection of the genes identified by elastic net and feature selection through random forest. This could be achieved by utilizing the variable importance as a metric. Previous research has indicated that the regulation of a few genes may exhibit differential patterns at varying stages of gliomas. However, this information is not currently available within our data set. The availability of data regarding the stages of gliomas may have enabled the development of a more specific prognostic model for GBM severity. Therefore, integrating data from GEO is considered to be a significant step towards improving the model performance and generalizing the findings.

It is recommended that future studies should also aim to incorporate additional information on glioma stages and further investigate the subtypes of gliomas, including 1p/19q non-codeletion cases, isocitrate dehydrogenase (IDH) wild-type cases and mesenchymal subtypes. This may fa-

cilitate the development of more tailored and robust prognostic models for GBM severity. Furthermore, despite the evidence that XGBoost is more effective than random forest in GBM, the prediction models did not yield optimal results. Furthermore, it would be beneficial to explore advanced machine learning techniques and algorithms, such as deep learning, as these may offer promising avenues to enhance the predictive capabilities of the model, ultimately leading to better patient stratification and personalized treatment strategies for gliomas.

Conclusion

Ferroptosis is relatively a newly discovered programmed cell death, and the relevant research especially in gliomas are still ongoing. In our study, we pointed out that the majority of the ferroptosis-related genes were higher in LGG, and the future study could gather more clinical information as the status of the progression to malignant glioma in LGG and find the potential biomarker and therefore implement targeted therapy to get a better prognosis for patients. We explored the exclusively enriched functions and pathways in GBM and LGG by filtering out the exclusively expressed genes in the two stages of gliomas. Although the performance of prediction model was not ideal, but we test-

ed that XGBoost is generally a better algorithm in predicting the survival status in GBM. With a more careful adjustment of the regularization parameter, a model with better prediction of the survival is possible, which could be a potential prognostic tool for clinicians and researchers.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

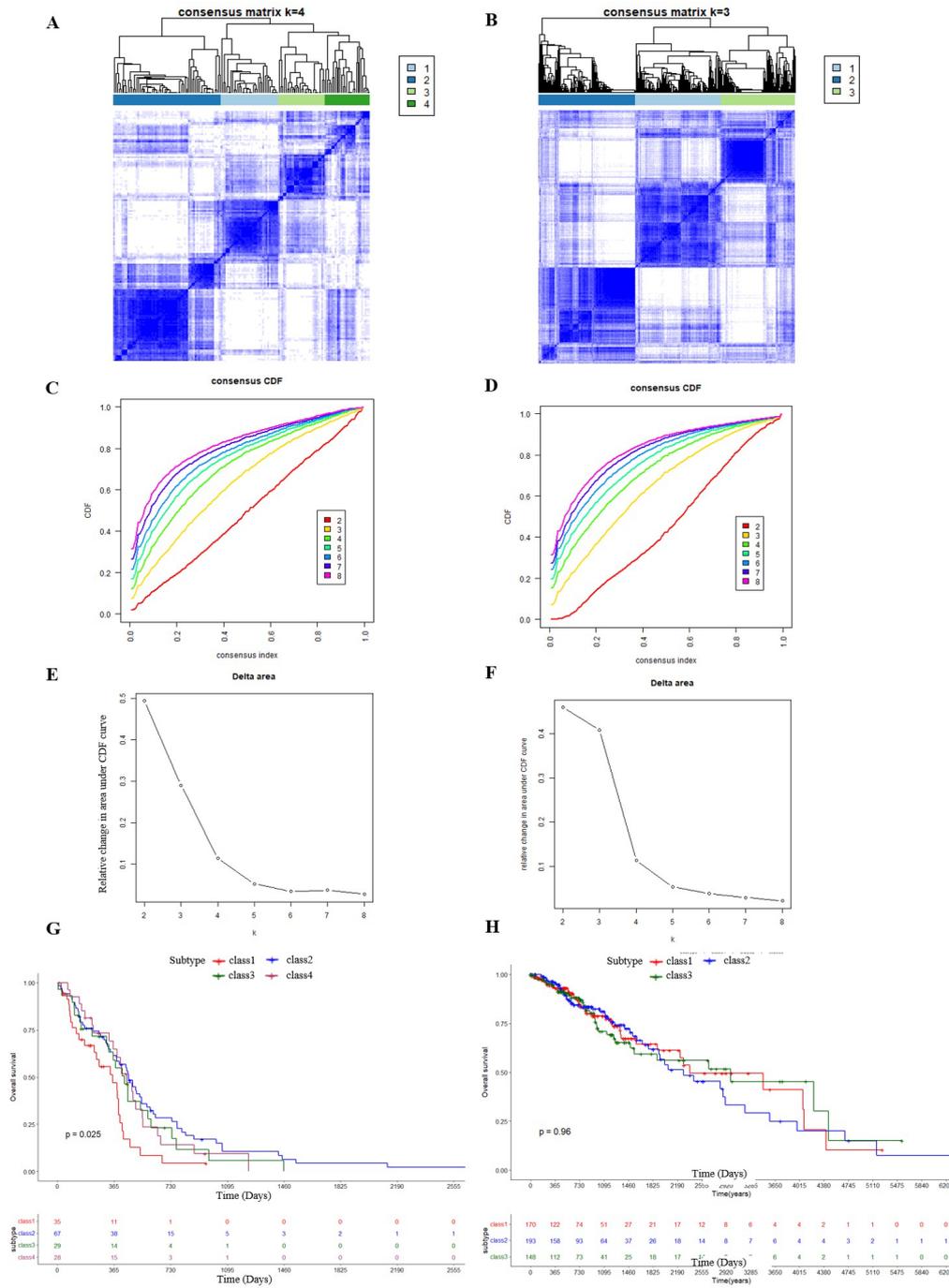
Author Contributions

Yanliang Tang: Writing – original draft, Conceptualization, Methodology, Data curation, Formal analysis, Validation, Software; Xiaoli Wang: Data curation, Formal analysis, Software; Xiaofei Tang: Data curation, Formal analysis, Software; Ye Yuan: Data curation, Formal Analysis. Wenwen Wang: Writing - original draft, Resource, Project administration.

Funding Information

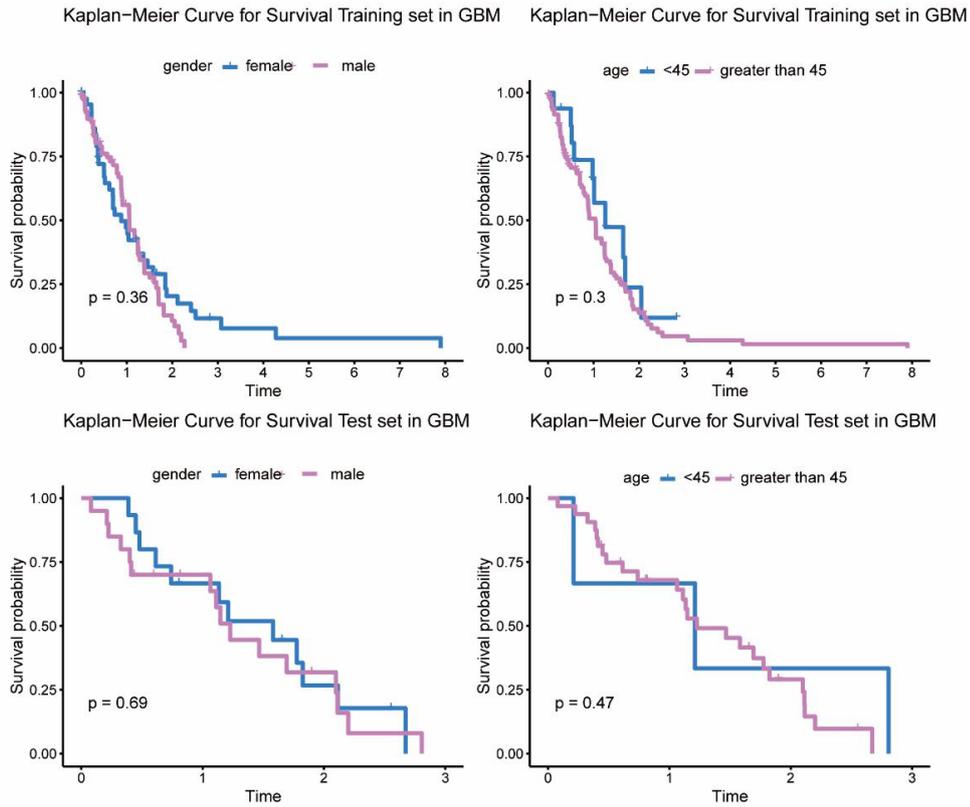
No funding.

Supplemental Figures

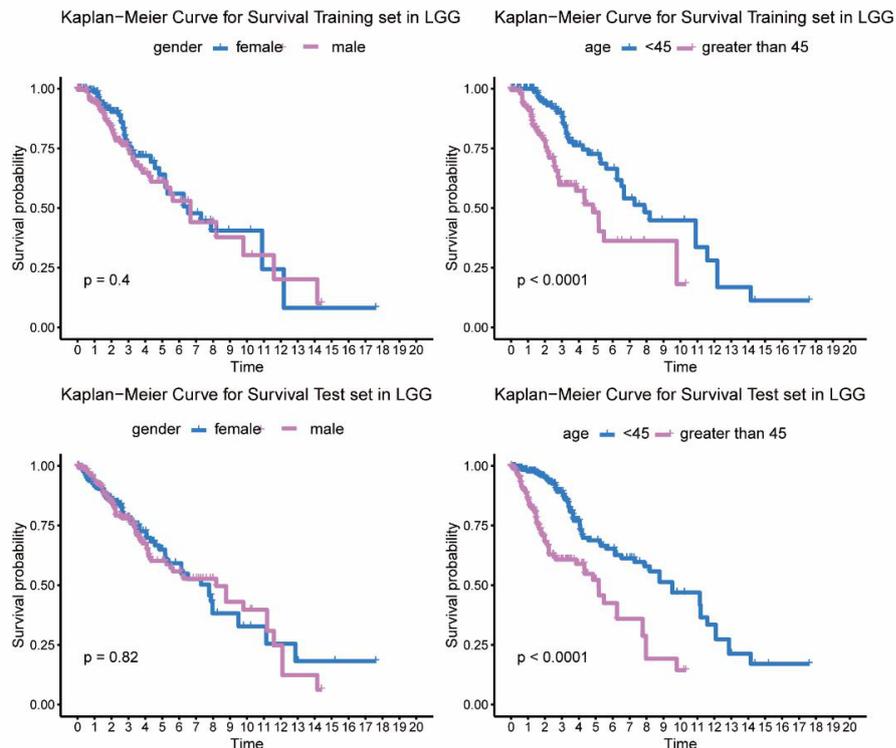


Supplementary Figure 1: Consensus clustering analysis in GBM and LGG. (A) According to the consensus clustering algorithm, the patients in GBM were clustered into 4 groups; (B) and (F) With the plot of CDF under k=2 to 8, we were able to pick the number of cluster between consensus index at 0.1 to 0.9. Ideally we tend to choose the number with the smallest difference between the two consensus index cutoff; (C) and (G) Following the CDF plotted in B and F, each point in C and G represents the area under the CDF at particular number of cluster. We usually pick the number after the sudden drop of the area from previous point, and the next k does not change much as the current number. In this case, we picked k=4 and k=3 for GBM and LGG respectively; (D) and (H) The Kaplan Meier curves of each class clustered by the algorithm, with overall survival on the y axis and time by days, whereas only the clusters in GBM showed group differences (p =0.025).

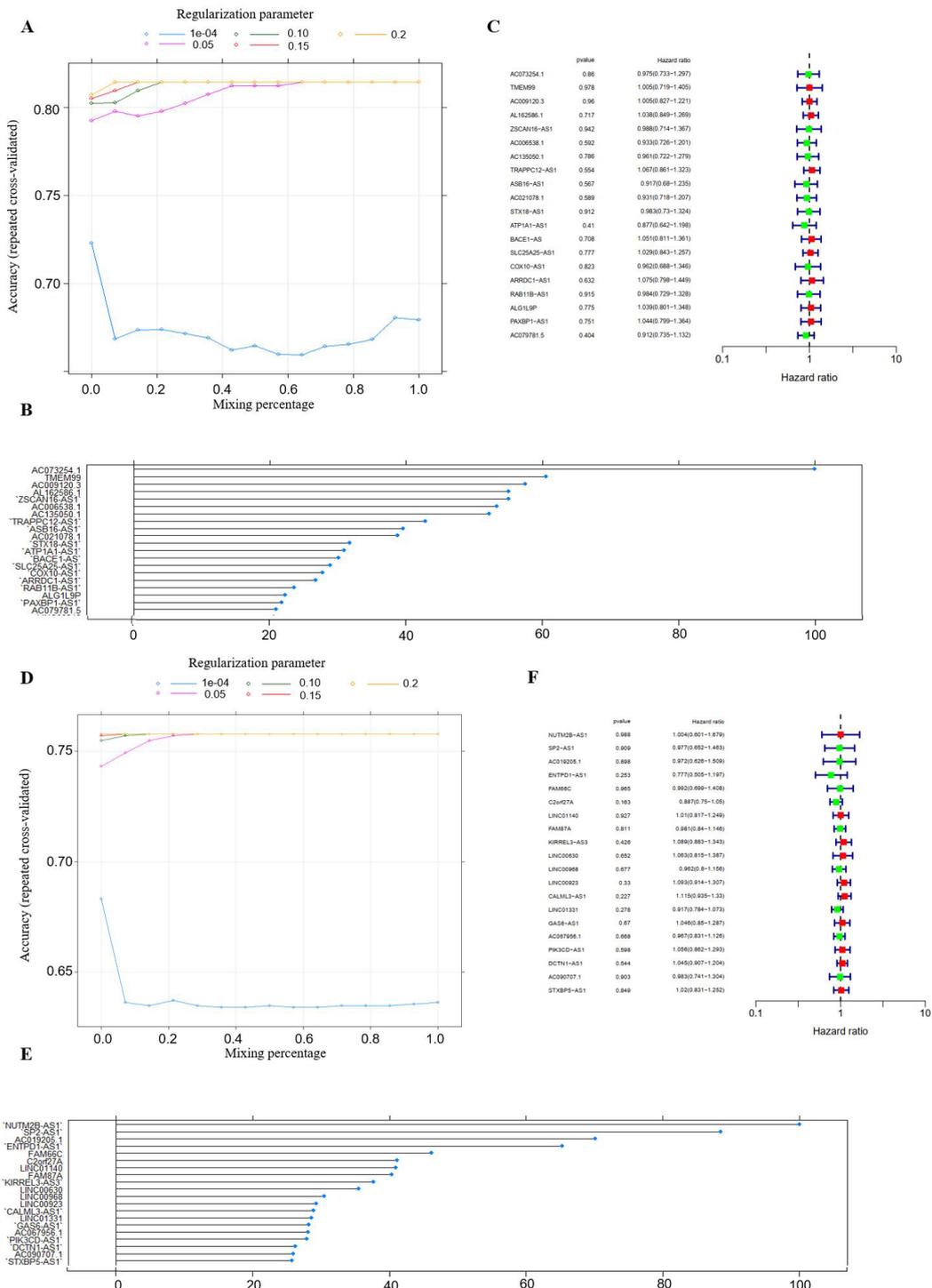
A



B



Supplement Figure 2: Assessment of the difference in survival outcome between clinical subtypes in glioma patients. A). Comparison of the survival outcome of the difference in gender and age groups in the training and test sets in GBM patients, respectively. Gender and age groups did not show survival differences in GBM patients. B). Comparison of the survival outcome of the difference in gender and age groups in the training and test sets in LGG patients, respectively. Patients under the age of 45 have a better survival rate than those over 45.



Supplementary Figure 3: Parameters of feature selection using elastic net. (A) The plot showed the changes in accuracy when the algorithm takes in different values of regularization parameter where the blue dotted line represent a series of parameter that was equally cut between 1e-04 to 0.05, the same to all the other lines; (B) The variable importance of the features selected by elastic net after we obtained the optimal value from previous plot; (C) The forest plot of the 20 selected variables and ran them into survival analysis, none of the 20 genes showed survival related ($p < 0.05$); (D) The plot of the changes in accuracy when the mixing percentage changes at different value of regularization parameters; (E) The top 20 ferroptosis related lncRNAs in LGG that were selected from elastic net, and ranked with variable importance. F). None of the 20 lncRNAs were significantly related to survival.

References

- Guo QL, Dai XL, Yin MY, Cheng HW, Qian HS, Wang H et al. (2022) Nanosensitizers for sonodynamic therapy for glioblastoma multiforme: current progress and future perspectives. *Mil Med Res.* 9: 26.
- Gusyatiner O, Hegi ME (2018) Glioma epigenetics: From subclassification to novel treatment options. *Seminars in cancer biology.* 51: 50-8.
- Kumthekar P, Raizer J, Singh S (2015) Low-grade glioma. *Cancer treatment and research.* 163: 75-87.
- Youssef G, Miller JJ (2020) Lower Grade Gliomas. *Current neurology and neuroscience reports.* 20: 21.
- Carabenciov ID, Buckner JC (2019) Controversies in the Therapy of Low-Grade Gliomas. *Current treatment options in oncology.* 20: 25.
- Miller JJ (2022) Targeting IDH-Mutant Glioma. *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics.* 19: 1724-32.
- Wu F, Li GZ, Liu HJ, Zhao Z, Chai RC, Liu YQ, et al. (2020) Molecular subtyping reveals immune alterations in IDH wild-type lower-grade diffuse glioma. *The Journal of pathology.* 251: 272-83.
- Berzero G, Di Stefano AL, Ronchi S, Bielle F, Villa C, Guillerm E, et al. (2021) IDH-wildtype lower-grade diffuse gliomas: the importance of histological grade and molecular assessment for prognostic stratification. *Neuro-oncology.* 23: 955-66.
- Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. (2021) The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-oncology.* 23: 1231-51.
- Le Rhun E, Preusser M, Roth P, Reardon DA, van den Bent M, Wen P, et al. (2019) Molecular targeted therapy of glioblastoma. *Cancer treatment reviews.* 80: 101896.
- Fletcher-Sananikone E, Kanji S, Tomimatsu N, Di Cristofaro LFM, Kollipara RK, Saha D, et al. (2021) Elimination of Radiation-Induced Senescence in the Brain Tumor Microenvironment Attenuates Glioblastoma Recurrence. *Cancer Res.* 81: 5935-47.
- Hegi ME, Genbrugge E, Gorlia T, Stupp R, Gilbert MR, Chinot OL, et al. (2019) MGMT Promoter Methylation Cutoff with Safety Margin for Selecting Glioblastoma Patients into Trials Omitting Temozolomide: A Pooled Analysis of Four Clinical Trials. *Clinical cancer research: an official journal of the American Association for Cancer Research.* 25: 1809-16.
- Ma R, Taphoorn MJB, Plaha P (2021) Advances in the management of glioblastoma. *Journal of neurology, neurosurgery, and psychiatry.* 92: 1103-11.
- Weller M, Tabatabai G, Kästner B, Felsberg J, Steinbach JP, Wick A, et al. (2015) MGMT Promoter Methylation Is a Strong Prognostic Biomarker for Benefit from Dose-Intensified Temozolomide Rechallenge in Progressive Glioblastoma: The DIRECTOR Trial. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 21: 2057-64.
- Li Z, Meng X, Wu P, Zha C, Han B, Li L, et al. (2021) Glioblastoma Cell-Derived lncRNA-Containing Exosomes Induce Microglia to Produce Complement C5, Promoting Chemotherapy Resistance. *Cancer immunology research.* 9: 1383-99.
- Liang J, Liu C, Xu D, Xie K, Li A (2022) LncRNA NEAT1 facilitates glioma progression via stabilizing PGK1. *Journal of translational medicine.* 20: 80.
- Lu C, Wei Y, Wang X, Zhang Z, Yin J, Li W, et al. (2020) DNA-methylation-mediated activating of lncRNA SNHG12 promotes temozolomide resistance in glioblastoma. *Molecular cancer.* 19: 28.
- Wei C, Wang B, Peng D, Zhang X, Li Z, Luo L, et al. (2022) Pan-Cancer Analysis Shows That ALKBH5 Is a Potential Prognostic and Immunotherapeutic Biomarker for Multiple Cancer Types Including Gliomas. *Front Immunol.* 13: 849592.
- Yu W, Ma Y, Hou W, Wang F, Cheng W, Qiu F, et al. (2021) Identification of Immune-Related lncRNA Prognostic Signature and Molecular Subtypes for Glioblastoma. *Front*

Immunol. 12: 706936.

20. Barbieri I, Kouzarides T (2020) Role of RNA modifications in cancer. *Nature reviews Cancer*. 20: 303-22.
21. Beck JD, Reidenbach D, Salomon N, Sahin U, Türeci Ö, Vormehr M, et al. (2021) mRNA therapeutics in cancer immunotherapy. *Molecular cancer*. 20: 69.
22. Goodall GJ, Wickramasinghe VO (2021) RNA in cancer. *Nature reviews Cancer*. 21: 22-36.
23. Statello L, Guo CJ, Chen LL, Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nature reviews Molecular cell biology*. 22: 96-118.
24. Liu J, Shang G (2022) The Roles of Noncoding RNAs in the Development of Osteosarcoma Stem Cells and Potential Therapeutic Targets. *Frontiers in cell and developmental biology*. 10: 773038.
25. Chen T, Liu J, Zhang H, Li J, Shang G (2022) Long Intergenic Noncoding RNA00265 Enhances Cell Viability and Metastasis via Targeting miR-485-5p/USP22 Axis in Osteosarcoma. *Frontiers in oncology*. 12: 907472.
26. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. (2012) Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell*. 149: 1060-72.
27. Liu HJ, Hu HM, Li GZ, Zhang Y, Wu F, Liu X, et al. (2020) Ferroptosis-Related Gene Signature Predicts Glioma Cell Death and Glioma Patient Progression. *Frontiers in cell and developmental biology*. 8: 538.
28. Chen Y, Mi Y, Zhang X, Ma Q, Song Y, Zhang L, et al. (2019) Dihydroartemisinin-induced unfolded protein response feedback attenuates ferroptosis via PERK/ATF4/HSPA5 pathway in glioma cells. *Journal of experimental & clinical cancer research: CR*. 38: 402.
29. Jiang X, Stockwell BR, Conrad M (2021) Ferroptosis: mechanisms, biology and role in disease. *Nature reviews Molecular cell biology*. 22: 266-82.
30. Wolpaw AJ, Shimada K, Skouta R, Welsch ME, Akavia UD, Pe'er D, et al. (2011) Modulatory profiling identifies mechanisms of small molecule-induced cell death. *Proceedings of the National Academy of Sciences of the United States of America*. 108: E771-80.
31. Chen J, Li T, Zhou N, He Y, Zhong J, Ma C, et al. (2023) Engineered Salmonella inhibits GPX4 expression and induces ferroptosis to suppress glioma growth in vitro and in vivo. *Journal of neuro-oncology*. 163: 607-22.
32. Lei G, Zhuang L, Gan B (2022) Targeting ferroptosis as a vulnerability in cancer. *Nature reviews Cancer*. 22: 381-96.
33. Shi Z, Zheng J, Liang Q, Liu Y, Yang Y, Wang R, et al. (2022) Identification and Validation of a Novel Ferroptotic Prognostic Genes-Based Signature of Clear Cell Renal Cell Carcinoma. *Cancers*. 14.
34. Xiong R, He R, Liu B, Jiang W, Wang B, Li N, et al. (2021) Ferroptosis: A New Promising Target for Lung Cancer Therapy. *Oxidative medicine and cellular longevity*. 2021: 8457521.
35. Yang L, Chen Z, Liu Y, Wang X, Li J, Ye Q (2022) Immunization Combined with Ferroptosis Related Genes to Construct a New Prognostic Model for Head and Neck Squamous Cell Carcinoma. *Cancers*. 14.
36. Zuo YB, Zhang YF, Zhang R, Tian JW, Lv XB, Li R, et al. (2022) Ferroptosis in Cancer Progression: Role of Non-coding RNAs. *International journal of biological sciences*. 18: 1829-43.
37. Zhang F, Wu L, Feng S, Zhao Z, Zhang K, Thakur A, et al. (2023) FHOD1 is upregulated in glioma cells and attenuates ferroptosis of glioma cells by targeting HSPB1 signaling. *CNS neuroscience & therapeutics*. 29: 3351-63.
38. Chen X, Wang Z, Li C, Zhang Z, Lu S, Wang X, et al. (2024) SIRT1 activated by AROS sensitizes glioma cells to ferroptosis via induction of NAD⁺ depletion-dependent activation of ATF3. *Redox biology*. 69: 103030.
39. Legendre C, Garcion E (2015) Iron metabolism: a double-edged sword in the resistance of glioblastoma to therapies. *Trends Endocrinol Metab*. 26: 322-31.
40. Yee PP, Wei Y, Kim SY, Lu T, Chih SY, Lawson C, et al. (2020) Neutrophil-induced ferroptosis promotes tumor ne-

crisis in glioblastoma progression. *Nat Commun.* 11: 5424.

41. Tu Z, Li J, Long X, Wu L, Zhu X, Huang K (2022) Transcriptional Patterns of Lower-Grade Glioma Patients with Distinct Ferroptosis Levels, Immunotherapy Response, and Temozolomide Sensitivity. *Oxidative medicine and cellular longevity.* 2022: 9408886.

42. Li X, Cheng Y, Yang Z, Ji Q, Huan M, Ye W, et al. (2024) Glioma-targeted oxaliplatin/ferritin clathrate reversing the immunosuppressive microenvironment through hijacking Fe(2+) and boosting Fenton reaction. *Journal of nano biotechnology.* 22: 93.

43. Sun Q, Xu Y, Yuan F, Qi Y, Wang Y, Chen Q, et al. (2022) Rho family GTPase 1 (RND1), a novel regulator of p53, enhances ferroptosis in glioblastoma. *Cell & bioscience.* 12: 53.

44. Yuan F, Sun Q, Zhang S, Ye L, Xu Y, Deng G, et al. (2022) The dual role of p62 in ferroptosis of glioblastoma according to p53 status. *Cell & bioscience.* 12: 20.

45. He Y, Ye Y, Tian W, Qiu H (2021) A Novel lncRNA Panel Related to Ferroptosis, Tumor Progression, and Microenvironment is a Robust Prognostic Indicator for Glioma Patients. *Frontiers in cell and developmental biology.* 9: 788451.

46. Shi J, Lai D, Zuo X, Liu D, Chen B, Zheng Y, et al. (2022) Identification of Ferroptosis-Related Biomarkers for Prognosis and Immunotherapy in Patients With Glioma. *Frontiers in cell and developmental biology.* 10: 817643.

47. Wan RJ, Peng W, Xia QX, Zhou HH, Mao XY (2021) Ferroptosis-related gene signature predicts prognosis and immunotherapy in glioma. *CNS neuroscience & therapeutics.* 27: 973-86.

48. Wang S, Nan B, Rosset S, Zhu J (2011) RANDOM LASSO. *The annals of applied statistics.* 5: 468-85.

49. Chen W, Lu Q, Lu L, Guan H (2017) Increased levels of alphaB-crystallin in vitreous fluid of patients with proliferative diabetic retinopathy and correlation with vascular endothelial growth factor. *Clinical & experimental ophthalmology.* 45: 379-84.

50. Cherneva R, Petrov D, Georgiev O, Slavova Y, Toncheva D, Stamenova M, et al. (2010) Expression profile of the small heat-shock protein alpha-B-crystallin in operated-on non-small-cell lung cancer patients: clinical implication. *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery.* 37: 44-50.

51. Wang SN, Luo S, Liu C, Piao Z, Gou W, Wang Y, et al. (2017) miR-491 Inhibits Osteosarcoma Lung Metastasis and Chemoresistance by Targeting α B-crystallin. *Molecular therapy : the journal of the American Society of Gene Therapy.* 25: 2140-9.

52. Augustus M, Pineau D, Aimond F, Azar S, Lecca D, Scamps F, et al. (2021) Identification of CRYAB(+) KCN-N3(+) SOX9(+) Astrocyte-Like and EGFR(+) PDGFRA(+) OLIG1(+) Oligodendrocyte-Like Tumoral Cells in Diffuse IDH1-Mutant Gliomas and Implication of NOTCH1 Signalling in Their Genesis. *Cancers.* 13.

53. Du Z, Wang Y, Liang J, Gao S, Cai X, Yu Y, et al. (2022) Association of glioma CD44 expression with glial dynamics in the tumour microenvironment and patient prognosis. *Computational and structural biotechnology journal.* 20: 5203-17.

54. He C, Sheng L, Pan D, Jiang S, Ding L, Ma X, et al. (2021) Single-Cell Transcriptomic Analysis Revealed a Critical Role of SPP1/CD44-Mediated Crosstalk Between Macrophages and Cancer Cells in Glioma. *Frontiers in cell and developmental biology.* 9: 779319.

55. Gao M, Yi J, Zhu J, Minikes AM, Monian P, Thompson CB, et al. (2019) Role of Mitochondria in Ferroptosis. *Molecular cell.* 73: 354-63.e3.

56. Fang X, Wang H, Han D, Xie E, Yang X, Wei J, et al. (2019) Ferroptosis as a target for protection against cardiomyopathy. *Proceedings of the National Academy of Sciences of the United States of America.* 116: 2672-80.

57. Kasukabe T, Honma Y, Okabe-Kado J, Higuchi Y, Kato N, Kumakura S (2016) Combined treatment with cotylenin A and phenethyl isothiocyanate induces strong anti-tumor activity mainly through the induction of ferroptotic cell death in human pancreatic cancer cells. *Oncology re-*

ports. 36: 968-76.

58. Wilms C, Lepka K, Häberlein F, Edwards S, Felsberg J, Pudenko L, et al. (2022) Glutaredoxin 2 promotes SP-1-dependent CSPG4 transcription and migration of wound healing NG2 glia and glioma cells: Enzymatic Taoism. *Redox biology*. 49: 102221.

59. Su J, Li Y, Liu Q, Peng G, Qin C, Li Y (2022) Identification of SSBP1 as a ferroptosis-related biomarker of glioblastoma based on a novel mitochondria-related gene risk model and in vitro experiments. *Journal of translational medicine*. 20: 440.

60. Wang Z, Ding Y, Wang X, Lu S, Wang C, He C, et al. (2018) Pseudolaric acid B triggers ferroptosis in glioma cells via activation of Nox4 and inhibition of xCT. *Cancer letters*. 428: 21-33.

61. Al-Deri N, Okur V, Ahimaz P, Milev M, Valivullah Z, Hagen J, et al. (2021) A novel homozygous variant in TRAPPC2L results in a neurodevelopmental disorder and disrupts TRAPP complex function. *Journal of medical genetics*. 58: 592-601.

62. Milev MP, Graziano C, Karall D, Kuper WFE, Al-Deri N, Cordelli DM, et al. (2018) Bi-allelic mutations in TRAPPC2L result in a neurodevelopmental disorder and have an impact on RAB11 in fibroblasts. *Journal of medical genetics*. 55: 753-64.

63. Deng X, Chen K, Ren J, Zeng J, Zhang Q, Li T, et al. (2022) A B7-CD28 Family-Based Signature Demonstrates Significantly Different Prognosis and Immunological Characteristics in Diffuse Gliomas. *Frontiers in molecular biosciences*. 9: 849723.

Submit your manuscript to a JScholar journal and benefit from:

- ¶ Convenient online submission
- ¶ Rigorous peer review
- ¶ Immediate publication on acceptance
- ¶ Open access: articles freely available online
- ¶ High visibility within the field
- ¶ Better discount for your subsequent articles

Submit your manuscript at
<http://www.jscholaronline.org/submit-manuscript.php>