

doi:10.3969/j.issn.1005-3697.2024.07.002

❖ 基础研究 ❖

利用 TCGA 数据库构建胃癌铁死亡相关 LncRNA 的预后模型

熊黎¹, 向尹¹, 刘燕群², 谭劲松³, 张若兰³, 阳丽红³, 刘康³

(1. 乐山市人民医院检验科, 四川 乐山 614000; 2. 郫都区中医院检验科, 四川 成都 610000; 3. 川北医学院第二临床医学院·南充市中心医院组织工程与干细胞研究所, 四川 南充 637000)

【摘要】目的: 通过筛选与胃癌铁死亡相关的长链非编码 RNA (LncRNA), 构建预后风险模型, 协助评估胃癌患者预后。**方法:** 在癌症基因组数据库 (The Cancer Genome Atlas TCGA) 下载胃癌的转录组和临床数据, 从 NCBI 网站获得 103 个铁死亡相关基因。采用 R 软件进行 Kaplan-Meier (KM) 生存分析和单因素 COX 回归分析筛选出关键的 LncRNA, 将临床样本按 7:3 的比例随机分为训练集和验证集, 利用多因素 COX 回归分析建立胃癌铁死亡相关 LncRNA 的预后模型。两组分别进行生存分析、风险曲线分析、单因素和多因素独立预后分析、多指标受试者操作特征 (ROC) 曲线的绘制及列线图分析, 最后对训练集进行 GSEA 的 KEGG 分析。**结果:** 单因素 Cox 回归分析得到 13 个 LncRNA (LINC00106、TNFRSF10A-AS1、AC244153.1、AC005586.1、AP001318.2、MAGI2-AS3、AP001528.2、MSC-AS1、PVT1、AC037198.1、AC010719.1、HAGLR), 经多因素回归分析筛选出 5 个胃癌铁死亡相关的 LncRNA (AC005586.1、AP001318.2、AP001528.2、AC010719.1、HAGLR), 训练集和验证集构建的预后模型趋势一致, 且 ROC 评估风险模型的预测能力大于其他临床指标。构建的 5 个 LncRNA 列线图, 预测 1 年、2 年、3 年校准曲线也基本和 45° 的对角线重合, 模型预测性能良好。**结论:** LncRNA AC005586.1、AP001318.2、AP001528.2、AC010719.1、HAGLR 胃癌铁死亡相关 LncRNA 可有效预测胃癌患者的预后。

【关键词】 胃癌; 铁死亡; 长链非编码 RNA; 预后风险模型

【中图分类号】 R735.2 **【文献标志码】** A

A prognostic model of ferroptosis-related lncRNA in gastric cancer was constructed using the TCGA database

XIONG Li¹, XIANG Yin¹, LIU Yan-qun², TAN Jin-song³, ZHANG Ruo-lan³, YANG Li-hong³, LIU Kang³

(1. Department of Clinical Laboratory, the People's Hospital of Leshan, Leshan 614000; 2. Department of Clinical Laboratory, Pidu District Hospital of Traditional Chinese Medicine, Chengdu 610000; 3. Institute of Tissue Engineering and Stem Cells, Nanchong Central Hospital, the Second Clinical Medical College, North Sichuan Medical College, Nanchong 637000, Sichuan, China)

【Abstract】Objective: The prognostic risk model was constructed to help evaluate the prognosis of patients with gastric cancer by screening long non-coding RNA (LncRNA) associated with ferroptosis in gastric cancer. **Methods:** The transcripts and clinical date of gastric cancer were downloaded from the Cancer Genome Atlas (TCGA) and 103 ferroptosis-related genes were downloaded from NCBI. The expression matrix of LncRNA associated with ferroptosis in gastric cancer was obtained through data processing by R and Perl software. Combined that with survival time and then the key LncRNAs were screened out by KM and univariate Cox regression an analysis using R software. Randomly divided clinical samples into training and validation sets in a 7:3 ratio. The prognostic prediction models of the Train and Test group were established by multivariate Cox regression model. Survival analysis, risk cure analysis, single-factor and multi-factor independent prognostic analysis, multi-ROC curve drawing, etc., were performed for the two groups respectively. Finally, line chart and KEGG analysis of GSEA were performed for the train. **Results:** 13 LncRNAs (LINC00106, TNFRSF10A-AS1, AC244153.1, AC005586.1, AP001318.2, MAGI2-AS3, AP001528.2, MSC-AS1, PVT1, AC037198.1, AC010719.1, HAGLR) were obtained by univariate Cox regression analysis, and five LncRNAs (AC005586.1, AP001318.2, AP001528.2, AC010719.1, HAGLR) related to ferroptosis were screened out by multivariate regression analysis. 5 LncRNAs related to ferroptosis in gastric cancer were screened through multiple regression analysis (AC005586.1, AP001318.2, AP001528.2, AC010719.1, HAGLR), the trend of the

基金项目: 国家自然科学基金青年基金 (82203851); 四川省科技厅省院省校合作项目 (2023YFSY0045); 四川省自然科学基金 (2023NSFSC0731)

作者简介: 熊黎 (1998 -), 男, 硕士, 初级检验技师。E-mail: 2639532361@qq.com

通讯作者: 刘康, 博士。E-mail: liukang@nsmc.edu.cn

prognostic models constructed from the training and validation sets was consistent. The predictive power of ROC assessment risk model was greater than that of other clinical indicators. The constructed 5 lncRNA histograms showed that the 1-year, 2-year and 3-year calibration curves coincided with the diagonal of 45°, indicating that the risk model had a good predictive ability. **Conclusion:** AC005586. 1, AP001318. 2, AP001528. 2, AC010719. 1 and HAGLR associated with iron death in gastric cancer can effectively predict the prognosis of gastric cancer patients.

[Key words] Gastric cancer; Ferroptosis; LncRNA; Prognostic risk model

胃癌是全球第五大最常见癌症和第三大致死性癌症,尽管历年发病率和死亡率有所降低,但总体生存率仍很低^[1-2],就诊患者多为胃癌晚期,死亡率居高不下,为我国死亡率第三高的恶性肿瘤^[3]。靶向作用于肿瘤细胞死亡过程是临床抗癌治疗的常用策略^[4],而铁死亡是一种由脂质过度氧化引起的铁依赖形式的调节性死亡^[5-6],或可成为抗癌药物研发和抗癌治疗的新靶点^[7]。铁死亡基因在胃癌细胞的增殖和死亡中发挥重要作用,如生长分化因子 15 (growth differentiation factor 15, GDF15) 可抑制铁死亡的发生^[8],而花生四烯酸 15 脂肪氧化酶 (arachidonate 15-Lipoxygenase, ALOX15) 则促进胃癌细胞的铁死亡^[9]。袁漫秋等^[10]已经成功构建了胃癌铁死亡相关基因的预后模型。

长链非编码 RNA (long non-coding RNA, LncRNA) 参与转录水平的调控,与胃癌的增殖、发展、转移和预后密切相关^[11-12],因此,LncRNA 也有可能成为胃癌新的生物标志物。LncRNA 肿瘤敏感因子 11 (cancer susceptibility 11, CASC11) 可稳定肝癌细胞中溶质载体 7A11 (solute carriers 7A11, SLC7A11) mRNA 表达水平进而调控铁死亡^[13]。因此,本研究拟构建胃癌铁死亡相关 LncRNA 的预后模型,旨在挖掘胃癌铁死亡相关的 LncRNA,为胃癌风险及预后评估提供具有预后价值的生物标志物。

1 材料与方法

1.1 数据获取与处理

从肿瘤基因组图谱 (Cancer Genome Atlas Program, TCGA) 下载胃癌的转录组和临床数据,转录组数据包括 32 个癌旁样品,375 个胃癌组织样品。通过 perl 软件将转录组数据分为 mRNA 和 LncRNA 两个数据集;从美国国家生物技术信息中心 (NCBI) 下载 103 个铁死亡相关基因,利用 R 软件获取了相应基因在 mRNA 数据集中的表达值。将铁死亡相关基因表达矩阵和 LncRNA 数据集用 R 语言“limma”包进行处理,得到铁死亡相关 LncRNA 表达矩阵 (相关系数 $R^2 > 0.3, P < 0.01$)。

1.2 Cox 模型的构建和验证

将临床生存数据和铁死亡相关 LncRNA 进行合

并,再通过 R 语言“survival”包进行 KM 和单因素 Cox 回归分析,筛选出具有预测价值的铁死亡相关 LncRNA ($P < 0.03$)。使用 R 软件将上述得到的铁死亡相关 LncRNA 表达矩阵以 7:3 的比例分为训练集 (train) 和验证集 (test) 并进行多因素 Cox 回归分析,进而构建胃癌铁死亡相关 LncRNA 预后模型。利用风险值计算公式得出训练集和验证集患者的风险值,通过训练集风险值的中位数将患者分为高风险组和低风险组,并在 R 软件绘制预后相关的 LncRNA 生存曲线。高低风险组运用“survival”包和“survminer”包分别构建训练集和验证集的生存曲线、独立预后森林图、列线图。采用 R 软件的“survivalROC”包来绘制受试者操作特性 (receiver operating characteristic curve, ROC) 曲线图来评估模型预测准确性。风险值计算公式:风险值 (RiskScore) = $\sum_{i=1}^n (\exp Gene_i \times coef_i)$

1.3 功能富集分析

通过 GSEA 软件进行 KEGG 通路分析,找出在高风险和低风险中高表达的通路。以低风险组作为对照组,阈值过滤条件设置为 $FDR < 0.05$ 。

1.4 统计学分析

采用 R 软件 v4.02 及 perl 软件 v5.32 进行统计学分析。胃癌铁死亡相关 LncRNA 相关性检验,阈值过滤条件为相关系数 > 0.3 ,且 $P < 0.001$ 。采用 Kaplan-Meier (KM) 方法进行不同组的生存差异比较,使用单因素和多因素 Cox 回归分析来确定独立预后因子。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 确定铁死亡相关 LncRNA

采用 KM 方法和单因素 Cox 回归分析筛选出 13 个有预后价值的 lncRNA,再通过多因素 Cox 回归分析过滤得到 AC005586. 1、AP001318. 2、AP001528. 2、AC010719. 1、HAGLR 共 5 个具有预后价值的 LncRNA。LncRNA 的生存曲线显示 HAGLR 和 AP001528. 2 低表达时生存率更高,而 AP001318. 2、AC005586. 1、AC010719. 1 三个 LncRNA 高表达预示生存率更高。见图 1。

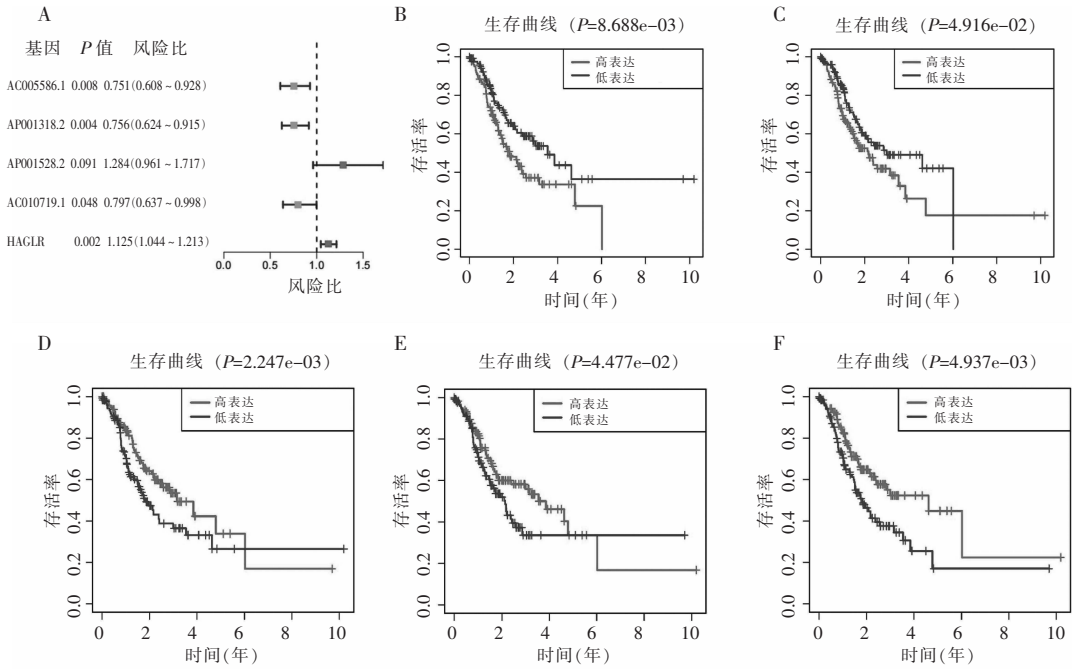


图 1 铁死亡相关 LncRNA 的生存曲线

A. 多因素回归分析; B-F. HAGLR、AP001528.2、AP001318.2、AC010719.1、AC005586.1 的生存曲线。

2.2 构建预后模型与验证

以风险值的中位数作为分界标准,将患者分为高风险组和低风险组。生存分析中,训练集和验证集生存曲线均显示低风险组生存率高于高风险组 ($P < 0.05$)。训练集组和验证集组 ROC 曲线下面

积(AUC)分别为 0.698 和 0.645,均大于其他临床指标。训练集和验证集的风险曲线、生存状态与生存分析趋势一致,风险值越高,死亡率越高。见图 2 及图 3。

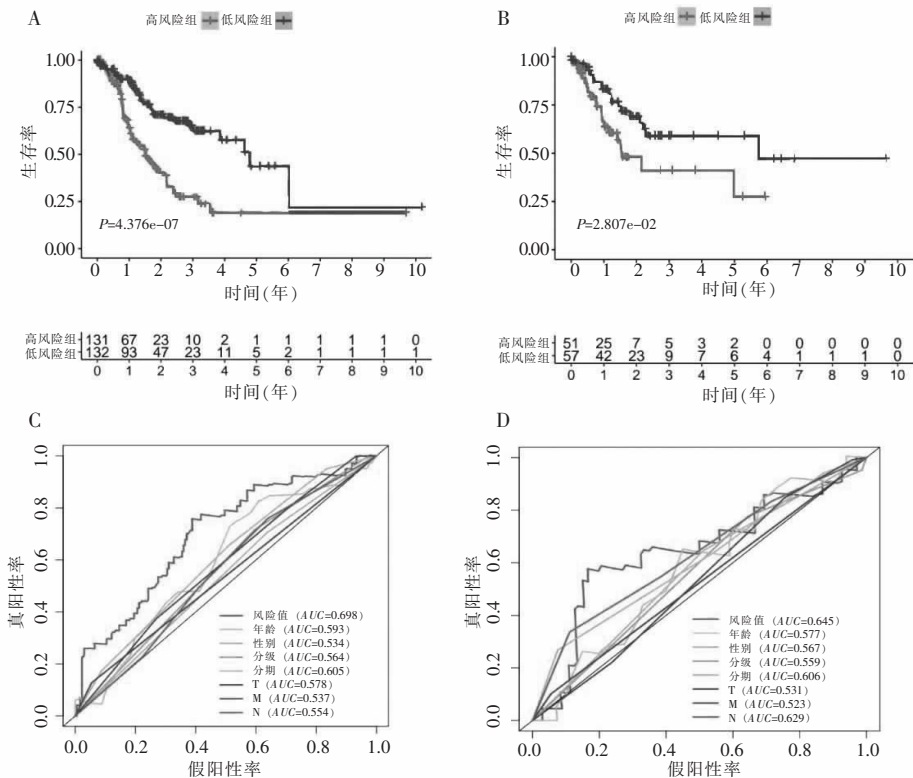


图 2 胃癌患者铁死亡相关 LncRNA 的风险生存曲线和预后指标

A. 训练集风险生存曲线; B. 验证集风险生存曲线; C. 训练集 ROC 曲线; D. 验证集 ROC 曲线。

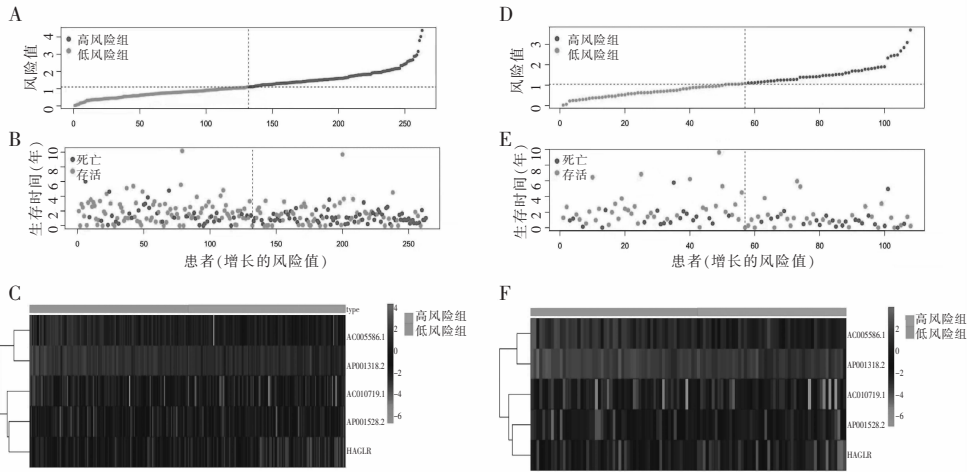


图3 胃癌患者铁死亡相关 LncRNA 特征性分布

A. 训练集风险评分曲线;B. 训练集生存状态图;C. 训练集 5 个铁死亡相关 LncRNA 表达热图;D. 验证集风险评分曲线验证集;E. 验证集生存状态图;F. 验证集 5 个铁死亡相关 LncRNA 表达热图。

2.3 单因素和多因素独立预后分析

将训练集中的每一位患者的风险值、年龄、性别、肿瘤分级、T、M、N 分期分别进行单因素和多因素 Cox 回归分析,风险值差异在单因素和多因素

Cox 回归分析中均具有统计学意义,趋势与验证集一致(均 $P < 0.05$)。构建模型的有效性较高。见图 4。

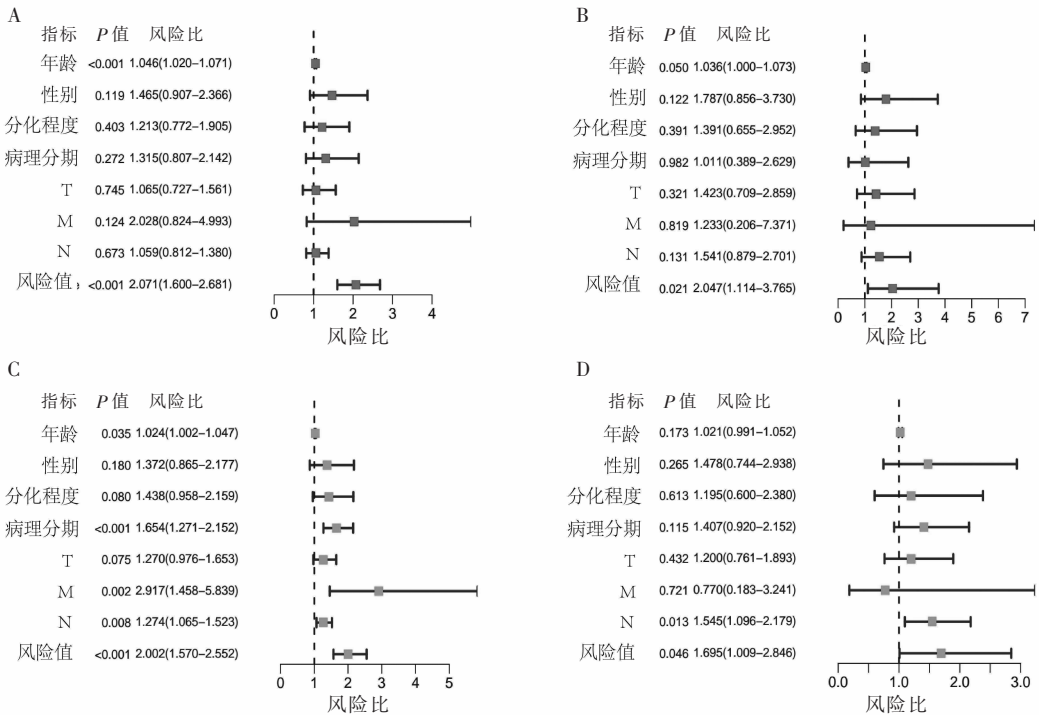


图4 独立预后分析

A. 训练集多因素 Cox 分析;B. 验证集多因素 Cox 分析;C. 训练集单因素 Cox 分析;D. 验证集单因素 Cox 分析。

2.4 胃癌预后预测的列线图

运用列线图计算风险值并预测生存率,训练集和验证集基本趋势一致;绘制预测 1 年、2 年、3 年生

存曲率的校准曲线与 45° 的对角线基本重合,表明列线图的实际生存和预测生存一致性良好。见图 5 及图 6。

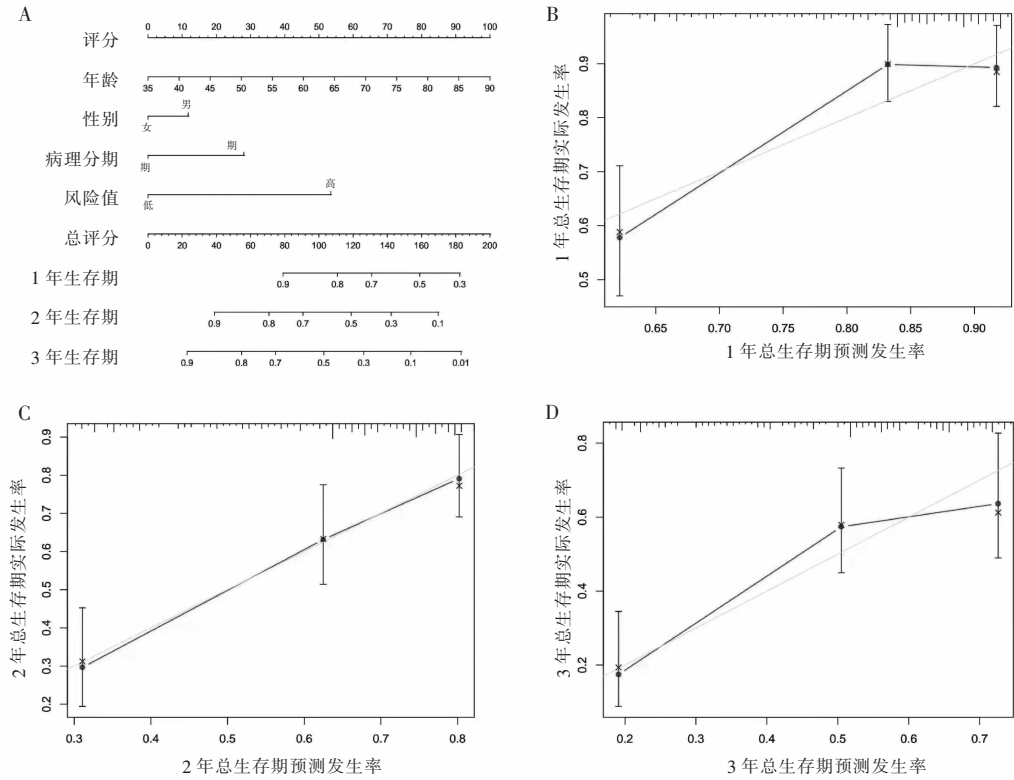


图 5 训练集列线图 and 校准曲线

A. 训练集预测 1 年、2 年、3 年总体生存率的列线图; B. 预测患者 1 年总体生存率的校准曲线; C. 预测患者 2 年总体生存率的校准曲线; D. 预测患者 3 年总体生存率的校准曲线。

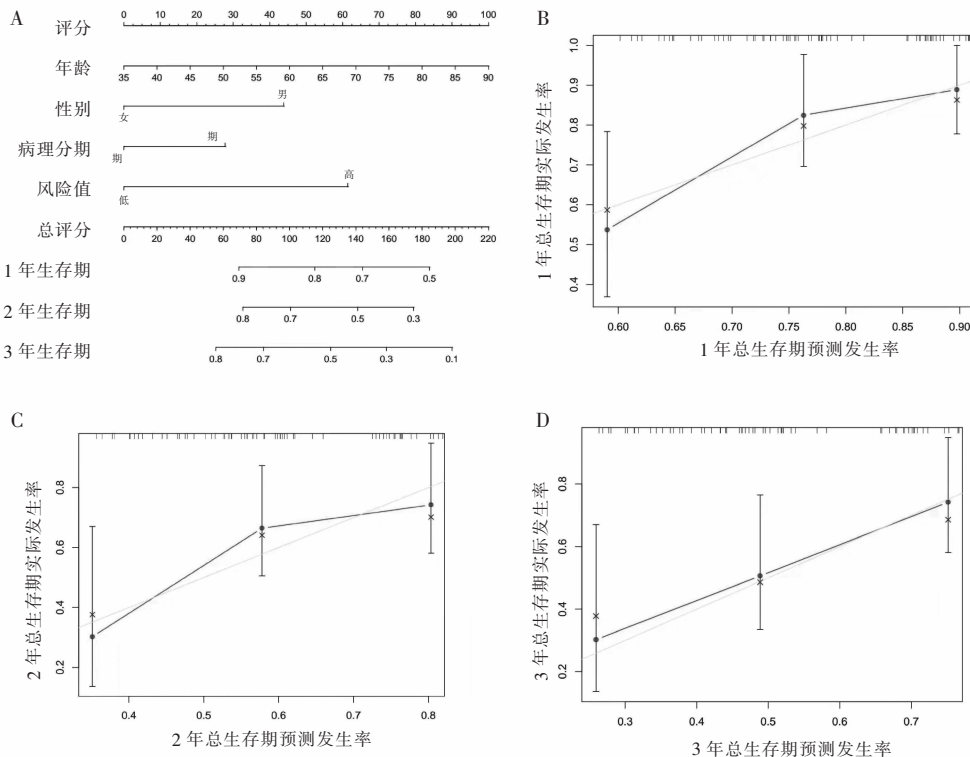


图 6 验证集列线图 and 校准曲线

A. 验证集预测 1 年、2 年、3 年总体生存率的列线图; B. 预测患者 1 年总体生存率的校准曲线; C. 预测患者 2 年总体生存率的校准曲线; D. 预测患者 3 年总体生存率的校准曲线。

2.5 功能富集分析

用GSEA软件进行KEGG富集分析高低风险患者的信号通路富集情况,得到25个信号通路,其中

差异最为显著的6条信号通路均为在高风险患者中活跃,阈值滤过标准为FDR p-val < 0.05。见表1。

表1 KEGG富集的25个信号通路

KEGG名称	SIZE	NOM p-val	FDR q-val	FWER p-val
肥厚型心肌病	83	0	8.88e-04	0.001
细胞外基质受体信号通路	84	0	4.44e-04	0.001
局部粘着斑激酶信号通路	199	0	0.001752505	0.005
扩张型心肌病	90	0	0.001314379	0.005
补体与凝血级联信号通路	69	0	0.00172258	0.006
右心室失常病信号通路	74	0	0.005479215	0.024
肾素_血管紧张素信号通路	17	0	0.006208129	0.03
神经受体配体相互作用信号通路	271	0	0.006608783	0.032
钙离子信号通路	177	0	0.006275093	0.033
肌动蛋白细胞骨架	213	0	0.006788095	0.038
神经节系列鞘糖脂生物合成	15	0	0.007070727	0.042
造血信号通路	85	0.002004008	0.007167199	0.046
细胞因子受体相互作用	264	0.002096436	0.007199816	0.047
葡萄糖基聚糖生物合成硫酸软骨素	22	0	0.007131334	0.051
细胞黏附分子	131	0.002066116	0.011356534	0.083
轴突向信号通路	129	0.002040816	0.019501675	0.129
血管平滑肌收缩	115	0.00209205	0.019087339	0.131
细胞间间隙连接	90	0	0.026759192	0.166
血管加压素信号通路	44	0.002040816	0.03639905	0.221
肾细胞癌	70	0.006134969	0.037541438	0.235
甘氨酸葡萄糖降解	21	0.007984032	0.036306035	0.241
HEDGEHOG信号通路	56	0.014084507	0.039941214	0.259
JAK_STAT信号通路	155	0.013972056	0.040563237	0.265

3 讨论

铁死亡机制在肿瘤领域的应用进展迅速,为胃癌治疗提供了新的理论基础和方法^[14]。胃癌组织中铁代谢异常诱导铁死亡,干预铁代谢通路,可对胃癌细胞进行选择性的杀伤,而调控铁死亡因子则有望改善预后^[15]。LncRNA和铁死亡具有密切相关,Mao等^[16]发现LncRNA P53RRA与Ras-GTP酶激活蛋白SH3结构域[GTPase activating protein (SH3 domain) binding protein 1, G3BP1]相互结合调控肿瘤抑制基因P53在胞核中的表达情况进而调节肿瘤细胞铁死亡。本研究中,5个LncRNA(AC005586.1、AP001318.2、AP001528.2、AC010719.1、HAGLR)生存曲线均具有显著差异。AC005586.1和AC010719.1可作为胃癌糖酵解相关LncRNA预测模型基因^[17];HAGLR通过海绵吸附miR-338-3p进而靶向LDHA-糖酵解途径促进胃癌5-Fu耐药^[18];AP001318.2、AP001528.2尚未见胃癌相关报道。HAGLR为差异最显著的高风险LncRNA,在癌症进展过程中起着重要的调控作用,具有作为胃癌靶向治疗研究的潜

在价值。食管癌中HAGLR充当microRNA-143-5p的竞争性内源RNA,增加了溶酶体相关膜糖蛋白(lysosome-associated membrane protein, LAMP)的表达,促进肿瘤侵袭转移^[19];HAGLR高表达可促进结肠癌和非小细胞肺癌进展^[20-21];此外,HAGLR可作为乳腺癌和肺腺癌治疗的作用靶点^[22-23]。Jia等^[24]通过RNA测序发现AP001528.2可能是乳腺癌发生发展的关键LncRNA,而AC010719.1在胰腺癌中可作为铜死亡相关LncRNA构建预后风险模型^[25]。

综上,由5个胃癌铁死亡相关LncRNA:AC005586.1、AP001318.2、AP001528.2、AC010719.1、HAGLR构建的胃癌预测预后模型均显示风险值越高生存率越低。

参考文献

- [1] Schrupf T, Behrens HM, Haag J, et al. FGFR2 overexpression and compromised survival in diffuse-type gastric cancer in a large central European cohort[J]. PLoS One, 2022, 17(2): e0264011.
- [2] Li J, Wang Y, Wang R, et al. Spectral CT for preoperative prediction of lymphovascular invasion in resectable gastric cancer: with external prospective validation[J]. Frontiers in Oncology, 2022,

- 12:942425.
- [3] 吴腾飞,赵瑾瑾,林建秀,等.胃癌靶向治疗研究进展[J].中国临床研究,2024,37(2):165-170.
- [4] Xing N, Du Q, Guo S, *et al.* Ferroptosis in lung cancer: a novel pathway regulating cell death and a promising target for drug therapy[J]. *Cell Death Discovery*, 2023, 9(1):110.
- [5] Chen X, Kang R, Kroemer G, *et al.* Broadening horizons: the role of ferroptosis in cancer[J]. *Nature Reviews Clinical Oncology*, 2021, 18(5):280-296.
- [6] Liu LX, Heng JH, Deng DX, *et al.* Sulconazole induces PANoptosis by triggering oxidative stress and inhibiting glycolysis to increase radiosensitivity in esophageal cancer [J]. *Molecular & Cellular Proteomics*, 2023, 22(6):100551.
- [7] Su Y, Zhao B, Zhou L, *et al.* Ferroptosis, a novel pharmacological mechanism of anti-cancer drugs [J]. *Cancer Letters*, 2020, 483:127-136.
- [8] Chen L, Qiao L, Bian Y, *et al.* GDF15 knockdown promotes erastin-induced ferroptosis by decreasing SLC7A11 expression [J]. *Biochemical and Biophysical Research Communications*, 2020, 526(2):293-299.
- [9] Zhang H, Deng T, Liu R, *et al.* CAF secreted miR-522 suppresses ferroptosis and promotes acquired chemo-resistance in gastric cancer[J]. *Molecular Cancer*, 2020, 19(1):43.
- [10] 袁漫秋,裴建赢,何鑫,等.胃癌预后相关铁死亡基因筛选及预后预测模型构建[J].山东医药,2021,61(30):6-9.
- [11] Sun TT, He J, Liang Q, *et al.* LncRNA GClnc1 promotes gastric carcinogenesis and may act as a modular scaffold of WDR5 and KAT2A complexes to specify the histone modification pattern[J]. *Cancer Discovery*, 2016, 6(7):784-801.
- [12] Ghafouri-Fard S, Taheri M. Long non-coding RNA signature in gastric cancer[J]. *Experimental and Molecular Pathology*, 2020, 113:104365.
- [13] Chen F, Wang L. Long noncoding RNA CASC11 suppresses sorafenib-triggered ferroptosis via stabilizing SLC7A11 mRNA in hepatocellular carcinoma cells[J]. *Discover Oncology*, 2023, 14(1):145.
- [14] Feng S, Tang D, Wang Y, *et al.* The mechanism of ferroptosis and its related diseases[J]. *Molecular Biomedicine*, 2023, 4(1):33.
- [15] Zheng Y, Wu S, Huang X, *et al.* Ferroptosis-related lncRNAs act as novel prognostic biomarkers in the gastric adenocarcinoma microenvironment, immunotherapy, and chemotherapy[J]. *Oxidative Medicine and Cellular Longevity*, 2023, 2023:9598783.
- [16] Mao C, Wang X, Liu Y, *et al.* A G3BP1-interacting lncRNA promotes ferroptosis and apoptosis in cancer via nuclear sequestration of p53[J]. *Cancer Research*, 2018, 78(13):3484-3496.
- [17] Liao T, Lu Y, Li W, *et al.* Construction and validation of a glycolysis-related lncRNA signature for prognosis prediction in Stomach Adenocarcinoma[J]. *Frontiers in Genetics*, 2022, 13:794621.
- [18] Hu J, Huang L, Ding Q, *et al.* Long noncoding RNA HAGLR sponges miR-338-3p to promote 5-Fu resistance in gastric cancer through targeting the LDHA-glycolysis pathway [J]. *Cell Biology International*, 2022, 46(2):173-184.
- [19] Yang C, Shen S, Zheng X, *et al.* Long noncoding RNA HAGLR acts as a microRNA-143-5p sponge to regulate epithelial-mesenchymal transition and metastatic potential in esophageal cancer by regulating LAMP3 [J]. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, 2019, 33(9):10490-10504.
- [20] Sun W, Nie W, Wang Z, *et al.* Lnc HAGLR Promotes Colon Cancer Progression Through Sponging miR-185-5p and Activating CDK4 and CDK6 in vitro and in vivo [J]. *OncoTargets and Therapy*, 2020, 13:5913-5925.
- [21] Lu C, Ma J, Cai D. Increased HAGLR expression promotes non-small cell lung cancer proliferation and invasion via enhanced de novo lipogenesis [J]. *Tumour Biology*, 2017, 39(4):1010428317697574.
- [22] Mahlab-Aviv S, Zohar K, Cohen Y, *et al.* Spliceosome-associated microRNAs signify breast cancer cells and portray potential novel nuclear targets [J]. *International Journal of Molecular Sciences*, 2020, 21(21):8132.
- [23] Guo X, Chen Z, Zhao L, *et al.* Long non-coding RNA-HAGLR suppressed tumor growth of lung adenocarcinoma through epigenetically silencing E2F1 [J]. *Experimental Cell Research*, 2019, 382(1):111461.
- [24] Jia X, Lei H, Jiang X, *et al.* Identification of crucial lncRNAs for luminal A breast cancer through RNA sequencing[J]. *International Journal of Endocrinology*, 2022, 2022:6577942.
- [25] Chen L, Zhang L, He H, *et al.* Systemic analyses of cuproptosis-related lncRNAs in pancreatic adenocarcinoma, with a focus on the molecular mechanism of LINC00853 [J]. *International Journal of Molecular Sciences*, 2023, 24(9):7923.

(收稿日期:2023-09-15

修回日期:2024-04-08)