

CD8⁺ T 细胞免疫相关基因标记有效预测胃癌预后和治疗反应性

孙学增¹, 苏日顺¹, 李文超², 王小群¹, 尹松成¹, 陈景耀^{1*}

1.中山大学附属第七医院 消化医学中心, 广东 深圳 518107

2.山东第一医科大学附属皮肤病医院 中心实验室, 山东 济南 250000

【摘要】 目的 探索 CD8⁺ T 细胞免疫相关基因与胃癌免疫治疗及患者预后的关系。**方法** 从癌症基因组图谱(the cancer genome atlas, TCGA)数据库下载胃癌相关转录组测序数据和对应的患者临床资料, 并将 TCGA 数据集的数据($n=407$)和基因表达综合(gene expression omnibus, GEO)数据集的数据($n=433$)分别作为训练集和验证集, 2 个数据集合并作为总体样本($n=840$)。基于单因素 Cox 回归、Lasoo 回归分析构建 CD8⁺ T 细胞免疫相关基因的风险评分模型并进行各数据集的验证。根据训练集的中位风险评分, 将训练集、验证集和总体样本分别划分为高风险和低风险。Kaplan–Meier 生存分析和受试者操作特征曲线(receiver operating characteristic curve, ROC 曲线)被用于验证模型的预测性能。通过单因素和多因素 Cox 回归分析评估风险评分等临床指标的预后价值。结合风险评分、年龄和肿瘤 TNM 分期构建预测胃癌患者总生存率的列线图, ROC 曲线与校准曲线被用于评估列线图的预测准确性。借助风险评分进行免疫细胞浸润分析、基因突变分析和免疫治疗疗效预测。使用 IMvigor 210 数据集验证风险评分模型用于预测其他肿瘤患者对免疫治疗反应的效能。收集 2023 年 8 月至 2024 年 4 月中山大学附属第七医院收治的 15 例胃癌患者的胃癌组织样本和匹配的癌旁组织样本, 采用荧光定量聚合酶链反应检测 CD8⁺ T 细胞免疫相关基因在胃癌组织中的差异表达。**结果** 共鉴定出 10 个具有预后价值的 CD8⁺ T 细胞免疫相关基因。风险评分是胃癌患者预后的独立影响因素($HR=3.495$, $95\%CI$ 2.072~5.894, $P<0.001$)。Kaplan–Meier 生存曲线显示高风险患者的总生存率低于低风险患者($P<0.05$)。构建的列线图预测胃癌患者的 1 年、3 年和 5 年总生存率与实际总生存率相接近。高风险患者具有较差的预后, 且伴有免疫抑制、低频基因突变和较低的免疫检查点分子表达。风险评分模型在 IMvigor 210 数据集中具有较好的预测效能。CD8⁺ T 细胞免疫相关基因在胃癌组织样本中的表达与癌旁组织样本存在差异 ($P<0.05$)。**结论** 本研究所构建的风险评分模型对于胃癌肿瘤微环境、预后和肿瘤突变负荷具有重要意义, 为改善胃癌患者预后和治疗反应性奠定基础。

【关键词】 CD8⁺ T 细胞; 胃癌; 免疫特征; 免疫浸润; 免疫治疗

CD8⁺ T cell immune-related gene signature effectively predicts prognosis and treatment responsiveness in gastric cancer

Sun Xuezheng¹, Su Rishun¹, Li Wenchao², Wang Xiaoqun¹, Yin Songcheng¹, Chen Jingyao^{1*}

1.Digestive Diseases Center, the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518107, Guangdong, China

2.Central Laboratory of Hospital for Skin Diseases, Shandong First Medical University, Jinan 250000, Shandong, China

*Corresponding author: Chen Jingyao, E-mail: chenjy589@mail.sysu.edu.cn

【Abstract】 Objective To explore the relationship between CD8⁺ T cell immune-related genes and immunotherapy and prognosis of gastric cancer. **Method** The transcriptome sequencing data of gastric cancer and the corresponding clinical data of patients were downloaded from the cancer genome atlas (TCGA)

基金项目: 广东省消化系统恶性肿瘤防治研究重点实验室(2021B1212040006)

* 通信作者: 陈景耀, E-mail: chenjy589@mail.sysu.edu.cn

database, and the data of TCGA dataset ($n=407$) and gene expression omnibus (GEO) dataset ($n=433$) were used as the training set and validation set, respectively, and the two datasets were combined as population samples ($n=840$). Based on univariate Cox regression and Lasso regression analysis, the risk scoring model of CD8⁺ T cell immune-related genes was constructed, and each dataset was validated. According to the median risk score of the training set, the training set, validation set, and population samples were classified as high-risk and low-risk, respectively. Kaplan-Meier survival analysis and receiver operating characteristic curve (ROC curve) were used to validate the predictive performance of the model. Univariate and multivariate Cox regression analyses were used to evaluate the prognostic value of clinical indicators such as risk scores. A nomogram for predicting the overall survival rate of gastric cancer patients was constructed based on risk score, age and tumor TNM stage, and the ROC curve and calibration curve were used to evaluate the prediction accuracy of the nomogram. Immune cell infiltration analysis, gene mutation analysis and immunotherapy efficacy prediction were conducted with the help of risk scores. IMvigor 210 dataset was used to validate the performance of risk scoring model on prediction for the response to immunotherapy of other tumor patients. Gastric cancer tissue samples and paired adjacent tissue samples of 15 gastric cancer patients from the Seventh Affiliated Hospital of Sun Yat-sen University between August 2023 and April 2024 were collected, and the differential expression of CD8⁺ T cell immune-related genes in gastric cancer tissues were detected by real-time quantitative polymerase chain reaction. **Result** Totally, 10 CD8⁺ T cell immune-related genes with prognostic value were identified. Risk score was an independent prognostic factor for gastric cancer patients ($HR=3.495$, 95% CI 2.072–5.894, $P<0.001$). The Kaplan-Meier survival curve showed that the overall survival rate of high-risk patients was lower than that of low-risk patients ($P<0.05$). The 1-year, 3-year, and 5-year overall survival rates of gastric cancer patients predicted by nomogram were close to the actual overall survival rates. The high-risk patients had a poor prognosis with immunosuppression, low-frequency gene mutations, and low expression of immune checkpoint molecules. The risk scoring model demonstrated good predictive performance in the IMvigor210 dataset. The expression levels of CD8⁺ T cell immune-related genes were different between gastric cancer tissue samples and adjacent tissue samples ($P<0.05$). **Conclusion** The risk scoring model constructed in this study is of great significance for tumor microenvironment, prognosis and tumor mutational burden of gastric cancer, and lays a foundation for improving the prognosis and treatment responsiveness of gastric cancer patients.

【Key words】 CD8⁺ T cell; Gastric cancer; Immune signature; Immune infiltration; Immunotherapy

胃癌是世界上导致癌症相关死亡的主要原因之一^[1]。尽管手术联合化疗已被广泛应用于胃癌的治疗,但胃癌的5年生存率仍然很低^[2]。文献报道,非转移性胃癌患者的5年生存率可达60%,而晚期转移性胃癌患者的5年生存率不足10%^[3]。因此,尽早识别具有高危转移风险的胃癌患者,并采取更积极的治疗方案具有重要的临床意义。目前化疗是晚期胃癌患者主要的治疗选择,然而对某些患者来说,化疗效果欠佳。晚期胃癌患者也可以选择免疫检查点阻断作为治疗方法,然而,虽然许多其他类型的肿瘤对免疫检查点抑制剂 (immune checkpoint inhibitor, ICI) 有积极反应,但该类药物对胃癌的效果并不十分理想。

ICI在胃癌治疗中的失败可能归因于几个因素,其中一个潜在的因素是胃癌的低肿瘤突变负

荷 (tumor mutation burden, TMB), 导致其免疫原性较低。高TMB通常与免疫治疗的良好反应相关^[4]。此外,相对较低的错配修复缺陷 (deficient mismatch repair, dMMR) 发生率也是一个影响因素,而胃癌的dMMR发生率相对较低(11.68%~33.82%)^[5],这提示胃癌可能难以诱导T细胞介导的免疫反应。CD8⁺ T细胞是一种重要的免疫监视细胞,能够检测肿瘤细胞产生的抗原。肿瘤中CD8⁺ T细胞高度浸润往往预示着良好的预后,并成为免疫检查点阻断治疗的靶点^[6]。

为了预测胃癌患者预后,本研究收集了先前发表过的CD8⁺ T细胞免疫相关基因,并分析了癌症基因组图谱 (the cancer genome atlas, TCGA) 和基因表达综合 (gene expression omnibus, GEO) 数据库中这些基因的表达水平。利用Lasso回归算法进

行模型开发,确定了一个基因模型,该模型具有良好的预测精度和可解释性。本研究还进行了生存分析以及免疫细胞相关性评估^[7],并使用 IMvigor 210 数据集验证了该模型。总之,本研究系统地调查了胃癌患者中的 CD8⁺ T 细胞基因群,为临床预后和治疗应用方面提供基础信息。

1 材料和方法

1.1 数据来源

研究用于生物信息学分析的数据来自 TCGA 和 GEO 数据库下载的所有胃癌相关原始数据。在 TCGA-胃癌 RNA 测序谱图中获取患者的临床病理特征,见表 1。将 TCGA 数据集($n=407$)和 GEO 数据集(GSE84437 系列, $n=433$)分别作为训练集和验证集,2 个数据集合并作为总体样本($n=840$)。

1.2 研究方法

本研究方法的主要设计思路:首先利用生物信息学分析获得具有预后价值的 CD8⁺ T 细胞免疫相关基因,依此构建风险评分模型并进行 TCGA 和 GEO 数据集的验证和分析,包括生存分析、基因突变分析、免疫浸润分析等。此外,使用外部数据集 IMvigor 210 对模型进行验证,并应用荧光定量聚合酶链反应(polymerase chain reaction, PCR)在临床样本中验证构建风险评分模型所选择的 CD8⁺ T 细胞免疫相关基因的表达模式。研究设计流程见图 1。

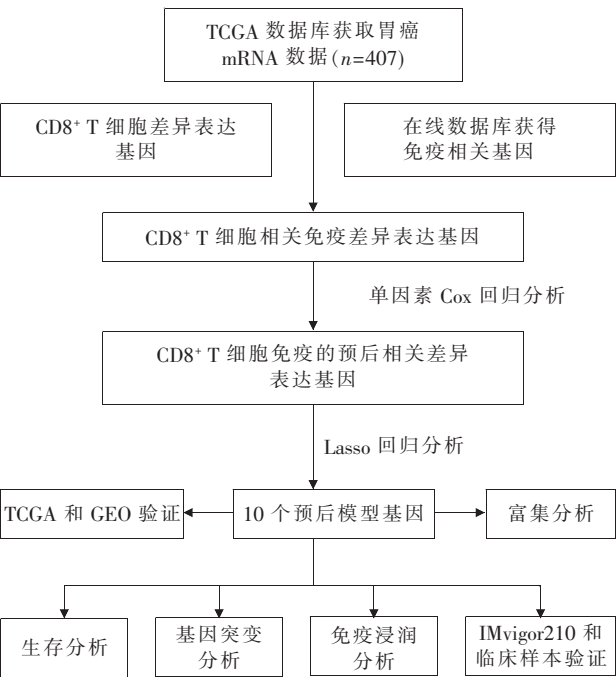


图 1 研究设计流程图

表 1 TCGA-胃癌患者的临床病理特征

临床病理特征	例数	百分比(%)
性别		
男	285	70.0
女	122	30.0
年龄		
≤65 岁	202	49.6
>65 岁	205	50.4
TNM 分期		
I 期	59	14.5
II 期	130	31.9
III 期	183	45.0
IV 期	35	8.6
T 分期		
T ₁ 期	23	5.7
T ₂ 期	93	22.9
T ₃ 期	162	39.8
T ₄ 期	119	29.2
T _x 期	10	2.4
N 分期		
N ₀ 期	96	23.6
N ₁ 期	119	29.2
N ₂ 期	85	20.9
N ₃ 期	88	21.6
N _x 期	19	4.7
M 分期		
M ₀ 期	355	87.2
M ₁ 期	30	7.4
M _x 期	22	5.4
生存状态		
生存	236	58.0
死亡	171	42.0
生存时间		
≥5 年	356	87.5
<5 年	51	12.5

1.2.1 胃癌 CD8⁺ T 细胞相关差异表达基因的鉴定

CD8⁺ T 细胞差异表达基因(differentially expressed genes, DEGs)是从肿瘤免疫单细胞中心(tumor immune single-cell hub, TISCH)(<http://tisch.comp-genomics.org/>)的细胞质小 RNA 数据集(GSE134520)中获得的。本研究根据标准(|log₂ 差异倍数|>1 和校正后 $P<0.05$),从 TISCH 中筛选出与 CD8⁺ T 细胞相关的基因。免疫相关基因来自 ImmPort 数据库(<https://www.immport.org/home>)和 InnateDB 数据库(<https://www.innatedb.ca/>)。将这些 CD8⁺ T 细胞相关基因与免疫相关基因进行比对,最终确定了 174 个与

CD8⁺ T 细胞免疫相关的基因。

1.2.2 CD8⁺ T 细胞免疫相关基因风险评分模型的构建与验证

基于单因素 Cox 回归分析对上述 174 个基因进行筛选,以 $P<0.05$ 作为筛选标准识别与 CD8⁺ T 细胞免疫相关的预后基因,随后采用 Lasso 回归分析进一步压缩风险模型,重复 1000 次 10 倍交叉验证,以 $P<0.05$ 作为筛选标准来最小化基因数量,并据此构建 CD8⁺ T 细胞免疫相关基因的风险评分模型。风险评分计算公式为:风险评分=Σ 基因的风险系数×基因表达值。根据训练集的中位风险评分 3.869 753 894 485 5,将训练集、验证集和总体样本分别划分为高风险组和低风险组:训练集低风险组 203 例,高风险组 204 例;验证集低风险组 206 例,高风险组 227 例;总体样本低风险组 409 例,高风险组 431 例。通过运用风险评分曲线、散点图、热图实现对不同风险患者生存状态及风险评分的刻画。采用 Kaplan-Meier 生存分析预测高、低风险患者的总生存(overall survival, OS)率。采用“survival ROC”R 包计算受试者操作特征曲线(receiver operating characteristic curve, ROC 曲线)的曲线下面积(area under the curve, AUC),确定风险评分模型预测 1 年、3 年和 5 年生存率的特异度和敏感度。

1.2.3 列线图的构建与验证

对风险评分和 TCGA-胃癌患者的临床病理特征行单因素和多因素 Cox 回归分析,探讨 CD8⁺ T 细胞免疫相关基因风险评分模型对胃癌患者预后的独立影响价值^[8]。各自变量赋值如下:风险评分(低风险=0,高风险=1),年龄(连续变量,赋值从小到大),性别(女=0,男=1),肿瘤分级(G_1 级=0, G_2 级=1, G_3 级=2),肿瘤 TNM 分期(I 期=0, II 期=1, III 期=2, IV 期=3),T 分期(T_1 期=0, T_2 期=1, T_3 期=2, T_4 期=3, T_x 期=4),N 分期(N_0 期=0, N_1 期=1, N_2 期=2, N_3 期=3, N_x 期=4),M 分期(M_0 期=0, M_1 期=1, M_x 期=2)。以风险评分和临床病理特征构建列线图,预测 TCGA-胃癌患者的 1 年、3 年和 5 年 OS 率,并使用“rms”R 包创建了校准曲线,评估预测结果和实际 OS 率之间的一致性。通过决策曲线分析(decision curve analysis, DCA)评估模型的可靠性^[9]。

1.2.4 通路富集分析

为了检查训练集高风险组和低风险组之间高

度相关的 HALLMARK 通路的富集程度,进行了基因集富集分析(gene set enrichment analysis, GSEA)。基因集从 GSEA 数据库(<https://www.gsea-msigdb.org/gsea>)检索获得。GSEA 被应用于研究高风险组和低风险组中调节肿瘤发生通路的潜在参与。该方法基于风险评分对高风险组和低风险组进行生物学功能差异的分析。首先,使用以下标准筛选出高风险组和低风险组之间的差异基因:log 差异倍数 ≥ 0.5 和错误发现率 <0.05 。其次,利用整个 TCGA 数据集的所有基因信息,在高风险组和低风险组中探索京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)通路。

1.2.5 基于风险评分的免疫治疗反应评价

从基因表达谱中确定细胞组成的有效方法是 CIBERSORT^[10],研究采用了 CIBERSORT 技术和 LM22 基因标记,在 TCGA 数据库测定了胃癌中 22 种不同类型肿瘤浸润免疫细胞(tumor-infiltrating immune cells, TIICs)的比例。对训练集高、低风险组的免疫细胞浸润到肿瘤微环境中的比例进行比较,并使用基于免疫细胞与基质细胞比例的 Estimate 算法计算了免疫评分、基质评分、Estimate 评分^[11]。

另外,由于 TMB 可以成为预测 ICI 疗效的生物标志物^[12-13],因此本研究进行了基因突变分析,将训练集的高、低风险数据与 TCGA 数据库中胃癌染色体的突变数据进行合并分析,利用 Perl 语言得到高风险亚组($n=177$)和低风险亚组($n=184$)的肿瘤突变数据,再使用 R 语言的“maftools”包获取突变频率最高的前 20 位基因进行展示,并分别绘制瀑布图。

为了对免疫治疗有效的患者进行分子亚型鉴别,本研究还分析了训练集高、低风险患者 TMB 的差异,并探讨了风险评分与 TMB、RNA 干性指数、微卫星不稳定(microsatellite instability, MSI)之间的关系。其中,RNA 干性指数是用于评估肿瘤细胞转移、复发和耐药性的一种指标,研究使用“TCGAbiolinks”R 包(4.1.1 版本),通过单类逻辑回归机器学习方法计算训练集每例患者的 RNA 干性指数。MSI 状态评估表则从 TCGA 数据库获取。研究还使用癌症免疫组图谱(the cancer immunome atlas, TCIA)(<https://tcia.at/home>)进行免疫表型评分,用于预测高、低风险患者对 ICI 的治疗反应。

为了证实 CD8⁺ T 细胞免疫相关基因风险评分模型同样适用于预测其他类型肿瘤患者对免疫治

疗的反应,本研究采用了 IMvigor210 队列数据(http://research-pub.gene.com/IM_vigor210CoreBiologies)进行验证。该队列包含 348 例接受了抗程序性死亡受体配体 1(programmed death-ligand 1, PD-L1)治疗的转移性尿路上皮癌患者数据,采用训练集的风险评分公式计算了 IMvigor210 队列中每例患者的风险评分,并根据评分中位数(0.278 777)将 IMvigor 210 队列患者划分为高风险($n=100$)和低风险($n=248$)。然后利用“survival”和“ggplot2”包绘制生存曲线和 ROC 曲线,分析两组患者的生存差异以及该风险评分模型预测 IMvigor210 队列患者对免疫治疗反应的能力。

1.2.6 荧光定量聚合酶链反应验证

人正常胃上皮细胞 GES-1 和胃癌细胞系细胞 AGS、MGC-803、HGC-27、SNU-719、MKN-45 购自中科院上海细胞库。将细胞置于含有 10%胎牛血清和 1%青/链霉素的 DMEM 培养基中,于 37℃、含体积分数 0.05 CO₂ 的条件下进行培养,用于后续分析。

收集中山大学附属第七医院 2023 年 8 月至 2024 年 4 月收治的 15 例胃癌患者的胃癌组织样本和匹配的癌旁组织样本,存放在-80℃冷冻。其中男性 9 例,女性 6 例,年龄范围 30~80 岁,平均年龄(59.8±16.0)岁。纳入标准:①内镜或术后病理结果明确为胃恶性肿瘤;②术前未接受过任何放疗、化疗;③自愿签署生物样本捐献同意书。排除标准:①不能明确肿瘤分化程度、浸润深度、病理类型等;②既往存在其他器官肿瘤病史。研究经过中山大学附属第七医院伦理审查委员会批准(伦理批准号:KY-2025-061-01)。

使用 TRIzol 试剂从上述组织和细胞中提取总 RNA,使用反转录酶将 RNA 转化为 cDNA。使用 SYBR Green I 试剂进行荧光定量 PCR。引物由中国华大基因公司合成,采用 GAPDH 为参照,所有引物序列见表 2。反应体系 10 μl,SYBR 5 μl,前后引物各 0.5 μl,cDNA 1 μl,焦碳酸二乙酯处理水 3 μl。利用荧光定量 PCR 设备(CFX96 Touch, Bio-Rad)检测组织和细胞中 CD8⁺ T 细胞免疫相关基因的表达水平。反应条件:预变性 95℃ 30 s,变性 95℃ 5 s,退火 60℃ 30 s,共 40 个循环。计算 2^{-ΔΔCt} 值为各基因的相对表达量。

1.3 统计学方法

使用 R 软件(4.3.1 版本)进行统计分析。生存分析采用 Kaplan-Meier 法以及 Log-rank 检验。胃癌细胞系和临床组织样本中的基因表达差异用配对 *t* 检验。组间的肿瘤微环境评分差异、免疫细胞浸润水平差异、TMB 差异、风险评分差异、免疫表型评分差异等采用单因素方差分析。组间不同 MSI 状态患者比例的差异采用卡方检验。应用 Cox 比例风险回归模型计算 HR 和 95%CI。采用 R 包“survival”进行单因素和多因素 Cox 回归分析。用 Spearman 相关性分析评估风险评分与 TMB、肿瘤 RNA 干性指数的相关性。双侧 *P*<0.05 为差异有统计学意义。

2 结果

2.1 CD8⁺ T 细胞免疫相关基因风险评分模型的构建与验证

从在线数据库获取的免疫相关基因共 2533 个,在 TISCH 中筛选出与 CD8⁺ T 细胞相关的

表 2 引物序列

基因	正向	反向
CXCR4	ACTACACCGAGGAAATGGGCT	CCCACAATGCCAGTTAAGAAGA
VIM	GACGCCATCAACACCGAGTT	CTTTGTCGTTGGTTAGCTGCT
TMSB4X	GGGCAATGGATTGGTCATCCT	TGCAGCCTGTACTTGTCCG
CMTM3	CGAGTCGGGTCTCTCATTCAT	CCCTGCCACTTGTCAATTCAG
ATF3	CCTCTGCGCTGGAATCAGTC	TTCTTTTCTCGTCCCTCTTTTT
CXCL3	CGCCCAAACCGAAGTCATAG	GCTCCCTTGTTCAGTATCTTTT
LSGAL9	CCAAACCGAAGTCATAGCCAC	TGCTCCCTTGTTCAGTATCT
ADM	ATGAAGCTGGTTTCCGTCG	GACATCCGCAGTTCCCTCTT
SOCS3	CCTGCGCCTCAAGACCTTC	GTCACTGCGCTCCAGTAGAA
RGS2	AAGATTGGAAGACCCGTTTGA	GCAAGACCATAATTGCTGGCT
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG

DEGs 共 896 个,将两组基因进行相交分析后得到 174 个重叠的基因(图 2A)。接着,通过单因素 Cox 回归分析筛选出 15 个与 CD8⁺ T 细胞免疫相关的预后基因(图 2B),再运用 Lasso 回归分析对这 15 个候选基因进行筛选(图 2C 和 2D),最终确定了 10 个与 CD8⁺ T 细胞免疫相关的关键基因,它们分别是 *CXCR4*、*VIM*、*TMSB4X*、*CMTM3*、*ATF3*、*CXCL3*、*LGALS9*、*ADM*、*SOCS3* 和 *RGS2*(表 3)。最后根据风险评分公式计算出风险评分: 风险评分= $0.123\times CXCR4+0.057\times VIM+0.182\times TMSB4X+0.014\times CMTM3+0.104\times ATF3-0.086\times CXCL3-0.176\times LGALS9+0.079\times ADM+0.020\times SOCS3+0.052\times RGS2$ 。

表 3 风险评分模型中 10 个 CD8⁺ T 细胞免疫相关基因的风险系数

CD8 ⁺ T 细胞免疫相关基因	风险系数
<i>CXCR4</i>	0.123
<i>VIM</i>	0.057
<i>TMSB4X</i>	0.182
<i>CMTM3</i>	0.014
<i>ATF3</i>	0.104
<i>CXCL3</i>	-0.086
<i>LGALS9</i>	-0.176
<i>ADM</i>	0.079
<i>SOCS3</i>	0.020
<i>RGS2</i>	0.052

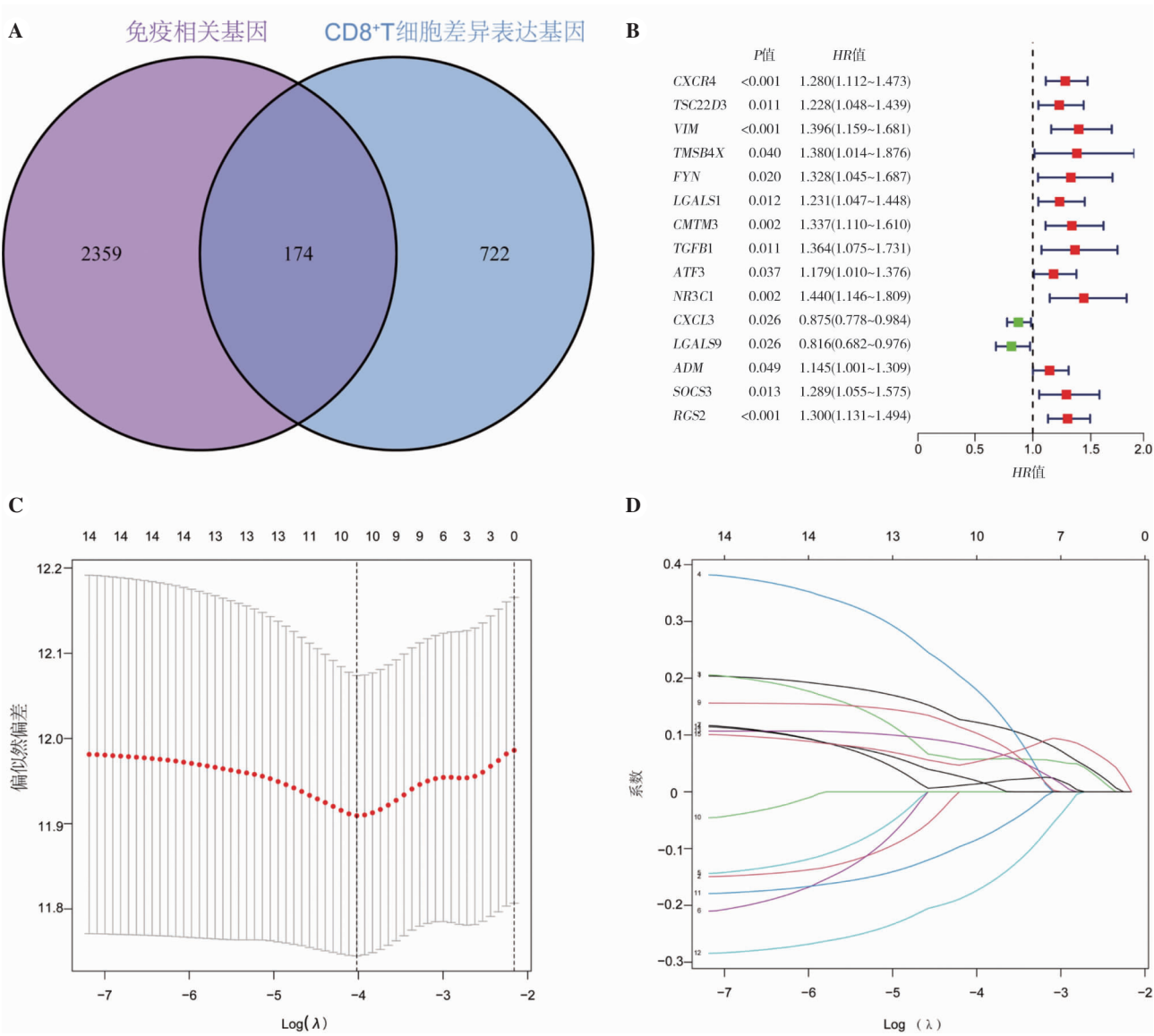


图 2 鉴定并筛选出 10 个 CD8⁺T 细胞免疫相关基因

注:A,CD8⁺ T 细胞差异表达基因与免疫相关基因的韦恩图;B,15 个 CD8⁺ T 细胞免疫相关差异表达基因的森林图;C,根据 Log(λ)绘制偏似然偏差图;D,由 Lasso 回归分析确定的基因系数曲线。

训练集的中位风险评分为 3.869 753 894 485 5, 训练集、验证集和总体样本的风险评分曲线见图 3A。散点图结果显示高风险患者的死亡率更高(图 3B), 热图则展示了这 10 个 CD8⁺ T 细胞免疫相关基因在不同风险组中的表达情况(图 3C)。与单因素 Cox 回归分析结果一致, *CXCL3* 和 *LGALS9* 在低风险组中表达较高, 而其他基因在高风险组中表达较高。此外, Kaplan-Meier 生存分析结果显示, 在训练集、验证集和总体样本中, 高风险患者的 OS 率较低(图 3D)。采用 ROC 曲线检验风险评分模型对预后预测的准确性, 在训练集、验证集和总体样本中 5 年生存率的 AUC 分别为 0.705、0.564 和 0.605(图 3E), 提示 CD8⁺ T 细胞免疫相关基因风险评分模型是一种有用的胃癌预后预测模型。

2.2 列线图的构建与验证

单因素 Cox 回归分析结果显示, 年龄、TNM 分期、肿瘤分级和风险评分是胃癌患者预后的影响因素($P<0.05$); 多因素 Cox 回归分析结果显示, 风险评分是胃癌患者预后的独立影响因素($HR=3.495, 95\%CI\ 2.072\sim5.894, P<0.001$), 见图 4A 和 4B。

基于风险评分和临床病理特征构建了列线图, 并通过计算风险评分来预测胃癌患者的 1 年、3 年和 5 年 OS 率(图 4C)。校准曲线显示, 1 年、3 年和 5 年的预测 OS 率与实际 OS 率高度一致(图 4D)。此外, DCA 提示 CD8⁺ T 细胞免疫相关基因风险评分在预测 1 年、3 年和 5 年 OS 率时比 TNM 分期、年龄和性别等更有优势(图 4E)。

2.3 通路富集分析

GSEA 结果显示, 高风险组显著富集的 KEGG 通路为扩张型心肌病、细胞外基质受体相互作用、黏着斑、肥厚型心肌病以及神经活性配体受体相互作用; 而低风险组则主要富集在碱基切除修复、DNA 复制、过氧化物酶体、蛋白酶体以及核糖体方面(图 5A 和 5B), 提示肿瘤微环境对胃癌的发展至关重要。

2.4 高、低风险组的免疫浸润分析

通过 Estimate 算法计算训练集的免疫评分、基质评分以及 Estimate 评分, 结果发现高风险组得分更高(图 6A)。采用 CIBERSORT 对高、低风险组的免疫微环境进行了详细分析, 并通过柱状图显示 22 种免疫细胞的浸润水平, 结果显示, 高风险组的 M2 巨噬细胞和嗜酸性粒细胞比例较高,

而低风险组的浆细胞比例较高(图 6B 和 6C)。

2.5 基于风险评分的免疫治疗反应评价

基因突变分析显示, 不同风险亚组存在不同的免疫特征。如图 7 所示, 错义突变是最常见的突变形式。*TTN*、*TP53* 和 *MUC16* 基因的突变率在两组中均超过 25%, 且为最常见的类型。而与低风险亚组相比, 高风险亚组具有较低频率的基因突变。

此外, 本研究还发现, 低风险患者 TMB 较高($P<0.01$, 图 8A), 并且风险评分与 TMB($r=-0.29, P<0.01$, 图 8B)、RNA 干性指数($r=-0.57, P<0.01$, 图 8C)呈负相关。不同风险评分的微卫星稳定(microsatellite stability, MSS)、高微卫星不稳定(microsatellite instability-high, MSI-H)、低微卫星不稳定(microsatellite instability-low, MSI-L)患者比例的差异有统计学意义, 且 MSS 和 MSI-L 患者的风险评分高于 MSI-H 患者(图 8D 和 8E)。针对 ICI 治疗效果方面, 本研究观察到低风险患者表现出更好的治疗效果, 说明这一群体具有更高的免疫原性(图 8F~8I)。

在 IMvigor210 队列中验证该风险评分模型时, 发现完全缓解和部分缓解患者的风险评分较低, 而疾病稳定和疾病进展的患者则相反(图 8J), 低风险患者的生存时间也长于高风险患者(图 8K)。ROC 曲线结果显示该风险评分模型能够有效预测 IMvigor210 队列患者对免疫治疗的反应($AUC=0.588$, 图 8L)。

2.6 胃癌细胞系和临床组织样本中差异表达基因的验证

荧光定量 PCR 结果显示, *CXCL3* 和 *LGALS9* 在胃癌组织中的表达量较癌旁组织低(均 $P<0.05$), 而 *CXCR4*、*CMTM3*、*ATF3*、*ADM*、*SOC3* 和 *RGS2* 在胃癌组织中的表达量较癌旁组织高(均 $P<0.05$); 这些基因在部分胃癌细胞系中的表达也与人正常胃上皮细胞 GES-1 存在差异, 见图 9。

3 讨论

胃癌具有多种致癌机制, 且在不同性别、种族和地理位置之间存在显著异质性^[14]。大多数胃癌患者确诊时已为晚期, 手术结果不理想^[15]。尽管有放疗、化疗、基因治疗和靶向治疗等替代疗法, 但 5 年生存率仍低于 30%^[16]。新兴的治疗方法, 如癌症疫苗、免疫调节剂、靶向抗体和过继细胞治疗已经显示出希望, 但胃癌患者对免疫治疗的反应率

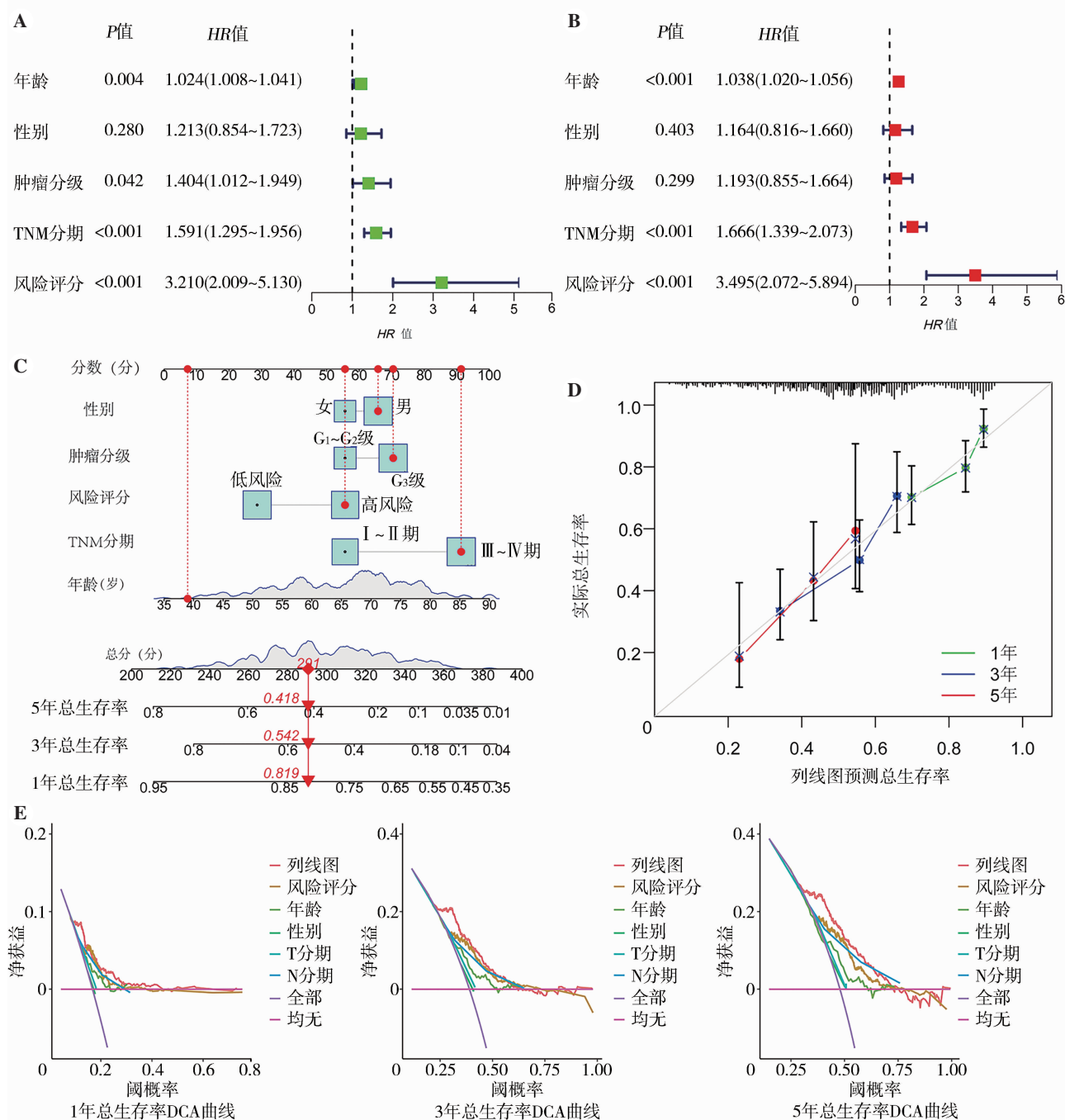


图4 预测胃癌患者1年、3年和5年总生存率的列线图构建与验证

注:A,胃癌患者预后影响因素的单因素Cox回归分析;B,胃癌患者预后影响因素的多因素Cox回归分析;C,基于年龄、性别、肿瘤分级、TNM分期、风险评分构建的列线图;D,列线图预测胃癌患者1年、3年和5年总生存率的校准曲线;E,列线图预测胃癌患者1年、3年和5年总生存率的DCA曲线。

前胃癌模型的局限性和不足,并强调了改进转化研究的必要性。为了解决这个问题,本研究聚焦在将CD8⁺T细胞免疫相关基因表达变化作为胃癌潜在预后指标。为了探索与CD8⁺T细胞免疫相关的潜在候选基因,本研究进行了免疫相关基因和CD8⁺T细胞特异性DEGs之间的重叠分析,并利用单因素Cox回归分析和Lasso回归分析将候选基因缩

小到10个。这10个基因的鉴定为胃癌预后模型的建立提供了潜在的基础。应用这些基因作为生物标志物,有可能更准确地预测胃癌患者的预后。本研究还观察到CD8⁺T细胞免疫相关基因在胃癌组织样本中的表达与癌旁组织样本存在差异。通过单因素和多因素Cox回归分析,确定高风险评分是胃癌患者预后的一个独立危险因素。总之,

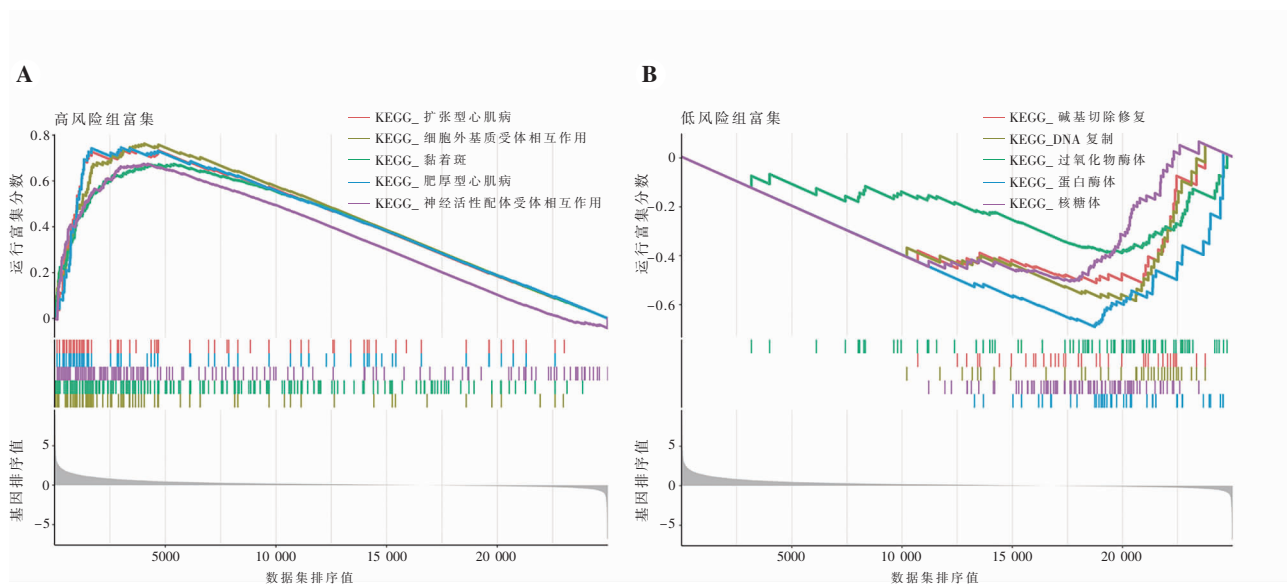


图5 高、低风险组的基因集富集分析

注:A,高风险组中富集的 HALLMARK 通路;B,低风险组中富集的 HALLMARK 通路。

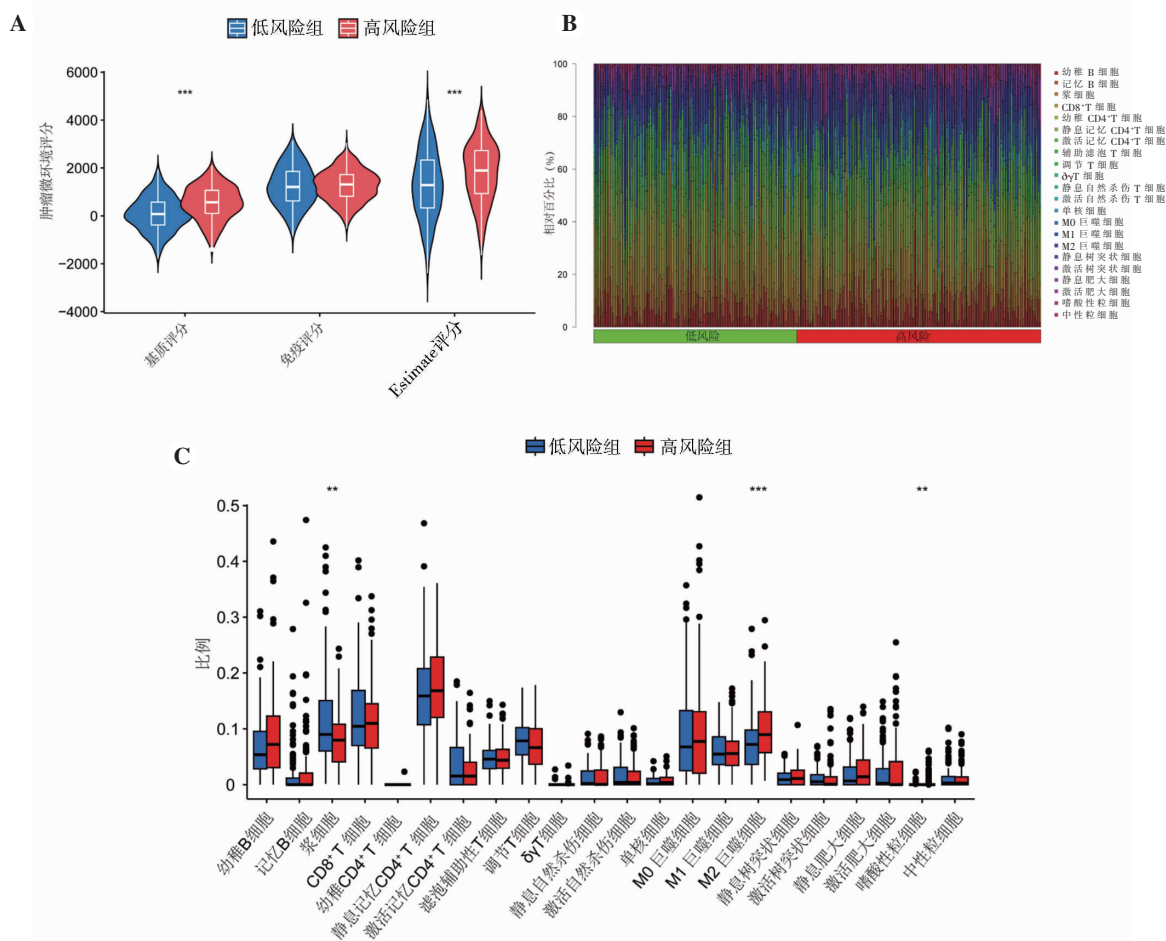


图6 高、低风险组的免疫细胞浸润分析

注:A,高、低风险组肿瘤微环境的差异分析;B,高、低风险组的免疫细胞组成;C,高、低风险组中22种免疫细胞的免疫浸润水平。 $^{**}P<0.01$, $^{***}P<0.001$ 。

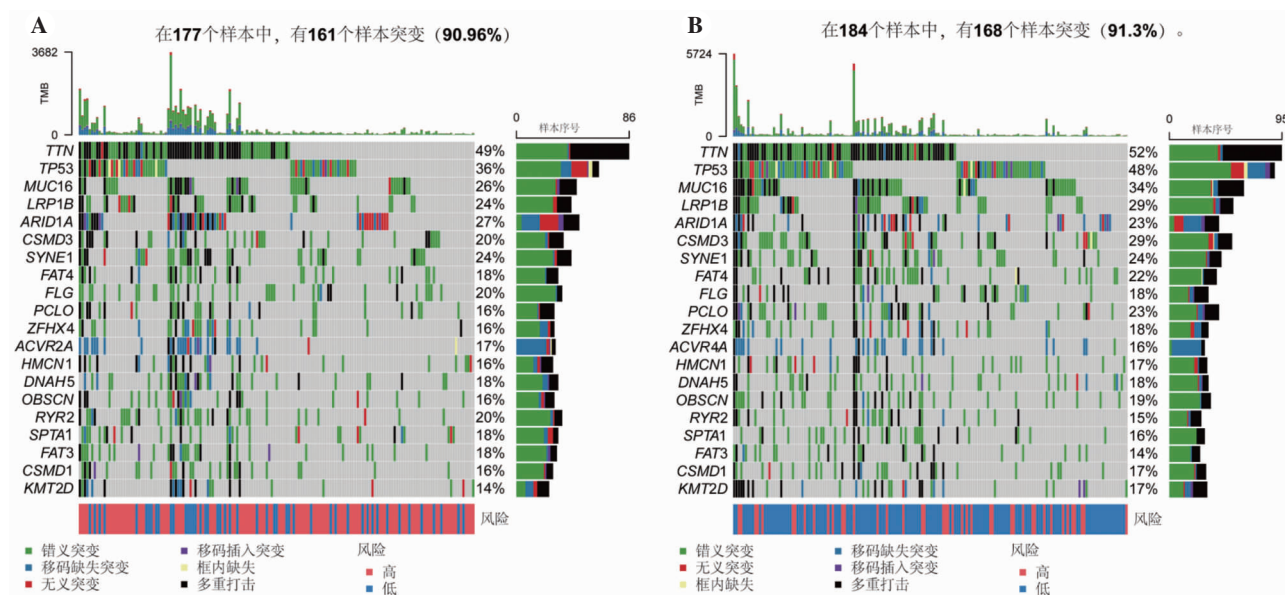


图7 风险评分与基因突变的关系

注:A,高风险亚组中显著突变的基因;B,低风险亚组中显著突变的基因。

本研究强调了 CD8⁺ T 细胞浸润及其免疫相关基因在胃癌预后中的意义。通过阐明这些关系,为胃癌患者改善预后和靶向治疗的发展做出贡献。

本研究结果显示,胃癌的预后与几个因素相关,包括年龄、风险评分和 TNM 分期,这与既往的研究结果相一致^[19-20]。因此,本研究将风险评分与年龄、TNM 分期相结合,形成预测胃癌患者 OS 率的列线图,以提高预后预测的准确性。结果表明,列线图具有很好的 OS 率预测精度,并且与理想模型非常接近。

高风险和低风险患者的预后差异显著可能有多种原因解释。在对各类肿瘤进行大量研究后发现,TMB 高的患者通常有较高的生存率。同样地,本研究也发现了类似趋势:TMB 越高,则风险评分越低。另外,*MUC16* 被发现是胃癌最常见的突变基因之一,并且其突变与更高的 TMB、更强的抗肿瘤免疫反应以及更好的临床预后相关^[21]。而 *TTN* 突变则与肿瘤的免疫检查点阻断相关^[22]。超过50%的人类肿瘤组织都带有 *TP53* 突变^[23],*TP53* 是最常见的肿瘤抑制基因突变之一。*TP53* 基因编码的蛋白在调节细胞周期、DNA 损伤与修复、细胞衰老与凋亡、调节自噬等方面发挥着重要作用^[24]。因此,TMB 可能有助于解释高风险和低风险患者预后的显著差异。分子机制也可能是解释这种现象的另一个原因。高风险组显著富集的 KEGG 通路为扩张型心肌病、细胞外基质受体相互作用、黏着

斑、肥厚型心肌病以及神经活性配体受体相互作用;而低风险组则主要富集在碱基切除修复、DNA 复制、过氧化物酶体、蛋白酶体以及核糖体方面。综上所述,不同的分子通路可能是高、低风险患者预后差异的潜在机制。

此外,免疫细胞浸润的因素也不容忽视。高风险患者 M2 巨噬细胞和嗜酸性粒细胞比例较高,而低风险患者则浆细胞比例较高。研究表明,M2 巨噬细胞和嗜酸性粒细胞可能促进恶性肿瘤免疫逃逸^[25-26]。相反,浆细胞具有强大的抗肿瘤免疫反应^[27]。而患者的免疫抑制状态会导致高风险患者预后较差。

近年来,免疫治疗已成为一种有前途的策略,可用于大量胃癌患者^[28]。然而,并非每例患者对免疫治疗都有良好的反应。研究表明,在肿瘤组织中适当表达免疫检查点对免疫治疗的有效性至关重要^[29-30]。MSI 和 TMB 被认为是预测胃癌抗程序性死亡受体 1(programmed death-1, PD-1)/PD-L1 治疗效果的有效生物标志物^[31]。MSI-H 是一种高突变负荷,占胃癌患者的 22%^[32]。研究发现,无论 PD-L1 状态如何,使用帕博利珠单抗治疗 MSI-H 患者均可获得更好的疗效^[33]。在本研究的低风险组中,MSI-H 患者的比例较高,且 MSI-H 患者的风险评分明显低于 MSS 和 MSI-L 患者,这可能是低风险胃癌患者预后较好的原因之一。TMB 作为一种新的生物标志物,被用来预测免疫治疗反应^[34]。

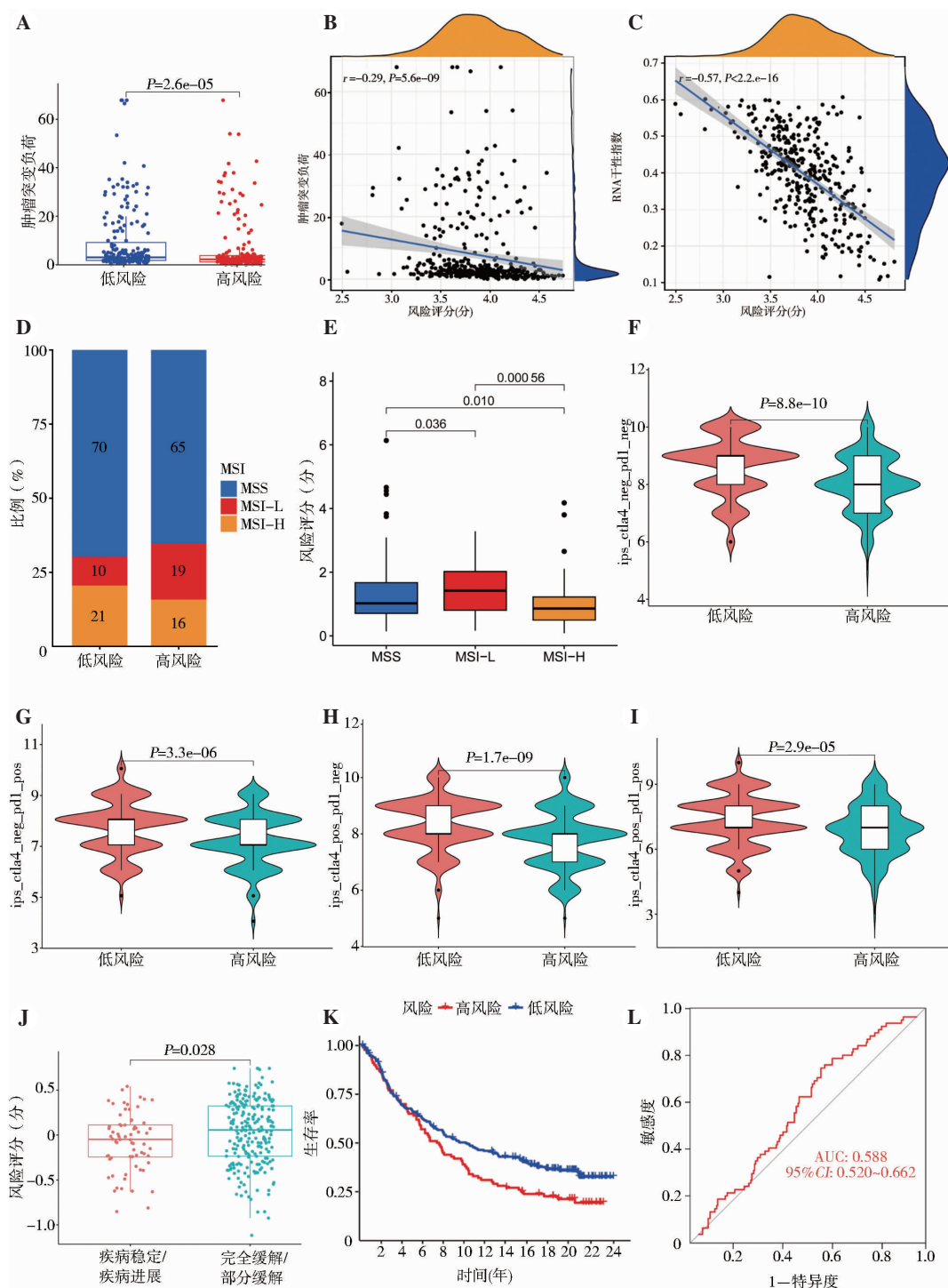


图8 基于风险评分的免疫治疗反应评价

注:A、高、低风险患者的肿瘤突变负荷差异比较;B、风险评分与肿瘤突变负荷的相关性散点图;C、风险评分与RNA干性指数的相关性散点图;D、高、低风险组中各类微卫星不稳定患者所占的比例;E、各类微卫星不稳定患者的风险评分差异比较;F~I、免疫表型评分在高、低风险患者中的分布情况,ips_ctla4_neg_pdl_neg(F)表示对细胞毒性T淋巴细胞相关蛋白4抑制剂和程序性死亡受体1抑制剂均无应答,ips_ctla4_neg_pdl_pos(G)表示对细胞毒性T淋巴细胞相关蛋白4抑制剂无应答而对程序性死亡受体1抑制剂应答,ips_ctla4_pos_pdl_neg(H)表示对细胞毒性T淋巴细胞相关蛋白4抑制剂应答而对程序性死亡受体1抑制剂无应答,ips_ctla4_pos_pdl_pos(I)表示对细胞毒性T淋巴细胞相关蛋白4抑制剂和程序性死亡受体1抑制剂均应答;J、IMvigor210队列中不同临床反应患者的风险评分差异比较;K、IMvigor210队列中高、低风险患者的Kaplan-Meier生存曲线;L、风险评分模型预测IMvigor210队列患者免疫治疗反应的ROC曲线。MSI,微卫星不稳定;MSS,微卫星稳定;MSI-H,高微卫星不稳定;MSI-L,低微卫星不稳定。

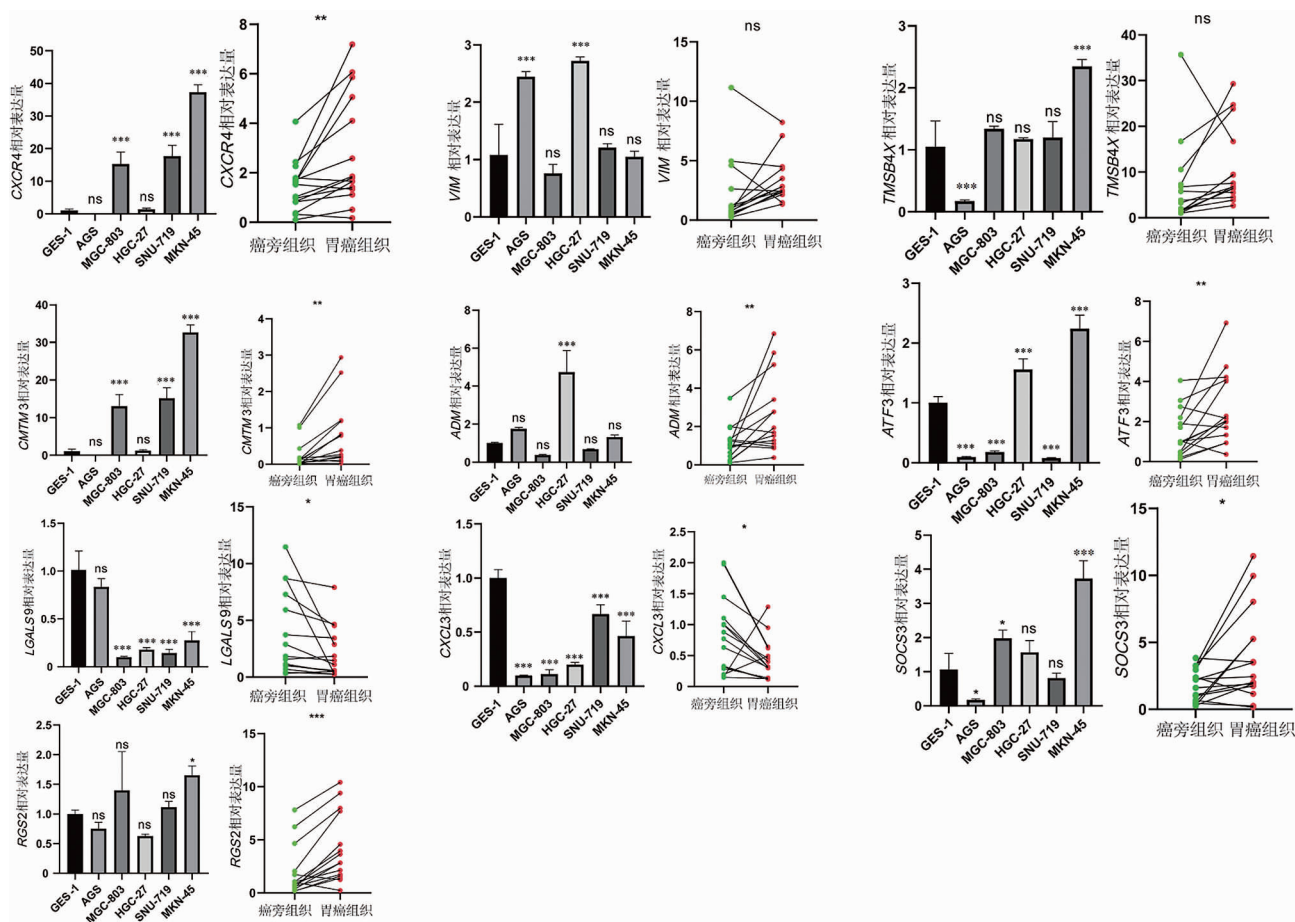


图9 分析胃癌细胞系和新鲜胃癌组织样本中10个CD8⁺T细胞免疫相关基因的表达情况

注:ns,差异无统计学意义;* $P<0.05$;** $P<0.01$;*** $P<0.001$ 。

本研究结果显示,低风险患者的TMB高于高风险患者,这提示了低风险患者更有可能从ICI治疗中获益。此外,本研究采用免疫表型评分检测抗细胞毒性T淋巴细胞相关蛋白4和抗PD-1方案的效果,结果发现在各治疗方案中,低风险患者均表现出优于高风险患者的效果。

然而,本研究的模型存在一些局限性。由于TCGA数据库主要包含高加索和非洲人口的信息,因此本研究缺乏来自亚洲人口的数据。虽然研究通过荧光定量PCR验证了候选基因在胃癌组织中的差异表达,但CD8⁺T细胞具体影响免疫治疗的机制尚不清楚,需要进行体内和体外研究。综上所述,本研究构建的风险评分模型是一种可靠的胃癌治疗敏感性和预后预测生物标志物,可为下一步的基础研究和临床实践提供新的视角。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 孙学增,负责实验设计、实验操作、数据收

集和初步分析,并撰写论文的大部分内容;苏日顺、李文超、王小群,参与实验设计和数据分析,并撰写论文部分章节;尹松成、陈景耀,负责数据的最终分析和论文的校对工作。

参考文献

- [1] SUNG H, FERLAY J, SIEGEL RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA Cancer J Clin, 2021, 71(3): 209-249.
- [2] SMYTH EC, NILSSON M, GRABSCHE HI, et al. Gastric cancer [J]. Lancet, 2020, 396(10251):635-648.
- [3] DARANG E, PEZESHKIAN Z, MIRHOSEINI SZ, et al. Bioinformatics and pathway enrichment analysis identified hub genes and potential biomarker for gastric cancer prognosis [J]. Front Oncol, 2023, 13: 1187521.
- [4] SAMSTEIN RM, LEE C, SHOUSHARI AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types [J]. Nat Genet, 2019, 51(2):202-206.

- [5] SIDAWAY P. Neoadjuvant pembrolizumab shows promise in MSI-H/dMMR solid tumours[J]. *Nat Rev Clin Oncol*, 2023, 20(3):138.
- [6] HUANG H, HUANG Z, GE J, et al. CD226 identifies functional CD8(+) T cells in the tumor microenvironment and predicts a better outcome for human gastric cancer [J]. *Front Immunol*, 2023, 14: 1150803.
- [7] TORRES R, JUDSON-TORRES RL. Research Techniques Made Simple: Feature Selection for Biomarker Discovery [J]. *J Invest Dermatol*, 2019, 139(10): 2068–2074.
- [8] IASONOS A, SCHRAG D, RAJ GV, et al. How to build and interpret a nomogram for cancer prognosis[J]. *J Clin Oncol*, 2008, 26(8):1364–1370.
- [9] VAN CALSTER B, WYNANTS L, VERBEEK JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators [J]. *Eur Urol*, 2018, 74(6): 796–804.
- [10] NEWMAN AM, LIU C, GREEN MR, et al. Robust enumeration of cell subsets from tissue expression profiles[J]. *Nat Methods*, 2015, 12(5): 453–457.
- [11] THORSSON V, GIBBS DL, BROWN SD, et al. The Immune Landscape of Cancer [J]. *Immunity*, 2018, 48(4): 812–830.
- [12] VALERO C, LEE M, HOEN D, et al. The association between tumor mutational burden and prognosis is dependent on treatment context [J]. *Nature Genetics*, 2021, 53(1):11–15.
- [13] MCGRAIL DJ, PILIE PG, RASHID NU, et al. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types[J]. *Ann Oncol*, 2021, 32(5): 661–672.
- [14] AKHAVAN -NIAKI H, SAMADANI AA. Molecular Insight in Gastric Cancer Induction: An Overview of Cancer Stemness Genes[J]. *Cell Biochem Biophys*, 2014, 68(3):463–473.
- [15] MARANO L, POLOM K, PATRITI A, et al. Surgical management of advanced gastric cancer: An evolving issue[J]. *Eur J Surg Oncol*, 2016, 42(1):18–27.
- [16] ZHU Z, GONG Y, XU H. Neoadjuvant therapy strategies for advanced gastric cancer: Current innovations and future challenges [J]. *Chronic Dis Transl Med*, 2020, 6(3): 147–157.
- [17] LONG J, LIN J, WANG A, et al. PD-1/PD-L blockade in gastrointestinal cancers: lessons learned and the road toward precision immunotherapy [J]. *J Hematol Oncol*, 2017, 10(1): 146.
- [18] 朱正纲. 围手术期联合免疫检查点抑制剂治疗局部进展期胃癌或胃食管结合部腺癌的临床意义[J]. *中华胃肠外科杂志*, 2024, 27(7): 659–665.
- [19] IN H, SOLSKY I, PALIS B, et al. Validation of the 8th Edition of the AJCC TNM Staging System for Gastric Cancer using the National Cancer Database [J]. *Ann Surg Oncol*, 2017, 24(12): 3683–3691.
- [20] 张捷, 蔡钦波, 陈焕杰, 等. SLC25A13 的表达与胃癌临床病理特征及预后的相关性分析 [J/CD]. *消化肿瘤杂志(电子版)*, 2023, 15(1):22–26.
- [21] LI X, PASCHE B, ZHANGW, et al. Association of MUC16 Mutation With Tumor Mutation Load and Outcomes in Patients With Gastric Cancer [J]. *JAMA Oncol*, 2018, 4(12): 1691–1698.
- [22] YANG Y, ZHANG J, CHEN Y, et al. MUC4, MUC16, and TTN genes mutation correlated with prognosis, and predicted tumor mutation burden and immunotherapy efficacy in gastric cancer and pan-cancer [J]. *Clin Transl Med*, 2020, 10(4):155.
- [23] LEROY B, ANDERSON M, SOUSSIT. TP53 Mutations in Human Cancer: Database Reassessment and Prospects for the Next Decade[J]. *Hum Mutat*, 2014, 35(6): 672–788.
- [24] AJORE R, RAISER D, MCCONKEYM, et al. Deletion of ribosomal protein genes is a common vulnerability in human cancer, especially in concert with TP53 mutations [J]. *EMBO Mol Med*, 2017, 9(4): 498–507.
- [25] LI J, SUN J, ZENG Z, et al. Tumour-associated macrophages in gastric cancer: From function and mechanism to application[J]. *Clin Transl Med*, 2023, 13(8):1386.
- [26] ZHENG P, LI W. Crosstalk Between Mesenchymal Stromal Cells and Tumor-Associated Macrophages in Gastric Cancer[J]. *Front Oncol*, 2020, 10: 571516.
- [27] PATIL NS, NABET BY, MUELLER S, et al. Intratumoral plasma cells predict outcomes to PD-L1 blockade in non-small cell lung cancer [J]. *Cancer Cell*, 2022, 40(3): 289–300.
- [28] 杨欢, 彭建军. 胃癌免疫治疗研究进展[J/CD]. *消化肿瘤杂志(电子版)*, 2022, 14(3):268–278.
- [29] JIA K, CHEN Y, SUN Y, et al. Multiplex immunohistochemistry defines the tumor immune microenvironment and immunotherapeutic outcome in CLDN18.2-positive gastric cancer[J]. *BMC Med*, 2022, 20(1):223.
- [30] YAN Y, ZHAO Y, ZHANG C, et al. Toxicity spectrum of immunotherapy in advanced lung cancer: A safety analysis from clinical trials and a pharmacovigilance

- system[J]. E Clinical Medicine, 2022, 50:101535.
- [31] KAWAZOE A, SHITARA K, BOKU N, et al. Current status of immunotherapy for advanced gastric cancer[J]. Jpn J Clin Oncol, 2021, 51(1): 20–27.
- [32] BASS AJ, THORSSON V, SHMULEVICH I, et al. Comprehensive molecular characterization of gastric adenocarcinoma[J]. Nature, 2014, 513(7517):202–209.
- [33] SHITARA K, OZGUROGLU M, BANG Y, et al. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial[J]. Lancet, 2018, 392(10142):123–133.
- [34] MELENDEZ B, VAN CAMPENHOUT C, RORIVE S, et al. Methods of measurement for tumor mutational burden in tumor tissue [J]. Transl Lung Cancer Res, 2018, 7(6): 661–667.

收稿日期:2024-09-25