

基于转录组分析的食管鳞状细胞癌性别相关预后基因的筛选和验证

张冬云*, 吴红芳, 王 静

南阳医学高等专科学校临床肿瘤病理重点实验室, 河南 南阳

收稿日期: 2025年8月25日; 录用日期: 2025年9月19日; 发布日期: 2025年9月28日

摘 要

目的: 基于转录组分析筛选食管鳞状细胞癌(Esophageal Squamous Cell Carcinoma, ESCC)性别相关预后基因(Gender-Related Prognostic Genes, GRPGs), 开发风险评分模型及预后列线图, 临床样本验证关键基因的生物学特征。方法: GSE53625数据集中, 加权基因共表达网络分析(Weighted Gene Co-expression Network Analysis, WGCNA)筛选性别相关模块基因, 与Kaplan-Meier (K-M)分析得到的预后基因取交集。最小绝对收缩与选择算子(Least Absolute Shrinkage and Selection Operator, LASSO)及多变量Cox回归分析方法识别GRPGs。构建风险评分模型, 并将独立预后因素纳入预后预测列线图。校准曲线、受试者工作特征曲线(Receiver Operating Characteristic, ROC)和决策曲线分析(Decision Curve Analysis, DCA)评估模型临床应用效能。结果: 风险评分将患者分为高风险组和低风险组, 两组总生存期(Overall Survival, OS)有显著差异。涵盖年龄、治疗方式和风险评分的预后列线图显示出良好的校准和鉴别能力。免疫组织化学分析(Immunohistochemistry, IHC)显示女性患者IL5RA表达显著升高, 其表达水平与免疫调控及雌激素代谢相关通路的活性正相关。结论: 本研究鉴定出的GRPGs有望成为新型预后生物标志物, 为ESCC的精准治疗及预后评估提供重要参考依据。

关键词

食管鳞状细胞癌, 列线图, 性别, 精准治疗

Screening and Validation of Gender-Related Prognostic Genes in Esophageal Squamous Cell Carcinoma Based on Transcriptomic Analysis

Dongyun Zhang*, Hongfang Wu, Jing Wang

*通讯作者。

文章引用: 张冬云, 吴红芳, 王静. 基于转录组分析的食管鳞状细胞癌性别相关预后基因的筛选和验证[J]. 临床医学进展, 2025, 15(10): 123-130. DOI: 10.12677/acm.2025.15102735

Abstract

Aims: To identify gender-related prognostic genes (GRPGs) in esophageal squamous cell carcinoma (ESCC) through transcriptome analysis, develop a risk scoring model and prognostic nomogram, and validate the biological characteristics of key genes in clinical samples. **Methods:** Weighted gene co-expression network analysis (WGCNA) was applied to identify gender-associated module genes from the GSE53625 dataset, which intersected with prognostic genes derived from Kaplan-Meier (K-M) survival analysis. Least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression analyses were applied to identify Gender-Related Prognostic Genes (GRPGs). We constructed a risk score model and incorporated independent prognostic factors into a nomogram to predict ESCC patient outcomes. Calibration curves, receiver operating characteristic (ROC) curves, and decision curve analysis (DCA) evaluated the clinical utility of the model. **Results:** The risk model stratified patients into high- and low-risk groups with significantly different overall survival (OS). The prognostic model, which integrates age, treatment modality, and risk score, exhibited excellent calibration and discriminative ability. Immunohistochemistry (IHC) revealed elevated IL5RA expression in female patients, positively correlating with activity in immune regulation and estrogen metabolism pathways. **Conclusion:** The identified GRPGs demonstrate potential as novel prognostic biomarkers, providing a critical foundation for precision therapy and survival assessment in ESCC.

Keywords

Esophageal Squamous Cell Carcinoma, Nomogram, Gender, Precision Therapy

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

食管鳞状细胞癌(Esophageal Squamous Cell Carcinoma, ESCC)是一种极具侵袭性的消化系统恶性肿瘤,患者死亡率较高[1]。目前临床以根治性手术联合辅助放化疗为主,但大多数患者就诊时已是中晚期,五年生存率不足 30% [2]。特异性治疗靶点的缺乏,使食管癌的精准治疗面临重大挑战[3]。多项临床研究证实,相同 TNM 分期的肿瘤患者预后存有差异,女性生存优势较为明显[4] [5]。初步机制研究提示可能与雌激素受体 β (ESR2)的高表达有关[6] [7],具体调控机制尚需进一步阐明。肿瘤患者的性别依赖性预后差异现象,凸显传统 TNM 分期在个体化预后评估中的局限性,亟需发现新型预后生物标志物以优化风险分层模型。

本研究利用转录组学分析,从基因表达综合数据库(Gene Expression Omnibus, GEO)中筛选并鉴定性别相关预后基因(Gender-Related Prognostic Genes, GRPGs),构建基因集以开发风险评分模型和预后预测列线图,采用临床组织样本验证关键基因的表达特征,旨在为 ESCC 的个体化靶向治疗和精准预后评估提供重要理论依据和转化工具。

2. 材料与方法

2.1. 数据获取

GEO 数据库下载 ESCC 转录组数据集 GSE53625 和 GSE53622, 获取相关的原始芯片数据和临床信息。使用 limma 包进行背景校正和分位数归一化处理后, 排除术后生存期少于 30 天及随访信息不完整的病例。训练队列(GSE53625 数据集)用于构建预后预测模型; 验证队列(GSE53622 数据集)测试模型的临床应用效能。

2.2. 性别相关预后基因(GRPGs)筛选

训练队列中选择第 75 百分位基因, 使用“WGCNA”构建加权共表达网络, 确定软阈值 $\beta = 7$ ($R^2 = 0.90$), 识别共表达模块(最小基因数 30), 评估基因与性别相关性, 合并相关系数 > 0.85 模块, 筛选性别相关关键模块, 提取前 10% 枢纽基因。K-M 分析确定的预后基因($P < 0.05$)和加权基因共表达网络分析(Weighted Gene Co-expression Network Analysis, WGCNA)筛选出的性别相关模块基因取交集。最小绝对收缩和选择算子(Least Absolute Shrinkage and Selection Operator, LASSO)回归结合 10-fold 交叉验证(迭代 2000 次)优化模型, 确定最优 GRPGs。

2.3. 列线图开发与验证

基于多基因表达谱建立风险评分, 公式为 $\Sigma(\beta_i \times \text{Expri})$, 其中 β_i 为 Cox 回归系数, Expri 为标准化表达值。根据风险评分的最佳截断值将样本分为高风险组和低风险组。将单因素分析中 $P < 0.20$ 的影响因素纳入多因素 Cox 比例风险回归, 构建预后预测模型。采用校准曲线、受试者工作特征(Receiver Operating Characteristic, ROC)曲线和决策曲线分析(Decision Curve Analysis, DCA)评价模型预测效能。

2.4. 关键基因的生物医学意义和临床特征

采用基因集富集分析(Gene Set Enrichment Analysis, GSEA)和免疫浸润分析等方法探讨关键基因的生物医学意义。征得伦理委员会批准, 收集南阳医专一附院 2010~2020 年间的 75 例食管鳞状细胞癌患者的石蜡包埋组织, 制备为组织芯片。详细记录并核实患者的临床病理信息, 电话随访了解其生存状态。采用免疫组织化学(Immunohistochemistry, IHC)方法检测关键基因的蛋白表达, 两名病理医师双盲评估 H-score 分值。

2.5. 统计分析

R 软件(4.5.0)进行数据分析, $P < 0.05$ 具有统计学意义。t 检验或 Mann-Whitney U 检验分析连续变量, ANOVA 或 Kruskal-Wallis 检验进行多组比较, Kaplan-Meier 法(log-rank 检验)进行生存分析, 多因素 Cox 回归分析筛选预后因素。

3. 结果

3.1. GRPGs 的筛选

WGCNA 筛选出与性别相关的模块基因, 提取前 10% 的 763 个基因与 4038 个预后基因取交集(图 1(A) & 图 1(B)), Lasso-Cox 回归模型进一步筛选出 7 个 GRPGs。基于 GRPGs 构建的风险评分为 $(0.202 \times \text{TSPAN2}) + (-0.203 \times \text{SLC25A48}) + (0.144 \times \text{F2RL2}) + (-0.414 \times \text{KLHDC8B}) + (0.223 \times \text{ANGPTL7}) + (-0.248 \times \text{IL5RA}) + (-0.473 \times \text{GRIN3A})$, 模型将患者分为高风险组、低风险组。风险分布图显示高风险组患者生存率显著降低(图 1(C)), K-M 生存曲线证实其预后较差($P < 0.001$, 图 1(D)), 验证队列也观察到高、低风险组间存有 OS 差异(图 1(E), 图 1(F))。

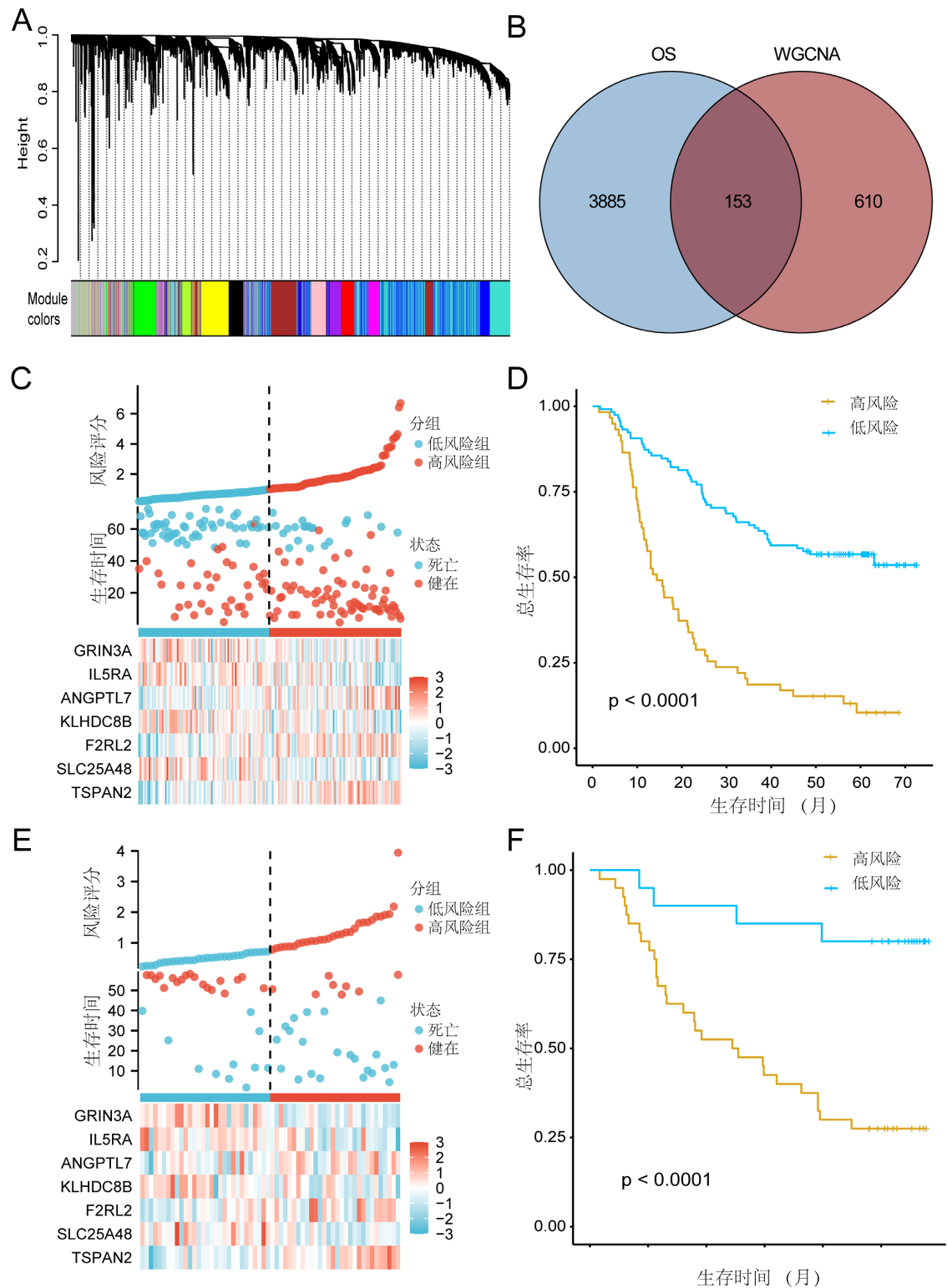


Figure 1. Identification of Gender-Related Prognostic Genes (GRPGs). (A) WGCNA analysis; (B) Venn diagram; (C) (D) Risk factor plot and Kaplan-Meier survival curves in the training cohort; (E) (F) Risk factor plot and Kaplan-Meier survival curves in the validation cohort

图 1. 性别相关预后基因的筛选。(A) WGCNA 分析; (B) 维恩图; (C) (D) 训练队列的风险因子图和 K-M 生存曲线; (E) (F) 验证队列的风险因子图和 K-M 生存曲线

3.2. 列线图的开发和验证

将单变量分析中 $P < 0.2$ 的因素(风险评分、年龄、肿瘤部位、肿瘤分化、N 分期、TNM 分期及治疗方式)纳入多变量 Cox 回归分析, 最终确定风险评分、年龄和治疗方式为独立预后因素。基于此, 构建包含独立因素的预后模型预测患者 OS (图 2(A))。校准曲线显示患者 1 年、3 年和 5 年的 OS 预测概率与实际生存概率一致, 时间依赖性 AUC 分别为 0.727、0.808 和 0.822 (图 2(B))。DCA 分析表明该模型的临床净收益优于单一变量, 包括 TNM 分期(图 2(C))。验证队列的评判结果进一步支持模型的临床预测效能(图 2(D), 图 2(E))。

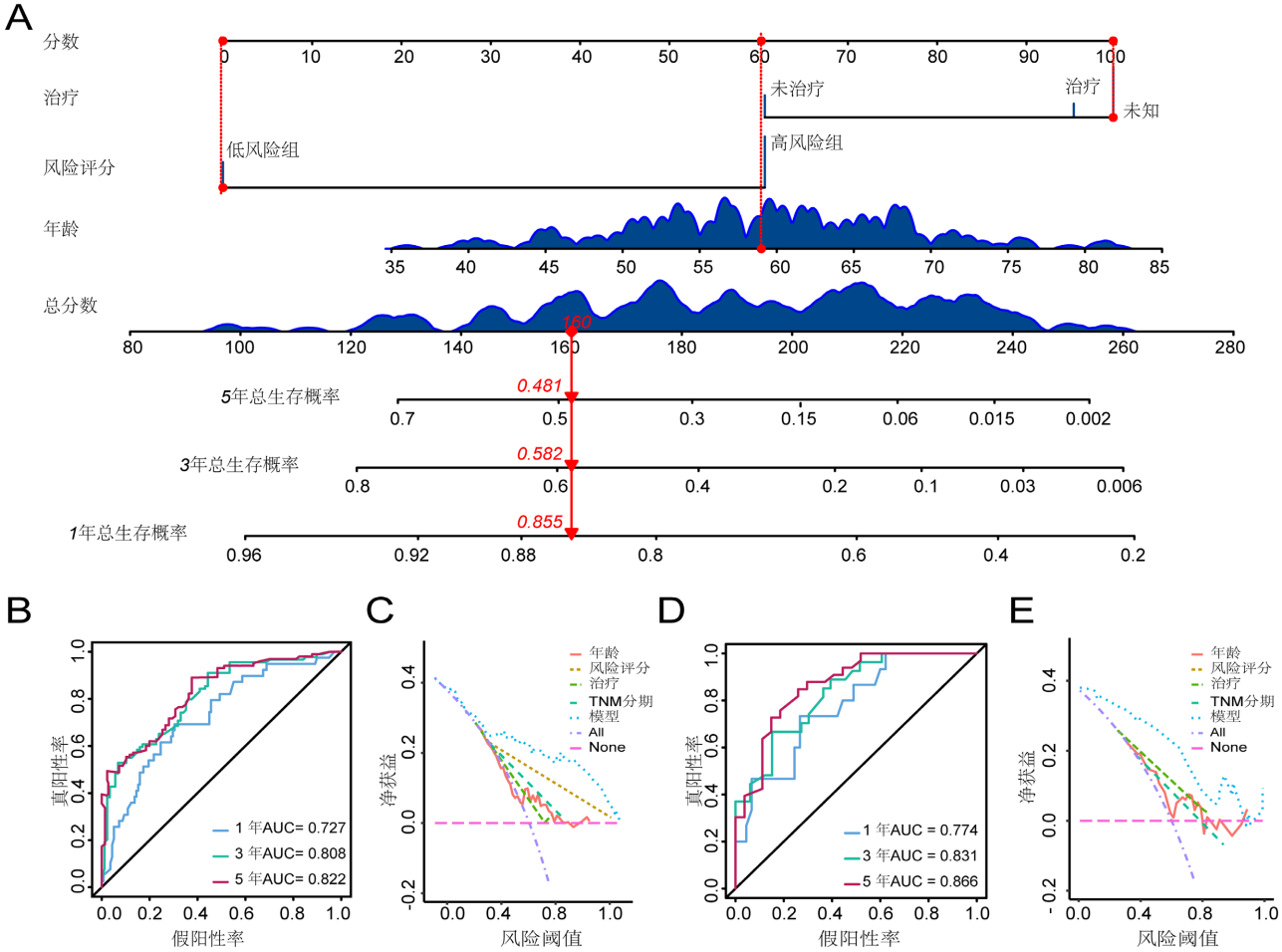


Figure 2. Construction and validation of the prognostic model. (A) Nomogram; (B) (C) receiver operating characteristic (ROC) analysis and decision curve analysis (DCA) in the training cohort; (D) (E) ROC analysis and DCA in the validation cohort

图 2. 预后预测模型的构建和验证。(A) 列线图; (B) (C) 训练队列中的受试者工作特征曲线(ROC)和临床决策曲线分析(DCA); (D) (E) 验证队列中的 ROC 曲线、DCA

3.3. IL5RA 生物学意义及临床特征

女性患者、低风险组患者高表达 IL5RA mRNA (图 3(A) & 图 3(B))。GSEA 分析显示 IL5RA 表达水平与雌激素代谢通路及干扰素代谢正相关(图 3(C), 图 3(D)); 免疫浸润分析提示女性 ESCC 肿瘤微环境中嗜酸性粒细胞浸润水平高于男性(图 3(E))。进一步临床检测显示, IL5RA 蛋白主要表达于间质细胞的细胞膜, 依据 H score 值将患者分为高表达组、低表达组(图 3(F))。两组间有显著差异, 且女性表达高于

男性(图 3(G)), 这和其 mRNA 表达水平一致。生存分析显示高表达 IL5RA 患者预后较好, 女性生存优势明显($P < 0.05$, 图 3(H), 图 3(I))。

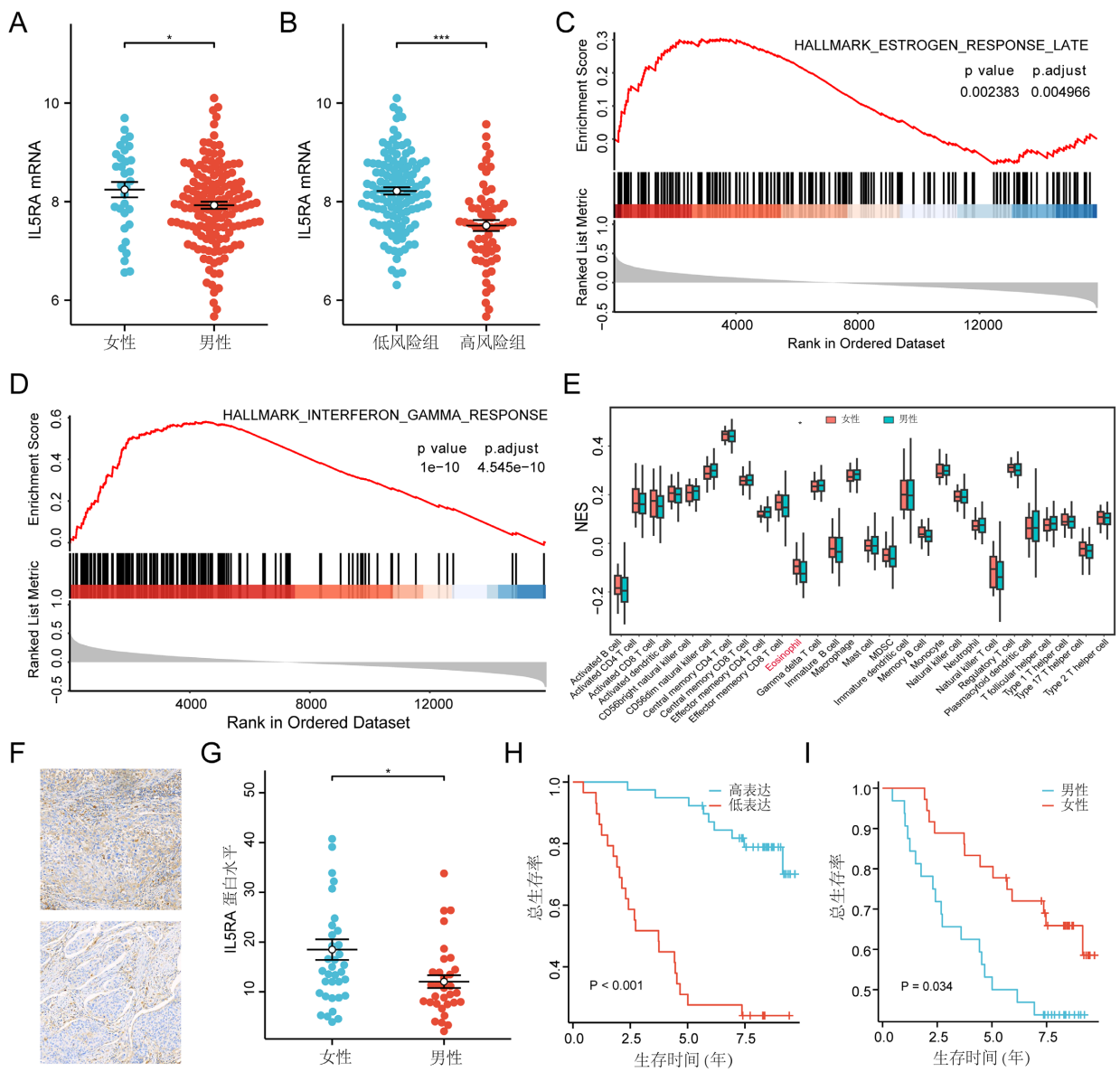


Figure 3. Biological implications and clinical characteristics of IL5RA. (A) IL5RA mRNA expression in male versus female patients; (B) IL5RA mRNA expression in high-risk versus low-risk groups; (C) (D) Gene set enrichment analysis (GSEA); (E) Immune infiltration analysis; (F) IL5RA protein expression; (G) IL5RA protein expression in male versus female patients; (H) Correlation between IL5RA protein expression and overall survival (OS); (I) Association between gender and overall survival (OS)

图 3. IL5RA 生物学意义及临床特征。(A) 男性、女性间 IL5RA mRNA 表达差异; (B) 高、低风险组间 IL5RA mRNA 表达差异; (C) (D) 基因集富集分析; (E) 免疫细胞浸润分析; (F) IL5RA 蛋白表达水平; (G) 男性、女性间 IL5RA 蛋白表达差异; (H) IL5RA 蛋白表达水平与总生存期关系; (I) 性别与 OS 关系

4. 讨论

临床实践中, 由于肿瘤异质性和肿瘤微环境调控网络的复杂性与多样性, ESCC 的分子靶向治疗仍然面临挑战[8] [9]。作为一种基于统计学的预测建模工具, 列线图提供了一种直观的图形化评判方式, 有效

帮助医生进行决策[10]。诸多肿瘤研究报道表明,临床列线图模型在预后预测的准确性和可靠性方面,优于传统 TNM 分期,这为临床治疗方案的制定提供了思路和依据[11][12]。本研究将基因组风险评分与患者年龄、治疗方案等临床参数结合构建了一个整合列线图模型,与传统 TNM 分期相比,该模型显著提升了预后预测准确性,其一致性指数(C-index)分别为 0.727、0.582。以一名 59 岁、低风险、接受新辅助治疗的患者为例,模型预测其 5 年总生存率为 48.1%,显著高于 TNM 分期预测的 32.7%。DCA 分析显示在一定阈值范围内,模型的临床适用性能覆盖 83%患者群体,且优于传统 TNM 分期。开发性别差异基因驱动的风险分层模型及列线图可在一定程度上突破传统 TNM 分期的预后预测评估瓶颈。

虽已有研究利用 SEER 数据库构建 ESCC 预后列线图以辅助风险分层和生存预测[13][14],但由于种族地域差异,这些模型对中国患者的适用性有限。本研究构建的模型的 C-index 优于上述两个模型(0.671 和 0.662);尽管 Yu 等模型的 C-index 更高(0.749) [15],但其仅适用于 T1 分期伴淋巴结转移的 ESCC 患者。我们前期开发的临床与生殖因素模型,通过整合雌激素受体 α (ESR1)和 β (ESR2)表达水平、绝经年龄与妊娠次数等生殖参数,显著提升了预测效能,但该模型仅适用于女性患者,尚缺乏对生殖因素作用机制的深入解析[16]。本研究在构建预后模型的基础上,进一步阐述了 ESCC 存有预后性别差异的分子机制,这符合癌症治疗中整合分子特征的精准诊疗趋势。

白细胞介素 5 受体 α (IL5RA)介导 IL-5 信号传导,在免疫炎症反应中起关键作用。本研究发现,无论是 mRNA 水平还是蛋白水平,女性患者中 IL5RA 表达均高于男性。生物信息学分析显示,IL5RA 通过调控雌激素响应通路和特异性募集嗜酸性粒细胞增强抗肿瘤免疫反应。这种激素-免疫交互作用与既往研究发现的嗜酸性粒细胞的浸润能改善 ESCC 患者预后的结果相吻合[17]。靶向 IL5RA 能够调节肿瘤免疫微环境,提升免疫治疗应答响应[18]。未来研究可聚焦于针对 ESCC 患者的联合靶向治疗策略,结合 IL-5 通路调节剂与雌激素受体调节剂,以协同调控 IL5RA 与雌激素的相互作用。这将重塑免疫微环境,有效应对 ESCC 异质性,但仍需进一步实验验证。

5. 结论

本研究成功构建了基于 GRPGs 的预后预测列线图,为临床实践中的个体化治疗和预后评估提供重要依据。IL5RA 介导的激素-免疫互作机制推动了性别分层策略的临床应用,强化了其作为潜在治疗靶点的应用价值。然而,样本量限制可能导致潜在生物标志物未被充分挖掘,回顾性分析难以完全排除混杂因素的干扰。未来研究需扩大中心、前瞻性队列的验证以增强结论可靠性,推动 ESCC 精准治疗策略向临床转化。

基金项目

河南省科技厅科技攻关项目(232102310137);河南省省部共建食管癌防治国家重点实验室开放课题(K2024-007)。

参考文献

- [1] Zhu, H., Ma, X., Ye, T., Wang, H., Wang, Z., Liu, Q., et al. (2023) Esophageal Cancer in China: Practice and Research in the New Era. *International Journal of Cancer*, **152**, 1741-1751. <https://doi.org/10.1002/ijc.34301>
- [2] Liu, S., Luo, L., Zhao, L., Zhu, Y., Liu, H., Li, Q., et al. (2021) Induction Chemotherapy Followed by Definitive Chemoradiotherapy versus Chemoradiotherapy Alone in Esophageal Squamous Cell Carcinoma: A Randomized Phase II Trial. *Nature Communications*, **12**, Article No. 4014. <https://doi.org/10.1038/s41467-021-24288-1>
- [3] Fatehi Hassanabad, A., Chehade, R., Breadner, D. and Raphael, J. (2020) Esophageal Carcinoma: Towards Targeted Therapies. *Cellular Oncology*, **43**, 195-209. <https://doi.org/10.1007/s13402-019-00488-2>
- [4] Xu, Q., Li, Q., Xu, Z., Lan, L., Hou, Z., Liu, J., et al. (2023) Prognostic Value of the Immunohistochemical Score Based

- on Four Markers in Head and Neck Squamous Cell Carcinoma. *Frontiers in Immunology*, **14**, Article ID: 1076890. <https://doi.org/10.3389/fimmu.2023.1076890>
- [5] Hu, J., Ma, Y., Ma, J., Yang, Y., Ning, Y., Zhu, J., *et al.* (2021) M2 Macrophage-Based Prognostic Nomogram for Gastric Cancer after Surgical Resection. *Frontiers in Oncology*, **11**, Article ID: 690037. <https://doi.org/10.3389/fonc.2021.690037>
- [6] 张冬云, 刘冉, 库建伟, 刘安丽, 赵学科, 张朋, 王静, 李寅, 吴红芳, 王立东. 婚育因素对中国女性食管鳞状细胞癌患者的生存影响[J]. 世界华人消化杂志, 2015(22): 3517-3525.
- [7] Zhang, D., Ku, J., Liu, R., Wu, H., Liang, G., Wei, Y., *et al.* (2017) Characterization of Serum Estradiol Level and Tissue Estrogen Receptor Immunostaining with Clinical Response and Reproductive Factor Changes in Chinese Female Patients with Esophageal Squamous Cell Carcinoma. *Biomedicine & Pharmacotherapy*, **93**, 879-884. <https://doi.org/10.1016/j.biopha.2017.07.020>
- [8] Dinh, H.Q., Pan, F., Wang, G., Huang, Q., Olingy, C.E., Wu, Z., *et al.* (2021) Integrated Single-Cell Transcriptome Analysis Reveals Heterogeneity of Esophageal Squamous Cell Carcinoma Microenvironment. *Nature Communications*, **12**, Article No. 7335. <https://doi.org/10.1038/s41467-021-27599-5>
- [9] Qi, L., Wang, J., Hou, S., Liu, S., Zhang, Q., Zhu, S., *et al.* (2025) Unraveling the Tumor Microenvironment of Esophageal Squamous Cell Carcinoma through Single-Cell Sequencing: A Comprehensive Review. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, **1880**, Article 189264. <https://doi.org/10.1016/j.bbcan.2025.189264>
- [10] Shariat, S.F., Capitanio, U., Jeldres, C. and Karakiewicz, P.I. (2009) Can Nomograms Be Superior to Other Prediction Tools? *BJU International*, **103**, 492-497. <https://doi.org/10.1111/j.1464-410x.2008.08073.x>
- [11] Ma, M., Wang, J., Hu, Y., Weng, M., Liu, X. and Wang, Y. (2019) Prognostic Value of Inflammatory Biomarkers in Gastric Cancer Patients and the Construction of a Predictive Model. *Digestive Surgery*, **36**, 433-442.
- [12] Feng, Y., Wang, Y., Xie, Y., Wu, S., Li, Y. and Li, M. (2021) Nomograms Predicting the Overall Survival and Cancer-specific Survival of Patients with Stage IIIC1 Cervical Cancer. *BMC Cancer*, **21**, Article No. 450.
- [13] Shao, C., Liu, X., Yao, S., Li, Z., Cong, Z., Luo, J., *et al.* (2021) Development and Validation of a New Clinical Staging System to Predict Survival for Esophageal Squamous Cell Carcinoma Patients: Application of the Nomogram. *European Journal of Surgical Oncology*, **47**, 1473-1480. <https://doi.org/10.1016/j.ejso.2020.12.004>
- [14] Du, F., Sun, Z., Jia, J., Yang, Y., Yu, J., Shi, Y., *et al.* (2020) Development and Validation of an Individualized Nomogram for Predicting Survival in Patients with Esophageal Carcinoma after Resection. *Journal of Cancer*, **11**, 4023-4029. <https://doi.org/10.7150/jca.40767>
- [15] Yu, J., Hu, W., Yao, N., Sun, M., Li, X., Wang, L., *et al.* (2021) Development and Validation of a Nomogram to Predict Overall Survival of T1 Esophageal Squamous Cell Carcinoma Patients with Lymph Node Metastasis. *Translational Oncology*, **14**, Article 101127. <https://doi.org/10.1016/j.tranon.2021.101127>
- [16] Zhang, D., Ku, J., Zhao, X., Zhang, H., Song, X., Wu, H., *et al.* (2022) Increased Prognostic Value of Clinical-Reproductive Model in Chinese Female Patients with Esophageal Squamous Cell Carcinoma. *World Journal of Gastroenterology*, **28**, 1347-1361. <https://doi.org/10.3748/wjg.v28.i13.1347>
- [17] 王立东, 马琳琳, 徐瑞华, 宋昕, 赵学科, 尤朵, 雷玲玲, 魏梦霞, 王苒, 胡景峰, 孟超龙, 李兴嵩, 周福有, 高社干. 食管软、硬癌分类和人工智能研究: 临床病理特征和嗜酸性粒细胞-TGF- β 1-纤维组织增生互作关系[J]. 郑州大学学报(医学版), 2023, 58(1): 1-6.
- [18] Zhu, Q., Chen, H., Li, X., Wang, X. and Yan, H. (2022) JMJD2C Mediates the MDM2/p53/IL5RA Axis to Promote CDDP Resistance in Uveal Melanoma. *Cell Death Discovery*, **8**, Article No. 227. <https://doi.org/10.1038/s41420-022-00949-y>