

# 基于机器学习预测新诊断喉癌患者的淋巴结转移

申帆<sup>1,2</sup>, 王永存<sup>3</sup>, 马万瑞<sup>1</sup>, 丁海峰<sup>1</sup>, 梁启廉<sup>3</sup>, 万崇华<sup>1,2</sup>

(1. 广东医科大学附属东莞第一医院, 广东 东莞 523808; 2. 广东医科大学生命质量与应用心理研究中心/生命质量与心理干预重点实验室, 广东 东莞 523808; 3. 广东医科大学附属医院, 广东 湛江 524023)

**摘要:** [目的] 建立一个机器学习预测模型, 可用于预测新诊断喉癌患者的淋巴结转移。[方法] 回顾性分析 2018—2020 年美国国立癌症研究所监测、流行病学和最终结果数据库 (SEER) 中新诊断喉癌患者的人口统计学变量和临床病理变量, 开发基于机器学习的淋巴结转移预测模型。利用 ROC 曲线下面积大小 (AUC)、准确率、召回率、精确率和特异度来评估和比较各机器学习模型的预测性能。[结果] 共纳入 4 462 例喉癌患者, 其中 1 444 例发生淋巴结转移。年龄、原发部位、临床分级和 T 分期可以作为独立的风险因素。原发部位、T 分期、年龄在淋巴结转移的重要性评估中排前 3 位。在所有的机器学习模型中, 除 KNN 模型预测效果较差外, Random Forest、Boost、GLM、Neuralnet、Decision Tree 这 5 种模型对于新诊断喉癌患者是否发生淋巴结转移都有较好的预测效果且分类能力近乎相同 (在测试集中 AUC 值均在 0.83 左右), 而准确率、召回率、特异度和精确率则在较小范围内各有优劣。[结论] 本研究构建的机器学习模型可以较好地预测新诊断喉癌患者是否发生淋巴结转移, 为临床医生的临床决策提供帮助。机器学习技术有可能帮助喉癌患者淋巴结转移预测模型的开发。

**关键词:** 淋巴结转移; 机器学习; 喉癌

中图分类号: R739.65 文献标识码: A 文章编号: 1004-0242(2023)12-0956-09

doi: 10.11735/j.issn.1004-0242.2023.12.A010

## Prediction Models for Lymphatic Metastasis in Newly Diagnosed Laryngeal Cancer Patients Based on Machine Learning

SHEN Fan<sup>1,2</sup>, WANG Yongcun<sup>3</sup>, MA Wanrui<sup>1</sup>, DING Haifeng<sup>1</sup>, LIANG Qilian<sup>3</sup>, WAN Chonghua<sup>1,2</sup>

(1. Dongguan First Hospital, Guangdong Medical University, Dongguan 523808, China; 2. Centre for Quality of Life and Applied Psychology Research/Key Laboratory of Quality of Life and Psychological Assessment Intervention, Guangdong Medical University, Dongguan 523808, China; 3. Affiliated Hospital of Guangdong Medical University, Zhanjiang 524023, China)

**Abstract:** [Purpose] To establish machine learning prediction models for lymphatic metastasis in patients with newly diagnosed laryngeal cancer. [Methods] The demographic and clinicopathological data of newly diagnosed laryngeal cancer patients obtained from the Surveillance, Epidemiology, and End Results database (SEER) during 2018 to 2020 were retrospectively analyzed. On this basis, the prediction models of lymphatic metastasis based on machine learning was developed. Area under ROC curve (AUC), accuracy, recall, specificity and precision were used to evaluate and compare the prediction performance of the established machine learning models. [Results] A total of 4 462 patients with newly diagnosed laryngeal cancer were included in the study, of whom 1 444 occurred lymphatic metastasis. Age, primary site, clinical grade and T stage were independent risk factors for lymphatic metastasis. The primary site, T stage and age were the top three risk factors of lymphatic metastasis. Among all machine learning models except the KNN model, including Random Forest, Boost, GLM, Neuralnet and Decision Tree had good prediction effects (AUC value in the test set was about 0.83), while accuracy, recall, specificity and precision had their own advantages and disadvantages in a small range. [Conclusion] The machine learning models constructed in this study can be used for predicting lymphatic metastasis in patients with newly diagnosed laryngeal cancer, indicating that machine learning technology has a potential to help clinicians for clinical decision making.

**Key words:** lymphatic metastasis; machine learning; laryngeal cancer

收稿日期: 2023-07-18; 修回日期: 2023-09-26

基金项目: 国家自然科学基金 (71974040, 81273185)

通信作者: 万崇华, E-mail: wanchh1964@qq.com

喉癌是头颈部的恶性肿瘤,是第二常见的呼吸系统恶性肿瘤,其中95%以上是鳞状细胞癌。2018年,全球有177 422例喉癌新发病例和94 771例死亡病例<sup>[1]</sup>。崔静等<sup>[2]</sup>在一项中国喉癌发病与死亡趋势分析和预测中发现,1990—2019年喉癌发病率与死亡率逐年升高,发病率由1.19/10万上升至3.09/10万,死亡率由0.97/10万上升至1.42/10万。根据其解剖亚部位,喉癌一般分为声门癌、声门上癌和声门下癌,鳞状细胞癌占组织学类型的绝大多数<sup>[3]</sup>。淋巴结转移是头颈部肿瘤最重要的预后指标,局部和远处转移是患者死亡的主要原因,而目前对于颈部淋巴结转移模式并未完全达成共识,特别是隐匿性转移,其中声门上型喉癌更为显著<sup>[4-5]</sup>。

许多淋巴结转移可能没有明显的症状或只表现为非特异性的颈部肿块、不适或压痛等,这种情况下,容易被误解为良性病变或其他非恶性原因,从而延误了诊断和治疗时机。并且对于初诊喉癌患者,医生可能更关注原发肿瘤的治疗,而对淋巴结转移的筛查和评估不够重视。因此在喉癌患者的初始诊断中往往忽视掉淋巴结转移的检测,或者是通过颈部触诊等非侵入性操作来进行评估,对淋巴结的评估主要依靠淋巴结短径、形态、钙化、囊变等征象,主观性较强,判断不准确<sup>[6]</sup>。目前的非侵入性操作检测手段主要是颈部触诊以及影像学检查,包括超声、计算机断层扫描(CT)、磁共振成像(MRI)和正电子发射计算机断层扫描(PET/CT)等。由于PET-CT价格高昂,超声检查受限于操作者的水平,影响其结果的稳定性,因此颈淋巴结转移的常规诊断仍不能依赖PET/CT和超声检测。虽然MRI在肿瘤侵犯范围及软组织显像方面优于CT,但增强MRI和增强CT对于头颈鳞状细胞癌淋巴结转移的诊断结果无显著差异。比较超声、超声引导下细针穿刺活检、CT和MRI检测,发现超声引导下细针穿刺活检对可疑淋巴结的诊断效果最好<sup>[3]</sup>。但是由于超声引导下细针穿刺活检存在辐射损伤、手术费用高、侵入性操作等缺陷,一般仅建议在喉癌患者出现淋巴结转移相关事件的情况下才会进行活检检查。此外对术前辅助检查未发现颈部淋巴结转移的患者,是否存在隐匿性颈部淋巴结转移的风险及是否需要常规行淋巴结清扫术仍存在争议<sup>[6]</sup>。由于癌症医疗数据的庞大数量和复杂性,机器学习得以在疾病诊断和结果预测中

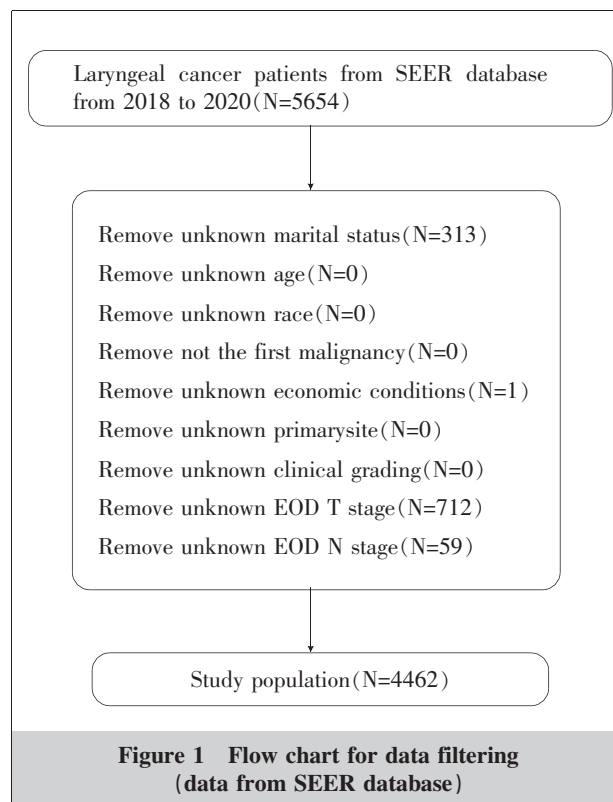
发挥巨大作用。Ho等<sup>[7]</sup>在研究中发现鼻咽癌淋巴结转移遵循可预测和有序的模式,这也为基于机器学习来进行喉癌淋巴结转移的预测提供了重要启示,相信机器学习可以为临床上进行喉癌患者淋巴结转移检查提供新思路。为此,本文基于机器学习进行初诊喉癌患者的淋巴结转移预测。

## 1 资料与方法

### 1.1 研究人群

本研究数据来自美国国立癌症研究所监测、流行病学和最终结果(surveillance, epidemiology and end results, SEER)数据库。使用SEER\*stat 8.3.9.2软件从“SEER Research data. 17 registries. Nov 2022, sub (2000—2020)”中下载。本研究仅限于2018—2020年的数据资料,因为该数据库最新的数据更新到2020年,新的EOD分期系统包含喉癌患者从2015—2020年的患者分期信息,但是患者临床分级的信息仅在2018年及以后可以使用。

排除标准:①年龄、性别、种族、婚姻状况、肿瘤分级、肿瘤部位等情况未知;②喉癌不是第一个肿瘤(Figure 1)。



## 1.2 数据选择

在这项研究中,共纳入了 10 个人口统计学变量和临床病理变量(Table 1)。人口变量包括性别、年龄、种族、婚姻状况、经济状况,临床病理变量包括临床分级、侧位性、组织学类型、EOD T 分期以及原发部位。以上所有的人口统计学变量以及临床病理变量都是来源于 SEER 数据库中喉癌患者初步诊断时的信息。在这些变量中,除年龄为数值型变量外,其余均为因子型变量。所有方法均根据 SEER 数据库的相关指南进行。

## 1.3 机器学习方法

本次研究总共采用 6 种常见的机器学习方法,包括逻辑回归(LR)、梯度提升(Boost)、决策树(Decision Tree)、随机森林(Random Forest)、k-最近邻(KNN)以及神经网络(Neuralnet)。逻辑回归模型用于研究性状变量对目标变量的影响,是一种特殊情况下的广义线性模型(GLM)。目标变量通常是二元分类器,例如喉癌患者中是否存在淋巴结转移<sup>[8]</sup>。决策树模型由决策点、策略点(事件点)及结果构成的树形图,一般应用于序列决策中,通常以最大收益期望值或最低期望成本作为决策准则,通过图解方式求解在不同条件下各类方案的效益值,然后通过比较,做出决策。决策树模型可以准确识别多种肿瘤组织病理,分类率高<sup>[9-10]</sup>。梯度提升以前用于预测 miRNA 疾病的关联,是一种在梯度提升框架下实现的机器学习算法<sup>[11]</sup>。随机森林可用于减少训练方差并改善积分和泛化,它是指使用多个树来训练和预测样本的机器学习分类器<sup>[12]</sup>。KNN 是使用最广泛的非参数分类方法之一,它基于这样的准则:如果样本附近的大多数 k 最近的样本都属于特征空间中的特定类,则该样本也属于此类别<sup>[13]</sup>。而神经网络是对生物神经系统的交互反应进行模拟,由具有适应性的简单单元组成广泛并行的互连网络。神经网络各层之间通过权值连接,通过激活函数对输出进行控制。其基本结构包括:a. 输入层,对预后特征进行表示;b. 权值矩阵,位于输入层与隐藏层之间,对预后特征进行加权处理;c. 激活函数,位于隐藏层与输出层之间,用于计算并输出最终的结果;d. 输出层,输出癌症的结果<sup>[14]</sup>。

## 1.4 模型建立

研究中的所有统计分析都使用 R (R Foundation

for Statistical Computing, version 4.3.1) 来完成,在构建机器学习模型中所使用的 R 语言函数包包括:①randomForest;②gbm;③neuralnet;④C50;⑤class;⑥ROCR。所有患者按照 7:3 随机分为训练集和测试集,并使用卡方检验和秩和检验的方法筛选了训练集和测试集中对于淋巴结转移有影响的变量,并将作为构建机器学习模型的纳入因素。训练集用于建立随机森林(Random Forest)、梯度提升(Boost)、神经网络(Neuralnet)、决策树(Decision Tree)、k-最近邻(KNN)和广义线性模型(GLM),并用测试集对其进行性能评估。所有模型均使用 caret 函数包通过十折交叉验证的方式,进行最佳参数确定。确定采用 ROC 曲线下面积大小(AUC)、准确率、召回率、特异度、精确率 5 种指标对模型的预测能力进行比较和评估。

# 2 结 果

## 2.1 人口统计和临床病理特征

共有 4 462 例喉癌患者参与了这项研究,其中有 1 444 例初次诊断时被发现淋巴结转移(32.36%),3 018 例患者没有发现淋巴结转移(67.64%)。所有患者以 7:3 的比例完全随机分为训练集 (n=3 123)和测试集(n=1 339)(Table 1)。在 10 个人口统计学和临床病理特征中,除种族、经济状况和侧位性外,其余变量无论在训练集还是测试集中均显示与是否发生淋巴结转移之间的关联差异存在统计学意义 ( $P$  均  $<0.05$ )。所以将除种族、经济状况和侧位性外的其他 7 个变量纳入机器学习模型进行分析。

## 2.2 模型分析及变量对预测的影响

测试所有被纳入变量之间的 Pearson 相关性,相关热图(Figure 2)显示所有纳入分析的变量之间的相关性强度均较弱。对于采用输入变量选择方法的多变量 LR 模型,有 4 个特征被确定为独立风险因素,包括年龄、原发部位、临床分级和 T 分期(Table 2)。对于随机森林模型,变量的重要性是以袋外(OOB)错误率来评估的,它可以反映出每个变量在分类淋巴结转移与无淋巴结转移时的贡献(Figure 3)。原发部位,其次是 T 分期和年龄是最重要的三个变量。在 Random Forest 模型中,最重要的三个变量与 LR 模型所筛选的对于淋巴结转移的影

**Table 1** Demographic and clinicopathological characteristics of the training and test sets[n(%)]

| Variable                         | Training set                        |                                  |          | Test set                           |                                 |          |
|----------------------------------|-------------------------------------|----------------------------------|----------|------------------------------------|---------------------------------|----------|
|                                  | No lymphatic metastasis<br>(n=2106) | Lymphatic metastasis<br>(n=1017) | <i>P</i> | No lymphatic metastasis<br>(n=912) | Lymphatic metastasis<br>(n=427) | <i>P</i> |
| Gender                           |                                     |                                  |          |                                    |                                 |          |
| Female                           | 368(17.5)                           | 267(26.3)                        | <0.001   | 148(16.2)                          | 112(26.2)                       | <0.001   |
| Male                             | 1738(82.5)                          | 750(73.7)                        |          | 764(83.8)                          | 315(73.8)                       |          |
| Race                             |                                     |                                  |          |                                    |                                 |          |
| Black                            | 276(13.1)                           | 188(18.5)                        | <0.001   | 136(14.9)                          | 80(18.7)                        | 0.208    |
| White                            | 1745(82.9)                          | 785(77.2)                        |          | 729(79.9)                          | 326(76.3)                       |          |
| Others                           | 85(4.0)                             | 44(4.3)                          |          | 47(5.2)                            | 21(4.9)                         |          |
| Marital status                   |                                     |                                  |          |                                    |                                 |          |
| Divorced                         | 279(13.2)                           | 154(15.1)                        | <0.001   | 121(13.3)                          | 71(16.6)                        | <0.001   |
| Married                          | 1184(56.2)                          | 443(43.6)                        |          | 510(55.9)                          | 179(41.9)                       |          |
| Unmarried                        | 448(21.3)                           | 323(31.8)                        |          | 209(22.9)                          | 136(31.9)                       |          |
| Widowed                          | 195(9.3)                            | 97(9.5)                          |          | 72(7.9)                            | 41(9.6)                         |          |
| Economic conditions(USD)         |                                     |                                  |          |                                    |                                 |          |
| <60000                           | 659(31.3)                           | 353(34.7)                        | <0.001   | 274(30.0)                          | 131(30.7)                       | 0.686    |
| 60000~75000                      | 484(23.0)                           | 280(27.5)                        |          | 234(25.7)                          | 117(27.4)                       |          |
| >75000                           | 963(45.7)                           | 384(37.8)                        |          | 404(44.3)                          | 179(41.9)                       |          |
| Primary site                     |                                     |                                  |          |                                    |                                 |          |
| Larynx and glottis               | 1562(74.2)                          | 228(22.4)                        | <0.001   | 689(75.5)                          | 115(26.9)                       | <0.001   |
| Supraglottis                     | 513(24.4)                           | 757(74.4)                        |          | 207(22.7)                          | 301(70.5)                       |          |
| Subglottis                       | 31(1.5)                             | 32(3.1)                          |          | 16(1.8)                            | 11(2.6)                         |          |
| EOD T stage                      |                                     |                                  |          |                                    |                                 |          |
| T <sub>1</sub>                   | 1177(55.9)                          | 151(14.8)                        | <0.001   | 510(55.9)                          | 70(16.4)                        | <0.001   |
| T <sub>2</sub>                   | 414(19.7)                           | 242(23.8)                        |          | 175(19.2)                          | 105(24.6)                       |          |
| T <sub>3</sub>                   | 335(15.9)                           | 329(32.4)                        |          | 150(16.4)                          | 141(33.0)                       |          |
| T <sub>4</sub>                   | 180(8.5)                            | 295(29.0)                        |          | 77(8.4)                            | 111(26.0)                       |          |
| Clinical grading                 |                                     |                                  |          |                                    |                                 |          |
| Low                              | 1154(54.8)                          | 453(44.5)                        | <0.001   | 496(54.4)                          | 206(48.2)                       | <0.001   |
| Medium                           | 205(9.7)                            | 207(20.4)                        |          | 86(9.4)                            | 92(21.5)                        |          |
| Others                           | 747(35.5)                           | 357(35.1)                        |          | 330(36.2)                          | 129(30.2)                       |          |
| Lateral                          |                                     |                                  |          |                                    |                                 |          |
| Unilateral                       | 2077(98.6)                          | 1005(98.8)                       | 0.650    | 895(98.1)                          | 423(99.1)                       | 0.203    |
| Bilateral                        | 29(1.4)                             | 12(1.2)                          |          | 17(1.9)                            | 4(0.9)                          |          |
| Histological type                |                                     |                                  |          |                                    |                                 |          |
| Squamous cell neoplasms          | 2072(98.4)                          | 985(96.9)                        | 0.005    | 902(98.9)                          | 415(97.2)                       | 0.021    |
| Others                           | 34(1.6)                             | 32(3.1)                          |          | 10(1.1)                            | 12(2.8)                         |          |
| Age( $\bar{x}\pm s$ )(years old) | 66.19±10.67                         | 62.48±9.52                       | <0.001   | 65.98±10.71                        | 62.64±9.76                      | 0.001    |

响显著水平最高的三个变量是一致的。

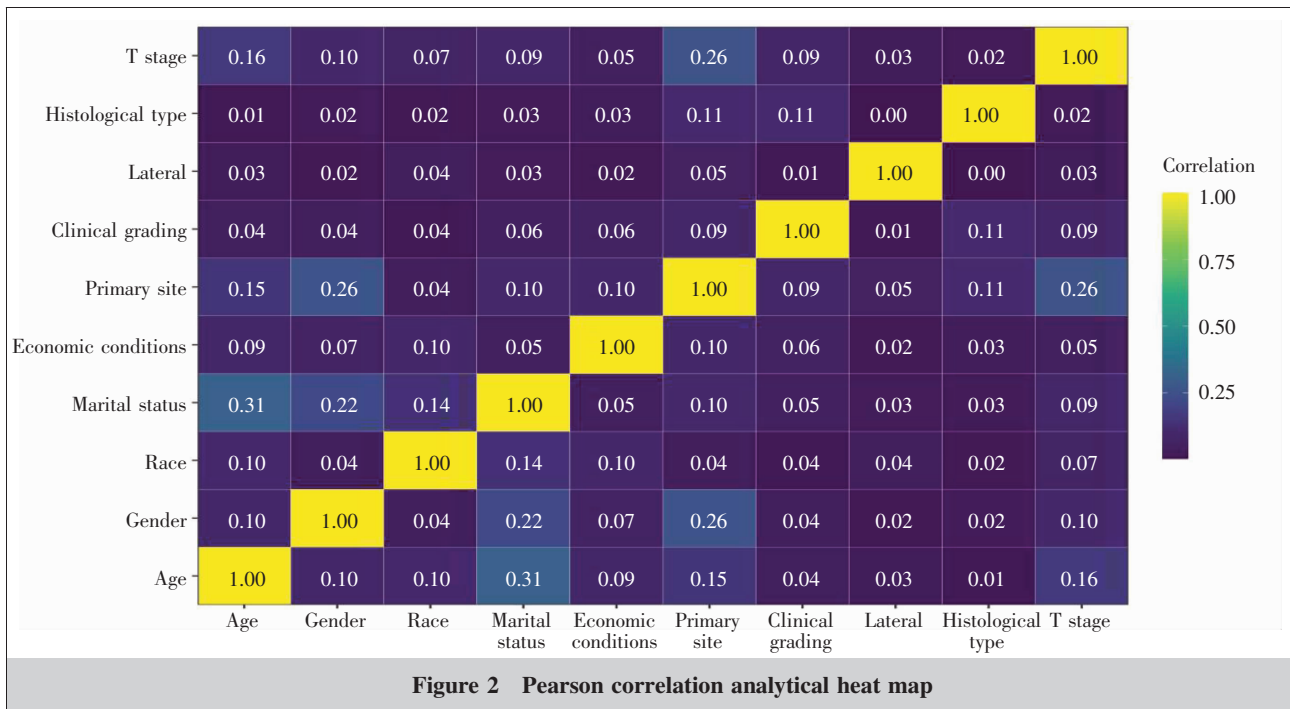
**2.3 模型性能**

应用训练集和测试集用于测试和比较所有模型的预测性能。使用 AUC、准确率、召回率、特异度和精确率 5 个指标来评估和比较模型性能。通过比较发现，在所有的机器学习模型中，Random Forest、Boost、GLM、Neuralnet、Decision Tree 分类能力近乎相同 (在测试集中 AUC 值均在 0.83 左右)，而准确率、召回率、特异度和精确率则在较小范围内各有优

劣。而在测试集中，KNN 模型的 AUC 值低于前面的 4 种模型并且在特异度指标方面也表现的并不理想。所以我们认为 Random Forest、Boost、GLM、Neuralnet、Decision Tree 这 5 种模型相较于 KNN 模型而言表现出了更好的预测能力(Table 3~4, Figure 4)。

**3 讨论**

通过对人类癌症的研究，淋巴结系统对肿瘤转



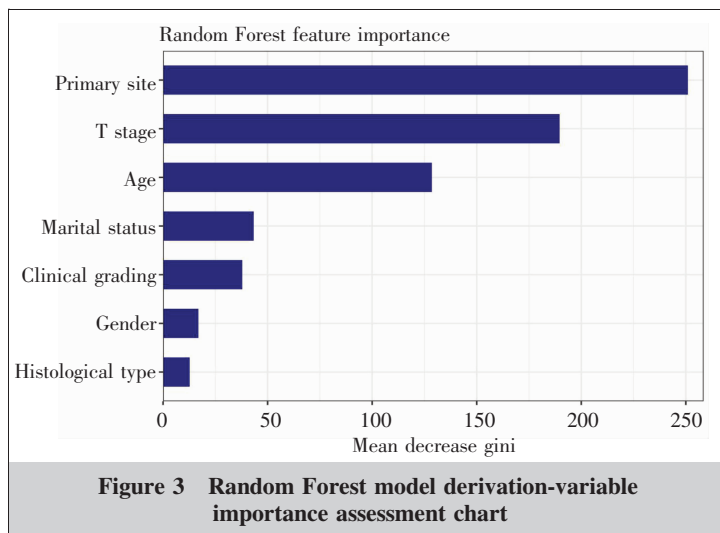
**Table 2 Multivariable Logistic regression model with enter variable selection**

| Variable                  | Estimated value | Z      | P      |
|---------------------------|-----------------|--------|--------|
| Constant term             | -2.0342         | -5.605 | <0.001 |
| Age                       | -0.0169         | -3.370 | <0.001 |
| Gender(Male)              | 0.0781          | 0.667  | 0.505  |
| Marital status(Married)   | -0.0327         | -0.232 | 0.816  |
| Marital status(Unmarried) | 0.1689          | 0.152  | 0.267  |
| Marital status(Widowed)   | -0.0854         | -0.423 | 0.672  |
| Primarysite(Supraglottis) | 2.1190          | 20.129 | <0.001 |
| Primary site(Subglottis)  | 1.3782          | 0.296  | <0.001 |
| Clinical grading(Others)  | 0.2803          | 0.106  | 0.008  |
| Clinical grading(Medium)  | 0.6917          | 0.137  | <0.001 |
| Histological type(Others) | 0.2021          | 0.697  | 0.486  |
| T Stage(T <sub>2</sub> )  | 0.9864          | 7.445  | <0.001 |
| T Stage(T <sub>3</sub> )  | 1.3604          | 10.484 | <0.001 |
| T Stage(T <sub>4</sub> )  | 2.3728          | 16.276 | <0.001 |

移的贡献正日益受到重视<sup>[15]</sup>。喉癌患者的淋巴结转移会导致转移部位的淋巴结肿大,出现声音嘶哑、喉咙疼痛、吞咽困难、体重下降、呼吸困难等症状,极大地影响患者的生命质量。颈部淋巴结转移是影响喉癌患者生存的重要因素之一<sup>[16]</sup>。在一项研究报道中发现,在病理学评估中早期(T<sub>1</sub>/T<sub>2</sub>)原发性患者在55%的病例中显示颈部淋巴结转移,而T<sub>3</sub>/T<sub>4</sub>期肿瘤患者的比例为39%<sup>[17]</sup>。而在本研究中喉癌患者淋巴结转移的发生率(32.36%)要低于报道中的

发生率,这可能是由于SEER数据库中记录的数据是喉癌患者初次诊断时的数据,而报道中的关于淋巴结转移的数据是不同时期的累积数据,因此在这项研究中淋巴结转移的发生率较低。综上所述,对于喉癌患者而言,在初次诊断时发生淋巴结转移的发生率相对较低,会有不少患者的淋巴结转移会在喉癌初次诊断后的临床随访中发生。因此,在喉癌患者被初步诊断后,对发生淋巴结转移可能性较高的患者进行进一步的随访检查对于后续的治疗和改善预后非常重要。影像学检查通常用于识别新诊断为喉癌的患者可能发现的淋巴结转移。然而,由于影像学检查昂贵且有辐射损伤,因此使用这种方法可能不适合进一步的随访检查。病理诊断被认为是金标准。然而,研究表明,活检不仅困难和痛苦,而且还会增加肿瘤细胞增殖的风险,这意味着它可能不适合常规诊断<sup>[18]</sup>。为了解决这个问题,我们使用先进的机器学习算法来构建模型用于识别淋巴结转移高风险喉癌患者。

随着人工智能技术的蓬勃发展和多学科交叉合作,人工智能技术在喉癌诊断、治疗决策和预后预测等领域都取得了良好的进展<sup>[19]</sup>。机器学习作为人工智能的一个分支,可以从一系列复杂的数据集中挖掘重要的模式,故可以应用到癌症数据的分析中<sup>[14]</sup>。例如李祖飞等<sup>[20]</sup>通过人工智能线性判别分析



**Table 3 Comparison of prediction performance of different machine learning models(based on the training set)**

| Model         | AUC    | Accuracy | Recall rate<br>(sensitivity) | Specificity | Precision |
|---------------|--------|----------|------------------------------|-------------|-----------|
| Random Forest | 0.8817 | 0.8120   | 0.8651                       | 0.7021      | 0.8574    |
| Boost         | 0.8358 | 0.7919   | 0.8234                       | 0.7266      | 0.8618    |
| Neuralnet     | 0.8488 | 0.7826   | 0.7953                       | 0.7561      | 0.8710    |
| Decision Tree | 0.8404 | 0.7704   | 0.8248                       | 0.6578      | 0.8331    |
| KNN           | 0.8718 | 0.7871   | 0.9102                       | 0.5319      | 0.8010    |
| GLM           | 0.8493 | 0.7746   | 0.8399                       | 0.6391      | 0.8281    |

**Table 4 Comparison of prediction performance of different machine learning models(based on the test set)**

| Model         | AUC    | Accuracy | Recall rate<br>(sensitivity) | Specificity | Precision |
|---------------|--------|----------|------------------------------|-------------|-----------|
| Random Forest | 0.8327 | 0.7625   | 0.8410                       | 0.5948      | 0.8160    |
| Boost         | 0.8358 | 0.7662   | 0.8180                       | 0.6557      | 0.8354    |
| Neuralnet     | 0.8350 | 0.7640   | 0.7961                       | 0.6956      | 0.8481    |
| Decision Tree | 0.8227 | 0.7595   | 0.8289                       | 0.6112      | 0.8200    |
| KNN           | 0.7883 | 0.7311   | 0.8859                       | 0.4005      | 0.7594    |
| GLM           | 0.8344 | 0.7603   | 0.8432                       | 0.5831      | 0.8120    |

(linear discriminant analysis, LDA)算法建立判断喉癌5年生存状态的预测模型,取得了较好的预测效果。本研究纳入了多种比较常见的机器学习算法纳入进行分析和比较,最后发现 Random Forest、Boost、GLM、Neuralnet、Decision Tree 这5种机器学习模型展现出了良好的预测能力。虽然这5种机器学习模型展现出的预测能力中规中矩,并不是十分优秀,但是依然可以给临床医生提供参考。

这5种机器学习模型虽然分类能力近乎相同,但是临床医生对模型的选择运用应该考虑现实条件、各个模型的特点和所获取到的资源。从现实条件来说,在本次研究正式展开之前,对于构建机器学习

模型的变量筛选时就考虑到了所建立模型的普适性,各级医院的各级临床医生都可以非常轻松地在初次诊断为喉癌患者身上获取到,以便用于患者的淋巴结转移预测。从各个模型的特点来说,Random Forest、Neuralnet、Decision Tree 和 Boost 通常更为复杂,而 GLM 模型往往更为简单且结果更容易解释。如果临床医生在建立模型时更为考虑模型的可解释性时,应该考虑选择 GLM 模型。此外,虽然这5种机器学习模型在测试集中的 AUC 值近乎相同,但是准确率、召回率、特异度和精确率则在较小范围内各有优劣。当临床医生进行模型选择时更倾向于特异度、精确率、准确率时,则应考虑 Neuralnet 和 Boost。当临床医生更倾向于召回率(灵敏度)时,则应考虑 Random Forest 和 GLM 模型。从所获取的资源来说,不同级别和不同等级医院的临床医生由于现实情况可以获取到的资源不同,而 Random Forest、Boost、Neuralnet、Decision Tree 相比较 GLM 来说通常需要更多的计算资源和时间资源来训练和进行预测。因此,对于拥有资源较少的临床医生来说,考虑 GLM 模型是较为合适的选择。此外,对于一些拥有大量资源而又对本研究中所建立的机器学习模型的预测效果不太满意的临床医生来说,由于 Random Forest、Boost、Neuralnet 都是集成学习方法,可以通过组合多个弱分类器来提高性能。因此可以选择 Random Forest、Neuralnet 和 Boost 来建立模型,可以在本研究所选取的

变量和建立模型的基础上,借助医院资源和自身资源添加新变量和新分类器。根据其测试效果可以将其纳入到这两种模型中从而进一步提高这两种模型的预测能力。针对性地选择模型可以为临床医生做出进一步的治疗诊断决策提供帮助。该模型的用途在于肿瘤科的临床医生对喉癌患者进行初次确诊时,借助现有的诊断资料通过机器学习模型来预测喉癌患者淋巴结转移的发生。对于机器学习模型预测为淋巴结转移阳性的人群,将其纳入到淋巴结转移高风险人群中去。告知患者机器模型预测结果并做进一步的病理学和影像学检查以确定其是否发生淋巴结转移。对于机器学习模型预测结果为阳性的

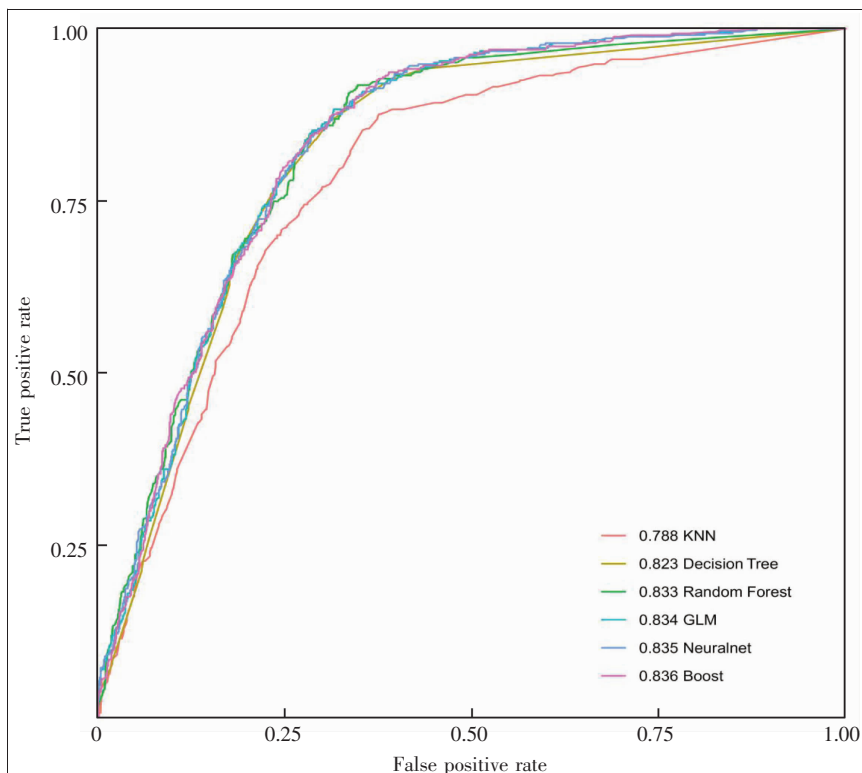


Figure 4 Machine learning model ROC curve

患者,但目前活检尚未显示淋巴结转移的患者,主诊医生也应该给予更大的关注,并且定期随访了解病情进展,必要时可对患者给予淋巴结转移高风险部位放疗的预防性照射。

在本次研究中,我们通过 LR 模型发现年龄、原发部位、临床分级和 T 分期可以作为淋巴结转移的独立风险因素( $P<0.05$ )。长期以来,只有喉癌患者出现明显的颈部疼痛以及颈部出现明显可触及的肿块时医生才会对患者进行靶向筛查,这明显是不合理的。因为这样会延迟患者的治疗,影响治疗效果。因此,有必要建立模型来预测淋巴结转移高风险的喉癌患者,并提供早期关注和筛查。

在临床上,为了预防喉癌患者出现淋巴结转移,会对淋巴结转移高风险的患者进行预防性的颈部淋巴结清扫。而临床淋巴结状态阴性患者的择期颈部清扫率取决于原发肿瘤的定位、分期和大小<sup>[21]</sup>。声门上癌、肿瘤大小 $>5$  cm 和 3~4 级分类是淋巴结转移的独立危险因素<sup>[22]</sup>。手术治疗的喉癌患者发生颈淋巴结转移的倾向因原发肿瘤部位和分期而异,对于 cN0 患者,隐匿性淋巴结转移的发生率均随肿瘤分期的增加而增加,且声门上肿瘤的发生率最高。隐匿

性淋巴结转移的预测因素包括声门上部位、病理分期 T<sub>3</sub> 和 T<sub>4</sub>、边缘阳性以及淋巴管侵犯<sup>[23]</sup>。在我们的研究中发现原发部位和 T 分期是喉癌患者淋巴结转移的重要风险因素。据报道,声门上癌和声门下癌比声门癌更常发生淋巴结转移,原因是声带淋巴引流稀少<sup>[21]</sup>。声门上型喉癌颈部转移率居各型喉癌之首,是影响患者预后的重要因素。原因是声门上区域具有丰富的淋巴网,其淋巴双向引流,故声门上型喉癌早期极易发生颈淋巴结转移,且双侧颈淋巴结转移风险较高<sup>[24]</sup>。这也为原发部位作为喉癌患者淋巴结转移的重要风险因素提供了理论支撑。在一项基于列线图建立喉癌患者淋巴结转移风险的预测模型研究中,肿瘤的临床分级、年龄、肿瘤大小成为

了重要的预测因素<sup>[25]</sup>。这也与本次研究中年龄和临床分级成为重要的风险因素相吻合。我们同样不能忽视婚姻状况对于新诊断喉癌患者淋巴结转移的影响,虽然未能成为淋巴结转移的独立风险因素,但是我们可以发现无论是训练集还是测试集,未婚群体在无淋巴结转移中的占比较少,而在出现淋巴结转移的占比中出现了较大的增幅。这表明新诊断喉癌的患者中,处于未婚状态相比较其他状态更容易发生淋巴结转移。此外,本次研究受制于 SEER 数据库的权限限制,并未将一些常见的生化指标纳入,因此临床医生在构建模型时,也可以试着将一些生化因子纳入。有研究淋巴细胞-单核细胞比值、中性粒细胞与淋巴细胞比值与血小板计数作为喉癌淋巴结转移的预测变量,也取得了较好的效果<sup>[25-26]</sup>。

本次研究是基于主流的机器学习方法和 SEER 数据库来建立喉癌患者淋巴结转移的预测模型,它扩展了许多研究人员经常使用的基于 LR 的列线图模型。本研究的 Random Forest、Boost、GLM、Neuralnet、Decision Tree 5 种模型相比较于 KNN 模型表现出了良好的预测能力。然而,也具有局限性。①筛选出的影响喉癌患者淋巴结转移的重要特征是基于机

器学习中的随机森林方法得到的,可能在临床上无法进行解释;②本次研究中的数据均来源于 SEER 数据库,是一项基于北美人群的研究,并没有收集到足够数量的本土喉癌患者从而对模型展开进一步临床验证;③SEER 数据库记录了患者初步诊断时的信息,意味着后续治疗数据缺失,我们无法将其纳入喉癌患者的淋巴结转移预测分析中去。

## 参考文献:

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6):394-424.
- [2] 崔静, 张倩. 中国喉癌发病与死亡趋势分析及预测[J]. *疾病监测*, 2023, 38(8):1000-1006.  
Cui J, Zhang Q. Analysis and prediction of the incidence and mortality trend of laryngeal cancer in China [J]. *Disease Surveillance*, 2023, 38(8):1000-1006.
- [3] Nahavandipour A, Jakobsen KK, Grønhoj C, et al. Incidence and survival of laryngeal cancer in Denmark: a nation-wide study from 1980 to 2014[J]. *Acta Oncol*, 2019, 58(7):977-982.
- [4] 胡晨, 张明, 薛继尧, 等. cN0 声门上型喉癌的颈淋巴结隐匿性转移及处理 [J]. *临床耳鼻咽喉头颈外科杂志*, 2020, 34(7):615-617.  
Hu C, Zhang M, Xue JY, et al. Concealed cervical lymph node metastasis and management of supraglottic laryngeal cancer in cN0[J]. *Journal of Clinical Otolaryngology Head and Neck Surgery*, 2020, 34(7):615-617.
- [5] 伍兴东, 傅然, 余林. 病理分化程度对声门上型喉癌颈清扫的影响[J]. *肿瘤防治研究*, 2021, 48(6):607-610.  
Wu XD, Fu R, Yu L. The effect of pathological differentiation on neck dissection of supraglottic laryngeal cancer [J]. *Cancer Prevention and Treatment Research*, 2021, 48(6):607-610.
- [6] 何知杨, 谷金铭, 李泉江, 等. 影像组学在喉癌诊疗中的应用综述[J]. *医疗卫生装备*, 2023, 44(3):102-107.  
He ZY, Gu JM, Li QJ, et al. A review of the application of imaging omics in the diagnosis and treatment of laryngeal cancer [J]. *Medical and Health Equipment*, 2023, 44(3):102-107.
- [7] Ho FC, Tham IW, Earnest A, et al. Patterns of regional lymph node metastasis of nasopharyngeal carcinoma: a meta-analysis of clinical evidence [J]. *BMC Cancer*, 2012, 12:98.
- [8] Nick TG, Campbell KM. Logistic regression [J]. *Methods Mol Biol*, 2007, 404:273-301.
- [9] Payabvash S, Aboian M, Tihan T, et al. Machine learning decision tree models for differentiation of posterior fossa tumors using diffusion histogram analysis and structural MRI findings[J]. *Front Oncol*, 2020, 10:71.
- [10] 殷子博. 基于决策树算法的癌症合成致死基因组合的预测和预后分析[D]. 南京:南京邮电大学, 2020.  
Yin ZB. Prediction and prognostic analysis of cancer synthetic lethal gene combinations based on decision tree algorithm [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020.
- [11] Chen X, Huang L, Xie D, et al. EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction[J]. *Cell Death Dis*, 2018, 9(1):3.
- [12] Jiang H, Mao H, Lu H, et al. Machine learning-based models to support decision-making in emergency department triage for patients with suspected cardiovascular disease[J]. *Int J Med Inform*, 2021, 145:104326.
- [13] Salvador-Meneses J, Ruiz-Chavez Z, Garcia-Rodriguez J. Compressed KNN: K-nearest neighbors with data compression[J]. *Entropy*, 2019, 21(3):234.
- [14] 高美虹, 尚学群. 利用人工智能预测癌症的易感性、复发性和生存期 [J]. *生物化学与生物物理进展*, 2022, 49(9):1687-1702.  
Gao MH, Shang XQ. Using artificial intelligence to predict cancer susceptibility, recurrence, and survival[J]. *Advances in Biochemistry and Biophysics*, 2022, 49(9):1687-1702.
- [15] Wakisaka N, Hirota K, Kondo S, et al. Induction of lymphangiogenesis through vascular endothelial growth factor-C/vascular endothelial growth factor receptor 3 axis and its correlation with lymph node metastasis in nasopharyngeal carcinoma[J]. *Oral Oncol*, 2012, 48(8):703-708.
- [16] Erdag TK, Guneri EA, Avincsal O, et al. Is elective neck dissection necessary for the surgical management of T2N0 glottic carcinoma? [J]. *Auris Nasus Larynx*, 2013, 40(1):85-88.
- [17] Kürten CHL, Zioga E, Gauler T, et al. Patterns of cervical lymph node metastasis in supraglottic laryngeal cancer and therapeutic implications of surgical staging of the neck[J]. *Eur Arch Otorhinolaryngol*, 2021, 278(12):5021-5027.
- [18] Zhang L, Xiang ZL, Zeng ZC, et al. A microRNA-based prediction model for lymph node metastasis in hepatocellular carcinoma[J]. *Oncotarget*, 2016, 7(3):3587-3598.
- [19] 刘秋蕊, 赵宇. 人工智能技术在喉癌诊疗领域中的应用[J]. *中国耳鼻咽喉颅底外科杂志*, 2021, 27(5):530-533.

- Liu QR,Zhao Y. Application of artificial intelligence technology in the diagnosis and treatment of laryngeal cancer [J]. Chinese Journal of Otolaryngology and Cranial Base Surgery, 2021, 27(5):530-533.
- [20] 李祖飞,黄志刚,房居高,等. 人工智能线性判别分析算法建立喉癌 5 年生存状态预测模型 [J]. 肿瘤预防与治疗, 2022, 35(2):120-126.
- Li ZF, Huang ZG, Fang JG, et al. Establishment of a 5-year survival prediction model for laryngeal cancer using artificial intelligence linear discriminant analysis algorithm [J]. Cancer Prevention and Treatment, 2022, 35(2):120-126.
- [21] Ketterer MC, Lemus Moraga LA, Beitingger U, et al. Surgical nodal management in hypopharyngeal and laryngeal cancer[J]. Eur Arch Otorhinolaryngol, 2020, 277(5):1481-1489.
- [22] Pan Y, Zhao X, Zhao D, et al. Lymph nodes dissection in elderly patients with T3-T4 laryngeal cancer [J]. Clin Interv Aging, 2020, 15:2321-2330.
- [23] Prasad A, Carey RM, Panara K, et al. Nodal metastasis in surgically treated laryngeal squamous cell carcinoma[J]. Head Neck, 2023, 45(9):2303-2312.
- [24] 马涛,吴干勋,赵瑞力,等. 声门上型喉癌颈清扫范围研究进展[J]. 中国医学文摘(耳鼻咽喉科学), 2023, 38(4):138-140.
- Ma T, Wu GX, Zhao RL, et al. Progress in the study of neck dissection range for supraglottic laryngeal cancer[J]. Chinese Medical Abstracts (Otolaryngology), 2023, 38(4):138-140.
- [25] Song L, Heng Y, Hsueh CY, et al. A predictive nomogram for lymph node metastasis in supraglottic laryngeal squamous cell carcinoma[J]. Front Oncol, 2022, 12:786207.
- [26] 高伟, 郭信芳, 孟朝瞰, 等. 中性粒细胞与淋巴细胞比值与血小板计数对声门上型喉癌淋巴结转移的预测价值[J]. 海南医学, 2021, 32(8):986-989.
- Gao W, Wu XF, Meng CT, et al. The predictive value of neutrophil to lymphocyte ratio and platelet count for lymph node metastasis in supraglottic laryngeal cancer[J]. Hainan Medical Journal, 2021, 32(8):986-989.

## 《中国肿瘤》关于伦理审查的要求

根据《世界医学协会赫尔辛基宣言》和我国《涉及人的生物医学研究伦理审查办法》等的相关规定以及国际通行的动物福利和伦理准则,为进一步规范医学领域研究程序,保护研究对象的合法权益,本刊对相关论文的投稿提出如下要求:

(1)当论文的主体是以人为研究对象时(包括前瞻性研究、横断面研究、回顾性研究等),作者应当说明是否经所在单位或地区伦理学委员会的批准,是否取得研究对象或其家属的知情同意,并提供该委员会的批准文件复印件以及研究对象或其家属的知情同意书复印件。除此之外,凡涉及临床试验研究(前瞻性研究),作者原则上均应在 WHO 国际临床试验注册中心 (<https://www.who.int/ictpr/en/>)或中国临床试验注册中心 (<http://www.chictr.org.cn/index.aspx>) 进行注册,并在论文中标注临床试验注册号。

(2)涉及实验动物的研究性论文,需遵守《实验动物管理条例》《实验动物质量管理办法》《善待实验动物指导性意见》的相关规定,并提供该项研究的伦理审查通过证明复印件及相应的动物合格证号。文中需注明所用动物的品种、品系、性别、日龄或月龄、体质量、数量、饲养条件、建模方法和时间、实验起点和终点、处死方法等必要信息。

(3)本刊伦理内容规范书写格式如下:(供参考)

本研究方案经\*\*\*医院伦理委员会(或实验动物伦理委员会)审批(编号:XXXX),在\*\*\* (临床试验注册机构)注册(注册号:XXXX),患者均签署知情同意书(或符合实验室动物管理与使用准则)。