DOI:10.16016/j.2097-0927.202301110

# HIV 感染/AIDS 患者结核感染风险预测模型的建立

陈倩立,郑小燕,陈瑞华,蒋 秋,蒋 红 404000 重庆 万州,重庆大学附属三峡医院感染科

[摘要] **目的** 调查人类免疫缺陷病毒(human immunodeficiency virus,HIV)感染/艾滋病(acquired immune deficiency syndrome,AIDS)患者结核分枝杆菌(tuberculosis,TB)感染情况,并建立其预测模型。**方法** 选择 2017 年 1 月至 2022 年 8 月在我院建档治疗的 HIV 感染/AIDS 患者 816 例进行回顾性队列研究。采用信封法将以 7∶3分为模型组 571 例与验证组 245 例。收集可能引起 HIV 感染/AIDS 患者 TB 感染的相关指标,根据有无 TB 感染将患者分为结核组与非结核组,比较 2 组各因素,以 LASSO 筛选变量后行多因素 Logistic 回归分析,根据多因素分析结果建立列线图模型并进行验证。**结果** 本研究模型组 571 例 HIV 感染/AIDS 患者中 59 例(10.33%)确诊为 TB 感染。经 LASSO 回归筛基础上行多因素 Logistic 回归分析,结果显示:年龄、吸毒史、卡介苗接种史、近期 $CD4^+T$ 细胞水平、结核病接触史、近 1 年咳嗽史是 HIV 感染/AIDS 患者肺结核的独立影响因素($P<0.05$)。基于上述影响因素构建 HIV 感染/AIDS 患者 TB 感染风险预测的列线图模型,ROC 分析结果显示模型组构建 HIV 感染/AIDS 患者 TB 感染风险预测值的曲线下面积(area under the curve,AUC)为 0.823,95%*CI* 为 0.761~0.884;敏感度为 77.97%,特异度为 96.48%,阳性预测值为 71.88%,阴性预测预测值为 97.44%。验证组 ROC 分析 AUC 为 0.817,95%*CI* 为 0.723~0.912;敏感度为 76.92%,特异度为 96.35%,阳性预测值为 71.43%,阴性预测预测值为 97.24%。校准曲线结果显示,模型组与验证组预测模型与理想模型基本拟合,提示模型准确度较好。**结论** HIV 感染/AIDS 患者 TB 感染主要受年龄、吸毒史、卡介苗接种史、近期 $CD4^+T$ 细胞水平、结核病接触史、近 1 年咳嗽史的影响,基于上述因素建立的预测模型预测 HIV 感染/AIDS 患者 TB 感染风险具有较高的准确度与区分度。

[关键词] 人类免疫缺陷病毒;艾滋病;结核分枝杆菌;多因素分析;预测模型
[中图法分类号] R181.32; R512.91; R52 [文献标志码] A

# Establishment of a risk prediction for tuberculosis infection model in HIV-infected/AIDS patients

CHEN Qianli, ZHENG Xiaoyan, CHEN Ruihua, JIANG Qiu, JIANG Hong Department of Infectious Diseases, Three Gorges Hospital Affiliated to Chongqing University, Wanzhou, Chongqing, 404000, China

[**Abstract**] **Objective** To investigate the infectious status of *Mycobacterium tuberculosis* (TB) in human immunodeficiency virus (HIV)-infected/acquired immune deficiency syndrome (AIDS) patients, and to establish a prediction model for infection risk. **Methods** A retrospective cohort study was conducted on 816 HIV-infected/AIDS patients who were registered in our hospital from January 2017 to August 2022. Then with the aid of envelopes, 571 cases were assigned into a testing group and 245 cases into a validation group at a ratio of 7∶3. Relevant indicators that may cause TB infection were collected in these patients. The patients were also divided into tuberculosis group and non-tuberculosis group. The relevant factors were compared between the 2 groups. After screening variables with LASSO, multivariate logistic regression analysis was performed and the obtained factors were used to establish a nomogram model, which was further verified.

[基金项目] 重庆市万州区科卫联合医学科研项目(WZSTC-KW2020028)
[通信作者] 蒋 红, E-mail:1982860828@qq.com

2097-0927/Copyright © 2023, 陆军军医大学(第三军医大学)。这篇开放获取文章遵循 CC BY 许可协议(https://creativecommons.org/licenses/by/4.0/)

**Results** Among the 571 HIV-infected/AIDS patients in the model group, 59 (10.33%) were diagnosed as TB infection. Multivariate logistic regression analysis showed that age, drug abuse history, BCG vaccination history, recent $CD4^+$ T cell level, tuberculosis contact history, and cough history in recent 1 year were independent influencing factors of TB infection in the patients ($P<0.05$). Then a nomograph model for predicting the risk of TB infection was constructed in the HIV-infected/AIDS patients. ROC curve analysis indicated that the area under the curve (AUC) of the predictive value of TB infection risk in HIV infected/AIDS patients was 0.823, and the 95%*CI* was 0.761~0.884; the sensitivity was 77.97%, the specificity was 96.48%, the positive predictive value was 71.88%, and the negative predictive value was 97.44% in the model group. The AUC value of the validation group was 0.817, the 95%*CI* was 0.723~0.912, the sensitivity was 76.92%, the specificity was 96.35%, the positive predictive value was 71.43%, and the negative predictive value was 97.24%. The calibration curve results suggested that the prediction models of the model group and the validation group were basically matched with the ideal model, indicating that the model accuracy is good. **Conclusion** TB infection of HIV-infected/AIDS patients is mainly affected by age, drug abuse history, BCG vaccination history, recent $CD4^+$ T cell level, tuberculosis contact history and cough history in the past 1 year. Our prediction model established based on the above factors has high accuracy and discrimination in predicting TB infection risk of these patients.

[**Key words**] human immunodeficiency virus; acquired immune deficiency syndrome; *mycobacterium tuberculosis*; multivariate analysis; prediction model

Supported by the Joint Scientific Research Project of Medial Science and Health of Wanzhou District, Chongqing (WZSTC-KW2020028). Corresponding author: JIANG Hong, E-mail: 1982860828@qq.com

艾滋病(acquired immune deficiency syndrome, AIDS)是因人类免疫缺陷病毒(human immunodeficiency virus,HIV)感染所致的慢性传染性疾病,该病可对T淋巴细胞造成严重损伤而致患者免疫功能下降,机会感染风险增加[1-2]。结核分枝杆菌(tuberculosis,TB)是HIV感染/AIDS患者最常见的条件致病菌原,过去虽然经过积极的预防措施使TB感染疫情曾一度受到控制,但随着近年HIV感染人数的骤升导致TB感染人数也相应增加[3]。有报道称HIV感染可导致TB感染率增加30倍,TB感染及HIV感染均为重点控制的传染病,两者可相互促进导致患者病情恶化,甚至可致患者死亡[4]。列线图模型为近年常用的预测模型,该模型是在Logistic回归模型基础上发展而来,与Logistic模型相比列线图模型操作简便、易行,更适用于个性化预测[5],因此本研究旨在通过建立HIV感染/AIDS患者TB感染风险预测模型,以期为指导临床对高风险人群进行个性化干预提供参考。

## 1 资料与方法

### 1.1 一般资料

选择2017年1月至2022年8月在我院建档治疗的HIV感染/AIDS患者816例进行回顾性队列研究。纳入标准:①经疾控中心确诊为HIV感染/AIDS患者;②定期随访,并在随访期间定期接受TB感染检查;③在我院行规范化抗HIV感染治疗。排除标准:①拟分析相关资料不全者;②在HIV感染确诊前即有TB感染史的患者。采用信封法将以7∶3分为模型组571例与验证组245例。其中模型组男382例,女189例;年龄32~71岁,平均53.18±11.04岁;WHO临床分期:Ⅰ~Ⅱ期486例,Ⅲ~Ⅳ期85例。验证组男171例,女74例;年龄30~75岁,平均54.65±11.38岁;WHO临床分期:Ⅰ~Ⅱ期211例,Ⅲ~Ⅳ期34例。两组一般资料差异无统计学意义($P>0.05$)。本研究已获我院伦理委员会审核通过(202344)。

### 1.2 方法

1.2.1 指标采集 收集以下可能引起HIV感染/AIDS患者TB感染的相关指标:年龄、性别、居住地、受教育程度、婚姻状况、吸毒史、吸烟史、饮酒史、职业、卡介苗接种史、近1年内咳嗽史、结核病患者接触史、合并其他机会感染情况、近期$CD4^+$T细胞水平、HIV感染确诊时长、商业性行为史、WHO临床分期。其中吸烟指吸食超过100支卷烟或相当数量的烟丝;饮酒指每月有1次以上饮酒超过50 mL。本研究采取由我院负责免疫接种的专业人员对患者双臂部卡介苗疤痕检查以评估患者卡介苗情况,有卡介苗疤痕者代表有卡介苗接种史;近1年咳嗽史指近1年内有连续2周及以上时间出现咳嗽或咳喘。

1.2.2 样本量估计 本研究采用Logsitic回归,根

据样本量不得少于变量数 10~15 倍,本研究拟分析 17 个因素,样本量不得少于 170 例,本研究根据我院实际情况纳入 816 例,满足样本量要求。

1.2.3 诊断标准 TB 感染诊断参照《肺结核诊断标准》[6] 的要求;以病原学检查为基础,结合症状体征、流行病学、影像学等相关检查。菌阳性肺结核:涂片镜检:抗酸杆菌 1~9 条/300 视野为涂片阳性;1~9 条/300 视野为 1+,1~9 条/10 视野为 2+,1~9 条/每视野为 3+,≥10 条/每视野为 4+。痰培养结果:菌落分散生长,占斜面面积不足 1/4 者为 1+;菌落分散生长,占斜面面积≥1/4 但<1/2 者为 2+;菌落密集生长或部分融合,面积>1/2 但<3/4 斜面者为 3+;菌落呈苔样分布,占据斜面面积≥3/4 者为 4+。菌阴性肺结核:经 3 次涂片+1 次痰培养均为阴性,但(a)有典型肺结核临床症状及 X 线影像学表现;(b)抗结核治疗有效;(c)除外其他非结核性肺部疾病;(d)结核菌素皮肤实验呈强阳性,血清抗结核抗体阳性;(e)痰结核分枝杆菌 PCR 检测阳性;(f)肺外组织病理结果证实为结核病变;(g)支气管肺泡灌洗液中出现抗酸杆菌;(h)支气管或肺部组织证实有结核性改变。符合 a~f 中 3 项和/或 g~h 中一项即可确诊。根据有无 TB 感染将患者分为结核组与非结核组,比较 2 组各因素,以 LASSO 筛选变量后行多因素 Logistic 回归分析,根据多因素分析结果建立列线图模型并进行验证。

1.3 统计学方法

采用 SPSS 22.0 统计学软件进行数据处理。计数资料以[$n$(%)]表示,组间比较采用 $\chi^2$ 检验;计量资料以 $\bar{x}\pm s$ 表示,两组间比较采用成组 $t$ 检验。采用 R 4.1.3 软件"glmnet"包进行 LASSO 回归分析以筛选协变量,在此基础上采用多因素 Logistic 回归模型分析 HIV 感染/AIDS 患者肺结核的影响因素。采用 R 4.1.3 软件"rms"包构 HIV 感染/AIDS 患者肺结核风险预测的列线图模型,绘制 ROC 曲线以评价该列线图模型的区分度,采用 Bootstrap 法进行验证,采用 H-L 拟合优度检验、校准曲线对该列线图模型进行评价和校准,绘制决策曲线以评价该列线图模型的临床有效性。以 $P$<0.05 为差异有统计学意义。

## 2 结果

2.1 两组患者各指标比较

本研究模型组 571 例 HIV 感染/AIDS 患者中 59 例(10.33%)确诊为 TB 感染。单因素分析结果显示:两组患者年龄、吸毒史、饮酒史、卡介苗接种史、近 1 年内咳嗽史、结核病患者接触史、近期 $CD4^+T$ 细胞水平差异均具有统计学意义($P$<0.05),见表 1。

**表 1 两组患者各指标比较[$n$(%)]**

| 因素 | TB 感染组($n$=59) | 非 TB 感染组($n$=512) | $\chi^2$ | $P$ |
|---|---|---|---|---|
| 年龄/岁 | | | | |
| ≤60 | 52(88.14) | 339(66.21) | 10.787 | 0.001 |
| >60 | 7(11.86) | 173(33.79) | | |
| 性别 | | | | |
| 男 | 37(62.71) | 345(67.38) | 0.332 | 0.565 |
| 女 | 22(37.29) | 167(32.62) | | |
| 居住地 | | | | |
| 农村 | 19(32.2) | 127(24.8) | 1.158 | 0.282 |
| 城市 | 40(67.8) | 385(75.2) | | |
| 受教育程度 | | | | |
| 高中以下 | 23(38.98) | 236(46.09) | 0.811 | 0.368 |
| 高中及以上 | 36(61.02) | 276(53.91) | | |
| 婚姻状况 | | | | |
| 非已婚 | 24(40.68) | 181(35.35) | 0.441 | 0.506 |
| 已婚 | 35(59.32) | 331(64.65) | | |
| 吸毒史 | | | | |
| 无 | 32(54.24) | 403(78.71) | 16.141 | <0.001 |
| 有 | 27(45.76) | 109(21.29) | | |
| 吸烟史 | | | | |
| 无 | 18(30.51) | 227(44.34) | 3.584 | 0.058 |
| 有 | 41(69.49) | 285(55.66) | | |
| 饮酒史 | | | | |
| 无 | 23(38.98) | 133(25.98) | 3.876 | 0.049 |
| 有 | 36(61.02) | 379(74.02) | | |
| 职业 | | | | |
| 公职或事业单位 | 21(35.59) | 218(42.58) | 0.793 | 0.373 |
| 其他 | 38(64.41) | 294(57.42) | | |
| 卡介苗接种史 | | | | |
| 无 | 18(30.51) | 65(12.7) | 12.117 | <0.001 |
| 有 | 41(69.49) | 447(87.3) | | |
| 近 1 年内咳嗽史 | | | | |
| 无 | 22(37.29) | 291(56.84) | 7.392 | 0.007 |
| 有 | 37(62.71) | 221(43.16) | | |
| 结核病患者接触史 | | | | |
| 无 | 37(62.71) | 432(84.38) | 15.477 | <0.001 |
| 有 | 22(37.29) | 80(15.62) | | |
| 合并其他机会感染情况 | | | | |
| 无 | 35(59.32) | 372(72.66) | 3.966 | 0.046 |
| 有 | 24(40.68) | 140(27.34) | | |
| 近期 $CD4^+T$ 细胞水平/(个/L) | | | | |
| <200 | 33(55.93) | 109(21.29) | 32.153 | <0.001 |
| ≥200 | 26(44.07) | 403(78.71) | | |
| HIV 感染确诊时长/年 | | | | |
| <1 | 34(57.63) | 314(61.33) | 0.169 | 0.681 |
| ≥1 | 25(42.37) | 198(38.67) | | |
| 商业性行为史 | | | | |
| 无 | 28(47.46) | 173(33.79) | 3.755 | 0.053 |
| 有 | 31(52.54) | 339(66.21) | | |
| WHO 临床分期 | | | | |
| Ⅰ~Ⅱ | 45(76.27) | 441(86.13) | 3.32 | 0.068 |
| Ⅲ~Ⅳ | 14(23.73) | 71(13.87) | | |

### 2.2 HIV 感染/AIDS 患者 TB 感染风险因素的筛选

将有无肺结核作为因变量,将表 1 中各指标作为自变量,通过 LASSO 回归模型的三折交叉验证确定最佳惩罚项系数 λ,在图 2 中分别于 λ(0.29)与 λ+SE(0.43)画线,于 λ+SE 处最终筛选出 6 个 LASSO 回归系数非 0 的影响因素为潜在因素分别为年龄、吸毒史、卡介苗接种史、近期 $CD4^+T$ 细胞水平、结核病接触史、近 1 年咳嗽史,见图 1~2。

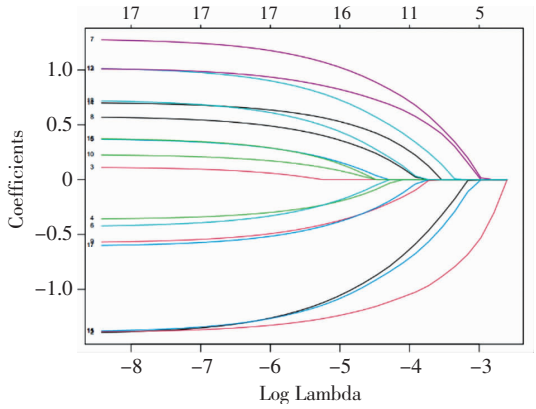



**图 1 LASSO 回归的系数路径**

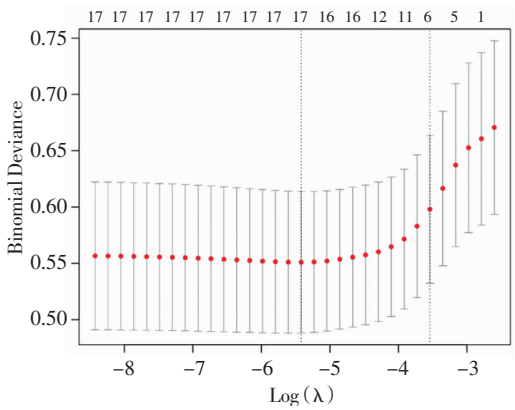



**图 2 LASSO 回归交叉验证结果**

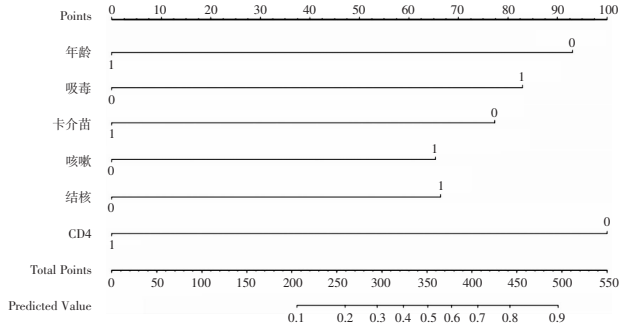
### 2.3 HIV 感染/AIDS 患者 TB 感染的独立影响因素

以有无肺结核作为作为因变量,将 LASSO 回归筛选出的影响因素作为自变量,进行多因素 Logistic 回归分析,结果显示:年龄、吸毒史、卡介苗接种史、近期 $CD4^+T$ 细胞水平、结核病接触史、近 1 年咳嗽史是 HIV 感染/AIDS 患者肺结核的独立影响因素($P<0.05$),见表 2。

### 2.4 列线图模型构建

基于上述影响因素构建 HIV 感染/AIDS 患者 TB 感染风险预测的列线图模型,见图 3。

根据表 3 中多因素分析中各因素 β 值构建预测模型:$Prob=1/[1+\exp(-Y)]$,$Y=-1.146-1.421\times$年龄$+1.211\times$吸毒$-1.233\times$卡介苗$+0.935\times$咳嗽$+1.068\times$结核$-1.470\times$CD4,根据该公式采用 R 4.1.3 软件构建列线图模型。根据多因素分析采用 R 4.1.3 软件的 rms 包构建 HIV 感染/AIDS 患者 TB 感染模型,见图 3。




各因素赋值情况:年龄(岁):≤60=0,>60=1;吸毒史:无=0,有=1;卡介苗接种史:无=0,有=1;近 1 年内咳嗽史:无=0,有=1;结核病患者接触史:无=0,有=1;近期 $CD4^+T$ 细胞水平(个/L):<200=0,≥200=1

**图 3 HIV 感染/AIDS 患者 TB 感染风险的列线图模型**

### 2.5 验证组患者与模型组各资料比较

验证组的 245 例患者共出现 26 例(10.61%)TB 感染者。经统计,验证组各因素分布情况与对照组差异无统计学意义($P>0.05$),见表 3。

**表 2 HIV 感染/AIDS 患者 TB 感染的多因素 Logistic 回归分析**

| 变量 | *B* | *S.E.* | Wald | *P* | Exp(*B*) | 95%*CI* | |
|---|---|---|---|---|---|---|---|
| | | | | | | 下限 | 上限 |
| 年龄 | −1.421 | 0.440 | 10.408 | 0.001 | 0.241 | 0.102 | 0.572 |
| 吸毒史 | 1.211 | 0.317 | 14.623 | <0.001 | 3.356 | 1.804 | 6.242 |
| 卡介苗接种史 | −1.233 | 0.357 | 11.942 | 0.001 | 0.291 | 0.145 | 0.586 |
| 近 1 年内咳嗽史 | 0.935 | 0.317 | 8.708 | 0.003 | 2.547 | 1.369 | 4.739 |
| 结核病患者接触史 | 1.068 | 0.330 | 10.489 | 0.001 | 2.908 | 1.524 | 5.550 |
| 近期 $CD4^+T$ 细胞水平/(个/L) | −1.470 | 0.309 | 22.583 | <0.001 | 0.230 | 0.125 | 0.422 |

各因素赋值情况:年龄(岁):≤60=0,>60=1;吸毒史:无=0,有=1;卡介苗接种史:无=0,有=1;近 1 年内咳嗽史:无=0,有=1;结核病患者接触史:无=0,有=1;近期 $CD4^+T$ 细胞水平(个/L):<200=0,≥200=1

**表 3 验证组患者与模型组各资料比较[$n$(%)]**

| 因素 | 模型组($n$=571) | 验证组($n$=245) | $\chi^2$ | $P$ |
|---|---|---|---|---|
| 年龄/岁 | | | | |
| ≤60 | 391(68.48) | 173(70.61) | 0.366 | 0.545 |
| >60 | 180(31.52) | 72(29.39) | | |
| 性别 | | | | |
| 男 | 382(66.90) | 171(69.80) | 0.658 | 0.417 |
| 女 | 189(33.10) | 74(30.20) | | |
| 居住地 | | | | |
| 农村 | 146(25.57) | 74(30.20) | 1.870 | 0.171 |
| 城市 | 425(74.43) | 171(69.80) | | |
| 受教育程度 | | | | |
| 高中以下 | 259(45.36) | 107(43.67) | 0.197 | 0.657 |
| 高中及以上 | 312(54.64) | 138(56.33) | | |
| 婚姻状况 | | | | |
| 非已婚 | 205(35.90) | 81(33.06) | 0.608 | 0.437 |
| 已婚 | 366(64.10) | 164(66.94) | | |
| 吸毒史 | | | | |
| 无 | 435(76.18) | 177(72.24) | 1.417 | 0.234 |
| 有 | 136(23.82) | 68(27.76) | | |
| 吸烟史 | | | | |
| 无 | 245(42.91) | 102(41.63) | 0.114 | 0.736 |
| 有 | 326(57.09) | 143(58.37) | | |
| 饮酒史 | | | | |
| 无 | 156(27.32) | 53(21.63) | 2.911 | 0.088 |
| 有 | 415(72.68) | 192(78.37) | | |
| 职业 | | | | |
| 公职或事业单位 | 239(41.86) | 95(38.78) | 0.673 | 0.412 |
| 自由职业 | 332(58.14) | 150(61.22) | | |
| 卡介苗接种史 | | | | |
| 无 | 83(14.54) | 47(19.18) | 2.765 | 0.096 |
| 有 | 488(85.46) | 198(80.82) | | |
| 近 1 年内咳嗽史 | | | | |
| 无 | 313(54.82) | 146(59.59) | 1.589 | 0.207 |
| 有 | 258(45.18) | 99(40.41) | | |
| 结核病患者接触史 | | | | |
| 无 | 469(82.14) | 208(84.90) | 0.925 | 0.336 |
| 有 | 102(17.86) | 37(15.10) | | |
| 合并其他机会感染情况 | | | | |
| 无 | 407(71.28) | 180(73.47) | 0.408 | 0.523 |
| 有 | 164(28.72) | 65(26.53) | | |
| 近期 $CD4^+T$ 细胞水平/(个/L) | | | | |
| <200 | 142(24.87) | 57(23.27) | 0.239 | 0.625 |
| ≥200 | 429(75.13) | 188(76.73) | | |
| HIV 感染确诊时长/年 | | | | |
| <1 | 348(60.95) | 166(67.76) | 3.410 | 0.065 |
| ≥1 | 223(39.05) | 78(32.24) | | |
| 商业性行为史 | | | | |
| 无 | 201(35.20) | 85(34.69) | 0.019 | 0.889 |
| 有 | 370(64.80) | 160(65.31) | | |
| WHO 临床分期 | | | | |
| Ⅰ~Ⅱ | 486(85.11) | 211(86.12) | 0.140 | 0.708 |
| Ⅲ~Ⅳ | 85(14.89) | 34(13.88) | | |

### 2.6 模型验证

内部验证:以模型组构建 HIV 感染/AIDS 患者 TB 感染预测模型预测值的曲线下面积(area under the curve,AUC)为 0.823,95%*CI* 为 0.761~0.884;敏感度为 77.97%(46/59),特异度为 96.48%(494/512),阳性预测值为 71.88%(46/64),阴性预测预测值为 97.44%(494/422)(图 4);提示本研究建立的模型区分度较好。再以验证组数据进行外部验证,由验证组构建模型的 AUC 为 0.817,95%*CI* 为 0.723~0.912,敏感度为 76.92%(20/26),特异度为 96.35%(211/219),阳性预测值为 71.43%(20/28),阴性预测预测值为 97.24%(211/217),与内部验证结果相接近,见图 5。

校准曲线结果显示,预测曲线与标准曲线基本拟合,提示模型预测准确度较高,见图 6~7。

### 2.7 列线图模型的临床有效性

决策曲线分析结果显示,当列线图模型预测 HIV 感染/AIDS 患者 TB 感染风险的概率阈值为 0.10~0.70 时,患者的净收益率大于 0,见图 8~9。

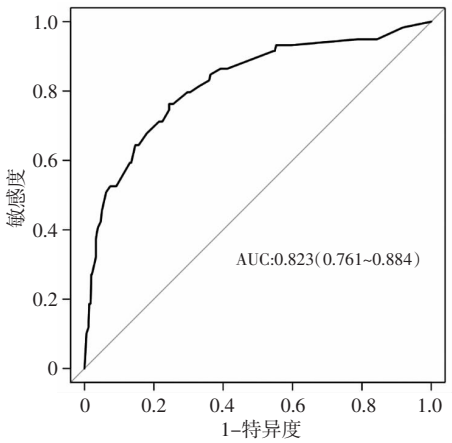



**图 4 模型组 ROC 曲线**

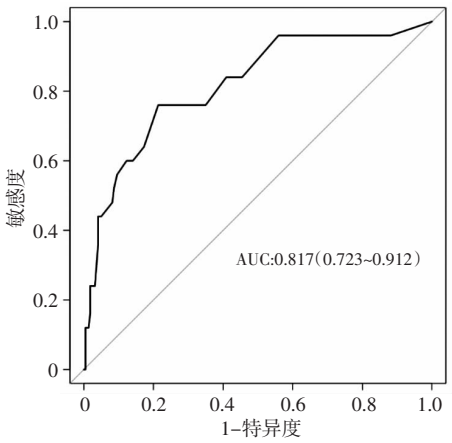



**图 5 验证组 ROC 曲线**

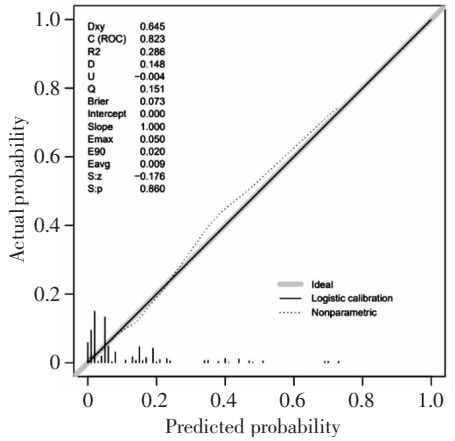



**图 6 模型组校准曲线**

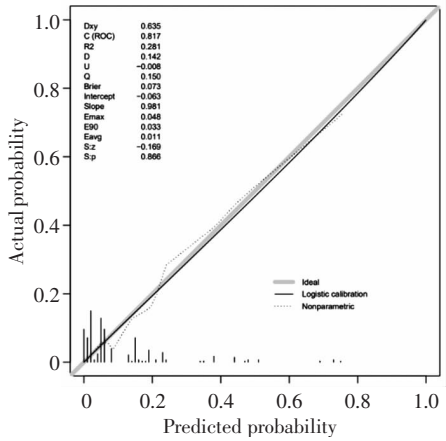



**图 7 验证组校准曲线**

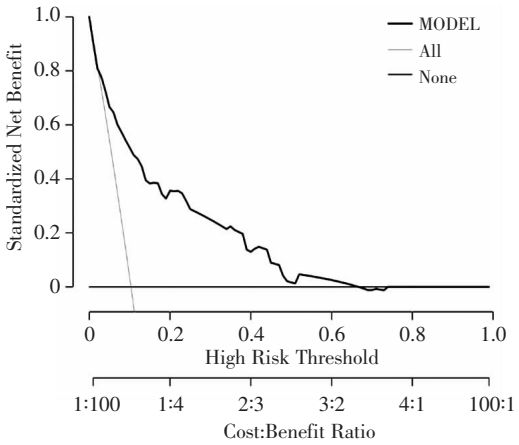



**图 8 模型组决策曲线**

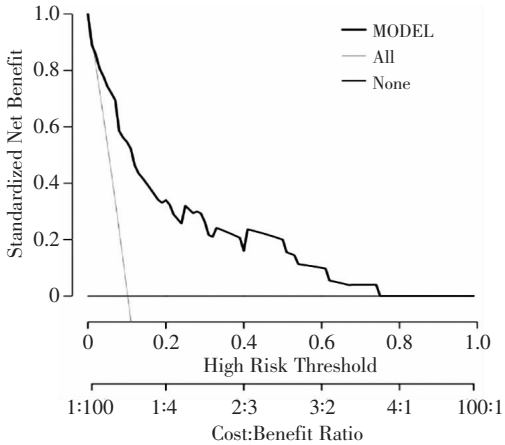



**图 9 验证组决策曲线**

## 3 讨论

HIV 病毒侵袭人体后可对 $CD4^+$T 细胞等免疫细胞造成严重损伤而致患者机体免疫功能骤降,机会感染风险增加[7]。TB 为 HIV 感染/AIDS 患者最常见的机会感染,上述两种疾病均可对机体健康造成严重的影响,当两者同时发病时对于机体的破坏性较任何一种疾病单独发生均更强,因此若能早期预测 HIV 感染/AIDS 患者 TB 感染风险具有重要的意义[8-9]。目前关于 HIV 感染/AIDS 患者 TB 感染影响因素的相关研究虽有报道[10],但多采用单因素分析基础上行多因素 Logistic 回归分析,以此方法筛选出的危险因素可能存在一定的交叉影响而影响临床事件预测的准确性。因此本研究采用 LASSO 回归以筛选潜在因素,LASSO 回归为一种基于惩罚函数的模型,主要通过压缩回归系数、保留子集收缩等优势而解决了复共线性数据的有偏估计问题[11-12]。

本研究将 LASSO 回归筛选出的影响因素作为自变量,进行多因素 Logistic 回归分析,结果显示:年龄、吸毒史、卡介苗接种史、近期 $CD4^+$T 细胞水平、结核病接触史、近 1 年咳嗽史是 HIV 感染/AIDS 患者肺结核的独立影响因素。①从年龄上来看,年龄≤60 岁者为家庭的主要劳动力,常因个人工作、学习等原因需要长期在外地生活,其流动性较年龄>60 岁的人群更大、社会交集更广,接触的人群也较复杂[13]。TB 为机会感染的一种,随着接触人群复杂性的增加,TB 感染风险也明显增加,年龄较大者则交际圈较小,出现 TB 感染的风险更低。另外 TB 的传播主要以飞沫传播为主,有结核病接触史的人群 TB 感染的风险高于无相关接触的人员[14]。②吸毒史对于 HIV 感染/AIDS 患者 TB 感染也有重要的影响,报道称吸毒可严重损害患者体液免疫及细胞免疫,使本就免疫力下降的 HIV 感染/AIDS 患者免疫功能进一步受损,增加 TB 等条件致病菌感染的风险[15-16]。另外,吸毒人员的社交也更为复

杂,接触的人群更广,TB 等机会感染的风险也更大。③卡介苗为牛结核杆菌制剂的一种,研究证实该疫苗可有效增强机体细胞免疫功能而用于 TB 感染的预防[17]。考虑到卡介苗接种多在新生儿时期,距离患者就诊时间已较为久远,患者难以通过记忆回想卡介苗接种情况,本研究采用卡疤评估卡介苗接种情况更具客观性,可信度更高[18]。④$CD4^+T$ 细胞为免疫细胞的一种,其水平的高低可用于反映机体免疫功能。当机体受到 HIV 侵袭时,$CD4^+T$ 细胞可因受到攻击而使其水平大量下降,细胞免疫功能受损而增加 TB 等机会感染的风险[19]。⑤咳嗽为 TB 感染的早期表现之一,若 HIV 感染/AIDS 患者在近 1 年内出现连续 2 周以上的咳嗽则应高度怀疑是否有 TB 感染的可能,必要时需进行痰培养等相关检查以诊断是否出现 TB 感染[20]。

列线图模型实现了 Logistic 回归的可视化,该模型有可视化、计算方便的优势而在近年得到了广泛的应用。列线图模型不仅可根据 HIV 感染/AIDS 患者 TB 感染的影响因素,还可方便地针对每位患者进行个性化预测,计算出 HIV 感染/AIDS 患者 TB 感染的风险,满足患者个性化医疗的需求。本研究还采用内部验证与外部验证的方法证明列线图模型具有较高的准确度与区分度。另外本研究决策曲线分析结果显示,当列线图模型预测 HIV 感染/AIDS 患者 TB 感染风险的概率阈值为 0.10~0.70 时,患者的净收益率大于 0,即列线图模型预测概率在 0.10~0.70 时该模型具有临床价值,医护人员可采用本研究建立的模型对患者进行个性化预测,并根据结果进行健康宣教,制订针对性的干预措施,可提高患者对自身病情的了解并提升治疗依从性。

综上所述,HIV 感染/AIDS 患者 TB 感染主要受年龄、吸毒史、卡介苗接种史、近期 $CD4^+T$ 细胞水平、结核病接触史、近 1 年咳嗽史的影响,基于上述因素建立的预测模型预测 HIV 感染/AIDS 患者 TB 感染风险具有较高的准确度与区分度。

**参考文献:**


[1] WONG K, NGUYEN J, BLAIR L, *et al*. Pathogenesis of human immunodeficiency virus-*Mycobacterium tuberculosis* co-infection[J]. J Clin Med, 2020, 9(11): 3575. DOI: 10.3390/jcm9113575.

[2] HE X B, EDDY J J, JACOBSON K R, *et al*. Enhanced human immunodeficiency virus-1 replication in CD4+ T cells derived from individuals with latent *Mycobacterium tuberculosis* infection[J]. J Infect Dis, 2020, 222(9): 1550-1560. DOI: 10.1093/infdis/jiaa257.

[3] MARTINEZ L, WOLDU H, CHEN C, *et al*. Transmission dynamics in tuberculosis patients with human immunodeficiency virus: a systematic review and meta-analysis of 32 observational studies[J]. Clin Infect Dis, 2021, 73(9): e3446-e3455. DOI: 10.1093/cid/ciaa1146.

[4] FUJITA J. Immune reconstitution inflammatory syndrome in the lung in non-human immunodeficiency virus patients[J]. Respir Investig, 2020, 58(1): 36-44. DOI: 10.1016/j.resinv.2019.11.001.

[5] MOGLAD E H O, AHMED D A O, AWAD AL-KAREEM S M M, *et al*. Prevalence of human immunodeficiency virus among pulmonary tuberculosis patients: a cross-sectional study[J]. Microbiol Immunol, 2020, 64(12): 810-814. DOI: 10.1111/1348-0421.12856.

[6] 中华人民共和国卫生部. 肺结核诊断标准: WS 288—2008[S]. 北京: 人民卫生出版社, 2008.
Ministry of Health of the People's Republic of China. Diagnostic criteria for pulmonary tuberculosis: WS 288—2008[S]. Beijing: People's Medical Publishing House, 2008.

[7] YAGHOUBI A, SALEHABADI S, ABDEAHAD H, *et al*. Tuberculosis, human immunodeficiency viruses and TB/HIV co-infection in pregnant women: a meta-analysis[J]. Clin Epidemiol Glob Health, 2020, 8(4): 1312-1320. DOI: 10.1016/j.cegh.2020.05.003.

[8] DAWIT Z, ABEBE S, DESSU S, *et al*. Incidence and predictors of mortality among children co-infected with tuberculosis and human immunodeficiency virus at public hospitals in Southern Ethiopia[J]. PLoS One, 2021, 16(6): e0253449. DOI: 10.1371/journal.pone.0253449.

[9] PHOSWA W N, ECHE S, KHALIQ O P. The association of tuberculosis mono-infection and tuberculosis-human immunodeficiency virus (TB-HIV) co-infection in the pathogenesis of hypertensive disorders of pregnancy[J]. Curr Hypertens Rep, 2020, 22(12): 104. DOI: 10.1007/s11906-020-01114-5.

[10] 方喆, 黄移生, 王起, 等. 2017—2019 年长沙地区 HIV 感染者/AIDS 患者合并肺结核感染的流行现状及相关危险因素[J]. 职业与健康, 2020, 36(13): 1784-1788. DOI: 10.14164/j.cnki.cn11-5581/r.2019.22.137.
FANG Z, HUANG Y S, WANG Q, *et al*. Prevalence and risk factors of AIDS/HIV patients complicated with pulmonary tuberculosis in Changsha area from 2017 to 2019[J]. Occup Health, 2020, 36(13): 1784-1788. DOI: 10.14164/j.cnki.cn11-5581/r.2019.22.137.

[11] 张韶辉, 苏强, 赵永亮, 等. 基于 LASSO 回归对冠心病相关血脂指标的筛选[J]. 中国综合临床, 2021, 37(2): 148-153. DOI: 10.3760/cma.j.cn101721-20200807-00060.

ZHANG S H, SU Q, ZHAO Y L, *et al*. Screening of lipid parameters in coronary artery disease based on LASSO regression[J]. Clin Med China, 2021, 37(2): 148-153. DOI: 10.3760/cma.j.cn101721-20200807-00060.

[12] DOOLEY K, KAPLAN R, MWELASE N, *et al*. Dolutegravir-based antiretroviral therapy for the treatment of patients co-infected with tuberculosis and HIV: a multicenter, phase 3B, noncomparative, open-label, randomized trial[J]. Clin Infect Dis, 2020, 70(4):549-556. DOI: 10.1093/cid/ciz256.

[13] KERAGALA B S D P, HERATH H M M T B, JANAPRIYA G H D C, *et al*. Coexistence of mycobacterial infections-*Mycobacterium* tuberculosis and *Mycobacterium leprae* - in Sri *Lanka*: a case series[J]. J Med Case Rep, 2020, 14(1): 101. DOI: 10.1186/s13256-020-02413-w.

[14] RAVIMOHAN S, AULD S C, MAENETJE P, *et al*. Lung injury on antiretroviral therapy in adults with human immunodeficiency virus/tuberculosis[J]. Clin Infect Dis, 2020, 70(9): 1845-1854. DOI: 10.1093/cid/ciz560.

[15] SNYMAN J, HWA S H, KRAUSE R, *et al*. Similar antibody responses against severe acute respiratory syndrome coronavirus 2 in individuals living without and with human immunodeficiency virus on antiretroviral therapy during the first South African infection wave[J]. Clin Infect Dis, 2022, 75(1): e249-e256. DOI: 10.1093/cid/ciab758.

[16] WALKER N F, OPONDO C, MEINTJES G, *et al*. Invariant natural killer T-cell dynamics in human immunodeficiency virus-associated tuberculosis[J]. Clin Infect Dis, 2020, 70(9): 1865-1874. DOI: 10.1093/cid/ciz501.

[17] HARIYANTO T I, ROSALIND J, CHRISTIAN K, *et al*. Human immunodeficiency virus and mortality from coronavirus disease 2019: a systematic review and meta-analysis[J]. South Afr J HIV Med, 2021, 22(1): 1220. DOI: 10.4102/sajhivmed.v22i1.1220.

[18] LAKSEMI D A, SUWANTI L T, MUFASIRIN M, *et al*. Opportunistic parasitic infections in patients with human immunodeficiency virus/acquired immunodeficiency syndrome: a review[J]. Vet World, 2020, 13(4): 716-725. DOI: 10.14202/vetworld.2020.716-725.

[19] ALEMU A, YESUF A, GEBREHANNA E, *et al*. Incidence and predictors of extrapulmonary tuberculosis among people living with human immunodeficiency virus in Addis *Ababa*, Ethiopia: a retrospective cohort study[J]. PLoS One, 2020, 15(5): e0232426. DOI: 10.1371/journal.pone.0232426.

[20] SUDHARSHAN S, NAIR N, CURI A, *et al*. Human immunodeficiency virus and intraocular inflammation in the era of highly active anti retroviral therapy-An update[J]. Indian J Ophthalmol, 2020, 68(9): 1787. DOI: 10.4103/ijo.ijo_1248_20.

(**收稿**:2023-01-28;**修回**:2023-05-03)

(编辑 刘秋池)