



安徽医科大学学报  
Acta Universitatis Medicinalis Anhui  
ISSN 1000-1492, CN 34-1065/R

## 《安徽医科大学学报》网络首发论文

题目：肺结核与社区获得性肺炎血液鉴别诊断模型的构建与验证  
作者：苏小兰，张瑞，杨泽青，王新蕊，毕梓瑜，刘念，韩云艳，任琦  
收稿日期：2026-02-09  
网络首发日期：2026-04-03  
引用格式：苏小兰，张瑞，杨泽青，王新蕊，毕梓瑜，刘念，韩云艳，任琦. 肺结核与社区获得性肺炎血液鉴别诊断模型的构建与验证[J/OL]. 安徽医科大学学报. <https://link.cnki.net/urlid/34.1065.r.20260403.1136.008>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

## 肺结核与社区获得性肺炎血液鉴别诊断模型的构建与验证

苏小兰<sup>1\*</sup>, 张瑞<sup>1\*</sup>, 杨泽青<sup>1</sup>, 王新蕊<sup>3</sup>, 毕梓瑜<sup>1</sup>, 刘念<sup>1</sup>, 韩云艳<sup>2</sup>, 任琦<sup>1\*</sup>

(<sup>1</sup>华北理工大学公共卫生学院, 唐山 063210; <sup>2</sup>唐山市第七医院检验科, 唐山

063021; <sup>3</sup>河北省石家庄市长安区疾病预防控制中心检验科, 石家庄 050011)

**摘要 目的** 构建基于常规血液参数的肺结核与社区获得性肺炎鉴别诊断模型并验证。

**方法** 回顾性纳入唐山市第七医院肺结核患者 642 例, 社区活动性肺炎患者 503 例, 按 7:3 比例随机分为训练集与内部验证集, 独立纳入中国人民解放军第 981 医院的 218 例患者作为外部验证集。采用 Boruta 算法与递归特征消除算法筛选 82 个血液参数变量、性别及年龄, 构建多因素 Logistic 回归模型, 并通过受试者工作曲线 (ROC)、校准曲线及决策曲线分析评估性能。**结果** 最终模型纳入 8 项指标, 分别为尿酸、尿素氮、碱性磷酸酶、单胺氧化酶、丙氨酸氨基转移酶、谷胱甘肽和中性粒细胞百分比及年龄, 训练集、内部验证集和外部验证集的 ROC 曲线下面积 (AUC) 分别为 0.800 (95%CI: 0.769~0.830)、0.787 (95%CI: 0.738~0.836) 和 0.736 (95%CI: 0.667~0.835)。训练集校准度良好, 决策曲线分析显示在 10%~80% 阈值概率范围内具有临床净收益。**结论** 该模型具有良好的鉴别能力和临床实用性, 可为基层医疗机构提供有效的早期筛查工具。

**关键词** 肺结核; 社区获得性肺炎; 血液参数; 鉴别模型; 外部验证; 多因素 Logistic 回归模型

中图分类号 R 183.3

文献标志码 A

### Development and validation of a blood-based differential diagnosis model for pulmonary tuberculosis and community-acquired pneumonia

Su Xiaolan<sup>1</sup>, Zhang Rui<sup>1</sup>, Yang Zeqing<sup>1</sup>, Wang Xinrui<sup>3</sup>, Bi Ziyu<sup>1</sup>, Liu Nian<sup>1</sup>, Han Yunyan<sup>2</sup>, Ren Qi<sup>1</sup>

(<sup>1</sup>School of Public Health, North China University of Science and Technology, Tangshan 063210; <sup>2</sup>Department of Clinical Laboratory, Tangshan Seventh Hospital, Tangshan 063021; <sup>3</sup>Department of Clinical Laboratory, Chang'an District Center for Disease Control and Prevention, Shijiazhuang 050011)

**Abstract Objective** To develop and validate a diagnostic prediction model for differentiating

2026-02-09 接收

基金项目: 河北省医学科学研究课题计划项目 (编号: 20260674)

作者简介: 苏小兰, 女, 硕士研究生;

任琦, 女, 副教授, 硕士生导师, 通信作者, E-mail:renqi@ncst.edu.cn

\*对本文具有同等贡献

pulmonary tuberculosis from community-acquired pneumonia based on routine blood parameters. **Methods** A total of 642 patients with pulmonary tuberculosis and 503 patients with community-acquired pneumonia were retrospectively enrolled from the Seventh Hospital of Tangshan. They were randomly divided into a training set and an internal validation set at a 7:3 ratio. Additionally, 218 patients from the 981st Hospital were independently included as an external validation set. The Boruta algorithm and recursive feature elimination were employed to select predictors from 82 blood parameters, sex, and age. A multivariate Logistic regression model was established and evaluated using receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis. **Results** The final model incorporated eight predictors, namely uric acid, urea nitrogen, alkaline phosphatase, monoamine oxidase, alanine aminotransferase, glutathione, neutrophil percentage, and age. The model achieved areas under the curve (AUCs) of 0.800 (95%CI: 0.769-0.830), 0.787 (95%CI: 0.738-0.836), and 0.736 (95%CI: 0.667-0.835) in the training, internal validation, and external validation sets, respectively. The model demonstrated good calibration, and decision curve analysis showed clinical net benefit within the threshold probability range of 10%-80%. **Conclusion** The developed model exhibits good discriminative ability and clinical utility, serving as an effective early screening tool for primary health care institutions.

**Key words** pulmonary tuberculosis; community-acquired pneumonia; blood parameters; diagnostic model; external validation; multivariate Logistic regression model

**Fund program** Medical Science Research Project of Hebei Provincial Health Commission (No. 20260674).

**Corresponding Author** Qi Ren, E-mail: renqi@ncst.edu.cn

结核病 (tuberculosis, TB) 是全球最主要的致死性传染病之一, 近 90% 的病例集中在包括中国在内的高负担国家<sup>[1-4]</sup>。肺结核 (pulmonary tuberculosis, PTB) 在临床上常与社区获得性肺炎 (community-acquired pneumonia, CAP) 表现相似, 导致鉴别诊断困难, 约 48% 的肺结核病例曾被误诊为肺炎而接受经验性抗感染治疗<sup>[5]</sup>。目前, 传统诊断方法 (如分子检测、培养) 存在周期长、成本高、诊断准确性不足等局限<sup>[6-8]</sup>。鉴于 PTB 由 T 淋巴细胞介导, 而 CAP 则以中性粒细胞浸润为主, 两者在免疫病理上存在根本差异。这些差异可反映为不同的血液学特征<sup>[9-10]</sup>, 因此利用易获取的常规血液指标进行鉴别具有潜在优势。然而, 现有基于血液参数的诊断模型主要集中于区分肺结核与非结核性肺病<sup>[6]</sup>, 专门针对肺结核与 CAP

仍较为有限。因此,本研究旨在构建并验证基于常规血液参数的鉴别诊断模型,以期为早期、精准区分肺结核与 CAP 提供实用工具。

## 1 材料与方法

### 1.1 研究对象

本研究为一项回顾性、多中心研究,数据来源于唐山市第七医院和中国人民解放军第 981 医院(以下简称“第 981 医院”),纳入了 2023 年 1 月至 2024 年 12 月期间于医院治疗的肺结核患者与 CAP 患者。肺结核的诊断严格依据《肺结核诊断标准 WS288-2017》<sup>[11]</sup>。CAP 的诊断则参照《成人社区获得性肺炎基层诊疗指南(2018 年)》<sup>[12]</sup>进行。纳入标准为:① 年龄 $\geq 18$  周岁;② 符合上述 PTB 或 CAP 诊断标准。排除标准为:① 仅依据 X 线胸片诊断为 PTB 或 CAP,缺乏病原学或更精确影像学(如 CT)证据的患者;② 关键基线资料或本研究所需的血液学指标存在缺失者;③ 合并有糖尿病、HIV 感染/艾滋病、因器官移植/自身免疫病等正在接受系统性免疫抑制治疗的患者。样本量估算采用医学统计中广泛应用的每变量事件数(events per variable, EPV)法则。为确保多因素回归分析结果的稳健性,本研究遵循 10 EPV 的经验准则。本研究方案已获得医院的伦理委员会审查批准(伦理批号:2024238、15、20241204),所有参与研究的患者均签署了书面知情同意书。由经过统一培训的研究人员,通过各医院的电子病历信息系统,系统性收集患者的临床资料。采集的数据包括:人口统计学特征、临床特征、实验室指标。

### 1.2 统计学处理

使用 Excel 2021 进行数据整理与数据库构建。所有统计分析均在 R 软件(版本 4.4.2)中完成。

基线资料描述:符合正态分布的连续变量以均数 $\pm$ 标准差  $\bar{x} \pm s$  表示,组间比较采用独立样本  $t$  检验;非正态分布的连续变量以  $M(P_{25}, P_{75})$  表示,组间比较采用 Mann-Whitney  $U$  检验。分类变量以  $n(\%)$  表示,组间比较采用卡方检验或 Fisher 精确检验(当理论频数 $<1$  或  $n < 40$ )。所有检验均为双侧,  $P < 0.05$  为差异有统计学意义。

数据集划分:将来自唐山市第七医院的患者数据按 7:3 的比例随机划分为训练集与内部验证集。将来自第 981 医院的患者数据作为独立的外部验证集,用于评估模型的泛化能力。

特征筛选:为构建稳健的预测模型,本研究采用两种互补的特征筛选策略。① Boruta 算法:该算法通过创建随机化的“影子特征”,经过 100 次迭代,识别出重要性持续高于影子特征的真实预测变量。② 递归特征消除结合随机森林:以随机森林的基尼重要性作为特征排序

依据，通过 5 折交叉验证确定最优特征数量为 10。最终，取两种方法共同筛选出的特征交集，用于后续多因素 Logistic 回归分析。

模型构建与评估：基于初步筛选出的变量，采用多因素 Logistic 回归构建鉴别诊断模型，并通过后向逐步回归法以最小化赤池信息准则 (akaike information criterion, AIC) 值为标准。候选变量经初步筛选后，进入逐步回归流程。最终模型保留整体 AIC 最小的变量组合。在建模前，对所有自变量进行多重共线性诊断，计算方差膨胀因子 (variance inflation factor, VIF)，以  $VIF < 10$  为标准，确保变量间不存在严重共线性。模型性能通过以下方面综合评估：通过受试者工作特征曲线下面积 (area under the receiver operating characteristic curve, AUC) 评估模型区分度；校准曲线评估预测概率与实际概率的一致性；临床决策曲线 (decision curve analysis, DCA) 评估模型的临床净获益。

## 2 结果

### 2.1 研究对象特征

本研究共纳入唐山市第七医院患者 1 145 例，其中 PTB 患者 642 例，CAP 患者 503 例，按 7 : 3 比例随机分为训练集 ( $n=802$ ) 与内部验证集 ( $n=343$ )，两组基线特征比较见表 1。

比较分析显示，在人口学特征方面，训练集和验证集 PTB 组患者的中位年龄明显低于社区获得性肺炎 (CAP) 组 ( $P < 0.001$ )；训练集中 PTB 组男性比例高于 CAP 组 ( $P = 0.007$ )，但验证集中该差异不显著 ( $P = 0.313$ )；两组文化程度分布差异显著 ( $P < 0.001$ )，训练集和验证集 PTB 组中大专及以上学历者比例均高于 CAP 组；内部验证集中，CAP 组城镇职工医疗保险比例高于 PTB 组 ( $P = 0.014$ )。在临床特征方面，训练集中 PTB 组非初治患者比例及中位住院时长均高于 CAP 组 ( $P < 0.001$ )，而两组在咳痰、发热等临床症状方面差异无统计学意义 ( $P > 0.05$ )。

表 1 训练集与内部验证集中 PTB 与 CAP 患者基线特征比较 [ $n$  (%),  $M$  ( $P_{25}$ ,  $P_{75}$ )]

Tab.1 Comparison of baseline characteristics between PTB and CAP patients in the training set and internal validation set [ $n$  (%),  $M$  ( $P_{25}$ ,  $P_{75}$ )]

Variable	Training set			Internal validation set		
	PTB $n=446$	CAP $n=356$	$P$ value	PTB $n=196$	CAP $n=147$	$P$ value
Age	59 (49, 69)	68 (57, 74)	<0.001	59 (50, 68)	70 (58, 77)	<0.001
Gender			0.007			0.313
Male	319 (71.52)	222 (62.36)		143 (72.96)	99 (67.35)	

Female	127 (28.48)	134 (37.64)		53 (27.04)	48 (32.65)	
Marital status			0.053			0.184
Single	25 (5.61)	8 (2.25)		6 (3.06)	2 (1.36)	
Married				176		
	369 (82.74)	301 (84.55)		(89.80)	127 (86.39)	
Divorced/Widowed	52 (11.65)	47 (13.20)		14 (7.14)	18 (12.25)	
Education level			<0.001			<0.001
Below high school	134 (30.04)	117 (32.87)		57 (29.08)	48 (32.65)	
High school graduate	172 (38.57)	221 (62.08)		76 (38.78)	93 (63.27)	
Post-secondary education	140 (31.39)	18 (5.05)		63 (32.14)	6 (4.08)	
Registered residence			0.526			
Urban	271 (60.76)	225 (63.20)		132	96 (65.31)	0.779
				(67.35)		
Rural	175 (39.24)	131 (36.80)		64 (32.65)	51 (34.69)	
Health insurance type			0.262			0.014
Urban resident	298 (66.82)	231 (64.89)		131	80 (54.42)	
				(66.84)		
Urban employee	137 (30.72)	121 (33.99)		60 (30.61)	66 (44.90)	
Self-pay	11 (2.46)	4 (1.12)		5 (2.55)	1 (0.68)	
Length of hospital stay (days)	27 (18, 41)	21 (8, 43)	<0.001	25 (17, 36)	23 (8, 43)	0.160
Sputum			0.506			0.223
Yes	364 (81.61)	283 (79.49)		161	112 (76.19)	
				(82.14)		
No	82 (18.39)	73 (20.51)		35 (17.86)	35 (23.81)	
Cough			0.976			0.987
Yes	267 (59.87)	214 (60.11)		120	89 (60.54)	
				(61.22)		
No	179 (40.13)	142 (39.89)		76 (38.78)	58 (39.46)	
Fever <sup>a</sup>			0.506			
Yes	364 (81.61)	283 (79.49)		161	112 (76.19)	0.223
				(82.14)		
No	82 (18.39)	73 (20.51)		35 (17.86)	35 (23.81)	

Initial treatment status			<0.001			0.145
Yes			189			
	423 (94.84)	355 (99.72)	(96.43)	146 (99.32)		
No	23 (5.16)	1 (0.28)	7 (3.57)	1 (0.68)		

a: Fever was defined as an axillary temperature  $\geq 37.3$  °C.

## 2.2 血液参数鉴别肺结核与肺炎指标筛选

为从初选的 84 个候选变量（包括 82 个血液学参数、性别和年龄）中，筛选出最具鉴别力的预测变量，本研究联合采用了 Boruta 算法与递归特征消除两种方法。Boruta 算法基于随机森林与影子特征比较，迭代 100 次后，共识别出 22 个“可接受”的重要特征与 10 个“可疑”特征（图 1A）。递归特征消除采用随机森林与 5 折交叉验证，根据变量重要性排序，筛选出排名前 10 的特征，依次为：尿酸（uric Acid, UA）、尿素氮（blood urea nitrogen, BUN）、碱性磷酸酶（alkaline phosphatase, ALP）、单胺氧化酶（monoamine oxidase, MAO）、丙氨酸氨基转移酶（alanine aminotransferase, ALT）、谷胱甘肽（glutathione, GSH）、直接胆红素（direct bilirubin, DBIL）、淋巴细胞百分比（lymphocyte percentage, LYM%）、中性粒细胞百分比（neutrophil percentage, NEU%）与年龄（图 1B）。为提升模型的稳健性与泛化能力，取两种方法筛选结果的交集，将共同选出的 10 个变量纳入后续的多因素 Logistic 回归分析。

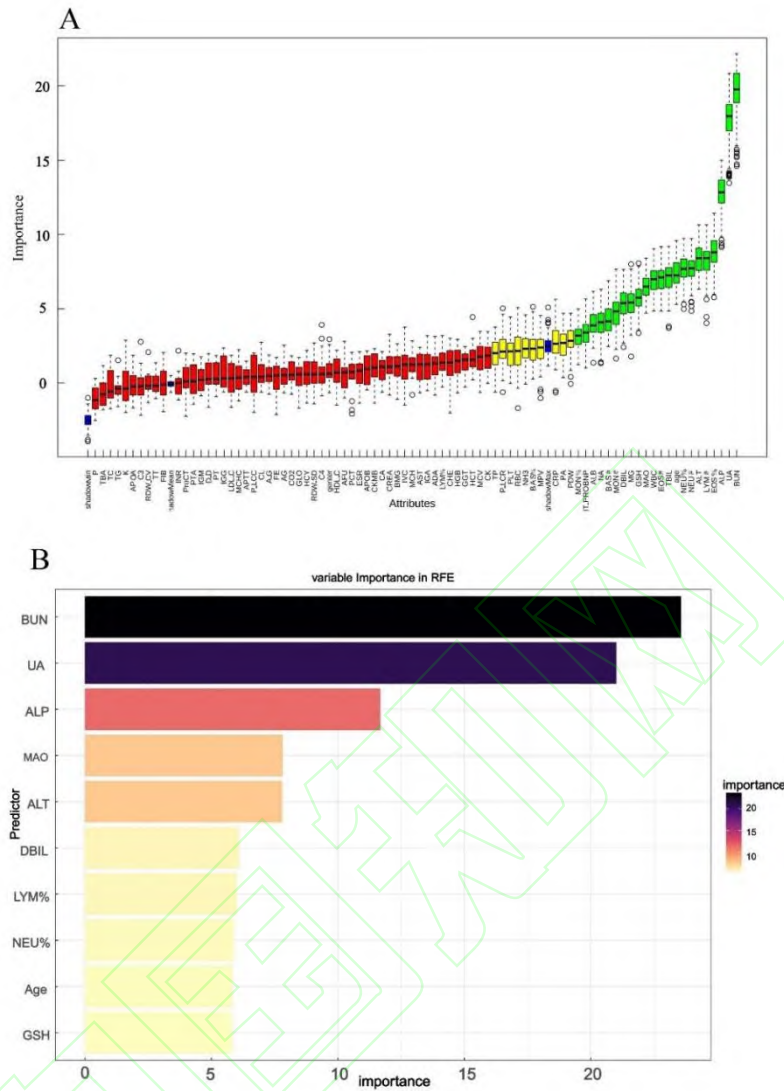


图 1 基于 Boruta 和递归特征消除 (RFE) 的血液参数特征筛选结果

**Fig. 1 Results of blood parameter feature selection based on Boruta and recursive feature elimination (RFE)**

A: Green indicates acceptable features, yellow indicates tentative features, and red indicates unimportant features; acceptable features are those whose importance is significantly higher than the maximum importance score of shadow features, while tentative features require further validation; B: Important variables are ranked; AG: anion gap; CL: chloride; CO<sub>2</sub>: carbon dioxide content; K: potassium; NA: sodium; CA: calcium; MG: magnesium; P: phosphorus; NT-PROBNP: n-terminal pro-b-type natriuretic peptide; ProCT: procalcitonin; CREA: creatinine; CK: creatine kinase; CKMB: creatine kinase-mb; HCY: homocysteine; BMG: beta-2-microglobulin; ALB: albumin; GLO: globulin; A/G: albumin-to-globulin ratio; TP: total protein; PA: prealbumin; TBA:

total bile acids; TBIL: total bilirubin; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; CHE: cholinesterase; NH<sub>3</sub>: ammonia; ADA: adenosine deaminase; AFU: alpha-l-fucosidase; IVC: indirect bilirubin; TC: total cholesterol; TG: triglycerides; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; APOA: apolipoprotein a1; APOB: apolipoprotein b; WBC: white blood cell count; RBC: red blood cell count; HGB: hemoglobin; HCT: hematocrit; MCV: mean corpuscular volume; MCH: mean corpuscular hemoglobin; MCHC: mean corpuscular hemoglobin concentration; RDW-CV: red cell distribution width-coefficient of variation; RDW-SD: red cell distribution width-standard deviation; PLT: platelet count; MPV: mean platelet volume; PDW: platelet distribution width; PCT: plateletcrit; P-LCR: platelet-large cell ratio; P-LCC: platelet-large cell count; NEU#: neutrophil count; LYM#: lymphocyte count; MON#: monocyte count; EOS#: eosinophil count; BAS#: basophil count; MON%: monocyte percentage; EOS%: eosinophil percentage; BAS%: basophil percentage; CRP: c-reactive protein; ESR: erythrocyte sedimentation rate; C3: complement C3; C4: complement C4; IGA: immunoglobulin A; IGG: immunoglobulin G; IGM: immunoglobulin M; FE: iron; PT: prothrombin time; INR: international normalized ratio; PTA: prothrombin activity; APTT: activated partial thromboplastin time; TT: thrombin time; FIB: fibrinogen; D-D: d-dimer.

### 2.3 血液参数鉴别 PTB 与 CAP 的构建与验证

基于前述变量初步筛选结果,采用多因素 Logistic 回归分析最终筛选出 8 个对鉴别 PTB 与 CAP 最具价值的独立预测变量并构建模型。共线性诊断显示所有变量 VIF 均小于 5,表明不存在多重共线性问题。其回归分析结果见表 2, UA 及 ALP 与 PTB 风险呈正相关 ( $OR=1.004$ 、 $1.014$ , 均  $P<0.001$ ); 而 BUN、MAO、NEU%及 GSH 与 PTB 风险呈负相关 ( $OR=0.863$ 、 $0.921$ 、 $0.917$ 、 $0.982$ , 均  $P<0.05$ )。

进一步通过 ROC 曲线、校准曲线与 DCA 评估模型性能。如图 2A 所示,模型在训练集的 AUC 为 0.800 (95% CI: 0.769~0.830), 最佳截断值为 0.525, 灵敏度为 68.5%, 特异度为 78.3%。内部验证集中的 AUC 为 0.787 (95% CI: 0.738~0.836), 最佳截断值为 0.587, 灵敏度为 71.4%, 特异度为 74.0%, 显示出良好的区分能力。校准度评估采用 Hosmer-Lemeshow 拟合优度检验(图 2B), 模型在训练集 ( $\chi^2=7.018$ ,  $P=0.535$ ) 与验证集 ( $\chi^2=5.016$ ,  $P=0.756$ ) 中均具有良好的校准性能。表明模型预测概率与实际观测概率之间无显著差异, 模型校准度良好。DCA 表明在 10%~80%的阈值概率范围内模型具有临床净收益(图 2C)。

为验证变量组合的合理性,构建 4 个嵌套模型进行性能比较:模型一:年龄+BUN+UA;模型二:模型一+ALP+ALT+MAO;模型三:模型二+NEU%;模型四:模型三+GSH。结果显示,包含全部 8 个指标的完整模型(模型四)在训练集与验证集中均表现最优。模型二、三、四的 AUC 均显著高于模型一(DeLong 检验均  $P<0.05$ ),其中模型四的 AUC 最高(训练集 0.810,验证集 0.786)。校准度采用 Hosmer-Lemeshow 拟合优度检验进行评估。模型一在训练集中校准不佳( $\chi^2=17.433, P=0.026$ ),验证集校准良好( $\chi^2=2.945, P=0.938$ )。模型二(训练集  $\chi^2=11.992, P=0.152$ ;验证集  $\chi^2=4.874, P=0.771$ )、模型三(训练集  $\chi^2=8.661, P=0.372$ ;验证集  $\chi^2=9.253, P=0.321$ )和模型四(训练集  $\chi^2=7.018, P=0.535$ ;验证集  $\chi^2=5.016, P=0.756$ )在训练集与验证集中均校准良好,综合评估,模型四为最优模型,且在多数阈值概率下净收益最大(图 3)。

**表 2 PTB 与 CAP 多因素 Logistic 回归分析结果**

**Tab. 2 The results of the multivariate Logistic regression analysis of PTB and CAP**

Variable	$\beta$	SE	Z value	P value	OR (95% CI)
Intercept	1.816	0.55 164	3.291	<0.001	6.145 (2.109-18.38)
Age (years)	-0.024	0.00 554	-4.346	<0.001	0.976 (0.965-0.986)
UA ( $\mu\text{mol/L}$ )	0.004	0.00 052	7.854	<0.001	1.004 (1.003-1.005)
BUN (mmol/L)	-0.146	0.0 321	-4.561	<0.001	0.863 (0.807-0.914)
ALP (U/L)	0.014	0.00317	4.492	<0.001	1.014 (1.008-1.020)
NEU% (%)	-0.086	0.03 202	-2.701	0.007	0.917 (0.859-0.975)
MAO (U/L)	-0.082	0.02 765	-2.951	0.003	0.921 (0.872-0.972)
GSH (mg/L)	-0.017	0.00 484	-3.605	<0.001	0.982 (0.973-0.991)
ALT (U/L)	-0.004	0.00 253	-1.558	0.119	0.996 (0.990-1.000)

OR for continuous variables are per 1-unit increase.

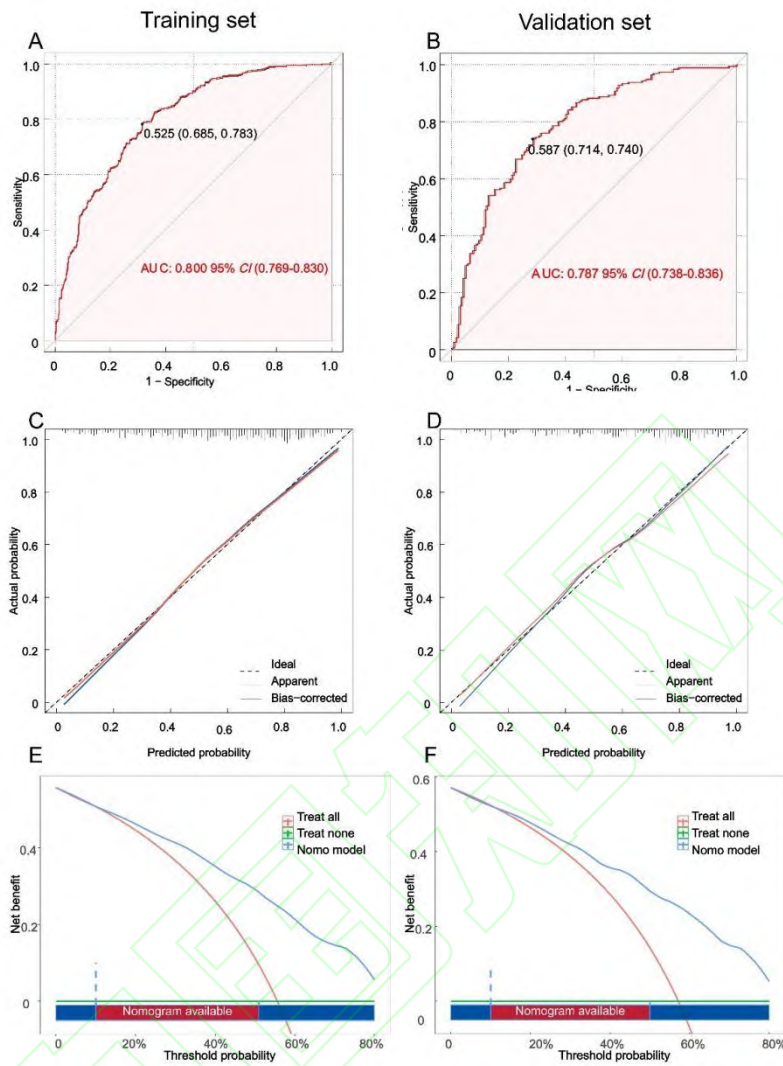


图 2 基于血液参数的多因素 Logistic 回归模型鉴别肺结核与肺炎的效能分析

Fig. 2 Performance of the multivariate Logistic regression model based on blood parameters for differentiating pulmonary tuberculosis from pneumonia

A, B: The receiver operating characteristic curve; C, D: The calibration curve; E, F: The clinical decision curve.

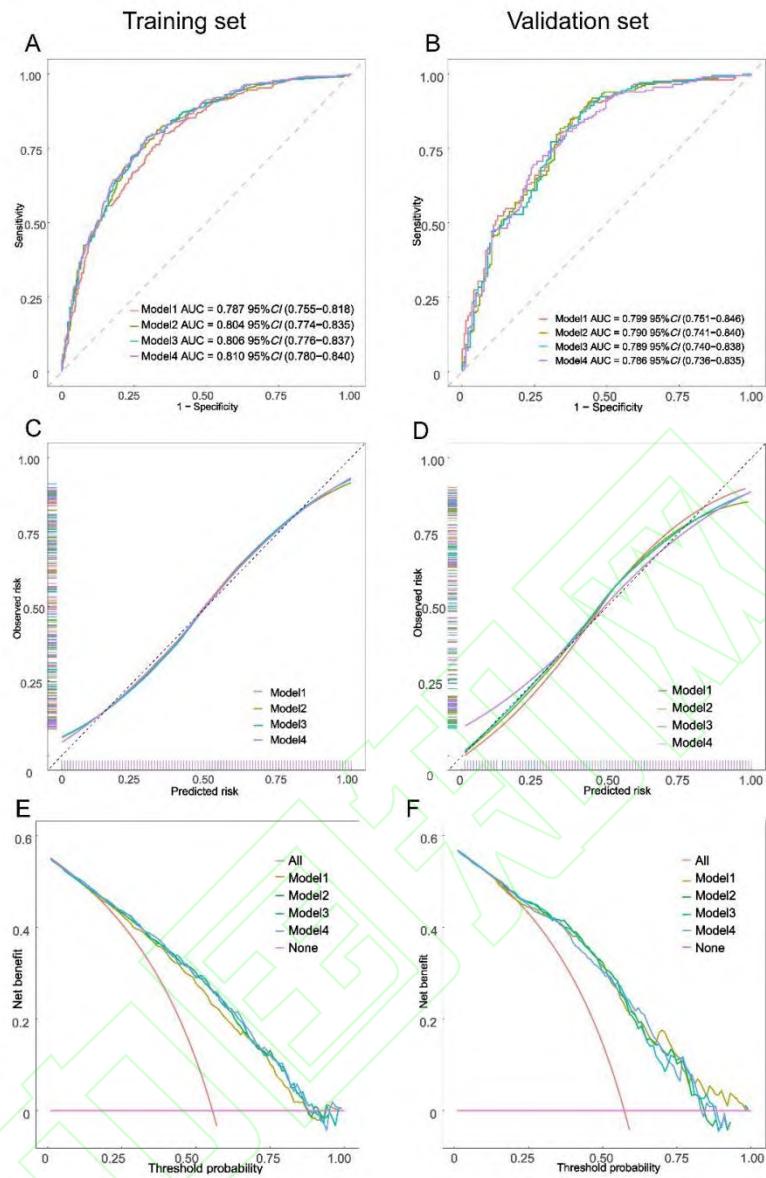


图 3 不同血液参数模型鉴别肺结核与肺炎的效能分析

**Fig. 3 Performance analysis of different blood parameter models for differentiating pulmonary tuberculosis from pneumonia**

A, B: The receiver operating characteristic curve; C, D: The calibration curve; E, F: The clinical decision curve.

## 2.4 外部验证结果

为评估模型的泛化能力，本研究采用来自中国人民解放军第 981 医院的 218 例患者数据进行外部验证（PTB 组 129 例，CAP 组 89 例）。PTB 组的中位年龄为 58 岁，男性占 78.29%，与训练集相比差异无统计学意义，外部验证集患者住院时长中位数（18 天）比训

训练集短（27天），咳嗽占比（78.29%）高于训练集（59.87%）。CAP组中位年龄为58岁，与训练集比差异有统计学意义，表3所示。模型的区分效能良好，其AUC值为0.736（95% CI: 0.667~0.805）模型预测概率的最佳截断值为0.594，灵敏度为68.9%，特异度为72.9%（图4A）

Hosmer-Lemeshow拟合优度检验结果显示模型预测概率与实际观测概率之间无显著差异（ $\chi^2 = 10.724$ ,  $P = 0.218$ ），表明模型的预测概率与实际情况具有良好的一致性（图4B）。决策曲线分析显示，在阈值概率低于40%时，模型的净收益低于“全部治疗”策略（图4C）。在40%–80%的阈值概率区间内，模型的净收益高于“全部治疗”与“均不治疗”两种极端策略。模型的阳性预测值为77.1%，阴性预测值为63.9%。

**表3 训练集与外部验证集中PTB与CAP患者基线特征比较** [ $n$  (%),  $M$  ( $P_{25}$ ,  $P_{75}$ )]

**Tab.3 Comparison of baseline characteristics between PTB and CAP patients in the training set and external validation set** [ $n$  (%),  $M$  ( $P_{25}$ ,  $P_{75}$ )]

Variable	PTB			CAP		
	Training set $n=446$	External validation set $n=129$	$P$ value	Training set $n=356$	External validation set $n=89$	$P$ value
Age	59 (49,69)	58 (52,66)	0.181	68 (57,74)	58 (42,68)	<0.001
Gender			0.158			0.031
Male				222 (62.36)	67 (75.28)	
Female	319 (71.52)	101 (78.29)		134 (37.64)	22 (24.72)	
Marital status	127 (28.48)	28 (21.71)	0.176			0.334
Single	25 (5.60)	11 (8.52)		8 (2.25)	3 (3.37)	
Married				301 (84.55)	79 (88.76)	
Divorced/Widowed	369 (82.74)	109 (84.50)				
d	52 (11.66)	9 (6.98)		47 (13.20)	7 (7.87)	
Education level			0.001			0.009
Below high school				117 (32.87)	40 (44.94)	
High school graduate	134 (30.04)	58 (44.96)		221 (62.08)	40 (44.94)	
Post-secondary education	172 (38.57)	48 (37.21)				
d	140 (31.39)	23 (17.83)		18 (5.05)	9 (10.12)	
Registered residence			0.110			0.106

Urban				225		
	271 (60.76)	89 (68.99)		(63.20)	65 (73.03)	
Rural				131		
	175 (39.24)	40 (31.01)		(36.80)	24 (26.97)	
Health insurance type			0.956			0.807
Urban resident				231		
	298 (66.82)	88 (68.22)		(64.89)	61 (68.54)	
Urban employee				121		
	137 (30.72)	38 (29.45)		(33.99)	27 (30.34)	
Self-pay	11 (2.46)	3 (2.33)		4 (1.12)	1 (1.12)	
Length of hospital stay(days)			<0.00			
	27 (18,41)	18 (11,27)	1	21 (8,43)	19 (12,25)	0.145
Initial treatment status			0.998			0.103
				355		
Yes	423 (94.84)	123 (95.35)		(99.72)	87 (97.75)	
No	23 (5.16)	6 (4.65)		1 (0.28)	2 (2.25)	
			<0.00			
Cough			1			0.011
				214		
Yes	267 (59.87)	101 (78.29)		(60.11)	67 (75.28)	
				142		
No	179 (40.13)	28 (21.71)		(39.89)	22 (24.72)	
Fever <sup>a</sup>			0.195			0.386
				283		
Yes	364 (81.61)	98 (75.97)		(79.49)	75 (84.27)	
No	82 (18.39)	31 (24.03)		73 (20.51)	14 (15.73)	
			<0.00			
Sputum			1			0.039
				283		
Yes	364 (81.61)	79 (61.24)		(79.49)	61 (68.54)	
No	82 (18.39)	50 (38.76)		73 (20.51)	28 (31.46)	

A: Fever was defined as an axillary temperature  $\geq 37.3$  °C.

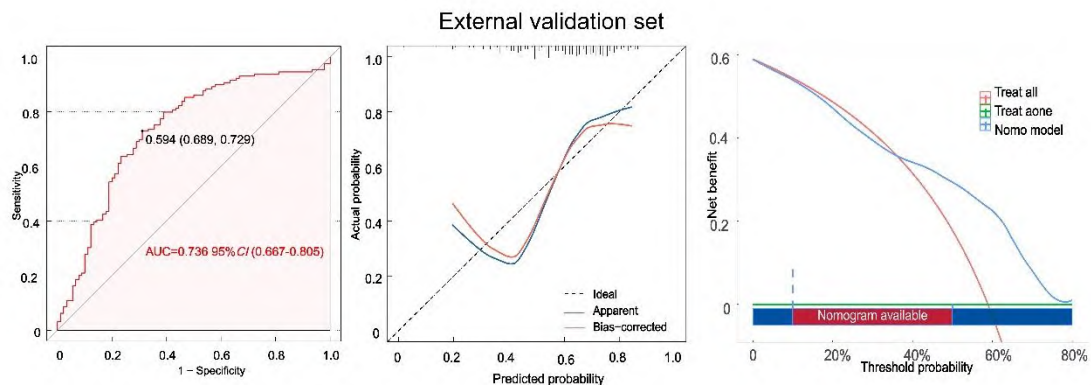


图 4 外部验证集中血液参数模型区分度、校准度与临床实用度综合评价

**Fig.4 Comprehensive evaluation of discrimination, calibration, and clinical applicability of blood parameter models in the external validation set**

A: The receiver operating characteristic curve; B: The calibration curve; C: The clinical decision curve.

### 3 讨论

PTB 与 CAP 在早期临床表现、影像学及炎症反应模式上高度重叠,导致临床鉴别困难,在基层医疗机构尤为突出。文献<sup>[13]</sup>显示,约 48%的 PTB 患者在确诊前曾被误诊为肺炎并接受抗生素治疗,延误抗结核治疗时机。因此,开发一种基于常规血液指标、便于在基层医疗机构实施的鉴别模型,对提高早期临床识别能力、减少误诊具有重要意义。

本研究联合采用 Boruta 算法与递归特征消除进行变量筛选,可避免传统单因素分析可能导致的过拟合与变量冗余。两者交集最终纳入多因素 Logistic 回归,确保变量稳定性与可解释性。分析显示,UA、ALP 与 PTB 风险呈正相关,而 BUN、MAO、NEU%、ALT、GSH 及年龄与 PTB 风险呈负相关,其中 7 项血液变量在感染、免疫、代谢与氧化应激方面具有明确生物学意义。UA 升高反映感染状态下蛋白质分解加剧及系统性炎症负荷<sup>[14]</sup>; ALP 与 ALT 参与肝脏免疫调节及急性期反应,结核肉芽肿形成可致肝酶轻度升高<sup>[15]</sup>; GSH 作为抗氧化分子,在结核感染氧化应激中作用显著<sup>[16]</sup>; NEU%在细菌性肺炎中通常升高,而 PTB 则以淋巴细胞/巨噬细胞反应为主,NEU%较低<sup>[17]</sup>; MAO 可能参与免疫调节及纤维化过程,其机制尚待探讨。

本模型在血液参数的基础上纳入年龄,提升模型的鉴别效能。训练集中 PTB 患者较 CAP 患者更年轻,这与结核病好发于青壮年、而 CAP 在老年人群中发病率更高的典型临床特征相符<sup>[11-12]</sup>,该差异支持将年龄纳入模型的合理性。

本研究建立的模型在训练集、内部验证集和外部验证集中均表现良好,AUC 分别为 0.800、0.787 和 0.736,校准良好。DCA 表明,模型在 10%~80%的阈值概率范围有明确的临床净收益,外部验证集中最佳适用阈值 40%~80%,建议结合患者初始临床评估确定适用场景。

由于本研究训练集与外部验证集在患者地域来源、诊疗流程等方面存在差异,这些差异可能导致血液参数的系统偏倚。但在外部验证中,该模型仍保持了良好性能,表明基于血液参数的鉴别模式具有较好的泛化能力。本模型基于相对变化模式而非绝对阈值,且血液检测在国内三级医院已相对标准化。表明该模型对这类差异具有较好的稳健性。未来多中心研究

可验证其在更广泛医疗环境下的适用性。同时,该模型保持了基本鉴别能力(灵敏度 68.9%,特异度 72.9%),但阴性预测值较低(63.9%),提示模型在不同中心的适用性需优化。尽管如此,其较高的阳性预测值(77.1%)支持在资源有限地区作为肺结核筛查辅助工具的应用潜力。

本研究尚存一定局限性:首先,仅纳入血液学指标,未来可结合影像学、基因组学与代谢组学等多维度数据以提升模型性能;其次,本研究为回顾性设计,尽管已严格控制了已知的强混杂因素(如糖尿病、HIV 感染),但仍可能无法完全识别所有潜在的免疫抑制状态,这可能在某种程度上影响结果的普适性;最后,外部验证样本量有限,模型的性能有待在大样本研究中验证。

综上,本研究构建并验证了一种基于常规血液参数的 PTB 与 CAP 鉴别预测模型,该模型具有良好的预测效能,适用于基层医疗机构的早期筛查与分诊。未来通过融合多组学数据、扩大样本规模及优化算法框架,有望进一步提升模型的稳定性与个体化诊疗价值,为 PTB 的精准防控提供支持。

#### 参考文献

- [1] Wen Z L, Wang L, Ma H, et al. Integrated single-cell transcriptome and T cell receptor profiling reveals defects of T cell exhaustion in pulmonary tuberculosis[J]. *J Infect*, 2024, 88(6): 106158. doi:10.1016/j.jinf.2024.106158.
- [2] Fang D N., Wang R H, Fan X T, et al. Recombinant BCG vaccine expressing multistage antigens of *Mycobacterium tuberculosis* provides long-term immunity against tuberculosis in BALB/c mice[J]. *Hum Vaccin Immunother*, 2024, 20(1): 2299607. doi:10.1080/21645515.2023.2299607.
- [3] Gunasekaran H, Sampath P, Thiruvengadam K, et al. A systematic review and meta-analysis of circulating serum and plasma microRNAs in TB diagnosis[J]. *BMC Infect Dis*, 2024, 24(1): 402. doi:10.1186/s12879-024-09232-0.
- [4] World Health Organization. Global tuberculosis report 2024[R]. Geneva: World Health Organization,2024.
- [5] Maze M J, Nyakunga G, Sakasaka P A, et al. Etiologic investigation of patients diagnosed with bacteriologically unconfirmed tuberculosis in Tanzania[J]. *Am J Trop Med Hyg*, 2023, 109(4): 733-9. doi:10.4269/ajtmh.23-0023.
- [6] 张莹, 张志苏, 施自伦, 等. 基于 CD161 构建肺结核的临床诊断模型[J]. *安徽医科大学学报*, 2025, 60(3): 515-23. doi:10.19405/j.cnki.issn1000-1492.2025.03.018.  
Zhang Y, Zhang Z S, Shi Z L, et al. Constructing a clinical diagnostic model for pulmonary tuberculosis based on CD161[J]. *Acta Univ Med Anhui*, 2025, 60(3): 515-23. doi:10.19405/j.cnki.issn1000-1492.2025.03.018.
- [7] 徐文辉, 张艳秋, 石洁, 等. 生物标志物在结核病诊断中的研究进展[J]. *中国防痨杂志*, 2024, 46(6): 713-21. . doi:10.19982/j.issn.1000-6621.20240046.

- Xu W H, Zhang Y Q, Shi J, et al. Advances in biomarker research for tuberculosis diagnosis[J]. Chin J Antituberc, 2024, 46(6): 713-21. doi:10.19982/j.issn.1000-6621.20240046.
- [8] Ahmed I A, Senan E M, Shatnawi H S A, et al. Multi-techniques for analyzing X-ray images for early detection and differentiation of pneumonia and tuberculosis based on hybrid features[J]. Diagnostics (Basel), 2023, 13(4): 814. doi:10.3390/diagnostics13040814.
- [9] Sia J K, Rengarajan J. Immunology of *Mycobacterium tuberculosis* infections[J]. Microbiol Spectr, 2019, 7(4). doi:10.1128/microbiolspec.GPP3-0022-2018.
- [10] Aliberti S, Dela Cruz C S, Amati F, et al. Community-acquired pneumonia[J]. Lancet, 2021, 398(10303): 906-19. doi:10.1016/S0140-6736(21)00630-9.
- [11] 周林, 刘二勇, 孟庆琳, 等. 《WS 288-2017 肺结核诊断》标准实施后肺结核诊断质量评估分析[J]. 中国防痨杂志, 2020, 42(9): 910-5. doi:10.3969/j.issn.1000-6621.2020.09.005.
- Zhou L, Liu E Y, Meng Q L, et al. Evaluation of the quality of pulmonary tuberculosis diagnosis after the implementation of the newly revised WS 288-2017 Diagnosis for pulmonary tuberculosis standards[J]. Chin J Antituberc, 2020, 42(9): 910-5. doi:10.3969/j.issn.1000-6621.2020.09.005.
- [12] 中华医学会, 中华医学会杂志社, 中华医学会全科医学分会, 等. 成人社区获得性肺炎基层诊疗指南(2018年)[J]. 中华全科医师杂志, 2019, 18(2): 117-26. doi:10.3760/cma.j.issn.1671-7368.2019.02.005.
- Chinese Medical Association, Chinese Medical Journals Publishing House, Chinese Society of General Practice, et al. Guideline for primary care of adult community acquired pneumonia: practice version(2018)[J]. Chin J Gen Pract, 2019, 18(2): 117-26. doi:10.3760/cma.j.issn.1671-7368.2019.02.005.
- [13] Yimer G, Gry M, Amogne W, et al. Evaluation of patterns of liver toxicity in patients on antiretroviral and anti-tuberculosis drugs: a prospective four arm observational study in Ethiopian patients[J]. PLoS One, 2014, 9(4): e94271. doi:10.1371/journal.pone.0094271.
- [14] 张明霞, 杨倩婷. 结核病诊断生物标志物的研究现状和展望[J]. 中国热带医学, 2019, 19(10): 907-11. doi:10.13604/j.cnki.46-1064/r.2019.10.01.
- Zhang M X, Yang Q T. Current status and prospects of tuberculosis diagnostic biomarkers[J]. China Trop Med, 2019, 19(10): 907-11. doi:10.13604/j.cnki.46-1064/r.2019.10.01.
- [15] Cao R, Kolloli A, Kumar R, et al. Effects of glutathione diminishment on the immune responses against *Mycobacterium tuberculosis* infection[J]. Appl Sci (Basel), 2021, 11(17): 8274. doi:10.3390/app11178274.
- [16] Shojaan H, Kalami N, Ghasempour Alamdari M, et al. Diagnostic value of the neutrophil lymphocyte ratio in discrimination between tuberculosis and bacterial community acquired pneumonia: a meta-analysis[J]. J Clin Tuberc Other Mycobact Dis, 2023, 33: 100395. doi:10.1016/j.jctube.2023.100395.
- [17] Collins G S, Moons K G M, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. BMJ, 2024, 385: e078378. doi:10.1136/bmj-2023-078378.