

成人血脂异常风险关联模型的构建与初步验证

李菡焱¹⁾, 孙华钰²⁾, 李欣²⁾, 袁逸爽³⁾, 陈莹²⁾, 王兴培⁴⁾,
张以涛⁵⁾, 张亚芬⁶⁾, 李丽菊³⁾

(1)石林彝族自治县疾病预防控制中心公共卫生科, 云南昆明 652200; 2)昆明医科大学公共卫生学院, 云南昆明 650500; 3)石林彝族自治县疾病预防控制中心慢性非传染性疾病防治科, 云南昆明 652200; 4)石林彝族自治县圭山镇卫生院, 云南昆明 652204; 5)石林彝族自治县大可乡卫生院, 云南昆明 652202; 6)石林彝族自治县石林卫生院, 云南昆明 652208)

[摘要] 目的 构建血脂异常风险关联模型, 为早期筛查及精准干预提供科学工具。方法 基于 2022 年 11 月至 2023 年 4 月云南省慢性病及其危险因素监测项目在石林县的横断面调查, 纳入 1577 例有效研究对象, 通过问卷调查、体格检查及实验室检测收集, 经预处理后纳入指标共 46 项, 按 7 : 3 划分训练集与测试集。对训练集数据先通过单因素分析筛选有统计学意义变量, 后采用 Lasso、Boruta、SWSFS 及 XGBoost 四种机器学习算法进行特征筛选, 选取四种算法共同识别的变量作为关键变量构建多因素 Logistic 回归模型并绘制列线图可视化。使用测试集数据通过混淆矩阵、ROC 曲线下面积、决策曲线分析及校准曲线系统评估模型性能。结果 本研究血脂异常患病率为 43.37%; 最终筛选出 6 个重要变量: BMI、腰围、糖化血红蛋白、心率、血尿酸及步行或骑行; 其中糖化血红蛋白呈显著非线性关联(非线性检验 $P = 0.003$), BMI ($P = 0.002$)、血尿酸 ($P < 0.001$) 呈线性正相关, 三者为血脂异常关键风险因素; 步行或骑行行为显著保护因素 ($P < 0.001$); 腰围 ($P = 0.384$) 与心率 ($P = 0.078$) 的关联未达统计学显著性。模型测试集 AUC 为 0.663, 灵敏度为 67.1%、特异度为 63.2%; DCA 显示在 0.25 ~ 0.65 阈值概率区间具有显著临床净获益, 校准曲线提示模型拟合良好。结论 本研究构建的血脂异常风险因素关联模型具有中等的鉴别效能及临床实用性, 可为血脂异常高危人群的早期识别及针对性干预提供参考。

[关键词] 血脂异常; 机器学习; 风险因素; 风险关联模型; 列线图

[中图分类号] R583.7 **[文献标志码]** A **[文章编号]** 2095 - 610X(2026)07 - 0062 - 09

Development and Preliminary Validation of a Risk Association Model for Dyslipidemia in Adults

LI Hanyi¹⁾, SUN Huayu²⁾, LI Xin²⁾, YUAN Yishuang³⁾, CHEN Ying²⁾, WANG Xingpei⁴⁾,
ZHANG Yitao⁵⁾, ZHANG Yafen⁶⁾, LI Liju³⁾

(1) Department of Public Health, Shilin Yi Autonomous County Center for Disease Control and Prevention, Kunming Yunnan 652200; 2) School of Public Health, Kunming Medical University, Kunming Yunnan 650500; 3) Department of Chronic Non-communicable Disease Prevention and Control, Shilin Yi Autonomous County Center for Disease Control and Prevention, Kunming Yunnan 652200; 4) Guishan Town Health Center, Shilin Yi Autonomous County, Kunming Yunnan 652204; 5) Dake Township Health Center, Shilin Yi Autonomous County, Kunming Yunnan 652202; 6) Shilin Health Center, Shilin Yi Autonomous County, Kunming Yunnan 652208, China)

[收稿日期] 2026 - 02 - 04

[基金项目] 国家自然科学基金(72164024)

[作者简介] 李菡焱(1986 ~), 女, 云南石林人, 医学学士, 主管护师, 主要从事慢性病管理工作。

[通信作者] 李丽菊, E-mail: 2223948432@qq.com

[Abstract] **Objective** To develop a risk association model for dyslipidemia, providing a scientific tool to support early screening and precision interventions. **Methods** Based on a cross-sectional survey of the Yunnan Provincial Chronic Disease and Risk Factor Surveillance Project conducted in Shilin Yi Autonomous County from November 2022 to April 2023, a total of 1577 eligible participants were enrolled. Data were collected through questionnaires, physical examinations, and laboratory tests. After preprocessing, 46 variables were included, and the dataset was randomly split into training and test sets at a ratio of 7 : 3. Univariate analysis was first performed on the training set to identify statistically significant variables. Subsequently, four machine learning algorithms—Lasso, Boruta, SWSFS, and XGBoost—were employed for feature selection. Variables commonly identified by all four algorithms were retained as key variables to construct a multivariable logistic regression model, which was visualized as a nomogram. Model performance was systematically assessed using the test set via the confusion matrix, area under the ROC curve (AUC), decision curve analysis (DCA), and calibration curves. **Results** The prevalence of dyslipidemia in this study was 43.37%. Six important variables were ultimately identified in the model: BMI, waist circumference, glycated hemoglobin, heart rate, serum uric acid, and walking or cycling. Among them, glycated hemoglobin exhibited a significant nonlinear association (P for nonlinearity=0.003), whereas BMI ($P = 0.002$) and serum uric acid ($P < 0.001$) showed linear positive correlations; establishing them as key risk factors for dyslipidemia. Walking or cycling was a significant protective factor ($P < 0.001$). The associations of waist circumference ($P = 0.384$) and heart rate ($P = 0.078$) with dyslipidemia did not reach statistical significance. The model achieved an AUC of 0.663 in the testing set; with a sensitivity of 67.1% and specificity of 63.2%. DCA showed a significant clinical net clinical benefit within the threshold probability range of 0.25 ~ 0.65, and the calibration curve indicated good model fit. **Conclusion** The proposed dyslipidemia risk association model demonstrates moderate discriminatory performance and clinical utility, offering a reference for early identification of high-risk individuals with dyslipidemia and targeted interventions.

[Key words] Dyslipidemia; Machine learning; Risk factors; Risk association model; Nomogram

血脂异常(Dyslipidemia)是以血浆总胆固醇、低密度脂蛋白胆固醇或甘油三酯升高,以及高密度脂蛋白胆固醇降低为特征的脂代谢紊乱^[1]。随着我国居民生活方式的显著改变,其患病率持续攀升^[2],我国成人血脂异常患病率已达38.1%^[1],且呈现年轻化趋势,给医疗卫生系统带来沉重负担。血脂异常作为心血管疾病独立可改变的危险因素^[3-4],通过促进动脉粥样硬化进展介导发病风险^[5],且其与动脉粥样硬化性心血管疾病风险的关联已被证实^[6];即使降脂治疗达标,患者仍存在残余心血管事件风险^[7]。加之其高患病率、低知晓率、低治疗率及低控制率的流行特征,已成为全球公共卫生领域的严峻挑战^[8]。

机器学习凭借复杂大数据处理与预测优势^[9],已在慢性病风险评估中展现应用价值^[10];而我国普通通用的成人血脂异常风险因素关联模型仍较缺乏,本研究整合Lasso、Boruta、SWSFS及XG-Boost四种算法,选取至少三种算法共同识别的因素作为关键变量;旨在开发精准简洁的风险评估工具,为临床早期筛查与个性化防治提供科学依据。

1 资料与方法

1.1 研究对象与样本量

本研究数据来源于2022年11月至2023年4月,云南省慢性病及其危险因素监测项目在石林县开展的横断面调查。研究对象为石林县18岁及以上的常住居民(调查前12个月内当地居住 ≥ 6 个月),已签署知情同意书。排除标准:孕妇、存在认知障碍或严重疾病可能影响调查者,以及居住于军营、学生宿舍等功能性区域的人口。本研究已通过昆明医科大学医学伦理委员会审批(KMMU2022MEC027)。

样本量计算公式:

$$n = \frac{Z_{\alpha/2}^2 \times \pi \times (1 - \pi)}{\delta^2} \times \text{deff}$$

基于我国成人血脂异常患病率为38.1%^[1],设定容许误差(δ)为4%,deff取1.3, $\alpha = 0.05$,计算得到每层最小样本量为680人。考虑10%无应答率,根据分层(城乡两层)本研究至少需调查1512人。

采用四阶段分层整群随机抽样法选取研究对象：首先从全县 7 个乡镇(街道)中随机抽取 4 个；随后在每个抽中乡镇(街道)中随机抽取 2 个村(居)委会；继而在每个村(居)委会中随机抽取 2 个居民小组；最终在 2 个居民小组中抽取 189 人(约 90 户)。通过统一培训调查员、预调查、现场督导和问卷核查等质量控制严格现场实施，共完成调查 1580 人，其中城市人口 792 人，农村人口 788 人。

1.2 研究方法

本研究资料收集包括问卷调查、体格检查及实验室检测三部分，均严格按照项目统一规范实施。问卷调查采用项目统一编制的标准化问卷，由经培训的调查员收集人口学特征、生活方式及疾病史等信息，数据实时同步录入并开展多级质量核查。体格检查采用标准化方法集中进行，测量身高、体重、腰围及血压；测量仪器均经国家计量认证。实验室检测采集调查对象空腹静脉血 8 mL，检测空腹血糖、糖化血红蛋白、血脂四项、尿酸、血肌酐、白蛋白及总蛋白；无糖尿病史者另予 75 g 无水葡萄糖口服后检测服糖后 2 h 血糖。所有血标本统一委托专业医学检验机构检测，血糖于采样后 24 h 内完成，各指标均按标准化方法测定，全程严格遵循省、市疾控中心实验室质控方案。符合以下至少一项标准，则可以诊断受试者患有血脂异常^[1]：(1)甘油三酯(triglyceride, TG)水平 ≥ 2.3 mmol/L；(2)总胆固醇(total cholesterol, TC)水平 ≥ 6.2 mmol/L；(3)低密度脂蛋白胆固醇(low-density lipoprotein cholesterol, LDL-C)水平 ≥ 4.1 mmol/L；(4)高密度脂蛋白胆固醇(high-density lipoprotein cholesterol, HDL-C)水平 < 1.0 mmol/L；(5)既往已被诊断为血脂异常。

1.3 统计学分析

本研究采用 R4.5.1 软件完成统计分析，含数据预处理、变量筛选、模型构建与验证等阶段。

1.3.1 数据预处理 研究初始调查对象 1580 例，经数据核查与质量控制，采用完全案例分析法剔除 3 例年龄信息缺失者，最终纳入有效样本 1577 例、共 46 项指标进行后续统计分析。连续变量以 $(\bar{x} \pm s)$ 或 $M(IQR)$ 表示，分类变量以构成比/率(%)表示；数据集按 7:3 比例分层随机划分为训练集与测试集，确保两组血脂异常患病率及变量结构一致；训练集用于变量筛选与模型构建，测试集用于模型性能评价。

1.3.2 统计方法 连续型自变量满足正态分布及

方差齐性采用两独立样本 t 检验，偏态资料采用 Mann-Whitney U 检验；无序分类自变量采用 χ^2 检验或校正 χ^2 检验或 Fisher 检验；有序分类自变量采用 Cochran-Armitage 趋势分析。 $\alpha=0.05$ ，双侧检验。

1.3.3 变量筛选 首先对变量进行单因素分析，单因素分析有统计学意义变量，采用以下四种算法筛选：(1)LASSO 回归(least absolute shrinkage and selection operator, LASSO)，通过 10 折交叉验证确定最优 λ ；(2) Boruta(boruta feature selection algorithm)，基于随机森林对比原始特征与影子特征重要性筛选；(3) 滑动窗口顺序前进选择法(sliding window sequential forward selection, SWSFS)，按特征重要性排序结合最小袋外误差率确定；(4) 极端梯度提升(eXtreme gradient boosting, XGBoost)，基于 SHAP 值评估特征重要性。最终纳入四种方法共同识别的变量。

1.3.4 模型构建与验证 基于纳入的关键变量，首先采用方差膨胀因子(variance inflation factor, VIF)评估自变量间多重共线性，以 $VIF < 5$ 判定为无严重多重共线性；对 $VIF \geq 5$ 的高度相关变量予以适当剔除；其次，对连续变量采用限制性立方样条(restricted cubic spline, RCS)条分析以评估其潜在非线性关系，若样条图形显示效应呈近似线性趋势则以线性项引入，若显示存在非线性关联则以样条函数(节点数=4)形式纳入。基于上述设定构建多因素 Logistic 回归全模型并绘制风险列线图。采用 ROC 曲线下面积(area under the ROC curve, AUC)及混淆矩阵评估鉴别能力，用决策曲线分析(decision curve analysis, DCA)评估临床实用性，校准曲线检验预估风险与实际风险的一致性^[11]。

2 结果

2.1 人口学特征比较

本研究共纳入 1577 例研究对象，血脂异常患者 684 例，患病率为 43.37%；血脂异常组在年龄及文化程度上与正常组比较差异有统计学意义($P < 0.05$)，而性别、民族、职业、地区分布均未见统计学差异($P > 0.05$)；血脂异常组年龄中位数为 61.00(54.00 ~ 70.00)岁，血脂正常组为 59.00(50.00 ~ 70.00)岁，差异具有统计学意义($P = 0.026$)，提示血脂异常组年龄偏高。文化程度分布差异亦有统计学意义($P = 0.013$)，异常组小学及以下学

历占比(64.77%)高于正常组(62.26%), 大专及以上学历仅占 1.90%, 低于正常组(5.04%), 提示血脂异常组文化程度偏低。性别构成上, 两组男性分别占 55.56% 和 55.21%, 女性占 44.44% 和 44.79%, 差异无统计学意义($P=0.890$)。两组均以汉族为主, 异常组和正常组汉族分别占 85.38% 和 87.01%, 彝族占 13.89% 和 11.98%, 分布无统计学差异($P=0.460$)。职业方面, 两组均以生产与服务类人员为主(分别占 77.63% 和 78.95%), 两组职业构成相似($P=0.890$)。城乡比例在两组间基本均衡, 异常组城市占比 50.29%、农村占比 49.71%, 正常组城市占比 49.94%、农村占比 50.06%, 差异无统计学意义($P=0.891$), 见表 1。

2.2 特征变量筛选

在进行变量筛选前对训练集(1 105 例)与测试集(472 例)的基线特征进行比较以评估数据划分的均衡性。结果显示, 除民族($P=0.035$)与戒烟($P=0.028$)两项外, 两组在其他人口学特征、生活方式、疾病史及生化指标等方面的分布差异均无统计学意义($P>0.05$), 表明数据划分基本均衡,

具有良好的可比性。

首先对纳入的 45 个变量进行单因素分析, 筛选出有统计学意义的变量 16 个, 分别为: BMI、高血压、糖尿病、腰围、糖化血红蛋白、血尿酸、低脂膳食、心率、体重控制、步行或骑行、低盐膳食、规律锻炼、总蛋白、每天蔬菜或水果、文化程度、血肌酐。将因素分析显著变量纳入后续特征筛选, LASSO 回归在 λ_{1se} 时筛选出 7 个变量: BMI、腰围、糖化血红蛋白、血尿酸、低脂膳食、心率、步行或骑行, 见图 1A; Boruta 算法基于随机森林和影子特征比较, 经 1000 次迭代后确认 10 个重要变量: BMI、糖尿病、腰围、糖化血红蛋白、血尿酸、低脂膳食、心率、步行或骑行、总蛋白、血肌酐, 见图 1B; SWSFS 按变量重要性排序逐步建模, 以最小 OOB 误差(0.342)为标准, 确定 12 个变量: BMI、血尿酸、血肌酐、总蛋白、腰围、糖化血红蛋白、心率、文化程度、步行或骑行、高血压、糖尿病、体重控制, 见图 1C~D; XGBoost 模型结合 SHAP 值分析, 识别出 14 个变量, 其中 BMI、糖化血红蛋

表 1 血脂异常组与血脂正常组人口学特征比较 [$M(IQR)/n(\%)$]

Tab. 1 Comparison of demographic characteristics between dyslipidemia and normolipidemic groups [$M(IQR)/n(\%)$]

变量	血脂异常组($n=684$)	血脂正常组($n=893$)	Z/χ^2	P
年龄(岁)	61.00(54.00~70.00)	59.00(50.00~70.00)	2.225	0.026*
性别			0.019	0.890
女	380(55.56)	493(55.21)		
男	304(44.44)	400(44.79)		
民族			1.553	0.460
汉族	584(85.38)	777(87.01)		
彝族	95(13.89)	107(11.98)		
其他少数民族	5(0.73)	9(1.01)		
文化程度			2.474	0.013*
小学及以下	443(64.77)	556(62.26)		
初中	191(27.92)	236(26.43)		
高中/中专/技校	37(5.41)	56(6.27)		
大专及以上	13(1.90)	45(5.04)		
职业			3.984	0.890
生产与服务类	531(77.63)	705(78.95)		
编制与专业技术类	10(1.46)	23(2.58)		
其他从业类	22(3.22)	30(3.36)		
非从业类	121(17.69)	135(15.12)		
地区			0.018	0.891
城市	344(50.29)	446(49.94)		
农村	340(49.71)	447(50.06)		

* $P<0.05$ 。

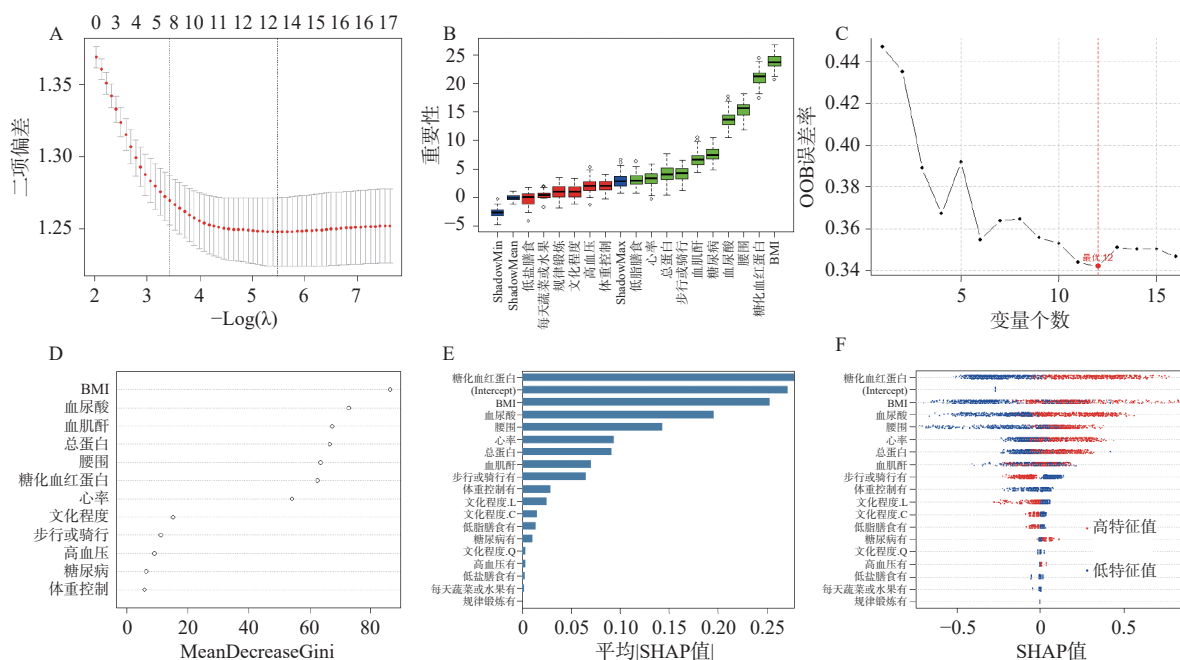


图 1 成人血脂异常四种算法特征筛选结果

Fig. 1 Feature selection results for dyslipidemia in adults using four algorithms

A: LASSO 交叉验证曲线; B: Boruta 分析结果; C ~ D: SWSFS 分析结果; E ~ F: XGBoost 结合 SHAP 分析结果。

白、血尿酸、腰围、总蛋白、血肌酐、心率、步行或骑行的 SHAP 值贡献最为显著, 见图 1F ~ G。综合四种算法, 以“四种方法共同选取”为标准, 最终确定 6 个重要变量: BMI、腰围、糖化血红蛋白、心率、血尿酸及步行或骑行。

2.3 构建多因素 logistic 回归模型和列线图

对最终确定的 6 个变量进行多重共线性检验, 结果显示所有自变量的 VIF 均低于 5 (范围 1.007 ~ 2.353), 提示各变量间不存在严重的多重共线性。后对纳入的 5 个连续变量进行了限制性立方样条分析, 结果显示 BMI、腰围及血尿酸呈近似线性趋势, 以线性项引入; 糖化血红蛋白、心率则存在显著非线性关联, 以样条函数形式纳入。见图 2。基于上述探索构建多因素 Logistic 回归全模型并绘制列线图, 见表 2、图 3A。

2.4 模型性能评估

模型测试集 AUC 为 0.663 (0.613, 0.713), 见图 3B; 灵敏度为 0.671, 特异度为 0.632; 显示模型具备一定的泛化能力, 但区分效能仍处于中等水平。DCA 分析显示该模型在 0.25 至 0.65 的阈值概率区间内能带来净获益, 见图 3C; 校准曲线平均绝对误差为 0.006, 反应推算风险与实际发生率拟合良好, 见图 3D。

按性别和地区进行亚组分析, 结果表明模型在女性亚组的 AUC 为 0.628 (95%CI: 0.558 ~ 0.697), 男性亚组为 0.705 (95%CI: 0.634 ~ 0.775);

城市地区 AUC 为 0.649 (95%CI: 0.577 ~ 0.721), 农村地区为 0.679 (95%CI: 0.611 ~ 0.748)。DeLong 检验显示, 不同性别 ($P=0.127$) 及不同地区 ($P=0.548$) 亚组间 AUC 差异均无统计学意义, 提示该风险关联模型在不同人群中具有稳定一致的区分能力。

3 讨论

血脂异常患病率升高已成为全球性公共卫生问题^[12]。本研究显示石林县人群血脂异常患病率为 43.37%, 显著高于国内报道水平 (38.1%)^[1], 或与西南地区动物油-畜肉-粗粮膳食模式及传统南方膳食模式相关^[13], 提示当地亟需强化血脂异常的预防与控制工作。本研究基于多因素分析构建的血脂异常风险列线图, 将传统 Logistic 回归结果转化为直观的可视化评估工具, 清晰呈现各因素对血脂异常发生的风险贡献度。该工具不仅可作为临床咨询的高效沟通媒介, 更利于在基层医疗与健康管理场景中普及应用。研究中模型测试集 AUC 为 0.663, 灵敏度与特异度分别为 0.671 和 0.632。上述指标提示模型具备一定的泛化稳定性, 但其判别区分效能整体处于中等水平, 表明模型在独立人群中的判别能力较为有限。而模型校准度分析显示, 其预估风险与实际发病风险一致性良好, 临床决策曲线分析亦证实, 模型在适宜风险阈值范围内可产生临床净获益, 具备一

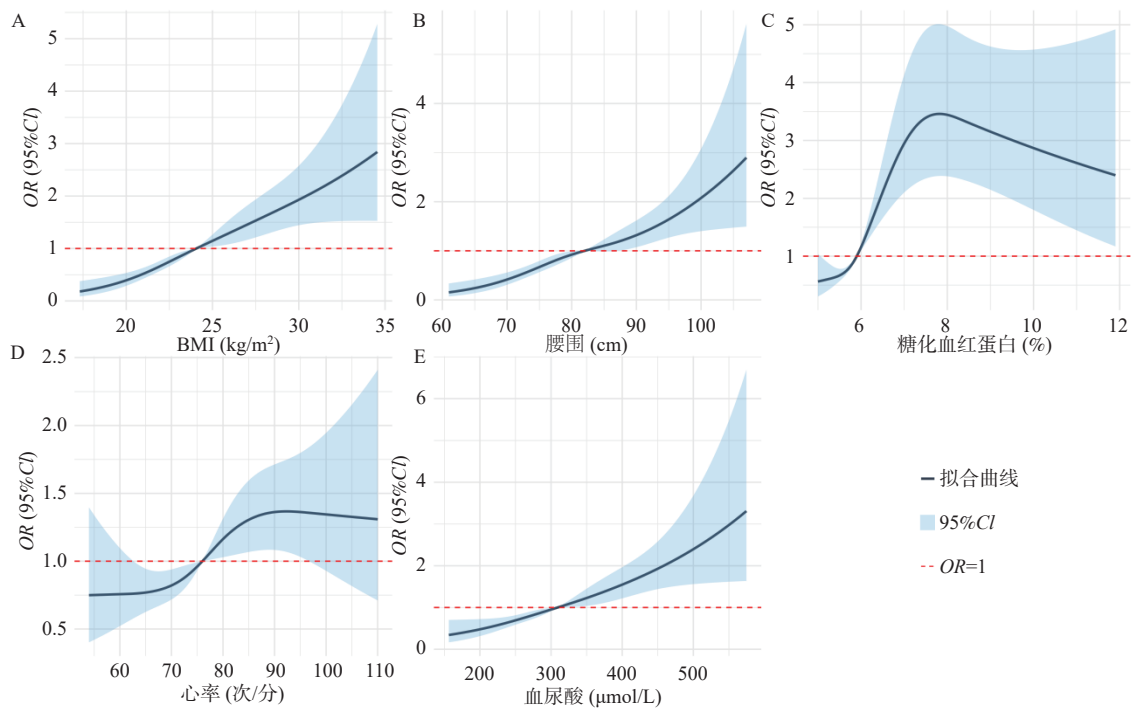


图 2 各连续变量与血脂异常发生风险的限制性立方样条拟合曲线

Fig. 2 Restricted cubic spline plots of the association between each continuous variable and dyslipidemia risk in adults
A: BMI 的 RCS 拟合曲线; B: 腰围的 RCS 拟合曲线; C: 糖化血红蛋白的 RCS 拟合曲线; D: 心率的 RCS 拟合曲线; E: 血尿酸的 RCS 拟合曲线。

表 2 成人血脂异常风险关联模型多因素 logistic 回归分析结果

Tab. 2 Multivariate logistic regression analysis of the risk association model of adult dyslipidemia						
变量	β	SE	Z	OR(95%CI)	P	非线性检验P
糖化血红蛋白(%)	—	—	—	—	<0.001***	0.003**
心率(次/min)	—	—	—	—	0.078	0.508
BMI(kg/m ²)	0.086	0.028	3.090	1.089(1.032 ~ 1.150)	0.002**	—
腰围(cm)	0.009	0.011	0.870	1.009(0.989 ~ 1.031)	0.384	—
血尿酸(μmol/L)	0.004	0.001	4.880	1.004(1.003 ~ 1.006)	<0.001***	—
步行或骑行(以无为参照)						
有	-0.491	0.138	-3.570	0.612(0.467 ~ 0.801)	<0.001***	—

“—”代表不适用;连续变量OR值的单位变化说明: BMI对应每增加1 kg/m²;腰围对应每增加1 cm;血尿酸对应每增加1 μmol/L; *P<0.05; **P<0.01; ***P<0.001。

定临床应用可行性。与同类研究相比,王俊等^[14]基于机器学习构建的老年人高脂血症预测模型 AUC 为 0.690~0.820,杜文雯等^[15]利用肠道菌群特征构建的成人血脂异常预测模型 AUC 为 0.784,若单纯以 AUC 判别效能衡量,本研究模型并无优势。但需明确的是,高判别度预测模型通常需以提升变量获取成本、牺牲临床可操作性为代价,例如机器学习模型依赖专业算力支持、模型可解释性欠佳,肠道菌群检测在基层医疗机构难以常规开展,落地应用受限。而本研究所纳入的指标均为临床及社区健康体检常规可获取的简易指标,

在保证一定准确性的前提下显著降低了应用门槛与资源依赖。

3.1 非线性关联变量的发现与生物学机制

本研究结果显示,糖化血红蛋白(glycated hemoglobin, HbA1c)与血脂异常的发生风险呈现显著的非线性关联(整体 P<0.001,非线性检验 P=0.003)。区间效应分析表明,HbA1c 从 5.6% 升至 6.3% 时风险增加约 85%(OR=1.848,95%CI: 1.426~1.2.394),提示血糖控制恶化初期对脂质代谢的扰动最为剧烈;当 HbA1c 水平超过 8% 后,RCS 曲线呈现边际递减模式,风险上升速率有所

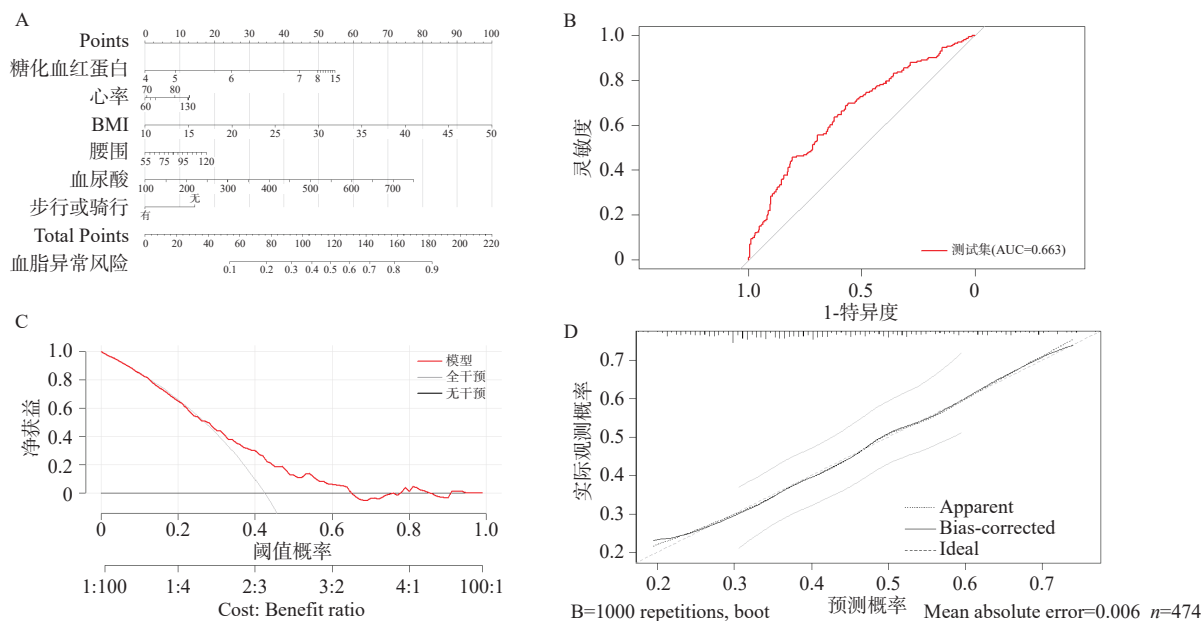


图 3 成人血脂异常风险因素关联模型列线图及性能评估结果

Fig. 3 Nomogram of the risk factor association model for adult dyslipidemia and model performance evaluation results

A: 成人血脂异常风险因素关联模型列线图; B: 成人血脂异常风险因素关联模型 ROC 曲线; C: 成人血脂异常风险因素关联模型决策曲线分析; D: 成人血脂异常风险因素关联模型校准曲线。

减缓, 但整体风险仍维持在较高水平。这一趋势与 Hong 等^[16] 利用非参数回归技术获得的发现一致, 该研究证实 HbA1c 与多项脂质代谢指标存在复杂的非线性交互作用。已有研究证实, 糖脂代谢紊乱的发生发展与机体炎症反应激活、氧化应激失衡密切相关, 异常的糖代谢状态可通过介导炎症与氧化应激损伤, 进一步加剧脂质代谢紊乱, 推动血脂异常的发生进展^[17]。综上所述, 糖化血红蛋白不仅是血脂异常风险关联模型中的强关联指标, 更是反映代谢网络稳态失衡的核心标志物; 重视糖化血红蛋白的监测与早期达标, 可能是阻断血脂异常发生发展、防控心血管代谢疾病的关键防线。

本研究未观察到静息心率与血脂异常发生风险间存在统计学显著的整体关联(整体检验 $P = 0.078$, 非线性检验 $P = 0.508$, 无显著非线性证据)。但在特定区间对比中发现, 当静息心率从 70 次/min 升至 84 次/min, 点估计提示风险增加约 44% ($OR = 1.440$, 95%CI: 1.029 ~ 2.014); 然而需注意, RCS 整体检验需联合估计所有样条基函数项, 多自由度可能稀释检验效能致使整体 P 值未达传统的显著性水平。这一点估计的风险升高方向, 与交感神经过度激活驱动代谢紊乱的经典框架相符: 交感神经兴奋可促进脂肪组织脂解, 并增强肝脏甘油三酯的合成与分泌, 进而导致血脂代谢异常^[18]。静息心率作为反映交感-副交感

神经张力平衡的无创、简易指标, 其与致动脉粥样硬化血脂谱的相关性已在多项不同人群研究中获得独立验证^[19]。综上所述, 本研究结果尚不支持将静息心率列为血脂异常风险评估的独立参考指标, 该阳性点估计的稳健性尚待更大样本研究验证。

3.2 线性关联变量的病理生理意义

BMI 与血脂异常风险呈单调递增的线性趋势 ($P = 0.002$), BMI 每增加 1 kg/m², 风险相应上升 8.9% ($OR = 1.089$, 95%CI: 1.032 ~ 1.150)。这一结果与既往文献报道一致: Shahraz 等^[20] 基于 PERSIAN Guilan 队列的研究同样发现, BMI 每升高 1 kg/m², 血脂异常风险增加 9%。但不同人群 BMI 与血脂异常发生风险的剂量-反应关系存在差异: 四川某研究呈 J 型曲线^[21], 潮安区研究与本研究一致为单调递增^[22], 提示该关联受人群特征修饰。其机制可能与肥胖引发脂质代谢紊乱相关, 包括肠道菌群改变、脂肪因子分泌异常等共同破坏脂质稳态^[23]。腰围变化趋势与 BMI 相近, 但多因素校正后无统计学意义 ($P = 0.384$); 腰围每增加 1 cm, 风险升高 0.9% ($OR = 1.009$, 95%CI: 0.989 ~ 1.031), 与代谢综合征优先采用腰围评估的观点存在差异。原因可能为: 腰围测量易受呼吸状态及操作者差异影响, 误差可能大于 BMI; 腰围与 BMI 存在一定共线性 ($VIF = 2.353$), 校正后独立作用减弱。但在模型中腰围由 75 cm 增至

89 cm 时 $OR = 1.139$ (95%CI: 0.850 ~ 1.528), 虽未达到统计学显著性, 但点估计值提示存在一定的正向关联趋势, 不显著结果可能与研究检验效能不足有关。综上所述, BMI 作为反映全身性肥胖的稳健指标, 适用于基层筛查与风险分层; 尽管本研究中腰围未显示出独立关联价值, 但在个体化风险评估中仍具有辅助参考意义。

已有研究表明, 血尿酸与脂代谢异常存在明确关联^[24]。本研究结果显示, 血尿酸与血脂异常风险呈线性正相关($P < 0.001$), 每增加 $1\mu\text{mol/L}$ 风险升高 0.4% ($OR = 1.004$, 95%CI: 1.003 ~ 1.006); Mohammadi 等^[25] 的队列研究证实高尿酸血症者血脂异常风险显著升高。值得注意的是, 袁空军等^[26] 基于 RCS 模型报道血尿酸与血脂异常呈非线性关联, 低尿酸区间风险较低而高水平区间风险骤升, 与本研究观察到的线性递增趋势存在差异, 提示不同人群尿酸代谢特征及风险阈值可能存在异质性。机制上, 高尿酸通过激活 NLRP3 炎症小体、诱导氧化应激及加剧胰岛素抵抗, 驱动“尿酸-代谢紊乱-血脂异常”的病理级联^[27]。血尿酸在本模型中效应独立于 BMI 与腰围等, 可捕捉传统肥胖指标之外的残余代谢风险, 为血脂异常早期筛查提供增量价值。

3.3 可干预生活方式因素的保护效应

模型还纳入步行或骑行这项行为因素, 结果显示步行或骑行与血脂异常风险呈显著负相关($P < 0.001$), 存在该习惯者发生血脂异常的风险较无此习惯者降低约 39% ($OR = 0.612$, 95%CI: 0.467 ~ 0.801); 凸显了可干预行为因素在血脂异常防控中的核心价值。体力活动对脂质谱的改善效应已获广泛循证支持: 基于 CHARLS 的纵向研究发现高水平体力活动可使血脂异常风险降低 19%^[28], Smart 等^[29] 的荟萃分析进一步确认运动训练可全面改善 TC、LDL-C 及 TG 水平。这项可干预因素的纳入, 可为后续开展个体化生活方式指导提供明确靶点。

本研究存在一些局限: 作为一项基于西南地区单一人群的研究, 样本区域代表性有限, 且受限于研究设计周期与数据获取条件, 现阶段尚未在其他独立人群中完成外部验证, 模型在中国更广泛人群中的泛化能力有待进一步评估; 其次, 研究未纳入遗传、肠道菌群等潜在变量, 可能限制了模型效能的进一步提升; 此外, 本研究为横断面设计, 仅能分析变量间的关联关系, 无法确立因果推断, 其对血脂异常发病的长期预测效能,

有待前瞻性队列研究进一步确认。

本研究基于西南地区特定人群, 采用多算法筛选构建血脂异常风险关联模型, 模型具有一定的鉴别效能与临床参考价值。其中, 糖化血红蛋白呈显著非线性关联, BMI、血尿酸呈线性正相关, 三者为血脂异常关键风险因素; 步行或骑行具保护效应; 而腰围与心率的关联意义有限。该模型可补充区域人群血脂异常研究数据, 也可为西南地区血脂异常高危人群的筛查、风险分层及早期干预提供参考工具, 助力优化血脂管理与心血管疾病防控。

[参考文献]

- [1] 李建军, 赵水平, 高润霖, 等. 中国血脂管理指南 (2023 年) [J]. 中华心血管病杂志, 2023, 51(3): 221-255.
- [2] Opoku S, Gan Y, Yobo E A, et al. Awareness, treatment, control, and determinants of dyslipidemia among adults in China [J]. *Scientific Reports*, 2021, 11: 10056.
- [3] 全面管理血脂相关心血管风险专家共识制订专家组, 苏州工业园区东方华夏心血管健康研究院. 全面管理血脂相关心血管风险专家共识 (2025) [J]. *中华心血管病杂志(网络版)*, 2025, 8(1): 1-23.
- [4] 周慧敏, 冉利梅, 杨大刚, 等. 1838 例健康体检人群血脂异常流行特征及影响因素分析 [J]. 贵州医科大学学报, 2023, 48(10): 1186-1192.
- [5] Kalwicz M, Roth M. A comprehensive review of the genetics of dyslipidemias and risk of atherosclerotic cardiovascular disease [J]. *Nutrients*, 2025, 17(4): 659.
- [6] Bi L, Yi J, Wu C, et al. Atherosclerotic cardiovascular disease risk and lipid-lowering therapy requirement in China [J]. *Frontiers in Cardiovascular Medicine*, 2022, 9: 839571.
- [7] Liu H H, Guo Y L, Zhu C G, et al. Synergistic effect of the commonest residual risk factors, remnant cholesterol, lipoprotein(a), and inflammation, on prognosis of statin-treated patients with chronic coronary syndrome [J]. *Journal of Translational Medicine*, 2022, 20(1): 1-10.
- [8] Xia Q, Chen Y, Yu Z, et al. Prevalence, awareness, treatment, and control of dyslipidemia in Chinese adults: A systematic review and meta-analysis [J]. *Frontiers in Cardiovascular Medicine*, 2023, 10: 1186330.

- [9] Sadiq I Z, Abubakar F S, Saliu M A, et al. Machine learning algorithms for predictive modeling of dyslipidemia-associated cardiovascular disease risk in pregnancy: A comparison of boosting, random forest, and decision tree regression[J]. *Bulletin of the National Research Centre*, 2025, 49(1): 1–13.
- [10] 彭莲, 马璐, 徐玉艳, 等. 血脂异常风险评估方法研究进展[J]. *现代预防医学*, 2022, 49(3): 385–389.
- [11] Zhang L X, Hou S B, Zhao F F, et al. Development and validation of a hyperlipidemia risk prediction model for middle-aged and older adult Chinese using 2015 CHARLS data[J]. *Frontiers in Public Health*, 2025, 13: 1420596.
- [12] Chew N W S, Mehta A, Goh R, et al. The global cardiovascular-liver-metabolic syndemic: epidemiology, trends and challenges[J]. *Nature Reviews Cardiology*, 2026, 23(4): 219–238.
- [13] Li J, Zhao L, Yu D, et al. Regional dietary pattern associated with the risk of hypertensive dyslipidemia multimorbidity in Chinese elderly: Results from China nutrition and health surveillance in 2015–2017[J]. *Nutrients*, 2025, 17(5): 852.
- [14] 王俊, 王培承, 郭琼, 等. 基于机器学习的老年人高脂血症风险预测[J]. *现代预防医学*, 2023, 50(18): 3408–3413.
- [15] 杜文雯, 周玮玮, 关方旭, 等. 基于肠道菌群特征的中国四省(自治区)成年人血脂异常发生风险预测模型构建及评价[J]. *卫生研究*, 2025, 54(5): 738–743.
- [16] Hong S F, Ahmad W M A W, Adnan M N, et al. Advanced health assessment of HbA1c through nonparametric regression: Investigating TG, FBS, BMI, TC, and LDL interactions[J]. *JP Journal of Biostatistics*, 2025, 25(1): 215–221.
- [17] 甄云蕾, 张贺芳, 王一冰. 血浆致动脉硬化指数联合血尿酸/高密度脂蛋白胆固醇比值预测 2 型糖尿病发生早期 DPN 的价值分析[J]. *河北医科大学学报*, 2025, 46(8): 953–958.
- [18] Grassi G, Drager L F. Sympathetic overactivity, hypertension and cardiovascular disease: state of the art[J]. *Current Medical Research and Opinion*, 2024, 40(sup1): 5–13.
- [19] Zafir B. Resting heart rate in patients with diabetes: A risk marker or direct mediator of increased cardiovascular risk?[J]. *Kardiologia Polska*, 2025, 83(11): 1247–1249.
- [20] Shahraz J, Joukar F, Sheida F, et al. Associations between body mass index (BMI) and dyslipidemia: results from the PERSIAN guilan cohort study (PGCS) [J]. *Obesity Science and Practice*, 2025, 11(1): e70055.
- [21] 蒋莉华, 陈颖, 徐新茂, 等. 四川省三县(市)成年居民肥胖与血脂异常关联研究[J]. *四川大学学报(医学版)*, 2026, 57(1): 192–201.
- [22] 蒋蓉, 陈思彤, 吴玉霞, 等. 潮州市潮安区居民 BMI、腰围与血脂异常患病的剂量-反应关系[J]. *华南预防医学*, 2026, 52(1): 18–24.
- [23] Zhang X, Guan F, Gou W, et al. Multi-trajectories of BMI, waist circumference, gut microbiota, and incident dyslipidemia: A 27-year prospective study[J]. *mSystems*, 2025, 10(5): e0024325.
- [24] 沈亚洁, 杨亚丽, 余志清, 等. 云南省贡山县少数民族高尿酸血症患病率调查[J]. *昆明医科大学学报*, 2025, 46(3): 39–43.
- [25] Mohammadi H, Mirjalili S R, Marques-Vidal P M, et al. The association between serum uric acid levels and the 10-year prospective risk of dyslipidemia: A cohort study from healthy heart project (YHHP)[J]. *Health Science Reports*, 2025, 8(5): e70857.
- [26] 袁空军, 赵创艺, 杨媛, 等. 基于限制性立方样条模型分析体检人群血清尿酸和血脂异常关联性研究[J]. *中国卫生统计*, 2023, 40(5): 664–669.
- [27] Deji-Olorunboba O O, Balogun J O, Elufioye T O, et al. Hyperuricemia and insulin resistance: interplay and potential for targeted therapies[J]. *Journal of Translational Medicine*, 2025, 5(3): 1–30.
- [28] Wan K, Wang R W, Yan H M, et al. Physical activity dose-response and long-term effects on dyslipidemia in Chinese adults: A CHARLS study[J]. *BMC Public Health*, 2025, 25(1): 1–14.
- [29] Smart N A, Downes D, van der Touw T, et al. The effect of exercise training on blood lipids: A systematic review and meta-analysis[J]. *Sports Medicine*, 2025, 55(1): 67–78.