

老年动脉粥样硬化性心血管疾病患者估算葡萄糖处置率与长期死亡风险的关系：一项回顾性队列研究与机器学习预测模型构建

安莉¹, 张帷², 付治卿², 吕月³, 李珊^{3*}

¹解放军总医院第二医学中心呼吸与危重症医学科, 北京 100853; ²解放军总医院第二医学中心心血管内科, 北京 100853; ³解放军总医院第二医学中心保健三科, 北京 100143

[中图分类号] R541.4 [文献标志码] A [DOI] 10.11855/j.issn.0577-7402.2436.2026.0625

[声明] 本文所有作者声明无利益冲突

[引用本文] 安莉, 张帷, 付治卿, 等. 老年动脉粥样硬化性心血管疾病患者估算葡萄糖处置率与长期死亡风险的关系: 一项回顾性队列研究与机器学习预测模型构建[J]. 解放军医学杂志, 2026, 51(19): 43-51.

[收稿日期] 2025-11-27

[录用日期] 2026-05-02

[上线日期] 2026-06-25

[摘要] **目的** 探讨估算葡萄糖处置率(eGDR)与老年动脉粥样硬化性心血管疾病(ASCVD)患者全因及心血管死亡的关联, 构建全因死亡风险预测模型并进行验证。**方法** 收集2008年1月—2022年12月解放军总医院第二医学中心≥60岁、确诊ASCVD的患者2479例进行回顾性分析。根据eGDR四分位数水平将所有患者分为4组: Q₁组($n=617$)、Q₂组($n=620$)、Q₃组($n=622$)、Q₄组($n=620$), 以评估eGDR与死亡风险的剂量-反应关系。采用COX比例风险模型和限制性立方样条(RCS)分析eGDR与死亡风险的关联。比较7种机器学习算法[K最近邻、支持向量机、随机森林、最小绝对收缩和选择算子回归(LASSO)回归、朴素贝叶斯、极限梯度提升和决策树]对全因死亡的预测效能。最优模型采用夏普利加性解释(SHAP)对变量重要性进行排序。**结果** 在中位随访时间的8.1年期间, 发生全因死亡1101例(44.4%)和心血管死亡429例(17.3%)。与eGDR水平最低四分位数(Q₁)患者相比, 最高四分位数(Q₄)患者的全因死亡风险降低28%($HR=0.72$, 95%CI 0.57~0.90), 心血管死亡风险降低42%($HR=0.58$, 95%CI 0.44~0.78)。RCS分析显示, eGDR与全因和心血管死亡呈明显负相关($P<0.05$)。7种算法模型中, 随机森林的曲线下面积(AUC)最高($AUC=0.748$, 95%CI 0.740~0.756)。SHAP分析显示, 年龄、低密度脂蛋白胆固醇、肾小球滤过率、eGDR是预测全因死亡的4个最重要特征。**结论** eGDR水平降低与≥60岁ASCVD患者全因及心血管死亡风险增加独立相关。纳入eGDR的随机森林模型对全因死亡具有较好的预测能力, 可将eGDR作为该人群死亡风险分层的有潜力的生物标志物。

[关键词] 胰岛素抵抗; 估算葡萄糖处置率; 动脉粥样硬化性心血管疾病; 死亡; 机器学习; 风险预测

Estimated glucose disposal rate and long-term mortality in older adults with atherosclerotic cardiovascular disease: a retrospective cohort study and machine-learning prediction

An Li¹, Zhang Wei², Fu Zhi-Qing², Lyu Yue³, Li Shan^{3*}

¹Department of Respiratory and Critical Care Medicine, the Second Medical Center & National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital, Beijing 100853, China

²Department of Cardiology, the Second Medical Center & National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital, Beijing 100853, China

³Third Department of Healthcare, the Second Medical Center, the Second Medical Center & National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital, Beijing 100853, China

*Corresponding author, E-mail: lishan301301@163.com

This work was supported by the National Clinical Research Centre for Geriatric Diseases Open Project (NCRCG-PLAGH-2024016)

[基金项目] 国家老年疾病临床医学研究中心开放课题(NCRCG-PLAGH-2024016)

[作者简介] 安莉, 医学博士, 副主任医师, 从事呼吸系统危重症与代谢风险管理方面的研究

[通信作者] 李珊, E-mail: lishan301301@163.com

[Abstract] Objective To investigate the association between estimated glucose disposal rate (eGDR) and all-cause and cardiovascular mortality in older adults with established atherosclerotic cardiovascular disease (ASCVD), and to develop and validate a prediction model for all-cause mortality. **Methods** This retrospective study enrolled 2,479 hospitalized and outpatient patients aged ≥ 60 years with established ASCVD at the Second Medical Center of PLA General Hospital from January 2008 to December 2022. Participants were divided into four groups according to the eGDR quartiles: Q_1 group ($n=617$), Q_2 group ($n=620$), Q_3 group ($n=622$), and Q_4 group ($n=620$), to evaluate the dose-response relationship between eGDR and mortality risk. COX proportional hazards models and restricted cubic spline curve analysis were employed to examine the association between eGDR levels and mortality risk. The predictive performance of seven machine learning models (K-nearest neighbours, support vector machines, random forests, LASSO regression, naive Bayes, gradient-boosted regression and decision trees) on all-cause mortality was compared. The SHapley Additive exPlanations (SHAP) method was used to interpret variable importance in the optimal model. **Results** During a median follow-up of 8.1 years, 1,101 (44.4%) all-cause deaths and 429 (17.3%) cardiovascular deaths occurred. Compared with patients in the lowest eGDR quartile, those in the highest quartile had a 28% reduction in all-cause mortality ($HR=0.72$, 95% CI 0.57-0.90) and a 42% reduction in cardiovascular mortality ($HR=0.58$, 95% CI 0.44-0.78). Restricted cubic spline analysis demonstrated a significant negative association between eGDR and all-cause and cardiovascular mortality ($P<0.05$). Of the seven machine learning models tested, the random forest model achieved the highest area under the curve ($AUC=0.748$, 95%CI 0.740-0.756). SHAP analysis identified age, low-density lipoprotein cholesterol (LDL-C), estimated glomerular filtration rate (eGFR), and eGDR as the four most important predictors of all-cause mortality. **Conclusions** Lower eGDR levels were independently associated with increased risks of both all-cause and cardiovascular mortality in patients aged ≥ 60 years with ASCVD. The random forest model incorporating eGDR demonstrated good predictive performance for all-cause mortality risk, suggesting that eGDR may serve as a promising biomarker for mortality risk stratification in this population.

[Key words] insulin resistance; estimated glucose disposal rate; atherosclerotic cardiovascular disease; mortality; machine learning; risk prediction

动脉粥样硬化性心血管疾病(atherosclerotic cardiovascular disease, ASCVD)是全球非传染性疾病中导致死亡的首要原因。目前全球约有5亿ASCVD患者,每年死亡人数高达1900万^[1]。我国ASCVD患病人数已突破7000万,其中冠心病及缺血性卒中导致的死亡占我国心血管总死亡的85%^[2]。随着人口老龄化的加剧,疾病负担持续上升,对医疗卫生体系形成严峻挑战。因此,寻找可早期识别、精准量化的风险标记物,并构建适用于老年人群的个体化风险评估工具,已成为ASCVD防控的迫切需求。

胰岛素抵抗(insulin resistance, IR)是ASCVD发生与进展的核心代谢异常。IR通过诱发高胰岛素血症、糖脂代谢紊乱、系统性炎症及内皮功能障碍,可直接加速动脉粥样硬化的进程^[3]。高胰岛素-正葡萄糖钳夹试验是IR定量的金标准,但其操作复杂、费用昂贵,难以在大规模临床及科研环境中推广。近年来,基于常规实验室变量计算得出的估算葡萄糖处置率(estimated glucose disposal rate, eGDR)被提出并被验证为IR的可靠替代指标^[4-6]。多项研究证实,eGDR与代谢性疾病及心血管事件风险呈明显负相关^[7-10]。然而,现有的证据多集中于普通人群或糖尿病患者,eGDR在已确诊ASCVD的老年患者中的预后价值仍缺乏系统评价。此外,老年ASCVD患者的临床表型呈现高度异质性,传统风险评分在死亡风险预测方面难以满足个体化需求^[11]。最近,人工智

能技术的快速发展为疾病风险预测提供了新的思路。机器学习(machine learning, ML)可自动捕获高维度数据中的非线性交互规律,显著提升预测模型的精度。可解释性分析,如夏普利加性解释(SHapley Additive exPlanations, SHAP)可量化特征对预测结果的边际贡献,增强模型的透明度与临床可信度^[12]。因此,本研究基于回顾性队列分析eGDR与老年ASCVD长期死亡风险的独立关系,并构建及验证纳入eGDR的多算法ML模型,再用SHAP方法解析关键的驱动因素,旨在为老年ASCVD患者的个体化风险管理提供依据。

1 资料与方法

1.1 研究对象 选取2008年1月—2022年12月在解放军总医院第二医学中心心血管内科住院或门诊就诊的5879例年龄 ≥ 60 岁且符合ASCVD标准的患者进行回顾性分析。纳入标准:(1)年龄 ≥ 60 岁;(2)临床或影像资料(CT扫描或冠状动脉造影)诊断为ASCVD,包括稳定型心绞痛、不稳定型心绞痛、心肌梗死、冠状动脉血运重建术(经皮冠状动脉介入治疗或冠状动脉旁路移植术),或缺血性卒中。排除标准:(1)缺失关键实验室变量,包括血糖、糖化血红蛋白(HbA_{1c})、血脂、肝肾功能、血常规等;(2)缺失人体测量数据,包括身高、体重、腰围等;(3)缺失死亡终点信息。最终共纳入2479例,根据eGDR四

分位数水平分为4组： Q_1 组($n=617$)、 Q_2 组($n=620$)、 Q_3 组($n=622$)、 Q_4 组($n=620$)，用于评估eGDR与死亡风险的剂量-反应关系。随后按7:3比例将参与者随机拆分为训练集($n=1735$)和验证集($n=744$)，用于机

器学习模型构建与验证。研究流程图见图1。本研究获解放军总医院医学伦理委员会审批(伦审第S2024-200-02号)。因回顾性采集，知情同意豁免，研究过程严格遵守《赫尔辛基宣言》。

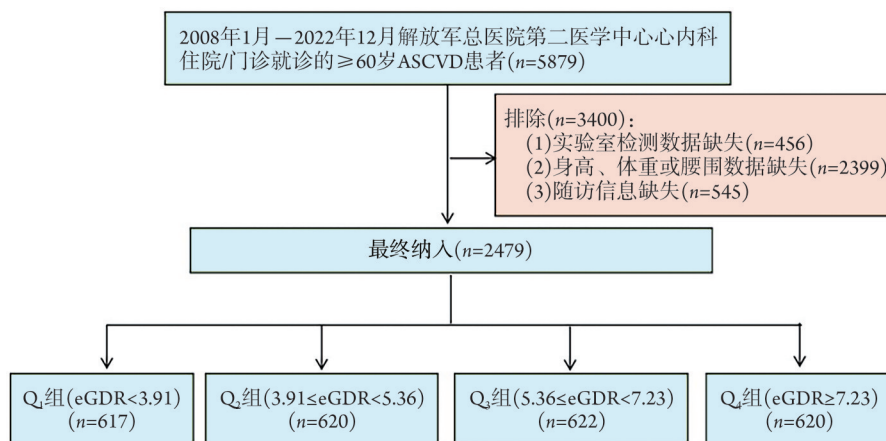


图1 老年ASCVD患者筛选及分组流程图

Fig.1 Flowchart of patient enrollment and eGDR quartile grouping in elderly patients with ASCVD

ASCVD. 动脉粥样硬化性心血管疾病; eGDR. 估算葡萄糖处置率

1.2 数据收集 查阅研究对象住院或门诊病历，收集下面的相关临床资料。(1)人口学信息及生活方式：年龄、性别、吸烟状态；(2)体格检查：身高、体重、腰围、收缩压(SBP)和舒张压(DBP)；(3)实验室指标：血清肌酐、尿酸、总胆固醇、甘油三酯、低密度脂蛋白胆固醇(LDL-C)、高密度脂蛋白胆固醇(HDL-C)、空腹血糖、HbA_{1c}、血红蛋白、白蛋白；(4)临床合并症：高血压(≥ 2 次静息SBP ≥ 140 mmHg或DBP ≥ 90 mmHg，或已应用降压药)、糖尿病(根据2021年美国糖尿病协会指南^[13]，或已应用降糖药物)、慢性阻塞性肺疾病(COPD)、慢性肾脏疾病(CKD)、估算的肾小球滤过率(eGFR) < 60 ml/(min $\cdot 1.73$ m²)、肿瘤等；(5)服用药物情况：阿司匹林、他汀类、胰岛素及口服降糖药；(6)eGDR计算：根据公式 $eGDR = 21.158 - 0.09 \times \text{腰围}(\text{cm}) - 3.407 \times \text{高血压}(\text{是}=1, \text{否}=0) - 0.551 \times \text{HbA}_{1c}(\%)$ ，eGDR水平越低表示IR越严重。

1.3 资料采集方法 人口学特征、生活方式信息、临床病史及用药情况均由训练有素的医护人员通过病史采集获得。身高、体重和腰围由医务人员按照标准化方案进行测量。血压采用经校正的台式水银血压计或电子血压计(大连欧姆龙有限公司，HEM-8102A)测量。在实验室检测方面，研究对象禁食8 h后，于入院次日或门诊就诊当日采集肘静脉血，使用罗氏全自动生化分析仪(Cobas 702)检测常规生化指标。HbA_{1c}采用高效液相色谱法(TOSOH HLC-723[®]G8)测定。所有检测项目均在解放军总医院第二医学中心临床检验科完成，由专业检验技师严格按

照临床检验操作规程执行。

1.4 研究终点及随访 研究终点为全因死亡和心血管死亡。全因死亡定义为任何原因导致的死亡。心血管死亡定义为由以下原因导致的死亡：缺血性心脏病、脑血管疾病、心力衰竭、心律失常或心源性猝死。上述事件均由两名具有主治医师及以上职称的心内科医师于医院电子病历系统中确认，仅当医师做出相同的死亡诊断时，病例才被确定纳入队列。随访起点为入院/入组日期，随访终点为发生死亡或随访终止日期。

1.5 eGDR与死亡风险的关联分析及预测模型构建

1.5.1 关联分析 采用多因素COX比例风险模型分析eGDR与全因和心血管死亡风险的关联。eGDR纳入模型的形式有3种：(1)按eGDR每增加1个标准差和1个单位作为连续变量，估计连续变化的死亡风险效应。(2)按四分位数(Q_1 - Q_4)作为分类变量，以 Q_1 为参照，分析 Q_2 、 Q_3 、 Q_4 组与参照组的死亡风险差异。(3)将 Q_1 - Q_4 组作为有序分类变量，纳入COX模型进行线性趋势检验，分析eGDR水平对死亡风险的整体变化趋势。剔除共线性显著的变量(协方差因子VIF ≥ 5)后，模型调整年龄、性别、SBP、DBP、BMI、吸烟、糖尿病、COPD、CKD、肿瘤、eGFR、血红蛋白、白蛋白、尿酸、HDL-C、LDL-C、甘油三酯、阿司匹林、他汀类、胰岛素或口服降糖药。采用限制性立方样条(RCS)分析eGDR与死亡风险的剂量-反应关系及非线性趋势。

1.5.2 预测模型构建与解释 基于ML算法建立全

因死亡风险的预测模型。首先按照7:3比例将研究人群队列随机拆分为训练集($n=1735$)与验证集($n=744$),分别用于模型训练和性能评估。利用Boruta算法和最小绝对收缩和选择算子回归(LASSO)算法从候选变量中识别关键预测因子,同时结合临床重要性进行综合判断。Boruta算法通过创建随机阴影特征作为噪声参照,迭代比较真实特征与阴影特征的重要性,筛选出重要性显著优于随机噪声的特征。设定最大迭代次数为100次,直至所有特征均被“确认”或“排除”。LASSO回归基于L1正则化将部分系数压缩至零以实现变量选择, λ 为L1正则化参数, λ 增大则惩罚增强。基于5折交叉验证的二项式偏差曲线,用于选择最优正则化参数 λ 。以 $\lambda=0.01$ (偏差最小处)为最优 λ 值,对应系数非零的变量为最终入选特征。基于筛选后的特征,在训练集中训练了7个ML模型,包括K最近邻(K-nearest neighbors, KNN),支持向量机(support vector machines, SVM),随机森林(random forests, RF),LASSO回归,朴素贝叶斯(naive bayes, NB),极限梯度提升(extreme gradient boosting, XGBoost)和决策树,模型进行5折交叉验证并通过重抽样方法优化模型超参数;最后在验证集中采用受试者工作特征(ROC)曲线下面积(AUC)、精准度、召回率和F1分数等指标评估模型性能,并采用SHAP方法量化各特征对预测结果的贡献,识别影响死亡风险的主要驱动因素。

1.6 统计学处理 采用R4.3.2软件进行统计分析。计量资料呈正态分布时以 $\bar{x}\pm s$ 表示,多组间比较采用单因素方差分析,进一步两两比较采用SNK- q 检验(方差齐时)或Dunnett T_3 法(方差不齐时);呈偏态分布时以 $M(Q_1, Q_3)$ 表示,组间比较采用Kruskal-Wallis检验;计数资料以例(%)表示,组间比较采用 χ^2 检验。双侧 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 各组研究对象的基线特征比较 研究队列总体年龄为(74.5 ± 5.6)岁,总体eGDR为 5.60 ± 2.34 ;中位随访8.1年期间,发生1101例(44.4%)全因死亡和429例(17.3%)心血管死亡病例。按eGDR水平四分位数分组后,4组间年龄、体格指标、临床疾病(高血压、糖尿病、慢性肾脏病)、实验室检查、药物治疗方面均存在明显差异($P<0.05$),呈剂量-反应趋势。其中,与 Q_1 组比较, Q_4 组患者年龄较大,BMI、腰围较小,高血压、糖尿病和CKD患病率较低,尿酸、甘油三酯水平、糖化血红蛋白较低,血红蛋白、白蛋白、总胆固醇、LDL-C、HDL-C、eGFR水平较高,阿司匹林、降糖和降脂药物的应用比例较低,差异均有统计学意义($P<0.05$),但冠心病、陈旧性心肌梗死、

血运重建史、缺血性卒中、COPD、肿瘤等病史在4组间差异无统计学意义($P>0.05$)(表1)。

2.2 eGDR与死亡风险的关联 多因素COX回归分析显示,以eGDR最低四分位数(Q_1)组为参照,最高四分位数(Q_4)组全因死亡风险降低28%($HR=0.72$, 95%CI 0.57~0.90, $P<0.001$),心血管死亡风险降低42%($HR=0.58$, 95%CI 0.44~0.78, $P<0.001$)。eGDR每增加1个单位,全因死亡风险下降4%($HR=0.96$, 95%CI 0.93~0.99, $P=0.040$),心血管死亡风险下降8%($HR=0.92$, 95%CI 0.86~0.97, $P=0.004$)(表2)。RCS分析显示,eGDR与全因死亡风险呈显著的非线性负相关($P_{\text{总体}}<0.001$, $P_{\text{非线性}}=0.004$),与心血管死亡呈显著的线性负相关($P_{\text{总体}}=0.020$, $P_{\text{非线性}}=0.308$)。亚组分析显示,eGDR与全因死亡风险的关联在 <70 岁人群较 ≥ 70 岁人群更为突出($HR: 0.83$ vs. 1.01 , $P_{\text{交互}}=0.001$);心血管死亡风险亦如此($HR: 0.71$ vs. 0.91 , $P_{\text{交互}}=0.022$)(附表S1)。

2.3 候选变量的筛选 排除共线性显著的变量后,纳入21项候选变量(后简称为特征)。在训练集内依次进行Boruta算法及LASSO回归。Boruta算法通过创建随机阴影特征作为噪声参照,经100次迭代后确认14项重要变量;LASSO回归基于L1正则化,将部分系数压缩至零以实现变量选择(系数非零者为选中变量),保留15项非零系数变量。基于上述两种特征筛选方法并考虑临床重要性和精简模型的需要,最终纳入10项特征:年龄、性别、eGFR、eGDR、LDL-C、SBP、尿酸、他汀、CKD和糖尿病(图3)。

2.4 模型训练与性能评估 在训练集内对7种算法进行5折交叉验证及超参优化后,于验证集上评价预测效能。随机森林表现出最高的预测效能($AUC=0.748$, 95%CI 0.740~0.756),随后依次是LASSO($AUC=0.739$, 95%CI 0.732~0.746)、SVM($AUC=0.738$, 95%CI 0.732~0.744)、NB($AUC=0.731$, 95%CI 0.726~0.736)、XGBoost($AUC=0.708$, 95%CI 0.700~0.716)、决策树($AUC=0.683$, 95%CI 0.676~0.690)和KNN($AUC=0.640$, 95%CI 0.630~0.650)。此外,随机森林在准确性(0.701)、F1评分(0.700)、精准度(0.689)和特异度(0.692)等指标上均表现最优(附图1、附表2)。

2.5 模型解释 由SHAP条形图中的平均SHAP值可见,各变量对模型预测风险的相对重要性依次为:年龄贡献最大,其次是LDL-C、eGFR、eGDR、SBP及尿酸(图4A)。SHAP蜂群图展示了各变量的影响方向及其强度:高龄、高LDL-C、高尿酸、男性、CKD、糖尿病以及未服用他汀类药物均与预测风险升高相关(SHAP值升高,风险增加);而高eGFR和高eGDR则与预测风险降低相关(SHAP值降低,风险降低)(图4B)。SHAP依赖图用于分析各变量的效应

表1 按eGDR水平四分位数分组的研究人群基线特征比较

Tab.1 Baseline characteristics of individuals stratified by quartiles of eGDR

项目	Q ₁ 组(n=617)	Q ₂ 组(n=620)	Q ₃ 组(n=622)	Q ₄ 组(n=620)	P
eGDR($\bar{x}\pm s$)	2.79±0.90	4.69±0.40 ⁽¹⁾	6.08±0.47 ⁽¹⁾⁽²⁾	8.83±1.14 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
年龄(岁, $\bar{x}\pm s$)	73.4±6.5	75.9±7.1 ⁽¹⁾	74.7±7.3 ⁽¹⁾⁽²⁾	74.8±7.5 ⁽¹⁾⁽²⁾	<0.001
男性[例(%)]	456(73.9)	448(72.2)	435(69.9) ⁽¹⁾	453(73.1)	<0.001
体重指数[kg/m ² , $\bar{x}\pm s$]	36.4±6.0	30.7±4.0 ⁽¹⁾	27.2±4.5 ⁽¹⁾⁽²⁾	26.2±4.5 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
腰围(cm, $\bar{x}\pm s$)	121.4±11.7	108.1±7.2 ⁽¹⁾	98.1±10.4 ⁽¹⁾⁽²⁾	95.9±11.9 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
SBP(mmHg, $\bar{x}\pm s$)	136.2±21.5	137.6±22.1	137.9±23.3	130.8±22.4 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
DBP(mmHg, $\bar{x}\pm s$)	67.2±15.3	68.7±13.8 ⁽¹⁾	68.1±15.5	66.1±14.6 ⁽²⁾⁽³⁾	0.017
吸烟情况[例(%)]					<0.001
从不吸烟	222(36.0)	231(37.3)	257(41.3) ⁽¹⁾	217(35.0) ⁽³⁾	
先前吸烟	315(51.1)	289(46.6)	232(37.3) ⁽¹⁾⁽²⁾	266(42.9) ⁽¹⁾	
当前吸烟	80(13.0)	100(16.1)	133(21.4) ⁽¹⁾⁽²⁾	137(22.1) ⁽¹⁾⁽²⁾	
临床疾病[例(%)]					
冠心病	547(88.6)	556(89.7)	570(91.6)	560(90.3)	0.209
陈旧性心肌梗死	121(19.6)	92(14.8)	108(17.4)	98(15.8)	0.101
血运重建史	83(13.5)	61(9.8)	77(12.4)	73(11.8)	0.517
缺血性卒中	159(25.8)	169(27.2)	149(23.9)	158(25.5)	0.175
高血压	615(99.7)	601(96.9) ⁽¹⁾	546(87.8) ⁽¹⁾⁽²⁾	94(15.2) ⁽¹⁾⁽²⁾⁽³⁾	<0.001
糖尿病	482(78.1)	306(49.4) ⁽¹⁾	208(33.4) ⁽¹⁾⁽²⁾	166(26.8) ⁽¹⁾⁽²⁾	<0.001
COPD	117(19.0)	116(18.7)	93(15.0)	100(16.1)	0.168
肿瘤	115(18.6)	144(23.2)	119(19.1)	118(19.0)	0.145
慢性肾脏病	220(35.7)	210(33.9)	184(29.6) ⁽¹⁾	136(21.9) ⁽¹⁾⁽²⁾⁽³⁾	<0.001
实验室检查					
血红蛋白(g/ml, $\bar{x}\pm s$)	13.8±1.6	14.0±1.6 ⁽¹⁾	13.8±1.6	14.1±1.5 ⁽¹⁾⁽³⁾	0.010
白蛋白(g/L, $\bar{x}\pm s$)	40.4±3.3	41.5±3.3 ⁽¹⁾	41.7±3.4 ⁽¹⁾	42.0±3.3 ⁽¹⁾⁽²⁾	<0.001
尿酸(mg/dl, $\bar{x}\pm s$)	6.5±1.7	6.3±1.6 ⁽¹⁾	5.9±1.5 ⁽¹⁾⁽²⁾	5.7±1.5 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
总胆固醇(mg/dl, $\bar{x}\pm s$)	180.1±46.7	181.1±44.7	184.1±47.6	188.3±45.7 ⁽¹⁾⁽²⁾	0.009
LDL-C(mg/dl, $\bar{x}\pm s$)	96.6±36.4	100.5±39.3 ⁽¹⁾	103.0±38.3 ⁽¹⁾	106.8±39.0 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
HDL-C(mg/dl, $\bar{x}\pm s$)	45.1±12.8	47.3±13.8 ⁽¹⁾	52.1±16.6 ⁽¹⁾⁽²⁾	52.4±15.5 ⁽¹⁾⁽²⁾	<0.001
甘油三酯[mg/dl, M(Q ₁ , Q ₃)]	165.0(112.0, 247.0)	144.5(100.0, 213.0) ⁽¹⁾	124.0(84.0, 181.5) ⁽¹⁾⁽²⁾	120.5(87.0, 180.0) ⁽¹⁾⁽²⁾	<0.001
eGFR(ml/min, $\bar{x}\pm s$)	66.8±22.6	67.7±22.0	70.2±22.5 ⁽¹⁾⁽²⁾	74.6±19.8 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
糖化血红蛋白(% , $\bar{x}\pm s$)	7.3±1.8	6.2±1.0 ⁽¹⁾	5.9±1.0 ⁽¹⁾⁽²⁾	5.8±0.8 ⁽¹⁾⁽²⁾⁽³⁾	<0.001
药物治疗[例(%)]					
阿司匹林	227(36.8)	205(33.1)	198(31.8)	158(25.5) ⁽¹⁾⁽²⁾⁽³⁾	<0.001
他汀类药物	400(64.8)	381(61.5)	380(61.1)	302(48.7) ⁽¹⁾⁽²⁾⁽³⁾	<0.001
胰岛素或口服降糖药	307(49.8)	176(28.4) ⁽¹⁾	112(18.0) ⁽¹⁾⁽²⁾	98(15.8) ⁽¹⁾⁽²⁾	<0.001
全因死亡[例(%)]	320(51.9)	269(43.4) ⁽¹⁾	261(42.0) ⁽¹⁾	251(40.5) ⁽¹⁾	0.009
心血管死亡[例(%)]	140(22.7)	105(16.9) ⁽¹⁾	98(15.8) ⁽¹⁾	86(13.9) ⁽¹⁾	0.022

与Q₁组比较, (1) $P<0.05$; 与Q₂组比较, (2) $P<0.05$; 与Q₃组比较, (3) $P<0.05$ 。eGDR 估计的葡萄糖处置率; SBP 收缩压; DBP 舒张压; COPD 慢性阻塞性肺疾病; LDL-C 低密度脂蛋白胆固醇; HDL-C 高密度脂蛋白胆固醇; eGFR 估算的肾小球滤过率

模式: 年龄与预测风险呈显著正相关; LDL-C 在约 100 mg/dl 处存在阈值效应, 超过该值后风险急剧上升; eGFR 在 60 ml/min 附近呈反J型曲线, 提示肾功能中度下降即可显著推高风险; eGDR 与风险呈轻度负相关, 未见明确拐点; SBP 与风险呈U型关联, 血压过高或过低均增加预测风险; 尿酸呈正J型风险

曲线, 浓度>7 mg/dl(420 μ mol/L)后风险增幅明显; 未使用他汀类药物、合并CKD或糖尿病、男性性别均与风险升高相关(附图2)。

3 讨 论

随着全球人口老龄化进程的加速, ASCVD 疾病

表2 eGDR与老年ASCVD患者全因和心血管死亡风险关联的COX回归分析

Tab.2 COX proportional hazards regression analyses for the association between eGDR and all-cause and cardiovascular mortality in elderly patients with ASCVD

变量	模型1		模型2		模型3	
	HR(95% CI)	P	HR(95% CI)	P	HR(95% CI)	P
eGDR与全因死亡风险						
eGDR 每增加1个标准差	0.90(0.85~0.96)	0.001	0.88(0.82~0.93)	<0.001	0.92(0.86~0.98)	0.040
eGDR 每增加1个单位	0.96(0.93~0.98)	0.001	0.94(0.92~0.97)	<0.001	0.96(0.93~0.99)	0.040
eGDR四分位数						
Q ₁	1(参照)		1(参照)		1(参照)	
Q ₂	0.70(0.59~0.83)		0.62(0.52~0.73)		0.65(0.54~0.76)	
Q ₃	0.72(0.61~0.85)		0.67(0.56~0.79)		0.69(0.55~0.85)	
Q ₄	0.73(0.62~0.86)		0.67(0.57~0.79)		0.72(0.57~0.90)	
P 趋势性检验	0.001		0.001		0.019	
eGDR与心血管死亡风险						
eGDR 每增加1个标准差	0.86(0.78~0.95)	0.003	0.83(0.75~0.92)	0.001	0.82(0.71~0.94)	0.004
eGDR 每增加1个单位	0.94(0.90~0.98)	0.003	0.93(0.89~0.97)	0.001	0.92(0.86~0.97)	0.004
eGDR四分位数						
Q ₁	1(参照)		1(参照)		1(参照)	
Q ₂	0.81(0.63~1.06)		0.76(0.58~0.99)		0.74(0.53, 0.97)	
Q ₃	0.78(0.60~1.01)		0.68(0.52~0.89)		0.66(0.51~0.87)	
Q ₄	0.62(0.47~0.82)		0.59(0.46~0.76)		0.58(0.44~0.78)	
P 趋势性检验	0.002		<0.001		0.009	

模型1: 未调整; 模型2: 调整年龄和性别; 模型3: 进一步调整SBP、DBP、BMI、吸烟、糖尿病、COPD、CKD、肿瘤、eGFR、血红蛋白、白蛋白、尿酸、HDL-C、LDL-C、甘油三酯、阿司匹林、他汀、胰岛素和口服降糖药。SBP.收缩压; DBP.舒张压; BMI.体重指数; COPD.慢性阻塞性肺疾病; CKD.慢性肾脏疾病; eGFR.估算的肾小球滤过率; HDL-C.高密度脂蛋白胆固醇; LDL-C.低密度脂蛋白胆固醇

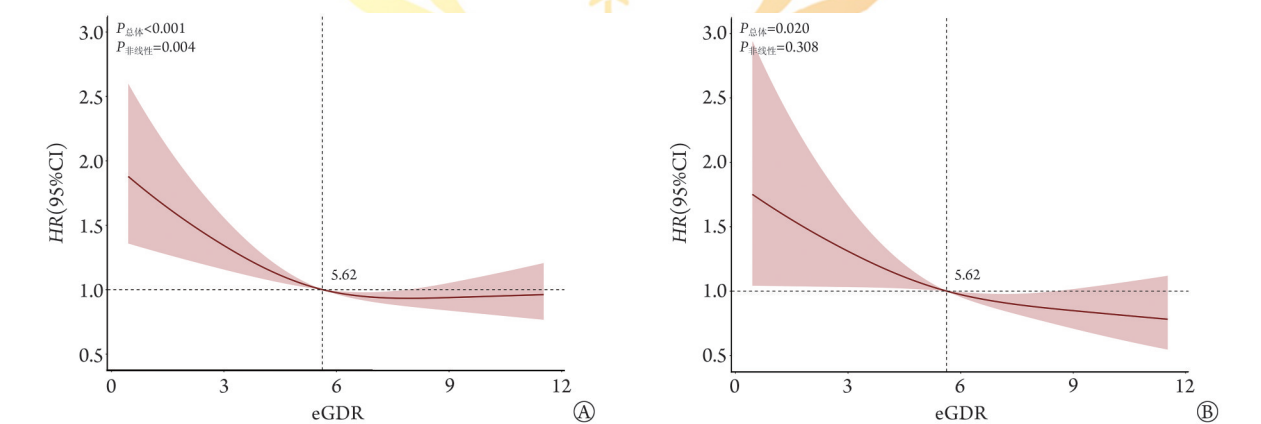


图2 eGDR与老年ASCVD患者全因和心血管死亡风险关联的限制性立方样条(RCS)曲线

Fig.2 Restricted cubic spline (RCS) curves for the association between eGDR and all-cause and cardiovascular mortality in elderly patients with ASCVD

A. 全因死亡; B. 心血管死亡。红色实线为eGDR与死亡风险的剂量-反应关系曲线, 粉色阴影为95%置信区间(CI)。调整年龄、性别、SBP、DBP、BMI、吸烟、糖尿病、COPD、CKD、肿瘤、eGFR、血红蛋白、白蛋白、尿酸、HDL-C、LDL-C、甘油三酯、阿司匹林、他汀、胰岛素和口服降糖药。SBP.收缩压; DBP.舒张压; BMI.体重指数; COPD.慢性阻塞性肺疾病; CKD.慢性肾脏疾病; eGFR.估算的肾小球滤过率; HDL-C.高密度脂蛋白胆固醇; LDL-C.低密度脂蛋白胆固醇

负担持续攀升。ASCVD是脂质代谢紊乱、慢性炎症与IR相互交织的复杂病理生理过程^[14-15]。尽管循证抗栓及强化降脂策略明显降低了ASCVD的远期不良

风险, 但残余风险仍高达70%~80%, 提示亟须挖掘新的干预靶点^[16]。IR通过对代谢、炎症和血管功能的多重损害, 在ASCVD的发生与进展中发挥关键作

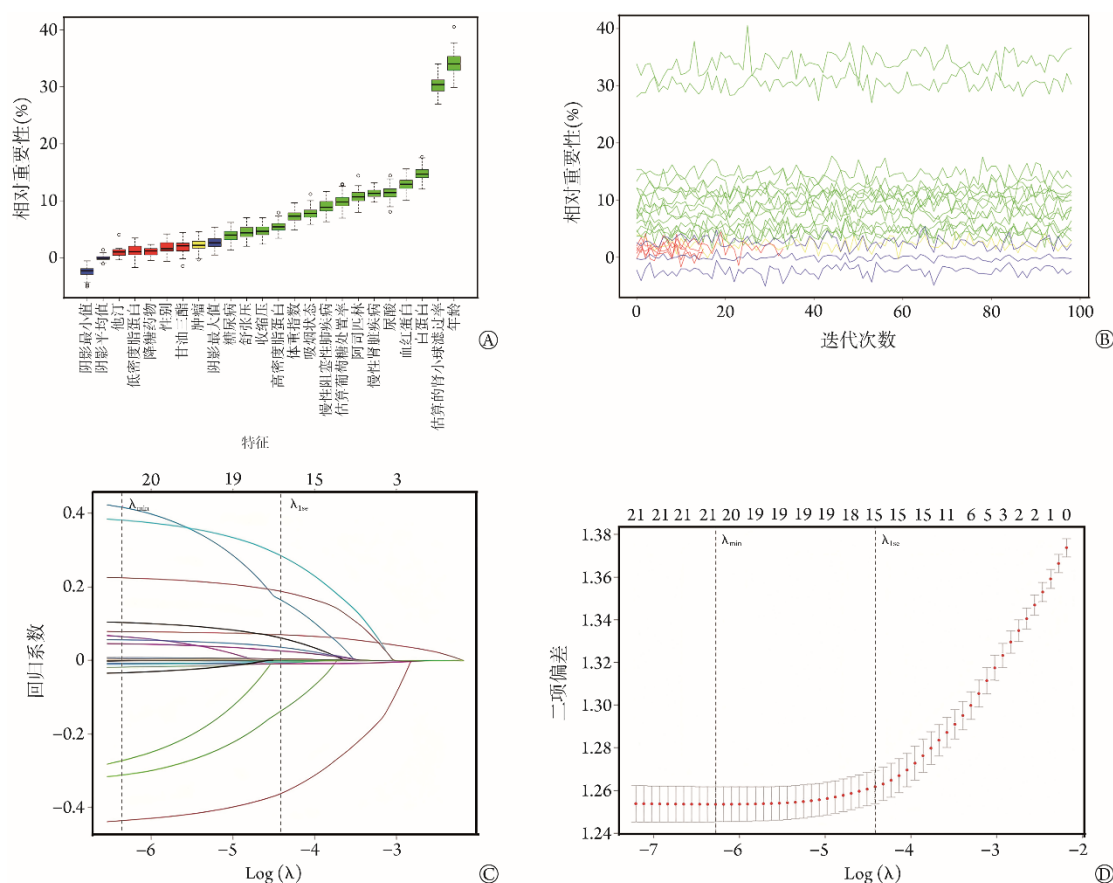


图3 基于Boruta算法和LASSO回归的老年ASCVD患者全因死亡风险预测模型特征筛选

Fig.3 Feature selection of the all-cause mortality prediction model in elderly patients with ASCVD based on Boruta algorithm and LASSO regression

A. Boruta算法特征选择, 筛选出重要性显著优于随机噪声的特征。绿色: 确认特征; 黄色: 暂定特征; 红色: 拒绝特征; 蓝色: 阴影特征; B. Boruta算法特征重要性迭代轨迹, 展示各特征重要性随迭代次数的演变; C. LASSO算法系数路径图, 展示各变量回归系数随正则化参数 λ (对数刻度)的变化轨迹。垂直虚线表示关键 λ 值: $\lambda_{\min}$ 对应预测误差最小的模型, λ_{1se} 则是按“一个标准差”原则选取的更简约模型; D. LASSO算法交叉验证偏差曲线, 基于5折交叉验证的二项式偏差曲线, 用于选择最优正则化参数 λ 。曲线为偏差均值, 误差为其标准误。垂直虚线含义与图3C一致

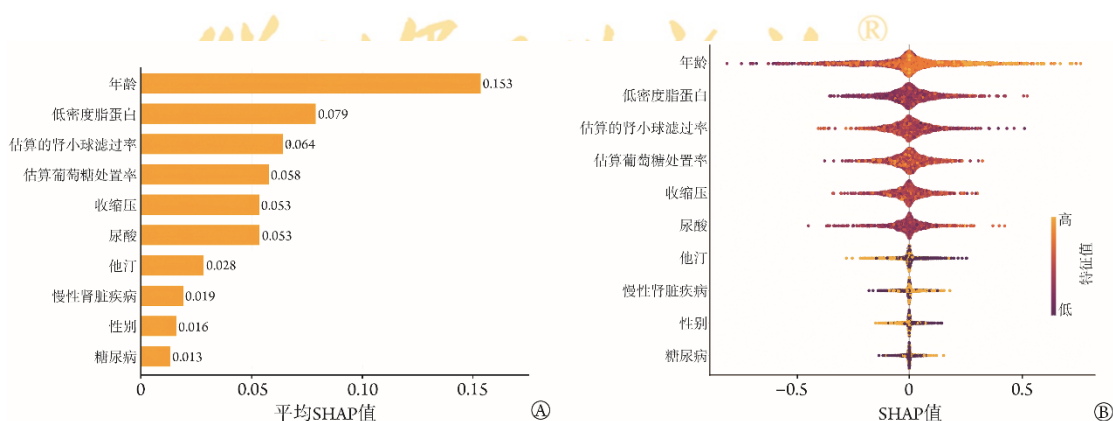


图4 随机森林模型预测老年ASCVD患者全因死亡的SHAP模型解释

Fig.4 SHAP interpretation of the random forest model for predicting all-cause mortality in elderly ASCVD patients

A. SHAP特征重要性条形图(平均绝对SHAP值越大表示贡献越大); B. SHAP值蜂群图。每个点代表一名患者, 颜色示该患者的特征实测值(黄色: 高值; 紫色: 低值)。SHAP值用于量化各特征对模型预测的贡献方向与强度, 正值提示该特征推动死亡风险增高, 负值提示该特征降低死亡风险; 例如, 年龄越高、eGFR越低、eGFR越低, 对应的SHAP值越大, 提示死亡风险越高

用,并显著增加患者的远期死亡风险。近年的研究指出,IR的可靠替代指标可早期识别代谢功能障碍相关的不良心血管结局^[17-18]。尤其是,eGDR已被证实与不良心血管结局呈独立负相关,且与远期死亡风险增加相关^[8-9]。与其他IR指标如甘油三酯-葡萄糖指数(triglyceride-glucose index, TyG)相比,eGDR对代谢综合征人群全因死亡的预测能力更优^[19]。NHANES前瞻性研究进一步显示,低eGDR可增加ASCVD患者的死亡风险^[20]。然而,上述证据多来自以欧美中年人群为主要研究对象的队列,尚缺乏针对中国老年ASCVD人群的长期调查研究。基于此,本研究在单中心回顾性中国老年ASCVD人群队列中证实,eGDR与全因和心血管死亡风险呈独立负相关;将eGDR纳入可解释ML框架后,随机森林模型的判别效能最佳(AUC=0.748);SHAP分析进一步明确了年龄、LDL-C、eGFR及eGDR为驱动全因死亡风险概率的4大关键变量。这些发现不仅验证了eGDR作为新型风险标志物的临床价值,也为老年ASCVD患者的个体化风险分层提供了新的思路。

本研究发现,eGDR与老年ASCVD患者的全因和心血管死亡均呈显著负相关。与Q₁组比较,Q₄组的全因和心血管死亡分别降低28%和42%,与先前研究结果一致^[19-20]。值得注意的是,这种负相关在<70岁亚组人群中更为突出。可能的原因包括:(1)≥70岁的亚组人群存在更多合并症及多重用药,稀释了单一代谢因素的独立效应;(2)<70岁的亚组人群对IR相关内皮炎症和交感激活的代偿能力相对较差,导致IR的病理损害被放大。性别、糖尿病、阿司匹林、他汀等亚组分析未发现显著交互作用,增强了eGDR在老年ASCVD人群中应用的普适性和可推广性。

随着大数据和算力技术的迭代进步,ML已成功用于心血管疾病的风险预测和个性化治疗。ML模型的表现高度依赖数据质量与特征工程。本研究采用Boruta与LASSO进行双重特征筛选。Boruta算法基于随机森林进行特征排序,LASSO回归采用L1-正则化惩罚使特征系数归零并逐步剔除。二者原理迥异但在本研究中的特征筛选结果高度重合,提示结果稳健可信。为兼顾临床重要性与模型精简,最终纳入10个特征进入分析。本研究在预测模型构建的算法选择上遵循互补性、数据适应性及避免过拟合的原则。7种候选算法分别擅长不同的建模任务:LASSO回归适用于线性特征筛选;KNN与SVM擅长捕捉非线性关系;随机森林、XGBoost、NB及决策树能够自动处理特征间的交互作用。上述算法在中等样本量下均表现稳健,不易过拟合。通过系统比较不同原理的算法,可客观选出最优模型,避免单一方法

偏好导致的偏倚。本研究的7种候选算法中,随机森林在验证集上判别力最优(AUC=0.748, 95%CI 0.740~0.756)。本研究不仅将ML用于ASCVD死亡风险的评估,还将eGDR这一简易IR指标嵌入ML框架,为早期识别高危患者并启动强化干预提供了依据,拓展了ASCVD残余风险管理的临床路径。

机器学习模型虽具有较高的预测性能,但其决策过程通常缺乏透明性,如同“黑箱”运行。为增强结果的临床可解释性,本研究采用SHAP分析对模型输出进行分解,将整体预测风险量化为各临床变量的独立贡献,从而使原本不可解释的模型具有临床可理解性。结果显示,对模型贡献最大的前4个变量依次为年龄、LDL-C、eGFR和eGDR,累积贡献率达66.0%,揭示了模型高度聚焦的3个关键病理环节:血脂、肾功损伤和IR。从变量作用方向看,SHAP值为正提示该变量水平升高与死亡风险增加相关(危险因素),SHAP值为负则提示与风险降低相关(保护因素),绝对值大小反映影响强度。本研究SHAP分析结果与现有临床认知高度一致。具体而言,年龄增长与死亡风险升高密切相关(SHAP正值增加);LDL-C水平升高同样增加死亡风险,且在>100 mg/dl后风险持续增加,支持“LDL-C越低越好”的管理理念在老年ASCVD中的适用性;eGFR<60 ml/min时死亡风险升高,提示慢性肾脏病是ASCVD残余风险的重要组成部分;eGDR降低(即IR加重)与死亡风险升高相关,反映了IR所致内皮损伤对死亡风险增加的促进效应;SBP与老年ASCVD呈U型关联,血压过高或过低均可增加死亡概率;尿酸>7 mg/dl后老年ASCVD患者死亡风险上升,支持高尿酸通过氧化应激、血小板聚集及内皮功能障碍加剧动脉粥样硬化;而老年ASCVD患者使用他汀、性别为女性以及非糖尿病状态,则均可持续降低风险(SHAP负值增加)。综上,SHAP分析将机器学习“黑箱”转化为可视、可量化的临床信息,揭示了IR(低eGDR)与肾功能下降、血脂异常、血压异常及高尿酸血症协同驱动老年ASCVD全因死亡的病理生理网络,为多靶点干预提供了机制依据和优先级参考。

此外,本研究联合Boruta与LASSO进行双重特征初筛,再结合临床重要性确定最终特征,避免了数据驱动模型过拟合;此外,本研究采用的SHAP方法提供了一个强大的、可解释的框架,避免了机器学习模型的“黑箱效应”,增加了模型应用的透明度。

综上所述,本研究证实,eGDR升高是老年ASCVD患者全因及心血管死亡的独立保护因素,低eGDR者死亡风险明显增加;纳入eGDR的机器学习模型展现出较好的死亡风险预测效能(AUC=0.748)。

尽管模型表现稳健,但本研究仍存在一定的局限性:(1)为单中心回顾性设计,存在潜在选择偏倚,且无法确定因果关联。(2)不能完全排除残余混杂对效应的稀释或放大。(3)所有死亡事件均从本院电子病历系统中提取,无法获取院外死亡数据,因此存在对死亡事件漏捕的可能。(4)最优模型的预测能力仅为中等,这是由于ML的预测性能是多因素依赖的,临床特征、算法差异、模型复杂度和超参数调整共同决定了最终性能。后续需纳入多模态数据(影像、组学)并扩大样本量以进一步提升区分度。(5)构建的预测模型未经过独立外部队列的验证,其泛化能力尚须进一步评估。因此,未来应开展前瞻性、多中心队列研究,并纳入更广泛的人群进行外部验证,以推动该模型向临床决策支持系统的转化。

【附加材料】

附图 1—2,附表 1—2 见 <http://dx.doi.org/10.11855/j.issn.0577-7402.2436.2026.0625FJ>。
请扫描右方二维码获取。



【参考文献】

- [1] GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021[J]. *Lancet*, 2024, 403 (10440): 2133-2161.
- [2] 胡盛寿,王增武.《中国心血管健康与疾病报告2022》概述[J]. *中国心血管病研究*, 2023, 21(7): 577-600.
- [3] Ndumele CE, Rangaswami J, Chow SL, *et al*. Cardiovascular-kidney-metabolic health: a presidential advisory from the American Heart Association[J]. *Circulation*, 2023, 148(20): 1606-1635.
- [4] Dong B, Chen Y, Yang X, *et al*. Estimated glucose disposal rate outperforms other insulin resistance surrogates in predicting incident cardiovascular diseases in cardiovascular-kidney-metabolic syndrome stages 0-3 and the development of a machine learning prediction model: a nationwide prospective cohort study[J]. *Cardiovasc Diabetol*, 2025, 24(1): 163.
- [5] Yan L, Zhou Z, Wu X, *et al*. Association between the changes in the estimated glucose disposal rate and new-onset cardiovascular disease in middle-aged and elderly individuals: a nationwide prospective cohort study in China[J]. *Diabetes Obes Metab*, 2025, 27(4): 1859-1867.
- [6] Huo G, Yao Z, Yang X, *et al*. Association between estimated glucose disposal rate and stroke in middle-aged and older Chinese adults: a nationwide prospective cohort study[J]. *J Am Heart Assoc*, 2025, 14 (9): e039152.
- [7] Zheng X, Han W, Li Y, *et al*. Changes in the estimated glucose disposal rate and incident cardiovascular disease: two large prospective cohorts in Europe and Asia[J]. *Cardiovasc Diabetol*, 2024, 23(1): 403.
- [8] Tao S, Yu L, Li J, *et al*. Insulin resistance quantified by estimated glucose disposal rate predicts cardiovascular disease incidence: a nationwide prospective cohort study[J]. *Cardiovasc Diabetol*, 2025, 24(1): 161.
- [9] Peng J, Zhang Y, Zhu Y, *et al*. Estimated glucose disposal rate for predicting cardiovascular events and mortality in patients with non-diabetic chronic kidney disease: a prospective cohort study[J]. *BMC Med*, 2024, 22(1): 411.
- [10] Zhang Z, Zhao L, Lu Y, *et al*. Insulin resistance assessed by estimated glucose disposal rate and risk of incident cardiovascular diseases among individuals without diabetes: findings from a nationwide, population based, prospective cohort study[J]. *Cardiovasc Diabetol*, 2024, 23(1): 194.
- [11] Nguyen PK, Zhao D, Okamura T, *et al*. Atherosclerotic cardiovascular disease risk prediction models in China, Japan, and Korea: implications for East Asians? [J]. *JACC Asia*, 2025, 5(3 Pt 1): 333-349.
- [12] Ali S, Akhlaq F, Imran AS, *et al*. The enlightening role of explainable artificial intelligence in medical & healthcare domains: a systematic literature review[J]. *Comput Biol Med*, 2023, 166: 107555.
- [13] American Diabetes Association. 12. Older adults: standards of medical care in diabetes-2021[J]. *Diabetes Care*, 2021, 44(Suppl 1): S168-S179.
- [14] Yarbeygi H, Farrokhi FR, Butler AE, *et al*. Insulin resistance: review of the underlying molecular mechanisms[J]. *J Cell Physiol*, 2019, 234(6): 8152-8161.
- [15] Szukiewicz D. Molecular mechanisms for the vicious cycle between insulin resistance and the inflammatory response in obesity[J]. *Int J Mol Sci*, 2023, 24(12): 9818.
- [16] Mach F, Visseren FLJ, Cater NB, *et al*. Addressing residual risk beyond statin therapy: new targets in the management of dyslipidaemias-a report from the European Society of Cardiology Cardiovascular Round Table[J]. *J Clin Lipidol*, 2024, 18(5): e685-e700.
- [17] Jiang L, Zhu T, Song W, *et al*. Assessment of six insulin resistance surrogate indexes for predicting stroke incidence in Chinese middle-aged and elderly populations with abnormal glucose metabolism: a nationwide prospective cohort study[J]. *Cardiovasc Diabetol*, 2025, 24(1): 56.
- [18] Yue Y, Li P, Sun Z, *et al*. Association of novel triglyceride-glucose-related indices with incident stroke in early-stage cardiovascular-kidney-metabolic syndrome[J]. *Cardiovasc Diabetol*, 2025, 24 (1): 301.
- [19] Chen X, Li A, Ma Q. Association of estimated glucose disposal rate with metabolic syndrome prevalence and mortality risks: a population-based study[J]. *Cardiovasc Diabetol*, 2025, 24(1): 38.
- [20] Li G, Yang W, Kuang Z, *et al*. Association of the estimated glucose disposal rate with and mortality risk in patients with atherosclerotic cardiovascular disease: a cohort study from the NHANES 1999 - 2018[J]. *Diabetes Res Clin Pract*, 2025, 225: 112263.

(责任编辑: 张小利)