

本文引用:夏桂溪,马俊杰,赵子涵,等.糖尿病并发冠心病风险预测模型的构建、评价与验证[J].牡丹江医学院学报,2024,45(3):47-54.

糖尿病并发冠心病风险预测模型的构建、评价与验证

夏桂溪¹,马俊杰¹,赵子涵²,王子婕¹,陆进^{3,4}

(1.蚌埠医学院临床医学院;2.蚌埠医学院公共卫生学院;3.蚌埠医学院人体解剖学教研室;4.数字医学与智慧健康安徽省重点实验室,安徽 蚌埠 233000)

摘要: **目的** 基于大数据筛选糖尿病(Diabetes Mellitus, DM)并发冠心病(Coronary Heart Disease, CHD)的关键因素,构建风险预测模型,为DM并发CHD患者的早期诊断和干预提供依据。**方法** 首先,采用“mice”软件包对原始数据进行数据清洗,构建新数据集并对其进行差异分析;其次,对新数据集按7:3随机分为建模组和验证组两组;再次,对建模组依次进行Lasso、单因素与多因素Logistic分析,筛选DM并发CHD的独立影响因素,并构建风险预测模型;最后,利用受试者工作特征曲线(Receiver Operating Characteristic, ROC)、决策曲线分析法(Decision Curve Analysis, DCA)、校准曲线(Calibration Curve, CC)对建模组和验证组进行内部评价和内部验证。**结果** 共纳入2 980例DM患者的临床数据(建模组2 086例;验证组894例);建模组数据采用机器学习算法筛选出的影响因素有患病种类、高血压、高脂血、动脉粥样硬化、颈动脉狭窄、脂肪肝、其他慢性肝病、心肌梗死、心功能不全及心力衰竭、心律失常、其他内分泌疾病、年龄、高密度脂蛋白胆固醇($1 \sim 1.6 \text{ mmol/L}$)、低密度脂蛋白胆固醇($0 \sim 3.4 \text{ mmol/L}$)、总蛋白、磷脂以及部分活化凝血酶原时间等,并利用其成功构建风险预测模型;该模型的内部评价与内部验证的ROC曲线下的面积(Area Under Curve, AUC)分别是0.948和0.946, DCA净获益率分别为($1\% \sim 99\%$)和($1\% \sim 100\%$),且两组的校准曲线高度一致。**结论** DM并发CHD风险预测模型具有准确性、一致性、区分性与实用性等价值,可为患者的预防、诊断与治疗提供重要依据。

关键词: 糖尿病;冠心病;风险预测;模型

中图分类号:R587.1;R542.22 文献标识码:A 文章编号:1001-7550(2024)03-0047-08

DOI:10.13799/j.cnki.mdjxyxb.2024.03.015

Analysis of associated factors of diabetes mellitus complicated with coronary heart disease and construction, evaluation and validation of risk prediction model

XIA Gui-xi et al

(Bengbu Medical College, Bengbu 233000, Anhui, China)

Abstract: **Objective** To screen the key factors of diabetes mellitus (DM) complicated with coronary heart disease (CHD) based on big data, construct a risk prediction model to provide a basis for early diagnosis and intervention of DM complicated with CHD. **Methods** Firstly, mice package was used to clean the original data, build a new data set and analyze its difference. Secondly, the new data sets were randomly grouped to modeling group and verification group according to 7:3. Thirdly, LASSO, univariate, and multivariate Logistic analysis was performed on the modeling group to screen the independent influencing factors of DM complicated with CHD, then construct the risk prediction model. Finally, use receiver operating characteristic curve (ROC), decision curve analysis (DCA), and calibration curve for internal evaluation and verification of the modeling group and verification group. **Results** Clinical data of two thousand nine hundred and eighty DM patients were included (Modeling group: 2 086 cases, Verification group: 894 cases). The influencing factors screened by the machine learning algorithm for the data of the modeling group include label, hypertension, hyperlipidemia, atherosclerosis, catheter-stenosis, fatty liver disease, chronic liver diseases, myocardial infarction, cardiac dysfunction, arrhythmias and endocrine disease, age, HDL-C, LDL-C, TP, PT and APTT were successfully used to construct risk prediction models. The AUC of internal evaluation and internal verification of this model were 0.948 and 0.946. The DCA net benefit rates were ($1\% \sim 99\%$) and ($1\% \sim 100\%$), respectively, and the calibration curves of the two groups were highly consistent. **Conclusion** The risk prediction model of DM complicated with CHD has the value of accuracy, consistency, differentiation, and practicability, which can provide an important basis for the prevention, diagnosis, and treatment of patients.

Key words: Diabetes mellitus; coronary heart disease; risk prediction; model

基金项目:蚌埠医学院厅级重点实验室开放课题基金项目(AHCM2022Z004)

作者简介:夏桂溪(2002-),女,本科在读,研究方向:大数据分析。

通讯作者:陆进,实验师,研究方向:数据挖掘,生物信息学。

DM 在世界范围内已达到流行的程度,且其患病率呈上升趋势^[1]。其发生与家族遗传、生活方式及环境密切相关,主要发病机制是胰岛素分泌紊乱或效应紊乱以及胰岛素抵抗^[2],从而导致脂代谢紊乱与凝血功能异常等出现。随着病情的进展,DM 可导致多种并发症,如心血管疾病(CVD)、肾脏疾病、视网膜病变和神经系统疾病等^[3],甚至可致残致死,具有较高的危害性,且死亡风险明显高于单一 DM,严重影响患者的健康水平和生存质量。

DM 已被证实是一种严重影响冠状动脉的危险因素,常累及心血管,从而导致 CHD 发生^[4]。据文献报道,DM 患者 CHD 增加的机制是多因素的,包括内皮功能障碍、血管壁改变、高血糖毒性、氧化应激和炎症、胰岛素抵抗^[5]。其发病原因复杂,且是 DM 最常见、最严重的一种并发症^[6]。因此,对于 DM 合并 CHD 的无症状或症状不典型患者的早期诊治至关重要。

本文依次通过 Lasso、单因素与多因素 Logistic 回归分析筛选出具有代表性的特征变量,构建风险列线图预测模型,并分别对建模组和验证组采用 DCA、ROC 及 AUC 和 CC 进行内部评价和内部验证,以判定模型的准确性和可行性,为临床 DM 合并 CHD 患者提供有价值的预测模型及诊断依据,有利于提高患者的生存质量和生活水平。

1 数据与方法

1.1 数据来源 所有分析数据来源于中国人民解放军总医院《糖尿病并发症预警数据集》^[7],主要包括一般体格检查、实验室检查及糖尿病并发症三个方面数据,最终纳入数据库样本量为 3 000 例。

1.2 数据处理 运用 R 4.2.2 软件对原始数据进行删除、填补并构建为新数据集。流程图如图 1。

1.2.1 数据剔除 (1)变量删除:剔除缺失值超过 10%的变量;(2)个案删除:剔除缺失值超过 10%的个案;(3)数据删除:剔除严重偏离均值的数值。

1.2.2 数据填补 通过“mice”软件包对缺失值进行多重回归插补法填充。

1.2.3 新数据集 整理新数据集包括 2 980 例样本,66 个变量。

1.3 差异分析 对整理后数据中的分类变量进行卡方检验;连续变量先进行正态性分布检验,符合正态分布的数据,进行 t 检验,否则利用非参数 u 检验,以比较各变量之间的差异。

1.4 随机与风险预测模型构建

1.4.1 随机分组与特征变量筛选 将新数据集按 7:3 随机分为建模组和验证组;采用 Lasso、单因素与多因素 Logistic 回归分析等机器学习算法对建模组数据进行特征变量筛选,分析 DM 并发 CHD 的风险因素。

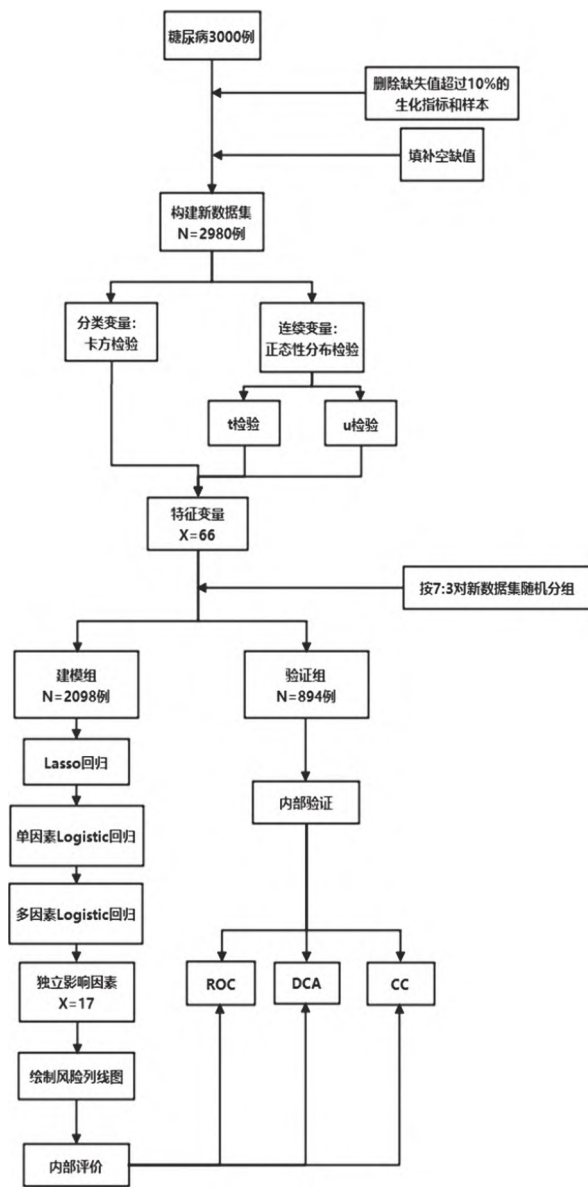


图1 流程图

1.4.2 模型构建 根据多因素 Logistic 回归分析所筛选的特征变量绘制风险列线图预测模型。

1.5 模型评价

1.5.1 内部评价 采用 ROC 及曲线下面积 AUC、DCA、校准曲线对建模组进行内部评价,检测风险预测列线图的效能:(1)AUC 是 ROC 曲线下与坐标轴围成的面积,一般取值范围在 0.5~1 之间,AUC 越接近 1.0,检测方法真实性越高,等于 0.5 时则真实性最低,无应用价值。(2)DCA 是评价临床受益的

一种评估方法,曲线越往右上角效果越好。(3)校准曲线是预测风险和实际风险的拟合线,校准曲线与预测标准曲线越贴合说明预测和实测发生率越接近,说明该模型越准确。

1.5.2 内部验证 利用验证组数据对风险预测模型采用 ROC 曲线下面积 AUC、DCA、CC 三种方式进行内部验证,以检验模型的准确性。

1.6 统计学方法 采用 R 4.2.2 软件对原始数据进行删除和填补。符合正态分布的连续变量采用 *t* 检验,否则采用非参数 *u* 检验;分类变量采用卡方检验。采用 Lasso、单因素与多因素 Logistic 回归分析影响 DM 并发 CHD 的危险因素,并利用其构建风险预测模型。利用建模组和验证组数据,采用 ROC、DCA、CC 对风险预测模型进行内部评价和内部验证。以 *P*<0.05 为差异有统计学意义。

2 结果

2.1 差异分析 差异分析发现,共筛选出 66 个显著影响 DM 合并 CHD 发病的特症指标 (*P*<0.05)。见表 1 和表 2。

表 1 分类变量差异分析

变量	分类	分组		χ^2	<i>P</i>
		有 CHD	无 CHD		
患病种类	0	869	1 113	82.613	<0.001
	1	601	373		
性别	0	726	1 256	1.357	0.244
	1	379	595		
民族	0	1 866	116	1.695	0.193
	1	929	45		
婚姻状态	0	40	1 942	0.223	0.637
	1	23	951		
高血压	0	743	1 239	95.747	<0.001
	1	191	783		
高脂血	0	1 683	299	170.263	<0.001
	1	620	354		
动脉粥样硬化	0	1 398	584	1194.641	<0.001
	1	28	946		
脑卒中	0	1 844	138	2.989	0.084
	1	888	86		
颈动脉狭窄	0	1 912	70	7.997	0.005
	1	917	57		

表 1 续表

变量	分类	分组		χ^2	<i>P</i>
		有 CHD	无 CHD		
脂肪肝	0	1 262	720	65.355	<0.001
	1	764	210		
肝硬化	0	1 945	37	3.199	0.074
	1	965	9		
其他慢性肝病	0	1 673	309	19.064	<0.001
	1	880	94		
胰腺外分泌疾病	0	1 946	36	2.168	0.141
	1	964	10		
胆道疾病	0	1 687	295	1.478	0.224
	1	846	128		
肾病	0	1 071	911	22.734	<0.001
	1	617	357		
肾衰	0	1 860	122	0.001	0.982
	1	915	59		
神经系统疾病	0	1 881	101	6.358	0.012
	1	901	73		
心肌梗死	0	1 972	10	343.319	<0.001
	1	796	178		
心功能不全及心力衰竭	0	1 924	58	153.446	<0.001
	1	824	150		
心律失常	0	1 920	62	77.974	<0.001
	1	864	110		
呼吸系统疾病	0	1 677	305	0.558	0.455
	1	813	161		
下肢动脉病变	0	1 663	319	0.047	0.829
	1	821	153		
血液病	0	1 665	317	5.576	0.018
	1	851	123		
风湿免疫疾病	0	1 915	67	0.036	0.848
	1	939	35		
妊娠哺乳期	0	1 973	9	3.066	0.08
	1	974	0		
其它内分泌疾病	0	1 242	740	37.276	<0.001
	1	721	253		
内分泌腺瘤	0	1 901	81	3.365	0.067
	1	948	26		
多囊卵巢综合征	0	1 980	2	0	1
	1	973	1		

表 1 续表

变量	分类	分组		χ^2	<i>P</i>
		有 CHD	无 CHD		
消化系肿瘤	0	1 863	119	8.634	0.003
	1	941	33		
泌尿系肿瘤	0	1 967	15	3.246	0.072
	1	959	15		
妇科肿瘤	0	1 917	65	0.402	0.526
	1	947	27		
乳腺肿瘤	0	1 973	9	1.463	0.226
	1	973	1		

表 1 续表

变量	分类	分组		χ^2	<i>P</i>
		有 CHD	无 CHD		
肺部肿瘤	0	1 951	31	0.656	0.418
	1	954	20		
颅内肿瘤	0	1 973	9	0.429	0.513
	1	967	7		
其他肿瘤	0	1 800	182	8.105	0.004
	1	915	59		

注:0=未患病;1=患病。其中未患病为参照组。

表 2 连续变量的差异分析

变量	未合并 CHD	合并 CHD	<i>P</i>	正态分布检验 <i>P</i>
年龄	55.461(44.589~66.333)	62.39(52.221~72.559)	<0.001	<0.001
身高	166.614(159.166~174.062)	166.055(158.945~173.165)	0.031	<0.001
收缩压	138.23(117.128~159.332)	139.756(118.777~160.735)	0.021	<0.001
舒张压	81.661(69.851~93.471)	78.18(66.271~90.089)	<0.001	<0.001
空腹血糖	8.334(4.416~12.252)	8.677(4.849~12.505)	0.005	<0.001
糖化血红蛋白	7.827(6.014~9.640)	7.755(6.187~9.323)	0.819	<0.001
甘油三酯	2.072(0.38~3.764)	1.962(0.416~3.508)	0.178	<0.001
总胆固醇	4.755(3.271~6.239)	4.343(3.103~5.583)	<0.001	<0.001
高密度脂蛋白胆固醇	1.083(0.752~1.414)	1.030(0.749~1.311)	<0.001	<0.001
低密度脂蛋白胆固醇	2.969(1.766~4.172)	2.622(1.643~3.601)	<0.001	<0.001
纤维蛋白原	7.117(~29.923~44.157)	9.449(~32.3~51.198)	0.076	<0.001
血尿素	7.159(2.096~12.222)	7.042(2.251~11.833)	0.413	<0.001
血肌酐	107.457(~12.239~227.153)	105.38(~14.420~225.18)	0.072	<0.001
血清尿酸	327.862(223.490~432.234)	329.485(231.893~427.077)	0.501	<0.001
血红蛋白	131.56(107.626~155.494)	132.086(110.569~153.603)	0.887	<0.001
红细胞压积(红细胞比积测定)	0.385(0.319~0.451)	0.388(0.328~0.448)	0.338	<0.001
血小板	218.844(148.476~289.212)	215.936(151.166~280.706)	0.392	<0.001
总胆红素	10.876(~1.314~23.066)	11.12(0.468~21.772)	0.009	<0.001
直接胆红素	3.568(~6.043~13.179)	3.461(~3.952~10.874)	0.01	<0.001
总蛋白	64.68(56.974~72.386)	67.088(60.464~73.712)	<0.001	<0.001
血清白蛋白	39.049(32.776~45.322)	40.085(35.184~44.986)	0.002	<0.001
乳酸脱氢酶	172.871(112.935~232.807)	175.269(106.784~243.754)	0.715	<0.001
谷丙转氨酶	24.083(~6.66~54.826)	23.873(4.809~42.937)	0.482	<0.001
谷草转氨酶	19.972(~3.489~43.433)	20.917(1.109~40.725)	0.316	<0.001
谷氨酰胺转移酶	44.7(~33.971~123.371)	37.322(~10.304~84.948)	0.885	<0.001
碱性磷酸酶	76.792(24.565~129.019)	71.409(46.36~96.458)	0.119	<0.001
凝血酶原时间	13.068(11.877~14.259)	13.252(11.943~14.561)	<0.001	<0.001

表 2 续表

变量	未合并 CHD	合并 CHD	P	正态分布检验 P
凝血酶原活动度	100.323(81.916~118.73)	96.739(77.191~116.287)	<0.001	<0.001
部分活化凝血酶原时间	36.335(31.272~41.398)	37.324(27.446~47.202)	0.004	<0.001
间接胆红素	7.308(3.053~11.563)	7.652(2.517~12.787)	0.021	<0.001
球蛋白	25.626(20.822~30.43)	27.008(22.076~31.94)	<0.001	<0.001

注:与对照组比较,*P<0.05。

2.2 随机分组 按 7 : 3 把数据随机分建模组 2 086例,验证组 894 例。

2.3 特征变量筛选与模型构建

2.3.1 Lasso 回归分析 采用 Lasso 回归分析,得到参数 λ 的最优解,获得 22 个特征变量,分别为患病种类:(0—糖尿病;1—糖尿病并发视网膜病变)、高血压、高脂血、动脉粥样硬化、颈动脉狭窄、脂肪肝、其他慢性肝病、肾病、心肌梗死、心功能不全及心力衰竭、心律失常、血液病、其它内分泌疾病、年龄、收缩压、舒张压、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、直接胆红素、总蛋白、凝血酶原时间、部分活化凝血酶原时间,见图 2。

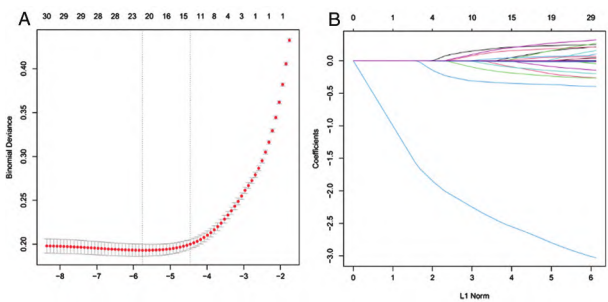


图 2 Lasso 模型选择变量

注:A 图为 Lasso 相关变量的系数曲线;B 图用于调整 Lasso 回归中的参数的交叉验证。

2.3.2 Logistic 回归分析 对 Lasso 分析筛选的变量先进行单因素 Logistic 逻辑回归,共筛选得到 20 特征变量($P<0.05$);再把其纳入到多因素 Logistic 分析,共筛选出 17 个特征变量($P<0.05$),分别为患病种类、高血压、高脂血、动脉粥样硬化、颈动脉狭窄、脂肪肝、其他慢性肝病、心肌梗死、心功能不全及心力衰竭、心律失常、其他内分泌疾病、年龄、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、总蛋白、凝血酶原时间、部分活化凝血酶原时间,见表 3 和表 4。

表 3 单因素 Logistic 逻辑回归

变量	OR	95%CI		P
		下限	上限	
患病种类	0.431	0.357	0.521	<0.001

表 3 续表

变量	OR	95%CI		P
		下限	上限	
高血压	2.178	1.76	2.694	<0.001
高脂血	3.464	2.791	4.3	<0.001
动脉粥样硬化	69.678	44.875	108.192	<0.001
颈动脉狭窄	1.627	1.041	2.542	0.032
脂肪肝	0.478	0.386	0.592	<0.001
其他慢性肝病	0.595	0.447	0.792	<0.001
肾病	0.597	0.493	0.721	<0.001
心肌梗死	43.808	21.331	89.968	<0.001
心功能不全及心力衰竭	6.659	4.562	9.719	<0.001
心律失常	3.786	2.557	5.605	<0.001
血液病	0.704	0.538	0.921	0.01
其它内分泌疾病	0.561	0.458	0.687	<0.001
年龄	1.067	1.057	1.077	<0.001
收缩压	1.003	0.998	1.007	0.213
舒张压	0.975	0.967	0.983	<0.001
高密度脂蛋白胆固醇	0.549	0.405	0.743	<0.001
低密度脂蛋白胆固醇	0.726	0.66	0.799	<0.001
直接胆红素	0.998	0.989	1.007	0.676
总蛋白	1.05	1.036	1.064	<0.001
凝血酶原时间	1.109	1.028	1.196	0.007
部分活化凝血酶原时间	1.018	1.003	1.033	0.021

表 4 多因素 Logistic 逻辑回归

变量	OR	95%CI		P
		下限	上限	
患病种类	0.474	0.34	0.661	<0.001
高血压	2.189	1.528	3.135	<0.001
高脂血	2.339	1.653	3.309	<0.001
动脉粥样硬化	95.336	56.415	161.108	<0.001
颈动脉狭窄	0.518	0.272	0.984	0.044
脂肪肝	0.404	0.288	0.566	<0.001

表 4 续表

变量	OR	95% CI		P
		下限	上限	
其他慢性肝病	0.563	0.367	0.863	0.008
肾病	0.785	0.548	1.125	0.188
心肌梗死	21.338	8.219	55.4	<0.001
心功能不全及心力衰竭	3.201	1.664	6.16	<0.001
心律失常	2.109	1.094	4.065	0.026
血液病	0.819	0.522	1.285	0.385
其它内分泌疾病	0.55	0.398	0.76	<0.001
年龄	1.059	1.041	1.077	<0.001
舒张压	0.995	0.981	1.009	0.477
高密度脂蛋白胆固醇	0.469	0.28	0.788	0.004
低密度脂蛋白胆固醇	0.813	0.708	0.934	0.004
总蛋白	1.037	1.014	1.061	0.002
凝血酶原时间	0.832	0.722	0.959	0.011
部分活化凝血酶原时间	1.052	1.015	1.09	0.005

2.3.3 风险列线图预测模型构建 利用多因素 Logistic 分析筛选出的 17 个特征变量构建风险列线图预测模型(图 3)。在风险列线图预测模型可视化图中, Terms 指的是不同的特征变量(患病种类、高血压、高脂血、动脉粥样硬化、颈动脉狭窄、脂肪肝、其他慢性肝病、心肌梗死、心功能不全及心力衰竭、心律失常、其他内分泌疾病、年龄、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、总蛋白、磷脂以及部分活化凝血酶原时间),其对应的数值均不同,然后将每一变量对应的得分相加,对应的总得分可以在下方的 Odds 上读出。因此,该模型可有助于实现临床患者的个体化预测。

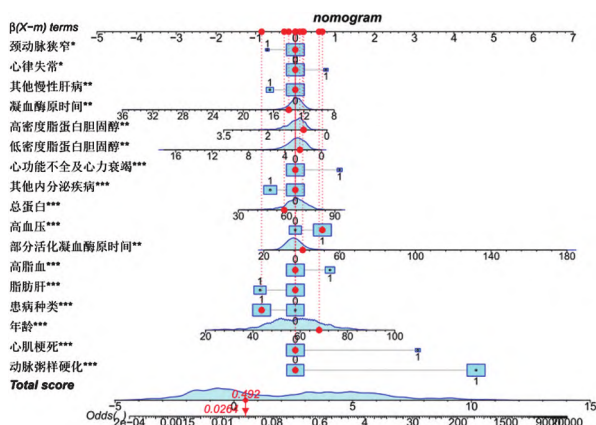


图 3 风险列线图预测模型

2.4 风险列线图预测模型的验证 分析发现,内部

评价与内部验证的 DCA 净获益率分别为(1%~99%)和(1%~100%),且均向右上角分布,显示预测模型有良好的临床净收益(图 4A~4B);内部评价与内部验证的 AUC 分别为 0.948 和 0.946(图 4C),显示该预测模型的区分能力较强,均具有显著的预测价值;内部评价与内部验证的校准曲线较贴合,揭示模型的预测概率和实际情况之间吻合度良好,显示预测模型的准确性(图 4D)。

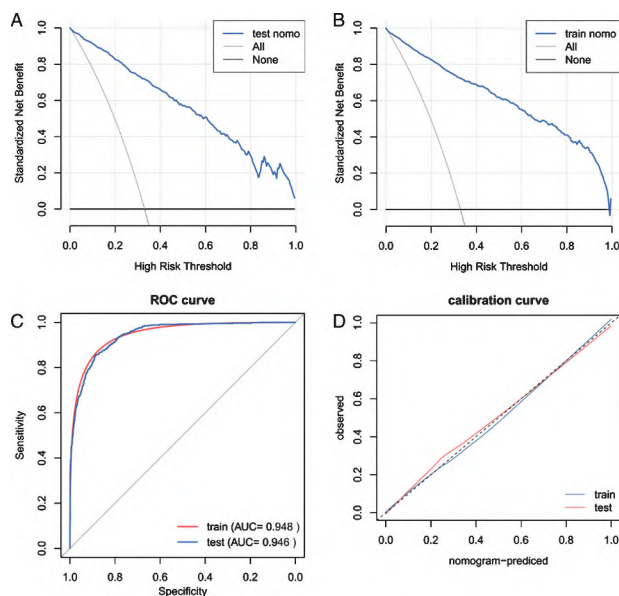


图 4 模型验证

注:A:验证组 DCA;B:建模组 DCA;C:ROC;D:CC。

3 讨论

随着中国逐步步入老龄化社会,糖尿病及心血管疾病的发病率逐年增高^[8]。根据世界卫生组织(世卫组织)和国际糖尿病联合会(IDF)的数据,2019 年 DM 患病人数达 4.63 亿人(420 万人死亡),在过去十年中迅速增长^[9]。全球 DM 患病率为 9%,预计到 2030 年将成为第七大死亡原因^[10]。DM 发病机制复杂,可累及多器官多系统,是冠状动脉疾病和其他血管并发症的主要危险因素^[11],DM 患者继发 CHD 的概率大大提高,其死亡率远远高于单一 DM 患者。因此,早期预防 DM 合并 CHD 尤为重要,然而由于 DM 患者痛阈较高,不能及时发现心绞痛症状,因此需要找到一种新型监测手段。

本研究利用 Lasso 与 Logistic 分析等机器学习算法筛选出影响 DM 合并 CHD 的特征变量,通过构建风险预测模型对 DM 合并 CHD 的发病情况进行预测,并利用内部评价与内部验证,表明该模型具有较高的准确性和区分度,显示重要的临床实用价值。研究发现,高血压是 DM 合并 CHD 的危险因

素。有研究表明,DM 患者的高血压与高血糖密切相关^[12]。高血糖常伴随血管损伤,包括内皮功能障碍、炎症反应和动脉粥样硬化;究其原因,可能是由于肾素-血管紧张素-醛固酮系统激活、晚期糖基化终末产物增加和免疫系统激活等造成的^[12-14],该损伤可影响患者的血压,因此会对 CHD 的发病有一定影响,这与本研究结论相吻合。同时,本研究还发现,动脉粥样硬化是 DM 合并 CHD 危险因素,且与非 DM 患者相比,DM 患者的冠状动脉粥样硬化可能更大^[4]。究其原因,可能是由于长期高血糖状态,刺激患者血管内皮细胞,从而促进动脉粥样硬化的形成,从而影响心血管系统的结构和功能病变。Lu^[15]等人的研究表明,遗传风险影响 DM 患者冠心病风险的主要机制就是通过增加患者动脉粥样硬化。既往研究认为,DM 患者低密度脂蛋白胆固醇水平会显示异常。低密度脂蛋白胆固醇升高可能促进 CHD 的发生,是动脉粥样硬化的危险因素^[16-17]。故血清低密度脂蛋白胆固醇可能是监测 DM 患者冠心病的潜在生物标志物,DM 患者低密度脂蛋白胆固醇水平与 CHD 发生及严重程度显著相关^[18]。有研究表明^[19],低水平高密度脂蛋白胆固醇也是动脉粥样硬化的独立危险因素,其发生机制为通过增加斑块不稳定性、促进动脉粥样硬化而诱导血栓形成,导致冠状动脉狭窄,最终增加 CHD 发生风险,这与本研究结论一致。因此,动态检测 DM 患者的血压、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇以及动脉粥样硬化等指标,可预测 DM 患者 CHD 的发病情况。

脂肪肝和其他慢性肝病同样是 DM 合并 CHD 的重要影响因素。有研究表明,DM 患者肝脂肪含量升高和肝纤维化,可能预示患者心血管风险加重,是独立于代谢综合征的风险,是 CHD 发病风险的可靠标志物^[20]。在脂肪肝和其他慢性肝病中,巨噬细胞发挥重要作用^[21]。慢性脂肪超载诱导肝细胞死亡,导致细胞释放相关危险因子,从而触发巨噬细胞激活^[22]。在慢性肝损伤期间,单核细胞来源的巨噬细胞和 Kupffer 细胞群的招募增加,导致系统炎症标志物的循环水平增加^[23]。这一过程可通过内皮功能障碍、血管张力改变、斑块形成和凝血障碍等加剧 CHD 的患病风险^[24]。此外,年龄也是 DM 合并 CHD 的独立影响因素,且随着患者年龄的增加,DM 合并 CHD 的发病风险增加。CHD 的发病年龄通常是根据访谈或医疗记录确定的,但 DM 的症状可以是渐进的或无症状,因此,很难确定 DM 的发病时间,而且年龄的影响在某种程度上与病程相关。此

外,随着患者年龄的增加,机体自身的机能变化也会对结果产生干扰^[25]。因此,年龄在预测 DM 合并 CHD 方面发挥重要作用。

本研究结果显示,凝血酶原时间与部分活化凝血酶原时间对 DM 合并 CHD 发病有重要影响,且随着凝血酶原时间的减少和部分活化凝血酶原时间的增加,患者 DM 合并 CHD 发病风险增高。有研究表明^[26],血液高凝可导致内皮功能障碍和血管并发症,而凝血酶原时间与部分活化凝血酶原时间可转化为机体凝血系统功能和凝血酶活性。CHD 的形成是血管内皮慢性炎性损伤的过程,部分活化凝血酶原时间是内源性凝血系统的筛选试验,凝血酶原时间是外源性凝血系统的筛选试验,各种因素导致的内皮下组织暴露、组织因子释放等,激活内、外源性凝血系统;考虑形成严重、复杂冠状动脉病变是血管内皮慢性损伤的炎症过程,而这些改变又会对血栓驱动的动脉粥样硬化、血管形态和心血管事件产生多重影响。因此,凝血酶原时间和部分活化凝血酶原时间在预测 DM 合并 CHD 发病方面,起到多重作用。

随着 DM 并发 CHD 在我国发病率的逐年提高,其早期诊断显得更为重要。本文基于大数据与机器学习算法,对 DM 合并 CHD 的独立影响因素展开探索,建立风险预测模型并对其进行评价与验证,结果显示该模型有较好个性化诊断价值。然而,由于数据处理方式以及数据分析方法的不同,可能与真实的临床数据存在一定的误差。此外,介于缺乏外部数据的进一步验证,模型的临床适用性还需要进一步验证。但是,本研究构建的模型仍然具有可靠的循证医学力度。

参考文献

- [1] SENYIGIT A, UZUN H, GULTEPE I, et al. The relationship between carotid intima-media thickness and serum secreted frizzled-related protein-4 and dipeptidyl peptidase-4 in diabetic patients with cardiovascular diseases[J]. Bratisl Lek Listy, 2019, 120(3): 188-194.
- [2] PETERSMANN A, MULLER-WIELAND D, MULLER U A, et al. Definition, Classification and Diagnosis of Diabetes Mellitus[J]. Exp Clin Endocrinol Diabetes, 2019, 127(S1): 1-7.
- [3] ZHENG Y, LEY S H, HU F B. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications[J]. Nat Rev Endocrinol, 2018, 14(2): 88-98.
- [4] DAS R N. Relationship between diabetes mellitus and coronary heart disease[J]. Curr Diabetes Rev, 2016, 12(3): 285-296.
- [5] WANG Y, YU Q, FAN D, et al. Coronary heart disease in type 2 diabetes: mechanisms and comprehensive prevention strategies[J]. Expert Review of Cardiovascular Therapy, 2014, 10(8): 1051-1060.

- [6] FAN R,ZHANG N,YANG L,et al.Ai-based prediction for the risk of coronary heart disease among patients with type 2 diabetes mellitus[J].Scientific Reports,2020,10(1):14457.
- [7] 中国人民解放军总医院.糖尿病并发症预警数据集.国家人口健康科学数据中心数据仓储 PHDA,2022.<https://doi.org/>.
- [8] JIMENEZ J,LAVINE K J.The dynamic role of cardiac macrophages in aging and disease[J].Curr Cardiol Rep,2022,24(8):925-933.
- [9] International Diabetes Federation [M].IDF Diabetes Atlas Ninth (IDF,Dunia,2019).
- [10] AMMOURI A A,ABU RADDHAHA A H,NATARAJAN J,et al.Perceptions of risk of coronary heart disease among people living with type 2 diabetes mellitus[J].International Journal of Nursing Practice,2018,24(1):12610.
- [11] NUSCA A,TUCCINARDI D,ALBANO M,et al.Glycemic variability in the development of cardiovascular complications in diabetes[J].Diabetes/Metabolism Research and Reviews,2018,34(8):3047.
- [12] PETRIE J R,GUZIK T J,TOUYZ R M.Diabetes,hypertension,and cardiovascular disease:clinical insights and vascular mechanisms[J].Can J Cardiol,2018,5(34):575-584.
- [13] SIBARANI M H R,WIJAYA I P,RIZKA A.Cardiovascular disease prediction model for indonesian adult population with prediabetes and diabetes mellitus;the bogor cohort study of noncommunicable diseases risk factors[J].Diabetes Metab Syndr.,2022,1(12):102330.
- [14] MARZ W,KLEBER M E,SCHARNAGL H,et al.Hdl cholesterol:reappraisal of its clinical relevance[J].Clinical Research in Cardiology,2017,106(9):663-675.
- [15] LU T,FORGETTA V,YUO H Y,et al.Polygenic risk for coronary heart disease acts through atherosclerosis in type 2 diabetes[J].Cardiovascular Diabetology,2020,19(1):12.
- [16] DURAN E K,ADAY A W,COOK N R,et al.Triglyceride-rich lipoprotein cholesterol,small dense ldl cholesterol,and incident cardiovascular disease[J].J Am Coll Cardiol,2022,17(75):2122-2135.
- [17] TALEBI S,BAGHERNIYA M,ATKIN S L,et al.The beneficial effects of nutraceuticals and natural products on small dense ldl levels,ldl particle number and ldl particle size;a clinical review[J].Lipids in Health and Disease,2020,19(1):66.
- [18] HUANG J,GU J X,BAO H Z,et al.Elevated serum small dense low-density lipoprotein cholesterol may increase the risk and severity of coronary heart disease and predict cardiovascular events in patients with type 2 diabetes mellitus[J].Disease Markers,2021(2021):5597028.
- [19] 肖丽,周军,王栋.外显子 8 基因变异及高密度脂蛋白胆固醇与老年冠心病的关系[J].中华老年心脑血管病杂志,2022,24(10):1044-1047.
- [20] BROUHA S S,NGUYEN P,BETTENCOURT R,et al.Increased severity of liver fat content and liver fibrosis in non-alcoholic fatty liver disease correlate with epicardial fat volume in type 2 diabetes;a prospective study[J].European Radiology,2018,28(4):1345-1355.
- [21] DONGIOVANNI P,PAOLINI E,CORSINI A,et al.Nonalcoholic fatty liver disease or metabolic dysfunction-associated fatty liver disease diagnoses and cardiovascular diseases:From epidemiology to drug approaches[J].Eur J Clin Invest,2021,51(7):13519.
- [22] KRENKEL O,TACKE F.Liver macrophages in tissue homeostasis and disease[J].Nature Reviews Immunology,2017,17(5):306-321.
- [23] HE Y,HWANG S,AHMED Y A,et al.Immunopathobiology and therapeutic targets related to cytokines in liver diseases[J].Cell Mol Immunol,2021,18(1):18-37.
- [24] BACK M,YURDAGUL A.Inflammation and its resolution in atherosclerosis:mediators and therapeutic opportunities[J].Nat Rev Cardiol,2019,16(7):389-406.
- [25] GAJARDO-NAVARRETE J,IBIETA G,CONCHA M,et al.Enfermedad coronaria en mujeres pre y postmenopáusicas y su relación con la diabetes mellitus tipo 2 Coronary artery disease in pre and postmenopausal women[J].Rev Med Chil,2020,148(2):145-150.
- [26] 刘晓霞,亢瑞娜.冠心病患者血浆纤维蛋白原、活化部分凝血酶时间、凝血酶原时间及血清 Ca²⁺水平的临床研究[J].现代预防医学,2012,39(14):3725-3727.

(收稿日期:2023-09-21 本文编辑:郭丽双)

(上接 46 页)

- [9] 孙福礼,吕丽,吕秀鹏,等.乳腺癌中 ER α 36、HER2 的表达与临床病理特征及预后的关系[J].大连医科大学学报,2023,45(1):7-12,18.
- [10] 帅鸽,郁义星,董佳,等.乳腺 X 线影像组学标签在预测乳腺癌 HER2 表达中的价值[J].放射学实践,2022,37(1):41-47.
- [11] MOVCHAN O V,BAGMUT I Y,SHIPKO A F,et al.HER2/positive and HER2/low in inflammatory breast cancer recurrence[J].J Med Life,2022,15(12):1573-1578.
- [12] SUNG H,FERLAY J,SIEGEL R L,et al.Global Cancer Statistics 2020;GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries[J].CA Cancer J Clin,2021,71(3):209-249.
- [13] ROSKOSKI J R.The ErbB/HER family of protein-tyrosine kinases and cancer[J].Pharmacol Res,2014(79):34-74.
- [14] SCHLESSINGER J.Ligand-induced, receptor-mediated dimerization and activation of EGF receptor[J].Cell,2002,110(6):669-672.
- [15] BEHRAVAN J,RAZAZAN A,BEHRAMAN G.Towards Breast Cancer Vaccines,Progress and Challenges[J].Curr Drug Discov Technol,2019,16(3):251-258.
- [16] QIAN L,JIEQIONG L,ZEFEI J,et al.CSCO breast cancer guideline:precise,economical and oriental[J].Sci China Life Sci,2020,63(9):1410-1412.
- [17] MEMON R,GRANADA C P,HARADA S,et al.Discordance Between Immunohistochemistry and In Situ Hybridization to Detect HER2 Overexpression/Gene Amplification in Breast Cancer in the Modern Age:A Single Institution Experience and Pooled Literature Review Study[J].Clin Breast Cancer,2022,22(1):123-133.

(收稿日期:2023-12-13 本文编辑:郭丽双)