

论著·临床研究

基于决策树模型和Logistic回归模型的 2型糖尿病发生的影响因素分析[▲]

孟继娴¹ 刘蕾² 吴薇³ 王嘉钰² 甄紫伊¹ 傅一婷¹ 马小慧¹ 孙金菊¹(1 沈阳医学院公共卫生学院,辽宁省沈阳市 110034;2 辽宁中医药大学护理学院,
辽宁省沈阳市 116600;3 沈阳医学院全科医学院,辽宁省沈阳市 110034)

【摘要】 目的 基于决策树模型和Logistic回归模型探讨2型糖尿病(T2DM)发生的影响因素,为早期筛查高危人群提供科学依据。方法 选取3 235例常规体检人群作为研究对象。应用SPSS26.0软件中决策树卡方自动交互检验和Logistic回归建立预测模型,比较两个模型的受试者工作特征曲线的曲线下面积(AUC)、灵敏度、特异度和约登指数。结果 共检出T2DM患者186例。决策树模型分析结果显示,年龄、体质指数、甘油三酯水平、收缩压、合并高血压情况和性别是T2DM发生的影响因素;Logistic回归模型分析结果显示,性别、年龄、体质指数、收缩压和甘油三酯水平是T2DM发生的影响因素;两种模型得到的结果基本一致。决策树模型预测T2DM发生风险的AUC为0.832,大于Logistic回归模型的0.800($P<0.05$),两个模型的预测性能均良好;决策树模型的特异度(0.688)、约登指数(0.537)高于Logistic回归模型(0.626、0.481),Logistic回归模型的灵敏度(0.855)高于决策树模型(0.849)。结论 本研究构建的决策树模型和Logistic回归模型均可较好地预测T2DM的发生风险,实际应用中建议将两种模型结合,最大限度地发挥两种模型的优势,为进一步预防和治疗T2DM提供参考依据。

【关键词】 2型糖尿病;影响因素;Logistic回归模型;决策树模型;预测效能**【中图分类号】** R 587.1 **【文献标识码】** A **【文章编号】** 0253-4304(2024)11-1656-06

DOI: 10.11675/j.issn.0253-4304.2024.11.04

Influencing factors for the occurrence of type 2 diabetes mellitus based on decision tree and Logistic regression models: an analytic study

MENG Jixian¹, LIU Lei², WU Wei³, WANG Jiayu², ZHEN Ziyi¹, FU Yiting¹, MA Xiaohui¹, SUN Jinju¹

(1 School of Public Health, Shenyang Medical College, Shenyang 110034, Liaoning, China; 2 Nursing School, Liaoning University of Traditional Chinese Medicine, Shenyang 116600, Liaoning, China; 3 School of General Practice, Shenyang Medical College, Shenyang 110034, Liaoning, China)

【Abstract】 **Objective** To explore the influencing factors for the occurrence of type 2 diabetes mellitus (T2DM) based on decision tree and Logistic regression models, so as to provide scientific basis for early screening of high-risk population. **Methods** A total of 3235 individuals with routine check-up were selected as the research subjects. The prediction model was established by employing the chi-square automatic interaction detector of decision tree and Logistic regression in SPSS 26.0 software. Area under the curve (AUC) of receiver operating characteristic, sensitivity, specificity, and Youden index were compared between the two models. **Results** A total of 186 patients suffered from T2DM. The results of decision tree model analysis revealed that age, body mass index, triglyceride, systolic blood pressure, concomitant hypertension, and gender were the influencing factors for the occurrence of T2DM. The results of Logistic regression model analysis indicated that gender, age, body mass index, systolic blood pressure, and triglyceride level were the influencing factors for the occurrence of T2DM. The results obtained by the two models were basically the same. AUC of

[▲]基金项目:沈阳医学院硕士研究生科技创新基金项目(Y20220515)

第一作者简介:孟继娴,硕士,研究方向为慢性病管理。

通信作者简介:刘蕾,博士,教授,研究方向为慢性病管理。

decision tree model for predicting the occurrence risk of T2DM was 0.832, which was larger than that of Logistic regression model (0.800, $P < 0.05$), and the prediction performance of the two models was favorable. The specificity (0.688) and Youden index (0.537) of decision tree model were higher than those of Logistic regression model (0.626, 0.481), and the sensitivity (0.855) of Logistic regression model was higher than that of decision tree model (0.849). **Conclusion** The models of decision tree and Logistic regression established in this study can favorably predict the occurrence risk of T2DM. In practice, it is recommended to combine the two models to maximize the advantages of the two models and provide reference basis for the further prevention and treatment of T2DM.

【Key words】 Type 2 diabetes mellitus, Influencing factors, Logistic regression model, Decision tree model, Prediction efficiency

随着全球经济的迅猛发展及人口老龄化进程的加速,糖尿病的患病率不断攀升。据2021年国际糖尿病联盟发布的数据显示,全球约有5.37亿人患有糖尿病,预计到2045年将增加46%,与2019年的数据相比,糖尿病患病率总体增加12.9%;我国糖尿病患者总数已达1.4亿,居全球首位,预计到2045年将超过1.74亿^[1]。我国糖尿病患者以2型糖尿病(type 2 diabetes mellitus, T2DM)为主,占糖尿病患者总数的90%以上^[2],其并发症发生率和死亡率较高。面对这一严峻形势,早期识别T2DM的高危人群,进而采取有针对性的预防和治疗措施显得尤为重要。

随着医学模式从经验医学到循证医学再到精准医学的不断发展,风险预测模型在疾病预防等方面发挥越来越重要的作用。预测模型是基于疾病病因的统计模型,通过寻找危险因素来预测特定人群在未来发生某种疾病的概率^[3]。目前主要采用Logistic回归模型进行影响因素分析,通过自变量与因变量之间的数量依存关系来探索疾病的危险因素^[4]。虽然Logistic回归模型能够较好地量化危险因素,但是当自变量数量较多时,Logistic回归模型无法解决自变量之间的共线性问题。决策树模型作为一种非参数统计方法,能够更好地处理数量较多的自变量之间的交互作用,弥补Logistic回归模型的不足^[5]。研究表明,联合运用决策树和Logistic回归两种模型可以实现优势互补,从而提高分析效果^[6]。因此,本研究综合决策树和Logistic回归两种模型的分析结果来探讨影响T2DM发生的影响因素,为T2DM高危人群的干预提供参考。

1 对象与方法

1.1 研究对象 选取2021年8—12月某院体检部的常规体检人群3 235例作为研究对象,对其电子病历资料进行调查分析。纳入标准:(1)年龄 ≥ 18 周岁;

(2)电子病历资料完整。排除标准:(1)因严重创伤、急性感染或其他应激情况出现短暂性血糖升高者;(2)长期服用某些影响血糖药物的患者;(3)1型糖尿病患者。本研究已通过沈阳医学院医学伦理审查(编号:SYMC-20201126-01)。

1.2 资料收集 研究人员通过查阅体检人群的电子病历,收集体检人群的社会人口学资料和实验室指标。其中,社会人口学资料包括年龄、性别、身高、体重、体质指数、收缩压、舒张压和既往疾病史,实验室指标包括甘油三酯、总胆固醇、HDL-C、LDL-C、肌酐和尿酸。

1.3 诊断标准及相关定义 (1)T2DM诊断标准参照《中国2型糖尿病防治指南(2020年版)》^[2]:典型糖尿病症状加上空腹血糖 ≥ 7.0 mmol/L,或加上随机血糖 ≥ 11.1 mmol/L,或加上OGTT 2 h血糖 ≥ 11.1 mmol/L。(2)高血压诊断标准参照《中国高血压防治指南2018年修订版》^[7]:在未使用降压药物的情况下,收缩压 ≥ 140 mmHg和(或)舒张压 ≥ 90 mmHg,或患者既往有高血压病史,目前正在使用降压药物治疗。(3)人体胖瘦程度的评价参照中国成人体质指数分类标准:体质指数=体重(kg)/身高²(m²),体质指数18.5~23.9 kg/m²为体重正常,24.0~27.9 kg/m²为超重, ≥ 28.0 kg/m²为肥胖。(4)血脂的评价参照《中国成人血脂异常防治指南(2016年修订版)》^[9]的诊断标准:甘油三酯水平 < 1.70 mmol/L为正常,总胆固醇水平 < 5.18 mmol/L为正常,LDL-C水平 < 3.37 mmol/L为正常,HDL-C水平 ≥ 1.04 mmol/L为正常。

1.4 统计学分析 应用SPSS 26.0软件建立Logistic回归模型和决策树模型。符合正态分布的计量资料以 $(\bar{x} \pm s)$ 表示,组间比较采用两独立样本 t 检验;不符合正态分布的计量资料以 $[M(P_{25}, P_{75})]$ 表示,组间比较采用秩和检验;计数资料以频数和百分率表示,组间比较采用 χ^2 检验。将单因素分析中差异有统计学意义的因素纳入多因素Logistic回归分析,建立Logistic回归

模型。采用卡方自动交互检验(chi-squared automatic interaction detector, CHAID)建立决策树模型,模型中拆分节点显著性水平 $\alpha=0.05$,最大生长深度为3层,父节点最小样本量为100,子节点为50。根据受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under the curve, AUC)、灵敏度、特异度和约登指数比较两个模型的预测性能。其中, AUC<0.5,说明模型的预测性能较差;AUC在0.5~0.7范围内,说明模型的预测性能一般;AUC在>0.7~0.9范围内,说明模型的预测性能良好;AUC>0.9,说明模型的预测性能非常好^[10]。以 $P\leq 0.05$ 为差异有统计学意义。

2 结果

2.1 研究对象的一般资料 纳入的3 235名研究对象中,男性1 843例(56.97%)、女性1 392例(43.03%),平均年龄为(45.38±13.82)岁,有高血压病史者215例(6.65%)。检出T2DM患者186例(5.75%),纳入T2DM组,其余3 049例纳入对照组。

2.2 单因素分析 T2DM组和对照组的年龄、性别、高血压病史、体质指数,以及收缩压、舒张压、总胆固醇、甘油三酯、HDL-C水平比较,差异有统计学意义($P<0.05$),见表1。

表1 3 235名体检人群发生T2DM的单因素分析结果

变量	T2DM组(n=186)	对照组(n=3049)	$t/\chi^2/z$ 值	P值
年龄($\bar{x}\pm s$,岁)	55.62±13.14	44.75±13.61	-10.930	<0.001
性别[n(%)]			23.880	<0.001
男性	138(74.19)	1705(55.92)		
女性	48(25.81)	1344(44.09)		
高血压病史[n(%)]			65.241	<0.001
是	39(20.97)	176(5.77)		
否	147(79.03)	2873(94.23)		
体质指数($\bar{x}\pm s$, kg/m ²)	26.91±4.03	25.05±3.71	6.615	<0.001
收缩压($\bar{x}\pm s$, mmHg)	145.47±19.65	128.14±18.56	-12.317	<0.001
舒张压($\bar{x}\pm s$, mmHg)	84.91±12.21	78.18±11.89	-7.488	<0.001
总胆固醇($\bar{x}\pm s$, mmol/L)	5.05±1.19	4.76±0.89	-3.258	0.001
HDL-C($\bar{x}\pm s$, mmol/L)	1.14±0.29	1.29±0.35	6.750	<0.001
LDL-C($\bar{x}\pm s$, mmol/L)	3.13±0.99	3.04±0.80	-1.238	0.217
肌酐($\bar{x}\pm s$, μ mol)	74.67±16.83	74.58±15.53	-0.078	0.938
尿酸($\bar{x}\pm s$, μ mol)	351.07±95.41	341.71±91.73	-1.340	0.180
甘油三酯[M(P_{25}, P_{75}), mmol/L]	1.90(1.26, 3.05)	1.24(0.86, 1.86)	-8.285	<0.001

2.3 多因素 Logistic 回归分析 以是否发生T2DM作为因变量(否=0,是=1),将单因素分析中差异有统计学意义的变量作为自变量(性别:女性=0,男性=1;高血压病史:否=0,是=1;其余变量以原值输入),纳入

多因素 Logistic 回归模型进行分析。结果显示,年龄、体质指数、收缩压、甘油三酯水平、性别是T2DM发生的影响因素($P<0.05$),见表2。

表2 多因素 Logistic 回归分析结果

指标	回归系数	标准误	Wald χ^2 值	OR值(95% CI)	P值
年龄	0.046	0.007	47.951	1.047(1.034, 1.061)	<0.001
体质指数	0.052	0.025	4.389	1.053(1.003, 1.106)	0.036
性别	0.543	0.199	7.411	1.721(1.164, 2.543)	0.006
高血压病史	0.133	0.232	0.327	1.142(0.724, 1.801)	0.568
收缩压	0.028	0.006	21.963	1.029(1.017, 1.041)	<0.001
舒张压	-0.013	0.009	2.099	0.987(0.969, 1.005)	0.147
总胆固醇	0.078	0.093	0.703	1.081(0.901, 1.296)	0.402
甘油三酯	0.191	0.045	18.017	1.211(1.108, 1.323)	<0.001
HDL-C	-0.671	0.354	3.599	0.511(0.256, 1.022)	0.058
常量	-9.578	1.066	80.763	0.001(—)	<0.001

2.4 决策树模型 以T2DM的发生情况为因变量,将单因素分析中差异有统计学意义的指标作为自变量(性别:女性=0,男性=1;高血压病史:否=0,是=1;其余变量以原值输入),采用CHAID构建决策树模型。最终构建的决策树模型包含3层,共19个结点,12个终端结点,共筛选出年龄、体质指数、甘油三酯、收缩压、高血压病史和性别6个影响因素,见图1。由决策树模型可知,根节点中T2DM的发生率为5.7%。模型第一层筛选出的因素是年龄,表明年龄对T2DM发生风险影响最大,其中,年龄>64岁发生T2DM的可能性最大,占该结点的15.4%。年龄>64岁且有高血压

病史的人群更有可能发生T2DM,占该结点的24.1%;年龄≤39岁且体质指数>29.753 kg/m²的人群更有可能发生T2DM,占该结点的8.7%;年龄为39~49岁且甘油三酯水平>2.780 mmol/L的人群更有可能发生T2DM,占该结点的17.3%;年龄为49~64岁,收缩压>136.0 mmHg且为男性的人群更有可能发生T2DM,占该结点的18.5%;年龄为49~64岁,收缩压≤136.0 mmHg且甘油三酯水平>2.120 mmol/L的人群更有可能发生T2DM,占该结点的12.8%。综上,年龄越大,体质指数越高、有高血压病史且收缩压、甘油三酯水平越高的男性人群更容易发生T2DM。

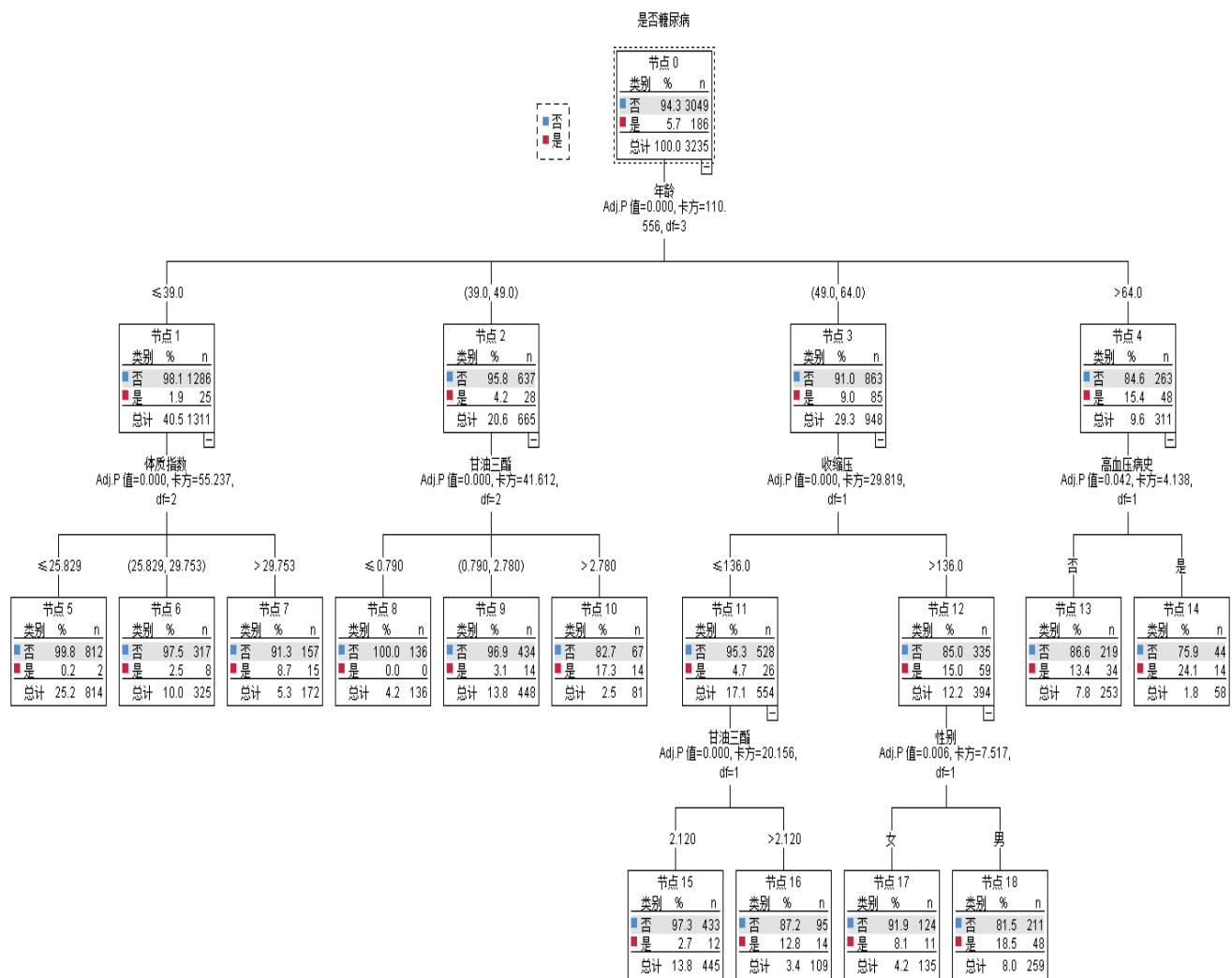


图1 体检人群发生T2DM的决策树模型

2.5 Logistic 回归模型与决策树模型预测效能的比较 ROC 曲线分析结果显示,决策树模型预测T2DM发生的AUC为0.832(95% CI: 0.808, 0.855),大于Logistic回归模型的0.800(95% CI: 0.773, 0.827),差异有统计学意义(Z=2.126, P=0.034),两个模型的AUC

均在0.7~0.9范围内,预测性能良好。决策树模型的特异度(0.688)、约登指数(0.537)高于Logistic回归模型(0.626, 0.481),Logistic回归模型的灵敏度(0.855)高于决策树模型(0.849),见图2、表3。

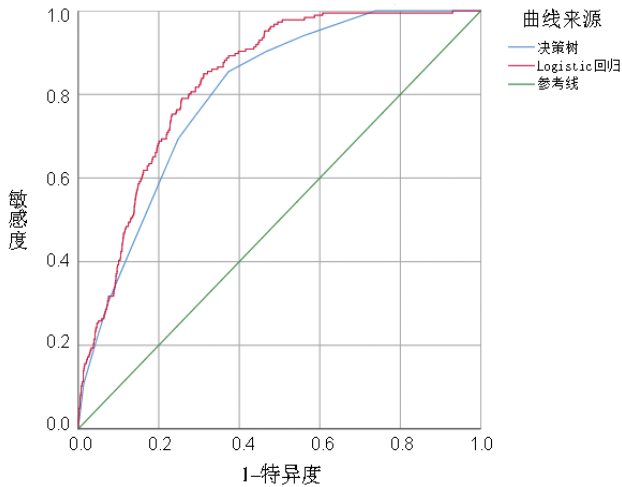


图2 Logistic回归模型与决策树模型
预测T2DM发生风险的ROC曲线图

表3 Logistic回归模型和决策树模型的预测性能比较

模型	AUC(95% CI)	灵敏度	特异度	约登 指数	SE值	P值
Logistic 回归	0.800(0.773, 0.827)	0.855	0.626	0.481	0.014	<0.001
决策树	0.832(0.808, 0.855)	0.849	0.688	0.537	0.012	<0.001

3 讨论

本研究应用决策树模型和Logistic回归模型分析T2DM发生的影响因素,两种模型得到的结果基本一致,综合两模型的分析结果,提示年龄、体质指数、甘油三酯水平、收缩压、高血压病史和性别是T2DM发生的影响因素($P<0.05$)。(1)年龄。相关调查显示,60岁以上人群T2DM的患病率超过20%,较20~30岁人群的患病率高10倍^[12],考虑可能是随着年龄的增长,人体组织细胞功能和糖类分解利用能力减弱,导致机体血糖水平升高^[11]。(2)体质指数。Yang等^[13]研究显示,超重和肥胖是T2DM发生的危险因素,超重或肥胖会引起机体产生胰岛素抵抗,胰岛 β 细胞将代偿性分泌更多的胰岛素以满足机体的需求,进而引起机体发生高胰岛素血症。随着体质指数的升高,胰岛 β 细胞代偿性分泌的胰岛素越多, β 细胞的负担越重。然而,这种代偿性的胰岛素分泌不能长期保持,胰岛 β 细胞的分泌能力会逐渐下降,最终导致T2DM的发生^[14]。(3)甘油三酯水平。陈兰波等^[15]研究表明,甘油三酯水平是T2DM的影响因素,在控制其他混杂因素的情况下,高甘油三酯水平组比低甘油三酯水平组的T2DM发生风险增加80%,这是因为甘油三酯是脂肪组织中含量丰富的一种脂质,脂肪组织是内分泌器官之一,具有影响糖脂代谢的作用^[16]。

当脂肪堆积时,脂肪组织会释放大量游离脂肪酸诱发或进一步加重胰岛素抵抗,引发T2DM^[17]。(4)收缩压。Yang等^[18]研究发现,收缩压在120~124 mmHg范围内的研究对象,发生T2DM的风险是最低的,随着收缩压的升高,T2DM的发生风险上升。这是因为血糖正常和糖尿病前期的个体发生胰岛素抵抗与收缩压之间存在线性关联,且胰岛素抵抗是T2DM发生与发展的重要原因,因此T2DM的发生与收缩压显著相关^[19-20]。(5)高血压病史。研究发现,T2DM患者的高血压患病率高于一般人群,45岁T2DM患者高血压的患病率约为40%,而75岁T2DM患者高血压的患病率为60%^[21],这是因为高血压与T2DM有相似的遗传背景、危险因素和病理基础,且二者可同时发生,互为危险因素^[22]。(6)性别。赵玉杰^[23]研究表明,相较于女性,成年男性T2DM的发生风险更高,这可能是由于女性更注重身材、控制体重,而男性较少关注体重管理,因而更容易出现超重和肥胖。同时,由于男性承受着更大的工作压力和社会心理应激,易引起胰岛素抵抗^[24],因而更容易发生T2DM。

本研究构建的决策树模型和Logistic回归模型的预测性能良好,其中决策树模型预测T2DM发生的AUC(0.832)、特异度(0.688)和约登指数(0.537)大于或高于Logistic回归模型(0.800、0.626和0.481),而Logistic回归模型的灵敏度(0.855)高于决策树模型(0.849)。两种模型各有所长,优势互补。决策树模型是一种“非参数”统计方法,其在自变量筛选过程中能够有效地体现变量之间的交互作用,消除变量间的共线性,并以树形图的方式直观地展示结果^[25]。而Logistic回归模型能够计算出危险因素的具体OR值,清晰地展示每个有意义的自变量与因变量之间的数量依存关系,但其结果展示不如决策树模型直观^[26-27]。因此,在选择较优模型时不应只考虑决策树模型和Logistic回归模型的优缺点,而应将二者结合,最大限度地发挥两种方法的优势。

综上所述,本研究构建的决策树和Logistic回归两种模型显示,年龄、体质指数、甘油三酯水平、收缩压、高血压病史和性别是T2DM发生的影响因素,且两种模型的预测效果良好,将两种模型结合能够更全面地分析T2DM的危险因素,从而为T2DM的预防和治疗提供更有针对性的建议。由于本研究采用回顾性方法来收集体检人群的基本资料和实验室指标,相关信息记录缺失或错误可能会导致在资料收集过程中产生信息偏倚。因此,今后的研究应该采

用更为严谨的研究设计和数据收集方法,如前瞻性研究、随机对照试验等,以减少信息偏倚的影响,提高研究结果的可靠性,进一步优化和完善T2DM发生风险的预测模型,为临床实践提供简便清晰的评估工具。

参 考 文 献

- [1] Sun H, Saeedi P, Karuranga S, et al. IDF diabetes Atlas : global, regional and country -level diabetes prevalence estimates for 2021 and projections for 2045[J].Diabetes Res Clin Pract, 2022, 183: 109119.
- [2] 中华医学会糖尿病学分会. 中国2型糖尿病防治指南(2020年版)[J]. 国际内分泌代谢杂志, 2021, 41(5): 482-548.
- [3] 谷鸿秋, 周支瑞, 章仲恒, 等. 临床预测模型: 基本概念、应用场景及研究思路[J]. 中国循证心血管医学杂志, 2018, 10(12): 1454-1456, 1462.
- [4] 李爽境, 黄闪闪, 刘熹润, 等. 决策树模型与logistic回归模型在代谢综合征影响因素分析中的应用——以凉山地区为例[J]. 现代预防医学, 2019, 46(13): 2319-2323, 2328.
- [5] 张 燕, 彭伶俐, 张 磊, 等. 基于logistic回归和决策树法预测颅脑肿瘤患者术后病情恶化风险[J]. 护理学杂志, 2022, 37(7): 15-19.
- [6] Dennison EM, Sayer AA, Cooper C. Epidemiology of sarcopenia and insight into possible therapeutic targets [J]. Nat Rev Rheumatol, 2017, 13(6): 340-347.
- [7] 《中国高血压防治指南》修订委员会. 中国高血压防治指南2018年修订版[J]. 心脑血管病防治, 2019, 19(1): 1-44.
- [8] 陈春明, 国际生命科学学会中国办事处中国肥胖问题工作组联合数据汇总分析协作组. 中国成人体质指数分类的推荐意见简介[J]. 中华预防医学杂志, 2001, 35(5): 349-350.
- [9] 《中国循环杂志》编辑部, 中国成人血脂异常防治指南修订联合委员会. 《中国成人血脂异常防治指南(2016年修订版)》补充说明[J]. 中国循环杂志, 2017, 32(1): 53.
- [10] Greiner M, Pfeiffer D, Smith RD. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests[J]. Prev Vet Med, 2000, 45(1/2): 23-41.
- [11] 谢 平, 黄 坚. 2型糖尿病危险因素的研究进展[J]. 应用预防医学, 2016, 22(1): 88-91.
- [12] 《中国老年型糖尿病防治临床指南》编写组. 中国老年2型糖尿病防治临床指南(2022年版)[J]. 中国糖尿病杂志, 2022, 30(1): 2-51.
- [13] Yang WY, Lu JM, Weng JP, et al. Prevalence of diabetes among men and women in China[J]. N Engl J Med, 2010, 362(12): 1090-1101.
- [14] 葛惠娜, 孙华磊, 于 斐, 等. 2型糖尿病患者BMI与胰岛β细胞功能及胰岛素抵抗的相关性研究[J]. 预防医学情报杂志, 2017, 33(8): 723-728.
- [15] 陈兰波, 蔡昕添, 顾小红. 甘油三酯与2型糖尿病关系的回顾性队列研究[J]. 重庆医科大学学报, 2022, 47(4): 417-421.
- [16] Cui J, Ma P, Sun JP, et al. The ability of baseline triglycerides and total cholesterol concentrations to predict incidence of type 2 diabetes mellitus in Chinese men and women: a longitudinal study in Qingdao, China [J]. Biomed Environ Sci, 2019, 32(12): 905-913.
- [17] Cheng C, Liu Y, Sun X, et al. Dose-response association between the the rural Chinese cohort study and meta-analysis [J]. J Diabetes, 2019, 11(3): 183-192.
- [18] Yang Q, Zheng R, Wang S, et al. Systolic blood pressure control targets to prevent major cardiovascular events and death in patients with type 2 diabetes: a systematic review and network meta-analysis [J]. Hypertension, 2023, 80(8): 1640-1653.
- [19] Quesada O, Claggett B, Rodriguez F, et al. Associations of insulin resistance with systolic and diastolic blood pressure: a study from the HCHS/SOL [J]. Hypertension, 2021, 78(3): 716-725.
- [20] Hao X, Li W, Shi R, et al. Investigating the causal mediating effect of type 2 diabetes on the relationship between traits and systolic blood pressure: a two-step Mendelian randomization study [J]. Front Endocrinol (Lausanne), 2022, 16(10): 867-890.
- [21] UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38 [J]. BMJ, 1998, 317(7160): 703-713.
- [22] 符 健, 韩耀风, 王 萌, 等. 中国居民2型糖尿病危险因素Meta分析[J]. 中国老年学杂志, 2018, 38(6): 1303-1307.
- [23] 赵玉杰. 基于Logistic回归模型分析T2DM的危险因素与应对策略[J]. 实用中西医结合临床, 2022, 22(23): 6-9.
- [24] 王 佳, 韩晓燕, 王晓宁, 等. 北京市朝阳区成年居民超重肥胖流行现状及其与主要慢性病的关系[J]. 现代预防医学, 2016, 43(24): 4425-4427, 4436.
- [25] 李 倩, 刘芸宏, 吴晓慧, 等. 基于决策树和Logistic回归预测出血性脑卒中手术后医院感染风险[J]. 中华医院感染学杂志, 2021, 31(23): 3556-3561.
- [26] 李良俊, 翟 荣, 邬闻文, 等. Logistic回归及决策树模型在CCU老年病人睡眠障碍影响因素分析中的应用[J]. 护理研究, 2022, 36(16): 2874-2879.
- [27] 刘 兵, 李 苹, 朱玖焯, 等. 决策树模型与logistic回归模型在胃癌高危人群干预效果影响因素分析中的应用[J]. 中国卫生统计, 2018, 35(1): 70-73.

(收稿日期: 2024-07-11 修回日期: 2024-09-14)