

基于贝叶斯网络的2型糖尿病患者 并发脑卒中风险预测

陈莹莹^{1,2,3}, 王鲁⁴, 胡锡峰², 朱高培^{1,2,3}, 薛付忠^{1,2,3}

(1.山东大学齐鲁医学院公共卫生学院医学数据学系, 山东 济南 250012;

2.国家健康医疗大数据研究院, 山东 济南 250003; 3.山东大学齐鲁医院, 山东 济南 250012;

4.山东健康医疗大数据管理中心, 山东 济南 250002)

摘要: **目的** 构建简便、经济且适用于临床场景的2型糖尿病(type 2 diabetes mellitus, T2DM)并发脑卒中的预测模型, 精确预测糖尿病并发脑卒中的发病风险及相关危险因素。 **方法** 基于齐鲁全生命周期电子健康研究型数据库(Cheeloo Lifespan Electronic Health Research Data-library, Cheeloo LEAD)数据, 采用单因素Cox回归分析筛选与糖尿病并发脑卒中相关的危险因素, 应用联合Cox与贝叶斯网络模型构建风险预测模型, 并从鉴别和校准两方面评价模型的预测性能。 **结果** 本研究共纳入Cheeloo LEAD中2015年1月1日至2017年10月31日的糖尿病患者15 528例, 随访至2023年1月1日, 期间发生脑卒中2 552例, 单因素分析筛选出67个与脑卒中发病有关的潜在危险因素并用以构建贝叶斯网络, 多因素Cox回归模型筛选出4个独立危险因素, 分别为年龄、短暂和突发性疾病、循环系统疾病、脑血管病后遗症。联合Cox与贝叶斯网络模型构建T2DM并发脑卒中的预测模型, 预测个体3年发生脑卒中的风险, 训练集AUC为0.814, 测试集AUC为0.816, 结果基本一致。 **结论** T2DM患者中, 年龄、短暂和突发性疾病、循环系统疾病、脑血管病后遗症是导致脑卒中风险增加的重要危险因素。在临床实践中, 应重视T2DM患者脑部病变的发生, 识别相关的危险因素, 加强监测和管理, 以减少脑卒中的发生率并改善患者的预后。

关键词: 2型糖尿病; 脑卒中; 贝叶斯网络; Cox回归模型; 风险预测

中图分类号: R730.1

文献标志码: A

A Bayesian network-based risk prediction study of stroke in patients with type 2 diabetes mellitus

CHEN Yingying^{1,2,3}, WANG Lu⁴, HU Xifeng², ZHU Gaopei^{1,2,3}, XUE Fuzhong^{1,2,3}

(1. Department of Medical Dataology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, Shandong, China; 2. National Institute of Health and Medical Big Data, Jinan 250003, Shandong, China;

3. Qilu Hospital of Shandong University, Jinan 250012, Shandong, China;

4. National Administration of Health Data, Jinan 250002, Shandong, China)

Abstract: Objective To construct a simple, economical and clinically applicable prediction model for T2DM complicated stroke, so as to accurately predict the risk of diabetes mellitus and stroke and obtain relevant risk factors.

Methods Based on the data of Cheeloo Lifespan Electronic Health Research Data-library (Cheeloo LEAD), univariate Cox regression analysis was used to screen the risk factors associated with diabetes mellitus complicated with stroke, and the combined Cox and Bayesian network models were used to construct a risk prediction model, then the prediction performance of the model was evaluated from two aspects of identification and calibration. **Results** A total of 15,528 diabetic patients in Cheeloo LEAD database from January 1, 2015 to October 31, 2017 were included in this study, and 2,552 cases of stroke occurred from January 1, 2015 to January 1, 2023. Sixty-seven potential risk factors related to

stroke were screened out by univariate analysis and used to construct Bayesian networks, and 4 independent risk factors, including age, transient and sudden-onset diseases, circulatory system diseases, and sequelae of cerebrovascular diseases, were screened by multivariate Cox regression model. Combined with the Cox model and Bayesian network model, the prediction model of T2DM complicated stroke was constructed to predict the risk of stroke in individuals in 3 years, and the AUC of the training set and test set were 0.814 and 0.816, respectively, exhibiting the basically consistent results. **Conclusion** Age, transient and sudden-onset diseases, circulatory diseases, and sequelae of cerebrovascular diseases are important risk factors for the increased risk of stroke in patients with T2DM. In clinical practice, attention should be paid to the occurrence of brain lesions in patients with T2DM. The identification of relevant risk factors and the strengthening of monitoring and management should be carried out to reduce the incidence of stroke and improve the prognosis of patients.

Key words: Type 2 diabetes mellitus; Stroke; Bayesian networks; Cox regression model; Risk prediction

糖尿病是一种较为严重的慢性代谢性疾病。目前,全球每 11 人中就有 1 人被诊断患有糖尿病,其中约 90% 的病例为 2 型糖尿病 (type 2 diabetes mellitus, T2DM)^[1]。国际糖尿病协会的调查报告显示,糖尿病发病率在我国呈现出逐年上升趋势^[2],中国的糖尿病患者人数截至 2021 年已突破 1.4 亿^[3]。而且,随着饮食和生活方式的改变,糖尿病已经成为我国的重大公共卫生问题^[4],病程延长、年龄增长或血糖控制不佳都会导致糖尿病患者的心脑血管、肾脏、眼及神经等多器官和组织受到损伤,从而引起多种并发症^[5],造成致残、致死率居高不下^[6]。

众所周知,脑卒中是 T2DM 的主要脑血管并发症,T2DM 患者存在高血压、血液高凝、血液流变学异常、血脂紊乱等多种致动脉粥样硬化的因素,容易增加 T2DM 患者心脑血管疾病风险,严重影响糖尿病患者的生活和生命质量^[7-8]。因此,早发现、早治疗对于提高 T2DM 并发脑卒中患者的生存率具有重大意义。虽然有研究探讨了 T2DM 合并脑卒中患者的相关危险因素和临床特点,但其主要集中于识别危险因素和临床表现分析^[9-10],尚缺乏适用于临床场景的 T2DM 并发脑卒中的预测模型。

因此,本研究基于齐鲁全生命周期电子健康研究型数据库 (Cheeloo Lifespan Electronic Health Research Data-library, CHEELOO LEAD) 数据,采用联合 Cox 与贝叶斯网络模型的方法,构建 T2DM 并发脑卒中的预测模型,以精确预测 T2DM 并发脑卒中的发病风险并获得相关危险因素,为 T2DM 患者防控脑卒中提供理论参考。

1 资料与方法

1.1 调查对象

本研究基于 CHEELOO LEAD 数据。Cheeloo

LEAD 是从山东省 136 个县 (区) 1.01 亿人口中,采用整群随机抽样方法抽取 39 个县 (区) 的 500 万人 (5 152 597),链接融合 2009 年 1 月 1 日至今的全生命周期居民健康档案、基本公共卫生、电子病历、健康体检、疾病监测、医疗保险记录、死因监测、环境生态等 149 个数据表而构建的大型研究型数据库。Cheeloo LEAD 以国家健康医疗大数据研究院的数据采集融汇通用数据模型 (reconciled common data model, RCDM) 和科学通用数据模型 (scientific common data model, SCDM) 为标准建立纵向队列,样本量达 5 152 597 人,总记录数 1 227 393 条,总变量数 76 712 个;依托联邦健康大数据联盟网络,进行智能化、电子化更新随访,随访时间达 13 年以上。联邦健康大数据联盟网络是一种运用联邦学习和分布式技术,旨在促进健康医疗数据共享和分析。它通过允许不同机构在不共享原始数据的情况下进行联合分析,既保护了数据隐私和合规性,又提高了医疗数据的利用效率,为跨机构医疗数据分析提供了有效的解决方案。本研究纳入 2015 年 1 月 1 日至 2017 年 10 月 31 日期间确诊的 62 432 例成年 (≥ 18 岁) T2DM 患者,T2DM 判定标准按照《中国 2 型糖尿病防治指南 (2020 年)》^[11],典型的糖尿病症状加随机血糖 ≥ 11.1 mmol/L 或空腹血糖 ≥ 7.0 mmol/L 或口服无水葡萄糖 75 g 后 2 h 血糖 ≥ 11.1 mmol/L 或糖化血红蛋白 $\geq 6.5\%$ 。对于没有典型糖尿病症状患者,则需改日再次检查;脑卒中经检查符合脑卒中医学诊断标准^[12],并已通过头颅 CT 或核磁共振成像等影像学检查手段明确诊断。排除入队列前患有以下疾病的 46 904 例患者:①脑卒中患者;②恶性肿瘤、严重感染、免疫系统疾病及严重肝、肾功能不全的患者;③资料不完整者;④死亡患者。最终共纳入 15 528 例,随访截止至 2023 年 1 月 1 日。本研究终点事件定义为脑卒中发病。

1.2 方法

1.2.1 数据预处理

对年龄等连续变量根据三分位数进行三分类处理,方便后续构建联合模型。疾病变量通过国际疾病分类编码 ICD-10 来确定,其中 T2DM 的编码为 E11,脑卒中的编码为 I60~I64、G45。将数据按照 7:3 的比例随机分配到训练集和测试集,训练集包含 10 870 例样本,用于进行 T2DM 并发脑卒中预测模型的构建;测试集包含 4 658 例样本,用于评估在训练集基础上所构建模型的表现。

表 1 变量赋值

Table 1 Variable assignment

变量类别	赋值
性别	0=男性;1=女性
年龄	0=18~45岁;1=46~73岁;2=74~100岁
疾病诊断	据 ICD-10 诊断编码有过某疾病诊断记录的记为 1,否则记为 0。如 T2DM 的诊断编码为 E11
检验	检验指标水平处于正常值范围之内记为 0,否则记为 1
用药	有过某种药物使用记录的记为 1,否则记为 0
手术操作	有过某种手术操作记录的记为 1,否则记为 0
结局	在 2015 年 1 月 1 日至 2017 年 10 月 31 日期间确诊为脑卒中的 T2DM 患者记为 1,否则记为 0

1.3 统计学处理

1.3.1 统计描述

采用 R 4.3.1 软件对数据进行统计学处理。计数资料以 $n(\%)$ 形式表示,组间比较采用 χ^2 检验。检验水准 $\alpha=0.05$ 。

1.3.2 贝叶斯网络模型

本研究使用贝叶斯网络模型构建一个由危险因素和脑卒中构成的网络。它是一种用于描述变量间不确定性因果关系的图形网络模型,由节点、有向连线和节点概率表组成。由于网络结构要求节点之间不能形成任何闭环,所以贝叶斯网络模型也被称为有向无环图(directed acyclic graph, DAG)。在贝叶斯网络中,节点表示变量,而有向边表示节点之间的关系,节点和边形成 DAG,表示为 $G=(V,A)$ 。如果用贝叶斯网络表示一组变量 x_1, x_2, \dots, x_n 隐含的条件关系,则联合概率分布 $P(x_1, x_2, \dots, x_n)$ 由所有变量的条件概率的乘积给出:

$$\theta = P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P\{x_i | PA(x_i)\}, \quad (1)$$

其中, $PA(x_i)$ 表示 x_i 的父节点变量集。贝叶斯网络包括两个阶段:结构学习和参数学习。在结构学习阶段,本研究构造了一个由节点表示变量和有向边表示条件依赖的 DAG。在有向边连接的节点对中,箭头指向的节点称为子节点,箭头起始节点称为子节点的父节点。通常,父节点的状态会影响其子节

1.2.2 变量赋值

纳入变量包括:①一般人口学特征(年龄、性别等);②基线疾病诊断信息及疾病史(哮喘、肺炎等);③基线检验指标(血常规、血糖、血脂、肝功能等);④用药信息;⑤手术信息。变量赋值见表 1。为保证研究人群的疾病诊断信息完整详尽,根据研究人群的身份号码(由国家健康医疗大数据研究院数据平台加密)补充关联了门诊和住院的疾病诊断数据以及山东省平台疾病诊断数据。以研究对象身份号码为唯一索引将上述暴露因素合并。

点的状态^[13]。在参数学习阶段,构造了一个条件概率表来反映 DAG 上节点之间的关系。使用禁忌搜索算法^[14]学习脑卒中所有潜在预测因素的网络结构,采用极大似然估计法进行参数学习,计算网络中每个节点的条件概率表。

1.3.3 联合 Cox 与贝叶斯网络模型

为了确定脑卒中的危险因素,拟合单因素和多因素 Cox 回归模型,变量选择采用逐步回归方法(入选标准 0.05,剔除标准 0.10)。利用单因素分析筛选出的危险因素构建贝叶斯网络,利用逐步回归筛选出的预测因子构建多因素 Cox 回归模型,联合 Cox 与贝叶斯网络模型构建 T2DM 并发脑卒中的预测模型。检验水准 $\alpha=0.05$ (双侧)。具体来说,为了得到条件生存贝叶斯网络的网络结构,添加多因素 Cox 回归模型的预测因子指向脑卒中患病结局节点的有向边。其中逐步回归方法筛选出的所有预测因子作为新加入的脑卒中节点的父节点集合。新加入的脑卒中节点的条件概率表用多因素 Cox 回归模型预测得到,估计 t 年肺癌累计风险的计算公式为:

$$P_{E|X}(\text{stroke} = 1 | X, T = t) = P(\text{stroke} = 1 | \Pi_E, T = t) - 1 - S_0(t)^{\exp(\sum_{i=1}^p \beta_i x_i)}, \quad (2)$$

其中, stroke 表示每个 T2DM 个体是否患脑卒中, $\text{stroke} = 1$ 表示患脑卒中, $\text{stroke} = 0$ 表示不患脑卒中;

$S_0(t)$ 代表第 t 年的基线风险函数; β_i 为对应变量 x_i 的 Cox 模型的回归系数; P 为多因素 Cox 回归模型中脑卒中的预测因子的个数, Π_E 为脑卒中节点的父节点集合。

1.3.4 模型评价

模型构建完成后,对比训练集和测试集预测的受试者工作特征(receiver operating characteristic, ROC)曲线和曲线下面积(area under the curve, AUC)及校准曲线,评估模型的性能。

2 结果

2.1 一般人口学资料

本研究共纳入 15 528 例 T2DM 患者,其中男 7 929 例(51.06%)、女 7 599 例(48.94%),平均 53 岁。为发生脑卒中的 T2DM 患者人群构建队列,平均随访时间为 4.85 年,期间共发生脑卒中 2 552 例,

发病密度为 28.43‰。

2.2 联合 Cox 与贝叶斯网络模型的构建

2.2.1 风险因素的识别

单因素回归分析筛选出 67 个与脑卒中发病风险有关的潜在预测因子,基于此构建多因素 Cox 回归模型,结果显示:与 18~45 岁人群相比,46~73 岁人群发生脑卒中的风险增加 1.773 倍(HR:2.773, 95%CI:1.816~4.246),74~100 岁人群发生脑卒中的风险增加 4.112 倍(HR:5.112, 95%CI:3.077~7.901);与未患短暂和突发性疾病人群相比,患短暂和突发性疾病人群发生脑卒中的风险增加 3.106 倍(HR:4.106, 95%CI:3.540~4.763);与未患循环系统疾病人群相比,患循环系统疾病人群发生脑卒中的风险增加 7.184 倍(HR:8.184, 95%CI:6.586~10.168);与未患脑血管病后遗症人群相比,患脑血管病后遗症人群发生脑卒中的风险增加 0.247 倍(HR:1.247, 95%CI:1.085~1.432)。见表 2~3。

表 2 脑卒中的单因素 Cox 模型分析结果/ n (%)
Table 2 Results of univariate Cox model analysis for stroke/ n (%)

因素	未患脑卒中者 ($n=12\ 976$)	脑卒中患者 ($n=2\ 552$)	P	HR	95%CI
年龄/岁					
18~45	1 662(12.81)	42(1.65)		1.000	
46~73	9 767(75.27)	1 818(71.24)	<0.001	2.773	1.816~4.246
74~100	1 547(11.92)	692(27.12)	<0.001	5.112	3.077~7.901
性别					
女	6 356(48.98)	1 243(48.71)		1.000	
男	6 620(51.02)	1 309(51.29)	<0.001	2.206	1.602~3.038
心脏瓣膜病	1 398(10.77)	641(25.12)	<0.001	3.482	2.513~4.826
肺源性心脏病	1 398(10.77)	1 253(49.10)	<0.001	6.485	4.667~9.010
梗塞和短暂性脑缺血发作	74(0.57)	56(2.19)	<0.001	1.677	1.462~1.924
肾小球疾病	106(0.82)	126(4.94)	<0.001	4.185	3.373~5.192
除脂血症外的营养障碍	667(5.14)	870(34.09)	<0.001	1.899	1.645~2.191
使用前列腺地尔	176(1.36)	80(3.13)	<0.001	2.441	1.804~3.303
使用甲钴胺	193(1.49)	213(8.35)	<0.001	1.907	1.588~2.291
高血压病	799(6.16)	232(9.09)	<0.001	2.614	2.316~2.950
短暂和突发性疾病	202(1.56)	77(3.02)	<0.001	1.564	1.345~1.819
脑血管病后遗症	193(1.49)	75(2.94)	<0.001	1.484	1.280~1.722
局部缺血性心脏病	5 183(39.94)	1 357(53.17)	<0.001	12.280	10.739~14.042
肾小管-间质疾病	2 464(18.99)	1 144(44.83)	<0.001	2.816	2.329~3.406
神经系统的其他紊乱	1 895(14.60)	792(31.03)	<0.001	1.623	1.391~1.894
阿卡波糖	1 861(14.34)	786(30.80)	<0.001	1.622	1.383~1.902
多索茶碱	1 861(14.34)	786(30.80)	<0.001	3.006	1.913~4.723
环磷腺苷葡胺	1 839(14.17)	780(30.56)	<0.001	3.318	2.115~5.205
肺心病和肺循环疾病(包括肺栓塞、肺动脉高压和肺心病)	736(5.67)	280(10.97)	<0.001	6.452	4.646~8.959
心律失常	719(5.54)	265(10.38)	<0.001	3.480	2.967~4.083
外周动脉疾病	912(7.03)	794(31.11)	<0.001	1.735	1.437~2.094

续表

因素	未患脑卒中者 (n=12 976)	脑卒中患者 (n=2 552)	P	HR	95%CI
氯化钠溶液	608(4.69)	300(11.76)	<0.001	2.614	2.316~2.950
缺血性心脏病	674(5.19)	333(13.05)	<0.001	12.184	10.660~13.926
循环系统疾病	689(5.31)	339(13.28)	<0.001	1.564	1.345~1.819
神经、神经根和神经丛疾患	956(7.37)	236(9.25)	<0.001	1.484	1.280~1.722
慢性下呼吸道疾病	3 923(30.23)	971(38.05)	<0.001	1.581	1.367~1.828
消化系统疾病	3 353(25.84)	862(33.78)	<0.001	1.337	1.184~1.510
慢性肾炎和肾病	1 705(13.14)	458(17.95)	<0.001	2.893	2.386~3.508
流行性感风和肺炎	1 215(9.36)	389(15.24)	<0.001	1.813	1.542~2.131
再生障碍和其他贫血	1 311(10.10)	404(15.83)	<0.001	2.403	2.021~2.857
使用硝苯地平	1 748(13.47)	464(18.18)	<0.001	2.375	1.824~3.093
使用银杏叶中药	1 040(8.01)	316(12.38)	<0.001	2.814	2.088~3.792
使用低分子肝素钙	806(6.21)	244(9.56)	<0.001	2.905	1.849~4.566
使用胰岛素	5 992(46.18)	1 398(54.78)	<0.001	1.645	1.413~1.916
高血压心脏病和肾脏病	6 990(53.87)	1 598(62.62)	<0.001	4.086	2.449~6.815
脊椎病	622(4.79)	381(14.93)	<0.001	1.553	1.338~1.804
周围性血管疾病	749(5.77)	80(3.13)	<0.001	1.796	1.503~2.147
肾功能衰竭	3 149(24.27)	513(20.10)	<0.001	4.886	3.890~6.138
血液和血液形成器官疾病以及某些涉及免疫机制的疾病	1 860(14.33)	268(10.50)	<0.001	1.733	1.505~1.995
使用胰激肽原酶	1 115(8.59)	143(5.60)	<0.001	2.428	1.757~3.356
使用瑞舒伐他汀	2 298(17.71)	1316(51.57)	<0.001	2.065	1.569~2.717
使用单硝酸异山梨酯	515(3.97)	40(1.57)	<0.001	2.893	2.123~3.944
神经症以及与压力有关的躯体形式疾病	91(0.70)	171(6.70)	<0.001	12.280	10.739~14.042
慢性心包疾病	1 604(12.36)	551(21.59)	<0.001	4.307	2.992~6.201
视神经和视路疾患	2 392(18.43)	660(25.86)	<0.001	1.612	1.396~1.862
肺炎	1 621(12.49)	250(9.80)	<0.001	1.815	1.543~2.134
慢性阻塞性肺病	3 394(26.16)	878(34.40)	<0.001	1.876	1.533~2.295
胸膜疾病	120(0.92)	142(5.56)	<0.001	1.895	1.583~2.269
涉及循环系统的症状和体征	1 710(13.18)	428(16.77)	<0.001	1.688	1.396~2.040
使用阿托伐他汀	1 710(13.18)	428(16.77)	<0.001	2.049	1.612~2.605
心内膜炎和心瓣膜疾患	331(2.55)	115(4.51)	<0.001	2.652	1.822~3.860
肾疾病	739(5.70)	287(11.25)	<0.001	1.578	1.365~1.824
症状和体征	4 146(31.95)	1 191(46.67)	<0.001	1.325	1.176~1.493
贫血	1 519(11.71)	420(16.46)	<0.001	2.373	2.006~2.808
使用缬沙坦	1 256(9.68)	386(15.13)	<0.001	2.335	1.767~3.086
使用阿司匹林	214(1.65)	116(4.55)	<0.001	1.884	1.575~2.254
使用呋塞米	42(0.32)	36(1.41)	<0.001	2.574	1.896~3.495
使用曲美他嗪	214(1.65)	108(4.23)	<0.001	3.839	2.436~6.050
多神经病和周围神经系统的其他疾患	157(1.21)	85(3.33)	<0.001	13.193	11.512~15.120
心包炎	251(1.93)	116(4.55)	<0.001	4.307	2.992~6.201
慢性肾脏疾病	748(5.76)	296(11.60)	<0.001	6.919	5.140~9.313
累及循环系统的症状和体征	192(1.48)	94(3.68)	<0.001	2.124	1.682~2.682
呼吸系统的其他疾病	76(0.59)	50(1.96)	<0.001	1.918	1.642~2.241
代谢紊乱	165(1.27)	93(3.64)	<0.001	1.597	1.411~1.806
肝脏疾病	28(0.22)	29(1.14)	<0.001	1.553	1.338~1.804
肺源性心脏病和肺循环疾病	51(0.39)	33(1.29)	<0.001	1.564	1.345~1.819
使用螺内酯	162(1.25)	96(3.76)	<0.001	3.276	2.184~4.913

2.2.3 联合 Cox 与贝叶斯网络模型的验证与评价

利用上述方法构建的联合模型,预测个体3年脑卒中的发病风险。训练集的 AUC 为0.814(95%CI: 0.804~0.824),最佳临界值为0.063。测试集的 AUC 为

0.816(95%CI:0.800~0.832),最佳临界值为0.060。模型在训练集和测试集的校准曲线表明预测值与观测值基本一致。见图2~3。

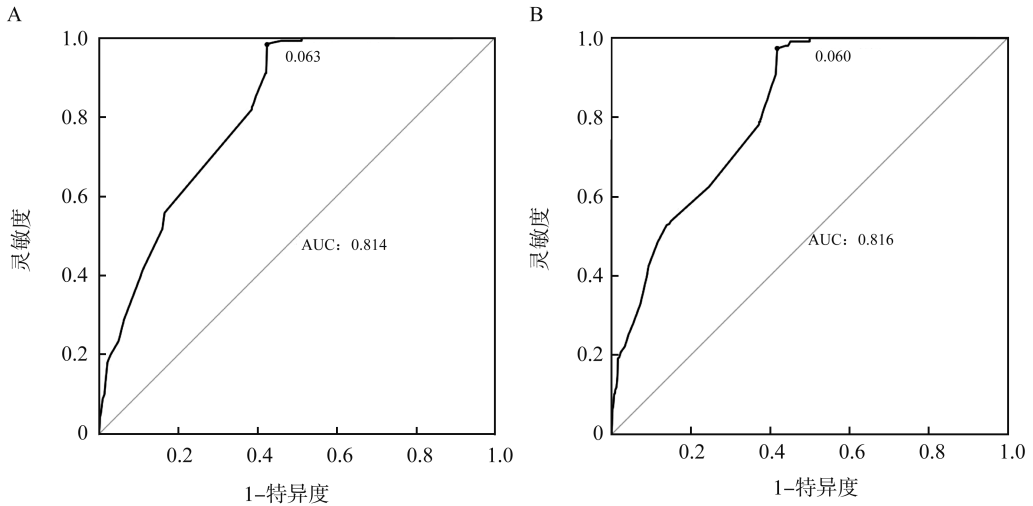


图2 脑卒中的联合 Cox 与贝叶斯网络预测模型的 ROC 曲线

A: 训练集; B: 测试集。

Figure 2 ROC curves of the combined Cox and Bayesian network prediction model for stroke

A: Training set; B: Test set.

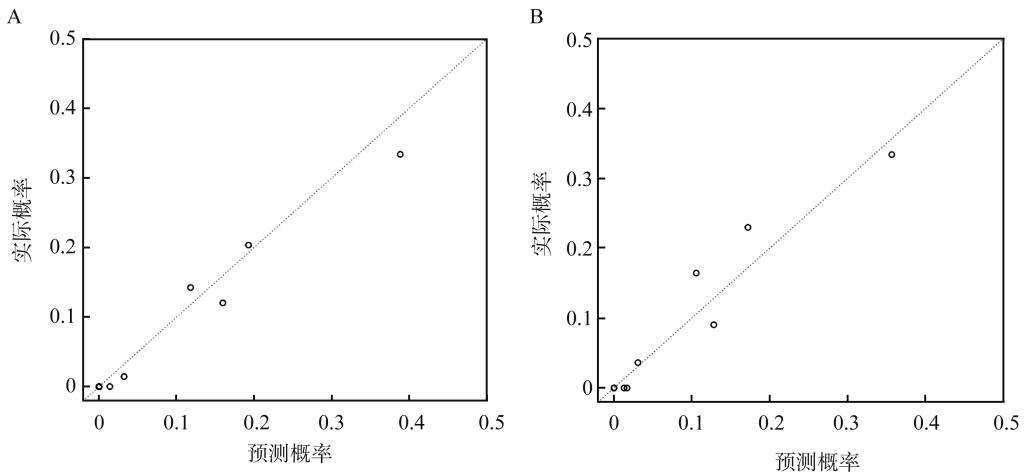


图3 脑卒中的联合 Cox 与贝叶斯网络预测模型的校准曲线

A: 训练集; B: 测试集。

Figure 3 Calibration curves of the combined Cox and Bayesian network prediction model for stroke

A: Training set; B: Test set.

3 讨论

T2DM 并发脑卒中会严重损伤患者的神经功能,造成 T2DM 患者致残甚至死亡^[15]。本研究使用山东省 15 528 例 T2DM 患者的数据,首先分析脑卒中发病的危险因素,基于所有潜在预测因子构建贝叶斯网络模型,将 Cox 模型与贝叶斯网络模型联合起来构建 T2DM 并发脑卒中的预测模型。最终 Cox 模型筛选出来的预测因子包括年龄、短暂和突

发性疾病、循环系统疾病、脑血管病后遗症。

老年 T2DM 患者的脑卒中发病率是正常人群的 2~4 倍^[16],主要发病表现为头痛、消瘦、口渴等,部分患者会因出现昏迷、偏瘫造成误诊,增加患者死亡风险^[17-18],原因可能是随着年龄的增长,老年人的生理机能逐渐下降,各组织器官功能逐渐退化,各种因素导致血管内皮损伤,加速了脑血管病的发生^[19]。不仅如此,年龄较大的 T2DM 患者更容易出现认知功能障碍^[20]及更多的慢性并发症,如高血压、高血脂和肾脏病变等^[21],这些都会进一步加剧

脑卒中的发病风险。

常见的循环系统疾病(也被称为心血管疾病)包括高血压、冠心病、心律失常、心力衰竭等。糖尿病是一种慢性代谢障碍性疾病,主要特征是持续性的高血糖状态^[22],长期的血糖控制不佳可引起微循环异常,导致各种并发症的发生,尤其是眼病、肾病、糖尿病足、神经病变、脑卒中等^[23]。此外,循环系统疾病中的高血压、冠心病和动脉粥样硬化等是脑卒中的主要危险因素之一,这些疾病通过影响脑部的血液供应,增加了脑卒中的发病风险^[24-25]。

短暂和突发性疾病是神经系统疾病中的一个分类,这类疾病可能包括但不限于短暂性脑缺血发作(transient ischemic attack, TIA)、偏头痛等。研究表明 T2DM 患者在确诊时所伴有的血管病变和神经病变是导致脑卒中发病的危险因素^[26], T2DM 患者的血糖波动会影响 T2DM 患者并发脑卒中的神经功能预后^[27],这些研究结果证实 T2DM 患者的神经系统受损是脑卒中的一个重要风险因素。然而,关于短暂和突发性疾病直接导致 T2DM 患者发生脑卒中的证据较少,阵发性神经系统疾病的症状可能与脑卒中有一定的关联,但这种关联更多地是通过糖尿病的整体病理生理机制来间接影响的。

脑血管病包括一系列与脑血管相关的疾病,这些疾病通过影响血管健康,与脑卒中的发生产生显著关联,增加了脑卒中的发病风险^[28-30]。并可能导致脑卒中后严重的并发症,如认知障碍^[31]。多项研究显示,伴发脑血管病危险因素对急性期和康复期脑卒中认知障碍的影响较大^[32]。因此,应密切监测糖尿病患者的健康状况,并采取适当的诊断和治疗措施,以减少并发症的风险和影响。

本研究构建的联合预测模型优势在于:①传统的 Cox 回归模型通常假设数据是完整的,但在现实应用中,数据往往不完整,缺失数据可能导致参数估计出现偏差或降低统计效率。贝叶斯网络可处理不完全数据。它通过推断缺失值的可能分布来弥补缺失数据,避免了对缺失数据的直接删除或填充。将贝叶斯网络与 Cox 回归结合,可以在模型中同时考虑缺失数据的影响,从而提高模型在缺失数据条件下的准确性和鲁棒性。②Cox 回归模型假设协变量与生存时间之间存在线性关系,且未考虑协变量之间的相互依赖。贝叶斯网络可以表示复杂的依赖关系,并且能够捕捉协变量之间的非线性和条件独立性,这对于多变量分析尤其重要。通过结合贝叶斯网络,能够扩展其对协变量间复杂关系的建模能力,尤其是当协变量之间存在非线性或高阶依赖时。③

Cox 回归依赖于点估计,通常没有直接处理模型参数的不确定性。贝叶斯网络则可以通过概率分布而非单一点估计来描述不确定性,使得结果更加灵活,能够提供关于结果的概率性解释。将贝叶斯网络与 Cox 回归结合,可以有效地处理和量化模型中的不确定性,为预测结果提供概率性解释,从而使决策更加可靠。④Cox 回归模型本身是半参数的,可以灵活地处理生存数据,但对于复杂的协变量关系建模能力却是有限的。贝叶斯网络是一种 DAG,可以用于表达更加复杂的因果关系,并允许通过学习来自动调整模型结构。结合 Cox 回归与贝叶斯网络,能够在处理生存数据的同时,增加模型的灵活性和扩展性,适应更多样化的现实数据情况。

本研究尚存在不足之处:①本研究在数据处理过程中,将连续变量年龄按照三分位数进行了三分类处理。这种分类方法虽然有助于简化数据结构,便于分析不同年龄段的特征,但它不可避免地导致了原始连续数据中信息的丢失,可能会对结果的全面性和精确性产生一定影响。②仅侧重临床诊疗数据作为数据来源,尚未进一步补充考虑患者的日常生活习惯、环境因素、遗传背景等方面的影响因素。

综上所述,本研究在真实世界大数据背景下,发现山东省 T2DM 患者并发脑卒中的主要影响因素为年龄、短暂和突发性疾病、循环系统疾病、脑血管病后遗症。因此,在临床实践中,应重视 T2DM 患者脑卒中的发生,可考虑根据以上因素给予患者针对性干预措施,加强危险因素的监测和管理,减少脑卒中的发病并改善患者预后。除此之外,对于 T2DM 患者,应加强早期监测,一旦发现脑卒中迹象,应立即采取有效治疗措施,以期达到最佳治疗效果,提高患者的生存率和生活质量^[33]。

参考文献:

- [1] Bellary S, Kyrou I, Brown JE, et al. Type 2 diabetes mellitus in older adults: clinical considerations and management[J]. Nat Rev Endocrinol, 2021, 17(9): 534-548.
- [2] 李建辉,徐波.老年人Ⅱ型糖尿病合并脑卒中危险因素分析[J].中国实用医药,2007,2(16):61-62.
LI Jianhui, XU Bo. Analysis of risk factors of type 2 diabetes mellitus complicated with stroke in the elderly[J]. China Practical Medicine, 2007, 2(16): 61-62.
- [3] Sun H, Saedi 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. doi: 10.1016/j.diabres.2021.109119
- [4] Graham H, White PL. Social determinants and lifestyles; integrating environmental and public health perspectives [J]. Public Health, 2016, 141: 270-278. doi:10.1016/j.puhe.2016.09.019
- [5] Kolarić V, Svirčević V, Bijuk R, et al. Chronic complications of diabetes and quality of life [J]. Acta Clin Croat, 2022, 61(3): 520-527.
- [6] 王富军, 王文琦. 《中国老年 2 型糖尿病防治临床指南(2022 年版)》解读 [J]. 河北医科大学学报, 2022, 43(12): 1365-1370.
WANG Fujun, WANG Wenqi. Interpretation of clinical guidelines for the prevention and treatment of type 2 diabetes in the elderly in China (2022 edition) [J]. Journal of Hebei Medical University, 2022, 43(12): 1365-1370.
- [7] 朱婧, 罗彩凤, 倪益益, 等. 2 型糖尿病病人并发缺血性脑卒中急性期血糖波动趋势及影响因素分析 [J]. 安徽医药, 2019, 23(12): 2395-2399.
ZHU Jing, LUO Caifeng, NI Yiyi, et al. Analysis of the tendency of blood glucose fluctuation and its influencing factors in type 2 diabetic patients with acute ischemic stroke [J]. Anhui Medical and Pharmaceutical Journal, 2019, 23(12): 2395-2399.
- [8] Zhou ZE, Lindley RI, Rådholm K, et al. Canagliflozin and stroke in type 2 diabetes mellitus [J]. Stroke, 2019, 50(2): 396-404.
- [9] 许俊杰, 陆霞, 张林燕, 等. 嘉兴市城南社区 2 型糖尿病患者并发脑卒中患病状况及影响因素调查分析 [J]. 实用预防医学, 2022, 29(11): 1358-1361.
- [10] 吴云虹, 林雅明, 符薇薇, 等. 糖尿病患者并发急性缺血性脑卒中的影响因素研究 [J]. 实用心脑血管病杂志, 2021, 29(5): 59-63.
WU Yunhong, LIN Yaming, FU Weiwei, et al. Influencing factors of diabetes patients complicated with acute ischemic stroke [J]. Practical Journal of Cardiac Cerebral Pneumal and Vascular Disease, 2021, 29(5): 59-63.
- [11] 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)(上) [J]. 中国实用内科杂志, 2021, 41(8): 668-695.
- [12] Nathanson MH, Andrzejowski J, Dinsmore J, et al. Guidelines for safe transfer of the brain-injured patient: trauma and stroke, 2019: guidelines from the Association of Anaesthetists and the Neuro Anaesthesia and Critical Care Society [J]. Anaesthesia, 2020, 75(2): 234-246.
- [13] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers [J]. Mach Learning, 1997, 29: 131-163. doi: 10.1023/A:1007465528199
- [14] Glover F. Artificial intelligence, heuristic frameworks and tabu search [J]. Manage Decis Econ, 1990, 11(5): 365-375.
- [15] 陈凯庭, 安艳荣, 董学勤, 等. 新诊断 2 型糖尿病患者采用短期胰岛素泵强化治疗的效果 [J]. 中国老年学杂志, 2021, 41(5): 945-948.
CHEN Kaiting, AN Yanrong, DONG Xueqin, et al. Effect of short-term insulin pump intensive treatment on newly diagnosed type 2 diabetes mellitus patients [J]. Chinese Journal of Gerontology, 2021, 41(5): 945-948.
- [16] 全会标, 高勇义, 李唐瑛. 老年 2 型糖尿病合并脑卒中危险因素分析 [J]. 中国热带医学, 2012, 12(5): 579-581.
QUAN Huibiao, GAO Yongyi, LI Tangying. Risk factors associated with cerebral stroke in seniole type 2 diabetes militus patients [J]. China Tropical Medicine, 2012, 12(5): 579-581.
- [17] 张竞文, 陈颀, 赵紫楠, 等. 依达拉奉治疗糖尿病合并急性缺血性脑卒中的有效性和安全性系统评价 [J]. 中国药业, 2022, 31(3): 108-114.
ZHANG Jingwen, CHEN Di, ZHAO Zinan, et al. Efficacy and safety of edaravone in the treatment of diabetes mellitus complicated with acute ischemic stroke: a systematic review [J]. China Pharmaceuticals, 2022, 31(3): 108-114.
- [18] 黄永锋, 罗若佳, 范思铭. 胰岛素泵强化甲巯咪唑治疗 2 型糖尿病合并甲状腺功能亢进症的疗效及预后分析 [J]. 吉林医学, 2021, 42(1): 90-92.
HUANG Yongfeng, LUO Ruojia, FAN Siming. Efficacy and prognosis analysis of insulin pump strengthening methimazole in the treatment of type 2 diabetes mellitus complicated with hyperthyroidism [J]. Jilin Medical Journal, 2021, 42(1): 90-92.
- [19] 尤艺, 韩迪. 2 型糖尿病合并非酒精性脂肪肝与缺血性脑血管病的关系分析 [J]. 临床医学进展, 2024, 14(3): 1504-1513.
YOU Yi, HAN Di. Analysis of the relationship between type 2 diabetes mellitus with nonalcoholic fatty liver disease and ischemic cerebrovascular disease [J]. Advances in Clinical Medicine, 2024, 14(3): 1504-1513.
- [20] 宗铮. T2DM 与认知功能的相关危险因素分析 [D]. 大连: 大连医科大学, 2021.
- [21] 周新华, 邱伟, 范改焕. T2DM 患者肾脏病变与血脂代谢的相关性 [J]. 医药论坛杂志, 2020, 41(11): 88-91.
ZHOU Xinhua, QIU Wei, FAN Gaihuan. Correlation between renal lesions and lipid metabolism in patients with T2DM [J]. Journal of Medical Forum, 2020, 41(11): 88-91.
- [22] 庞宗然, 苏晓慧, 刘祖涵, 等. 微循环障碍与糖尿病及其并发症关系 [J]. 时珍国医国药, 2011, 22(4): 988-989.
PANG Zongran, SU Xiaohui, LIU Zuhan, et al. Relationship between microcirculation disturbance and diabetes mellitus and its complication [J]. Lishizhen Medicine and Materia Medica Research, 2011, 22(4): 988-989.