·论著·

分类树模型在缺血性脑卒中危险因素筛选中的应用研究

姚爽 李浩 刘开祥 冷光朋 于健

541001 广西壮族自治区桂林，桂林医学院附属医院内分泌科（姚爽、于健），神经内科（李浩、刘开祥）；541100 广西壮族自治区桂林，桂林医学院第二附属医院神经内科（冷光朋）

通讯作者：于健，Email：duduyu1623@qq.com

DOI：10.3760/cma.j.issn.2095-4352.2018.10.014

【摘要】 目的 应用分类树模型构建缺血性脑卒中（IS）发病风险的预测模型，并评价其应用价值。**方法** 采取整群抽样的方法，选取2017年1月至12月桂林医学院附属医院临床资料完善的858例IS患者（IS组），并选择同期与IS患者性别、年龄相匹配的844例健康体检者作为对照（健康对照组），比较分析两组人群的代谢特征。应用分类树模型构建IS发病风险的预测模型，并采用增益图、索引图、错分概率Risk值和受试者工作特征曲线（ROC）评价该模型的应用价值。**结果** 与健康对照组比较，IS组患者体重指数（BMI）、空腹血糖（FPG）、三酰甘油（TG）、总胆固醇（TC）、低密度脂蛋白胆固醇（LDL-C）水平明显升高〔BMI（kg/m^2）：25.34±3.70比24.24±3.10，FPG（mmol/L）：6.79±2.89比5.73±1.17，TG（mmol/L）：1.62±1.06比1.44±1.06，TC（mmol/L）：4.70±2.73比4.35±0.79，LDL-C（mmol/L）：3.18±0.94比2.73±0.73，均$P<0.01$〕，高密度脂蛋白胆固醇（HDL-C）水平明显降低（mmol/L：1.12±0.33比1.35±0.36，$P<0.01$），高血压及有吸烟史、饮酒史的比例明显升高（69.0%比41.9%，23.1%比16.8%，19.2%比13.4%，均$P<0.01$）。对各因素赋值〔IS：无为0，有为1；BMI：＜24.0 kg/m^2为0，≥24.0 kg/m^2为1；FPG：＜7.0 mmol/L为0，≥7.0 mmol/L为1；TG：＜2.26 mmol/L为0，≥2.26 mmol/L为1；TC：＜6.22 mmol/L为0，≥6.22 mmol/L为1；LDL-C：＜4.14 mmol/L为0，≥4.14 mmol/L为1；HDL-C：＜1.04 mmol/L为0，≥1.04 mmol/L为1；高血压：无为0，有为1；吸烟史：无为0，有为1；饮酒史：无为0，有为1〕，建立分类树模型分析IS的危险因素，得出分类树模型共包括4层（第一层为高血压，第二层为FPG和HDL-C，第三层为HDL-C和FPG，第四层为LDL-C和吸烟史）、17个结点，最终筛选出5个解释变量，即高血压、FPG、HDL-C、LDL-C、吸烟史。树的第一层显示，高血压人群IS的发生概率（62.6%）明显高于无高血压人群（35.2%）。树的第二层显示，在有高血压人群中，HDL-C≥1.04 mmol/L者的IS发生概率（53.6%）低于HDL-C＜1.04 mmol/L者（78.5%）；在无高血压人群中，FPG≥7.0 mmol/L者的IS发生概率（71.1%）明显高于FPG＜7.0 mmol/L者（28.3%）。树的第三层显示，在无高血压、FPG＜7.0 mmol/L的人群中，HDL-C≥1.04 mmol/L者的IS发生概率（21.8%）低于HDL-C＜1.04 mmol/L者（48.7%）；在有高血压、HDL-C≥1.04 mmol/L的人群中，FPG≥7.0 mmol/L者的IS发生概率（78.6%）明显高于FPG＜7.0 mmol/L者（46.7%）。树的第四层显示，在无高血压、FPG＜7.0 mmol/L和HDL-C≥1.04 mmol/L的人群中，LDL-C≥4.14 mmol/L者的IS发生概率（53.8%）高于LDL-C＜4.14 mmol/L者（19.0%）；在无高血压、FPG＜7.0 mmol/L和HDL-C＜1.04 mmol/L的人群中，吸烟者的IS发生概率（76.9%）高于非吸烟者（39.1%）；在有高血压、HDL-C≥1.04 mmol/L和FPG＜7.0 mmol/L的人群中，LDL-C≥4.14 mmol/L者的IS发生概率（72.5%）高于LDL-C＜4.14 mmol/L者（44.4%）。IS分类树模型增益图显示增益值从0%开始先向100%快速增长后趋于平稳，索引图显示索引值从100%以上开始沿移动方向保持平稳状态，然后快速向100%下降，说明该模型良好。分类树模型错分概率Risk值为0.291，表明该模型对IS患者危险因素预测正确率是70.9%。分类树模型预测IS危险因素的ROC曲线下面积（AUC）为78.0%〔95%可信区间（95%*CI*）=75.9%～79.9%，$P<0.001$〕，敏感度为62.5%（95%*CI*=59.1%～65.7%），特异度为79.4%（95%*CI*=76.5%～82.1%）。**结论** 分类树模型能有效拟合IS患者危险因素的预测，其中高血压、高血糖、高LDL-C、吸烟史是IS的主要危险因素。

【关键词】 缺血性脑卒中；分类树；危险因素；预测模型

基金项目：广西自然科学基金（2016GXNSFAA380310）；广西科技计划项目（桂科AD16450020）；广西壮族自治区临床重点专科建设项目（2018-6）；广西桂林市科技攻关计划项目（20170109-11）

Study on the application of classification tree model in screening the risk factors of ischemic stroke

Yao Shuang, Li Hao, Liu Kaixiang, Leng Guangpeng, Yu Jian

Department of Endocrinology, the Affiliated Hospital of Guilin Medical College, Guilin 541001, Guangxi Zhuang Antonomous Region, China (Yao S, Yu J); Department of Neurology, the Affiliated Hospital of Guilin Medical College, Guilin 541001, Guangxi Zhuang Antonomous Region, China (Li H, Liu KX); Department of Cardiovascular Medicine, the Second Affiliated Hospital of Medical College, Guilin Medical University, Guilin 541100, Guangxi Zhuang Antonomous Region, China (Leng GP)

Corresponding author: Yu Jian, Email: duduyu1623@qq.com

【Abstract】 **Objective** To construct a prediction model for the risk of ischemic stroke (IS) by classification tree model, and evaluate its application value. **Methods** By cluster sampling, 858 IS patients with perfect clinical data from January to December 2017 in the Affiliated Hospital of Guilin Medical College (IS group) were enrolled, and 844 health checkups matched with the gender and age of IS patients in the same period were enrolled as controls (healthy control group). The metabolic characteristics of the two groups were compared and analyzed. The classification tree model was used to construct the prediction model of the risk of IS, and the gain diagram, index chart, risk value of misclassification probability and receiver operating characteristic curve (ROC) were used to evaluate the application value of the model. **Results** Compared with the healthy control group, body mass index (BMI), fasting blood glucose (FPG), triglyceride (TG), total cholesterol (TC), low density lipoprotein cholesterol (LDL-C) in IS group were significantly increased [BMI (kg/m^2): 25.34±3.70 vs. 24.24±3.10, FPG (mmol/L): 6.79±2.89 vs. 5.73±1.17, TG (mmol/L): 1.62±1.06 vs. 1.44±1.06, TC (mmol/L): 4.70±2.73 vs. 4.35±0.79, LDL-C (mmol/L): 3.18±0.94 vs. 2.73±0.73, all $P < 0.01$], high density lipoprotein cholesterol (HDL-C) was significantly decreased (mmol/L: 1.12±0.33 vs. 1.35±0.36, $P < 0.01$), and the proportion of hypertension, smoking and drinking were significantly increased (69.0% vs. 41.9%, 23.1% vs. 16.8%, 19.2% vs. 13.4%, all $P < 0.01$). By assigning values to each factor [IS: No = 0, Yes = 1; BMI: < 24.0 kg/m^2 = 0, ≥ 24.0 kg/m^2 = 1; FPG : < 7.0 mmol/L = 0, ≥ 7.0 mmol/L = 1; TG: < 2.26 mmol/L = 0, ≥ 2.26 mmol/L = 1; TC: < 6.22 mmol/L = 0, ≥ 6.22 mmol/L = 1; LDL-C: < 4.14 mmol/L = 0, ≥ 4.14 mmol/L = 1; HDL-C: < 1.04 mmol/L = 0, ≥ 1.04 mmol/L = 1; hypertension: No = 0, Yes = 1; smoking: No = 0, Yes = 1; drinking: No = 0, Yes = 1], a classification tree model was established to analyze the risk factors of IS. The classification tree model consisted of 4 layers and 17 nodes: the first layer was hypertension, the second layer was FPG and HDL-C, the third layer was HDL-C and FPG, and the fourth layer was LDL-C and smoking. There were five explanatory variables screened out in the model, including hypertension, FPG, HDL-C, LDL-C and smoking. The first layer of the tree showed that the incidence of IS in hypertensive population (62.6%) was significantly higher than that in non-hypertensive population (35.2%). The second layer of the tree showed that the incidence of IS in people with hypertension with HDL-C ≥ 1.04 mmol/L (53.6%) was lower than that in people with HDL-C < 1.04 mmol/L (78.5%). However, in the population without hypertension, the probability of IS occurrence in the population with FPG ≥ 7.0 mmol/L (71.1%) was significantly higher than that in the population with FPG < 7.0 mmol/L (28.3%). The third layer of the tree showed that the IS incidence of HDL-C ≥ 1.04 mmol/L (21.8%) was lower than that of HDL-C < 1.04 mmol/L (48.7%) in the population without hypertension and FPG < 7.0 mmol/L. However, in the population with hypertension and HDL-C ≥ 1.04 mmol/L, the probability of IS occurrence in the population with FPG ≥ 7.0 mmol/L (78.6%) was significantly higher than that in the population with FPG < 7.0 mmol/L (46.7%). The fourth layer of the tree showed that the IS incidence of people with LDL-C ≥ 4.14 mmol/L (53.8%) was higher than that of people with LDL-C < 4.14 mmol/L (19.0%) in the population without hypertension, FPG < 7.0 mmol/L and HDL-C ≥ 1.04 mmol/L. In the population without hypertension, the incidence of IS in smokers (76.9%) was higher than that in non-smokers (39.1%) of people with FPG < 7.0 mmol/L and HDL-C < 1.04 mmol/L. In the population with hypertension, the probability of IS occurrence in the population with LDL-C ≥ 4.14 mmol/L (72.5%) was higher than that in the population with LDL-C < 4.14 mmol/L (44.4 %) of people with HDL-C ≥ 1.04 mmol/L and FPG < 7.0 mmol/L. The gain diagram of IS classification tree model shown that the gain value increased rapidly from 0% to 100% and then tended to be stable. The index chart shown that the index value kept stable in the moving direction from above 100% and then dropped rapidly to 100%, indicating the model was very well. The risk value of misclassification probability of the classification tree model was 0.291, and the correct rate of risk factor for IS patients was 70.90%. The area under ROC curve (AUC) was 78.0% [95% confidence interval (95%*CI*) = 75.9%-79.9%, $P < 0.001$], the sensitivity was 62.5% (95%*CI* = 59.1%-65.7%) and the specificity was 79.4% (95%*CI* = 76.5%-82.1%). **Conclusion** Classification tree model can properly predict the risk factor of IS, and the most important risk factors are hypertension, hyperglycemia, high LDL-C and smoking.

【Key words】 Ischemic stroke; Classification tree; Risk factor; Prediction model

Fund program: Guangxi Natural Science Foundation of China (2016GXNSFAA380310); Guangxi Science and Technology Project of China (AD16450020); Guangxi Clinical Key Specialty Construction Project (2018-6); Guilin Technologies R & D Program Plan Project of Guangxi of China (20170109-11)

缺血性脑卒中(IS)是临床比较常见、多发的神经系统疾病,可由多种因素引起局部血液循环系统出现障碍、缺血及缺氧等,导致局限性脑组织出现缺血性软化或坏死[1]。IS 是脑卒中最常见的类型,在脑卒中患者中约占 85%[2],而且仍以每年 8.7% 的速率增长。IS 的发病率和病死率比心血管疾病高 3 倍以上,是复杂多因素共同作用的结果,既往研究通常采用 Logistic 回归、Cox 回归等传统方法对该疾病进行分析[3],但各种分析方法均存在自身缺陷,即无法处理共性问题,而且容易忽视各因素之间的层次关系。一种新的分析方法——分类树模型,能有效弥补传统分析方法的不足,可以有效、直观、有层次地显示疾病的危险因素以及各因素之间的交互作用[4],但该模型甚少应用于 IS 危险因素的预测。本研究旨在对 IS 进行分类树模型构建,并对该模型进行评测,结果报告如下。

1　资料与方法

1.1　研究对象:采取整群抽样的方法,选择 2017 年

1 月至 12 月桂林医学院附属医院神经内科住院部收治的 IS 患者（IS 组）；并选择同期与 IS 患者性别、年龄相匹配的健康体检者作为对照（健康对照组）。

1.1.1 纳入标准：① 符合全国第四届脑血管病会议制定的 IS 诊断标准（发病时间超过 4.5 h，不在溶栓治疗范围内）；② 头颅 CT 或磁共振成像（MRI）显示主病灶存在。

1.1.2 排除标准：① 有严重肝肾功能不全或心力衰竭等症状；② 严重的全身感染；③ 近期经历过外科手术；④ 有恶性肿瘤；⑤ 既往有外伤史；⑥ 临床资料不完整。

1.2 伦理学：本研究符合医院伦理学标准，并经医院伦理委员会审批通过（审批号：GLMu1A2017019）。

1.3 病例收集指标：记录患者的性别、年龄、吸烟史、饮酒史、身高、体重、体重指数（BMI）、血压、三酰甘油（TG）、总胆固醇（TC）、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（FPG）等。

1.4 定义：① 高血压[5]：3 次测量收缩压（SBP）平均值≥140 mmHg（1 mmHg =0.133 kPa）和（或）舒张压（DBP）≥90 mmHg；或者近 2 周内服用降压药血压正常者，排除继发性高血压者。② 血脂异常[6]：TC≥6.22 mmol/L，TG≥2.26 mmol/L，LDL-C≥4.14 mmol/L，HDL-C＜1.04 mmol/L，出现以上任何一项即可诊断。③ 成人 BMI[7]：24 kg/m^2≤BMI＜28 kg/m^2 为超重，BMI≥28 kg/m^2 为肥胖。④ 吸烟史：连续吸烟≥6 个月，且在调查日前 30 d 内吸过烟。⑤ 饮酒史：平均每周饮酒≥1 次，包括白酒、啤酒等。

1.5 统计学分析：使用 SPSS 18.0 软件和 Medcalc 软件分析数据。正态分布的计量资料以均数 ± 标准差（$\bar{x}\pm s$）表示，两组间比较采用 t 检验；计数资料用率表示，两组间比较采用 χ^2 检验；P＜0.05 为差异有统计学意义。建立分类树模型，并应用增益图、索引图和错分概率对模型结果进行评价，用受试者工作特征曲线（ROC）对模型准确性进行估算，树深度设定为 5，为充分挖掘影响 IS 患者的危险因素及其相互之间的关系，父结点中的最小样本数设定为 50，子结点设定为 20。

2 结 果

2.1 患者的基本资料和代谢特征（表 1）：共入选 858 例 IS 患者，844 例健康体检者。两组患者性别、年龄差异无统计学意义。与健康对照组比较，IS 组患者 BMI、FPG、TG、TC、LDL-C 水平明显升高，HDL-C 水平明显降低，高血压及有吸烟史、饮酒史的患者比例明显升高（均 P＜0.01）。

表 1 IS 患者与健康体检者基本资料和代谢特征比较

变量	IS 组（n=858）	健康对照组（n=844）	χ^2/t 值	P 值
性别（例，男 / 女）	515/343	538/306	2.497	0.122
年龄（岁，$\bar{x}\pm s$）	64.41±11.36	63.52±9.87	1.718	0.086
BMI（kg/m^2，$\bar{x}\pm s$）	25.34±3.70	24.24±3.10	6.653	0.000
吸烟史〔例（%）〕	198（23.1）	142（16.8）	10.404	0.001
饮酒史〔例（%）〕	165（19.2）	113（13.4）	10.626	0.001
高血压〔例（%）〕	592（69.0）	354（41.9）	126.141	0.000
FPG（mmol/L，$\bar{x}\pm s$）	6.79±2.89	5.73±1.17	9.942	0.000
TG（mmol/L，$\bar{x}\pm s$）	1.62±1.06	1.44±1.06	3.467	0.001
TC（mmol/L，$\bar{x}\pm s$）	4.70±2.73	4.35±0.79	3.543	0.000
LDL-C（mmol/L，$\bar{x}\pm s$）	3.18±0.94	2.73±0.73	10.989	0.000
HDL-C（mmol/L，$\bar{x}\pm s$）	1.12±0.33	1.35±0.36	−13.885	0.000

注：IS 为缺血性脑卒中，BMI 为体重指数，FPG 为空腹血糖，TG 为三酰甘油，TC 为总胆固醇，LDL-C 为低密度脂蛋白胆固醇，HDL-C 为高密度脂蛋白胆固醇

2.2 IS 分类树模型的建立：建立分类树模型分析 IS 的危险因素，将 IS 作为因变量，其余各因素结合专业知识进行赋值，见表 2。为了防止分类树模型过度拟合，分类树模型的修剪和生长根据上述规则进行约束后，得出模型共 4 层、17 个结点、9 个终末结点，共 5 个解释变量，即高血压、FPG、HDL-C、LDL-C、吸烟史，见图 1。

表 2 IS 患者危险因素主要变量及赋值

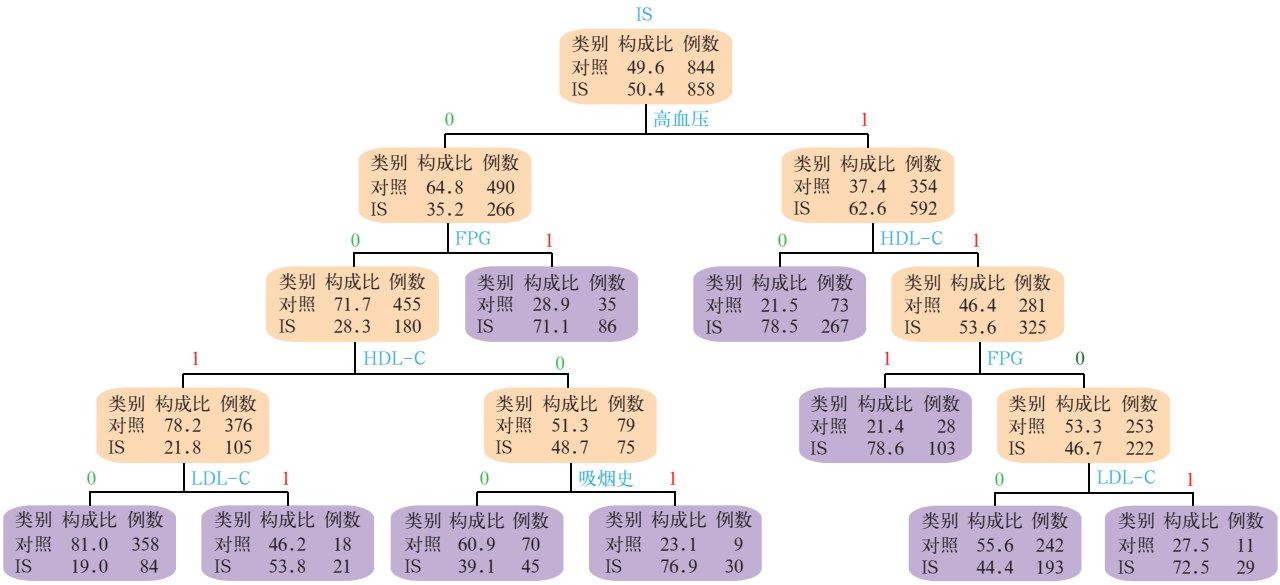
变量	赋值
IS	无＝0；有＝1
BMI	＜24.0 kg/m^2＝0；≥24.0 kg/m^2＝1
FPG	＜7.0 mmol/L＝0；≥7.0 mmol/L＝1
TG	＜2.26 mmol/L＝0；≥2.26 mmol/L＝1
TC	＜6.22 mmol/L＝0；≥6.22 mmol/L＝1
LDL-C	＜4.14 mmol/L＝0；≥4.14 mmol/L＝1
HDL-C	＜1.04 mmol/L＝0；≥1.04 mmol/L＝1
高血压	无＝0；有＝1
吸烟史	无＝0；有＝1
饮酒史	无＝0；有＝1

注：IS 为缺血性脑卒中，BMI 为体重指数，FPG 为空腹血糖，TG 为三酯甘油，TC 为总胆固醇，LDL-C 为低密度脂蛋白胆固醇，HDL-C 为高密度脂蛋白胆固醇

2.3 IS 分类树模型的评测

2.3.1 增益图（图 2）：累计增益图从 0% 开始且以 100% 结束。IS 分类树模型符合良好模型的标准，即增益值先向 100% 快速增长，最终趋于平稳，提示该模型良好。

2.3.2 索引图（图 3）：索引图用来评价该模型的优劣。IS 分类树模型的索引值从 100% 以上开始，沿移动方向保持平稳状态，然后快速地向 100% 下降，说明此次建立的模型良好。



注：IS 为缺血性脑卒中，FPG 为空腹血糖，HDL-C 为高密度脂蛋白胆固醇，LDL-C 为低密度脂蛋白胆固醇

图 1　筛选 IS 患者危险因素的分类树模型

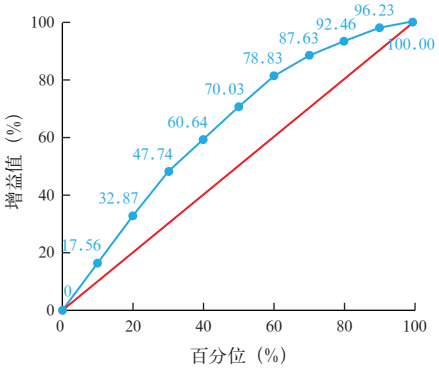


图 2　缺血性脑卒中（IS）发病风险分类树模型增益图

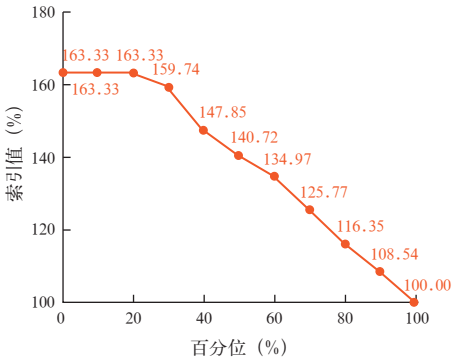


图 3　缺血性脑卒中（IS）发病风险分类树模型索引图

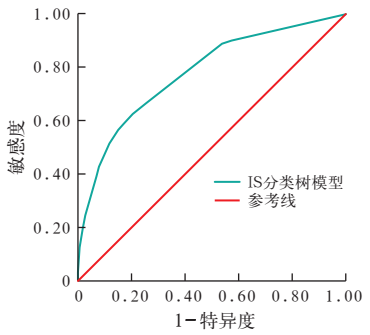


图 4　缺血性脑卒中（IS）分类树模型预测 IS 患者危险因素的受试者工作特征曲线（ROC）

2.3.3　IS 分类树模型的错分矩阵和 Risk 统计量（表 3）：分类树模型的 Risk 值为 0.291，即错估概率为 29.1%，说明该模型对 IS 患者危险因素预测正确率为 70.9%；错分矩阵分类表的预测结果与风险表一致，为 70.9%，正确率较高，说明该模型预测 IS 患者危险因素的效果良好。

表 3　IS 分类树模型错分矩阵和 Risk 统计量

预测分类	错分矩阵（真实分类）		Risk 统计量	
	IS 组（例）	对照组（例）	估计值	$s_{\bar{x}}$
IS 组（例）	536	322	0.291	0.011
对照组（例）	174	670		

注：IS 为缺血性脑卒中

2.4　ROC 曲线的绘制（图 4）：根据 IS 分类树模型得到预测变量，并将这个变量作为测试变量，IS 有无分组作为状态变量。结果显示，ROC 曲线远离参考线，说明 IS 分类树模型可信。另外，分类树模型预测 IS 危险因素的 ROC 曲线下面积（AUC）为 78.0%，95% 可信区间（95%*CI*）为 75.9%～79.9%，标准误为 0.011，差异有统计学意义（$P<0.001$），提示 IS 分类树模型预测准确性较高，其敏感度为 62.5%（95%*CI*=59.1%～65.7%），特异度为 79.4%（95%*CI*=76.5%～82.1%）。

3　讨　论

IS 的发生发展受遗传因素、环境因素和生活习惯等复杂因素以及相互作用的影响。既往研究用传统方法对这些多因子复杂疾病进行分析，但无法有效处理共线性问题；而分类树以树形图的方式可直观展示出多因素之间的交互作用，具有独特优势。

分类回归树（CART）模型是一种对进入模型的变量无特殊要求、连续变量也可作为自变量进入模型的非参数回归模型。树模型是根据目标变量自我分层的树状结构，每一种可能的组合在建立过程中都被考查过，因此，分类树模型是不断优化后的模型，呈现出的树形图易于理解，清晰明了，

最终结果也可以展现出分析过程以及各因素之间的相互作用，这一点相对于传统的统计分析方法具有明显优势[8]。本研究中 IS 分类树模型结果显示：高血压、高血糖、高 LDL-C、吸烟史是 IS 发病的危险因素，高 HDL-C 是保护因素，与以往分析结果基本一致[9-11]。分类树模型不仅能有效地对 IS 发病风险因素进行预测，也可有效筛选各因素之间的交互作用。本研究中分类树模型的第一层是高血压，所以 IS 发病风险最重要的影响因素是高血压，结果显示：高血压人群 IS 的发生概率为 62.6%，明显高于无高血压人群（35.2%），提示高血压患者为 IS 发病的高危人群。树的第二层，在有高血压人群中，HDL-C≥1.04 mmol/L 者的 IS 发生概率（53.6%）低于 HDL-C＜1.04 mmol/L 者（78.5%），提示 HDL-C 是 IS 患者发病的保护因素；而在无高血压人群中，FPG≥7.0 mmol/L 者的 IS 发生概率（71.1%）明显高于 FPG＜7.0 mmol/L 者（28.3%）。树的第三层，在无高血压、FPG＜7.0 mmol/L 的人群中，HDL-C≥1.04 mmol/L 者的 IS 发生概率（21.8%）低于 HDL-C＜1.04 mmol/L 者（48.7%）；而在有高血压、HDL-C≥1.04 mmol/L 的人群中，FPG≥7.0 mmol/L 者的 IS 发生概率（78.6%）明显高于 FPG＜7.0 mmol/L 者（46.7%）。树的第四层是吸烟和 LDL-C，在无高血压、FPG＜7.0 mmol/L 和 HDL-C≥1.04 mmol/L 的人群中，LDL-C≥4.14 mmol/L 者的 IS 发生概率（53.8%）高于 LDL-C＜4.14 mmol/L 者（19.0%）；在无高血压、FPG＜7.0 mmol/L 和 HDL-C＜1.04 mmol/L 人群中，吸烟者的 IS 发生概率（76.9%）高于非吸烟者（39.1%）；在有高血压、HDL-C≥1.04 mmol/L 和 FPG＜7.0 mmol/L 人群中，LDL-C≥4.14 mmol/L 者的 IS 发生概率（72.5%）高于 LDL-C＜4.14 mmol/L 者（44.4%）。因此，要加强对血压、血糖和血脂水平的检测，并积极进行治疗。

本研究用 Risk 值对分类树模型的预测结果进行评估，得出模型对 IS 危险因素预测的正确率为 70.9%；错分矩阵分类表的预测结果也为 70.9%。表明用树模型依靠风险因素可对 70.9% 的 IS 进行正确分类，模型预测结果良好。ROC 曲线分析显示，AUC 为 78.0%，标准误为 0.011，也说明分类树模型预测 IS 危险因素的准确性较高。以上结果表明本研究建立的 IS 风险因素预测模型具有较高的可信度。

分类树模型也有其局限性，例如在分析小样本量数据时稳定性较差，更适合分析样本量大的数据。

综上，分类树模型是分析样本数量大、多因子复杂疾病的有力工具，本研究将其应用于 IS 发病风险因素的分析，效果较好，可作为提供疾病预防以及住院方案制定的依据。

参考文献

[1] 中国中西医结合学会急救医学专业委员会. 中国急性缺血性脑卒中中西医急诊诊治专家共识 [J]. 中华危重病急救医学, 2018, 30 (3): 193-197. DOI: 10.3760/cma.j.issn.2095-4352.2018.03.001.
Professional Committee of First Aid Medicine of Chinese Society of Integrated Traditional and Western Medicine. Expert consensus of Chinese and Western medicine emergency treatment for acute ischemic stroke in China [J]. Chin Crit Care Med, 2018, 30 (3): 193-197. DOI: 10.3760/cma.j.issn.2095-4352.2018.03.001.

[2] Lloyd-Jones D, Adams RJ, Brown TM, et al. Heart disease and stroke statistics: 2010 update: a report from the American Heart Association [J]. Circulation, 2010, 121 (7): e46-215. DOI: 10.1161/CIRCULATIONAHA.109.192667.

[3] 周汝娟, 何龙锦. 江苏泰兴地区人群缺血性脑卒中的危险因素 [J]. 临床神经病学杂志, 2015, 28 (5): 385-387.
Zhou RJ, He LJ. Risk factors of ischemic stroke in Jiangsu Taixing population [J]. J Clin Neurol, 2015, 28 (5): 385-387.

[4] 蔡晓楠, 张丹丹, 严亚琼, 等. 应用分类树模型构建耐多药结核病发病风险模型 [J]. 中华疾病控制杂志, 2016, 20 (1): 91-95. DOI: 10.16462/j.cnki.zhjbkz.2016.01.023.
Cai XN, Zhang DD, Yan YQ, et al. A study on the application of classification tree model in building the risk model for multidrug resistant tuberculosis [J]. Chin J Dis Control Prevent, 2016, 20 (1): 91-95. DOI: 10.16462/j.cnki.zhjbkz.2016.01.023.

[5] 中国高血压防治指南修订委员会. 中国高血压防治指南 2010 [J]. 中华心血管病杂志, 2011, 39 (7): 579-616. DOI: 10.3760/cma.j.issn.0253-3758.2011.07.002.
Writing Group of 2010 Chinese Guidelines for the Management of Hypertension. 2010 Chinese guidelines for the management of hypertension [J]. Chin J Cardiol, 2011, 39 (7): 579-616. DOI: 10.3760/cma.j.issn.0253-3758.2011.07.002.

[6] 中国成人血脂异常防治指南制订联合委员会. 中国成人血脂异常防治指南 [J]. 中华心血管病杂志, 2007, 35 (5): 390-419. DOI: 10.3760/j.issn.0253-3758.2007.05.003.
Writing Group of Chinese Guidelines on Prevention and Treatment of Dyslipidemia in Adults. Chinese guidelines on prevention and treatment of dyslipidemia in adults [J]. Chin J Cardiol, 2007, 35 (5): 390-419. DOI: 10.3760/j.issn.0253-3758.2007.05.003.

[7] 中华人民共和国卫生部疾病控制司. 中国成人超重和肥胖症预防控制指南 [M]. 北京：人民卫生出版社, 2006: 2-4.
Department of Disease Control, Ministry of Health, China. Guidelines for prevention and control of overweight and obesity in Chinese adults [M]. Beijing: People's Medical Publishing House, 2006: 2-4.

[8] 傅传喜, 马文军, 梁建华, 等. 高血压危险因素 logistic 回归与分类树分析 [J]. 中华疾病控制杂志, 2006, 10 (3): 256-259. DOI: 10.3969/j.issn.1674-3679.2006.03.013.
Fu CX, Ma WJ, Liang JH, et al. Logistic regression and classification tree analysis on risk factors of hypertension [J]. Chin J Dis Control Prevent, 2006, 10 (3): 256-259. DOI: 10.3969/j.issn.1674-3679.2006.03.013.

[9] 朱明慧, 王辉, 徐艳, 等. 脑出血、脑梗死急性发作期患者血压、血脂、凝血纤溶功能差异及意义 [J]. 山东医药, 2017, 57 (4): 66-68. DOI: 10.3969/j.issn.1002-266X.2017.04.022.
Zhu MH, Wang H, Xu Y, et al. Difference and significance of blood pressure, blood lipid, coagulation and fibrinolysis in patients with cerebral hemorrhage and acute cerebral infarction [J]. Shandong Med J, 2017, 57 (4): 66-68. DOI: 10.3969/j.issn.1002-266X.2017.04.022.

[10] 孔亚婷, 朱正庭. 年龄和高血压及糖尿病对短暂性脑缺血发作患者发展为急性脑梗死的影响 [J]. 中国全科医学, 2017, 20 (A01): 15-17. DOI: 10.3969/j.issn.1007-9572.2017.z1.006.
Kong YT, Zhu ZT. Influence of age, hypertension and diabetes on transient ischemic attack developing into acute cerebral infarction [J]. Chin Gen Pract, 2017, 20 (A01): 15-17. DOI: 10.3969/j.issn.1007-9572.2017.z1.006.

[11] 宋晴, 高彩霞, 周齐, 等. 缺血性脑卒中危险因素的多元回归分析 [J]. 脑与神经疾病杂志, 2014, 22 (5): 350-353.
Song Q, Gao CX, Zhou Q, et al. Logistic regression analysis of risk factors of ischemic stroke [J]. J Brain Nerv Dis, 2014, 22 (5): 350-353.

（收稿日期：2018-04-16）