

基于数据挖掘技术的肺癌危险度预测模型的构建

黄普超¹, 原慧洁², 张桂芳¹

1. 新乡市中心医院 新乡医学院第四临床学院, 河南 新乡 453000; 2. 郑州大学公共卫生学院, 河南 郑州 450001

摘要: **目的** 借助数据挖掘技术, 联合流行病学特征和临床症状资料构建肺癌危险度预测模型, 评价各模型用于肺癌危险度预测的性能, 并筛选出最优模型。 **方法** 选取 460 例肺癌患者和 560 例肺良性疾病患者为研究对象, 收集其流行病学特征和临床症状共 16 个自变量。将研究对象按照 3 : 1 的比例随机分为训练集与测试集, 应用支持向量机 (support vector machine, SVM)、决策树 C5.0 和人工神经网络 (artificial neural network, ANN) 分别建立肺癌危险度预测模型, 并比较不同模型的预测性能。 **结果** 经特征提取, 痰中带血、发热出汗和吸烟史等 9 个变量被筛选为有效变量, 用来构建肺癌危险度预测模型。测试集中 SVM、决策树 C5.0 和 ANN 模型的灵敏度分别为 74.1%、62.5% 和 92.9%; 特异度分别为 76.2%、80.4% 和 64.3%; 阳性预测值分别为 70.9%、71.4% 和 67.1%; 阴性预测值分别为 79.0%、73.2% 和 92.0%; 准确度分别为 75.3%、72.5% 和 76.9%; 曲线下面积分别为 0.752 (95% CI: 0.694~0.803)、0.715 (95% CI: 0.655~0.769) 和 0.786 (95% CI: 0.730~0.835)。 **结论** ANN 预测模型的整体性能优于 SVM 模型和决策树 C5.0 模型, 在肺癌高危人群的筛查中具有潜在的应用价值。

关键词: 肺癌; 支持向量机; 决策树; 人工神经网络; 危险度预测

中图分类号: R734.2 **文献标识码:** A **文章编号:** 1006-3110(2022)11-1390-05 DOI: 10.3969/j.issn.1006-3110.2022.11.028

Construction of risk prediction model for lung cancer based on data mining technology

HUANG Pu-chao¹, YUAN Hui-jie², ZHANG Gui-fang¹

1. Xinxiang Central Hospital, the Fourth Clinical College of Xinxiang Medical University, Xinxiang, Henan 453000, China;

2. School of Public Health, Zhengzhou University, Zhengzhou, Henan 450001, China

Corresponding author: ZHANG Gui-fang, E-mail: xxchzhangguifang@126.com

Abstract: **Objective** To establish risk prediction models for lung cancer based on the data about epidemiological characteristics and clinical symptoms by data mining technology, and to evaluate the performance of each model so as to screen out the optimal predictive model. **Methods** Four hundred and sixty patients with lung cancer and 560 patients with benign lung disease were

作者简介: 黄普超 (1988-), 男, 河南商丘人, 硕士, 主治医师, 主要从事肿瘤诊治工作。

通信作者: 张桂芳, E-mail: xxchzhangguifang@126.com。

- [11] Boumba VA, Mitselou A, Vougiouklakis T. Fatal poisoning from ingestion of datura stramonium seeds[J]. Vet Hum Toxicol, 2004, 46(2):81-82.
- [12] 张金杰, 齐术凤, 郑连勇. 曼陀罗花中毒致死 1 例[J]. 中国实用乡村医生杂志, 2004, 11(2):45.
- [13] 王喜胜, 牟玲. 洋金花中毒致死 1 例[J]. 锦州医学院学报, 1995, 25(1):76.
- [14] 姚松杰, 刘洁泉, 施震霄. 急性曼陀罗药酒群体中毒 4 例救治体会[J]. 中国乡村医药, 2020, 27(19):31-32.
- [15] 罗容, 肖佩荣. 以神经精神症状首发的曼陀罗中毒 24 例[J]. 中国神经精神疾病杂志, 2020, 46(2):95-97.
- [16] 苏玮玮, 李娟娟, 余思洋, 等. 云南省 2012—2019 年乌头类植物中毒流行特征[J]. 中国热带医学, 2020, 20(7):666-669.
- [17] 彭浩, 唐江, 杨汉策, 等. 一起居民聚餐饮用自制药酒引起中毒事件的调查分析[J]. 实用预防医学, 2020, 27(3):374-375.
- [18] Soni P, Siddiqui AA, Dwivedi J. Pharmacological properties of *Datura stramonium* L. as a potential medicinal tree: an overview[J]. Asian Pac J Trop Bio, 2012, 2(12):1002-1008.
- [19] 张虎. 急性曼陀罗中毒患者的临床诊治 (32 例)[J]. 医疗装备, 2017, 30(5):106-107.
- [20] 周权, 林馨, 刘建, 等. 一起自制曼陀罗药酒引起的食物中毒[J]. 海峡预防医学杂志, 2011, 17(6):32.
- [21] 陈加飞, 杨华芳, 熊拥军, 等. 2 例自制药酒中毒患者救治经验总结[J]. 重庆医学, 2017, 46(14):2013-2015.
- [22] 马金龙, 张峰, 刘萍. 一起由误食曼陀罗引起的食物中毒的调查分析[J]. 现代预防医学, 2015, 42(3):429-430.
- [23] 杨小华, 杨莹. 一起误食曼陀罗引起食物中毒的调查分析[J]. 职业卫生与病伤, 2020, 35(3):187-191.
- [24] 陈莉莉, 陈江, 廖宁波, 等. 2019 年浙江省食源性疾病暴发事件监测数据分析[J]. 中国卫生检验杂志, 2021, 31(20):2554-2558, 2560.

selected as the subjects, and 16 independent variables comprising of epidemiological characteristics and clinical symptoms were collected. All the subjects were randomly divided into the training set and the test set in a ratio of 3:1. Based on the variables and by the use of support vector machine (SVM), C5.0 decision tree and artificial neural network (ANN), 3 risk prediction models for lung cancer were established respectively, and the predictive performances of these models were compared. **Results** After feature extraction, 9 variables including blood in phlegm, fever and sweating and smoking history were selected as the effective variables in the establishment of risk prediction models for lung cancer. In the test set, the sensitivities of SVM, C5.0 decision tree and ANN models were 74.1%, 62.5% and 92.9%, respectively. The specificities were 76.2%, 80.4% and 64.3%, respectively. The positive predictive values were 70.9%, 71.4% and 67.1%, respectively. The negative predictive values were 79.0%, 73.2% and 92.0%, respectively. The accuracies were 75.3%, 72.5% and 76.9%, respectively. The areas under the curves were 0.752 (95% CI:0.694–0.803), 0.715 (95% CI:0.655–0.769) and 0.786 (95% CI:0.730–0.835), respectively. **Conclusion** The ANN prediction model has a better overall performance than SVM and C5.0 decision tree models, and it has potential application value in the screening of high-risk population for lung cancer.

Key words: lung cancer; support vector machine; decision tree; artificial neural network; risk prediction

当前,全球肺癌的发病人数和病死人数持续增加,严重威胁着人类的健康^[1]。据报道,肺癌 I 期患者的 5 年生存率高达 90%,而发展到 IV 期肺癌患者的 5 年生存率不到 5%^[2]。因此,采用肺癌危险因素建立肺癌危险度预测模型对肺癌进行早期诊断与规范化治疗,在提高肺癌患者生存率上具有重要价值。数据挖掘技术在数据处理方面具有独特的优势,已在医学模式预测与诊断中得到广泛应用^[3]。常见的数据挖掘技术主要有支持向量机(support vector machine,SVM)^[4]、决策树^[5]与人工神经网络(artificial neural network,ANN)^[6]等。因此,该研究基于流行病学特征与临床症状资料,采用 SVM、决策树 C5.0 与 ANN 构建肺癌危险度预测模型,并比较不同模型的性能,筛选出优势模型,为肺癌早期筛查与临床辅助诊断提供依据。

1 资料与方法

1.1 研究对象 选择 2016 年 9 月—2021 年 9 月在新乡市中心医院就诊的肺良性疾病患者和原发性肺癌患者作为研究对象。所有被调查者均知情同意并自愿参加。纳入标准:①肺良性疾病组:由新乡市中心医院诊断为肺良性疾病患者;②肺癌组:由病理学或细胞学诊断为肺癌;入组前未接受免疫治疗、放化疗、分子靶向治疗与手术治疗;具有完善的临床资料。排除标准:①合并其他器官的肿瘤;②存在器官功能衰竭的症状。

1.2 观察指标 经过培训并考核合格的调查员,采用问卷访谈的形式收集研究对象的年龄、肺癌家族史和咯血等流行病学和临床症状资料。根据《中华医学会肺癌临床诊疗指南(2019 版)》^[7]将年龄以 45 岁为界限进行分组。预测变量包括 6 个流行病学变量和 10 个临床症状变量,各变量赋值见表 1。

表 1 肺癌危险度研究的变量赋值

变量名称	赋值说明
诊断	肺良性疾病=0;肺癌=1
年龄(岁)	0~44=0;45 及以上=1
吸烟史	否=0;是=1
饮酒史	否=0;是=1
肺基础性疾病	否=0;是=1
粉尘接触史	否=0;是=1
肺癌家族史	否=0;是=1
心慌	否=0;是=1
乏力	否=0;是=1
畏寒	否=0;是=1
发热出汗	否=0;是=1
胸闷	否=0;是=1
胸痛	否=0;是=1
咳嗽	否=0;是=1
咳痰	否=0;是=1
痰中带血	否=0;是=1
咯血	否=0;是=1

1.3 统计学分析 使用 SPSS 21.0 软件对两组样本进行统计分析,采用 χ^2 检验比较肺良性疾病组与肺癌患者组各风险因素组间分布的差异,检验水准 $\alpha=0.05$ 。使用变量平均事件发生数(events per variable,EPV)进行样本量估计,EPV=10。使用 SPSS Clementine 12.0 软件中的随机抽样功能,将样本按照 3:1 的比例随机分为训练集和测试集,其中训练集用来拟合模型,测试集用来评估模型的预测性能,并使用该软件建立三种数据挖掘预测模型。采用 MedCalc V11.6.0.0 软件绘制受试者工作特征(receiver operator characteristic curve,ROC)曲线。三种模型的性能比较使用灵敏度、特异度、预测值、准确度和 ROC 曲线下面积(area under curve,AUC)进行评估。

2 结 果

2.1 研究对象基线特征分析 用 EPV 估计预测模型的样本量,至少需要 160 例肺癌患者和 160 例肺良性疾病患者。本研究共收集 1 020 例患者的资料,其中肺良性疾病患者 560 例,肺癌患者 460 例。两组患者在年龄、吸烟史、饮酒史、肺基础性疾病、粉尘接触史、肺癌家族史、乏力、发热出汗、胸痛、咳嗽、痰中带血和咯血的构成差异均有统计学意义($P<0.05$),而两组患者在心慌、畏寒、胸闷和咳痰的构成差异均无统计学意义($P>0.05$),见表 2。

表 2 肺良性疾病组和肺癌组的基线特征比较($n, \%$)

基线特征	肺良性疾病组($n=560$)	肺癌组($n=460$)	χ^2 值	P 值
年龄(岁)			25.060	<0.001
<45	334(62.3)	202(37.7)		
≥45	226(46.7)	258(53.3)		
吸烟史			37.725	<0.001
否	425(61.5)	266(38.5)		
是	135(41.0)	194(59.0)		
饮酒史			12.435	<0.001
否	497(57.2)	372(42.8)		
是	63(41.7)	88(58.3)		
肺基础性疾病			9.329	0.002
否	45(73.8)	16(26.2)		
是	515(53.7)	444(46.3)		
粉尘接触史			9.816	0.002
否	560(55.3)	452(44.7)		
是	0(0.0)	8(100.0)		
肺癌家族史			18.533	<0.001
否	560(55.7)	445(44.3)		
是	0(0.0)	15(100.0)		
心慌			0.770	0.380
否	540(54.7)	448(45.3)		
是	20(62.5)	12(37.5)		
乏力			13.454	<0.001
否	514(53.5)	447(46.5)		
是	46(78.0)	13(22.0)		
畏寒			3.495	0.062
否	536(54.4)	450(45.6)		
是	24(70.6)	10(29.4)		
发热出汗			57.567	<0.001
否	347(47.5)	384(52.5)		
是	213(73.7)	76(26.3)		
胸闷			1.105	0.293
否	307(56.4)	237(43.6)		
是	253(53.2)	223(46.8)		
胸痛			18.778	<0.001
否	462(58.6)	327(41.4)		

续表 2

基线特征	肺良性疾病组($n=560$)	肺癌组($n=460$)	χ^2 值	P 值
是	98(42.4)	133(57.6)		
咳嗽			17.565	<0.001
否	191(65.2)	102(34.8)		
是	369(50.8)	358(49.2)		
咳痰			0.042	0.838
否	208(55.3)	168(44.7)		
是	352(54.7)	292(45.3)		
痰中带血			59.413	<0.001
否	538(59.1)	373(40.9)		
是	22(20.2)	87(79.8)		
咯血			39.704	<0.001
否	543(57.8)	397(42.2)		
是	17(21.3)	63(78.7)		

2.2 肺癌危险度预测模型的构建 使用 SPSS Clementine 12.0 软件中的随机抽样功能,将肺良性疾病组和肺癌组患者按照 3:1 的比例随机分为两部分,其中训练集包括 765 例样本,测试集包括 255 例样本。分别建立 SVM 模型、决策树 C5.0 模型和 ANN 模型。操作界面如图 1 所示。经过多次训练和优化,SVM 模型的参数设置如下: Use partitioned data: yes; Mode: Expert; Stopping criteria: $1.0E-3$; Bias: 0; Gamma: 0.5。决策树 C5.0 模型的参数设置如下: Group symbolics: no; Use boosting: yes; Pruning severity: 75; Minimum records per child branch: 2; Use global pruning: yes。ANN 模型的参数设置如下: Method: Exhaustive prune; Sample: 50%; Stop on: time (mins) 1 min; Optimize: Memory; Use binary set encoding: yes; Show feedback graph: yes; Mode: Expert。

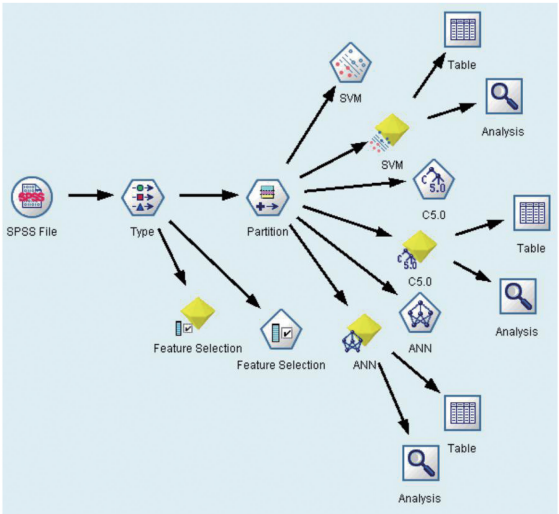


图 1 数据挖掘的流程图

2.3 基线特征在模型构建中的意义 采用 SPSS

Clementine 12.0 对 1 020 例样本的基线资料进行特征提取,并对其重要性进行排序,结果显示痰中带血、发热出汗、吸烟史、年龄、胸痛、咳嗽、饮酒史、胸闷与咳嗽这 9 个变量被筛选出来,用于构建基于数据挖掘技术的肺癌危险度预测模型。其中痰中带血、发热出汗、吸烟史、年龄、胸痛、咳嗽与饮酒史在模型的构建中具有重要作用。而心慌、畏寒、咯血、粉尘接触史、肺基础性疾病、肺癌家族史和乏力这 7 个变量因单个类别分类太大未纳入模型,见表 3。

表 3 基线特征对模型构建的重要性排序

重要性排序	变量	重要性数值
1	痰中带血	1.000
2	发热出汗	1.000
3	吸烟史	1.000
4	年龄	1.000
5	胸痛	1.000
6	咳嗽	1.000
7	饮酒史	1.000
8	胸闷	0.707
9	咳嗽	0.162

2.4 肺癌危险度预测模型的性能比较 基于 SVM、决策树 C5.0 和 ANN 三种数据挖掘技术构建的肺癌危险度预测模型,训练集中 SVM 模型、C5.0 模型和 ANN 模型的准确率分别是 80.7%、77.9%和 77.3%;测试集中 SVM 模型、C5.0 模型和 ANN 模型的准确率分别是 75.3%、72.5%和 76.9%,见表 4。

表 4 三种数据挖掘模型训练集和测试集的分类结果

模型		训练集(<i>n</i> = 765)		测试集(<i>n</i> = 255)	
		肺癌	肺良性	肺癌	肺良性
SVM	阳性	271	71	83	34
	阴性	77	346	29	109
	准确度(%)	80.7	75.3		
C5.0	阳性	226	47	70	28
	阴性	122	370	42	115
	准确度(%)	77.9	72.5		
ANN	阳性	321	147	104	51
	阴性	27	270	8	92
	准确度(%)	77.3	76.9		

测试集中 SVM、决策树 C5.0 和 ANN 模型的灵敏度分别为 74.1%、62.5%和 92.9%;特异度分别为 76.2%、80.4%和 64.3%;阳性预测值分别为 70.9%、71.4%和 67.1%;阴性预测值分别为 79.0%、73.2%和 92.0%;准确度分别为 75.3%、72.5%和 76.9%;AUC

(95%CI) 分别为 0.752(0.694~0.803)、0.715(0.655~0.769)和 0.786(0.730~0.835),见表 5。由此可知,ANN 模型的灵敏度、阴性预测值、准确度和 AUC 均高于 SVM 模型和决策树 C5.0 模型。测试集中三种模型的 ROC 曲线显示 ANN 模型的预测性能最好,见图 2。

表 5 三种数据挖掘模型对测试集的分类效果评估

模型	灵敏度 (%)	特异度 (%)	阳性预测值 (%)	阴性预测值 (%)	准确度 (%)	AUC (95%CI)
SVM	74.1	76.2	70.9	79.0	75.3	0.752(0.694~0.803)
C5.0	62.5	80.4	71.4	73.2	72.5	0.715(0.655~0.769)
ANN	92.9	64.3	67.1	92.0	76.9	0.786(0.730~0.835)

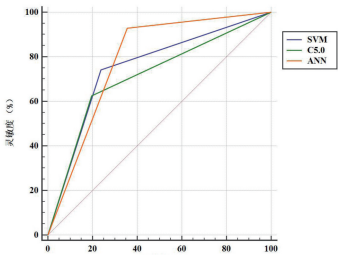


图 2 测试集中三种数据挖掘模型的 ROC 曲线

3 讨论

肺癌是严重威胁人类健康的全球性重大公共卫生问题^[8-9]。流行病学研究发现,对肺癌高危人群开展体检筛查,在可医治的阶段早期发现并进行规范化的治疗,可以有效提高肺癌患者的生存率,并且节约医疗资源^[10-12]。目前临床公认的肺癌高风险人群主要考虑年龄和吸烟状况两个因素^[13],随着流行病学的深入研究发现,肺癌家族史、粉尘接触史等流行病学特征^[14]以及胸痛、咯血等临床症状^[15]与肺癌的发生密切相关。因此综合利用多种肺癌相关危险因素,构建肺癌危险度预测模型,是全球公认经济有效的筛查肺癌高危人群的方法。该研究共纳入年龄、吸烟史和饮酒史等 16 项与肺癌相关的基线资料,使用 SPSS Clementine12.0 对 1 020 例样本的基线资料进行特征提取,结果显示痰中带血、发热出汗和吸烟史等 9 个变量被筛选出来,用于肺癌危险度评价模型的构建。

数据挖掘是融合人工智能、统计学和高性能计算等多个范畴的理论和 技术 发展 起来 的 信息 处理 方法。数据挖掘技术能够从海量和复杂的数据中提取有效的信息,尤其适用于医学预测和诊断^[16]。SVM、决策树 C5.0 和 ANN 是目前常用的数据挖掘模型。该研究分别建立了基于这三种数据挖掘技术的肺癌危险度预测模型,比较发现,ANN 模型的灵敏度、准确度和 AUC 值最高,该模型可以尽可能多地发现肺癌可疑患者,避

免漏诊肺癌患者带来严重后果,是三种模型中最理想的肺癌危险度预测模型。

已有研究将 SVM、决策树 C5.0 和 ANN 用于肺癌风险预测相关模型的构建。Duan 等^[17]基于 372 例肺癌患者和 470 例肺良性疾病患者的 14 项肺癌危险因素特征,分别利用 SVM、决策树 C5.0 和 ANN 三种数据挖掘技术建立了肺癌广谱筛查系统,结果表明 ANN 模型的性能最好, AUC 为 0.736 (95% CI: 0.670 ~ 0.795)。高孜博等^[18]利用 180 例肺癌患者和 240 例肺良性疾病患者的基线特征数据,运用决策树 C5.0 和 ANN 构建了肺癌风险预测模型,结果显示 ANN 模型效果最佳, AUC 为 0.675 (95% CI: 0.628 ~ 0.720), 与本研究结果一致,均显示 ANN 模型的预测性能最优,其原因可能是, ANN 模型与 SVM、决策树 C5.0 相比,能更好地模拟人类大脑神经元的学习方式,充分挖掘医疗大数据中隐藏有用的信息,具有较强的鲁棒性 (robustness) 和泛化能力^[19]。与以往研究相比,本研究建立的肺癌危险度预测模型准确度高、预测效果好,主要得益于扩大了样本量,且纳入的肺癌危险因素较为全面。

综上所述,该研究依据肺癌相关的流行病学和临床症状资料,建立的 ANN 模型用于肺癌风险预测的效果优于 SVM 和决策树 C5.0 模型,在肺癌高危人群筛查与肺癌早诊早治方面具有较大的临床价值。本研究存在的局限性:一方面,纳入研究的样本量相对较少;另一方面,纳入的临床症状过于常规且变量种类有限。因此,未来需要扩大研究样本数,纳入环境、遗传、行为生活方式等因素、并结合肿瘤标志物等客观检测指标,进一步验证并优化 ANN 模型。

参考文献

- [1] Sung H, Ferlay J, Siegel RL, 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.
- [2] Henley SJ, Ward EM, Scott S, et al. Annual report to the nation on the status of cancer, part I: national cancer statistics[J]. Cancer, 2020, 126(10):2225-2249.
- [3] Saberi-Karimian M, Khorasanchi Z, Ghazizadeh H, et al. Potential value and impact of data mining and machine learning in clinical diagnostics[J]. Crit Rev Clin Lab Sci, 2021, 58(4):275-296.
- [4] Valluru D, Jeya IJS. IoT with cloud based lung cancer diagnosis model using optimal support vector machine[J]. Health Care Manag Sci, 2020, 23(4):670-679.
- [5] Vossenhric J, Zech CJ, Heye T, et al. Response prediction of hepatocellular carcinoma undergoing transcatheter arterial chemoembolization; unlocking the potential of CT texture analysis through nested decision tree models[J]. Eur Radiol, 2021, 31(6):4367-4376.
- [6] Xu S, Guan LJ, Shi BQ, et al. Recurrent hemoptysis after bronchial artery embolization: prediction using a nomogram and artificial neural network model[J]. AJR Am J Roentgenol, 2020, 215(6):1490-1498.
- [7] 中华医学会, 中华医学会肿瘤学分会, 中华医学会杂志社. 中华医学会肺癌临床诊疗指南(2019 版)[J]. 中华肿瘤杂志, 2020, 42(04):257-287.
- [8] Henley SJ, Thomas CC, Lewis DR, et al. Annual report to the nation on the status of cancer, part II: progress toward healthy people 2020 objectives for 4 common cancers[J]. Cancer, 2020, 126(10):2250-2266.
- [9] 郑荣寿, 孙可欣, 张思维, 等. 2015 年中国恶性肿瘤流行情况分析[J]. 中华肿瘤杂志, 2019, 41(1):19-28.
- [10] Oudkerk M, Liu S, Heuvelmans MA, et al. Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives[J]. Nat Rev Clin Oncol, 2021, 18(3):135-151.
- [11] de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial[J]. N Engl J Med, 2020, 382(6):503-513.
- [12] 于丽娅, 郭薇, 吕艺, 等. 辽宁省城市肺癌患者 10 年生存率及其影响因素分析[J]. 实用预防医学, 2021, 28(12):1432-1436.
- [13] Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2019: a review of current American Cancer Society guidelines and current issues in cancer screening[J]. CA Cancer J Clin, 2019, 69(3):184-210.
- [14] Hidaka A, Sawada N, Svensson T, et al. Family history of cancer and subsequent risk of cancer: a large-scale population-based prospective study in Japan[J]. Int J Cancer, 2020, 147(2):331-337.
- [15] Wood DE, Kazerooni EA, Baum SL, et al. Lung cancer screening, version 3. 2018, NCCN clinical practice guidelines in oncology[J]. J Natl Compr Canc Netw, 2018, 16(4):412-441.
- [16] Duggirala HJ, Tonning JM, Smith E, et al. Use of data mining at the Food and Drug Administration[J]. J Am Med Inform Assoc, 2016, 23(2):428-434.
- [17] Duan S, Cao H, Liu H, et al. Development of a machine learning-based multimode diagnosis system for lung cancer[J]. Aging (Albany NY), 2020, 12(10):9840-9854.
- [18] 高孜博, 李迪, 段书音, 等. 数据挖掘技术在肺癌危险度预测模型中的应用[J]. 肿瘤防治研究, 2021, 48(5):479-483.
- [19] Yang GR, Wang XJ. Artificial neural networks for neuroscientists: a primer[J]. Neuron, 2020, 107(6):1048-1070.

收稿日期:2021-12-17