

构建并验证反向传播神经网络模型 对筛选重症手足口病影响因素的性能

陈琳¹,冯慧芬²,屈质³,马驰³

(郑州大学第五附属医院¹ 消化内科,² 感染科,郑州 450052;

³新乡医学院第一附属医院全科医学科,新乡 453000)

摘要 目的 通过构建反向传播神经网络(BPNN)模型,筛选重症手足口病(HFMD)临床早期预警指标,探讨神经网络技术在临床中的应用价值。方法 收集河南省新乡医学院第一附属医院感染科及儿科2019年1月至2023年1月收治的HFMD患儿临床资料,使用SPSS Modeler18.0将数据分为70%的训练样本和30%的测试样本,并构建BPNN模型和Logistic模型,对比评估模型预测准确性及筛选效果。结果 共收集589例患儿临床资料进行分析,轻症组324例,重症组265例。BPNN模型和Logistic回归模型的测试集($n=178$)预测正确率为82.02%、84.83%;ROC曲线下面积及95%CI分别为0.791(0.749~0.834)和0.625(0.577~0.674)。BPNN模型输出的预测变量中,对分组影响最大的前5位因素为:最高体温、发热持续时间、谷氨酰转氨酶、天冬氨酸氨基转移酶和球蛋白。两模型输出预测变量重要性结果前十位中重合的有3个,分别为:最高体温、发热持续时间和肢体抖动。结论 BPNN模型和Logistic回归模型在筛选验证重症手足口病危险因素方面均表现良好,但BPNN模型的综合预测性能更好,BPNN模型筛选出的前五名重症HFMD影响因素为最高体温、发热持续时间、谷氨酰转氨酶、天冬氨酸氨基转移酶和球蛋白。

关键词 重症手足口病;反向传播神经网络模型;预测;人工神经网络;Logistic回归模型;机器学习

中图分类号 R 512.57

文献标志码 A **文章编号** 1000-1492(2024)12-2222-08

doi:10.19405/j.cnki.issn1000-1492.2024.12.022

手足口病(hand-foot-mouth disease, HFMD)是由多种肠道病毒(enterovirus, EV)引起的急性发热出疹性传染病,主要通过密切接触、呼吸道和消化道传播,以5岁以内的散居儿童为主,病程多为1周,少数可并发无菌性脑膜炎、急性弛缓性麻痹、心肺衰竭等,不同年份HFMD病原谱中的优势病原呈现动态变化^[1-3]。国家疾病预防控制局法定传染病疫情概况显示, HFMD居2023年全国丙类传染病发病数第2位,发病例数为1 673 481例,约占丙类传染病报告总数的10.5%^[4]。

反向传播神经网络(back propagation neural network, BPNN)是一种常见的人工神经网络,它是一种通过反向传播算法来训练的多层前馈神经网络。BPNN广泛应用于模式识别、数据挖掘、人工智能等

领域,具有较强的非线性映射能力和准确性^[5]。本研究拟通过构建BPNN模型及传统Logistic模型,探究重症HFMD的危险因素,并对构建的临床预测模型进行验证与比较,以期对重症HFMD早期识别提供可靠的临床依据。

1 材料与方法

1.1 病例资料 收集河南省新乡市某医院感染科及儿科2019年1月至2023年1月收治的HFMD患儿的临床资料,根据《手足口病诊疗指南(2018年版)》确定轻症组、重症组诊断标准,并以出院第一诊断为参考依据。本研究通过新乡医学院第一附属医院伦理委员会审批(编号:EC-023-018)。

1.2 资料收集及标准 使用住院信息管理系统导出临床资料并进行汇总分析,将患儿的检验检查结果人工录入Excel软件完成数据收集。纳入标准:①首次诊断为HFMD的病例;②入院前3天的基本住院信息以及各种实验室检查结果(血常规、外周血免疫炎症指标、肝肾功能、电解质、血糖、血脂、凝血功能、呼吸道肠道病原体检测等);③既往无类似传染病接触及感染史;④入院无重叠类似传

2024-11-19 接收

基金项目:国家自然科学基金项目(编号:81473030);河南省医学科技攻关计划项目(编号:LHGJ20190426)

作者简介:陈琳,女,硕士研究生;

冯慧芬,女,博士,主任医师,博士生导师,通信作者, E-mail: hui fen. feng@163.com

染性疾病。排除标准:①入院诊断和出院诊断不一致患儿;②存在免疫力缺陷、先天性心脏病、抑郁症等特殊疾病。

1.3 资料整理 将 Excel 软件数据导入 SPSS 26.0 中,对变量进行命名,根据属性添加标签并定义值。使用描述统计功能,剔除数据缺失比例大于 15% 的自变量,对余变量进行筛查并剔除极值,分类型变量进行众数插补,连续数值型变量应用期望最大化法(EM 法)进行插补。通过利特尔完全随机缺失 MCAR 检验, $P > 0.05$ 。

1.4 统计学处理

1.4.1 组间比较 在 SPSS 26.0 进行数据分析,满足正态分布的连续性变量以 $(\bar{x} \pm s)$ 表示,通过 Levene's 方差齐性检验,使用独立样本 t 检验或校正 t 检验,检验统计量为 t 值;非正态分布的连续性变量以中位数和四分位数间距 $[M(P_{25}, P_{75})]$ 表示,使用 Mann-Whitney (U 检验),检验统计量为 Z 值。计数资料以 $n(\%)$ 表示,行 χ^2 检验或 Fisher 确切概率检验,检验统计量为 Z 值。 $P < 0.05$ 为差异有统计学意义。

1.4.2 模型构建和模型性能诊断评估 使用 SPSS Modeler18.0 软件,将单因素分析筛选出有统计学差异的自变量纳入模型,为增强模型的泛化能力以及防止过拟合,随机将数据划分为 70% 的训练样本和 30% 的测试样本,分区随机种子为 1234567,训练集求得模型参数,测试集评估结果,训练集 $n = 411$,测试集 $n = 178$ 。构建 BPNN 模型,指定防止过度拟合

集合的记录百分比为 30%,随机种子为 229176228;并构建 Logistic 回归模型。使用两模型预测数据的敏感度及特异度并绘制 ROC 曲线,对比模型预测变量的重要性及准确性。

2 结果

2.1 一般基本信息 共收集 638 份病例,因信息不全未被纳入的有 37 例,余 601 份病例,再排除危重症入住 ICU 的患儿 12 人。最终纳入 589 个病例,其中轻症组为 324 例,重症组 265 例。共纳入患儿一般资料、临床表现及实验室检验相关的 65 个变量,单因素分析后发现 43 个变量有统计学差异,见表 1、表 2。

表 1 两组 HFMD 一般资料对比 $[n(\%)/M(P_{25}, P_{75})]$

Tab. 1 Comparison of general data of HFMD between the two groups $[n(\%)/M(P_{25}, P_{75})]$

Variables	Mild ($n = 324$)	Severe ($n = 265$)	χ^2/Z value	P value
Age (months)	27(17,39)	18(12,29)	-5.990	<0.001
Sex			0.866	0.396
Male	192(59.26)	167(63.02)		
Female	132(40.74)	98(36.98)		
Location			0.017	0.933
Village	189(58.33)	156(58.87)		
City	135(41.67)	109(41.13)		
Aggregation situation			7.235	0.009
Kindergarten children	63(19.44)	30(11.32)		
Scattered living	261(80.56)	235(88.68)		

表 2 两组 HFMD 临床资料对比 $[n(\%)/M(P_{25}, P_{75})/(\bar{x} \pm s)]$

Tab. 2 Comparison of clinical data of HFMD between the two groups $[n(\%)/M(P_{25}, P_{75})/(\bar{x} \pm s)]$

Variables	Mild ($n = 324$)	Severe ($n = 265$)	$\chi^2/t/Z$ value	P value
Duration of fever (d)	2(1,3)	3(2,4)	-9.623	<0.001
Highest temperature ($^{\circ}\text{C}$)	38.3(37.5,38.7)	39.0(38.5,39.5)	-11.852	<0.001
Pulse rate(%)	100(92,108)	108(100,115)	-7.521	<0.001
Respiratory rate(%)	23(21,25)	24(22,26)	-4.742	<0.001
Poor spirit	152(46.91)	201(75.85)	50.825	<0.001
Sleep uneasiness	131(40.43)	163(61.51)	25.904	<0.001
Poor eating	227(70.06)	215(81.13)	9.539	0.002
Fever	257(79.32)	259(97.74)	45.527	<0.001
Lethargy	1(0.31)	18(6.79)	19.631	<0.001
Rash	317(97.84)	261(98.49)	0.337	0.762
Vomiting	19(5.86)	57(21.51)	31.750	<0.001
Cough	33(10.19)	44(16.60)	5.285	0.027
Diarrhea	8(2.47)	21(7.92)	9.268	0.003
Easy shock	2(0.62)	124(46.79)	184.830	<0.001
Limb-shaking	1(0.31)	118(44.53)	176.804	<0.001
Tic	0	7(2.64)		0.004 ^a
EV71 positive	69(21.30)	75(28.30)	3.873	0.054

续表 2

Variables	Mild (n = 324)	Severe (n = 265)	$\chi^2/t/Z$ value	P value
WBC ($10^9/L$)	8.80 (6.90, 11.10)	10.30 (7.9, 14.05)	-5.461	<0.001
RBC ($10^{12}/L$)	4.47 (4.26, 4.68)	4.42 (4.19, 4.66)	-1.921	0.055
Hb (g/L)	118.25 ± 10.45	113.74 ± 10.55	5.192	<0.001
PLT ($10^9/L$)	308.50 (257.25, 354.75)	328.00 (274.00, 395.00)	-3.416	0.001
Neutrophils (%)	43.90 (32.35, 56.58)	51.50 (35.70, 66.20)	-3.721	0
Lymphocytes (%)	47.10 (34.93, 56.70)	40.30 (26.90, 55.75)	-3.287	0.001
Monocytes (%)	6.01 (4.80, 7.90)	6.00 (4.50, 7.95)	-0.889	0.374
Eosinophils (%)	1.65 (0.60, 3.18)	0.63 (0.10, 1.70)	-6.884	<0.001
Basophils (%)	0.40 (0.20, 0.80)	0.30 (0.20, 0.50)	-4.343	<0.001
Neutrophils ($10^9/L$)	3.52 (2.41, 5.44)	5.19 (2.69, 8.18)	-5.191	<0.001
Lymphocytes ($10^9/L$)	3.77 (2.85, 5.07)	3.88 (2.64, 5.21)	-0.687	0.492
Monocytes ($10^9/L$)	0.52 (0.37, 0.72)	0.63 (0.41, 0.89)	-3.093	0.002
Eosinophils ($10^9/L$)	0.14 (0.05, 0.25)	0.08 (0.01, 0.19)	-5.102	<0.001
Basophils ($10^9/L$)	0.03 (0.02, 0.07)	0.03 (0.02, 0.05)	-1.451	0.147
CRP (mg/L)	6.80 (2.50, 13.61)	9.24 (3.14, 24.43)	-3.644	<0.001
K ⁺ (mmol/L)	4.74 ± 0.47	4.75 ± 0.52	-0.121	0.904
Na ⁺ (mmol/L)	139 (137, 140)	138 (136, 140)	-3.187	0.001
Ca ²⁺ (mmol/L)	2.44 (2.36, 2.51)	2.45 (2.37, 2.54)	-1.417	0.156
Cl ⁻ (mmol/L)	104.60 (103.10, 105.90)	103.80 (101.70, 105.55)	-3.919	<0.001
GLU (mmol/L)	4.28 (3.79, 4.80)	4.65 (3.91, 5.19)	-3.749	<0.001
Total protein (g/L)	64.36 ± 4.97	65.15 ± 5.28	-1.859	0.063
Albumin (g/L)	44.40 (42.53, 46.10)	44.50 (42.40, 46.50)	-0.403	0.687
Globulin (g/L)	19.70 (17.33, 22.30)	20.40 (17.70, 23.30)	-2.094	0.036
A/G	2.26 (1.97, 2.55)	2.20 (1.89, 2.47)	-1.850	0.064
TBIL ($\mu\text{mol/L}$)	6.80 (4.90, 8.80)	6.10 (4.50, 8.15)	-2.453	0.014
DBIL ($\mu\text{mol/L}$)	1.90 (1.30, 2.70)	1.80 (1.30, 2.40)	-1.196	0.232
IBIL ($\mu\text{mol/L}$)	4.80 (3.50, 6.48)	4.20 (3.12, 5.85)	-2.467	0.014
AST (U/L)	30 (26, 35)	32 (27, 39)	-3.010	0.003
ALT (U/L)	14 (12, 18)	15 (11, 19)	-1.036	0.300
AST/ALT	2.00 (1.70, 2.50)	2.10 (1.68, 2.59)	-0.995	0.320
GGT (U/L)	10 (8, 12)	10 (9, 13)	-2.119	0.034
ALP (U/L)	199.50 (169.00, 237.75)	196.00 (164.50, 233.00)	-0.613	0.540
Bile acid ($\mu\text{mol/L}$)	5.20 (3.15, 8.78)	6.19 (3.67, 9.65)	-2.233	0.026
LDH (U/L)	262.50 (233.73, 293.75)	269.00 (245.00, 299.50)	-2.326	0.020
α -HBDH (U/L)	213.00 (193.00, 238.75)	228.00 (204.00, 253.50)	-3.624	<0.001
CK (U/L)	75.00 (56.25, 98.75)	70.00 (51.00, 96.50)	-1.250	0.211
CK-MB (U/L)	21 (17, 26)	19 (15, 24)	-2.718	0.007
UA ($\mu\text{mol/L}$)	244.11 (202.25, 296.50)	237.00 (189.00, 283.50)	-1.878	0.060
SCR ($\mu\text{mol/L}$)	22.00 (18.00, 28.00)	22.00 (17.85, 28.65)	-0.740	0.459
BUN (mmol/L)	3.67 (3.04, 4.43)	3.42 (2.66, 4.23)	-2.331	0.020
Urine PH	6.00 (5.50, 6.50)	6.00 (5.50, 6.50)	-2.233	0.026
PRO positive	10 (3.09)	6 (2.26)	0.373	0.617
Urine GLU positive	6 (1.85)	18 (6.79)	9.103	0.003
KET positive	91 (28.09)	68 (25.66)	0.435	0.516

^a: Fisher exact test.

2.2 模型构建 将单因素分析中有统计学差异的变量纳入构建模型,均使用分区数据。BPNN 模型的参数包括:建模选择类神经网络,构建新模型多层感知器,Modeler 自动计算隐藏层,并输出预测变量重要性,前十位分别为:最高体温、发热持续时间、谷氨酰转肽酶、天冬氨酸氨基转移酶、球蛋白、呼吸频

率、中性粒细胞绝对数、脉搏、尿素、肢体抖动,见图 1。自动构建后的模型中输入层有 11 个输入因子,隐藏层为一个,包含 3 个隐藏单元(含常数项),见图 2。将所有单因素分析中具有统计学意义的 43 个变量全部纳入到多因素 Logistic 回归模型,对 Logistic 回归模型的逐步回归方法进行多次尝试,最终

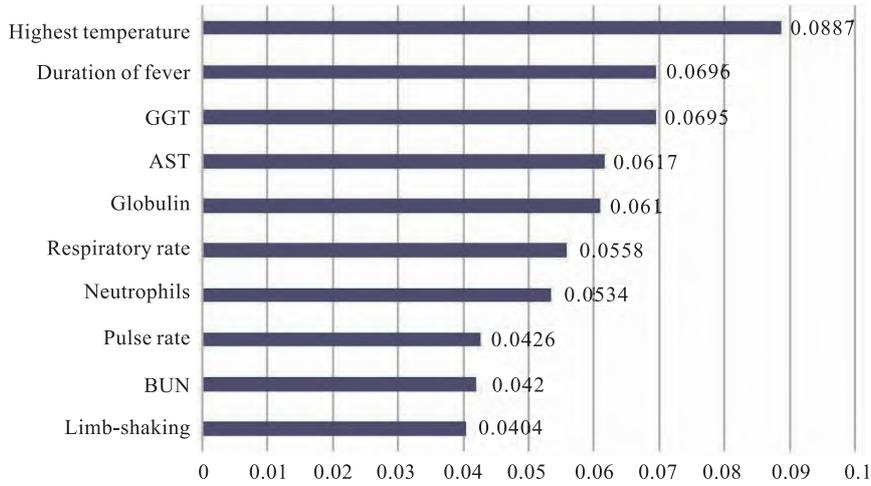


图 1 BP 神经网络模型预测变量重要性排序

Fig. 1 Importance of BPNN model predictor variables

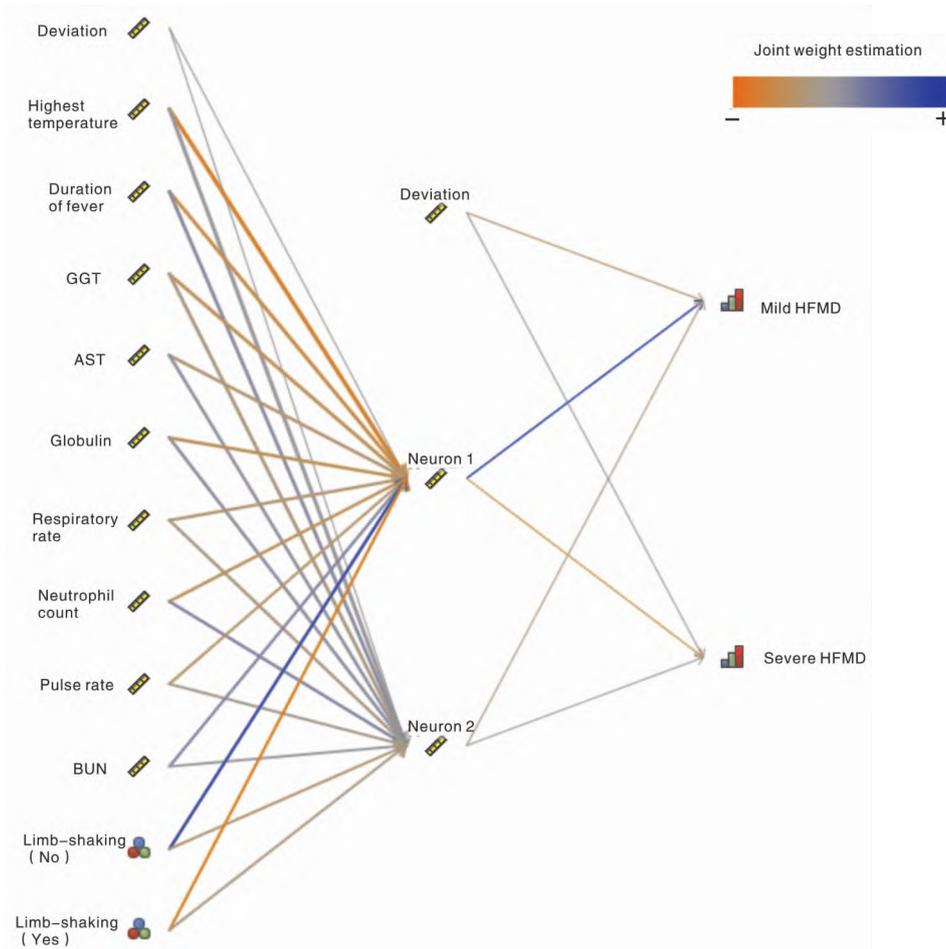


图 2 BP 神经网络模型结构图

Fig. 2 BPNN model structure diagram

选择进入法,结果显示发热持续时间、最高体温、血小板、C 反应蛋白、血糖影响手足口病由轻症发展为

重症的危险因素,而嗜酸性粒细胞百分比、嗜睡、呕吐、易惊、肢体抖动为保护因素(表 3)。

表 3 影响手足口病由轻症发展为重症的多因素 Logistic 回归分析

Tab. 3 Multivariate Logistic regression analysis of the factors affecting the development of hand-foot-mouth disease from mild to severe

Variables	β	SE	Wald	P value	OR	95% CI
Duration of fever	0.411	0.160	6.623	0.010	1.508	1.103 – 2.061
Highest temperature	2.249	0.461	23.805	<0.001	9.480	3.841 – 23.399
PLT	0.007	0.003	4.720	0.030	1.007	1.001 – 1.014
Percentage of eosinophils	-1.244	0.586	4.504	0.034	0.288	0.091 – 0.909
CRP	0.039	0.016	6.250	0.012	1.040	1.008 – 1.072
GLU	0.837	0.244	11.798	0.001	2.309	1.432 – 3.722
Lethargy	-4.423	1.834	5.818	0.016	0.012	0.000 – 0.436
Vomiting	-1.814	0.766	5.612	0.018	0.163	0.036 – 0.731
Easy shock	-6.026	1.307	21.262	<0.001	0.002	0.000 – 0.031
Limb-shaking	-7.243	1.802	16.164	<0.001	0.001	0.000 – 0.024
Constant term	0.906	34.446	0.001	0.979		

2.3 模型预测效果评价 利用 70% 的训练样本构建模型得出可靠的参数,并用 30% 的测试样本对所建立的模型进行验证,使用 Modeler 分析节点对两种模型进行整体预测性能评价,输出预测模型的重合矩阵,Logistic 回归模型训练集、测试集正确率分别为 94.16% 和 84.83%,BPNN 模型正确率分别为 89.78% 和 82.02%。两种模型有较佳的学习能力,模型稳定性良好。见表 4。绘制两种模型受试者工作特征 (ROC) 曲线,BPNN 模型在筛选验证重症手足口病危险因素方面均表现良好,见图 3。并计算相应诊断性能评价指标,可以看出 BPNN 模型在特异度方面与 Logistic 回归模型差异不大,而灵敏度则远大于 Logistic 回归模型。BPNN 模型 AUC 值较高,效果更好,由此可得出其综合预测性能更好,见表 5。

表 4 Logistic 回归模型与 BP 神经网络模型预测的重合矩阵

Tab. 4 Coincidence matrix predicted by Logistic regression model and BPNN model

Model	HFMD	Training samples (n = 411)		Testing samples (n = 178)	
		Mild	Severe	Mild	Severe
Logistic	Mild	213	7	90	14
	Severe	17	174	13	61
	Accurate rate	94.16%		84.83%	
BPNN	Mild	196	24	84	20
	Severe	18	173	12	62
	Accurate rate	89.78%		82.02%	

表 5 模型诊断性能评价

Tab. 5 Evaluation of model diagnostic

Model	AUC	95% CI	Max Youden index	Sensitivity	Specificity
Logistic	0.625	0.577 – 0.674	0.325	36.2%	96.3%
BPNN	0.791	0.749 – 0.834	0.669	74.0%	92.9%

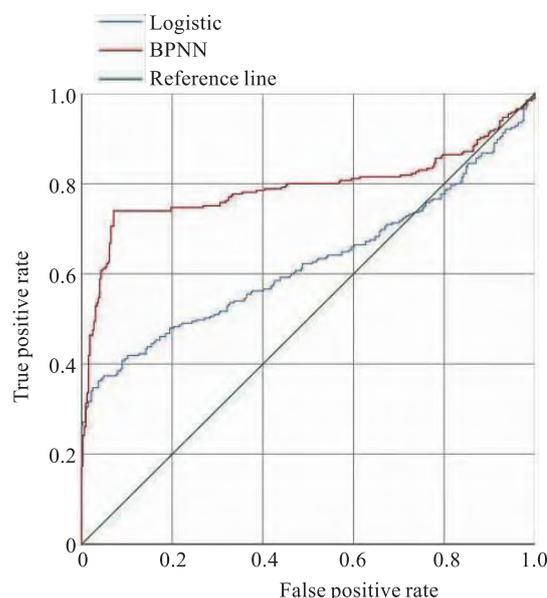


图 3 Logistic 回归模型与 BP 神经网络模型的 ROC 曲线

Fig. 3 ROC curve of Logistic regression model and BPNN model

3 讨论

本研究通过构建 BPNN 模型、Logistic 回归模型实现对重症 HFMD 的预测变量筛选。两模型筛选预测变量重要性结果前十位中,重合的 3 个变量分别为:最高体温、发热持续时间、肢体抖动。这与指南中临床分期、分型与重症手足口病的早期识别中的要素相符合^[1]。在 BPNN 模型的预测变量中,天冬氨酸氨基转移酶、球蛋白,中性粒细胞绝对数列前 10。EV71 感染所致的 HFMD 患儿会出现体液免疫、细胞免疫以及先天免疫功能紊乱,柯江维等^[6]对 HFMD 患儿免疫状态在轻、重、危重 3 组与健康组对照的研究中发现,免疫球蛋白浓度和免疫因子含量会随病情的加重不断增高,与本研究结果相符。

Zhang et al^[7]的实验室研究数据表明,中性粒细胞激活参与了严重 HFMD 临床病例的急性心肌损伤。重症患者外周血中性粒细胞显著升高,并且天冬氨酸氨基转移酶、肌酸激酶、肌酸激酶同工酶、乳酸脱氢酶水平也显著升高,这与 BPNN 模型预测的心肌酶改变相符合。陈静等^[8]研究发现,关注重症患儿早期预警指标如持续高热、呕吐、肢体抖动、心肌酶等,并采取针对性干预如隔离、消毒、定时通风、保持口腔清洁,加强对呼吸系统监护、提高血糖监测频率等,有助于阻断病情进展,改善预后。Zhou et al^[9]在关于 2019 年冠状病毒病大流行期间的研究发现,2020 年至 2022 年实验室确诊的 HFMD 病例数较 2018 年和 2019 年急剧下降,非药物干预措施如使用口罩、保持手部卫生、保持社交距离和关闭学校,也在缓解 HFMD 传播方面发挥作用。

马晓梅等^[10]在对重症 HFMD 临床早期预警指标的研究中发现,BPNN 型预测能力优于算法模型和 Logistic 回归模型。此外,BPNN 模型在风疹、疟疾、戊肝等传染性疾病的预测工作中也有良好的应用价值^[11-13]。近年来,有很多关于 HFMD 病情预测方面的机器学习研究,如决策树、判别分析、随机森林模型、ARIMA-GRNN 组合模型等,可发现不同机器学习较传统 Logistic 回归算法具有更高的准确性和总体预测优越性^[14-16]。关于本研究中 Logistic 模型测试集的正确率稍高于 BPNN 模型,但正确率只是某个随机样本的一个属性指标,会随着不同分类阈值的改变而不同,无法在统计学上得出模型更优的结论,仅作为参考对比。而 ROC 曲线需要计算各种阈值情况下的特异度和敏感度。计算曲线下的面积 AUC 值,它不关注某个截断值的表现如何,而综合所有截断值的预测性能,当模型预测正确率偏低而 AUC 值较高时,一般情况下选择较高的 AUC,而不是较高的正确率模型^[16]。HFMD 重症化过程各因素间存在着复杂的非线性关系,BPNN 模型比线性模型或 Logistic 回归模型对资料的性质要求较低,可以拟合任意非线性关系^[17]。

使用传统统计学分析方法时,需考虑资料类型、分布特征、样本量大小等因素,如数据处理分组不当会导致统计分析结果和结论出现偏差。神经网络模型可以不受限制地学习样本和训练样本,调整自身参数,输出信息完整有序,尽可能接近真实情况,对于未知关系变量和复杂关系处理具有优势^[18]。但 BPNN 训练时间较长,训练过程需要大量的数据和计算资源;网络结构和参数的选择对训练

结果有较大影响,容易产生过拟合,需要采用一些优化算法来提高训练效果。

综上所述,本研究结果显示 BPNN 模型可以用于预测重症 HFMD 患者病情进展的影响因素,且相较于传统 Logistic 回归模型具有一定的优越性。对于人工神经网络的集成学习算法及组合模型升级优化的预测价值有待进一步挖掘。

参考文献

- [1] Li X W, Ni X, Qian S Y, et al. Chinese guidelines for the diagnosis and treatment of hand, foot and mouth disease (2018 edition)[J]. *World J Pediatr*, 2018, 14(5): 437-47. doi:10.1007/s12519-018-0189-8.
- [2] 马婉婉, 龚磊, 肖永康, 等. 2015—2022 年安徽省手足口病流行病学特征及病原谱变化[J]. *安徽医科大学学报*, 2023, 58(10): 1763-8. doi:10.19405/j.cnki.issn1000-1492.2023.10.026.
- [2] Ma W W, Gong L, Xiao Y K, et al. Epidemiological characteristics and pathogen spectrum changes of hand-foot-mouth disease in Anhui province from 2015 to 2022[J]. *Acta Univ Med Anhui*, 2023, 58(10): 1763-8. doi:10.19405/j.cnki.issn1000-1492.2023.10.026.
- [3] Solomon T, Lewthwaite P, Perera D, et al. Virology, epidemiology, pathogenesis, and control of enterovirus 71[J]. *Lancet Infect Dis*, 2010, 10(11): 778-90. doi:10.1016/S1473-3099(10)70194-8.
- [4] 国家疾病预防控制中心. 2023 年全国法定传染病疫情概况[R/OL]. (2024-09-18) [2024-09-18]. <https://www.ndcpa.gov.cn/>
- [4] National Bureau of Disease Control and Prevention. National notifiable infectious disease epidemic situation in 2023 [R/OL]. (2024-09-18) [2024-09-18]. <https://www.ndcpa.gov.cn/>.
- [5] 刘璐. 基于神经网络技术的老年龋预测模型的构建及预测方法学比较的泛化能力验证研究[D]. 沈阳: 中国医科大学, 2021. doi:10.27652/d.cnki.gzyku.2021.000270.
- [5] Liu L. Construction of prediction models of dental caries in the elderly based on artificial neural network technology and validation of the generalization ability of methodology comparisons[D]. Shenyang: China Medical University, 2021. doi:10.27652/d.cnki.gzyku.2021.000270.
- [6] 柯江维, 陈强, 段荣, 等. EV71 病毒引起手足口病患儿免疫状态变化的研究[J]. *实验与检验医学*, 2014, 32(1): 5-7, 40. doi:10.3969/j.issn.1674-1129.2014.01.002.
- [6] Ke J W, Chen Q, Duan R, et al. Study of the changes of immune status in children with HFMD caused by EV71 virus[J]. *Exp Lab Med*, 2014, 32(1): 5-7, 40. doi:10.3969/j.issn.1674-1129.2014.01.002.
- [7] Zhang Y, Chen S, Sun T, et al. Abundant neutrophil-initiated acute myocardial injury following coxsackievirus A6 infection[J]. *J*

- Infect Dis, 2024, 229 (5): 1440 – 50. doi:10.1093/infdis/jiad407.
- [8] 陈 静, 孟 慧. 重症早期预警护理对手足口病患儿病情恢复和并发症发生的影响研究[J]. 中国医刊, 2021, 56(1): 88 – 90. doi:10.3969/j.issn.1008-1070.2021.01.025.
- [8] Chen J, Meng H. Study on the influence of early warning nursing on the recovery and complications of children with hand-foot-mouth disease[J]. Chin J Med, 2021, 56(1): 88 – 90. doi:10.3969/j.issn.1008-1070.2021.01.025.
- [9] Zhou K, Ding Z, Hu B, et al. Circulating trends of hand, foot, and mouth disease in Hubei province, China; impact from the COVID-19 pandemic[J]. Heliyon, 2023, 9(12): e22872. doi:10.1016/j.heliyon.2023.e22872.
- [10] 马晓梅, 徐学琴, 闫国立, 等. BP神经网络和决策树分析在重症手足口病临床早期预警指标中的应用[J]. 中国卫生统计, 2019, 36(3): 381 – 3.
- [10] Ma X M, Xu X Q, Yan G L, et al. Application of BP neural network and decision tree analysis in clinical early warning index of severe hand-foot-mouth disease[J]. Chin J Health Stat, 2019, 36 (3): 381 – 3.
- [11] Verma A K, Kuppili V, Srivastava S K, et al. A new backpropagation neural network classification model for prediction of incidence of malaria[J]. Front Biosci (Landmark Ed), 2020, 25 (2): 299 – 334. doi:10.2741/4808.
- [12] Chen Q, Zhao H, Qiu H, et al. Time series analysis of rubella incidence in Chongqing, China using SARIMA and BPNN mathematical models[J]. J Infect Dev Ctries, 2022, 16(8): 1343 – 50. doi:10.3855/jide.16475.
- [13] Ren H, Li J, Yuan Z A, et al. The development of a combined mathematical model to forecast the incidence of hepatitis E in Shanghai, China[J]. BMC Infect Dis, 2013, 13(1): 421. doi:10.1186/1471-2334-13-421.
- [14] Sui M, Huang X, Li Y, et al. Application and comparison of laboratory parameters for forecasting severe hand-foot-mouth disease using logistic regression, discriminant analysis and decision tree [J]. Clin Lab, 2016, 62(6): 1023 – 31. doi:10.7754/clinlab.2015.150917.
- [15] 霍荣瑞, 吕榜军, 于雪婷, 等. 基于 ARIMA-GRNN 组合模型预测手足口病发病率[J]. 应用预防医学, 2016, 22(6): 481 – 5.
- [15] Huo R R, Lv B J, Yu X T, et al. Forecasting incidence of hand-foot-mouth disease with ARIMA-GRNN hybrid model[J]. J Appl Prev Med, 2016, 22(6): 481 – 5.
- [16] 王 斌, 冯慧芬, 黄 平, 等. 随机森林模型在重症手足口病预测中的应用价值[J]. 郑州大学学报(医学版), 2018, 53 (6): 747 – 51. doi:10.13705/j.issn.1671-6825.2018.03.135.
- [16] Wang B, Feng H F, Huang P, et al. Application value of random forest model in prediction of severe hand-foot-mouth disease[J]. J Zhengzhou Univ Med Sci, 2018, 53 (6): 747 – 51. doi:10.13705/j.issn.1671-6825.2018.03.135.
- [17] 李浩月. 基于 BP 神经网络的骨质疏松预测模型构建及效果评价[D]. 武汉: 武汉科技大学, 2019.
- [17] Li H Y. The establishment and effect evaluation of osteoporosis prediction model based on BPNN[D]. Wuhan: Wuhan University of Science and Technology, 2019.
- [18] Ding H, Lu Q, Gao H, et al. Non-invasive prediction of hemoglobin levels by principal component and back propagation artificial neural network[J]. Biomed Opt Express, 2014, 5(4): 1145 – 52. doi:10.1364/BOE.5.001145.

To construct and verify the performance of the back propagation neural network model in screening the influencing factors of severe hand-foot-mouth disease

Chen Lin¹, Feng Huifen², Qu Zhi³, Ma Chi³

(¹Dept of Gastroenterology, ²Dept of Infectious Diseases, The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou 450052; ³General Medical Department, The First Affiliated Hospital of Xinxiang Medical University, Xinxiang 453000)

Abstract Objective To explore the clinical application value of neural network technology, and to screen the clinical early warning indicators of severe hand-foot-mouth disease (HFMD) by constructing a back propagation neural network (BPNN) model. **Methods** The clinical data of children with HFMD admitted to the Department of Infectious Diseases and Pediatrics of the First Affiliated Hospital of Xinxiang Medical University from January 2019 to January 2023 were collected. The data were divided into 70% training samples and 30% test samples using SPSS Modeler 18.0. The BPNN model and Logistic model were constructed to compare and evaluate the prediction accuracy and screening effect of the model. **Results** The clinical data of 589 children were collected and analyzed, including 324 cases in the mild group and 265 cases in the severe group. The prediction accuracy of the test set ($n = 178$) of BP neural network model and Logistic regression model was 82.02% and 84.83%,

网络出版时间:2024-12-05 14:07:11 网络出版地址:10.19405/j.cnki.issn1000-1492.2024.12.023

CST4、HP 抗体分型及 HER2 在胃癌中的表达及其诊断价值

金水¹,王瑞²,张丹丹¹,路亮¹(安徽医科大学附属巢湖医院¹ 消化内科、² 病理科,巢湖 238000)

摘要 **目的** 探讨人半胱氨酸蛋白酶抑制剂 S(CST4)、幽门螺杆菌(HP)抗体分型及人表皮生长因子受体 2(HER2)在胃癌患者中的表达及其对于预测胃癌淋巴结转移和 TNM 分期的预测价值。**方法** 选取 2021 年 12 月至 2023 年 12 月入住安徽医科大学附属巢湖医院接受手术的 150 例胃癌患者作为研究对象,术前行血清 CST4 及 HP 抗体分型检测,术后采用免疫组化检测胃癌组织中 HER2 表达水平。比较不同血清 CST4、HP 抗体分型及 HER2 表达的胃癌患者手术后临床病理资料并进行相关性分析。通过 ROC 曲线比较血清 CST4、HP 抗体分型及 HER2 表达单项及联合检测对于胃癌淋巴结转移及 TNM 分期的预测价值。**结果** CST4、HPI 型及 HER2 阳性表达患者的肿瘤浸润深度更深,更易发生远处淋巴结转移,TNM 分期更高($P < 0.05$),同时 CST4 阳性表达患者肿瘤更易发生脉管和神经侵犯($P < 0.05$),HPI 型阳性表达患者肿瘤多发生于贲门、胃底部,且肿瘤直径较大、更易侵犯脉管($P < 0.05$),而 HER2 与肿瘤分化程度呈负相关,多发生高中分化肿瘤($P < 0.05$)。ROC 曲线分析显示,患者血清 CST4、HPI 型及 HER2 表达单项及联合检测预测胃癌淋巴结转移的 AUC 分别为 0.780、0.676、0.611 和 0.872($P < 0.05$),诊断胃癌 TNM 分期的 AUC 分别为 0.762、0.635、0.613 和 0.801($P < 0.05$),联合检测的诊断效能均高于各单项检测。**结论** 血清 CST4、HPI 型及 HER2 表达与肿瘤浸润深度、淋巴结转移、TNM 分期密切相关,血清 CST4、HPI 型和 HER2 表达可作为胃癌患者有无淋巴结转移及 TNM 分期的重要预测指标,为临床诊断提供新的思路。

关键词 人半胱氨酸蛋白酶抑制剂 S;幽门螺杆菌抗体分型;人表皮生长因子受体 2;胃癌;病理分期;胃癌淋巴结转移

中图分类号 R 446.11

文献标志码 A **文章编号** 1000-1492(2024)12-2229-08

doi:10.19405/j.cnki.issn1000-1492.2024.12.023

2024-10-22 接收

基金项目:安徽省高校自然科学研究重点项目(编号:2023AH050581);安徽医科大学校科研基金项目(编号:2022xkj208)

作者简介:金水,男,副主任医师,硕士;

路亮,男,主任医师,硕士生导师,通信作者,E-mail:105982304@qq.com

胃癌仍然是全球范围内的一种重要癌症,2020 年研究^[1]显示,每年有超过 100 万新发病例,占有新发恶性肿瘤的 5.6%,约有 77 万人死亡(相当于全球每 13 名患者死亡中就有一名胃癌患者),其发病率在全球排名第五,病死率在全球排名第四。胃癌是中国第二大癌症,也是癌症相关死亡的第二大原因^[2]。中国的病死率/发病率(0.845)和 5 年患

respectively. The area under the ROC curve (95% CI) was 0.791 (0.749 - 0.834) and 0.625 (0.577 - 0.674), respectively. Among the predictive variables output by the BPNN model, the top five factors that had the greatest impact on the grouping were highest body temperature, duration of fever, glutamyl transpeptidase, aspartate aminotransferase, and globulin. There were 3 of the top 10 overlaps in the importance of the predictive variables output by the two models, which were the highest body temperature, duration of fever and limb shaking. **Conclusion** The BPNN model and Logistic regression model perform well in screening and verifying the risk factors of severe hand-foot-mouth disease, but the comprehensive prediction performance of BP neural network model is better. The top five influencing factors of severe HFMD screened by the BPNN model are the highest body temperature, duration of fever, glutamyl transpeptidase, aspartate aminotransferase and globulin.

Key words severe hand-foot-mouth disease; back propagation neural network; prediction; artificial neural network; Logistic regression model; machine learning

Fund programs National Natural Science Foundation of China (No. 81473030); Medical Science and Technology Research Plan Project of Henan Province (No. LHGJ20190426)

Corresponding author Feng Huifen, E-mail: huifen.feng@163.com