

手足口病重症化危险因素 BP 神经网络模型预测分析*

马晓梅, 隋美丽, 段广才, 陈帅印, 张荣光, 郝园林, 范清堂

摘要:目的 探讨 BP 神经网络(BPNN)模型分析在儿童手足口病重症化危险因素预测中的作用,为手足口病的临床诊断提供参考依据。方法 整群抽取河南省郑州市某医院 2013 年 4—6 月收治的 233 例出现发病症状到入院时间 < 72 h 的手足口病患者作为调查对象进行问卷调查,采用 MATLAB 7.0 神经网络工具箱构建 BPNN 模型,分析影响手足口病重症化危险因素的平均影响值(MIV),按 MIV 值的绝对值的大小排出因子顺位,与多因素 logistic 回归模型进行比较。结果 BPNN 模型最终的网络结构为 26→8→1,影响手足口病重症化的前 10 位危险因素(MIV 绝对值)依次为:易惊(0.691 4)、颈强直(0.537 3)、呕吐(0.453 8)、血糖升高(0.429 9)、手足抖动(0.381 3)、精神差(0.328 3)、热峰 ≥ 39 °C(0.308 6)、白细胞升高(0.290 2)、热程 ≥ 3 d(0.262 1)、嗜睡(0.242 7);多因素 logistic 回归分析的主要危险因素(OR 值)依次为:颈强直(183.633)、易惊(158.868)、呕吐(59.347)、血糖升高(23.133)、白细胞升高(12.243)、热程 ≥ 3 d(7.765)、手足抖动(5.738)、精神差(4.452);饱和对数线性模型分析结果显示,热峰 ≥ 39 °C 与精神差和白细胞升高均有交互作用($P < 0.05$)。结论 BPNN 模型可以较好地反映手足口病与各影响因素间复杂的非线性关系,网络的拟合效果和分类正确率均较好,可用于手足口病重症化危险因素的分析研究。

关键词:BPNN 模型;手足口病;重症化;危险因素

中图分类号:R 512.5 文献标志码:A 文章编号:1001-0580(2014)06-0758-04 DOI:10.11847/zgggws2014-30-06-20

Application of back propagation neural network model in prediction of risk factors of severe hand-foot-mouth disease

MA Xiao-mei, SUI Mei-li, DUAN Guang-cai, et al (Department of Molecular Epidemiology, College of Public Health, Zhengzhou University, Zhengzhou, Henan Province 450001, China)

Abstract: **Objective** To predict risk factors of severe progression of hand-foot-mouth disease (HFMD) with back propagation neural network (BPNN) model and to provide a reference for the diagnosis of severe HFMD in children. **Methods** Clinical data on 233 child HFMD cases hospitalized in Zhengzhou Children's Hospital during April 2013 through June 2013 were selected and surveyed with a questionnaire within 72 hours of the hospitalization. A BPNN model was established with neural network toolbox of MATLAB 7.0 software and mean impact values (MIV) of the risk factors for severe progression of HFMD were obtained and ranked descendingly according to their absolute values. Then the results of MIV were compared with those of multi-factor logistic regression. **Results** The final network structure of BPNN was 26→8→1. The top ten risk factors for the severity of HFMD (MIV absolute value) were panic tendency (0.691 4), stiff neck (0.537 3), vomiting (0.453 8), elevated blood sugar (0.429 9), shake of hands and feet (0.381 3), poor spirit (0.328 3), high body temperature above 39 °C (0.308 6), increase of white blood cell count (0.290 2), fever lasting for more than 3 days (0.262 1), and drowsiness (0.242 7); the risk factors ascertained with multi-factor logistic regression analysis (odds ratio) were stiff neck (183.633), panic tendency (158.868), vomiting (59.347), elevated blood sugar (23.133), increase of white blood cells count (12.243), fever lasting for more than 3 days (7.765), shake of hands and feet (5.738), and poor spirit (4.452). High body temperature above 39 °C, poor spirit and the increase of white blood cell count interacted with each other based on the results of saturated log linear model. **Conclusion** BPNN model can well describe the complex nonlinear relationship between risk factors and the severity of HFMD with good fitness and correct discrimination and could be used to analyze the risk factors of severe progression of HFMD.

Key words: back propagation neural network model; hand-foot-mouth disease; deterioration; risk factor

手足口病属于小 RNA 肠道病毒属病毒引起的急性肠道传染病,发病急、传染性强,近年来在我国较为流行。手足口病多为轻症病例,部分会发展为重症病例,甚至死亡^[1-3]。因此掌握手足口病重症化预警指标,对防止手足口病的暴发流行非常重要。由于影响手足口病重症化过程的各因素间存在着复杂的非线性关系,传统方法如线性模型或者 logistic

回归模型存在局限性,模型拟合难度较大;而神经网络有表达任意非线性关系以及学习能力的特点它无需套用某种理论框架或假定各种前提条件,只需直接输入数据,其网络就能建立各影响因素与结论之间的高非线性映像,提供比传统模型更为丰富的信息,可为解决这些问题提供了新的思路和方法^[4-6]。为探讨 BP 神经网络(back propagation neural net-

* 基金项目:国家自然科学基金(81172740)

作者单位:郑州大学公共卫生学院分子流行病学教研室,河南郑州 450001

作者简介:马晓梅(1988-),女,河南周口人,硕士在读,研究方向:分子流行病学。

通讯作者:段广才, E-mail:gcduan@zzu.edu.cn

数字出版日期:2014-4-10 13:56

数字出版网址: <http://www.cnki.net/kcms/detail/21.1234.R.20140410.1356.013.html>

work, BPNN) 在儿童手足口病重症化危险因素预测中的应用, 为手足口病的临床诊断提供参考依据, 对河南省郑州市某医院 2013 年 4-6 月收治的 233 例出现发病症状到入院时间 < 72 h 的手足口病患儿作为调查对象进行问卷调查, 采用 MATLAB 7.0 神经网络工具箱构建 BPNN 模型, 分析影响手足口病重症化危险因素的平均影响值 (mean impact value, MIV), 按 MIV 值绝对值的大小排出因子顺位, 与多因素 logistic 回归模型进行比较。结果报告如下。

1 对象与方法

1.1 对象 采用整群抽样方法, 抽取河南郑州某医院 2013 年 4-6 月收治的 233 例手足口病患儿作为调查对象进行问卷调查。所有患儿年龄均 ≤ 108 个月, 从出现发病症状到入院时间均 < 72 h。

1.2 方法 参照《手足口病预防控制指南 (2009 版)》中《手足口病重症或死亡病例个案调查表》^[7], 经过多次预调查及专家论证后, 自行设计调查问卷。内容包括患者及家庭的一般人口学特征、疾病史、临床症状及体征以及生化室检查的相关资料。依据卫生部颁发的《手足口病诊疗指南 (2010 年版)》^[8], 将其中 137 例重症病例作为病例组, 96 例轻症病例作为对照组, 分析手足口病重症化的危险因素。调查表由经统一培训的调查员分别对病例组和对照组的家长进行面访调查, 从调查问卷中选取其中 26 个因素进行危险因素分析, 分别为: 性别、月龄、易惊、呼吸节律不齐、手足抖动、嗜睡、四肢发凉、血糖升高、热程 ≥ 3 d、热峰 ≥ 39 °C、心率 ≥ 140 次/min、呼吸 ≥ 30 次/min、发热、精神差、烦躁、抽搐、呕吐、颈强直、肢体无力、白细胞升高、中性细胞比率升高、淋巴细胞比率升高、C 反应蛋白升高、降钙素原升高、心电图异常、胸部 X 线改变。

巴细胞比率升高、C 反应蛋白升高、降钙素原升高、心电图升高、胸部 X 线改变。

1.3 统计分析 采用 Epi Data 3.0 建立数据库, 应用 SPSS 18.0 软件进行 χ^2 检验、t 检验、单因素和多因素非条件 logistic 回归模型分析。在进行神经网络训练时, 先对原始数据进行预处理归一化操作, 应用 MATLAB 7.0 软件中自带的 mapminmax 函数将原始数据归一化到 [0.1, 0.9] 之间。以 26 个危险因素作为输入变量, 轻症与重症作为输出变量, 训练样本 163 例, 测试样本 70 例, (即训练样本: 测试样本 = 7:3) 建立 BPNN 模型, 结构为 26 → 8 → 1, 输入层和隐含层间的传输函数采用正切 Sigmoid 函数 - tansig, 隐含层到输出层采用的传输函数为对数 sigmoid 函数 - logsig。网络的训练函数采用函数 trainlm。最大训练步数 10 000 步, 训练误差为 1e - 25, 学习速率为 0.01, 采用 Levenberg - Marquardt 法。网络构建和训练应用 MATLAB 7.0 完成。

2 结果

2.1 一般情况 病例组 137 例重症手足口病患儿中, 男童 84 例 (61.3%), 女童 53 例 (38.7%), 平均月龄为 (24.38 ± 16.73) 个月; 对照组 96 例轻型手足口病患儿中, 男童 57 例 (59.4%), 女童 39 例 (40.6%), 平均月龄为 (23.52 ± 15.19) 个月; 病例组与对照组患儿比较, 2 组患儿性别、年龄差异均无统计学意义 (P > 0.05)。

2.2 手足口病重症化危险因素 logistic 回归分析

2.2.1 单因素非条件 logistic 回归分析 (表 1) 单因素 logistic 回归分析结果显示, 易惊、呼吸节律不

表 1 手足口病重症化危险因素单因素非条件 logistic 回归分析

因素	病例组		对照组		P 值	OR 值	95% CI
	例数	%	例数	%			
易惊	111	81.0	57	59.3	0.000	32.990	15.436 ~ 70.502
呼吸节律不齐	25	18.2	11	11.5	0.037	2.455	1.056 ~ 5.708
手足抖动	24	17.5	8	8.3	0.003	4.028	1.601 ~ 10.132
嗜睡	28	20.4	6	6.2	0.001	12.073	2.802 ~ 52.031
四肢发凉	21	15.3	0	0	0.006	17.198	2.272 ~ 130.210
血糖升高	69	50.3	1	1.0	0.000	6.749	3.303 ~ 12.708
热程 ≥ 3 d	93	67.9	13	13.5	0.000	4.728	2.182 ~ 10.246
热峰 ≥ 39 °C	92	67.2	35	36.4	0.004	2.222	1.300 ~ 3.799
心率 ≥ 140 次/min	57	41.6	19	19.8	0.000	2.871	1.623 ~ 5.081
呼吸 ≥ 30 次/min	15	10.9	3	3.1	0.039	3.811	1.072 ~ 13.533
发热	127	92.7	90	93.8	0.755	0.847	0.297 ~ 2.414
精神差	87	63.5	17	17.7	0.000	22.123	9.509 ~ 51.468
烦躁	23	16.8	5	5.2	0.011	3.672	1.343 ~ 10.037
抽搐	14	10.2	2	2.1	0.029	5.350	1.187 ~ 24.113
呕吐	44	32.1	1	1.0	0.000	44.946	6.067 ~ 332.983
颈强直	42	30.7	1	1.0	0.000	42.000	5.665 ~ 311.411
肢体无力	5	3.6	1	1.0	0.246	3.598	0.414 ~ 31.302
白细胞升高	71	51.8	37	38.5	0.046	1.175	1.009 ~ 2.915
中性细胞比率升高	41	29.9	22	22.9	0.237	1.437	0.788 ~ 2.618
淋巴细胞比率升高	67	48.9	54	56.3	0.270	0.744	0.441 ~ 1.257
C 反应蛋白升高	49	35.8	40	41.7	0.648	1.003	0.989 ~ 1.018
降钙素原升高	128	93.4	92	95.8	0.979	1.004	0.765 ~ 1.316
心电图异常	51	37.2	40	41.7	0.494	0.830	0.487 ~ 1.415
胸部 X 线改变	117	85.4	71	74.0	0.254	1.435	0.771 ~ 2.668

齐、手足抖动、嗜睡、四肢发凉、血糖升高、热程 ≥ 3 d、热峰 ≥ 39 °C、心率 ≥ 140 次/min、呼吸 ≥ 30 次/min、精神差、烦躁、抽搐、呕吐、颈强直、白细胞升高等 16 个因素可能是手足口病重症化的危险因素。

2.2.2 多因素非条件 logistic 回归分析(表 2) 多因素 logistic 回归分析结果显示,颈强直、易惊、呕吐、血糖升高、白细胞升高、热程 ≥ 3 d、手足抖动、精神差是手足口病重症化的危险因素。

表 2 手足口病重症化危险因素多因素非条件 logistic 回归分析

因素	β	S_x	Wald χ^2 值	P 值	OR 值	95% CI
颈强直	5.213	1.392	14.017	0.000	183.633	11.988 ~ 2812.853
易惊	5.068	0.889	8.206	0.000	158.868	27.814 ~ 907.412
呕吐	4.083	1.271	10.321	0.001	59.347	4.915 ~ 716.627
血糖升高	3.141	1.341	5.485	0.019	23.133	1.670 ~ 320.537
白细胞升高	2.505	0.669	14.032	0.000	12.243	3.301 ~ 45.406
热程 ≥ 3 d	2.050	0.916	5.006	0.025	7.765	1.289 ~ 46.762
手足抖动	1.747	0.622	7.891	0.005	5.738	1.698 ~ 19.415
精神差	1.493	0.705	4.493	0.034	4.452	1.119 ~ 17.714
常数项	-3.496	0.568	37.923	0.000		

2.3 BPNN 模型手足口病重症化危险因素分析

应用 MATLAB 7.0 软件建立 BPNN 模型,根据设置的程序训练样本,训练结果将全部保存。本次训练好的网络训练样本对训练数据的分类正确率为 100%,测试样本对测试数据的分类正确率 $> 90\%$ 。BPNN 模型拟合较好。BPNN 模型最终的网络结果为 $26 \rightarrow 8 \rightarrow 1$,影响手足口病重症化的前 10 位危险因素(MIV 绝对值)依次为:易惊(0.691 4)、颈强直(0.537 3)、呕吐(0.453 8)、血糖升高(0.429 9)、手足抖动(0.381 3)、精神差(0.328 3)、热峰 ≥ 39 °C(0.308 6)、白细胞升高(0.290 2)、热程 ≥ 3 d(0.262 1)、嗜睡(0.242 7)。

2.4 因子间交互作用分析 饱和对数线性模型分析结果显示:热峰 ≥ 39 °C 在 BPNN 模型分析和多因素 logistic 回归分析中存在差异,热峰 ≥ 39 °C 与精神差、白细胞升高均具有交互作用($P < 0.05$)。

3 讨论

手足口病是一个复杂多因子疾病,研究其病因直接关系到疾病的治疗和预防^[2,9]。在医学科研实践工作中有很多分析疾病危险因素的方法。传统的多重线性回归、Cox 比例风险回归、logistic 回归分析、对数线性模型等方法均有适用条件和注意事项,如果资料不满足或选择方法失误,则会错误预测信息。非参数检验虽然不依赖总体分布的具体形式,但检验效率偏低^[10]。与传统分析方法比较,BPNN 模型对资料的要求较为简单,其适用于任何分布形式的变量,能包含变量间完整的信息,无限逼近任意非线性函数,容错性强。因而在对疾病的危险因素进行分析时,BPNN 模型有着 logistic 回归分析方法所不能比拟的优势^[11-13]。

本研究首次将 BPNN 模型用于手足口病重症化过程危险因素预测,并与传统的 logistic 回归分析结果比较,探讨影响手足口病重症化过程的危险因素。结果显示:经 2 种方法筛选的危险因素基本一致。但 BPNN 模型纳入了在 logistic 回归分析中予以剔除的变量“热峰 ≥ 39 °C”,可能与该变量与其他因素存在交互作用有关。采用饱和对数线性模型分析结果显示,热峰 ≥ 39 °C 与精神差、白细胞升高均存在交互作用,提示热峰 ≥ 39 °C 是一个重要的协同变量。主要原因可能是热峰 ≥ 39 °C 时,重症患儿都有不同程度的精神萎靡不振,出现手足抖动、呕吐、嗜睡等中枢神经系统症状,机体代谢增强,出现一系列炎症变化,此解释与《手足口病诊疗指南(2010 版)》^[8]一致。

综上所述,BPNN 模型同样能够完成手足口病重症化危险因素的预测研究,而且优于传统统计模型处理方法。BPNN 模型分析不受制于变量的具体形式,能够较好的处理因素间复杂的关系,为流行病学病因分析提供了一个得心应手的新工具。此外,鉴于 logistic 回归分析有着明确的变量准入和剔除标准,BPNN 模型是否也可以实现此功能尚有待进一步的研究探讨。

参考文献

- [1] Mou J, Dawes M, Li Y, et al. Severe hand, foot and mouth disease in Shenzhen, South China; what matters most? [J]. *Epidemiol Infect*, 2014, 142(4): 776-788.
- [2] 袁国平, 郭祖鹏, 杨兴堂, 等. 手足口病发病影响因素病例对照研究[J]. *中国公共卫生*, 2011, 27(11): 1407-1409.
- [3] 姚文清, 陈静乙, 李鑫, 等. 辽宁省 2009 年手足口病流行病学分析[J]. *中国公共卫生*, 2010, 26(12): 1589-1590.
- [4] Dey P, Lamba A, Kumari S, et al. Application of an artificial neural network in the prognosis of chronic myeloid leukemia [J]. *Anal Quant Cytol Histol*, 2011, 33(6): 335-339.
- [5] Saylam B, Keskek M, Ocak S, et al. Artificial neural network

- analysis for evaluating cancer risk in multinodular goiter [J]. J Res Med Sci, 2013, 18(7): 554-557.
- [6] 兰莉, 高菡璐, 乔冬菊, 等. 应用 BP 神经网络模型预测急性心肌梗死率与气象因素的关系 [J]. 中国公共卫生管理, 2012, 28(5): 642-644.
- [7] 卫生部. 手足口病预防控制指南 (2009 版) [J]. 全科医学临床与教育, 2010, 8(2): 125-133.
- [8] 卫生部. 手足口病诊疗指南 (2010 年版) [J]. 柳州医学, 2012, 25(2): 140-143.
- [9] 李亮, 许可, 祁贤, 等. 儿童手足口病影响因素病例对照研究 [J]. 中国公共卫生, 2011, 27(1): 16-18.
- [10] 颜虹. 医学统计学 [M]. 北京: 人民卫生出版社, 2010: 56-85.
- [11] Chiu HC, Ho TW, Lee KT, et al. Mortality predicted accuracy for hepatocellular carcinoma patients with hepatic resection using artificial neural network [J]. Scientific World Journal, 2013, 4: 66-76.
- [12] Andersson B, Andersson R, Ohlsson M, et al. Prediction of severe acute pancreatitis at admission to hospital using artificial neural networks [J]. Pancreatology, 2011, 11(3): 328-335.
- [13] 王伟, 李霓, 许伟, 等. 应用 BP 神经网络分析子宫肌瘤危险因素 [C]. 全国肿瘤流行病学和肿瘤病因学学术会议论文集, 大连, 2007. 天津: 中国抗癌学会, 2007: 60-62.

收稿日期: 2013-12-30

(郭薇编辑 刘铁校对)

· 调查报告与分析 ·

宁夏窖水有机氯农药残留现状与健康风险评价*

李丰宝, 李玲, 田雨, 张鹏举

摘要:目的 了解宁夏南部山区饮用窖水中有机氯农药的残留状况及对人体健康产生的潜在危害。方法 于 2012 年 9 月初, 在宁夏南部仍以窖水作为饮用水的山区设置 12 个采样点, 每个采样点采集窖水 6 L, 测定其中六六六类 (HCHs) 和滴滴涕类 (DDTs) 各异构体残留量; 评价饮用窖水中 HCHs、DDTs 所引起的健康风险。结果 宁夏南部山区饮用窖水中检出 HCHs 和 DDTs, HCHs 含量范围在 0~5.20 ng/L, 主要残留物为 γ -HCHs, DDTs 含量范围在 0~73.50 ng/L, 只检测到 p,p'-DDE, DDTs 残留量高于 HCHs; HCHs、DDTs 所引起健康危害的个人年风险分别为 $1.74 \times 10^{-12}/a$ 、 $3.96 \times 10^{-11}/a$; 2 种有机氯农药对人体健康危害的年风险为 $4.14 \times 10^{-11}/a$ 。结论 宁夏南部山区饮用窖水中仍有 HCHs 和 DDTs 残留, 但低于地表水质量标准限值, 对人体潜在危害很小。

关键词: 六六六; 滴滴涕; 健康风险评估; 窖水

中图分类号: R 169.1 文献标志码: A 文章编号: 1001-0580(2014)06-0761-03 DOI: 10.11847/zgggws2014-30-06-21

Residues of organochlorine pesticides in cellar water and their health risk in south mountainous areas of Ningxia

LI Feng-bao, LI Ling, TIAN Yu, et al (Department of Environmental and Occupational Health, Public Health School, Ningxia Medical University, Key Laboratory of Reproduction and Genetics of Ningxia, Key Laboratory of Fertility Preservation and Maintenance of Ministry of Education, Yinchuan, Ningxia Hui Autonomous Region 750004, China)

Abstract: Objective To examine residues of organochlorine pesticides in cellar water and their potential health hazards of in south mountainous areas of Ningxia. **Methods** In September 2012, the samples of cellar water were collected at 12 points in the south mountainous areas of Ningxia. Six L cellar water was collected at each site. The residual levels of hexachloro-cyclohexane (HCH), dichloro-diphenyl-trichloroethane (DDT) and their isomers were determined. Health risks of HCHs and DDTs in drinking cellar water were assessed. **Results** The detection rates of HCH and DDT in the drinking cellar water samples was 100%. The HCHs content ranged from not detectable to 5.20 ng/L, and the main residue was γ -HCHs. The DDTs content ranged from not detectable to 73.50 ng/L, and the only residue detected was p,p'-DDT. The DDTs' content was higher than HCHs. The health risk of HCHs and DDTs of the residues in cellar water for an individual were 1.74×10^{-12} and 3.96×10^{-11} , respectively. The total health risk caused by the organochlorine pesticides was 4.14×10^{-11} per year. **Conclusion** The residues of organochlorine pesticide in cellar drinking water in the south mountainous areas of Ningxia were detected but the contents are lower than the surface water quality standard limits and the potential hazards to human health is weak.

Key words: HCHs; DDTs; health risk assessment; cellar water

有机氯农药是氯代烃类化合物, 易溶于脂肪及大多数有机溶剂, 挥发性小, 化学性质稳定, 通过食

物链和生物富集放大作用, 可对人体产生潜在的危害^[1-2]。六六六 (hexachloro-cyclohexane, HCHs) 和

* 基金项目: 宁夏医科大学“博士学位建设学科”开放课题 (2001140102)

作者单位: 宁夏医科大学公共卫生学院劳动卫生与环境卫生学系 宁夏回族自治区生殖与遗传重点实验室 生育力保持教育部重点实验室, 宁夏银川 750004

作者简介: 李丰宝 (1987-), 男, 山东人, 硕士在读, 研究方向: 环境有害因素与健康。

通讯作者: 李玲, E-mail: 13995307525@163.com

数字出版日期: 2014-4-16 17:03

数字出版网址: <http://www.cnki.net/kcms/detail/21.1234.R.20140416.1703.031.html>