

内分泌腺疾病及代谢病临床研究

DOI:10.13406/j.cnki.cyx.002103

糖尿病足患者预后预测相关模型研究

陈 静¹,程庆丰²,陈 悦³,易 静⁴

(1. 重庆市璧山区疾病预防控制中心,重庆 402760;2. 重庆医科大学附属第一医院内分泌科,重庆 400016;

3. 重庆市人民医院胸外科,重庆 400010;4. 重庆医科大学公共卫生与信息管理学院

卫生统计与信息管理教研室,重庆 400016)

【摘要】目的:比较 BP 神经网络模型(back propagation,BP)、遗传算法优化的 BP 神经网络模型(BP neural network model trained with genetic algorithm,GA-BP)、COX 比例风险模型在糖尿病足患者截肢及生存预后的预测效果,选择最优预测模型。**方法:**收集重庆医科大学附属第一医院 2014 年 1 月至 2016 年 1 月收治的 273 例糖尿病足患者住院资料,采用电话随访的形式对患者随访至 2016 年 12 月。建立 BP 神经网络模型、遗传算法优化的 BP 神经网络模型、COX 比例风险模型,以 ROC 曲线下面积大小,灵敏度、特异度等值判断 3 种模型在糖尿病足患者截肢及生存预后中的预测效果。**结果:**结局分别为截肢及死亡时,BP 神经网络模型(截肢: $\chi^2=7.692, P=0.005$;死亡: $\chi^2=12.071, P=0.000$)和遗传算法优化的 BP 神经网络模型(截肢: $\chi^2=10.083, P=0.001$;死亡: $\chi^2=12.071, P=0.000$)的预测性能均优于 COX 比例风险模型,而 BP 神经网络模型和遗传算法优化的 BP 神经网络模型的预测结果间差异均无统计学意义(截肢: $\chi^2=0.200, P=1.000$;死亡: $\chi^2=0.000, P=1.000$)。**结论:**BP 神经网络模型及遗传算法优化的 BP 神经网络模型均可应用于糖尿病足等慢性疾病的生存预后分析。

【关键词】糖尿病足;遗传算法;预测模型**【中图分类号】**R195.1**【文献标志码】**A**【收稿日期】**2018-12-10

A study of prognostic prediction model for diabetic foot patients

Chen Jing¹, Cheng Qingfeng², Chen Yue³, Yi Jing⁴

(1. Center for Disease Control and Prevention in Bishan District; 2. Department of Endocrinology,

The First Affiliated Hospital, Chongqing Medical University; 3. Department of Thoracic Surgery of Chongqing People's Hospital; 4. Department of Medical Statistics, College of Public Health, Chongqing Medical University)

【Abstract】Objective: To compare the results of the back propagation (BP) neural network model, BP neural network model trained with genetic algorithm (GA-BP), and Cox proportional hazards model in the prediction of amputation and survival prognosis of diabetic foot (DF) patients, and to choose the optimal prediction model. **Methods:** Hospitalization data of 273 patients with DF were collected who were admitted to The First Affiliated Hospital of Chongqing Medical University from January 2014 to January 2016. These patients were also followed up by telephone until December 2016. Then three models were established, namely the BP neural network model, BP neural network model trained with genetic algorithm, and Cox proportional hazards model. The area under the receiver operating characteristic curve, sensitivity, and specificity were used to judge the performance of the three models in the prediction of amputation and survival prognosis of DF patients. **Results:** When the outcomes were amputation and death, the BP neural network model (amputation: $\chi^2=7.692, P=0.005$; death: $\chi^2=12.071, P=0.000$) and BP neural network model trained with genetic algorithm (amputation: $\chi^2=10.083, P=0.001$; death: $\chi^2=12.071, P=0.000$) were better than the Cox proportional hazards model. However, there was no significant difference between the BP neural network model and the BP neural network model trained with genetic algorithm (amputation: $\chi^2=0.200, P=1.000$; death: $\chi^2=0.000, P=1.000$). **Conclusion:** Both the BP neural network model and the BP neural network model trained with genetic algorithm can be applied to the analysis of survival and prognosis of chronic diseases such as DF.

【Key words】diabetic foot; genetic algorithm; prediction model

作者介绍:陈 静, Email: 278794565@qq.com,

研究方向:流行病学与卫生统计。

通信作者:易 静, Email: yijinga@sina.com。

基金项目:重庆市科委基金资助项目(编号:cstc2015jcyj10092)。

优先出版: <http://kns.cnki.net/kcms/detail/50.1046.R.20190505.1724.008.html>

(2019-05-06)

糖尿病足(diabetic foot, DF)作为糖尿病最常见的并发症之一,具有高发病率、高截肢率、高死亡率的特点。据报道,糖尿病患者一生发生 DF 的风险高达 25%,每年因 DF 截肢的患者占非外伤截肢患者的 50%以上,5 年死亡率高达 43%~55%^[1]。DF 预后非常重要。目前用于构建 DF 等相关慢性病的风险预测主要是以统计学为基础的传统风险预测模型,如多元线性回归、多元 COX 回归和多元 logistic 回归等,但这些统计方法均有其各自的限制^[2-4]。在 DF 领域的生存分析中,BP 神经网络模型及遗传算法优化的 BP 神经网络模型均运用较少,故运用这 2 种模型与传统预测模型进行 DF 患者截肢及生存预后状况分析比较,以期为医师的临床决策提供科学依据。

1 资料与方法

1.1 研究对象

收集重庆医科大学附属第一医院 2014 年 1 月至 2016 年 1 月收治的 273 例 DF 患者人口学资料、临床资料、并发症、生化指标等。周围动脉疾病指由于动脉狭窄或闭塞导致肢体供血逐渐减少的疾病。其诊断标准除具有相应的下肢缺血症状和体格检查外,彩超、MRA、CTA 和 DSA 等影像学检查可用于诊断,其中以 DSA 为诊断金标准。冠心病为有明确冠状动脉影像学证据的患者,且符合 2007 年版《中国成人血脂异常防治指南》中有关冠心病的相关诊断标准。根据美国心脏病协会和美国疾病控制与预防中心建议,可以按照患者心血管病危险程度对超敏 C 反应蛋白的分类进行界定: $<1\text{ mg/L}$ 、 $1.0\sim 3.0\text{ mg/L}$ 和 $>3.0\text{ mg/L}$ 分别定义为相对低危险、中度危险及高度危险。足背动脉搏动分为正常、较弱、消失 3 组;糖尿病足溃疡分级采用最常用的 DF 感染 Wagner 法分类,将较轻微 0、1、2 级为第一组,较严重的 3、4、5 级为第二组。DF 诊断符合 2004 年美国感染协会推荐诊断标准。

1.2 变量筛选

本研究共收集人口学资料、临床资料、并发症、生化指标等变量 90 个,其中生化指标 69 个,遂采用聚类分析将 69 个生化指标聚类为肝功能指标、营养指标、炎症指标、凝血指标、肾功能指标 5 大类,每类选取 1~2 个欧氏距离 <2 及临床专业认为有意义的指标,共 12 个变量,即血红蛋白、碱性磷酸酶、血小板分布宽度、HbA1c(hemoglobin A1c, HbA1c)、肌酐、尿素、超敏 C 反应蛋白、白细胞总数、血小板、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、载脂蛋白 A1。最后进入研究的人口学资料、临床资料、并发症、生化指标等变量 33 个。

1.3 结局判定

将患者结局分为 2 类。第一类为 3 年截肢情况,包含小截肢与大截肢。小截肢是指下肢踝关节以下的部分被截肢;大截肢是指下肢踝关节以上的部分被截除。第二类为 3

年死亡情况,指与 DF 有关的死亡,无论有无截肢。本研究以入院时间作为起点,发生结局事件为终点,通过与患者或家属进行电话随访至 2016 年 12 月,获得患者的生存情况及结局事件发生时间。

1.4 研究方法

DF 患者资料被随机分为训练集和测试集,其比例为 4:1。训练集的数据用于构建模型,测试集的数据用于构建模型间的预测性能比较。采用独立样本 t 检验及卡方检验比较测试集与训练集间的差异。COX 比例风险模型以连续性变量预后指数作为预测的结果变量,并以训练集的最高预测一致率为预后指数分类的界值的选择标准,定义测试集中预后指数大于其界值为 1(结局事件发生),小于其界值为 0(结局事件未发生)。BP 神经网络模型和遗传算法优化的 BP 神经网络模型基于 Softmax 策略以二分类变量生存 0(结局事件未发生)和 1(结局事件发生)作为结果变量。以 AUC、灵敏度特异度对模型进行评价,配对卡方检验比较 3 个模型间的预测结果。以上计算均采用 SPSS 21.0、matlab R2015b 完成。

2 结果

2.1 一般情况

273 名 DF 患者中,男性所占比例为 66%,女性所占比例为 34%,平均年龄为 65 岁。观察期间,有 37 人进行了截肢,其中 83.8%是小截肢,16.2%是大截肢。一年生存率是 88.3%,一年截肢率是 13.4%。3 年生存率是 79.2%,3 年截肢率是 13.9%。根据研究目的,273 名患者被随机分为训练集(218, 75%)和测试集(55, 25%),训练集和测试集的数据差异无统计学意义(表 1)。

2.2 COX 比例风险模型

结局为截肢时,单因素 COX 分析结果显示坏疽严重程度、糖化血红蛋白值、血小板、超敏 C 反应蛋白、低密度脂蛋白胆固醇等 11 个变量有统计学意义($P<0.05$)。采用多因素 COX 模型分析,最后进入主方程的 3 个变量为坏疽严重程度、糖化血红蛋白、低密度脂蛋白胆固醇($P<0.05$, 表 2)。根据多因素分析结果建立预后函数,预后指数界值为 6.39 时,训练集最大预测一致率为 85.32%。结局为死亡时,单因素分析有统计学意义的 11 个变量,如坏疽严重程度、糖化血红蛋白值等带入多因素 COX 模型,最后进入主方程的 3 个变量为脑血管疾病、溃疡严重程度、周围动脉疾病($P<0.05$, 表 2)。预后指数界值为 5.78,训练集最大预测一致率为 77.98%。

2.3 BP 神经网络模型及遗传算法优化的 BP 神经网络模型

BP 神经网络模型及遗传算法优化的 BP 神经网络模型根据多因素 COX 模型分析有意义的变量为输入神经元,模型输入层数据结构皆为 3×219 ,即 219 个观测的 3 个变量数据。以凑试法^[5],即最大 AUC 来确定最佳隐含层节点数,建立 BP 神经网络模型及遗传算法优化的 BP 神经网络模型。并得出隐含层节点数分别为 4 和 2 时,结局为截肢及死亡的模型 AUC 最大。隐含层节点数改变时,模型 AUC 减小。以结

表 1 训练集与测试集中不同指标数据对比

变量	测试集	训练集	χ^2 值	P 值
高密度脂蛋白胆固醇	1.520 ± 0.439	1.481 ± 0.373	0.603	0.547
低密度脂蛋白胆固醇	2.180 ± 1.039	2.585 ± 1.994	-1.320	0.188
碱性磷酸酶	114.130 ± 63.395	98.010 ± 55.179	1.859	0.064
载脂蛋白 A1	1.081 ± 0.352	1.155 ± 0.739	-0.648	0.518
血小板	250.840 ± 158.084	241.330 ± 96.876	0.563	0.574
血红蛋白	109.220 ± 21.474	115.240 ± 21.089	-1.884	0.061
白细胞总数	9.325 ± 4.926	9.269 ± 4.361	0.082	0.935
HbA1c	10.348 ± 2.535	9.379 ± 2.462	2.292	0.025
糖尿病病程	10.332 ± 7.479	8.721 ± 6.962	1.510	0.132
足背动脉搏动(消失)(n,%)	13(28.3)	48(24.1)	0.485	0.785
超敏 C 反应蛋白(>3.0 mg/L)(n,%)	25(55.6)	91(48.4)	0.956	0.620
周围动脉疾病(n,%)	9(24.3)	58(38.2)	2.489	0.129
脑血管疾病(n,%)	9(16.4)	36(16.5)	0.001	0.979
冠心病(n,%)	8(14.5)	44(20.2)	0.905	0.443
坏疽(n,%)	17(30.9)	64(29.4)	0.051	0.869
溃疡严重程度(n,%)	22(40.0)	106(48.6)	1.312	0.291
一年截肢率(%)	10.9	13.5	0.374	0.541
一年生存率(%)	81.8	89.0	0.369	0.544

表 2 多因素 COX 回归分析结果

结局		B	SE	P	HR	95%CI
截肢	低密度脂蛋白胆固醇	0.127	0.061	0.037	1.135	1.007~1.279
	溃疡严重程度	1.551	0.508	0.002	4.715	1.742~12.763
	HbA1c	0.239	0.085	0.005	1.270	1.076~1.499
死亡	溃疡严重程度	1.352	0.652	0.038	3.866	1.076~13.887
	脑血管疾病	1.110	0.506	0.028	3.034	1.125~8.183
	周围动脉疾病	0.983	0.558	0.078	2.673	0.895~7.986

局变量的发生与否作为输出变量。最终建立的 BP 神经网络模型及遗传算法优化的 BP 神经网络模型如图 1、图 2 所示。

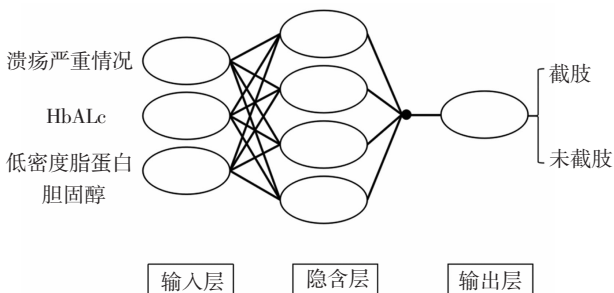


图 1 结局为截肢的 BP 神经网络模型结构

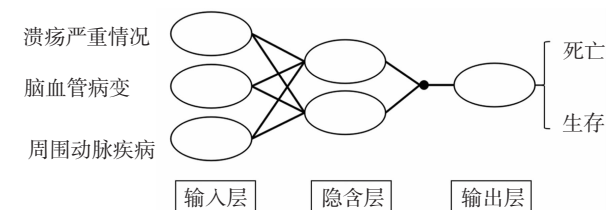


图 2 结局为死亡的 BP 神经网络模型结构

2.4 3 个模型的对比

将 BP 神经网络模型及遗传算法优化的 BP 神经网络模型、COX 比例风险模型的预测结果进行比较。结局为截肢时, BP 神经网络模型预测结果的 AUC 大于 COX 比例风险模型 (0.898 vs. 0.543, $\chi^2=7.692$, $P=0.005$, 表 3、表 4, 图 3)。遗传算法优化的 BP 神经网络模型预测结果的 AUC 也大于 COX 比例风险模型 (0.888 vs. 0.543, $\chi^2=10.083$, $P=0.001$, 表 3、表 4, 图 3)。而 BP 神经网络模型和遗传算法优化的 BP 神经网络模型的预测结果差异没有统计学意义 ($\chi^2=0.200$, $P=1.000$, 表 4)。通过表 3 可知, BP 神经网络模型及遗传算法优化的 BP 神经网络模型预测结果的灵敏度和特异度也大于 COX 比例风险模型。结局为死亡时, BP 神经网络模型和遗传算法优化的 BP 神经网络模型预测结果几乎完全相同 (0.705 vs. 0.705, $\chi^2=0.000$, $P=1.000$, 表 3、表 4, 图 4), 均大于 COX 比例风险模型 (0.705 vs. 0.580, $\chi^2=12.071$, $P=0.000$, 表 3、表 4, 图 4)。

表 3 3 种模型预测性能比较

结局	模型	曲线下面积	95%CI	灵敏度 (%)	特异度 (%)	P 值
截肢	BP	0.898	0.813~0.982	100.00	79.60	0.002
	GA-BP	0.888	0.798~0.977	100.00	77.60	0.002
	COX	0.543	0.285~0.800	16.70	91.80	0.736
死亡	BP	0.705	0.531~0.878	72.70	68.20	0.037
	GA-BP	0.705	0.531~0.878	72.70	68.20	0.037
	COX	0.580	0.379~0.780	27.30	88.60	0.419

表 4 3 种模型预测结果两两比较

对比模型	截肢		死亡	
	χ^2 值	P 值	χ^2 值	P 值
BP & GA-BP	0.200	1.000	0.000	1.000
COX & GA-BP	10.083	0.001	12.071	<0.001
COX & BP	7.692	0.005	12.071	<0.001

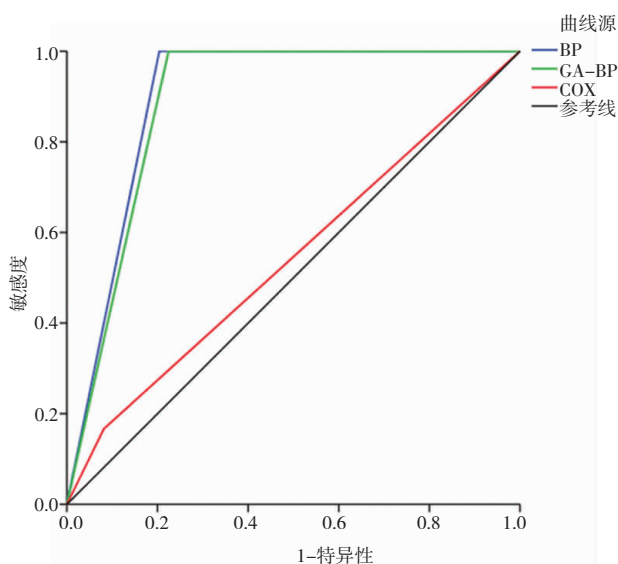


图 3 结局为截肢的 3 种模型 ROC 曲线

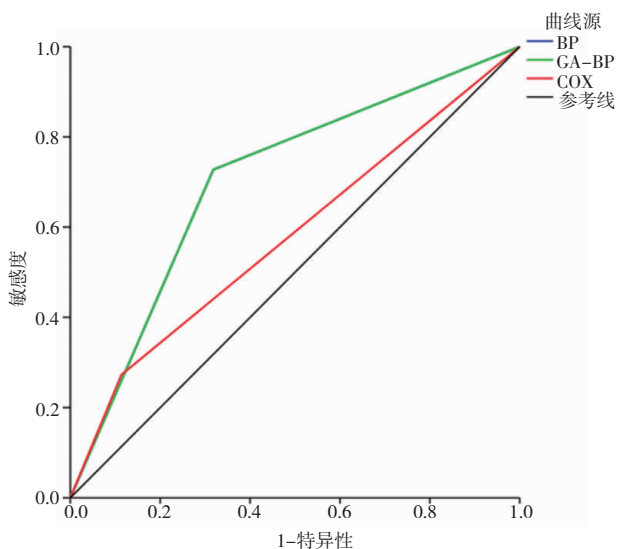


图 4 结局为死亡的 3 种模型 ROC 曲线

3 讨论

在结局为 DF 患者截肢的预测模型中,以溃疡严重情况、HbA1c、低密度脂蛋白胆固醇为输入变量建立 BP 神经网络模型等预测模型,取得良好结果。大量研究表明,溃疡情况越严重,DF 患者截肢率就越高^[6-7]。而血液中低密度脂蛋白胆固醇的升高会导致其在血管内膜中的积聚,从而导致血管痉挛、动脉粥样硬化等至血管阻塞,最终引起血运障碍,这可能是导致糖尿病足患者截肢的重要原因^[8]。HbA1c 是 DF 疾病发展的重要检测指标,与相关研究结果一致^[9]。在结局为 DF 患者死亡的预测模型中,预测变量为溃疡严重情况、脑血管病变和周围动脉疾病。在以往国内外相关研究中,大量学者提出与本文相似的结论。Brennan 等^[10]认为糖尿病足溃疡严重情况与死亡风险增加有关;Rubio 等^[11]则提出脑血管病变也是引起 DF 患者死亡的一个重要影响因素;Morbach 等^[12]和莫泽伟等^[13]的研究证实周围动脉阻塞性疾病对 DF 患者的死亡具有预测意义。

BP 神经网络模型在目前医学领域的应用日益广泛,但大多应用于癌症患者的预后评估,应用于生存分析领域较少,在慢性病领域更是鲜有报道^[14-16]。遗传算法优化的 BP 神经网络在医学方面的应用相对较少。近年来,Aalaei 等^[17]、Jeya 等^[18]、Hou 等^[19]将其应用于癌症的分类诊断,Fei 等^[20]则应用于重症急性胰腺炎并发急性肺损伤的预测,但关于慢性病的预后预测应用还有待研究。

以往国内外多选取 logistic、COX 等传统统计学模型进行 DF 预后的研究讨论^[21-23]。本文将人工智能模型 BP 神经网络及遗传算法优化的 BP 神经网络模型应用于 DF 患者的预后预测,并取得较好的结果,为 DF 等慢性疾病的预后预测提供了新思路。

本研究建立的 3-4-1 和(或)3-2-1 结构的 BP 神经网络模型较好地完成了个体预后的预测工作。神经网络属于非线性模拟系统^[24],比传统的统计模型更能够合理地描述危险因素与疾病之间的关系。由于其经过适当修改,并拥有强大的容错性,使得它在本研究中表现出良好的预测效果。BP 神经网络模型明显优于 COX 比例风险模型,表明它可以很好地用于慢性病的生存分析。

本研究应用遗传算法优化的 BP 神经网络模型

进行个体预后的预测,并在研究中表现出较好的拟合效果。其优点在于即使当适应度函数是不规则的、不连续的或者有噪声时,也可以最大可能地寻找到全局状态下的最优解,而不至于陷入局部的最优^[6]。但遗传算法优化的 BP 神经网络模型与 BP 神经网络模型的预测结果之间并无统计学差异。这可能与遗传算法容易出现过早收敛且本文样本量较少、慢性病结局变量也较少等因素有关。

本研究发现 COX 比例风险模型的预测效果并不理想。因其需要结局的相关变量按照一定方式对生存时间产生影响,大多数资料不能达到理想的效果是因为它们仅可以近似地符合此类型假设^[25]。也可能与预后指数分类界值点的选取有关,需在今后的研究中尝试多种分类界值选取标准,以作进一步判定。

综上所述,对于糖尿病足患者截肢及死亡的预测,BP 神经网络模型及遗传算法优化的 BP 神经网络模型均能够胜任,并且它们都优于 COX 比例风险模型。

参 考 文 献

- [1] 李翔,肖婷,王玉珍,等. 139 例糖尿病足溃疡患者的死亡率及伴有并发症分析[J]. 中华内分泌代谢杂志,2011,27(2):128-132.
- [2] 李文琦,黄水平,李海朋. BP 神经网络和 Cox 比例风险模型在生存分析应用中的比较[J]. 郑州大学学报(医学版),2014,49(6):822-825.
- [3] Khosravi B,Pourahmad S,Bahreini A,et al. Five years survival of patients after liver transplantation and its effective factors by neural network and cox poroportional hazard regression models[J]. Hepat Mon, 2015,15(9):1-7.
- [4] Ho WH,Lee KT,Chen HY,et al. Disease-free survival after hepatic resection in hepatocellular carcinoma patients:a prediction approach using artificial neural network[J]. PLoS One,2012,7(1):e29179.
- [5] 雷英杰. Matlab 遗传算法工具箱及应用[M]. 西安:西安电子科技大学出版社,2014.
- [6] Musa IR,Ahmed MON,Sabir EI,et al. Factors associated with amputation among patients with diabetic foot ulcers in a Saudi population[J]. BMC Research Notes,2018,11(1):260.
- [7] 胡慧萍,陈明卫. 糖尿病足溃疡愈合及截肢的相关因素分析[J]. 安徽医科大学学报,2016,51(11):1634-1637.
- [8] Yang Q,Wang JH,Huang DD,et al. Clinical significance of analysis of the level of blood fat,CRP and hemorheological indicators in the diagnosis of elder coronary heart disease[J]. Saudi J Biol Sci,2018,25(8):1812-1816.
- [9] 杨殷,杨彩哲. 糖尿病足发生截肢的相关危险因素的 meta 分析[J]. 山西医科大学学报,2016,47(7):618-623.
- [10] Brennan MB,Hess TM,Bartle B,et al. Diabetic foot ulcer severity predicts mortality among veterans with type 2 diabetes[J]. J Diabetes Complications,2017,31(3):556-561.
- [11] Rubio JA,Jiménez S,Ivarez J,et al. Clinical characteristics and mortality in patients treated in a multidisciplinary diabetic foot unit[J]. Endocrinol Diabetes Nutr,2017,64(5):241-249.
- [12] Morbach S,Furchert H,Ute Gröblichhoff U,et al. Long-term prognosis of diabetic foot patients and their limbs:amputation and death over the course of a decade[J]. Diabetes Care,2012,35(10):2021-2027.
- [13] 莫泽伟,陈道雄,全会标,等. 初发糖尿病足溃疡患者的生存预后研究[J]. 中国现代医学杂志,2018,28(28):77-84.
- [14] Ansari D,Nilsson J,Andersson R,et al. Artificial neural networks predict survival from pancreatic cancer after radical surgery[J]. Am J Surg,2013,205(1):1-7.
- [15] Amiri Z,Mohammad K,Mahmoudi M,et al. Assessing the effect of quantitative and qualitative predictors on gastric cancer individuals survival using hierarchical artificial neural network models[J]. Iran Red Cresc Med J,2013,15(1):42-48.
- [16] Qiao G,Li J,Huang A,et al. Artificial neural networking model for the prediction of post-hepatectomy survival of patients with early hepatocellular carcinoma[J]. J Gastroenterol Hepatol,2014,29(12):2014-2020.
- [17] Aalaei S,Shahraki H,Rowhanimanesh A,et al. Feature selection using genetic algorithm for breast cancer diagnosis:experiment on three different datasets[J]. Iran J Basic Med Sci,2016,19(5):476-482.
- [18] Jeya IJS,Deepa SN. Lung cancer classification employing proposed real coded genetic algorithm based radial basis function neural network classifier[J]. Comput Math Methods Med,2016,2016:7493535.
- [19] Hou Q,Bing ZT,Hu C,et al. Rankprod combined with genetic algorithm optimized artificial neural network establishes a diagnostic and prognostic prediction model that revealed C1QTNF3 as a biomarker for prostate cancer[J]. Ebiomedicine,2018,32:234-244.
- [20] Fei Y,Gao K,Li WQ. Artificial neural network algorithm model as powerful tool to predict acute lung injury following to severe acute pancreatitis[J]. Pancreatolgy,2018,18(8):892-899.
- [21] Won SH,Chung CY,Park MS,et al. Risk factors associated with amputation-free survival[J]. Yonsei Med J,2014,55(5):1373-1378.
- [22] 王丹丹,林海燕,戴文珍,等. 糖尿病足溃疡患者预后特点及危险因素分析[J]. 中国慢性病预防与控制,2017,25(11):866-868.
- [23] 陈静,程庆丰,陈悦,等. 糖尿病足患者截肢及生存预后影响因素分析[J]. 中国糖尿病杂志,2018,26(2):123-127.
- [24] Huang SH,Loh JK,Tsai JT,et al. Predictive model for 5-year mortality after breast cancer surgery in Taiwan residents[J]. Chin J Cancer, 2017,36(4):23.
- [25] 朱永明,邵庚云. 基于 Cox 生存模型的上市公司财务困境预测研究[J]. 商业研究,2013,56(11):108-116.

(责任编辑:张辉洁)