

3 种统计模型在预测肺癌术后并发症中的比较

宋 健¹ 苏 虹¹ 周洋洋¹ 郭亮亮¹ 王保龙²

摘要 目的 探讨 BP 神经网络模型在预测肺癌术后并发症中的应用价值。方法 调查肺癌患者术后并发症发生情况。分别应用 Logistic 回归、BP 神经网络模型和经 Logistic 回归筛选变量后的 BP 神经网络模型 3 种办法建立预测模型,并比较 3 种模型的预测准确度。结果 Logistic 回归、BP 神经网络模型和经 Logistic 回归筛选变量后的 BP 神经网络模型的预测一致率分别为 81.6%、89.7%、90.8%。3 种模型受试者工作特征曲线(ROC 曲线)下面积(AUC)分别为 0.636、0.801、0.808。Logistic 模型的 AUC 与两种 BP 神经网络模型的差异有统计学意义($P < 0.05$)。结论 BP 神经网络对肺癌术后并发症预测的效果优于 Logistic 回归模型。

关键词 Logistic 模型; BP 神经网络; 肺癌; 并发症

中图分类号 R 195.1; R 734.2; R 619.9

文献标志码 A 文章编号 1000-1492(2014)04-0472-04

肺癌是严重威胁人类健康的主要恶性肿瘤之一,其发病率和病死率在全球范围内呈逐年上升趋势^[1]。手术治疗仍是现阶段治疗肺癌患者的主要手段,但肺癌术后并发症的发生却影响了患者的康复。因此有必要准确地进行术后并发症的预测,提供合适的治疗方案。以往多采用 Logistic 回归进行术后并发症危险因素的研究,较少对其预测的准确性进行评价。人工神经网络是一项模仿生物神经系统中神经元的一种新数学处理方法。其中,应用最广泛的是采用误差反向传递学习方法(error back propagation learning methods)的多层感知机神经网络,简称 BP 神经网络,具有很强的非线性映射能力和柔性的网络结构^[2]。该研究尝试探讨 BP 神经网络在预测肺癌术后并发症中的应用价值,并与传统的 Logistic 回归方程相比较,为寻求一种更加准确的预测方法提供参考依据。

2013-10-16 接收

基金项目:国家自然科学基金(编号:81172172)

作者单位:¹安徽医科大学公共卫生学院流行病与卫生统计学系,合肥 230032

²安徽医科大学附属省立医院检验科,合肥 230001

作者简介:宋 健,男,硕士研究生;

苏 虹,女,教授,硕士生导师,责任作者,E-mail: suhong5151@sina.com

1 材料与方法

1.1 一般资料 回顾性调查安徽医科大学附属省立医院 2009 年~2012 年的 349 例肺癌手术病例。利用病案数据库查找与研究相关的患者资料。所有数据资料经过核实校对后,运用 Epidata 软件建立数据库。

1.2 Logistic 回归 采用 SPSS 17.0 进行多因素分析。模型采用最大似然估计前进法进行逐步回归分析,入选变量的标准是 $P < 0.05$,剔除变量的标准是 $P > 0.10$ 。通过训练数据建立多因素回归预测模型,并带入测试数据进行预测。概率的判别标准为 0.5,即 $P > 0.5$ 时预测结果为发生并发症,否则为不发生,统计学显著性水平规定为 $\alpha = 0.05$ 。

1.3 BP 神经网络 BP 神经网络由输入层、若干隐藏层和输出层组成。每一层有若干神经元连接,同层神经元之间无连接。其基本原理^[3]是:输入信号 X_i 通过中间节点作用于输出节点,经过非线性变化,产生输出信号 Y_k ,网络输出值 Y 与期望输出值 T 之间的偏差,通过反向调整输入节点与隐藏层、输出层与隐藏层之间的权值系数及阈值,使误差沿梯度方向下降,经过反复学习训练,确定合适的连接权值与阈值,以使输出值 Y 与期望输出值 T 之间的误差最小。经过训练的神经网络能对类似样本信息的输入自行处理并进行合理的预测。

本研究采用 SPSS 17.0 统计软件中的神经网络模块的多层感知器(multilayer perceiver, MLP)。MLP 是基于 BP 神经网络的一种算法。本研究将数据随机分为训练样本、预测样本和校验样本。分别占总样本的 1/2、1/4 和 1/4。训练样本用于网络的训练,产生合适的网络参数。校验样本在训练过程中起监督作用,防止过度拟合。预测样本为最终评价神经网络的独立数据集,它不用于构建模型。输入层和输出层节点数分别有输入变量和结局变量决定。首先输入层变量为所有研究变量,考虑到输入层神经元个数过多对样本量的要求较高,笔者尝试将 Logistic 回归与 BP 神经网络相结合,输入变量为 Logistic 多因素分析结果中有意义的变量。结局变

量均为有无并发症发生。研究证实对于任何闭区间内的一个连续函数都可以用一个隐层的 BP 网络来逼近。对于隐藏层内神经元,目前尚无公认的计算公式。选择“自动体系结构选择”,隐含层中最小单位数为 1,最大单位数为 50,训练类型为“批处理”,优化算法选用“调整的共轭梯度”,初始 Lambda 值为 0.000 000 5,初始 Sigma 值为 0.000 05,间隔中心点为 0,间隔偏移量为 ± 0.5。两种 BP 神经网络模型隐含层激活函数为双曲正切函数,输出层激活函数为 Softmax。

1.4 受试者工作特征曲线(receiver operator characteristic curve, ROC 曲线)比较 ROC 曲线是根据一系列不同的二分类方式,以灵敏度为纵坐标,1 - 特异度为横坐标绘制的曲线。通过分别计算各个试验的 ROC 曲线下面积(area under curve, AUC) 进行比较, AUC 最大者表示该试验的诊断价值最佳。ROC 的 AUC 值在 0.5 和 1.0 之间。在 AUC > 0.5 的情况下, AUC 越接近于 1,说明诊断效果越好。AUC = 0.5 时,说明诊断方法完全不起作用,无诊断价值。所以对 AUC 要首先检验与 0.5 的差异有无统计学意义。本研究利用医学统计软件包 Medcalc, 使用先前获得预测样本的 Logistic 回归和人工神经网络的预测拟概率,绘制 ROC 曲线,并进行不同 AUC 的非参数显著性检验(统计量为 Z)。

2 结果

2.1 一般情况 2009 年 ~ 2012 年肺癌患者术后住院期间死亡 1 例,共 348 例患者纳入该研究。其中男 220 例,女 128 例, ≥ 70 岁患者 102 例, < 70 岁患者 246 例。术后发生心肺并发症的患者共 79 例,占手术患者的 22.7%。研究变量主要包括年龄、性别、术中输血、切除部位、病理类型、吸烟史、糖尿病史、高血压史、冠心病史、手术方式、术前生理指标 [白蛋白、血红蛋白和一秒用力呼气容积/用力肺活量(FEV₁%)]、术前化疗史、心电图检查共 15 项。变量赋值及计数统计见表 1。

2.2 Logistic 分析结果 由于 Logistic 模型建立方程不需要校验样本,故将训练样本和校验样本都用来拟合模型。将所有 15 个变量纳入 Logistic 多因素分析,结果有 5 个变量作为肺癌并发症发生的危险因素有统计学意义,分别为年龄、手术方式、高血压史、白蛋白、FEV₁%。见表 2。根据其模型参数建立的预测方程为:

$$P = \frac{e(-2.595 + 0.962X8 + 0.870X4 + 0.825X2 + 1.024X15 + 1.094X13)}{1 + e(-2.595 + 0.962X8 + 0.870X4 + 0.825X2 + 1.024X15 + 1.094X13)}$$

Nagelkerke 决定系数为 0.778,显示模型拟合程度较好。根据方程对预测集数据进行预测。

表 1 肺癌患者一般情况统计

变量	赋值	病例数(%)	术后并发症(%)
性别(X1)	男(X1=1)	220(63.2)	53(24.1)
	女(X1=0)	128(36.8)	26(20.3)
年龄(X2)	≥70(X2=1)	102(29.3)	15(14.7)
	<70(X2=0)	246(70.7)	64(26.0)
术前化疗史(X3)	有(X3=1)	31(8.9)	8(25.8)
	无(X3=0)	317(91.1)	71(22.4)
手术方法(X4)	全肺切除(X4=1)	56(14.1)	19(33.9)
	肺段或肺叶切除(X4=0)	292(85.9)	60(20.5)
术中输血(X5)	是(X5=1)	45(16.1)	15(33.3)
	否(X5=0)	303(83.9)	64(21.1)
切除位置(X6)	右(X6=1)	210(60.3)	51(24.2)
	左(X6=0)	138(39.7)	28(20.3)
病理类型(X7)	鳞癌(X7=1)	189(54.3)	47(24.9)
	腺癌及其他(X7=0)	159(45.7)	32(20.1)
高血压史(X8)	是(X8=1)	104(29.9)	51(49.0)
	否(X8=0)	244(70.1)	40(16.4)
糖尿病史(X9)	是(X9=1)	70(20.1)	42(60.0)
	否(X9=0)	278(79.9)	37(13.3)
冠心病史(X10)	是(X10=1)	63(18.1)	24(38.1)
	否(X10=0)	285(81.9)	55(19.3)
心律失常(X11)	是(X11=1)	91(26.1)	31(34.1)
	否(X11=0)	257(73.9)	48(18.7)
吸烟史(X12)	是(X12=1)	212(60.9)	60(28.3)
	否(X12=0)	136(39.1)	19(14.0)
FEV ₁ % (X13)	<70(X13=1)	70(20.1)	34(48.6)
	≥70(X13=0)	278(79.9)	45(16.2)
白蛋白(X14)	<正常值(X14=1)	78(22.4)	34(43.6)
	正常(X14=0)	270(77.6)	45(16.7)
血红蛋白(X15)	<正常值(X15=1)	100(28.7)	28(28.0)
	正常(X15=0)	248(71.3)	51(20.6)

表 2 肺癌术后心肺并发症发生 Logistic 多因素分析

变量	β	SE(β)	P 值	OR(95% CI)
高血压(X8)	0.962	0.337	0.004	2.62(1.35 ~ 5.06)
手术方法(X4)	0.870	0.415	0.036	2.39(1.06 ~ 5.38)
年龄(X2)	0.825	0.351	0.019	2.28(1.15 ~ 4.54)
白蛋白(X15)	1.024	0.348	0.003	2.79(1.41 ~ 5.51)
FEV ₁ % (X13)	1.094	0.363	0.003	2.99(1.47 ~ 6.09)
常数项	-2.595	0.310	-	-

2.3 BP 神经网络分析结果

2.3.1 BP 神经网络 将所有 15 个变量纳入神经网络模型中。通过“自动体系结构选择”模型,隐含层单元数为 4。对各个自变量进行敏感度分析,结果显示对肺癌并发症发生影响最大的前 5 位因素依次是白蛋白(0.112)、高血压(0.111)、FEV₁%

(0.104)、年龄(0.098)和吸烟史(0.096)。

2.3.2 经 Logistic 回归筛选变量后的 BP 神经网络

联合应用 Logistic 方程与 BP 神经网络,输入层变量为 Logistic 多因素分析结果中有意义的变量共 5 个。对各个自变量进行敏感度分析,结果显示对肺癌并发症发生影响最大的因素是 FEV₁% (0.302),其余依次是高血压(0.208)、年龄(0.205)、白蛋白(0.196)和手术方式(0.089)。

2.4 模型预测能力比较

2.4.1 预测结果的比较

Logistic 预测模型的一致率、灵敏度和特异度均低于 BP 神经网络模型和经 Logistic 模型筛选变量的 BP 神经网络模型。表明 Logistic 模型的预测能力低于 BP 神经网络模型。见表 3。

表 3 3 种模型预测能力比较 (%)

预测结果	一致率	灵敏度	特异度
Logistic 回归模型	81.6	31.6	95.6
BP 神经网络	89.7	63.2	97.1
Logistic 回归模型 + BP 神经网络	90.8	63.2	98.5

2.4.2 ROC 曲线面积比较

通过 Medcalc 计算, Logistic 模型预测的 AUC 为 0.636, 95% CI 为 0.526 ~ 0.736。BP 神经网络模型的 AUC 为 0.801, 95% CI 为 0.702 ~ 0.879。经 Logistic 模型筛选变量的 BP 神经网络模型的 AUC 为 0.808, 95% CI 为 0.710 ~ 0.885。见表 4。经检验,三者 AUC 与 0.5 比较差异有统计学意义 ($P < 0.05$)。3 种模型 AUC 两两比较结果显示 Logistic 模型预测的 AUC 与 BP 神经网络模型 ($Z = 2.696, P = 0.007$)、经 Logistic 模型筛选变量的 BP 神经网络模型的 AUC ($Z = 2.464, P = 0.014$) 的差异均有统计学意义 ($P < 0.05$)。表明 BP 神经网络模型在预测性能上优于 Logistic 回归模型。两种 BP 神经网络模型的 AUC 比较差异无统计学意义 ($Z = 0.191, P = 0.849$)。见图 1。

表 4 3 种模型 ROC 的 AUC

预测结果	AUC	SE	95% CI
Logistic 回归模型	0.636	0.075 5	0.526 ~ 0.736
BP 神经网络	0.801	0.064 6	0.702 ~ 0.879
Logistic 回归模型 + BP 神经网络	0.808	0.063 7	0.710 ~ 0.885

3 讨论

3.1 肺癌术后并发症发生的危险因素

肺癌患者多为中老年人群,身体免疫力和耐受力低下,多数患者还同时伴有其他疾病。肺癌手术往往创面大,影

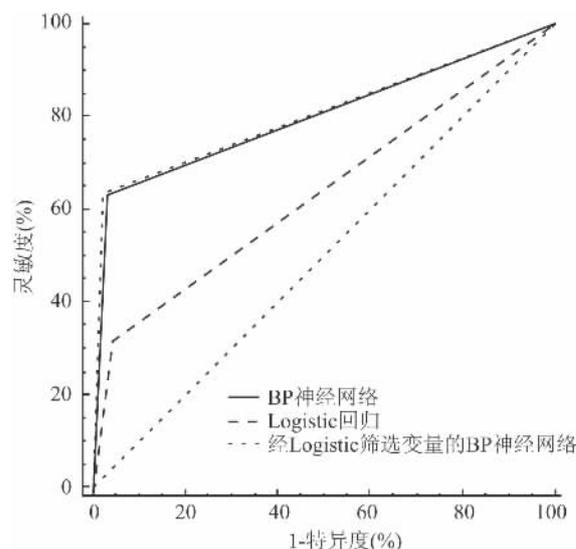


图 1 3 种预测模型 ROC 曲线

响患者呼吸功能。本研究数据显示术后并发症发生率为 22.7%,故有必要对肺癌术后并发症进行预测。本研究通过 Logistic 回归多因素分析显示,年龄 ≥ 70 岁和全肺切除是肺癌术后并发症的危险因素,这与钱永祥^[4]研究结果相同。随着年龄的增加,机体抵抗力下降,伴随疾病增多和恢复较慢会增加发生并发症的风险。肺切除的范围决定了对肺周围血管和神经丛损害的大小。全肺切除一般手术创面大,切除后会导致肺容量减少和心脏负荷增加等,进而导致发生并发症的可能性增加。FEV₁% 较低反映了患者的肺功能损减,与国外研究^[5]相同,本研究也证实了 FEV₁% 降低是术后并发症发生的危险因素。因此,对于 FEV₁% 偏低患者术后应予以重视,术前的肺功能恢复锻炼也是有利的。由于肿瘤是一种消耗性疾病,白蛋白减少患者一般存在营养不良的情况,会影响术后恢复的时间,增加患病的可能。对于伴有高血压的肺癌患者,高血压会损害患者的心脏,引起心脏结构异常,甚至造成心功能的损害。本研究结果证实高血压也是肺癌术后并发症的危险因素之一。因此,术前对血压的控制十分必要。

3.2 BP 神经网络模型优缺点

神经网络模型是生物神经网络在结构、功能及某些基本特性方面的理论抽象和模拟而构成的一种信息处理系统。目前已广泛应用于医学的各个领域,如疾病诊断及卫生管理等^[6-7]。BP 神经网络模型属于有监督学习的网络模型,被认为是稳定性较强的人工神经网络之一。与传统的统计学方法相比, BP 神经网络模型没有精

确的数学模型,对自变量和因变量均没有任何要求(如正态性和独立性)。在处理一系列复杂的非线性问题时,具有很强的容错性。但是,神经网络仍然存在一些不足需要解决。首先,建立神经网络模型的过程中各参数及函数的设置并没有统一的标准,这些设置的正确性缺乏理论依据^[2-3]。本研究中采用SPSS软件中的“自动体系结构选择”功能,有系统自行运算选择最优的参数设置。其次,神经网络模型并不能像Logistic回归那样给出明确的方程,每个自变量对因变量作用的合理解释尚无法知晓。变量准入和剔除的原则尚不明确。

3.3 BP神经网络与Logistic回归的预测应用 本研究结果显示神经网络预测性能优于Logistic回归,而两种纳入不同变量数目的BP神经网络的预测能力的差异并无统计学意义。虽然神经网络模型预测性能高于Logistic回归,但由于神经网络模型无法对自变量进行解释,所以在实际应用中,可将两者相互配合,充分发挥各自的优点,提高应用价值。同时,国内外许多学者也对神经网络模型与Logistic模型的预测性能进行了比较。Shi et al^[8]比较两种模型预测原发性肝癌术后住院死亡的准确性,结果同样提示神经网络模型具有更高的准确度。Biglarian et al^[9]使用这两种模型对大肠癌患者发生远处转移进行预测,Logistic回归和神经网络模型AUC分别为0.77和0.82,后者优于前者。有学者使用Logistic回归、决策树和神经网络模型预测胃癌术后并发症,结果显示神经网络模型是3种模型中预测准确性最

高^[10]。

参考文献

- [1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013 [J]. CA Cancer J Clin 2013 63(1): 11-30.
- [2] 李丽霞,王彤,范逢曦. BP神经网络与logistic回归的比较研究[J]. 中国卫生统计 2005 22(3): 138-40.
- [3] 元昌安. 数据挖掘原理与SPSS Clementine应用宝典[M]. 北京: 电子工业出版社 2009: 239-47.
- [4] 钱永祥. 非小细胞肺癌术后并发症的危险因素分析和预防措施[J]. 癌症进展, 2012, 10(6): 627-9.
- [5] Takamochi K, Oh S, Matsuoka J, et al. Risk factors for morbidity after pulmonary resection for lung cancer in younger and elderly patients [J]. Interact Cardiovasc Thorac Surg, 2011, 12(5): 739-43.
- [6] 余小兰,姚永忠,桑剑锋,等. 基于BP神经网络的甲状腺癌微创诊断模型的研究[J]. 现代生物医学进展, 2012, 12(36): 7104-8.
- [7] 叶健伟,沈亚诚,黄小玲. 基于BP神经网络的肺炎医保住院费用分析[J]. 卫生经济研究 2013, 6: 38-40.
- [8] Shi H Y, Lee K T, Lee H H, et al. Comparison of artificial neural network and logistic regression models for predicting in-hospital mortality after primary liver cancer surgery [J]. PLoS One 2012 7(4): e35781.
- [9] Biglarian A, Bakhshi E, Gohari M R, et al. Artificial neural network for prediction of distant metastasis in colorectal cancer [J]. Asian Pac J Cancer Prev, 2012, 13(3): 927-30.
- [10] Chien C W, Lee Y C, Ma T S, et al. The application of artificial neural networks and decision tree model in predicting post-operative complication for gastric cancer patients [J]. Hepatogastroenterology 2008 55(84): 1140-5.

A comparison between three statistical models in predicting post-operative complication for lung cancer patients

Song Jian, Su Hong, Zhou Yangyang, et al

(Dept of Epidemiology and Biostatistics, School of Public Health, Anhui Medical University, Hefei 230032)

Abstract Objective To explore the application value of BP neural network in predicting post-operative complication for lung cancer patients. **Methods** We applied Logistic regression, BP neural network model and BP neural network model screening variables by Logistic regression to establish prediction models and evaluate the practical application of each model in the prediction accuracy. **Results** The prediction accuracy of Logistic regression, BP neural network model and BP neural network model screening variables by Logistic regression were 81.6%, 89.7%, 90.8% and the AUC of ROC in the three models were 0.636, 0.801, 0.808, respectively. There were significant differences of the AUC of ROC between Logistic regression and two BP neural network models. **Conclusion**

The discrimination performance of BP neural network models is better than Logistic regression in the prediction of post-operative complication for lung cancer patients.

Key words Logistic regression; BP neural network; lung cancer; complication