

- [13] JOOSTEN LA, CRIŞAN TO, AZAM T, et al. Alpha-1-anti-trypsin-Fc fusion protein ameliorates gouty arthritis by reducing release and extracellular processing of IL-1 β and by the induction of endogenous IL-1Ra[J]. Ann Rheum Dis, 2016, 75(6): 1219-1227.
- [14] MARCHETTI C, SWARTZWELTER B, KOENDERS MI, et al. NLRP3 inflammasome inhibitor OLT1177 suppresses joint inflammation in murine models of acute arthritis[J]. Arthritis Res Ther, 2018, 20(1): 169.
- [15] 曾惠琼, 罗国辉, 卢小平, 等. 艾拉莫德对膝关节炎的疗效及作用机制初探[J]. 中华生物医学工程杂志, 2019, 25(2): 216-219.
- [16] 苏杉, 王蓉, 刘飞, 等. 秦艽不同配伍对风寒湿痹类风湿关节炎模型大鼠关节保护作用及其机制[J]. 中国临床药理学杂志, 2018, 34(21): 2542-2545.
- [17] TAKESHITA M, KUNO A, SUZUKI K, et al. Alteration of matrix metalloproteinase-3 O-glycan structure as a biomarker for disease activity of rheumatoid arthritis[J]. Arthritis Res Ther, 2016, 18(1): 112.
- [18] 李小峰, 樊凯, 安海转, 等. 骨髓间充质干细胞移植对胶原诱导关节炎大鼠血清及关节滑膜中白细胞介素-22及基质金属蛋白酶-3表达的影响[J]. 中华医学杂志, 2017, 97(9): 698-702.
- [19] 杨映雄. 骨创伤患者创伤后血浆D-二聚体水平与创伤程度的相关性研究[J]. 安徽医药, 2018, 22(3): 460-462.

(收稿日期: 2019-12-15, 修回日期: 2020-02-14)

引用本文: 张斌, 刘弋, 胡宗举. 人工神经网络在胃癌预后预测中的应用[J]. 安徽医药, 2021, 25(5): 987-991. DOI: 10.3969/j.issn.1009-6469.2021.05.035.

◇ 临床医学 ◇



人工神经网络在胃癌预后预测中的应用

张斌¹, 刘弋², 胡宗举³

作者单位:¹安徽省阜阳民生医院普外科, 安徽 阜阳 236000; ²安徽医科大学第一附属医院普外科, 安徽 合肥 230022; ³安徽省阜阳市人民医院普外科, 安徽 阜阳 236000

摘要: **目的** 应用人工神经网络(ANN)技术建立胃癌预后预测模型, 并与传统 logistic 回归模型相比较, 探讨其临床意义。**方法** 回顾性分析安徽省阜阳市人民医院 2011 年 1 月至 2016 年 12 月收治的 152 例经内镜活检或术后病理证实的胃癌病人的临床及病理资料, 单因素生存分析筛选出有统计学意义的指标, 用人工神经网络技术建立胃癌的预后模型对胃癌术后生存时间进行预测, 同时与传统 logistic 回归预测做对比。**结果** 癌胚抗原(CEA)水平、Borrmann 分型、浸润深度、淋巴结转移、分化程度、TNM 分期、手术方式、神经脉管浸润与胃癌病人的预后有关, ANN 对胃癌病人的 5 年生存率预测准确率 85.19%, 敏感度 88.24%, 特异度 80%; 二元 logistic 回归对胃癌病人的 5 年生存率预测准确率 74.07%, 敏感度 82.35%, 特异度 60%; ANN 模型在不同 TNM 分期中的预测准确率 77.78%, 敏感度 40%, 特异度 92.31%。**结论** 人工神经网络作为预测胃癌术后生存时间的一种新的方法, 具有良好的应用前景。

关键词: 胃肿瘤; 人工智能; 人工神经网络; 预后

Application of artificial neural networks in prognosis of gastric cancer

ZHANG Bin¹, LIU Yi², HU Zongju³

Author Affiliations: ¹Department of General Surgery, Fuyang Minsheng Hospital, Fuyang, Anhui 236000, China; ²Department of General Surgery, the First Affiliated Hospital of Anhui Medical University, Hefei, Anhui 230022, China; ³Department of General Surgery, Fuyang People's Hospital, Fuyang, Anhui 236000, China

Abstract: **Objective** Artificial neural network (ANN) technology was used to establish a prognostic prediction model for gastric cancer, and compared with the traditional logistic regression model to explore its clinical significance. **Methods** The clinical data of 152 cases which were confirmed to be gastric cancer after endoscopic biopsy and postoperative pathologic from January, 2011 to December, 2016 in Fuyang people's hospital of Anhui Province were analyzed retrospectively. Screening out statistically significant indicators by single factor survival analysis. To build a prediction model of gastric cancer prognosis by artificial neural network technology, compared with the traditional logistic regression model. **Results** CEA level, Borrmann typing, depth of invasion, lymph node metastasis, degree of differentiation, TNM stage, operation method and neurovascular invasion are related to the prognosis of gastric cancer patients. The accuracy, sensitivity and specificity of 5-year survival rate of gastric cancer patients predicted by artificial neural network was 85.19%, 88.24%, 80%; the accuracy, sensitivity and specificity of binary logistic regression was 74.07%, 82.35%, 60%; the accuracy, sensitivity and specificity of artificial neural network model in different TNM stages was 77.78%, 40%, 92.31%. **Conclusion** As a new method to predict the survival time of gastric cancer, ANN has a good application prospect.

Key words: Stomach neoplasms; Artificial intelligence; Artificial neural network; Prognosis

根据2019年国家癌症中心发布的最新数据显示,胃癌发病率居我国恶性肿瘤第2位,死亡率居第3位^[1]。国内的研究报告胃癌术后5年生存率普遍低于30%^[2],远低于欧美及日本等发达国家^[3]。因此,明确胃癌术后预后的影响因素及通过各种影响因素对胃癌术后生存时间进行预测,有着重要的临床意义。本研究用人工神经网络技术(ANN)建立胃癌的预后模型对胃癌术后5年生存率进行预测,与传统logistic回归模型预测进行比较,寻求胃癌预后最佳预测方法。

1 材料与方法

1.1 一般资料 收集2011年1月至2016年12月阜阳市人民医院收治的胃癌病人的临床资料。纳入标准:经胃镜活检或术后病理证实为胃癌。排除标准:(1)临床资料不全者;(2)未行手术治疗者;(3)合并有严重心肺功能疾病者;(4)术后随访期内失访者(以及随访期未及5年且存活者)。剔除后共选取152例病例,其中,男性125例,女性27例,男女比为4.63:1,年龄62.4岁,年龄范围为22~82岁,>60岁占69.7%。术后随访方式为电话及门诊随访,前2年每6个月随访1次,后每年随访1次,共随访5年,随访日期截至2019年12月。

1.2 筛选预后因素及量化数据 根据以往文献的报告,选取以下有可能影响胃癌预后的因素,包括性别、年龄、家族肿瘤史、血清癌胚抗原(CEA)水平、肿瘤发生部位、Borrmann分型、肿瘤最大直径、浸润

深度(T分期)、淋巴结转移数量(N分期)、TNM分期、分化程度、手术方式、术后化疗、神经脉管浸润共14项观察指标。将以上14项可能影响胃癌预后的观察指标量化处理,转化成满足ANN和logistic回归数据分析输入参数要求的数据。

1.3 统计学方法 本资料单因素生存分析采用SPSS 24.0软件包进行分析,Kaplan-Meier法计算各组生存率,log-rank法比较各组间差异, $P<0.05$ 为差异有统计学意义。在Matlab2017b的环境下采用人工神经网络(ANN)和logistic回归对胃癌术后5年生存概率进行预测,其中人工神经网络使用Matlab神经网络工具箱中的BP神经网络。

2 结果

2.1 总体生存率 152例病人中,截止到5年随访期结束共存活45例,死亡107例,中位生存期33个月,术后1年、3年、5年生存率分别为82.89%、48.68%、29.61%。

2.2 单因素分析 将152例病人的临床及病理资料分组,其中13例行姑息性手术(内引流、造瘘或大网膜切除活检术),缺少Borrmann分型、肿瘤大小、浸润深度、淋巴结转移、分化程度、神经脉管浸润等资料。采用Kaplan-Meier法计算各组生存率,Log-rank法比较各组间差异,结果显示:CEA水平、Borrmann分型、浸润深度、淋巴结转移、分化程度、TNM分期、手术方式、神经脉管浸润共8项指标差异有统计学意义($P<0.05$),见表1。

表1 影响胃癌预后的临床病理参数分析

临床病理参数	例数	百分比/%	中位生存时间/(月, $\bar{x} \pm s$)	χ^2 值	P值
性别				0.369	0.544
男	125	82	33±5		
女	27	18	38±13		
年龄				0.003	0.956
≤65岁	86	57	38±7		
>65岁	66	43	31±4		
家族肿瘤史				1.907	0.167
有	24	16	30±4		
无	128	84	41±8		
CEA水平				19.395	<0.05
正常	95	62	44±5		
升高	57	38	25±3		
肿瘤部位				0.595	0.743
贲门	50	33	37±6		
胃窦	33	22	38±6		
胃体	69	45	27±7		
Borrmann分型				9.982	<0.05
I	7	5	44±4		
II	33	24	52±5		

续表1 影响胃癌预后的临床病理参数分析

临床病理参数	例数	百分比/%	中位生存时间/(月, $\bar{x} \pm s$)	χ^2 值	P值
Ⅲ	86	62	33±5		
Ⅳ	13	9	22±5		
肿瘤大小				2.551	0.11
<5 cm	79	57	43±4		
≥5 cm	60	43	32±3		
浸润深度				11.641	<0.05
T1	12	9	—		
T2	14	10	45±2		
T3	98	70	33±5		
T4	15	11	22±6		
淋巴结转移				18.723	<0.05
N0	37	27	—		
N1	24	17	44±10		
N2	36	26	35±6		
N3	42	30	26±3		
分化程度				51.785	<0.05
高分化腺癌	5	4	—		
中分化腺癌	51	37	—		
低分化腺癌	63	45	31±6		
低分化黏液腺癌或印戒细胞癌	20	14	17±8		
TNM分期				30.154	<0.05
I a	11	7	—		
I b	7	5	54±10		
Ⅱ a	24	16	45±6		
Ⅱ b	21	14	37±13		
Ⅲ a	29	19	31±2		
Ⅲ b	37	24	30±5		
Ⅲ c	10	6	19±2		
Ⅳ	13	9	8±2		
手术方式				21.694	<0.05
胃大部切除术	41	27	45±3		
全胃切除术	89	59	32±3		
联合脏器切除术	9	6	—		
姑息手术	13	8	8±2		
术后化疗				1.423	0.233
有	121	80	38±5		
无	31	20	26±6		
神经脉管浸润				9.859	<0.05
有	33	24	26±5		
无	106	76	43±4		

2.3 建立ANN预测模型 因建立ANN训练模型不能存在缺失因素,将13例缺乏病理资料的姑息手术病人从152例总数中剔除,将剩下的139例病人按4:1随机分为训练集和测试集,即112例训练样本和27例测试样本。采用BP神经网络,输入层为单因素分析中8项有统计学差异的指标,记为 $x_j, j=1, 2, \dots, 8$;输出层为是否5年生存;隐含层含有 $k=\sqrt{N}+a$ 个节点个数。其中, N 为病例个数, a 为小于10的自然数,结构图见图1。

选取S型正切函数作为BPNN的隐含层及输出层神经元的传递函数,ANN预测模型如下: $y =$

$$\tan \operatorname{sig}\left(\sum_{i=1}^k w_{2i} a_i\right)=\tan \operatorname{sig}\left(\sum_{i=1}^k\left(w_{2i} \tan \operatorname{sig}\left(\sum_{j=1}^8 v_{ij} x_j\right)\right)\right)$$

其中, x_j 为输入指标, y 为输出值, a_i 为隐含层神经元, v_{ij} 为输入层到隐含层的权值, w_{2i} 为隐含层到输出层的权值。

设置神经网络训练的目标误差为 10^{-6} ,最大迭代次数为500,学习率为0.01, $a=3, k=15$ 。当 $y \geq$

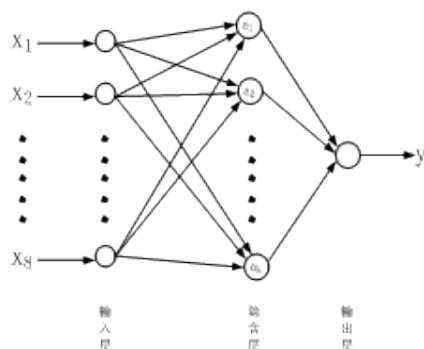


图1 BP神经网络结构图

0时,病例的预测结果为1,即是“5年死亡”,否则,病例的预测结果为0,即是“5年生存”。27例测试样本运算结果显示准确率 85.19% (23/27), 敏感度 88.24% (15/17), 特异度 80% (8/10), 如表2所示。为了更直观的展示预测效果,将预测结果与真实值进行比较,见图2。

表2 ANN测试集预测结果/例

ANN 预测结果	实际生存状态		合计
	死亡	生存	
死亡	15	2	17
生存	2	8	10
合计	17	10	27

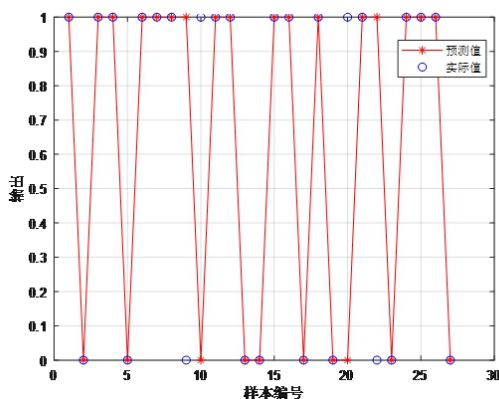


图2 测试集(27例随机样本)运算结果图

2.4 建立 logistic 回归预测模型 为了避免和 ANN 预测模型对比的人为误差,沿用 ANN 预测模型的 112 例训练样本和 27 例测试样本,将单因素生存分析中 8 项有统计学差异的指标纳入 logistic 回归预测作为分析的自变量。logistic 回归模型如下:

$$y = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x_1 + \dots + \theta_8 x_8)}}$$

设置初始参数 $\theta_0 = \theta_1 = \dots = \theta_8 = 1$, 步长为 0.01, 损失精度为 0.000 1。当 $y \geq 0.5$ 时,病例的预测结果为 1,即是“5年死亡”,否则,病例的预测结果为 0,即是“5年生存”。预测结果为是否 5 年生存,结果提示 27 例测试样本预测准确率 74.07% (20/

27), 敏感度 82.35% (14/17), 特异度 60% (6/10), 见表3。

表3 logistic回归测试集预测结果/例

logistic 回归预测结果	实际生存状态		合计
	死亡	生存	
死亡	14	4	18
生存	3	6	9
合计	17	10	27

2.5 ANN 模型在不同 TNM 分期中的预测 为了研究 ANN 预测模型对于不同 TNM 分期病人是否适用,将 139 例病人根据 TNM 分期定义训练集和测试集。沿用上文 ANN 所设置的模型参数,将 II 期和 III 期病人共 121 例设为训练集学习预测模型,对 18 例 I 期病人(测试集)的 5 年生存率进行预测,结果显示预测准确率 77.78% (14/18), 敏感度 40% (2/5), 特异度 92.31% (12/13), 具体结果见图3、表4。

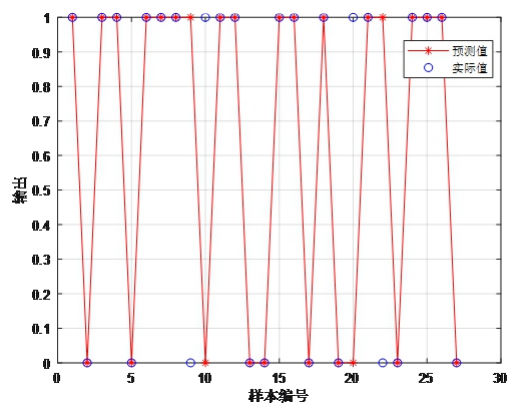


图3 测试集(I期病人)运算结果图

表4 ANN测试集预测结果/例

ANN 预测结果	实际生存状态		合计
	死亡	生存	
死亡	2	1	3
生存	3	12	15
合计	5	13	18

3 讨论

胃癌因起病隐匿、症状不典型、转移及复发几率高等特点,预后较差。在临床工作中,对于胃癌术后,其病人及家属不可避免的会问一个问题“我还能活多长时间?”。因影响胃癌预后的因素较为复杂,对于这个问题,临床医师往往根据自己的临床经验及主观因素给出一较模糊的答案。本研究通过分析 152 例胃癌术后的临床及病理资料,探索其影响胃癌预后的各种独立因素,分别用人工神经网络(BP神经网络)、logistic 回归两种方法建立胃癌的预后模型对胃癌术后生存时间进行预测,寻求胃癌预后最佳的预测方法。

本研究的单因素生存分析显示,CEA水平、Borrmann分型、浸润深度、淋巴结转移、分化程度、TNM分期、手术方式、神经管浸润8个因素与胃癌术后的总生存时间呈显著性相关($P<0.05$)。CEA水平越高、TNM分期越晚、分化程度越差、存在神经管浸润的病人其生存时间越短、5年生存率越低,根治性手术病人明显要比姑息性手术病人生存时间长。本研究结果与国内外目前研究结果相一致,Hou等^[4]研究发现病理分化程度越高、TNM分期越早、淋巴结转移率越低的胃癌病人,其生存时间越长。Ueno等^[5]研究发现,以75岁为界限,75岁以上病人病理结果和术后并发症是其生存率的重要预后因素,而75岁以下则病理结果和术后化疗是其重要的预后因素。Katai等^[3]对日本118 367例胃癌病人进行统计,其中I期胃癌病人占总数的58.7%,总5年生存率达到71.1%,远远超过国内研究统计的5年生存率,本研究的总体5年生存率偏低(29.61%),考虑原因与此次研究资料以进展期胃癌为主、未排除IV期行姑息性手术病人且早期胃癌占总样本比例较小(I期胃癌仅占比11.8%)有关。因此深入研究影响胃癌预后的相关因素极其重要,对于降低胃癌术后复发转移率有着重要的意义。

ANN是模拟人类大脑的神经系统,对复杂信息进行综合处理的一种数学模型,因其智能化的自适应学习能力特点,被广泛应用在语音识别、图像识别、医学诊断与预测等^[6-8]各个专业领域。人工神经网络包括BP神经网络、概率神经网络、卷积神经网络等,其中BP神经网络应用较为广泛。Walczak等^[9]使用ANN方法对219例胰腺癌病人进行术后7个月的生存预测,结果发现ANN的敏感度在90%以上。郑骏明、高新^[10]研究发现通过人工神经网络建立去势抵抗性前列腺癌病人的预后预测模型跟传统COX比例风险模型相比具有更高的准确性。唐雯等^[11]发现对656例腹膜透析病人的预后构建预测模型,循环神经网络相比传统logistic回归有着明显的优越性。通过本研究结果可以观察到,在随机分配的ANN预测模型中,27例胃癌测试样本中BP神经网络的预测准确率(85.19%)、敏感度(88.24%)和特异度(80%)高于传统logistic回归模型预测(74.07%、82.35%、60%),结果与国内外的研究较为一致;在不同TNM分期构建的ANN预测模型中,根据II期和III期病人的临床数据构建的ANN预测模型对I期病人同样有着较高的预测准确度和特异度,说明了ANN预测模型的普适性,不受TNM分期的影响。但预测结果敏感度相对较低,不能排除测试样本的总阳性数量偏少所致,在以后的研究中可以增加研究样本数量以求得更精确的效果。

本研究的结论提示人工神经网络相比传统logistic回归在胃癌预后方面有着更好的准确性,但在医学生存分析领域,并不能说明人工神经网络可以完全替代传统logistic回归、Cox回归等经典回归方法,但人工神经网络的优势在于处理复杂的非线性关系、较好的容错率、可学习性和适应性强等。故笔者认为,在目前的研究趋势中,人工神经网络方法可以和传统logistic回归、Cox回归等经典回归方法相互补充,进一步提高恶性肿瘤术后的预后预测。同时本研究中两种方法的预测准确率均未达到90%以上,考虑可能由于总样本量偏少、对胃癌预后分析所采用的影响因素不够完全等原因造成。在以后的研究中可以增加总样本量及增加多个可能影响胃癌预后的独立因素,以求达到更高的预测值,从而更好辅助临床进行决策及干预。

参考文献

- [1] 曹毛毛,陈万青.中国恶性肿瘤流行情况及防控现状[J].中国肿瘤临床,2019,46(3):145-149.
- [2] 左婷婷,郑荣寿,曾红梅,等.中国胃癌流行病学现状[J].中国肿瘤临床,2017,44(1):52-58.
- [3] KATAI H, ISHIKAWA T, AKAZAWA K, et al. Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007)[J]. Gastric Cancer, 2018, 21(1): 144-154.
- [4] HOU Y, WANG X, CHEN J. Prognostic significance of metastatic lymph node ratio: the lymph node ratio could be a prognostic indicator for patients with gastric cancer [J]. World Journal of Surgical Oncology, 2018, 16(1): 198-206.
- [5] UENO D, MATSUMOTO H, KUBOTA H, et al. Prognostic factors for gastrectomy in elderly patients with gastric cancer [J]. World Journal of Surgical Oncology, 2017, 15(1): 59-69.
- [6] OH SE, CHOI MG, SEO SW. ASO Author reflections: use of the survival recurrent network for prediction of overall survival in patients with gastric cancer [J]. Ann Surg Oncol, 2019, 26(Suppl 3):539-540.
- [7] WANG NB, CHEN JH, WU L, et al. Application of artificial neural network model in diagnosis of Alzheimer's disease [J]. BMC Neurology, 2019, 19(1): 154-161.
- [8] 王小亮,黄萍.基于近红外光谱技术和人工神经网络算法的佐匹克隆片定量分析[J].安徽医药,2019,23(1):50-54.
- [9] WALCZAK S, VELANOVICH V. An evaluation of artificial neural networks in predicting pancreatic cancer survival [J]. Journal of Gastrointestinal Surgery, 2017, 21(10): 1606-1612.
- [10] 郑骏明,高新.去势抵抗性前列腺癌人工神经网络预后预测模型的建立及其准确性分析[J/CD].中华腔镜泌尿外科杂志(电子版),2013,7(6):419-423. DOI: 10.3877/cma.j.issn.1674-3253.2013.06.003.
- [11] 唐雯,高峻逸,马辛宇,等.循环神经网络模型在腹膜透析临床预后预测中的初步应用[J].北京大学学报(医学版),2019,51(3):602-608.

(收稿日期:2020-02-13,修回日期:2020-03-19)