

前列腺癌人工神经网络诊断模型的应用研究

邱志磊¹ 徐立柱¹ 贾魁² 梁鑫¹ 胡伟¹ 郑帅帅¹

【摘要】 目的:结合患者临床信息和前列腺肿瘤标记物进行数据挖掘,建立基于人工神经网络的前列腺癌诊断模型,评估其准确性,并为前列腺癌的临床诊断提供客观的参考信息。**方法:**回顾分析 2010 年 1 月~2014 年 6 月在我院就诊的前列腺病患 310 例,其中前列腺癌组 100 例,非前列腺癌组 210 例。患者均接受经直肠前列腺穿刺活检。用 210 例样本(前列腺癌组 70 例,非前列腺癌组 140 例)建立人工神经网络(ANN)模型,并用 100 例样本(前列腺癌组 30 例,非前列腺癌组 70 例)盲法测试和评估此模型。**结果:**纳入分析的指标有年龄、前列腺特异性抗原(TPSA、FPSA/TPSA、PSAD)、直肠指检、前列腺体积、经直肠超声和前列腺核磁共振结果等,经过反复训练建立的 ANN 模型对测试样本预测的特异度为 80.0%,灵敏度为 93.3%。**结论:**ANN 在计算机辅助前列腺癌的诊断,评估患者患有前列腺癌的风险,以及对穿刺检测的指导等方面具有广阔的应用前景。

【关键词】 前列腺癌;人工神经网络;诊断模型

doi: 10.13201/j.issn.1001-1420.2015.10.011

【中图分类号】 R737.25 **【文献标识码】** A

Study on application of artificial neural network in the diagnosis model of prostate cancer

QIU Zhilei¹ XU Lizhu¹ JIA Kui² LIANG Xin¹ HU Wei¹ ZHENG Shuaishuai¹

(¹ Department of Urology, Qingdao Municipal Hospital, Qingdao, 266071, China; ² Hisense National Key Laboratory)

Corresponding author: Xu Lizhu, E-mail: qiuzhilei2006@163.com

Abstract Objective: To establish diagnostic model for prostate cancer based on artificial neural network (ANN) by combining the clinical information and tumor makers of prostate cancer, and evaluate the performance of ANN in order to provide references for clinical diagnosis of prostate cancer. **Method:** Clinical data of 310 patients with prostatic disease in our institution from January 2010 to June 2014 were collected and analyzed. The cases were divided into prostate cancer group (100 cases) and non-prostate cancer group (210 cases). All the patients underwent transrectal prostate biopsy. A total of 210 cases (70 cases from prostate cancer group and 140 cases from non-prostate cancer group) were used to establish the diagnostic model with ANN and 100 samples (30 samples from prostate cancer group and 70 samples from non-prostate cancer group) were used to test and evaluate this model. **Result:** Age, PSA (TPSA, F/TPSA, PSAD), rectal examination, prostatic volume, the results of prostate transrectal ultrasound and magnetic resonance were used for the analysis. After repeated training of the ANN model, the specificity and sensitivity of this model were 80.0% and 93.3%. **Conclusion:** The use of ANN shows great promise in computer-aided diagnosis, risk assessment of prostate cancer and prostate biopsy indication.

Key words prostate cancer; artificial neural network; diagnosis model

前列腺癌是男性泌尿生殖系统的常见恶性肿瘤,已成为西方国家男性的主要死亡原因之一。其发病率已超过肺癌居第 1 位,死亡率仅次于肺癌居第 2 位^[1]。我国的前列腺癌发病率虽然低于西方国家,但是近年来呈上升趋势^[2]。其诊断过程是在患者接受直肠指检和前列腺核磁共振检查之后,结合 PSA 水平对患者患病风险进行评估,最终通过前列腺穿刺活检来确诊。为了突破基于单一指标诊断模式的局限性和临床医生经验诊断的局限性,减少前列腺穿刺活检给患者带来的不必要创伤。

我们回顾性分析 2010 年 1 月~2014 年 6 月我院收治的前列腺病患 310 例。联合检测前列腺癌的多项相关指标,建立基于人工神经网络(artificial neural network, ANN)的前列腺癌诊断模型,综合研究患者的临床与实验室检查资料与其诊断结果之间的对应关系,提供较为准确的辅助诊断信息。现报告如下。

1 资料与方法

1.1 一般资料

本组 310 例,年龄 55~88 岁,平均年龄(71.3±15.8)岁,中位年龄 73.6 岁。其中前列腺癌组 100 例,非前列腺癌组 210 例。从本组患者中随机

¹ 青岛市市立医院泌尿外科(青岛,266071)

² 海信集团国家重点实验室

通信作者:徐立柱,E-mail: qiuzhilei2006@163.com

选取 100 例患者(前列腺癌组 30 例,非前列腺癌组 70 例)作为试验组。试验组年龄分布和 PSA 分布分别见表 1、表 2,患者均接受经直肠前列腺穿刺活检且经病理证实,前列腺癌患者的病理均提示为前列腺腺癌。

表 1 患者年龄分布				例
组别	50~ 59 岁	60~ 69 岁	70~ 79 岁	≥80 岁
前列腺癌组	1	6	10	13
非前列腺癌组	15	18	16	21
合计	16	24	26	34

表 2 患者 PSA 分布				例
组别	<4	4~10	11~20	>20
前列腺癌组	2	7	11	10
非前列腺癌组	46	12	8	4
合计	48	19	19	14

1.2 分析方法

随机选用 210 例样本(前列腺癌组 70 例,非前列腺癌组 140 例)训练人工神经网络,建立 ANN 模型,并用 100 例样本(前列腺癌组 30 例,非前列腺癌组 70 例)盲法测试和评估此模型。每例样本有 8 项输入指标,包括年龄、前列腺特异性抗原(TPSA、FPSA/TPSA、PSAD)、前列腺体积、直肠指检、经直肠超声和前列腺核磁共振结果,对应的病理检查结果为输出指标。对这些输入、输出指标进行量化处理,根据所有 8 项输入指标和对应 1 项输出指标,采用反向传播(back propogation, BP)方式进行训练,运用 Matlab 进行数据处理后,建立用于前列腺癌诊断的 ANN 模型。

典型的 BP 网络模型算法包括节点输入及输出、传递函数、训练误差和权值调整。

节点输入及输出:

节点输入: $net_i^{m+1} = \sum_j W_{ij}^m x_j^m$

节点输出: $x_i^{n+1} = f(net_i^{m+1})$

每一节点的输出由其节点输入、阈值和输入-输出传递函数决定,其中 x_i^m 为第 m 层第 j 个单元的输出, W_{ij}^m 表示第(m+1)层中的第 i 单元与第 m 层中第 j 个单元的连接权。

传递函数:传递函数一般取为(0,1)内连续取值 Sigmoid 函数:

$f(x) = \frac{1}{1 + e^{-x}}$

训练误差:训练误差是反映神经网络期望输出与计算输出之间误差大小的函数,通常取为:

$E_n = \frac{1}{2} \sum_i [x_i^M(n) - d_i(n)]^2$

对于第 n 个样本模式, $x_i^M(n)$ 为每一节点的网络预测输出, $d_i(n)$ 为期望输出。

权值调整:权值调整的过程,即连接下层节点和上层节点之间的权重矩阵 W_{ij}^m 的设定和误差修正过程,权值调整为:

$\Delta W_{ij}^m = \eta \sigma_i^{m+1} x_j^m$

其中 η 成为学习参数,取值范围 $0 \leq \eta \leq 1$ 。

$\sigma_i^M = (d_i - x_i^M)(1 - x_i^M),$

$\sigma_i^m = x_i^m(1 - x_i^m) \sum_j \sigma_i^{m+1} W_{ij}^m, m < M$

根据所有的 8 项输入指标和对应 1 项输出指标建立诊断前列腺癌的 ANN 模型(图 1)。其中神经网络的层数为 3 层。输入层 8 个节点对应各项参数指标,输出层可理解为诊断为前列腺癌的概率,隐藏层分别采用 2~4 个节点进行训练(根据文献,隐藏层一般采用输入层的一半),试验结果证明 3 个节点时效果最佳。训练过程中采用 MATLAB 提供的贝叶斯正则化方法防止过拟合。

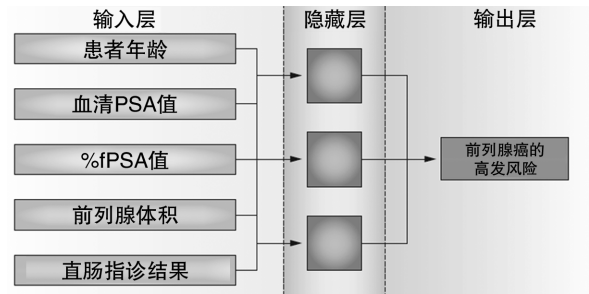


图 1 BP 神经网络模型结构示意图

将调查所得的 310 例前列腺疾病患者信息中的 210 例(前列腺癌组 70 例,非前列腺癌组 140 例)构成训练组,其余 100 例(前列腺癌组 30 例,非前列腺癌组 70 例)构成测试组对 ANN 模型进行交叉验证。以 PSAD≥0.16, f/tPSA 比值≤0.16 为可疑前列腺癌参考指标,分别计算 PSAD 和 f/tPSA 比值的预测价值,并与 BP 神经网络模型相对比。

2 结果

为了检验从训练组中得到的识别函数的可靠程度,对所建立的 8 项输入指标和 1 项输出指标的 BP 人工神经网络模型,用测试组 100 例(前列腺癌组 30 例,非前列腺癌组 70 例)数据进行测试,以检验其识别的可靠性。用受试者工作特征(receiver operating characteristic, ROC)分析衡量 BP 网络预测的准确性。通过 ROC 分析(图 2),ANN 的曲线下面积(AUC)为 0.808。对 100 例测试组样本,ANN 模型预测前列腺癌的结果所示(表 3),预测的特异度为 80.0%,灵敏度为 93.3%。在 ROC 分析中,ANN 得到了较高的 AUC,并且前列腺癌预

测准确率明显高于 PSAD(表 4)和 f/tPSA(表 5),证实了 ANN 在前列腺癌的预测中具有很高的参考价值。

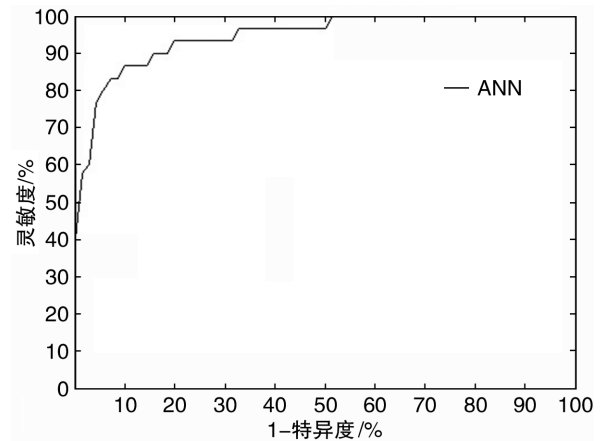


图 2 利用 BP 神经网络模型对前列腺癌进行预测的 ROC 曲线

表 3 两组 BP 神经网络模型预测测试样本的 ROC 分析结果

组别	例数	预测前列腺癌	预测非前列腺癌	准确率 / %
前列腺癌组	30	28	2	93.3
非前列腺癌组	70	14	56	80.0
合计	100	42	58	

表 4 两组 PSAD 预测结果

组别	例数	预测前列腺癌	预测非前列腺癌	准确率 / %
前列腺癌组	30	22	8	73.3
非前列腺癌组	70	7	63	90.0
合计	100	28	72	

表 5 两组 f/tPSA 比值

组别	例数	预测前列腺癌	预测非前列腺癌	准确率 / %
前列腺癌组	30	24	6	80.0
非前列腺癌组	70	9	61	87.1
合计	100	33	67	

3 讨论

ANN 是采用物理操作系统来模仿人脑神经细胞结构和功能的系统。它将生物神经网络的结构及工作方式用数学形式模拟出来,是用机器模拟人脑智能活动的杰出代表^[3]。ANN 具有许多优点,如自适应性、容错性、模式识别性能、外推性、自动抽提功能等。作为一个非线性的人工智能系统,它可以寻找变量之间的复杂相互关系,识别及建模输入变量和输出直接非线性关系等^[4]。通过学习过

程获得知识并利用这些知识进行预测。神经网络的类型可分为前馈神经网络、反馈神经网络、随机神经网络及自组织神经网络等^[5]。一个常见的 ANN 多层神经网络,由输入层、输出层和隐藏层三部分组成。输入层是众多神经元接受大量非线性输入信息,输出层是信息在神经元链接中传输、分析、权衡、形成输出结果,隐藏层是输入层和输出层间众多神经元和链接组成的各个层面。隐藏层可以有多个隐藏的节点,神经元数目不定^[6]。神经网络模型中,最常用的为误差反向传播(back-propagation, BP)神经网络模型,在欧美国家已经广泛应用于医学、药学、环境科学等方面。例如利用人工神经网络极强的智能模式识别能力,辅助阅读病理切片和放射片,以筛查宫颈癌、消化道癌、乳腺癌等,效果令人满意^[7]。近年来,国外研究人员将人工神经网络原理同前列腺癌诊治中的其他指标相结合,应用于提高前列腺癌的诊断效率,取得了很好的效果^[8]。

前列腺癌诊断过程是在患者经过直肠超声检查和直肠指检之后,结合 PSA 水平对患者患病风险进行评估,最终通过前列腺穿刺活检来确诊。在医学上,当血液中 TPSA <4 μg/L 的时候被认为是正常值,4~10 μg/L 是临界值,>10 μg/L 则被认为是高值^[9]。一般认为 PSA 水平越高,发生前列腺癌的风险越高。然而,PSA 水平高时不能认为患者一定患恶性肿瘤,因为前列腺炎、前列腺增生也可以导致 PSA 水平的升高。因此,患者还需要进行直肠指检。如果直肠指检发现异常,即使 PSA 水平处在正常水平也要进行前列腺穿刺活检来确定患者是否患前列腺癌。利用 ANN 来诊断前列腺癌是为了提高诊断准确率,而提高准确率的一个主要目的是减少不必要的前列腺穿刺。因为即使患者可以忍受前列腺穿刺,但是穿刺可能会引发感染、出血等并发症 而且一部分患者的穿刺结果为阴性。文献结果表明 ANN 具有较高的准确率,并在不同程度上避免了不必要穿刺^[10]。在美国,每年有大约一百万人做经直肠超声引导下的前列腺穿刺活检,但只能发现约二十万例的新发前列腺癌病例^[11]。Porter 等^[12]研究了当前广泛采用的 6 种预测模型,有 5 种是计算机学习模型,其中包括 ANN,另一种为 Logistic 回归。发现在预测前列腺穿刺活检的结果方面,ANN 的作用似乎与 Logistic 回归相同。但当采用更多的穿刺前输入变量时,ANN 比其他模型更为优秀,运用它可以减少大量不必要的穿刺活检。

为了改善穿刺活检的预测精确性,Matsui 等^[13]研究了在日本人群中运用 ANN 原理以提高其预测精确性的可能性。他们选取了 228 例血清 PSA 值在 2~10 μg/L 范围的患者。用两个人工

神经网络分析(ANNA), ANNA1 采用患者年龄、总 PSA(tPSA)、f/tPSA 比值、前列腺体积、移行带体积(TZ volume)、PSA 密度(PSAD)、PSA/TZ 作为输入变量, ANNA2 增加了假定面积比(PCAR)、直肠指诊两个指标。通过接收机运行曲线分析比较 ANN 与传统 PSA 方法和 Logistic 回归分析的差别。结论为在 PSA 范围为 4~10 $\mu\text{g/L}$ 时, ANN 模型可以减少不必要的穿刺活检, 并且不会漏诊癌症病例。Remzi 等^[14]利用人工神经网络的原理预测前列腺重复穿刺活检的结果。将 f/tPSA 比值、PSAD、PSA/TZ、前列腺总体积、移行带体积与 ANN 联合运用, 发现在较高灵敏度的情况下, ANN 可明显增加特异度。在 95% 灵敏度条件下, ANN 的特异度为 68%, 而多变量 Logistic 回归、f/tPSA 比值、PSA/TZ、PSAD、TPSA 各自的特异度分别为 54.0%、33.5%、21.4%、14.7%、8.3%。研究认为运用 ANN 可减少约 68% 的不必要的重复穿刺活检。

由于输入指标的选择不同, 缺少前瞻性的研究以及人工神经网络自身存在着一些局限性, 因而在很大程度上限制了 ANN 在世界范围的应用推广。但是随着 ANN 技术不断地应用于临床, 越来越多的数据可以用来训练 ANN, 从而不断地更新和改进 ANN 技术。毫无疑问, ANN 未来在前列腺癌诊断中必将发挥重要作用^[15]。在我国, 前列腺癌的发病率虽然不高, 但呈逐年递增的趋势, 为了体现三级预防的指导思想、结合我国实际情况, 重点就应该放在提高前列腺癌早期诊断效率上。因此, 结合临床诊断数据, 将人工神经网络运用于提高前列腺癌早期诊断精确度, 既可以减少患者痛苦, 又可以节约有限的卫生资源。

[参考文献]

- 1 Center M M, Jemal A, Lortet-Tieulent J, et al. International variation in prostate cancer incidence and mortality rates[J]. *Eur Urol*, 2012, 61(6): 1079—1092.
- 2 孙颖浩. 我国前列腺癌的研究现状[J]. *中华泌尿外科杂志*, 2004, 25(2): 77—80.
- 3 Manjunath G, Jaeger H. Echo state property linked to an input: exploring a fundamental characteristic of recurrent neural networks[J]. *Neural Comput*, 2013, 25(3): 671—696.
- 4 Lawrentschuk N, Lockwood G, Davies P, et al. Predicting prostate biopsy outcome: artificial neural networks and polychotomous regression are equivalent models[J]. *Int Urol Nephrol*, 2011, 43(1): 23—30.
- 5 Krogh A. What are artificial neural networks [J]? *Nat Biotechnol*, 2008, 26(2): 195—197.
- 6 Cammann H, Jung K, Meyer H A, et al. Avoiding pit-falls in applying prediction models, as illustrated by the example of prostate cancer diagnosis[J]. *Clin Chem*, 2011, 57(11): 1490—1498.
- 7 Patten S B. Problems encountered with the use of simulation in an attempt to enhance interpretation of a secondary data source in epidemiologic mental health research[J]. *BMC Res Notes*, 2010, 3: 231.
- 8 Panebianco V, Sciarra A, Marcantonio A, et al. Conventional imaging and multiparametric magnetic resonance (MRI, MRS, DWI, MRP) in the diagnosis of prostate cancer [J]. *Q J Nucl Med Mol Imaging*, 2012, 56(4): 331—342.
- 9 Regnier-Coudert O, McCall J, Lothian R, et al. Machine learning for improved pathological staging of prostate cancer: a performance comparison on a range of classifiers [J]. *Artif Intell Med*, 2012, 55(1): 25—35.
- 10 Stephan C, Vincendeau S, Houlgatte A, et al. Multi-center evaluation of [-2] proprostate-specific antigen and the prostate health index for detecting prostate cancer [J]. *Clin Chem*, 2013, 59(1): 306—314.
- 11 Stephan C, Cammann H, Deger S, et al. Benign prostatic hyperplasia-associated free prostate-specific antigen improves detection of prostate cancer in an artificial neural network [J]. *Urology*, 2009, 74(4): 873—877.
- 12 Porter C R, Crawford E D. Combining artificial neural networks and transrectal ultrasound in the diagnosis of prostate cancer[J]. *Oncology (Williston Park)*, 2003, 17(10): 1395—1399; discussion 1399, 1403—1406. Review.
- 13 Matsui Y, Utsunomiya N, Ichioka K, et al. The use of artificial neural network analysis to improve the predictive accuracy of prostate biopsy in the Japanese population[J]. *Jpn J Clin Oncol*, 2004, 34(10): 602—607.
- 14 Remzi M, Anagnostou T, Ravary V, et al. An artificial neural network to predict the outcome of repeat prostate biopsies[J]. *Urology*, 2003, 62(3): 456—460.
- 15 Meijer R P, Gemen E F, van Onna I E, et al. The value of an artificial neural network in the decision making for prostate biopsies [J]. *World J Urol*, 2009, 27(5): 593—598.

(收稿日期: 2015-02-27)