

基于血清紫外-可见吸收光谱的胆固醇含量研究

朱卫华^{1, 2}, 赵志敏^{1*}, 郭昕¹, 陈会¹

1. 南京航空航天大学理学院, 江苏 南京 210016

2. 河海大学理学院, 江苏 南京 210098

摘要 将血清紫外-可见吸收光谱与神经网络理论结合用于血液中胆固醇含量的分析。研究表明, 在 350 ~ 600 nm 波段范围内, 吸收光谱具有以下几个特征: (1) 不论胆固醇含量如何, 在 416 nm 处均有一个较强的吸收峰; (2) 在波段 450 ~ 500 nm 之间有一肩峰, 其中心波长位于 460 nm 处; (3) 在 578 nm 处还有一个较弱的峰; (4) 不同胆固醇含量的吸收光谱的形状是明显不同的。血清吸收光谱是胆固醇等多种组分光谱叠加的结果, 416 nm 处的吸光度与胆固醇含量之间无明显相关性, 呈现出一种随机的关系, 因此不能由 416 nm 处的吸光度来确定胆固醇的含量是否异常。基于波段 455 ~ 475 nm 范围内血清的吸光度与胆固醇含量有明显的相关性, 文章构建了血清吸收光谱与胆固醇含量之间关系的神经网络模型。

关键词 胆固醇含量; 紫外-可见吸收光谱; 神经网络; 血清

中图分类号: O433.4 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2009)04-1004-04

引言

2002 年世界卫生组织(WHO)公布的人类健康十大危险因素中, 胆固醇过高疾患位居第八位, 据统计由于体内胆固醇过高导致 440 万人死亡, 并可使 18% 的人群诱发心血管疾病。因此对血液中胆固醇的分析研究具有十分重要的意义。生物体尤其是人体, 是复杂的生命系统, 所携带的信息量大而丰富, 其中就隐藏着重要的光谱信息(荧光光谱、红外光谱和紫外光谱)。研究表明, 生物组织出现异常时光谱特征将发生变化。据报道, 瑞士科学家发现, 神经细胞组织的变性是导致早发性痴呆症的主要原因, 通过分析细胞组织变性过程的荧光信息, 可对该病进行早期诊断^[1]。因此, 充分利用生物组织的光谱信息, 对疾病进行有效的分析和诊断, 尤其是血液光谱分析技术是新的具有挑战性的研究课题, 已引起有关学者的关注, 并做出有意义的研究^[2-15]。文献^[6]表明, 在 370 ~ 600 nm 的波段范围内, 正常人血清和高胆固醇血症者血清的吸收光谱有明显的不同, 通过实验也发现, 对某一血清样本增加胆固醇含量, 416 nm 处的吸光度在一定的浓度范围内单调增加。我们通过大量实际血清样本的测试, 发现了 416 nm 处吸光度与胆固醇含量之间呈现一种比较复杂

的随机的关系, 分析其原因认为是实际血清中除了胆固醇含量变化外, 其他生化指标如血糖等均变化, 416 nm 处的吸光度是各组分吸光度的线性组合。基于 455 ~ 475 nm 波段范围内血清的吸收光谱与胆固醇含量有明显的相关性, 本文利用神经网络的可学习性和强大的非线性映射能力的特点, 构建了血清吸收光谱与胆固醇含量之间关系的神经网络模型, 并得到了比较好的预测结果。

1 神经网络结构与方法

神经网络能够模拟人类学习、感知事物的过程, 处理高度非线性问题, 这一优越性使得该技术得以广泛应用于复杂系统的预测计算问题上, 达到传统算法难以实现的非线性映射效果^[16, 17]。

图 1 为神经网络结构中研究得最多、也是应用最广泛的 BP 网络, 是一种前馈神经网络, 式中 I 为输入, O 为输出。这种网络结构中的神经元通常仅接收前一层的神经元的输出, 同一层内的神经元之间没有连接, 研究表明, 3 层前馈神经网络就能以任意精度逼近任何连续函数^[18]。

本文将 BP 神经网络用于血液中胆固醇含量测定。其构造含一个隐层的三层人工神经网络模型, 它们分别对应于系

收稿日期: 2007-10-20, 修订日期: 2008-01-26

基金项目: 国家自然科学基金项目(10172043), 南京航空航天大学科研创新基金项目(1008905704)和河海大学自然科学基金项目(2008430311)资助

作者简介: 朱卫华, 1963 年生, 河海大学理学院教授 e-mail: weihua_zhu@126.com

* 通讯联系人 e-mail: zhaozhimin@nuaa.edu.cn

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

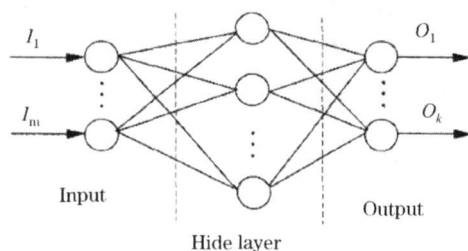


Fig 1 Diagram of a BP neural network

统的输入、过渡和输出。网络输入层由某一波段范围内的吸光度构成, 隐层节点数由试凑法来确定, 输出层含 1 个输出节点, 表示胆固醇含量(用浓度表示)。两层间的传递函数形式不同, 输入层与隐层之间采用 S 型函数, 此处具体函数为双曲正切形式, 隐层与输出层之间则采取线性函数形式。为了提高神经网络的训练效率, 对吸收光谱作了预处理: (1) 归一化处理, 使得归一化后的输入和目标数据均服从正态分布; (2) 对归一化后的样本数据进行主元分析, 消除样本数据中的冗余成分, 起到数据降维的目的。为了提高网络的推广能力和识别能力, 训练中采用“提前中止”的方法, 为此将样本数据划分为训练样本集、验证样本集和测试样本集, 验证样本和测试样本是从原样本数据中均匀选取 1/4 而生成的。

2 实验设备与测试条件

2.1 吸收光谱检测系统

研究所用的吸收光谱检测系统为日本岛津公司生产的 UV3600 紫外-可见近红外分光光度计。其光路图如图 2 所示。

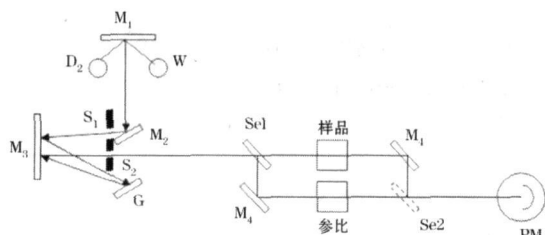


Fig 2 Diagram of absorption spectra testing optical path

采用双光束分光光路, 从光源 D_2 或 W 发射的辐射经反射镜 M_1 、平面反射镜 M_2 、入射狭缝 S_1 到达准直镜 M_3 形成平行光束, 到达光栅 G 被色散后, 经 M_3 到出射狭缝 S_2 , 再经扇形镜 Se_1 和反射镜 M_4 , 交替地进入样品槽和参比槽, 最后又交替地经过 S_2 , 最终到达光电倍增管 PM , 经过转换放大处理后, 显示在计算机上。实验扫描波长范围为 200~900 nm, 缝宽 0.2 nm, 扫描速度为快速, 间隔为 0.5 nm。

2.2 实验样品

实验样品选用人体血液, 选自南京航空航天大学医院。由受试者清晨空腹后采血得到。取 0.2 mL 血清, 加 2 mL 蒸馏水混合。取适量混合样品注入比色皿, 放入光谱仪检测其吸收光谱。

3 实验结果与分析

3.1 吸收光谱

图 3 为利用 UV3600 紫外-可见红外分光光度计获得的部分不同人体血清吸收光谱, 其胆固醇含量各不相同。

由图 3 分析可知, 在 350~600 nm 范围内, 吸收光谱具有以下几个特征: (1) 不论胆固醇含量如何, 在 416 nm 处均有一个较强的吸收峰; (2) 波长为 450~500 nm 处为一肩峰, 其中心波长位于 460 nm 处; (3) 在 578 nm 处还有一个较弱的峰, 但不明显; (4) 不同胆固醇含量血清的吸收光谱的形状是不同的。

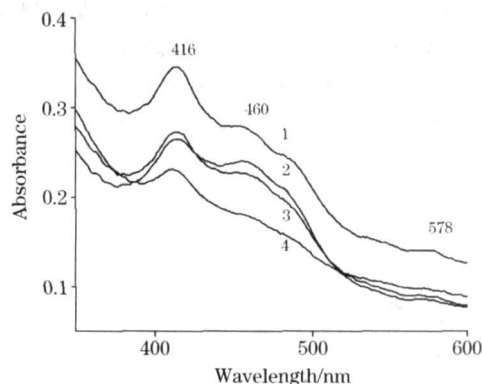


Fig 3 Absorption spectra of part human serum

1: 5.44 mmol·L⁻¹; 2: 5.19 mmol·L⁻¹;
3: 7.06 mmol·L⁻¹; 4: 4.36 mmol·L⁻¹

3.2 416 nm 处吸光度与胆固醇含量的关系

图 4 是 416 nm 处吸光度与胆固醇含量之间的关系图, 图中横坐标表示胆固醇含量, 纵坐标表示吸光度。

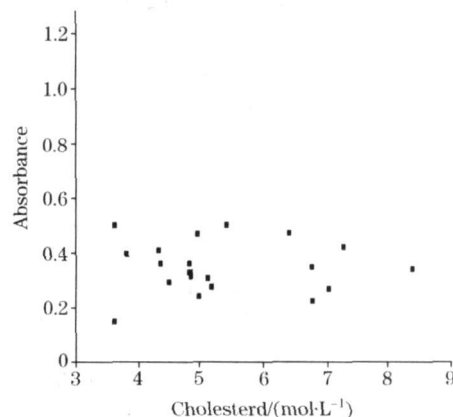


Fig 4 Relationship between absorbance at 416 nm and cholesterol concentration

分析图 4 可知, 两者之间呈现一种随机的关系。其原因是实际血液成分比较复杂, 整个吸收光谱是由多种成分共同作用的结果, 某一波长处的吸光度等于各组分吸光度之和。因此, 虽然某一样本的胆固醇含量比较小, 但可能血脂等其他组分的含量比较大, 从而导致 416 nm 处的吸光度反而会

比胆固醇含量高的样本的吸光度要大。因此, 对实际血液来讲, 仅考虑 416 nm 处峰值的吸光度, 还不能用来确定胆固醇含量是否异常。

3.3 神经网络预测胆固醇含量

如上所述, 血清吸收光谱是多种组分共同作用的结果, 每个波长处包含有许多成分的信息, 因此, 光谱分布中包含了血液中胆固醇含量等信息。本文构建的神经网络模型就是基于某一段吸收光谱来获取血液中胆固醇含量信息。通过计算发现当光谱波段取 455~475 nm 的时候, 可以找到一个比较好的网络预测模型。

图 5 是网络训练误差变化曲线, 图中横坐标表示训练次数, 纵坐标表示均方误差。由图可知, 验证误差、测试误差的变化趋势(图中虚线)基本一致, 说明样本集的划分基本合理(验证样本和测试样本是从原样本数据中均匀选取 1/4 而生成的)。网络训练迭代至第 6 步时提前停止, 这是由于测试误差已经开始变大。由训练误差曲线可见, 训练误差结果是比较满意的。

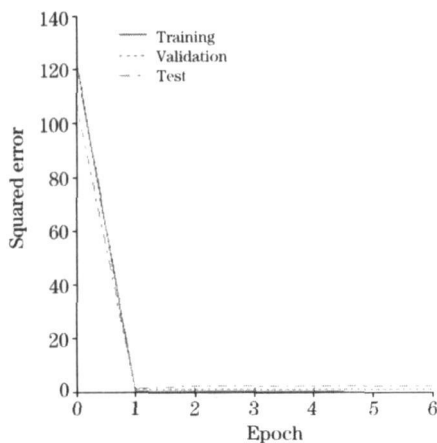


Fig 5 Error training curves

图 6 是网络输出胆固醇含量的回归分析结果, 胆固醇含量预测结果 A 与目标输出 T 的相关系数达到了 0.968, 达到

了比较好的预测效果。这也表明了血液中的胆固醇含量的信息确实隐含在吸收光谱中, 神经网络为光谱信息获取提供了一种较好的手段。

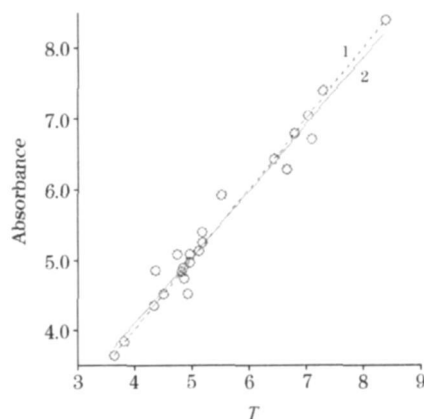


Fig 6 Regression analysis result of network output cholesterol concentration

1: Best linear fit; 2: $A = T$; $r = 0.968$

4 结 论

(1) 血清紫外-可见吸收光谱是胆固醇等多种组分吸收光谱叠加的结果, 每个波长处包含有多种成分的信息, 因此光谱分布中包含了胆固醇含量等信息, 这是光谱分析用于血液检测的基础。

(2) 由于不同血液样本的生化指标均不相同, 因此吸收光谱中的峰值一般受多种血液成分的影响, 416 nm 处峰值与胆固醇含量之间呈现出一种随机的非线性关系, 对于实际应用来说, 不能由 416 nm 峰值来确定胆固醇的含量是否异常。

(3) BP 神经网络用于血液胆固醇的预测中, 基于 455~475 nm 波段的吸光度作为网络的输入信息, 取得了良好的预测效果, 为胆固醇含量的测定提供一种光谱分析方法。

参 考 文 献

- [1] SUN Zhou(孙州编译). Nanjing Daily(南京日报), 2004-02-20.
- [2] WANG Le-xin, ZHAO Zhi-min, YAO Hong-bing, et al(王乐新, 赵志敏, 姚红兵, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2002, 22(6): 980.
- [3] ZHAO Zhi-min, CHEN Yu-ming, YU Xiao-lei, et al(赵志敏, 陈玉明, 俞晓磊, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2003, 23(5): 922.
- [4] ZHAO Zhi-min, GUO Lin-feng, YU Xiao-lei, et al(赵志敏, 郭林峰, 俞晓磊, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(12): 2056.
- [5] ZHU Wei-hua, ZHAO Zhi-min, ZHENG Min, et al(朱卫华, 赵志敏, 郑敏, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(12): 2531.
- [6] LAN Xi-feng, LIU Jian-gang, LIU Ying, et al(兰秀凤, 刘建刚, 刘莹, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(3): 467.
- [7] GU Hua-min, XING Da(谷怀民, 邢达). Acta Photonica Sinica(光子学报), 2003, 32(5): 640.
- [8] YU Chang-qing, LIU Fan-guang, GU Ying, et al(于常青, 刘凡光, 顾瑛, 等). Chin. J. Laser Med. Surg.(中国激光医学杂志), 2004, 13: 41.

- [9] LI Jian-dong, DING Jian-hua, WU Min, et al(李建东, 丁建华, 吴 敏, 等). Applied Laser(应用激光), 2004, 24(2): 112.
- [10] MENG Ji-wu, ZHENG Rong-er, ZHANG Yun-dong(孟继武, 郑荣儿, 掌蕴东). Chin. J. Laser Med. Surg.(中国激光医学杂志), 2003, 12(1): 43.
- [11] CHEN Rong, LI Bu-hong, CHEN Jin-kai, et al(陈 荣, 李步洪, 陈金铠, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2003, 23(4): 748.
- [12] GE Xiang-hong, ZHAO Yuan-li, ZHANG Feng-jiu, et al(葛向红, 赵元黎, 张凤秋, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(7): 841.
- [13] SHI Xiao-feng, MA Jun, MAO Wei-zheng, et al(史晓凤, 马 君, 毛伟征, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(2): 295.
- [14] GAO Shu-mei, LIU Ying, LAN Xiu-feng, et al(高淑梅, 刘 莹, 兰秀凤, 等). Chinese Journal of Lasers(中国激光), 2004, 31(7): 893.
- [15] QIN Shao-ping, DING Jian-hua, LI Jian-dong(秦少平, 丁建华, 李建东). Journal of Optoelectronics • Laser(光电子 • 激光), 2006, 17(5): 634.
- [16] WU Gen-hua, HE Chi-yang, CHEN Rong(吴根华, 何池洋, 陈 荣). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2002, 22(5): 813.
- [17] XU Fang-wei, CAI Kun-bao(徐方维, 蔡坤宝). Journal of Chongqing University(重庆大学学报), 2004, 27(8): 35.
- [18] MENG JOO E R. IEEE Transaction on Neural Network, 2002, 13(3): 15.

Study of Cholesterol Concentration Based on Serum UV-Visible Absorption Spectrum

ZHU Wei-hua^{1,2}, ZHAO Zhi-min^{1*}, GUO Xin¹, CHEN Hui¹

1. College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

2. College of Science, Hohai University, Nanjing 210098, China

Abstract In the present paper, UV-visible absorption spectrum and neural network theory were used for the analysis of cholesterol concentration. Experimental investigation shows that the absorption spectrum has the following characteristics in the wave band of 350-600 nm: (1) There is a stronger absorption peak at 416 nm for the test sample with different cholesterol concentration; (2) There is a shoulder peak between 450 and 500 nm, whose central wavelength is 460 nm; (3) There is a weaker peak at 578 nm; (4) Absorption spectrum's shape of different cholesterol concentration is different obviously. The absorption spectrum of serum is the synthesis result of cholesterol and other components (such as sugar), and the information is contained at each wavelength. There is no significant correlation between absorbance and cholesterol content at 416 nm, showing a random relation, so whether cholesterol content is abnormal is not determined by the absorbance peak at 416 nm. Based on the evident correlation between serum absorption spectrum and cholesterol concentration in the wave band of 455-475 nm, a neural network model was built to predict the cholesterol concentration. The correlation coefficient between predicted cholesterol content output A and objectives T reaches 0.968, which can be regarded as better prediction, and it provides a spectral test method of cholesterol concentration.

Keywords Cholesterol concentration; UV-visible absorption spectrum; Neural network(NN); Serum

(Received Oct. 20, 2007; accepted Jan. 26, 2008)

* Corresponding author