文章编号:1000-5404(2013)15-1613-03

论著

阿尔茨海默病和轻度认知损伤诊断的建模分析

罗万春 (400038 重庆,第三军医大学生物医学工程学院数学与生物数学教研室)

[摘要] **目的** 建立阿尔茨海默病和轻度认知损伤的诊断模型。**方法** 用ROC曲线进行指标筛选,再用判别分析进行诊断模拟,最后选出正确率最高的判别函数。**结果** ROC曲线筛选出4项指标。在包括正常人、阿尔茨海默患者以及轻度认知损伤患者的研究对象中,经过100次随机抽样诊断模拟,线性判别分析诊断的平均正确率达71.42%,非线性判别的平均正确率达74.75%,最高正确率为82.05%,并得到了相应的判别函数。**结论** 非线性判别函数可有效诊断阿尔茨海默病和轻度认知损伤。

[关键词] 阿尔茨海默病;轻度认知损伤;判别分析;ROC曲线
[中图法分类号] R311; R745.7 [文献标志码] A

Modeling for diagnosis of Alzheimer's disease and mild cognitive impairment

Luo Wanchun (Department of Mathematics and Biomathematics, School of Biomedical Engineering, Third Military Medical University, Chongqing, 400038, China)

[**Abstract**] **Objective** To establish a diagnostic model of Alzheimer's disease and mild cognitive impairment. **Methods** Receiver operator curve (ROC) curve was used to screen the indices, discriminant analysis was carried out for diagnosis simulation, and the discriminant function with the highest accuracy was selected. **Results** Four indices were selected by ROC curve. After 100 times of random sampling for diagnosis simulation in the objects including normal people, patients with Alzheimer's disease and patients with mild cognitive impairment, the average diagnosis accuracy of linear discriminant analysis was 71.42% and that of nonlinear discriminant analysis was 74.75%. The highest accuracy was 82.05%, and the corresponding discriminant function was obtained. **Conclusion** Nonlinear discriminant analysis can be used to diagnose Alzheimer's disease and mild cognitive impairment.

[**Key words**] Alzheimer's disease; mild congitive impairment; discrimination analysis; Receiver operator curve

Supported by the Natural Science Foundation of Chongqing(CSTC2013jcyjA10041). Corresponding author: Luo Wanchun, E-mail: wanchunluo@163.com

阿尔茨海默病(Alzheimer's disease, AD)是发生在老年期及老年前期的一种原发性退行性脑病,指的是一种持续性高级神经功能活动障碍,即在没有意识障碍的状态下,记忆、思维、分析判断、视空间辨认、情绪等方面的障碍。AD是由德国精神病医生和神经病理学家Alois Alzheimer于1906年首次描述并被由其名字命名[1],现已成为仅次于心血管病、脑血管病和癌症等,导致老年人死亡的"第四大杀手"。2010年AD患者已超过3 500万,预计到2050年,全世界每85人中将有1人患AD,患病人数将会超过1亿[2-3]。AD对患者、家庭、社会都带来了巨大的危害。不但危害人类的生命安全[4-5],其所需的直接和间接医疗费用还十分巨大,在欧洲每年超过1 200亿美元,美国AD协会公布的数据显示,2011年美国每年的此项开支已超过1 830亿美元,预计到2050年将达到1.1万亿美元[6]。因此,WHO已将AD定为21世纪五大重点疾病之一。

虽然AD正确的早期诊断对患者的评估、控制、护理和未来规划有重要价值[7-11],但由于轻度认知损伤(mild cognitive impairment, MCI)患者和早期的AD患者在临床上十分相像,使得AD的早期诊断变得十分困难[12]。本研究拟从数学模型的角度,建立能有效诊断AD和MCI的模型,旨在为AD和MCI的临床诊断提供一种辅助诊断的方法。

1 资料与方法

1.1 临床资料

数据源于The LONI Image Data Archive(https://ida.loni.ucla.edu/login.jsp)。选出所有指标数据齐全的样本391个,其中正常人101例,MCI患者200例,AD患者90例。数据矩阵的

[基金项目] 重庆市自然科学基金(CSTC2013jcyjA10041)
[通信作者] 罗万春, E-mail: wanchunluo@163.com
[优先出版] http://www.cnki.net/kcms/detail/51.1095.R.20130530.0838.011.html(2013-05-30)

维度为391×31,其中每个样本的指标30项,包括年龄、受教育时间、简易智力状况检查量表得分(mini-mental state examination score, MMSE)、ADAS-cog修正值(ADAS-cog modified measure, TOTALMOD)、大脑容积、心室容积、心室与大脑容积之比、左海马体积、右海马体积、海马体积与脑容量之比的1 000倍[hippocampus volume (left + right) divided by brain volume × 1 000, HV]、左脑的PRT值(regional PRT value of left hemisphere)、左右半脑的TMP均值、左脑的FRT值、左右半脑的Oct均值、左右半脑的PCING均值、左右半脑的ACING均值、左右半脑的FRT均值、左右半脑的PRT均值、左右半脑的SMC均值、左右半脑的VC均值、左右半脑的CAD均值、左右半脑的CBL均值、左右半脑的VER均值、左右半脑的HEMI均值、左右半脑的PNS均值、左右半脑的CTX均值、左右半脑的PTM均值、左右半脑的PARH均值、左右半脑的AMG均值和左右半脑的THL均值。每个指标分别记为$x1$、$x2$、$x3\cdots x30$,最后1列为患者类型,分别赋值为正常人(1)、轻度认知损伤(2)、AD患者(3)。

1.2 方法

用ROC曲线分析每项指标对AD的诊断敏感度,选择对诊断敏感的指标。再将所有样本随机分为训练样本和检测样本,其中训练样本集中包含正常人51例、MCI 100例、AD 45例,其余195例均为检测样本,检测样本和训练样本相互独立。最后用判别分析(discriminant analysis)做诊断模拟,并筛选出诊断效果最好的判别函数。

1.3 软件

ROC曲线在SPSS 13.0中完成,其余数据处理用数学软件MATLAB R2009a编程完成。

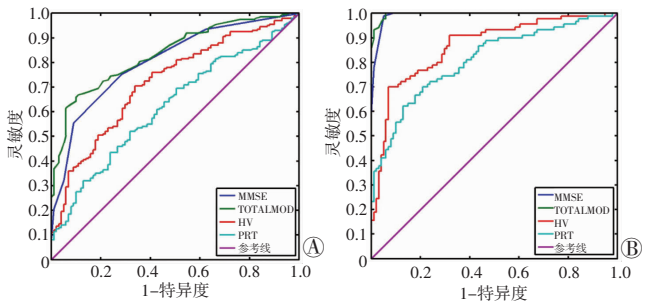
2 结果

2.1 ROC曲线分析

对每项指标作ROC曲线分析,发现简易智力状况检查量表得分、ADAS-cog修正值、海马体积与脑容量之比的1 000倍、左脑的PRT值对于MCI和AD的ROC曲线面积最大(图1)。因此,可以认为这些指标在AD和MCI的识别中有较强的作用,从而筛选这4项指标作为诊断模型的自变量。表1给出了3类人群4项指标的均值和标准差。

表1 正常人、MCI患者和AD患者的4项指标值($\bar{x}\pm s$)

研究对象	n	MMSE	TOTALMOD	HV	PRT
正常人	101	28.94±1.12	10.43±4.24	6.74±0.73	1.07±0.04
MCI患者	200	27.10± 1.68	17.81±6.06	6.12±0.88	1.04±0.06
AD患者	90	23.53±2.06	29.51±7.79	5.56±0.79	0.99±0.08



A:MCI患者;B:AD患者

图1 MCI患者和AD患者4项指标的ROC曲线分析

2.2 诊断模拟

筛选出的4项指标均为连续变量,4项指标完全独立且服从正态分布,符合使用判别分析的条件。按照1.2中训练样本集和检测样本集的组成方法作100次线性判别分析,检测样本的平均、最大、最小正确率分别为71.42%、76.41%和65.13%。进一步考虑增加非线性项来提高诊断正确率,通过多次模拟实验,发现当增加x_3x_4、x_3x_{10}和x_3x_{11}的非线性判别诊断正确率有较大提高,检测样本的平均、最大、最小正确率分别为74.75%、82.05%和68.72%(图2)。

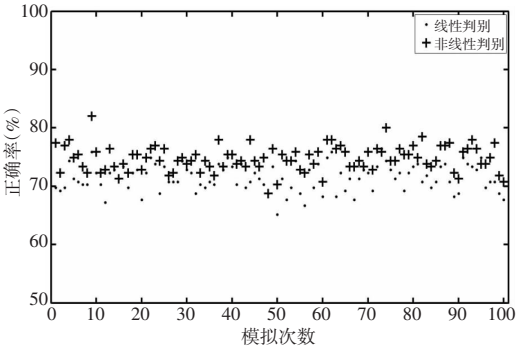


图2 线性和非线性判别分析的正确率

为了进一步比较非线性判别和线性判别正确率的差异,将100次模拟获得的线性和非线性判别的正确率作配对样本t检验,得到$t=-16.91$,$P=5.96\times10^{-31}$,可知非线性判别的正确率显著高于线性判别。在100次的非线性判别模拟中,正常人、MCI患者和AD患者的平均判别正确例数和正确率见表2。可知,当正常人、MCI和AD完全混杂在一起时,很大部分的MCI容易被误判为正常人或AD,部分AD也容易被诊断为MCI。但如果已经排除一类(MCI或AD)的可能性时,即只有两类人群的判别,则可达到较高的诊断正确率。从表3可以看出,如果3种人群能排除其中1种,则能非常好地判别其类型,尤其是当患者被排除MCI后,对于正常人和AD患者的100次模拟诊断的平均正确率达到了99.42%。当将正常人和MCI患者合并为一类(非AD)时,即只考虑是否患AD的情况下,其平均正确率达到了90.30%,诊断效果非常好。

表2 对正常人、MCI患者和AD患者非线性判别的正确效果

[平均例数(%)]

研究对象	判别为正常人	判别为MCI患者	判别为AD患者
正常人	41.07(82.14)	8.93(17.86)	0.00(0.00)
MCI患者	21.96(21.96)	70.88(70.88)	7.16(7.16)
AD患者	0.00(0.00)	11.19(24.87)	33.81(75.13)

表3 100次模拟两类的非线性判别诊断正确率(%)

类别	平均	最大	最小
非AD患者对AD患者	90.30	93.33	86.15
正常人对MCI患者	79.67	85.33	73.33
正常人对AD患者	99.42	100.00	95.79
MCI患者对AD患者	86.66	90.34	80.69

2.3 判别函数

既然非线性判别的诊断效果更好,则可筛选出判别正确率最高的情形(正确率82.05%),并得到其判别函数:

$$f_1=-6.15x_3x_4-80.31x_3x_{10}-1\,370.22x_3x_{11}+2\,041.63x_3+164.91x_4+2\,205.3x_{10}+36\,126.21x_{11}-27\,183.37$$

$$f_2=-6.11x_3x_4-80.58x_3x_{10}-1\,371.57x_3x_{11}+2\,043.26x_3+$$

$164.01x_4 + 2\,211.77x_{10} + 36\,158.96x_{11} - 27\,220.13$

$f_3 = -6.13x_3x_4 - 80.48x_3x_{10} - 1\,369.24x_3x_{11} + 2\,039.58x_3 + 164.57x_4 + 2\,208.92x_{10} + 36\,091.02x_{11} - 27\,120.37$

一旦获得了某人的这4项指标，只要代入上述3个方程比较大小即可。如果f_1最大，则为正常人；f_2最大，则为MCI患者；f_3最大，则为AD患者。

如果某患者能排除MCI，则可用下述函数判别：

$g_1 = -4.81x_3x_4 - 56.29x_3x_{10} - 1\,311.81x_3x_{11} + 1\,799.91x_3 + 127.53x_4 + 1\,537.08x_{10} + 33\,580.45x_{11} - 23\,346.77$

$g_2 = -4.7x_3x_4 - 56.55x_3x_{10} - 1\,315.48x_3x_{11} + 1\,800.84x_3 + 125.25x_4 + 1\,541.66x_{10} + 33\,656.14x_{11} - 23\,345.84$

将其4项指标的值代入上面的函数。若g_1更大，则为正常人；若g_2更大，则为AD患者。该判别函数对95个检测样本的诊断正确率为100%。

如果某患者能排除AD，则可用下述函数判别：

$h_1 = -20.11x_3x_4 - 150.13x_3x_{10} - 4\,300.52x_3x_{11} + 5\,759.24x_3 + 561.12x_4 + 4\,194.19x_{10} + 118\,173.81x_{11} - 79\,408.93$

$h_2 = -20.06x_3x_4 - 150.33x_3x_{10} - 4\,299.37x_3x_{11} + 5\,757.9x_3 + 560.08x_4 + 4\,199.5x_{10} + 118\,141.99x_{11} - 79\,372.91$

将其4项指标的值代入上面的函数。若h_1更大，则为正常人；若h_2更大，则为MCI患者。该判别函数对150个检测样本的诊断正确率为85.33%。

如果某患者能确定为MCI和AD之一时，则可用下述函数判别：

$l_1 = -9.5x_3x_4 - 106.45x_3x_{10} - 1\,857.97x_3x_{11} + 2\,759.36x_3 + 241.43x_4 + 2\,825.91x_{10} + 47\,810.48x_{11} - 35\,717.85$

$l_2 = -9.52x_3x_4 - 106.37x_3x_{10} - 1\,855.63x_3x_{11} + 2\,756.07x_3 + 241.91x_4 + 2\,823.94x_{10} + 47\,743.54x_{11} - 35\,630.65$

将其4项指标的值代入上面的函数。若l_1更大，则为MCI患者；若l_2更大，则为AD患者。该判别函数对145个检测样本的诊断正确率为90.34%。

3 讨论

本研究的分析对象包括正常人、MCI患者和AD患者，这更符合临床实际情况，患者前来就诊时不太可能马上可以排除某种疾病，尤其是在MCI的干扰下。因此，尽管在本研究中得到的非线性判别函数对独立于训练样本的检测样本的诊断正确率只有82.05%(160/195)，但对临床诊断极为困难的AD诊断方法提供了有益的探索，在临床上也可作为辅助诊断的参考。

研究[13]只考虑了正常人和AD患者，没有包含MCI患者，正确率为97.5%。如果仅仅考虑正常人和AD患者，本研究在100次模拟中的平均正确率达到了99.42%，最高可达100%，并可按此方法得到相应的判别函数。不但如此，我们还将只有两类样本的各种情形的最高正确率的判别函数求了出来，对于能排除其中任一种情形再判别为另外两类之一时可作辅助诊断。

MCI患者和正常人是最难区别诊断的，尤其是MCI患者的早期诊断，较大地影响了包含正常人、MCI患者和AD患者的诊断正确率。在临床诊断中，MMSE是诊断AD和MCI最重要的指标，本研究也得到了相似的结论。但分析ROC曲线发现，TOTALMOD更为敏感，曲线下面积更大。MMSE对AD患者和MCI患者非常敏感。虽然AD患者的MMSE均值低于MCI患者，但是许多MCI患者的MMSE值低于AD患者MMSE的最大值，部分AD患者的MMSE值又高于MCI患者，有许多交叉部分，因此，仅依靠MMSE不能很好地诊断早期AD和MCI。TOTALMOD的分析可能是以后AD和MCI诊断的一个重要研究方向，在ADNI的数据库中它已经作为了一项重要指标。海马与学习记忆，特别是近期记忆功能密切相关，是AD患者最易受累的脑区[14]，由ROC曲线筛选出的指标HV和海马容量有关，在AD患者和MCI患者的诊断中有较重要的作用。

参考文献：

[1] Berchtold N C, Cotman C W. Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s[J]. Neurobiol Aging, 1998, 19(3): 173-189.
[2] Brookmeyer R, Johnson E, Ziegler-Graham K, *et al*. Forecasting the global burden of Alzheimer's disease[J]. Alzheimers Dement, 2007, 3(3): 186-191.
[3] Querfurth H W, LaFerla F M. Alzheimer's Disease[J]. N Engl J Med, 2010, 362(4): 329-344.
[4] Molsa P K, Marttila R J, Rinne U K. Long-term survival and predictors of mortality in Alzheimer's disease and multi-infarct dementia[J]. Acta Neurol Scand, 1995, 91(3): 159-164.
[5] Dodge H H, Shen C, Pandav R, *et al*. Functional transitions and active life expectancy associated with Alzheimer disease[J]. Arch Neurol, 2003, 60(2): 253-259.
[6] Alzheimer's Association, Thies W, Bleiler L. 2011 Alzheimer's Disease Facts and Figures[J]. Alzheimer's Dement, 2011, 7(2): 208-244.
[7] Okonkwo O C, Griffith H R, Copeland J N, *et al*. Medical decision-making capacity in mild cognitive impairment: a 3-year longitudinal study[J]. Neurology, 2008, 71(19): 1474-1480.
[8] Carpenter B D, Xiong C, Porensky E K, *et al*. Reaction to a dementia diagnosis in individuals with Alzheimer's disease and mild cognitive impairment[J]. J Am Geriatr Soc, 2008, 56(3): 405-412.
[9] Bradford A, Kunik M E, Schulz P, *et al*. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors[J]. Alzheimer Dis Assoc Disord, 2009, 23(4): 306-314.
[10] Boise L, Neal M B, Kaye J. Dementia assessment in primary care: results from a study in three managed care systems[J]. J Gerontol A Biol Sci Med Sci, 2004, 59(6): M621-M626.
[11] Triebel K L, Martin R, Griffith H R, *et al*. Declining financial capacity in mild cognitive impairment: A 1-year longitudinal study[J]. Neurology, 2009, 73(12): 928-934.
[12] Mayeux R. Clinical practice. Early Alzheimer's disease[J]. N Engl J Med, 2010, 362(23): 2194-2201.
[13] Ramirez J, Chaves R, Gorriz J M, *et al*. Functional brain image classification techniques for early alzheimer disease diagnosis[J]. Lecture Notes in Computer Science, 2009, 5602: 150-157.
[14] 高长越，方传勤，许志强，等. 阻断内源性P75NTR对阿尔茨海默病小鼠海马神经干细胞增殖与学习记忆的影响[J]. 第三军医大学学报，2010，32(11)：1152-1154.

（**收稿**：2013-04-18；**修回**：2013-05-23）
（编辑 王小寒）