

多元分析在肺吸虫病诊断的应用^①

詹希美^{1,②} 邵向云² 徐小川¹ 李桂云¹ 路群¹ 李荣佳³

(1 中山医科大学寄生虫学教研室;广州,510089 2 浙江省温州市第五人民医院

3 广东省物资学校电脑室)

提 要 对浙江省温州市的 72 例肺吸虫病患者及 18 例正常人按 11 项检验结果作聚类分析及逐步判别分析;根据谱系分枝图及判别函数的提示,这些患者可分成 3 种类型,并与临床上的肺型、肝型、亚临床型对应;结合临床情况提出:诊断肺吸虫病只需 6 项指标(肺吸虫抗原皮试、血白细胞计数、血中嗜酸性细胞比例、血沉、 γ 球蛋白比例及 ELISA),而肝肿大、谷丙转氨酶、间接血凝试验等项可以省去而不影响判别效果。所得 2 组函数可能有助于诊断及分型。

关键词 肺吸虫病;聚类分析;判别分析

中图分类号 R383.24

肺吸虫病是由其病原肺吸虫(也称并殖吸虫)引起的有多个器官损害的寄生虫病。肺吸虫的种类较多,目前世界报道的已有 40 多种。但不是每一种都能使人致病,能致病的种也不定在人体发育成熟并把卵排出体外。因此,较难从痰或粪中找到虫卵。临床上常需靠免疫、生化检验等多项指标进行综合分析才能作出诊断。但对于如何分型则有不同主张,有的主张分为肺型及肺外型两大类,有的从主要侵犯部位将其分为肺型、腹型、肝型、脑型、皮下型等。对于每项检验结果的诊断价值如何,则很少从量上进行评价。作者应用多元分析的聚类分析及逐步判别方法,对温州市的 72 例临床上诊断为并殖病患者进行分析,报道如下。

1 材料与方 法

1.1 资料来源 1979~1991 年浙江省永嘉县人民医院收治了 200 多例“肺吸虫病”患者,从中选出 11 项指标齐全者 72 例。其中的肺型指征为肺部 X 线检查有明显改变,且在

痰中查到虫卵者。肝型主要指征为肝脏有明显损害表现而肺部改变不明显者。亚临床型则为多项免疫检查阳性而临床表现不明显者。这 72 例中属肺型者 33 例,肝型者 20 例,亚临床型者 19 例。11 项指标包括:肝肿大(单位 cm,在判别函数中用 x_1 表示)、血白细胞计数(万/mm³, x_2)、嗜酸性粒细胞在白细胞中比例(% , x_3)、血红细胞沉降率(mm/h, x_4)、谷丙转氨酶(U, x_5)、麝香草酚混浊度(U, x_6)、硫酸锌混浊度(U, x_7)、 γ 球蛋白(电泳蛋白,单位, x_8)、皮试(皮丘直径,cm, x_9)、间接血凝试验(x_{10})、ELISA(吸光度,OD 值, x_{11})。选取非疫区的 18 名健康人作为对照组。

1.2 方 法 对上述 90 人根据 11 项指标进行聚类分析(相关系数法)及逐步判别(Bayes 准则^[1])。

2 结 果

2.1 聚类分析 图 1 为 72 例患者及 18 名

① 本校科研基金资助项目;

② 第一作者,49 岁,男,教授

正常人根据 11 项指标计算结果所描绘出来的谱系分枝图。

从图 1 中可以看出,这 90 人被分成两大类,第一类主要为肺型患者,第二类包括肝型、亚临床型及正常组。在后一类中,正常组与亚临床组相似系数只有 0.31,这一类的组

合在 -0.01 处,可见这 3 组的相似系数相当小。

谱系分枝图的归类情况与原先的临床分型比较如表 1。

表 1 聚类图所示分型与临床分型比较¹⁾

临床分型	N ²⁾	n ³⁾	不符合而被判为			
			肺型	肝型	亚临床型	正常组
肺型	33	28		4		1
肝型	20	14	3		2	1
亚临床型	19	11	3	3		2
正常	18	17		1		

1)单位为例; 2)N表示总例数; 3)n表示符合例数

从表 1 中可以看出,分枝图上的 4 个组对应临床的 4 个类型。肺型的符合率为 85%,肝型符合率为 70%,正常组符合率为 94%,亚临床型符合率为 58%。

2.2 判别分析 将这 90 例按临床标准所分的 4 型进行逐步判别分析。选取 F 值(F 界

值)为 0、1、2、3、4、5、6 进行运算。 F 值越小,被选进函数的指标越多。反之, F 值越大,被选进的指标越小。表 2 是不同 F 值时被选取指标数及判别结果与原先临床分型比较的简表。

表 2 取不同 F 界值判别结果

F	选入函数指标 ¹⁾	判别前后相符				符合率(%)
		肺型 ²⁾	肝型 ²⁾	亚临床型 ²⁾	正常 ²⁾	
0	11	30	17	18	18	92
1	8	30	17	18	17	91
2	6	31	17	16	17	90
3	6	31	17	16	17	90
4	4	31	17	17	17	91
5	4	31	17	17	17	91
6	3	25	17	16	17	84

1)单位为项;2)单位为例

从表 2 可知,当 F 值为 0 时,11 个指标全部进入函数,判别前后符合率为 92%。当 F 值为 2、3 时,被选入函数的指标只有 6 个,符合率为 90%。被选入的指标是:血中白细胞计数、嗜酸性粒细胞比例、血沉、 γ 球蛋白比例、皮试和 ELISA。当 F 值为 4、5 时,被选入函数的指标又少了血中白细胞计数及 ELISA 两项,符合率为 91%。当 F 为 6 时,

又再剔除血沉一项,函数中只剩下嗜酸性粒细胞、 γ 球蛋白及皮试等 3 项,这时的符合率为 84%。

下面是取 F 值为 3 的一组判别函数

$$Y_1 = 21.4794 - 2.8606x_2 + 0.2682x_3 + 0.721x_4 + 0.8646x_8 + 6.0428x_9 + 6.9756x_{11}$$

$$Y_2 = -59.156 - 4.3231x_2 + 0.6392x_3 + 0.0871x_4 + 1.665x_8 + 7.6808x_9 + 6.6423x_{11}$$

$$Y_3 = -12.9667 - 1.4138x_2 + 0.0917x_3 + 0.0057x_4$$

$$+0.8455x_8+4.168x_9+4.1264x_{11}$$

$$Y_4=-9.556-1.2203x_2+0.032x_3+0.0028x_4$$

$$+0.8348x_8+2.2909x_9+2.1369x_{11}$$

下面是取 F 值为 5 的一组判别函数

$$Y_1=-16.9461+0.1777x_3+0.0668x_4+0.6853x_8$$

$$+5.3342x_9$$

$$Y_2=-52.2645+0.475x_3+0.077x_4+1.4141x_8$$

$$+3.7944x_9$$

$$Y_3=-11.595+0.0518x_3+0.0034x_4+0.7533x_8$$

$$+3.7944x_9$$

$$Y_4=-8.9524-0.0125x_3+0.0001x_4+0.7627x_8$$

$$+2.0174x_9$$

F 为临界值, Y_1 代表肺型, Y_2 代表肝型, Y_3 代表亚临床型, Y_4 代表正常组。

2.3 Jack-knife 法检验 以 Jack-knife 法作进一步检验。具体做法是在运算之前将某一例从总体样本中“切”下,以剩下的 89 例仍然取 F 值为 3,以频数作先验概率得出函

数,以这一组函数对未被包含进总体的那一例进行判别。如此循环,对本组的 90 例作了 90 次检验。然后,分别求出这 90 次运算所得的 90 个函数的自变量系数均数,以这些均数构成新的函数组,再对每一个样本进行检验。

下面是 Jack-knife 法所得系数平均值构成的一组判别函数。

$$Y_1=-20.368-2.848x_2+0.187x_3+0.0676x_4$$

$$+0.7828x_8+5.7779x_9+7.0507x_{11}$$

$$Y_2=-56.6829-4.4358x_2+0.4896x_3+0.097x_4$$

$$+1.5432x_8+7.2705x_9+6.7704x_{11}$$

$$Y_3=-12.7001-1.4382x_2+0.0495x_3+0.009x_4$$

$$+0.8057x_8+4.0427x_9+4.1687x_{11}$$

$$Y_4=-9.3592-1.2395x_2-0.0025x_3+0.0007x_4$$

$$+0.8007x_8+2.099x_9+2.1574x_{11}$$

表 3 是 Jack-knife 法与平均系数法的判别结果。

表 3 两种检验方法的判别结果^{1),2)}

临床分型	判 别 后				总 计
	肺 型	肝 型	亚临床型	正 常	
肺 型	31(27)		2(6)		33(33)
肝 型	3(3)	17(17)			20(20)
亚临床型		2(1)	15(18)	2	19(19)
正 常			1(1)	17 (17)	18(18)

1)表中括号外数字为 Jack-knife 法的判别结果,括号内数字为平均系数法的判别结果; 2)单位为例

从表 3 可以看出,Jack-knife 法判别结果与临床诊断符合率为 88.89%,平均系数法与临床诊断符合率为 87.78%。与表 2 中 F 值为 3 时的判别符合率为 90%相比,差异不明显。

3 讨 论

3.1 肺吸虫病诊断标准

在临床上能找到肺吸虫病病原体,诊断较易确定,如找不到病原体则要靠其它项目的检查。用哪些项目,众说不一,有些学者提出只要有吃溪蟹或喇蛄史,且有两项免疫检

查阳性即可诊断。有人则提出从多种资料进行分析^[2,3]。在本项研究中,判别分析在取 F 值为 2 或 3 时,被选进函数的指标有 6 个。而取 F 值为 4 或 5 时,被选进函数的指标只有 4 个,即从上述的 6 个指标中又剔除血白细胞计数及 ELISA 两项。尽管后者少了两项而符合率也没甚明显区别,但考虑到本次分析并没有将其它吸虫病病人列入,而这两项指标在鉴别诊断中又起着比较重要的作用,故认为仍应保留它们。也即是说,取 F 值为 2 或 3 时的函数可能更符合实际。综合不同 F 值的判别函数的筛选情况及临床实际应用,认为皮试阳性,血中白细胞增加、嗜酸性粒细

胞比例升高、血沉增快、ELISA 的吸光度(OD 值)大于 0.3、 γ 球蛋白比例增加或倒置,即可诊断为肺吸虫病,诊断正确率在 90%以上。6 项指标可以代替 11 项指标,判别效果一样,可节省时间与物力。

值得一提的是被剔除的 5 项指标中的麝香草酚混浊度与硫酸锌混浊度于两年前已被卫生部下令淘汰,但本文的实验数据均收集于其前。为保持资料完整性,在运算时仍将其计算在内。逐步判别结果将其剔除,更说明在今后的肺吸虫病诊断中没有检测此两项的必要。

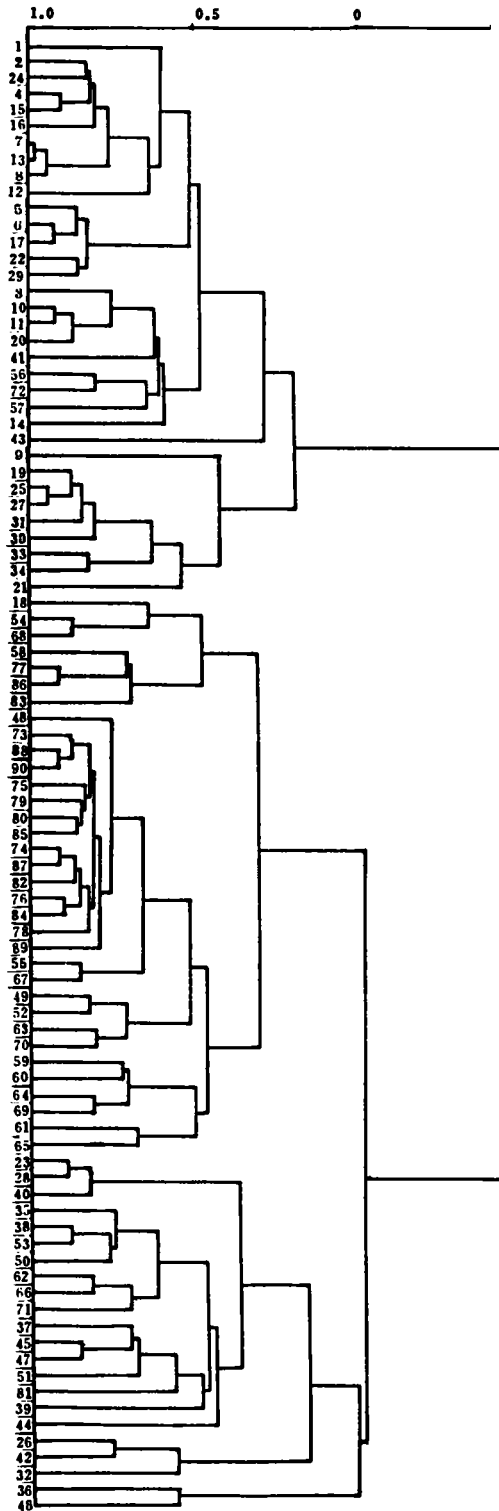
3.2 肺吸虫分型

关于肺吸虫病的临床分型,本文根据聚类分析,认为这些患者可分为 3 种类型,与邵向云^[4]提出的肺型、肝型、亚临床型相对应。在谱系分枝图上,肺型、肝型患者较集中,而亚临床型较分散,与原先临床分型符合率只有 58%。换句话说,临床上划为亚临床型者,可能有 42%属于肺型、肝型甚或正常者。究其原因,可能与这些患者的症状较轻且多样化,各项指标变化又较大有关。

在本次研究中,同时还作了主成分分析。从二维排序图中,也可见到这 90 例被分成 4 个组,与聚类分析基本相同。但在图上,很多点密集在一起,很难标示出来,故本文未将其列入。

3.3 判别函数的应用

F 值为 3 与 5 时所得的两组函数,每组各包含 4 个,分别代表肺型(Y_1)、肝型(Y_2)、亚临床型(Y_3)及正常组(Y_4)。若如要了解新就诊者是否是肺吸病患者?属哪一型?只要将检查结果代入函数,看哪一个函数的 Y 值大,该就诊者就属于那一型。至于用哪一组函数,则根据所用的检验项目而定。因这些函数所要求的检验项目不多,准确性较高。经过检验,3 种方法的符合率基本一致,说明这些函数较可靠,所以在临床诊断上可能具有一定意义。



附图 90 例 11 项指标谱系分枝图
1~33 肺型;34~53 肝型;54~72 亚临床型;
73~90 正常型

(本研究先后得到中山医科大学卫生学院洪楠副教授、方积乾教授的热心帮助和指导,深表谢意)

参 考 文 献

- 1 金丕焕,黄小兰,史秉璋,等. 医用统计程序集. 上海:上海科学技术出版社,1986. 225
- 2 陈翠娥,张悟澄. 肺吸虫病的研究与防治. 长沙:湖南科学出版社,1979. 56~57
- 3 曹维霖. 肺吸虫病的诊断和治疗. 中级医刊,1980,(7): 1
- 4 邵向云. 建立“肝型并殖吸虫病”临床类型的探讨. 浙江医学,1992, 14(3): 114

(1993-11-12 收稿 1994-06-30 修回)

APPLICATION OF MULTIVARIATE ANALYSIS TO DIAGNOSIS OF PARAGONIMIASIS

Zhan Ximei¹ Shao Xiangyun² Xu Xiaochuan¹
Li Guiyun¹ Lu Qun¹ Li Rongjia³

(1 Department of Parasitology, Sun Yat-Sen University of Medical Sciences, Guangzhou, 510089;
2 Fifth Peoples Hospital in Wenzhou, Zhejiang Province;
3 Guangdong School for Material Management, Guangzhou)

Cluster analysis and discrimination of 72 cases with paragonimiasis and 18 normal persons basing on 11 items of examination results in Wenzhou city, Zhejiang province indicate. These patients are divided into three groups responding with three clinical types: type of lung damage, type of liver damage and subclinical type. Basing numerical analysis in combination with clinical picture, 6 items of examination are required for diagnosis of paragonimiasis (skin test, white blood cell count, erythrocyte sedimentation rate, ratio of eosinocyte cells, ratio of gamma globulin and ELISA). Two groups of discrimination function may be useful for diagnosis paragonimiasis.

Key words paragonimiasis; cluster analysis; discrimination analysis

· 作者须知 ·

医学名词术语应以全国自然科学名词审定委员会公布的《医学名词》(科学出版社)等系列丛书为准。如尚未收入,则可参照《英汉医学词汇》(人民卫生出版社)和《中国医学百科全书》(上海科技出版社)。新兴学科的名词术语尚无通用译名者,于文内第1次出现时,应附原词或加注释。药名以《中华人民共和国药典》最新版本为依据,一般不宜用商品名,药典未收入者应附注拉丁名。