

# 基于基因打靶技术的中医脏腑辨证分型算法

吴芸<sup>1</sup>, 周昌乐<sup>2</sup>, 张志枫<sup>3</sup>

WU Yun<sup>1</sup>, ZHOU Chang-le<sup>2</sup>, ZHANG Zhi-feng<sup>3</sup>

1. 厦门理工学院 计算机科学与技术系, 福建 厦门 361024

2. 厦门大学 人工智能研究所, 福建 厦门 361005

3. 上海中医药大学 研究生院, 上海 201203

1. Department of Computer Science and Technology, Xiamen University of Technology, Xiamen, Fujian 361024, China

2. Institute of Artificial Intelligence, Xiamen University, Xiamen, Fujian 361005, China

3. Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

E-mail: yininwu2006@yahoo.com.cn

WU Yun, ZHOU Chang-le, ZHANG Zhi-feng. Genetic targeting-based Zangfu syndromes differentiation algorithm for TCM. Computer Engineering and Applications 2007, 43(27): 22-25.

**Abstract:** Based on an evocation of gene targeting in biology and the research to the specification type of syndrome, we offer a novel method for Zangfu syndromes differentiation. The main thoughts are making the symptoms about the Zangfu syndromes, which include the main symptoms, the minor symptoms and the symptoms about tongue and pulse, to a syndrome model target and making the symptoms of a clinic case to a symptoms genome about the clinic case, which is used to shoot at the syndrome model target. The experiment results about syndromes differentiation for the apoplexy and the chronic gastritis approve the method is effective. The method proposed provides a new research method in the objectivity and calculability of information process for TCM, and also has some clinical value of diagnosis and application.

**Key words:** information process for Traditional Chinese Medicine (TCM); gene targeting; syndrome differentiation algorithm

**摘要:** 受生物学基因打靶技术的启发和通过对中医脏腑证型分类标准的研究, 提出了一种新颖的中医智能脏腑辨证算法, 称之为中医脏腑辨证分型算法。主要思想是把中医脏腑证型的主要症状和次要症状, 以及舌象、脉象信息设计为脏腑证型模靶, 临床病例病症信息作为脏腑病症基因组, 并进行模靶射击式的辨证分型操作。通过对慢性胃炎疾病的辨证分型计算结果的分析, 说明中医脏腑辨证分型算法的有效性, 为中医辨证信息处理的客观化、可计算化提供了一种新的研究方法, 同时具有一定的临床诊断应用价值。

**关键词:** 中医辨证信息处理; 基因打靶; 辨证分型算法

文章编号: 1002-8331(2007)27-0022-04 文献标识码: A 中图分类号: TP391; TP18

生命科学是本世纪的热门研究领域之一, 人们长期以来不断探索生命的奥秘, 关于生命活体运转机制的研究已成为多种科学研究的焦点, 同时生命活体运转机制不断地启发工程技术人员, 不断地发掘和研究各种人工算法模拟自然机制对各种信息进行有效处理。基因打靶技术就是对生物的遗传物质进行修饰并研究其相关的表型<sup>[1]</sup>。受到打靶技术对生物的遗传物质进行表型研究的启发, 把中医诊断标准作为证型模靶, 将诊断病人的表现症状倒模形成病症基因组, 使用证型模靶达到对病症基因组显型的识别, 从而实现计算机能够对疾病进行中医智能辨证分型的目的。

## 1 中医脏腑辨证基因打靶技术相关概念

生物学上的基因打靶技术是一种定向研究生物活体遗传

信息的实验手段。根据实现需要, 即设计打靶载体通过对生物活体遗传信息的定向修饰, 并通过各种方法导入细胞基因组中, 并使修饰后的遗传信息在生物体内遗传, 表达突变的性状, 达到研究基因功能等目的。受到生物学基因打靶技术启发, 结合计算机模式识别技术和中医脏腑辨证特点, 设计了中医脏腑辨证基因打靶技术及其脏腑辨证分型算法。下面首先介绍中医脏腑辨证基因打靶技术中的几个重要概念:

**定义 1** 根据中医权威机构制定的中医脏腑证型分型标准设计的基因序列, 基因序列上的每一个基因对应表示中医脏腑证型分型标准的一个主要症状或次要症状, 该基因序列称为中医证型的基因模靶, 简称为证型模靶。

**定义 2** 将某脏腑疾病下的中医临床常见症状按一定规则排列形成的二进制-实数混合编码序列, 称为病症模型。

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60672018), 厦门大学院士基金。

作者简介: 吴芸(1979-), 女, 博士, 主要研究领域为人工智能, 信息处理, 软件工程; 周昌乐, 教授, 博士生导师, 主要研究领域为人工智能; 张志枫(1959-), 教授, 硕士生导师, 主要研究领域为中医诊断学。

定义3 将临床病例的所有临床表现症状依照病症模型排列形成病症基因组的过程称为病症倒模。

定义4 将临床病例的所有临床表现症状按病症模型倒模形成的症状序列称为中医脏腑病症基因组。

定义5 中医证型模靶与病症基因组进行模式判别的过程为辨证投射。

定义6 中医脏腑辨证基因打靶技术的疾病信息编码是指分别对某一疾病下常见临床症状用统一编号表示(除舌象和脉象以外的其他表现症状),即从0开始按序正整数编号,其中脉象信息化分为【脉位 脉率 脉宽 脉长 脉力 脉流利度 脉紧张度 脉均匀度】8个子项。舌象信息化分为【舌色 舌形 舌态 舌苔 舌苔色】5个子项。临床主要表现的脉象、舌象编码表,见表1和表2:

表1 脉象信息编码表

分类	脉名	编码	脉名	编码	脉名	编码	脉名	编码
脉位	浮脉	1	沉脉	2	散脉	3		
脉率	迟脉	1	缓脉	2	数脉	3	疾脉	4
脉宽	细脉	1	大脉	2	洪脉	3		
脉长	短脉	1	长脉	2				
脉力	虚脉	1	微脉	2	弱脉	3	实脉	4
脉流利度	滑脉	1	动脉	2	涩脉	3		
脉紧张度	弦脉	1	紧脉	2	革脉	3	濡脉	4
脉均匀度	结脉	1	代脉	2	促脉	3		

表2 舌象信息编码表

舌色	淡红	淡白	红	绛红	青紫				
编码	0	1	2	3	4				
舌形	正常	老	嫩	胖大	瘦薄	点刺	齿痕	裂纹	舌衄
编码	0	1	2	3	4	5	6	7	8
舌态	正常	痿软	短缩	颤动	歪斜	吐弄	强硬	舌纵	麻痹
编码	0	1	2	3	4	5	6	7	8
舌苔色	正常	白	黄	灰	黑				
编码	0	1	2	3	4				
舌苔质	正常	厚	薄	润	燥	腻	腐	剥落	
编码	0	1	2	3	4	5	6	7	

注:其中编码“0”表示“正常”或对应信息缺失。

## 2 中医脏腑辨证基因打靶技术理论概述

中医脏腑辨证基因打靶技术通过“倒模”和“辨证投射”两个主要操作和一个辨证知识库,实现计算机的中医智能脏腑辨证功能:中医脏腑辨证系统首先根据疾病的临床症状,在病症模型库中查找对应疾病病症模型,将临床症状倒模成病症基因组。再将倒模成的病症基因组与证型模靶库中对应的证型模靶进行基因投射操作,最终得到辨证分型结果,即中医脏腑证型。

### 2.1 中医脏腑辨证证型模靶

根据国家中医权威机构或协会制定的中医证型分型标准,首先设计了中医脏腑辨证模靶。假设中医证型A {主症 X x<sub>1</sub>, x<sub>2</sub>, ..., x<sub>n1</sub>; 次症 Y y<sub>1</sub>, y<sub>2</sub>, ..., y<sub>n2</sub>; 舌象 T t<sub>1</sub>, t<sub>2</sub>, ..., t<sub>5</sub>; 脉象 G g<sub>1</sub>, g<sub>2</sub>, ..., g<sub>8</sub>}, 证型分型标准 f={α+β+γ}。其中 x<sub>i</sub> {证型A的主要症状}且 1 ≤ i ≤ n<sub>1</sub>, y<sub>j</sub> {证型A的次要症状}且 1 ≤ j ≤ n<sub>2</sub>, t<sub>b</sub> {证型A的舌象信息}且 1 ≤ b ≤ 5, 舌象信息包括舌色 t<sub>1</sub>, 舌形 t<sub>2</sub>, 舌态 t<sub>3</sub>, 舌苔色 t<sub>4</sub>, 舌苔质 t<sub>5</sub>, g<sub>d</sub> {证型A的脉象信息}且 1 ≤ d ≤ 8, 脉象信息分为:脉位 g<sub>1</sub>, 脉率 g<sub>2</sub>, 脉宽 g<sub>3</sub>, 脉长 g<sub>4</sub>, 脉力 g<sub>5</sub>, 脉流利度 g<sub>6</sub>, 脉紧张度 g<sub>7</sub> 和脉均匀度 g<sub>8</sub>, m, n ∈ Z<sup>+</sup>, f 是证型A的分型标准,即表明有 α<sub>i</sub> 个主症和 β<sub>j</sub> 个次症(1 ≤ i ≤ K, K ∈ Z<sup>+</sup>, K 为

证型A的分型标准的数目),以及结合舌象和脉象进行综合判断。则中医脏腑辨证模靶 Γ 一般形式为 {num {X, Y, T, G} & f}, num 是对所有中医证型的统一编号,与中医证型名相对应。如图1所示:

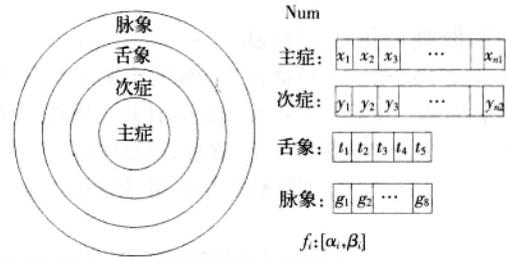


图1 中医脏腑证型模靶

有序排列证型的主症、次症和舌象、脉象信息,形成证型症状的序列,即中医脏腑辨证模靶表现形式,如图2所示:

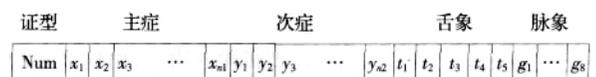


图2 证型模靶表现形式

### 2.2 中医脏腑病症基因组

中医病症基因组是将临床疾病的症状,按病症模型序列倒模成的疾病症状有效序列。设病例B {症状 p<sub>1</sub>, p<sub>2</sub>, ..., p<sub>s</sub>; 舌象: r<sub>1</sub>, r<sub>2</sub>, r<sub>3</sub>, r<sub>4</sub>, r<sub>5</sub>; 脉象 q<sub>1</sub>, ..., q<sub>8</sub> ∈ Z<sup>+</sup>}。每一类疾病(例如:胃病、心病、肺病等)都有中医病症模型,其中 p<sub>1</sub>, p<sub>2</sub>, ..., p<sub>s</sub> 是疾病 Γ 的症状的编码值(病症编码集合为 P), r<sub>1</sub>, ..., r<sub>5</sub> 分别表示舌象信息【舌色 舌形 舌态 舌苔色 舌苔质】的编码值 q<sub>1</sub>, ..., q<sub>8</sub> 分别是脉象信息【脉位 脉率 脉宽 脉长 脉力 脉流利度 脉紧张度 脉均匀度】的编码值。假设疾病 Γ 的病症模型 Φ 为 Φ {φ<sub>1</sub>, ..., φ<sub>m</sub>, φ<sub>m+1</sub>, ..., φ<sub>m+5</sub>, φ<sub>m+6</sub>, ..., φ<sub>m+13</sub>, m ∈ R<sup>+</sup>} 且 1 ≤ s ≤ m, φ<sub>i</sub> (1 ≤ i ≤ m) 表示疾病 Γ 的所有常见症状在病症基因组中的排列次序;其中 φ<sub>i</sub> (1 ≤ i ≤ m) 表示一般症状,其值为“0”或“1”, φ<sub>i</sub> (m+1 ≤ i ≤ m+5) 分别表示舌象信息【舌色 舌形 舌态 舌苔色 舌苔质】的编码; φ<sub>i</sub> (m+6 ≤ i ≤ m+13) 分别表示脉象信息【脉位 脉率 脉宽 脉长 脉力 脉流利度 脉紧张度 脉均匀度】的编码,病例B和疾病 Γ 的病症模型 Φ 经倒模 f 后,得到病例B的病症基因组 Φ。其中倒模 f 为: Φ = f(B, Φ) = [ψ<sub>1</sub>, ψ<sub>2</sub>, ..., ψ<sub>m</sub>, r<sub>1</sub>, ..., r<sub>5</sub>, q<sub>1</sub>, ..., q<sub>8</sub>], ψ<sub>i</sub> = { 1 j ∈ B(P) 0 j ∉ B(P) }, 1 ≤ i ≤ m, 其中“i ∈ B(P)”表示病例B中有表现症状 i, 则 ψ<sub>i</sub> = 1。因此,根据病例B包含的所有表现病症 i, 设置病症模型中对应的基因状态:若病例B中包含表现症状 i, 其对应基因状态为“1”, 病例B中不包含表现症状 i, 其对应基因状态为“0”。见图3所示:

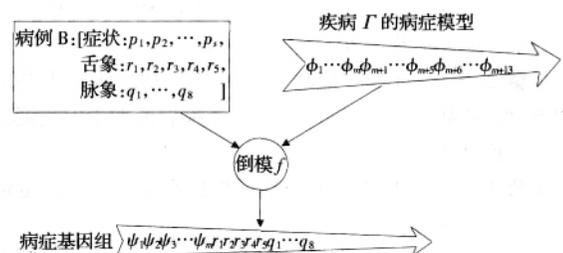


图3 中医脏腑辨证倒模

经倒模后,病症基因组的前面部分表示病人的临床表现症状,后面部分分别表示舌象信息和脉象信息。病症基因组是二

进制与实数的混合编码序列,其形式如图4所示:

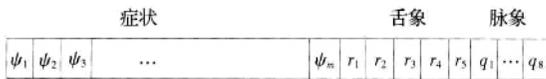


图4 病症基因组的表现形式

### 2.3 中医脏腑辨证投射机制

中医脏腑辨证投射是将病症基因组与中医证型模靶进行模式判别的过程。投射过程分为三部分:计算病症基因组的临床表现症状投射证型模靶的主症、次症,以及病症基因组中舌象信息和脉象信息投射证型模靶的舌象、脉象信息。即计算投射值  $f_i$ 。

(1)病症基因组症状投射证型模靶的主症,由病例B的病症基因组  $\Phi$  的症状投射  $num$  证型模靶  $\Gamma$  的主症,其投射值由参数  $\alpha$  表示,则

$$\alpha = \sum_{i=1}^{n1} \psi_x \quad (1)$$

(2)病症基因组症状投射证型模靶的次症,由病例B的病症基因组  $\Phi$  的症状投射  $num$  证型模靶  $\Gamma$  的次症,其投射值由参数  $\beta$  表示。则

$$\beta = \sum_{i=1}^{n2} \psi_y \quad (2)$$

(3)在中医辨证中,舌象和脉象信息对中医辨证有着重要的意义,且在中医辨证分型判定中,舌象信息和脉象信息是中医辨证分型判定内容的重要部分,因此,在中医脏腑辨证基因打靶中靶判定条件中,病例B的舌象、脉象信息投射证型模靶的舌象信息和脉象信息是中医脏腑辨证判定的第三步。由病症基因组的舌象、脉象信息投射  $num$  证型模靶  $\Gamma$  的舌象信息、脉象信息,其投射值由参数  $\eta$  表示:

$$\eta = e^{-\left(\sum_{i=1}^s (|r_i - t_i|) + \sum_{j=1}^q (|q_j - g_j|)\right)} + e \quad (3)$$

则辨证分型投射值为:

$$f = \max_{i=1}^K ((\alpha - \alpha_{num})^3 + (\beta - \beta_{num}) + \eta) \quad (4)$$

在中医脏腑辨证基因打靶技术中,病症基因组与疾病的所有证型模靶  $\Gamma$  进行投射计算,投射值越大,投射值对应证型的显性可能性越高。即  $num = NUM(\max_{i=1}^K f_i)$ ,  $K$  为证型模靶数目,函数  $NUM$  选择出最大  $f$  值所对应的证型模靶的  $num$ 。

### 2.4 中医脏腑辨证分型算法及其分析

#### 2.4.1 中医脏腑辨证分型算法

根据中医脏腑辨证基因打靶技术,设计了中医脏腑辨证分型算法。算法主要步骤描述如下:

步骤1 根据输入的“症状”集合,将表现症状转换为本文上述的统一编码值集合  $\Phi$ 。

步骤2 根据提供的“症状  $\phi(\phi - \Phi)$ ”,初步确定主要疾病类型  $\theta$ ,如心病、肺病、脾胃病等。并在疾病模型库中找到其对应的病症模型  $\Gamma$ 。

步骤3 根据疾病类型  $\theta$ ,在证型模靶库中查找确定疾病的证型模靶子库  $H$ 。

步骤4 将“症状”集合  $\Phi$  中的所有症状,按病症模型  $\Gamma$  进行倒模,形成疾病基因组  $\phi$ 。

步骤5 将病症基因组  $\phi$  与证型模靶子库  $H$  中的所有证型模靶进行打靶投射。计算投射值记为  $f_i$ ,其中  $1 \leq i \leq m, m$  为

模靶  $num$  号。

步骤6 比较投射值  $f_i$ 。最大的  $f_i$  值对应的  $i$  为辨证计算结果证型的  $num$ ,将  $num$  转换为中医证型名。

#### 2.4.2 中医脏腑辨证分型算法性能分析

中医脏腑辨证分型算法主要是将疾病的临床表现症状与按权威机构制定的证型分类标准设计的证型模靶进行基因投射,分型计算得到疾病的中医证型。分析算法的时间复杂度为:  $m \times length(\quad)$ ,即为  $n^2$  级。空间的消耗主要在于存储疾病基因组模型和证型模靶,所以空间复杂度也是  $n^2$  量级的。所以,从时间和空间角度来看,算法具有较好的时间和空间性能。

中医脏腑辨证分型算法的瓶颈问题在于对疾病“表现病症”的正确判断,以及中医证型分类标准的合理性,这也是算法是否能对疾病进行正确中医分型的关键因素。因此,中医脏腑辨证分型算法的辨证正确率与“病症”判断的准确率和中医证型分类标准的合理性成正比。即  $p_d, p_s, p_n, p_d$  为算法中医辨证正确率,  $p_s$  为评估“病症”判断的准确率,  $p_n$  为评估中医证型分类标准的合理性。这些因素正是目前中医的重要研究内容。

### 3 实验及其结果分析

由于社会生活快捷的节奏,慢性胃炎逐步成为了当代年轻人的常见疾病。之所谓的“百病之生,伤自脾胃”、“内伤脾胃,百病由生”之说。从中医角度看慢性胃炎,可分为五大证型,其中以中医症状为主症,西医胃镜检查结果为次症,结合舌象和脉象信息进行中医分型。因此,本文采用1989年“中国中西医结合研究会消化系统疾病专业委员会”制定的慢性胃炎中医分型标准为参照,设计慢性胃炎中医证型辨证的证型模靶和病症基因模型。在慢性胃炎中医分型标准中,次症主要是西医的胃镜检查结果,因此,缺省中医脏腑辨证证型模靶的次症部分。慢性胃炎症状编码见表3。

表3 慢性胃炎症状编码表

症状	胃脘胀痛	胃脘隐痛	胃脘灼痛	胃脘疼痛	痛窜两胁	喜按	喜暖
编码	1	2	3	4	5	6	7
症状	拒按	嗝气	嘈杂	泛酸	口苦	口臭	餐后胀闷
编码	8	9	10	11	12	13	14
症状	痞满	纳呆/少食	渴不欲饮	口干舌燥	便溏	腹泻	尿黄
编码	15	16	17	18	19	20	21
症状	大便干燥	便血	乏力/酸软	日久不愈			
编码	22	23	24	25			

慢性胃炎的中医证型模靶为:肝胃不和:[1,1 5 9 10 11,1 0 0 1 0, 0 0 0 0 0 1 0];脾胃虚弱:[2 2 6 7 14 15 16 19 20 24,1 0 6 1 2, 2 0 1 0 0 0 0 0];脾胃湿热:[3 3 1 12 13 21 15 17 2 0 0 2 5,0 0 0 0 0 1 2 0];胃阴不足:[4 3 18 22 2 7 0 0 0,0 0 1 0 0 0 1 0];胃络瘀血:[5 4 8 25 23,3 8 0 0 0, 0 0 0 0 0 1 1 0]。其中用“逗号”把证型号、主症和舌象脉象分割。慢性胃炎病症基因组模型为:【胃脘胀痛 胃脘隐痛 胃脘灼痛 痛窜两胁 喜按 喜暖 拒按 嗝气 嘈杂 泛酸 口苦 口臭 餐后胀闷 痞满 纳呆/少食 渴不欲饮 口干舌燥 便溏 腹泻 尿黄 大便干燥 便血 乏力/酸软 日久不愈 舌色 舌形 舌苔 舌苔色 舌苔质 脉位 脉率 脉宽 脉长 脉力 脉流力度 脉紧张度 脉均匀度】。

算法对77例慢性胃炎临床病例进行辨证分型计算(其中肝胃不和病人32例,脾胃虚弱病人25例,脾胃湿热病人7例;

胃阴不足病人 10 例,胃络瘀血病人 3 例)。部分实验结果见表 4,全部实验结果统计见表 5。

表 4 部分慢性胃炎病例中医证型对比表

编号	肝胃不和	脾胃虚弱	脾胃湿热	胃阴不足	胃络瘀血	算法		中医师	
						辨证证型	辨证证型	辨证证型	辨证证型
1	0.386 2	-7.931 9	-0.950 1	-0.864 5	-0.864 5	肝胃不和	肝胃不和	肝胃不和	肝胃不和
2	9.000 3	-7.981 3	0.185 1	-0.632 1	-0.632 1	肝胃不和	肝胃不和	肝胃不和	肝胃不和
3	0.135 3	1.135 3	-0.980 8	-0.632 0	-0.950 2	脾胃虚弱	脾胃虚弱	脾胃虚弱	脾胃虚弱
4	1.367 9	0.049 8	0.050 7	-0.864 5	-0.864 7	肝胃不和	脾胃虚弱	肝胃不和	脾胃虚弱
5	0.503 2	1.503 2	0.020 8	-0.631 8	-0.950 2	脾胃虚弱	脾胃虚弱	脾胃虚弱	脾胃虚弱
6	0.036 6	2.018 3	0.009 2	-0.950 2	-0.993 3	脾胃虚弱	脾胃虚弱	脾胃虚弱	脾胃虚弱
7	0.503 2	1.503 2	0.020 8	-0.631 8	-0.950 2	脾胃虚弱	脾胃虚弱	脾胃虚弱	脾胃虚弱
8	0.068 1	2.049 8	1.005 0	-0.950 2	-0.993 3	脾胃虚弱	脾胃虚弱	脾胃虚弱	脾胃虚弱
9	0.735 8	-7.582 3	0.050 1	2.000 3	-0.864 7	胃阴不足	胃阴不足	胃阴不足	胃阴不足
10	0.503 3	-7.814 9	1.050 7	1.000 9	-0.864 5	脾胃湿热	胃阴不足	脾胃湿热	胃阴不足
11	0.386 2	-7.931 9	0.056 5	1.000 1	-0.864 7	胃阴不足	胃阴不足	胃阴不足	胃阴不足
12	0.368 8	-7.949 3	1.049 8	2.000 0	-0.864 7	胃阴不足	胃阴不足	胃阴不足	胃阴不足
13	0.367 9	-7.950 2	0.049 8	2.018 3	-0.846 4	胃阴不足	胃阴不足	胃阴不足	胃阴不足
14	0.503 2	-0.814 9	-0.943 5	0.000 2	0.135 3	肝胃不和	胃阴不足	肝胃不和	胃阴不足
15	0.417 7	-0.900 4	0.068 1	1.000 0	-0.864 7	胃阴不足	胃阴不足	胃阴不足	胃阴不足
16	0.018 3	-7.963 4	-0.864 5	-0.632 0	-0.632 0	肝胃不和	胃络瘀血	肝胃不和	胃络瘀血
17	-0.963 4	-7.979 2	-0.981 6	-0.993 1	0.049 9	胃络瘀血	胃络瘀血	胃络瘀血	胃络瘀血
18	0.049 8	-7.931 9	-0.864 3	-0.631 8	8.368 2	胃络瘀血	胃络瘀血	胃络瘀血	胃络瘀血

表 5 77 例慢性胃炎辨证结果统计表

证型	肝胃不和	脾胃虚弱	脾胃湿热	胃阴不足	胃络瘀血	证型不详
辨证结果	37	22	9	7	2	0
正确结果	32	22	7	7	2	-

病例中缺乏与期待证型模靶的表现症状,同时舌象信息和脉象信息与证型模靶的舌象、脉象信息又不符是辨证计算结果与中医师辨证结果不符的主要原因。由方程(4)可知,与期待证型模靶的投射值为负值。如果病例中出现了能与其他证型模靶投射的症状,其投射值大于 0,则辨证分型不符中医师辨证结果。例如病 16# {症状: 胀痛,舌象: 青紫,脉象: 沉、弦},没有与“胃络瘀血”模靶相符的主症,但有与“肝胃不和”模靶相符的主症,因此,在不考虑胃镜检查结果的情况下,算法辨证误为“肝胃不和”,而不是“胃络瘀血”。如病例 17# {症状: 胀痛,舌象: 暗

红,脉象: 涩、细},有一个与“胃络瘀血”模靶相符的主症,其舌象信息基本符合,因而辨证计算得到“胃络瘀血”;另一方面,如果证型模靶的中靶条件较低,则计算得到该证型也相对容易。如病例 4# {症状: 胀痛、便溏,舌象: 淡红、白腻,脉象: 虚、弦},只出现两个主症信息相符,没有达到“脾胃虚弱”模靶要求至少出现 3 个相符主症,而“肝胃不和”模靶只要求出现两个以上相符主症。在脉象上与“肝胃不和”模靶的脉象信息相符,而与“脾胃虚弱”模靶的脉象信息不相符。因此在主症信息相对缺乏,在脉象上又更符合“肝胃不和”模靶的脉象信息,算法辨证计算结果为“肝胃不和”,与中医师诊断的“脾胃虚弱”结果不符。算法通过对 77 例临床病例进行辨证分型计算,辨证分型结果令人比较满意,辨证正确率达到 90.9%。

## 4 结论

本文研究为智能模拟中医辨证的症状-辨证-证型过程进行了一次有意义的尝试。中医辨证基因打靶技术有别于以往采用的智能技术,它从疾病表现症状出发,结合了中医辨证和中医师看病的特点。实验结果证明中医脏腑辨证分型算法能够对疾病的中医证型进行分型判定。因此,本文研究为中医辨证的客观化、可计算性提供了一种新的研究方法,也具有一定的临床诊断应用价值。(收稿日期: 2007 年 4 月)

## 参考文献:

- [1] 杨晓.基因打靶技术[M].北京:科学出版社,2001:1-100.
- [2] 模式识别[M].李晶皎,朱志良,译.北京:电子工业出版社,2004:209-226.
- [3] Jain A K,Duin.Statistical pattern recognition a review[J].Pattern Analysis and Machine Intelligence,IEEE,2000,22(1):4-37.
- [4] Fan Hui,Wen Jin-wei.A scalable pattern recognition method[C]//IEEE 2002 International Conference on Communication,Circuits and Systems and West Sino Exposition,2002,1:610-614.
- [5] Mottl V,Dvoenko S.Pattern recognition in interrelation data the problem fundamental assumptions recognition algorithms [C]//Proceeding of the 17th International Conference,2004,1:188-191.

(上接 12 页)

的分类性能。(收稿日期:2007 年 6 月)

## 参考文献:

- [1] Jain A K,Duin R P W,Mao J.Statistical pattern recognition a review[J].IEEE Trans Pattern Anal Mach Intell,2000,22(1):4-33.
- [2] Dasathry B V.Nearest neighbor(NN) norms:NN pattern classification techniques[M].Los Alamitos,CA:IEEE Computer Society Press,1990.
- [3] Qiang G Q,Zhang P.Neural networks for classification a survey[J].IEEE Trans Syst Man Cybern Part C,2000,30(4):451-458.
- [4] Burges J C.A tutorial on support vector machines for pattern recognition[J].Data Mining and Knowledge Discovery,1998,2:121-167.
- [5] Cristianini N,Shawe-Taylor J.An introduction to support vector

machines and other kernel-based learning methods[M].Cambridge U K:Cambridge Univ Press,2000.

- [6] Wolfe P.A duality theorem for nonlinear programming[J].Quarterly of Applied Mathematics,1961,19:239-244.
- [7] Kuhn H W,Tucker A W.Nonlinear programming[C]//Proceedings 2th Berkeley Symposium on Mathematical Statistics and Probability.Berkeley:University of California Press,1951:481-492.
- [8] Camastra F,Verri A.A novel kernel method for clustering[J].IEEE Trans Pattern Anal Mach Intell,2005,27(5):801-805.
- [9] Yager P R,Filev D P.Approximate clustering via the mountain method[J].IEEE Trans Syst Man Cybern,1994,24(8):1279-1284.
- [10] Statlog collection [DB/OL].http://www.niaad.liacc.up.pt/old/statlog/datasets.html.
- [11] LibSVM Website [DB/OL].http://www.csie.ntu.edu.tw/~cjlin/libsvm-tools/dataset.