

灰色系统预测模型对人体 正气随年龄变化的预测

□ 周 霞^{1*} 王 左²

(1. 山东中医药大学第二附属医院神经内科 山东 济南 250001)

(2. 上海中医药大学附属曙光医院急诊科 上海 200021)

摘 要 目的：应用灰色系统预测模型预测人体正气变化趋势。方法：通过对 883 名感冒患者首次就诊情况的临床流行病学调查，确定其正气随年龄变化情况并以数值量化。以每增长 10 岁为年龄分段，计算得出各年龄段正气虚损积分，形成原始数据列。以 70 岁以下 6 个年龄段积分原始数据为原始数据列，建立灰色预测数学模型，模型合格后外推预测 70~79 岁年龄段人体正气情况，并与原始数据相对比，以考察预测的准确性。结果：灰色预测数学模型可用于预测人体正气变化研究，预测准确率达 90% 以上。结论：灰色预测理论可用于中医药学研究，且预测准确性高。

关键词 灰色预测理论 中医药研究 正气变化

1 灰色系统理论应用于中医药研究的意义

在控制论中，人们常用颜色的深浅形容信息的明确程度，用“黑”表示信息未知，用“白”表示信息完全明确，用“灰”表示部分信息明确、部分信息不明确。相应地，信息完全明确的系统称为白色系统，信息未知的系统称为黑色系统，部分信息明确、部分信息不明确的系统称为灰色系统。灰色系统理论是上世纪 80 年代初由我国学者邓聚龙教授提出并发展的。该理论把一般系统论、信息论、控制论的观点

和方法延伸到社会、经济、生态等抽象系统，结合运用数学方法，发展了一套解决信息不完备系统即灰色系统的理论和方法，对未来学的研究具有重要意义。它通过对“部分”已知信息的生成、开发去了解、认识现实世界，实现对系统运行行为和演化规律的正确把握和描述，且对试验观测数据及其分布没有特殊的要求和限制，四个及四个以上的连续数据即可建立数学模型。

无论中医还是西医，人们对人体及疾病的认识，无不是一个从不知到部分已知，进而到已知的过程。故中医药对人体及疾病的认识，也是灰色的，是一个灰类白化的过程，可以应用灰色系统理论来研究。灰色系统理论应用于中医药的研究已在灰色关联分析的应用方面有所尝试，如吴忠等^[1]运用灰色关联度分析

* 作者简介 周霞，女，医学博士。主要从事灰色系统理论在中医药学中的应用研究。

来评价中药材连翘样品的质量。丁爱民等^[2]对近五年来有关中医药治疗老年性痴呆的文献进行了研究,归纳老年性痴呆为肝肾阴虚、脾肾阳虚、气血瘀滞、痰浊阻窍、髓海空虚五型。统计常用药物并归为补虚、化瘀、祛痰三类,通过对各类药物的灰色关联分析,来筛选主要有效药物。灰色预测是灰色系统理论的一个重要组成部分,是一种通过将无规律的原始数据生成、建模、拟合后,进而推测未来的新兴数学预测模型系统。至目前为止,尚未有灰色预测理论应用于中医药的研究报道。中医向来强调“上工治未病”,若能利用灰色系统预测模型来预测疾病的发展变化趋势,则可使传统对疾病发展趋势模糊的认识定量化、数字化,使原本处于“灰色”,对之了解不甚的规律明晰化、“白化”,则可提示医家临证遣方用药时,当以几分药力扶正以助祛邪,着力几分于“先安未受邪之地”,从而提高临床疗效。

2 建立灰色预测数学模型的方法及步骤

目前最常用的灰色预测模型是含一个变量的一阶微分方程,称之为 GM (1.1) 模型。GM (1.1) 模型的建立及其外推预测应分以下五步完成:

第一步:累加生成

设原始数列 $X = \{x(1), x(2), \dots, x(n)\}$, 对其进行一次累加生成,以弱化其随机性,强化其规律性,得累加生成列 $Y = \{y(t)\} \quad 1 \leq t \leq n$:

$$y(t) = \sum_{i=1}^t x(i) \quad t = 1, 2, \dots, n \quad (1)$$

第二步:均值生成

对累加数据列按公式 (2) 作均值生成,得均值数据列 $Z = \{z(t)\} \quad 1 \leq t \leq n$:

$$z(t) = \frac{1}{3}[y(t) + y(t-1)]^t = 2, 3, \dots, n \quad (2)$$

第三步:建立 GM (1.1) 模型

建立关于 $y(t)$ 的一阶线性微分方程,即其时间响应式:

$$\frac{dy(t)}{dt} + ay(t) = u \quad (3)$$

此式即为 GM (1.1) 预测模型,解该变量分离型微分方程得其特解为:

$$y(t) = \left[x(1) - \frac{u}{a} \right] e^{-a(t-1)} + \frac{u}{a} \quad (4)$$

式中为待定系数,根据最小二乘法估计参数向量,并由矩阵计算得其表达式为:

$$a = \frac{1}{D} \left\{ (n-1) \left[- \sum_{t=2}^n x(t)z(t) \right] + \left[\sum_{t=2}^n z(t) \right] \left[\sum_{t=2}^n x(t) \right] \right\} \quad (5)$$

$$u = \frac{1}{D} \left\{ \left[\sum_{t=2}^n z(t) \right] + \left[- \sum_{t=2}^n x(t)z(t) \right] + \left[\sum_{t=2}^n x(t) \right] \left[\sum_{t=2}^n z^2(t) \right] \right\} \quad (6)$$

$$\text{其中 } D = (n-1) \sum_{t=2}^n z^2(t) - \left[\sum_{t=2}^n z(t) \right]^2 \quad (7)$$

由式 (4) 所得估计值 $\hat{y}(t)$ 作累减还原生成,得原始数值 $x(t)$ 的估计值 $\hat{x}(t)$:

$$\hat{x}(t) = \hat{y}(t) - \hat{y}(t-1) \quad (8)$$

第四步:对 $\hat{x}(t)x$ 与 $x(t)$ 进行拟合效果检验(可靠性检验)

若两者拟合精度好,则模型可用于外推预测,若两者拟合精度不合格,则不可直接用于外推预测,须经残差修正后,再进行外推预测。确定灰色数列模型的可靠性可用平均相对误差、关联系数、后验差比值和小误差概率来检验。本论文拟用平均相对误差及后验差比值法来检验模型精度。

设原始数据列为 $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$, 模拟数据列为 $\hat{x}^{(0)} = (\hat{x}_1^{(0)}, \hat{x}_2^{(0)}, \dots, \hat{x}_n^{(0)})$ 则第 K 个数据 $x_k^{(0)}$ 的模拟残差 $\varepsilon(k) = x_k^{(0)} - \hat{x}_k^{(0)}$ (9)

$$\text{相对误差检验: } \alpha = \frac{1}{n} \sum_{k=1}^n \Delta k \quad (10)$$

$$\text{其中第 K 个数据的相对误差: } \Delta k = \left| \frac{\varepsilon(k)}{x_k^{(0)}} \right|_{k=(1,2,\dots,n)} \quad (11)$$

$$\text{后验差比值检验: } c_0 = s_2/s_1 \quad (12)$$

$$\text{其中:原始数据方差为 } s_1^2 = \frac{1}{n} \sum_{k=1}^n [x_k^{(0)} - \bar{x}]^2 \quad (13)$$

$$\text{残差方差为 } s_2^2 = \frac{1}{n} \sum_{k=1}^n [\varepsilon(k) - \bar{\varepsilon}]^2 \quad (14)$$

$$\text{原始数据均值为: } \bar{x} = \frac{1}{n} \sum_{k=1}^n x_k^{(0)} \quad (15)$$

残差均值为 $\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon_k$ (16)

根据精度检验等级参照表判断灰色数列的拟合优度。其中平均相对误差 $\bar{\Delta}$ 和模拟误差都要求越小越好，关联度 ε 要求越大越好，一般认为 ε 要大于 0.68 时，模型才令人满意，预测方可信。均方差比值 c 越小越好，及小误差概率 p 越大越好。见表 1。

表 1 精度检验等级参照表

预测精度等级	相对误差 (α)	关连度 (ε)	均方差比值 (c)	小误差概率 (p)
一级 (好)	0.01	0.9	0.35	0.95
二级 (合格)	0.05	0.8	0.5	0.8
三级 (勉强)	0.1	0.7	0.65	0.7
四级 (不合格)	0.2	0.6	0.8	0.6

第五步：外推预测

如果拟合优度高，即模型预测效果满意，可按下式进行外推预测： $\hat{x}(t) = \hat{y}(t) - \hat{y}(t-1)^{t=n+1, n+2, \dots}$ (17)

3 应用灰色预测模型预测人体正气变化趋势并检验

本研究以通过对 883 名感冒患者首次就诊情况的临床流行病学调查结果研究对象，以所得之随年龄增长而变化的正气虚损积分为原始数列，建立灰色预测数学模型，来预测人体正气虚损积分变化，并与调查所得原始数据比较以验证预测效果。调查对象为 2004 年 7 月 1 日至 2005 年 2 月 28 日在上海市曙光医院、上海市中医医院、上海市普陀区中医医院、上海市闸北区中医医院、上海市岳阳医院及上海市嘉定区中医医院急诊科首次就诊的感冒患者。

3.1 病例纳入标准 (1) 有外感诱因或感冒接触史；(2) 有感冒的症状体征：鼻咽部（鼻塞、流涕、咳嗽等）和/或全身症状（发热、恶寒、头身疼痛）等；(3) 白细胞计数低于 $10.0 \times 10^9/L$ ，且嗜中性粒细胞不升高（小于或等于 75%）；(4) 年满 16 周岁且就诊前未服药者。

3.2 病例排除标准 (1) 不符合上述标准任何一条者；(2) 有下呼吸道感染症情者；(3) 有严重心、肝、肾等脏器疾病或造血系统疾病者；(4) 无其它感

冒症状，仅表现为发热及咽痛之急性咽喉炎者；(5) 不愿意被调查者。

3.3 诊断标准 感冒中、西医诊断标准参照郑筱萸主编的《中药新药临床研究指导原则》之“中药新药治疗感冒的临床研究指导原则”制定。

3.4 调查内容 包括患者姓名、性别、年龄、发病及就诊时间、是否有基础疾病及感冒症情、（气、血、阴、阳）虚性症状表现。虚症赋值标准、分级标准参照《中药新药临床研究指导原则》，分别赋值轻、中、重度气虚或阳虚症状以 2、4、6 分。阴、津、血虚者不分级，对于津亏表现（大便干结、口燥咽干），任意一项或两项出现，均只计 2 分；阴虚表现（五心烦热、午后潮热）任意一项或两项出现，均只计 2 分，盗汗计 2 分；血虚表现：唇甲色淡计 2 分，头晕黑朦视具体情况计分；虚性舌象计 2 分；鉴于当今年青医师对脉诊不够重视，且体会不深，故脉象视整体辨证情况酌定取舍。符合整体症情表现之虚性脉象，计 2 分。此症情赋值标准参照王虎军等^[3]和赵铁良^[4]的临床研究赋值标准。

3.5 调查表的质量控制 调查之初由研究者对各医院课题组成员就如何调查及填写调查表进行培训。各医院课题组长由该院急诊科主任担任，负责和监控该院调查的进展。调查开始后研究者在每月初到各医院收集上月完成的病例，监控调查表完成的质量并协商解决所出现的问题。同时发放本月调查表。第一个月发放 60 份，以后每月发放 45 份，填完随时补发。

3.6 结果 自 2004 年 7 月 1 日至 2005 年 2 月 28 日，在上海市曙光医院等六家医院急诊科共收集了有效病例 883 份。其中上海市曙光医院 159 例，岳阳医院 80 例，嘉定区中医医院 340 例，市中医医院 153 例，闸北区中医医院 87 例，普陀区中医医院 64 例。883 例病人中，男性 439 例，女性 444 例。

根据《灵枢·天年》按十年分段的记载，本研究以每递增十岁为年龄分段标准，分别计算各年龄段正气虚损积分，取其算术平均值为该年龄段患者正气虚损积分。结果见表 2：

表 2 正气虚损积分随年龄变化表

年龄 (岁)	病例数	正虚积分 ($\bar{X} \pm SD$)
<20	35	10.24 ± 4.31
20-29	119	8.83 ± 4.55
30-39	90	8.53 ± 4.25
40-49	81	9.40 ± 4.82
50-59	92	10.31 ± 4.54
60-69	61	11.45 ± 5.44
70-79	54	13.39 ± 6.24
>80	15	18.91 ± 5.07

3.7 建立人体正气变化预测模型

3.7.1 确立原始数据列 X, 作累加生成, 得出累加生成列 Y

确立原始数据列及检验数值

取小于 20 岁、20~29 岁、30~39 岁、40~49 岁、50~59 岁、60~69 岁年龄段正气虚损指数 10.25、8.83、8.53、9.40、10.31、11.45 六个数值组成原始数据列, 以建立预测数学模型, 70~79 岁年龄段的正虚积分作为检验预测模型精度及可行性的检验数值。即:

表 3 正虚积分及 $Y(t) \geq Z(t)$ 计算值

年龄分段 (岁)	20	20-29	30-39	40-49	50-59	60-69	Σ
t	1	2	3	4	5	6	
$x(t)$	10.3	8.83	8.53	9.4	10.31	11.45	48.52
$y(t)$		19.08	27.61	37.01	47.32	58.77	
$z(t)$		14.665	23.345	32.31	42.17	53.05	165.53
$z^2(t)$		215.0622	544.989	1043.94	1777.89	2813.77	6395.65
$z(t)x(t)$		129.492	199.1329	303.714	434.7212	607.3653	1674.425

由公式(7)得:

$D = (6 - 1) \times 6395.65 - 165.532 = 4578.052$

由公式(5)得:

$\alpha = [(6 - 1)(-1674.425) + 165.53 \times 48.52] / 4578.052 = -0.0744$

由公式(6)得:

$u = [165.53 \times (-1674.425) + 48.52 \times 6395.65] / 4578.052 = 7.24089$

$\frac{u}{\alpha} = 7.24089 / (-0.0744) = -97.3228$

将参数 $x(1), \alpha, \frac{u}{\alpha}$, 代入公式(4)中得:

$X = (x(1):10.25, x(2):8.83, x(3):8.53, x(4):9.40, x(5):10.31, x(6):11.45)$

对原始数据列累加生成, 得出累加生成列

依据公式 (1), 对原始数据列的四个数据进行累加生成, 形成新的累加生成列, 以弱化其随机性, 增强其规律性。得累加生成列:

$Y = (y(2):19.08, y(3):27.61, y(4):37.01, y(5):47.32, y(6):58.77)$

3.7.2 建立均值生成列 依据公式 (2), 对累加生成列数据进行均值生成, 以增强数据的规律性, 形成均值生成列:

$Z = (z(2):14.665, z(3):23.345, z(4):32.31, z(5):42.17, z(6):53.05)$

3.7.3 建立 GM (1.1) 模型 依据公式 (4) 建立灰色预测一阶微分模型。由公式 (4) 可知, 建立灰色预测模型, 必需先计算出待定系数 a, u。由公式 (5)、(6)、(7) 可计算得出这两个待定系数的值。计算参数 $z^2(t)$ 及 $z(t)x(t)$ 的值并与 $y(t), z(t)$ 各数值组成列表, 见表 3。

整理得:

$y(t) = 107.5728e^{0.0744(t-1)} - 97.3228 \quad (18)$

式 (18) 即据正气虚损变化指数原始资料建立的灰色预测数学模型。

3.7.4 对所建立的模型进行拟合优度检验 利用相对误差及后验差比值法对所建立的模型进行拟合优度检验:

由式 (18) 可计算得出 $y(t)$ 的估计值 \hat{y}_t 。再依据公式 (8) 对 $y(t)$ 的估计值进行累减还原生成, 得出 $x(t)$ 的估计值 \hat{x}_t 。计算估计值与原始数值之间的残差。具体数值见表 4。

表 4 $y(t)$ 与 $\hat{y}t$, $x(t)$ 与 $\hat{x}t$ 的比较

年龄分段 (岁)	20	20-29	30-39	40-49	50-59	60-69
t	1	2	3	4	5	6
$x(t)$	10.25	8.83	8.53	9.4	10.31	11.45
$y(t)$		19.08	27.61	37.01	47.32	58.77
$\hat{y}t$		18.5588	27.5094	37.1513	47.5379	58.7268
$\hat{x}t$		8.3088	8.9506	9.6419	10.3866	11.1889
$\varepsilon(t)$		-0.5212	0.4206	0.2419	0.0766	-0.2611

相对误差检验:

平均相对误差可依公式 (9)、(10)、(11) 计算得出:

$\alpha=0.0348$
与“精度检验等级参照表”相对照, 0.00348 < 0.01, 模型等级为一级。

后验差比值法检验:
后验差比值可依公式 (12)、(13) (14)、(15) 及 (16) 计算得出:
后验差比值 $c_0=0.12$
与“精度检验等级参照表”相对照: 0.12 < 0.35, 模型精度为一级。

故模型预测效果满意, 可用之于外推预测。
3.7.5 外推预测 如前所验, 建立的 GM (1.1) 模型拟合优度高, 可用于正虚指数的外推预测。故用所建立的模型对 70~79 年龄段的正虚指数进行预测。并与原始数据中的该年龄段的正虚指数进行对比, 以验证利用该模型外推预测所得的数据的准确度, 考察利用灰色预测数学模型于中医药研究的可行性。

将 $t=7$ 代入所建立的模型式 (18) 中, 计算得:

表 5 正虚积分及 $y(t)$ 、 $z(t)$ 计算值

年龄分段 (岁)	30-39	40-49	50-59	60-69	Σ
t	1	2	3	4	
$x(t)$	8.53	9.4	10.31	11.45	31.16
$y(t)$		17.93	28.24	39.69	
$z(t)$		13.23	23.09	33.97	70.28
$z^2(t)$		175.03	532.92	1153.62	1861.57
$z(t)x(t)$		124.362	238.0064	388.8993	751.2676

由公式得各参数值如下:
 $\alpha=-0.09897$
 $u=8.068157$

$\hat{y}(7)=70.7799$
将所得 $\hat{y}(7)$ 值与 $\hat{y}(6)$ 值代入公式 (8) 中, 求得 $x(7)$ 的估计值 $\hat{x}(7):=12.0531$ 。
 $x(7)$ 的原始数值为 13.38636, 与估计值比较, 其预测准确度为:
(12.0531/13.38636) 100% = 90.04%

预测结果满意。说明所建立的灰色预测数学模型准确度高, 可以用于感冒患者正虚指数随年龄变化趋势的预测。
3.8 少数数据建模及外推预测 根据灰色系统理论: 灰色预测模型的特点是少数数据、贫信息建模, 用于随时间推移而变化的数列资料的外推预测。最少只需要 4 个相邻连贯的数据资料即可建立模型。故再对 4 个数据建模进行测试和探讨。

取 30~39 岁、40~49 岁、50~59 岁、60~69 岁年龄段正气虚损指数 8.53、9.40、10.31、11.45 四个数值组成原始数据列以建立预测数学模型。计算参数 $z^2(t)$ 及 $z(t)x(t)$ 的值并与 $y(t)$ 、 $z(t)$ 各数值组成列表, 见表 5:

$\frac{u}{\alpha}=8.068157/(-0.09897)=-81.52$
将参数 $x(1)$ 、 α 、 $\frac{u}{\alpha}$ 代入公式 (4) 中得灰色预测

数学模型:

$$y(t) = 90.05e^{0.09897(t-1)} - 81.52$$

对所建立的模型进行拟合优度检验

经计算, 得平均相对误差 $\alpha = 0.0094$, 关联系数 $\varepsilon = 0.99$,

与“精度检验等级参照表”相对照, 模型等级为一级。

外推预测

将 $t=5$ 代入所建立的预测模型中, 计算得:

$$\hat{y}(5) = 52.1054$$

将所得 $\hat{y}(5)$ 值与 $\hat{y}(4)$ 值代入公式 (8) 中, 求得 $x(5)$ 的估计值 $\hat{x}(5)$: 12.44507, 与原始数据值 13.3864 比较, 预测准确率为: 93%。

4 讨 论

灰色预测理论以“部分信息确定, 部分信息不确定”的、人们对之尚没有完全认识清楚的事物为研究对象, 通过对已知信息的生成、开发去了解、认识系统的运行行为和演化规律。在中医药理论中, 人们对发病过程中邪正盛衰变化态势的认识是灰色的。疾病的发生是邪正双方斗争的结果。虽然医家总结出了邪盛正不衰、邪盛正衰、邪正相持等多种对比形式, 然人们对邪正斗争过程中双方力量的具体对比状况是不了解的。假如初始发病时邪盛正不衰的力量对比是 10:0, 邪正相持时是 5:5, 邪盛正衰的双方力量对比则是多变的。从数学方面讲, 是 10:0 到 5:5 的一个闭区间, 有无穷的解, 故邪盛正衰的状态是一个灰色的状态。同样, 正盛邪退的状态也是一个灰色的状态。然疾病的发生发展是随着邪正双方力量对比的变化而推进的。故对邪正盛衰变化的进一步认识对认识疾病, 把握疾病的转变规律有着极其重要的意义。

在邪气与正气二者之间, 正气又是决定发病与否的更为重要的因素。如中医认为: “正气存内, 邪不可干”; “邪之所凑, 其气必虚”。对于外感病而言更是如此。“卒然逢疾风暴雨而不病者, 盖无虚, 故邪不能独伤人。此必因虚邪之风, 与其身形, 两虚相得, 乃客其形。”故对于外感病如感冒而言, 人体发病与否, 正气的盛衰起决定性作用 (毒性强劲的疫疔之邪当属别论)。外来邪气的性质、强弱变化多端,

不可人为控制, 故难于把握。而人体正气强盛与否却是可感知, 可调查研究分析的, 以积分的形式甚至可以将其初步量化。同时因为邪正双方是一个你消我长, 敌进我退的关系, 研究其中一个方面的变化也就明晰了其另一方面的消长变化。故本研究以积分的形式量化感冒患者人体正气虚损的状况进行研究分析, 来探讨患病机体的正气变化态势。

灰色数学理论的核心就是使“灰类白化”, 即通过数学的方式对知之不甚的事情或规律进行研究, 使之明晰化, 从而使人们更好地把握规律, 控制可控因素, 使事物向良性方面发展。灰色预测理论就是通过已知信息的分析, 找出隐藏其中的规律, 来预测将要发生的事件。根据所调查到的样本总体 883 例病人的正气虚损积分情况, 以 10 岁为年龄分段, 计算出各年龄段患者的正气虚损积分算术平均值。以各均值数组成原始数列, 依据灰色数学理论中的建模方法建立 GM (1.1) 预测模型, 并对模型进行精度检验。通过相对误差法和后验差比值法双重检验, 表明模型精度为一级, 可以进行外推预测。预测结果显示: 预测值与实际值相比, 准确率分别达 90.04% 及 93%。说明灰色预测理论完全可以应用于中医药研究, 而且建模需求信息较少, 预测精度高。中医强调“上工治未病”。若根据已知的某疾病的信息, 能更明晰地预测出其将要发生的情况, 及时控制可控因素, 当能明显减轻病人痛苦, 提高临床疗效。本项研究表明: 正确应用灰色预测理论于中医药研究, 能使中医“治未病”的学术思想明晰化、量化、现代化。据此遣方用药, 则可使用药更为精当, 或先安未受邪之地, 从而提高临床疗效。

参考文献

- [1] 吴 忠, 苏薇薇, 何新新. 中药连翘质量的灰色模式识别研究. 中药材 [J]. 2000, (23) 9: 536-538.
- [2] 丁爱民, 方永奇. 治疗老年性痴呆常用中药的灰色关联分析. 数理医药学杂志 [J]. 2000, 13 (6): 501-503.
- [3] 王虎军, 刘 汾, 刘丽萍, 等. 抗病毒口服液治疗感冒 (风热证) 120 例临床观察. 中华临床医药 [J]. 2002, 3 (8): 76-78.
- [4] 赵铁良. 感冒软胶囊治疗风寒感冒临床验证. 中国中医药信息杂志 [J]. 2001, 8 (7): 52-53.