

近红外光谱和聚类分析法无损快速鉴别小儿抽风散

王晶娟¹, 张贵君^{1*}, 周 群², 孙素琴^{2*}

1. 北京中医药大学中药学院, 北京 100102

2. 清华大学化学系, 北京 100084

摘 要 以小儿抽风散不同缺味复方的近红外图谱作为聚类分析的对象, 在建立混合成分模型的基础上, 采用 SIMCA 聚类分析法对小儿抽风散不同缺味复方进行了快速的分类研究。实验中, 在保证其他配伍组分完整的基础上, 分别设缺少蜈蚣、全蝎、僵蚕、土鳖虫和蝉蜕的 5 个阴性样品。结果表明: 尽管各样品的近红外图谱很相似, 难以直接区别, 但采用近红外光谱和聚类分析相结合的方法可以鉴定小儿抽风散的不同缺味复方。各样品之间光谱聚类结果较理想, 盲样检测的正确率可达 90 % 以上。此外, 如果能够获得足够的样本, 增加训练集样本数和采样的代表性, 加强操作的标准程度, 该法的准确率将会大大提高。因此, 近红外光谱与聚类分析法相结合可以快速、无损识别中药复方。

关键词 红外光谱; 小儿抽风散; 聚类分析法; 无损鉴别

中图分类号: R282.5 **文献标识码:** A **文章编号:** 1000-0593(2008)02-0327-04

引 言

动物类中药品种繁多, 来源复杂, 难以鉴别。特别是全蝎、蜈蚣等贵重药材, 疗效好, 价格高, 市场上以假乱真现象时有发生, 严重影响了临床用药的疗效与安全, 所以有必要寻找一种快速、准确鉴别动物类中药真伪的方法, 为临床合理用药和动物类中药的鉴别研究提供科学依据。

中药是一个复杂的混合物体系, 所含各化合物吸收强度的叠加具有难以解析的复杂性, 造成谱图有较大的相似性。但事实上不同种的中药材及不同配伍组分的中药复方, 只要各化学成分的含量不同, 红外的谱图就会有差异, 借助药材指纹谱图的差异, 建立相关的模式识别法, 实现谱图的辨认是可行的, 已有文献报道^[1-11]。在此基础上采用该技术尝试中药复方缺味药的智能识别。小儿抽风散是由蜈蚣、全蝎、蝉蜕、僵蚕、半夏、天南星、厚朴、橘红、枳壳、甘草、朱砂、土鳖虫、钩藤、薄荷十四味药组成的中药复方, 含五种动物药。本文运用模式识别中的聚类分析法对不同缺味药的小儿抽风散进行了快速识别, 达到了预期判别的结果。

1 实验部分

1.1 仪器设备与附件

收稿日期: 2007-02-08, 修订日期: 2007-05-12

基金项目: 国家教育部科学技术研究重点项目(00167)资助

作者简介: 王晶娟, 女, 1977 年生, 北京中医药大学中药学院博士研究生

近红外光谱仪(NIR)为美国 Perkin Elmer 公司的 Spectrum One, INGAAS 检测器, 分辨率 16 cm^{-1} , 测量范围: $4\,000\sim10\,000\text{ cm}^{-1}$, 扫描信号累加 16 次。

1.2 样品来源与检测

蜈蚣 (Scolopendra)、全蝎 (Scorpio)、僵蚕 (Bombyx Batryticatus)、土鳖虫 (Eupolyphaga seu Steleophaga)、蝉蜕 (Periostracum Cryptotympanae) 5 个组分购于北京同仁堂, 由北京中医药大学中药学院的张贵君教授鉴定。按照符合散剂的技术要求, 分别将药材研成粉末, 然后按组方的不同配比进行组方, 具体配伍组分如表 1 所示。分别取样品各 2 g, 置测量杯中, 进行扫描, 并做重现性测定, 共得到 200 张红外谱图。

Table 1 Alignments of different compound recipes

	蜈蚣	全蝎	僵蚕	土鳖虫	蝉蜕
FUFANG1	-	+	+	+	+
FUFANG2	+	-	+	+	+
FUFANG3	+	+	-	+	+
FUFANG4	+	+	+	-	+
FUFANG5	+	+	+	+	-

注: 各缺味复方中均含有半夏、天南星、厚朴、橘红、枳壳、甘草、朱砂、钩藤、薄荷; “+”为含有该药, “-”为不含该药

1.3 数据处理

红外光谱聚类分析采用 Perkin Elmer 公司的 Quant+ 和

*通讯联系人 e-mail: guijunzhang@163.com

SIMCA 应用软件^[12]。

2 结果与讨论

2.1 缺味复方的近红外谱图

各缺味复方的近红外谱图如图 1 所示, 各谱图相似度较高, 难以从表观上进行指纹特征提取, 故采用模式识别技术进行智能识别。

2.2 缺味复方聚类分析模型的建立

选 SIMCA (Soft independent modeling of class analogy) 聚类分析方法, 从 200 个缺味复方谱图中随机抽取 160 个组成训练集, 利用 Perkin Elmer 公司的 Quant⁺ 软件, 对其训练集样本进行训练。选用波数范围在 $7\,500 \sim 4\,000\text{ cm}^{-1}$ 的指纹特征较强的区域, 建立聚类模型。

2.3 缺味复方聚类模型的评价指标

2.3.1 类间距

在选用 SIMCA 进行聚类分析中, 一般采用聚类中心之间的距离 (类间距) 评价聚类效果, 而聚类结果好坏的关键

是合适特征的选取。得到类间距的数值后, 数值越大表明类与类之间差异越明显。本实验中 5 个缺味复方模型的类间距如表 2 所示。

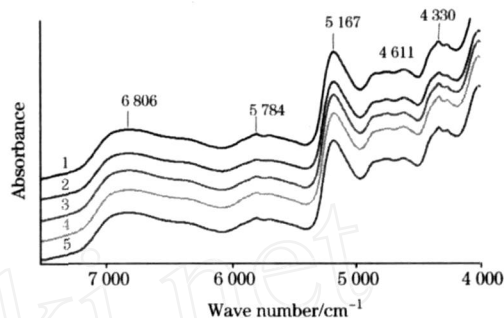


Fig 1 FTIR spectra of different compound recipes

- 1: the compound recipe without Scolopendra;
- 2: the compound recipe without Scorpio;
- 3: the compound recipe without Bombyx Batryticatus;
- 4: the compound recipe without Eupolyphaga seu Steleophaga;
- 5: the compound recipe without Periostracum Cryptotympanae

Table 2 Inter class distances of different compound recipes

	FUFANG1	FUFANG2	FUFANG3	FUFANG4	FUFANG5
FUFANG1	-	37.84	9.66	34.91	44.6
FUFANG2	-	-	76.59	137.71	180
FUFANG3	-	-	-	43.55	40.19
FUFANG4	-	-	-	-	65.8

2.3.2 识别率与拒绝率

识别率与拒绝率是反映类模型之间聚类可信度。所谓识别率就是考察某类供试品有多少落在该类模型的区域内, 而拒绝率是考察某类供试品模型对于其他不属于该类的未知供

试品的拒绝程度, 即是否落在该类模型的区域外。当 2 个值都为 100 % 时, 表明供试品之间没有重叠, 可以较好的将其聚类分开。表 3 为 5 个缺味复方的识别率与拒绝率。为了直观地表述, 可以将表 3 转换为识别率和拒绝率的柱形图, 如图 2 所示。

Table 3 Recognition rate and rejection rate of different compound recipes

	Recognition rate/ %	Rejection rate/ %
FUFANG 1	98 (39/ 40)	100 (160/ 160)
FUFANG 2	93 (37/ 40)	100 (160/ 160)
FUFANG 3	95 (38/ 40)	100 (160/ 160)
FUFANG 4	100 (40/ 40)	100 (160/ 160)
FUFANG 5	100 (40/ 40)	100 (160/ 160)

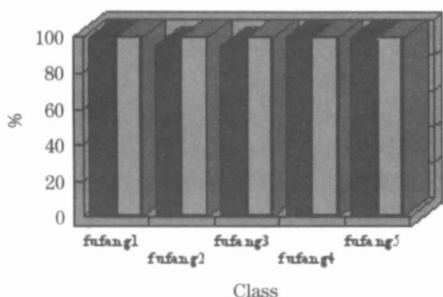


Fig 2 Histogram of recognition rate and rejection rate of different compound recipes

2.4 聚类分析结果图示

图 3 所示为 5 个缺味复方的聚类模型图。从模型图上看, 5 个缺味复方之间没有重叠, 互不干扰, 聚类结果比较理想。

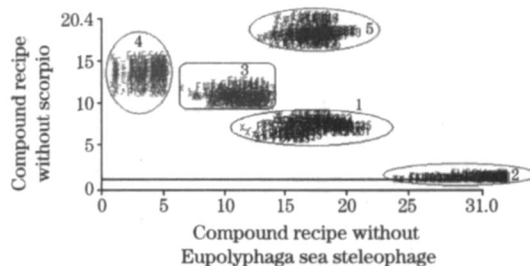


Fig 3 Classification results of different compound recipes

- 1: the compound recipe without Scolopendra;
- 2: the compound recipe without Scorpio;
- 3: the compound recipe without Bombyx Batryticatus;
- 4: the compound recipe without Eupolyphaga seu Steleophaga;
- 5: the compound recipe without Periostracum Cryptotympanae

Table 4 Class identification of unknown compound recipe sample (Critical Probability level : 0. 01)

	Residual	Model	Combined	Probability
FUFANG3	0. 802 7	0	0. 802 7	0. 934 3
FUFANG1	3. 556	6. 293	7. 228	0
FUFANG4	5. 794	6. 807	8. 94	0
FUFANG2	9. 641	7. 825	12. 42	0
FUFANG5	5. 843	12. 68	13. 96	0

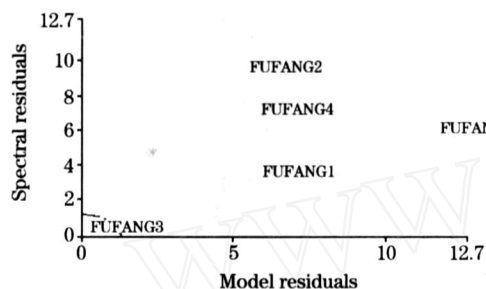


Fig 4 Classification results of unknown compound recipe sample

2.5 未知样品的预测

最后将重新配制的盲样进行分析,所有未知样本均都识别正确。经计算得出准确率为 90 %。如表 4 所示为某未知供试品的检测结果,同时给出了该未知供试品属于某类复方的可能性。图 4 所示为一未知供试品的聚类分析结果。图中横坐标为建模残差,纵坐标为光谱残差,图中弧线表示未知供试品落在该区域内,而 FUFANG3 的标准集合也在此区域内,所以可以认定该未知物为缺少僵蚕的复方,结果与实际是一致的。模型经过检验合格后,就可以用来对未知物进行分类。

3 结 论

总之,聚类分析法可以对小儿抽风散不同的缺味复方(蜈蚣、全蝎、僵蚕、土鳖虫和蝉蜕)进行快速鉴别。说明所建的模型基本上是合理的,并且有较强的识别能力。对于预测错误或不准确的样本,将谱图放大分析后,发现造成的原因主要是样本在扫描时出现操作上的问题,导致谱图不标准,为识别带来困难。为此,如果能够获得足够多的样本,增加训练集样本数,加强操作的标准程度,该法的准确率将会大大提高。因此,近红外光谱与聚类分析法相结合能无损快速识别中药复方的配伍组分。

参 考 文 献

- [1] LIU Hong-xia, SUN Su-qin, YANG Jun-shan(刘红霞,孙素琴,杨峻山). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(7): 1316.
- [2] CHEN Jian-bo, ZHOU Qun, SUN Su-qin, et al(陈建波,周群,孙素琴,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(8): 1493.
- [3] YANG Qun, WANG Yi-lin, YAO Jie, et al(杨群,王怡林,姚杰,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(12): 2219.
- [4] CHU De-ren, ZHOU Qun, YU Lu, et al(储德韧,周群,郁露,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(9): 1706.
- [5] LI Yan, WU Ran-ran, YU Bai-hua, et al(李燕,吴然然,于佰华,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(10): 1846.
- [6] WANG Li-qun, LI Ying-ming, ZHANG Li-ping, et al(王立群,李英明,张丽萍,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(6): 1061.
- [7] YU Lu, SUN Su-qin, ZHOU Qun, et al(郁露,孙素琴,周群,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(12): 2181.
- [8] A YIGULI ·Tashi, ZHOU Qun, DONG Xiao-ou, et al(阿依古丽·塔西,周群,董晓鸥,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(7): 1238.
- [9] LIU Shu-hua, ZHANG Xue-gong, ZHOU Qun, et al(刘沛华,张学工,周群,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(6): 878.
- [10] LIU Shu-hua, ZHANG Xue-gong, ZHOU Qun, et al(刘沛华,张学工,周群,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(4): 629.
- [11] PENG Yong, SUN Su-qin, ZHAO Zhong-zhen, et al(彭勇,孙素琴,赵中振,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(6): 679.
- [12] Perkin Elmer Corporation. SIMCA Procedures, 1996. 27.

FTIR and Classification Study on the Chinese Compound Recipe

WANG Jing-juan¹, ZHANG Gui-jun^{1*}, ZHOU Qun², SUN Si-qin^{2*}

1. School of Chinese Pharmacy, Beijing University of Chinese Medicine, Beijing 100102, China

2. Department of Chemistry, Tsinghua University, Beijing 100084, China

Abstract On the basis of building a model of main components and with the fingerprint infrared spectrum as the object, Xiaochoufeng powders were rapidly classified and closely studied by using the method of clustering analysis. Because of the tiny difference between the spectra of different compound recipes, including the compound recipe without scolopendra, the compound recipe without scorpio, the compound recipe without bombyx batryticatus, the compound recipe without eupolyphaga seu steleophaga and the compound recipe without periostracum cryptotympanae, the outcome of the clustering is ideal. When tested with the blind samples (which were picked out from the chosen samples) the accuracy of clustering reaches up to 90 %. In order to increase the clustering accuracy, the number of the samples should be increased and also these samples should be more characteristic. On the whole, with the combination of the fingerprint spectrum and the clustering analysis, the compound recipe can be distinguished rapidly.

Keywords FTIR spectroscopy; Xiaochoufeng Powder; Clustering analysis; Undamaged determination

* Corresponding author

(Received Feb. 8, 2007; accepted May 12, 2007)

中国化学会第 26 届学术年会(第一轮通知)

中国化学会第 26 届学术年会定于 2008 年 7 月 13~16 日在天津举行。会议由中国化学会主办,南开大学承办,天津师范大学、天津理工大学协办。

大会主题:化学与和谐社会

会议内容包括:大会特邀报告、分会邀请报告、专题报告与讨论、论文墙报展讲,同时设置专题学术论坛。会议期间还将组织展览展示。欢迎广大科技工作者积极参加,踊跃投稿。热烈欢迎相关企业、高校、科研院所利用此次契机,扩大影响,参与会展。

本次学术年设置绿色化学、环境化学、化学生物学等 20 个学术分会;专题论坛 1 个及新技术新仪器专场展示。其中分析化学为第 9 分会。

分析化学分会的征文范围

(1) 原子光谱分析方法;(2) 分子光谱分析方法;(3) 色谱法与分离技术;(4) 电分析化学与传感器;(5) 环境分析化学;(6) 生物分析化学;(7) 波谱法;(8) 质谱分析;(9) 分析仪器与联用技术。

征文联系及收集人:张毅

单位:南开大学化学系

地址:天津市卫津路 94 号

邮政编码:300071

电话:022 - 23506075

电子信箱:nkgcms@gmail.com