

DOI:10.3724/SP.J.1008.2010.00993

## 用于降糖类药物真伪快速检测的3种拉曼光谱处理方法的比较

宋清<sup>1,2</sup>, 翁欣欣<sup>1</sup>, 陆峰<sup>1\*</sup>

1. 第二军医大学药学院药物分析学教研室, 上海 200433

2. 解放军211医院药剂科, 哈尔滨 150080

**[摘要]** **目的** 比较3种用于药品真伪快速检测的拉曼光谱处理方法的适用性。**方法** 采用必达泰克便携式拉曼光谱仪测定6类降糖类药物对照品和相应的市售药品与12种常用制剂辅料的拉曼光谱, 采用上海技术物理研究所专利改进方法(MSFDSS)、局部直线筛选法(LSLS)和改进的局部直线筛选法(MLSLS)等3种处理方法对光谱进行处理。**结果** MLSLS法处理结果的灵敏度、专属性、阴性预报性和检出效率分别为96.77%、97.20%、99.29%和96.02%, 总体结果在3种处理方法中最佳。**结论** 发源于红外光谱的LSLS方法经过改进后, 可以用于药品真伪的拉曼光谱法快速检测。

**[关键词]** 局部直线筛选法; 拉曼光谱分析; 假药; 快检; 降血糖药

**[中图分类号]** R 927.1 **[文献标志码]** A **[文章编号]** 0258-879X(2010)09-0993-05

### Algorithms of Raman spectrum analysis in fast identification of counterfeit hypoglycemic agents

SONG Qing<sup>1,2</sup>, WENG Xin-xin<sup>1</sup>, LU Feng<sup>1\*</sup>

1. Department of Drug Analysis, School of Pharmacy, Second Military Medical University, Shanghai 200433, China

2. Department of Pharmacy, No. 211 Hospital of PLA, Harbin 150080, Heilongjiang, China

**[Abstract]** **Objective** To investigate the applicability of portable Raman spectroscopy for fast identification of counterfeit hypoglycemic agents. **Methods** The Raman spectra of 6 hypoglycemic agents, their corresponding reference standards and 12 common pharmaceutical excipients were detected by BW&Tek i-Raman portable spectrometer. The spectra were analyzed by three algorithms, namely, the minimized summation of first derivative of subtraction spectra algorithm, local straight-line screening (LSLS) algorithm, and modified LSLS (MLSLS) algorithm. **Results** Among the three algorithms, MLSLS algorithm showed the best performance, with the total sensitivity, specificity, accuracy and efficiency being 96.77%, 97.20%, 99.29%, and 96.02%, respectively. **Conclusion** The MLSLS algorithm, which is based on infrared spectroscopy, can be adapted to identify suspected counterfeit drugs based on portable Raman spectroscopy.

**[Key words]** local straight-line screening algorithm; Raman spectrum analysis; counterfeit drugs; fast detection; hypoglycemic agents

[Acad J Sec Mil Med Univ, 2010, 31(9):993-997]

假药不仅严重危害人民群众的健康, 而且破坏了医药产业的良性发展, 是当今世界各国特别是发展中国家共同面临的问题之一。传统的假药识别方法如化学反应、薄层色谱、液相色谱及联用技术均须破坏样品, 且耗时较长, 难以满足实际工作对药品快速检测的需求<sup>[1]</sup>。作为一种近年来发展迅速的光谱分析方法, 拉曼光谱有着独特的优势, 如光谱特征性较强, 测试样品非接触性、非破坏性、时间短, 样品所需量小及样品无需制备等<sup>[2-3]</sup>。

目前拉曼光谱用于药物真伪鉴别的方法有直观判别法、化学模式识别法和无模型判别法等。无模型分析方法不需要同质样品、阴性样品或模拟样品, 也不需要建立、验证模型的过程, 仅利用待测样品和相对应对照品各一张光谱信息, 经过数学和统计学解析, 对样品中可能存在的成分进行分析鉴定。该方法最初应用于中药掺杂样品的红外光谱分析<sup>[4-7]</sup>, 取得了较好的效果。本研究将上海技术物理研究所专利改进方法 MSFDSS(minimized summation of first deriva-

**[收稿日期]** 2010-05-02 **[接受日期]** 2010-06-24

**[基金项目]** 国家科技支撑计划(2008BAI55B06), 上海市技术标准专项(05dz05006)。Supported by National Scientific & Technology Supporting Program(2008BAI55B06) and Special Foundation for Technical Standard of Shanghai Municipality(05dz05006)。

**[作者简介]** 宋清, 硕士生。E-mail: song\_king@live.cn

\* 通讯作者(Corresponding author)。Tel: 021-81871261-83, E-mail: fenglufeng@hotmail.com

tive of subtraction spectra)、局部直线筛选法(local straight-line screening, LLSL)和改进的局部直线筛选法(modified LLSL, MLLSL)由红外光谱拓展应用到拉曼光谱,对样品(由原来的中药基质拓展到市售降糖药物、辅料和类似结构药物)中是否存在某种特定成分(由原来的掺杂成分拓展到主药成分)进行判别,以比较几种光谱处理方法在假药判别方面的适用性。

## 1 原理

1.1 上海技术物理研究所专利改进方法(MS-FDSS) 中国科学院上海技术物理研究所陆卫等<sup>[4]</sup>发明的专利“用红外光谱检测中药中化学药物的方法”单纯依靠掺杂化学药物的纯光谱与待检的可能发生掺杂的中药红外光谱,不经化学提取分离,无需模式识别与模型校正,直接得到其中的掺杂量。该方法按式  $S = \sum_x \{y[\ln T_{1(x)}]' - [\ln T_{2(x)}]'\}^2$  进行计算,其中  $T_{1(x)}$  为掺杂化学药物光谱,  $T_{2(x)}$  为待检中药光谱,  $x$  为波长或波数,  $S$  为对 2 个光谱的对数值微分的差值平方和,  $y$  为含量因子。  $S$  值关于  $y$  求最小值是一维最小值问题,该专利采用随机采样法求得  $y_m$ 。该专利确定一个临界值  $y_c$  (0.005 到 0.01),如果  $y_m$  值大于  $y_c$ ,则表示被测药物中含有某种特定成分,如果  $y_m$  值小于  $y_c$ ,则表明不含有。曹岩等<sup>[8]</sup>通过选择求导计算方法和扩大随机采样范围,对该专利方法做了改进,使得该方法的稳健性和有效性进一步提高。

本研究利用该改进的专利方法处理降糖药物、辅料与相应对照品的拉曼光谱数据,如果  $y$  值大于  $y_c$  值(设为 0.01),则被测样品判为阳性(真药);如果  $y$  值小于 0.01,则判为阴性(假药);如果  $y$  值等于 0.01,则判为可疑结果。

1.2 局部直线筛选法(LLSL) 根据观察,大部分的中药红外光谱都是变化平缓、尖锐特征信息少的图谱,在其中可以找到很多直线或近直线(相关系数  $r > 0.99$ )。而化合物的图谱则含有许多特征峰(一般在 50 个左右),是变化明显、急促的谱图。红外光谱本身具有加和性,如果化学药物在上述(近)直线谱段上有吸收峰的话,由于信号叠加,必然使这些直线段的线性降低,即  $|r|$  下降。反过来从掺杂中药光谱中差减一定比例的化合物光谱信息,这些线段就会回复(近)直线。因此可以通过筛选各个谱段上扣除比例的密集区域,确定(近)直线谱段的所在。密集区域所对应的扣除比例,即含量因子区间,也即化学药物的相对掺杂含量( $z_p$ )。该法的关键在于筛选化合物特征峰所对应的(近)直线谱段,即含量因子

(区间)的集中区域,从而估计真值所在范围,因此将该法命名为“局部直线筛选法”<sup>[5-6,9]</sup>。

拉曼光谱与红外光谱一样具有光谱特征信息丰富的特点。本研究利用该方法处理降糖药物、辅料与相应对照品的拉曼光谱数据,如果  $z_p$  值大于 0,则被测样品判为阳性(真药),如果  $z_p$  值小于或等于 0,则被测样品判为阴性(假药)。

1.3 局部直线筛选法的改进(MLLSL) 本研究对 LLSL 方法进行一系列的改进,使其更适应便携式拉曼光谱(而非红外光谱)的特点,更适应样品基质是化药辅料(而非中药基质)的特点,也更适应真药的活性药物组分(active pharmaceutical ingredient, API)含量已知(而非掺杂物含量未知)的特点,从而得到更为稳健合理的处理结果。改进算法包含 3 个模块:半数峰比较模块、改进的 LLSL 模块、主成分分析模块。

1.3.1 半数峰比较模块 本模块将先依照原始的 LLSL 算法对待检的对照化合物进行选峰,然后计算这些化合物峰所对应的差减之后的理论基线光谱(对应的待测药片拉曼光谱的峰—主药成分对照品对应的峰  $\times$  主药成分含量  $P$ ),再对这些理论基线光谱上近直线的峰的计数( $n'$ ),最后将  $n'$  与待检化合物的峰的个数的一半( $n$ )作比较,如果  $n' < n$ ,则将其判为假药。相反,如果  $n' \geq n$ ,则将其判为可能的真药样品,需要进行进一步的计算处理。

1.3.2 改进的 LLSL 模块 该模块通过小波去噪、插值、峰位校正、含量区间平移、含量区间加权对光谱数据进行处理来实现<sup>[10]</sup>。选择  $1730 \sim 175 \text{ cm}^{-1}$  的光谱谱段,用 MLLSL 进行计算,若计算结果  $\leq 0$ ,直接被判定为假药(可疑样品除外),其他样品将进行下一模块的考察。可疑样品是在改进的 LLSL 模块中计算产生的,一般分为 2 种情况:一种是含量区间图的最高平台部分覆盖 0 点;另一种情况是含量区间图在 0 点的两侧各有一个或多个近似的最高平台(图略)。

1.3.3 主成分分析模块 将以上模块中被判定为可疑样品或者可能的真药样品与已知此类药品的真药样品进行主成分分析。以主成分 1 与主成分 2 分别为横纵坐标作主成分分析图(图略),包含可疑样品和可能的真药样品在内的所有数据点在图中表现为集中点和离散点,将图上的离散点判别为假药,这样样品在上述 3 个模块都有可能被判别为假药。若可疑样品出现在集中点处,仍视其为可疑样品,留待进一步检测,出现在集中点处的其他样品则判别为真药样品。

半数峰模块和主成分分析模块的引入,进一步

提高了结果的准确度,降低了假阳性率和假阴性率。

## 2 材料和方法

**2.1 试剂和样品** 格列吡嗪、格列齐特、格列本脲、格列美脲、格列喹酮、吡格列酮对照品均由中国药品生物制品检定所提供;样品中包括市售不同厂家和批号的6类共32种降糖药(均为片剂,见表1)和12种常用制药辅料(包括L-羟丙基纤维素、聚乙烯吡咯烷酮-PVPK30、淀粉、糊精、滑石粉、羟丙基甲基纤维素、乳糖、羧甲基淀粉、微晶纤维素、硬脂酸镁、预胶化淀粉和蔗糖)。相关的药品与辅料均由山东省食品药品检验所提供。

表1 降糖类药品列表

Tab 1 Detailed information of hypoglycemic agents

Sample	Type and content of the API
Glipizide tablet	7 types (2.3%)
Gliclazide tablet	6 types (2 types: 29%, 4 types: 48%)
Glibenclamide tablet	6 types (3.3%)
Glimepiride tablet	5 types (3.2%)
Gliquidone tablet	2 types (12% and 26%)
Pioglitazone tablet	6 types (3 types: 12%, 1 type: 15%, 1 type: 18%, 1 type: 21%)

API: Active pharmaceutical ingredient

**2.2 光谱采集与信息处理** 实验所用仪器为必达泰克 i-Raman 系列 BWS415-785H 型拉曼光谱仪。光谱覆盖范围  $175\sim 3\,200\text{ cm}^{-1}$ ,分辨率为  $3\text{ cm}^{-1}$ 。采用热电制冷温控装置(TEC)和  $2\,048$  像素的 CCD

检测器。分别测定6种对照品和32种药片以及12种制剂辅料的拉曼光谱(药片去除包衣,刮除表面后采集,对照品和制剂辅料则直接采集)。激发光波长为  $785\text{ nm}$ ,积分时间为  $5\,000\text{ ms}$ 。将采集的光谱数据截取  $175\sim 1\,730\text{ cm}^{-1}$  波数范围纳入计算,3种算法均由 MATLAB 编程实现。

## 3 结果和讨论

本研究模拟的假药系统有2种,第一种假药系统是用12种常用的制药辅料分别冒充每一种降糖药片,用3种算法共针对6种对照品与44个样品进行运算,这样就产生了  $12\times 6=72$  个阴性样品,加上真药样品32种,MSFDSS 和 LSLs 在这一部分的样品数均为104个,在MLSLS中,主药含量在该算法中是一个非常重要的参数,由于API含量信息的引入,MLSLS算法有132个阴性样品,在这一部分的所有样品数为164个;在格列系列降糖药片中,格列本脲片较为便宜,而格列吡嗪片和格列美脲片价格略高,而这3种药片中主药成分的结构非常相似,所以本研究模拟的另一种假药系统是用格列本脲片来冒充格列吡嗪片或者格列美脲片作为研究对象,这样,每一种算法增加了  $6+6=12$  种阴性样品;MSFDSS、LSLS 和 MLSLS 的样品总数分别为116、116和176个。为了更好地评价各方法在上述的408个计算结果总量的表现,按照前文所述判别标准,将常用于方法评价的各项指标来评价、考察方法,结果见表2。

表2 3种方法计算结果的效能评价

Tab 2 Performance evaluation of the three algorithms

Performance	MSFDSS	LSLS	MLSLS			Subtotal
			API content > 5%	API content 3%-5%	API content 0-3%	
Total (N)	116	116	110	47	19	176
Positive-total (n)	32	32	14	11	7	32
Negative-total (n)	84	84	96	36	12	144
True negative (n)	39	39	96	34	9	139
True positive (n)	32	17	14	10	6	30
False negative (n)	0	15	0	1	0	1
False positive (n)	40	45	0	1	3	4
Suspect (n)	5	0	0	1	1	2
Rate of false negative (%)	0.00	46.88	0.00	9.09	0.00	3.13
Rate of false positive (%)	47.62	53.57	0.00	2.78	25.00	2.78
Sensitivity (%)	100.00	53.13	100.00	90.91	100.00	96.77
Specificity (%)	49.37	46.43	100.00	97.14	75.00	97.20
Accuracy (%)	100.00	72.22	100.00	97.14	100.00	99.29
Efficiency (%)	61.21	48.28	100.00	93.62	78.95	96.02

The definition of each performance was recorded in the references<sup>[11-12]</sup>. MSFDSS: Minimized summation of first derivative of subtraction spectra; LSLs: Local straight-line screening; MLSLS: Modified LSLs; API: Active pharmaceutical ingredient

评价方法考察了3种方法的4个指标——灵敏度、专属性、阴性预报性及检出效率。对于化学药物快速检测来说,一个合适的初筛方法应尽量保证足够高的专属性,即达到较低的假阳性率,以实现低漏检率,在此前提下兼顾较低的假阴性率,以及尽量少的可疑样品(降低后续筛选的工作量)。

从专属性来看,MLSLS算法里的API含量>5%的样品达到了100.00%,没有假阳性结果,在API含量为3%~5%和0%~3%区间分别出现了1个和3个假阳性结果,总体专属性达到了为97.20%,而前2种方法的专属性只有49.37%和46.43%,在专属性上,MLSLS表现最优,实现了低漏检率。

在灵敏度和阴性预报性方面,MSFDSS由于没有假阴性结果,所有实际阳性样品全部判别为阳性结果,灵敏度和阴性预报性结果都为100.00%,MLSLS算法里的API含量>5%和在0%~3%区间的样品也达到了100.00%的灵敏度和阴性预报性,没有假阴性结果,在API含量为3%~5%区间只出现了1个假阴性结果,总体灵敏度和阴性预报性分别达到了为96.77%和99.29%,在这两方面,MSFDSS最优,MLSLS算法稍逊。

在检出效率方面,MLSLS算法达到了96.02%,远远超过了前2种算法的61.21%和48.28%,所有结果与实际样品相符的检出率最高;在可疑样品方面,MSFDSS以临界点确定为可疑样品的情况下,出现了5个可疑值,LSLS算法未定义可疑样品,而MLSLS在经过PCA模块分析之后,剩余可疑值总数为2个,较MSFDSS算法可疑样品数要少。

总体来看,MSFDSS虽然假阴性结果为零,但其假阳性率较高,假药漏检率较大,此为假药初筛方法的大忌;而LSLS的各评价指标表现平平,表明该算法不能直接应用于拉曼光谱-化学药物这一假药检测系统;MLSLS算法由于对原始LSLS算法进行了一系列改进优化,引入了API含量信息与半数峰比较模块、主成分分析模块,在评价指标的各个方面总体结果最优,可以看出,MLSLS对3种API含量样品的结果来说,灵敏度和阴性预报性区别不大,但专属性和检出效率就随着API含量的下降有明显的降低趋势,体现出此算法对主药含量信息的敏感性。

为了进一步考察3种算法的稳健性,在首次采集数据3个月后,用同样的仪器重新采集了同批次样品光谱,结果见表3。

表3 3种方法应用于3个月后光谱数据计算结果的效能评价  
Tab 3 Performance evaluation of the three algorithms after three months

Performance	MSFDSS	LSLS	MLSLS			Subtotal
			API content > 5%	API content 3%-5%	API content 0-3%	
Total (N)	116	116	110	47	19	176
Positive-total (n)	32	32	14	11	7	32
Negative-total (n)	84	84	96	36	12	144
True negative (n)	54	43	91	34	8	133
True positive (n)	32	20	12	7	7	26
False negative (n)	0	12	2	3	0	5
False positive (n)	27	41	5	2	2	9
Suspect (n)	3	0	0	1	2	3
Rate of false negative (%)	0.00	37.50	14.29	27.27	0.00	15.63
Rate of false positive (%)	32.14	48.81	5.21	5.56	16.67	6.25
Sensitivity (%)	100.00	62.50	85.71	70.00	100.00	83.87
Specificity (%)	66.67	51.19	94.79	94.44	80.00	93.66
Accuracy (%)	100.00	78.18	97.85	91.89	100.00	96.38
Efficiency (%)	74.14	54.31	93.64	87.23	78.95	90.34

MSFDSS: Minimized summation of first derivative of subtraction spectra; LSLS: Local straight-line screening; MLSLS: Modified LSLS; API: Active pharmaceutical ingredient

由结果可以看出,MSFDSS和LSLS的计算结果和3个月前相比,各项指标有所提高,但其绝对数值仍处在较低的水平;相比之下,MLSLS的每个指标的数值尽管有所下降,可疑样品数有所增加,但总

体上4个指标还是维持在比较高的水平,基本能够满足假药快速检测的要求。

由于拉曼光谱自身的特点和仪器硬件技术的不断完善,拉曼光谱在假药快检中已经表现出了很好

的应用前景。本研究采用 3 种光谱处理方法,对降糖类药物的拉曼光谱进行处理,按各自标准对计算结果进行判别。结果表明,发源于红外光谱的 LSLS 方法经过改进后,使得计算结果具有较高的灵敏度、专属性、阴性预报性及检出效率,可以用于拉曼光谱法对化学药物真伪的快速鉴别。本研究用 3 个月后的光谱数据模拟现实应用的情况,其计算结果满意度虽然在一定程度上有所下降,但总体仍相对稳定。如何进一步消除其他潜在因素的干扰,使该方法更加成熟、稳健,仍在继续改进中。

[参考文献]

[1] 张 新,刘朝霞,倪坤仪,张启明. 便携式拉曼光谱仪用于部分头孢菌素类药品识别的初步研究[J]. 中国药事, 2008, 22: 555-557.

[2] 伍 林,欧阳兆辉,曹淑超,易德莲,秦晓蓉,孙少学,等. 拉曼光谱技术的应用及研究进展[J]. 光散射学报, 2005, 17: 180-186.

[3] 王 玉,李忠红,张正行,安登魁. 拉曼光谱在药物分析中的应用[J]. 药学学报, 2004, 39: 764-768.

[4] 陆 卫,王依婷,朱自强,李志锋,王金斌,陈贵良,等. 用红外光谱法检测中药中化学药物的方法: 中国, CN 1487280A [P].

2004-04-07.

[5] Lu F, Li S, Le J, Chen G, Cao Y, Qi Y, et al. A new method for testing synthetic drugs adulterated in herbal medicines based on infrared spectroscopy[J]. Anal Chim Acta, 2007, 589: 200-207.

[6] 李 树,曹 岩,乐 健,陈桂良,柴逸峰,陆 峰. 局部直线筛选法检测中药非法添加化学药物的研究[J]. 光谱学与光谱分析, 2009, 29: 327-330.

[7] 李 树. 中药非法添加化学物质快速检测新方法的研究[D]. 上海: 第二军医大学, 2007.

[8] 曹 岩,李 树,陆 峰. 红外光谱法测定中药掺杂化学药品的方法改进[J]. 药学实践杂志, 2007, 25: 292-294.

[9] 陈艳华. 中药非法掺伪快速检测的方法学研究[D]. 上海: 第二军医大学, 2007.

[10] Weng X, Yin L, Chai Y, Lu F. Model-free analysis of counterfeit drugs based on portable Raman spectroscopy[J]. Appl Spectr, (revised No. 10-05986).

[11] 朱向荣,李 娜,史新元,乔延江,张卓勇. 最小二乘支持向量机算法与紫外光谱法用于鉴别清开灵注射液四混中间体[J]. 分析化学, 2008, 36: 770-774.

[12] 陈艳华,陆 峰,尹利辉. 局部直线筛选法的建立及其改进研究[J]. 中国科学 B 辑, 2010, 40: 1142-1148.

[本文编辑] 尹 茶

• 更正启事 •

关于《奈福泮复合舒芬太尼对混合痔患者术后镇痛及应激反应的影响》一文的更正

本刊 2010 年 31 卷第 8 期 918-920 页刊登的孙琰等的论文《奈福泮复合舒芬太尼对混合痔患者术后镇痛及应激反应的影响》因作者的疏忽,表 3 中的数据出现差错,现作如下更正(论文网络版已进行了相应更正)。

表 3 PCIA 镇痛组患者各时间段舒芬太尼的消耗量以及 PCA 按压次数

(n=20,  $\bar{x} \pm s$ )

时间点	舒芬太尼的消耗量 $m/\mu\text{g}$		实际 PCA 按压次数	
	B 组	C 组	B 组	C 组
0~2 h	3.44±0.79	2.75±0.53**	2.1±0.6	1.2±0.4**
2~4 h	3.56±0.89	2.83±0.48**	2.0±0.5	1.3±0.5**
4~6 h	3.67±0.95	3.07±0.67**	2.2±0.4	1.5±0.8**
6~8 h	3.76±0.78	3.28±0.82*	2.3±0.6	1.7±0.9*
8~12 h	6.64±1.59	5.66±1.12*	4.3±0.9	3.2±1.6*
12~24 h	22.67±2.36	17.59±1.85*	5.1±2.7	4.8±2.4*
24~48 h	39.87±3.65	33.76±2.34*	12.1±4.2	9.7±3.6*
48 h 总量	82.35±2.82	69.74±1.93**	28.2±6.9	19.3±4.6**

\*  $P < 0.05$ , \*\*  $P < 0.01$  与 B 组相比