

白芍水提过程中芍药苷含量的 NIRS 快速测定研究

熊明华¹, 方少敏², 饶毅^{3*}, 王信², 刘安⁴, 田紫平⁵

(1. 江西省宜春市人民医院, 江西 宜春 336000; 2. 江西中医学院, 南昌 330004;

3. 中药固体制剂制造技术国家工程研究中心, 南昌 330006;

4. 中国中医科学院中药研究所, 北京 100700; 5. 贵州益佰制药股份有限公司, 贵阳 550008)

[摘要] 目的: 基于近红外光谱(NIRS)技术的, 建立白芍提取过程的快速质量控制方法。方法: 针对白芍提取过程, 以 HPLC 测定提取液样品中芍药苷含量, 同时采集其 NIRS, 采用偏最小二乘(PLS)回归法建立 NIRS 与 HPLC 分析值之间多元校正模型。结果: 校正模型的内部交叉验证决定系数 R^2 为 99.86, 内部交叉验证均方差 RMSECV 为 0.023 9。对提取过程中的白芍样本进行预测, 结果令人满意。结论: NIRS 可准确预测白芍提取过程中芍药苷含量, 可推广运用于中药提取过程的快速、实时质量监控。

[关键词] 近红外光谱; 白芍; 芍药苷; 提取过程; 偏最小二乘法

[中图分类号] R284.1 **[文献标识码]** A **[文章编号]** 1005-9903(2011)21-0052-05

Study for Rapid Determination of the Paeoniflorin Content in Water Extracting Process of Paeoniae Alba by NIRS

XIONG Ming-hua¹, FANG Shao-min², RAO Yi^{3*}, WANG Xin², LIU An⁴, TIAN Zi-ping⁵

(1. Yichun Hospital, Yichun 336000, China; 2. Jiangxi University of Chinese Traditional

Medicine, Nanchang 330004, China; 3. The National Pharmaceutical Engineering Center for Solid Preparation in Chinese Herbal Medicine, Nanchang 330006, China;

4. Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China;

5. Guizhou Ebay Pharmaceutical Co., Ltd., Guiyang 550008, China)

[Abstract] **Objective:** To set up a novel method for fast quality control of Paeoniae Alba's extracting process based on near infrared spectroscopy (NIRS). **Method:** Aiming at the process of extracting, the content of paeoniflorin was determined by HPLC as the reference values, and the NIRS data were collected, and then partial least square (PLS) regression was applied to build the calibration model of paeoniflorin. **Result:** RMSECV of the model for paeoniflorin was 0.023 9, the correlation coefficient of the prediction samples between the predictive value and the real value was 99.86. The calibration model was used for testing samples, and the results were satisfied. **Conclusion:** NIRS can accurately determine the content of paeoniflorin, and indicates potential to be widely used in the extracting process of traditional Chinese medicine.

[Key words] NIRS; Paeoniae Alba; paeoniflorin; extracting process; PLS

在中药制药过程中, 提取过程作为中药生产的

特色环节之一, 决定着药品疗效。目前, 我国中药提取过程的质量控制仍十分被动, 只对终端产品进行质检, 很少考虑工艺过程中的成分变化, 生产工艺很难得到精确控制, 从而导致中药产品批次间质量差异较大, 这也是中药难以现代化、国际化的主要原因之一。近红外光谱 (near-infrared spectroscopy, NIRS) 技术发展迅速, 引人注目, 已在农业、烟草、石

[收稿日期] 20110123(003)

[基金项目] “重大新药创制”科技重大专项(2009ZX09313-030)

[通讯作者] * 饶毅, 教授, 研究方向为中药复方质量控制, Tel:0791-7119609, E-mail:raoyi99@126.com

油化工、食品和药物等各领域得到广泛应用,具有分析速度快、操作方便、无污染、无破坏性、一般不需前处理,并特别适合于过程分析等特点^[1],在中药生产过程关键环节的质量控制中具有巨大的应用潜力^[2-5]。白芍为毛茛科植物芍药 *Paeoria lactiflora* Pall. 的干燥根,具有养血调经、敛阴止汗、柔肝止痛的功效^[6],是一种常用中药。本文以白芍为研究对象,运用 NIRS 技术对其提取过程进行快速质量控制研究,从而为实现白芍产业化提取过程的在线检测提供实验依据。

1 仪器与试剂

MPA 傅立叶变换近红外光谱仪(德国 BRUKER 公司),液体样品池为 2 mm 光程的石英比色皿,TE-InGaAs 检测器。BRUKER OPUS 软件中的定量分析软件(OPUS/QUANT)。Waters 2695 高效液相色谱仪,Waters 2996 DAD 检测器。3 批白芍饮片均由宁波立华制药有限公司提供,批号分别为 20091110,20091025,091104,产地分别为浙江、浙江、河南,在实验中分别记为 A,B,C。芍药苷对照品(中国药品生物制品检定所,批号 110736-200732,供 HPLC 含量测定用),乙腈为色谱纯,其余为分析纯。

2 方法

2.1 白芍提取液的制备 称取白芍饮片约 150 g,置于 3 000 mL 三颈烧瓶中,加 12 倍量水,加热回流,待水开始沸腾时,开始第 1 次取样,取约 8 mL 左右至样品瓶中。之后每隔 8 min 取样 1 次,煎煮 2 h,得 16 份样品。取一煎液残渣,加 8 倍量水进行第 2 次煎煮,待水开始沸腾时,开始第 1 次取样,之后每隔 10 min 取样 1 次,煎煮 2 h,得 13 份样品。每批白芍如此提取 2 次,得 29 份样品,3 批白芍提取得 A,B,C 3 组,每组 29 份样品,共 87 份样品用于近红外分析。

2.2 光谱采集 测定方式样品腔透射(2 mm 光程),扫描范围 12 500 ~ 4 000 cm^{-1} ,分辨率 8 cm^{-1} ,扫描次数 64 次,环境温度 17.2 ~ 18 $^{\circ}\text{C}$,湿度 44% ~ 52%。

取各个编号下的样品依次倒入样品瓶中,按实验条件依次进行近红外扫描,每个样品测定 3 次,取平均光谱用于近红外建模,共得 87 张光谱图。样品光谱图见图 1。

2.3 芍药苷的 HPLC 分析^[6] 精密量取各样品溶液 1 mL 至 10 mL 量瓶中,加水稀释至刻度,摇匀,按

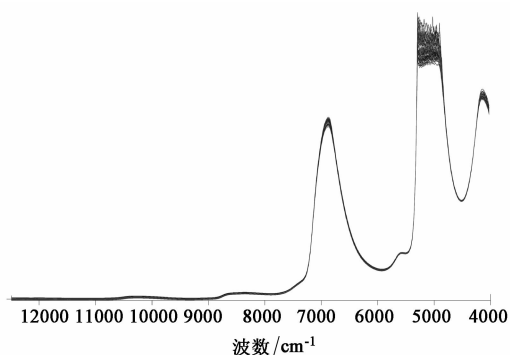


图 1 白芍水提液的近红外光谱叠加图

2010 年版《中国药典》白芍项下含量测定方法,采用 HPLC 测定样品中芍药苷含量。结果 87 份样品中芍药苷的浓度分布范围为 A 组 0.252 ~ 1.767 $\text{g}\cdot\text{L}^{-1}$, B 组 0.163 ~ 1.268 $\text{g}\cdot\text{L}^{-1}$, C 组 0.306 ~ 2.109 $\text{g}\cdot\text{L}^{-1}$,含量测定结果的分布图见图 2。

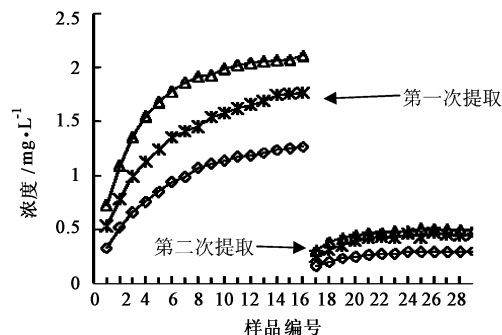


图 2 87 份样品中芍药苷含量测定结果分布

3 结果与分析

将芍药苷的名称、HPLC 含量测定值和提取过程样品的 NIRS 输入 OPUS QUANT 定量分析软件包中,进行数据处理,采用偏最小二乘(partial least square, PLS)回归法建立定量校正模型,根据 HPLC 测定的浓度梯度分布情况,选择 B,C 组样品用于建模,A 组样品作为预测集,注意所选的建模样品要能涵盖预测样品的浓度范围(0.252 ~ 1.767 $\text{g}\cdot\text{L}^{-1}$),即能代表组分的浓度信息。

3.1 光谱区间的选择 水分子所含的 OH 基,在近红外谱区的 6 944,5 155 cm^{-1} 附近有很强的合频与倍频吸收谱带,而其他各种物质分子的倍频与合频吸收相对较弱,因此在水溶液物质体系中,水就成为近红外光谱分析时的强干扰因素。所以一般采用厚度较薄的样品池和较高浓度的溶液,同时在建模过程中避开水分的强干扰吸收波段进行建模。本文中采用了光程仅为 2 mm 的样品腔,同时对多个波段

进行了建模试验。波段选取对所得模型的 RMSECV 的影响见表 1。

表 1 不同光谱波段的建模

光谱波段/ cm^{-1}	R^2	RMSECV
12 493.2 ~ 3 999.8	24.75	0.552
11 135.5 ~ 7 741.2	98.42	0.080 1
6 101.9 ~ 5 446.2	99.74	0.032 7
8 987 ~ 7 301.5, 6 283.2 ~ 5 457.8	99.86	0.023 9
12 493.2 ~ 7 494.3, 6 101.9 ~ 5 442.4	99.43	0.048 1

从表 2 可见,随着水的 2 个强吸收带依次纳入 PLS 建模谱段,模型的精度没有提高, R^2 和 RMSECV 极不理想。当合理去除水在 6 944,5 155 cm^{-1} 的 2 个强吸收波段时,模型的精度得到显著提高。通过比较,本文最终选定 8 987 ~ 7 301.5, 6 283.2 ~ 5 457.8 cm^{-1} 为最佳建模波段,从而避免了过多的冗余信息和干扰信息,改善模型性能,提高计算速度,HPLC 与 NIR 预测值之间的相关性 R^2 达到 99.86, RMSECV = 0.023 9。

表 2 不同光谱预处理方法的建模

光谱预处理方法	R^2	RMSECV
消除常数偏移量	98.97	0.064 8
矢量归一化	99.52	0.044 2
一阶导数	99.81	0.027 9
二阶导数	99.55	0.042 8
一阶导数 + 矢量归一化	99.86	0.023 9

3.2 光谱预处理方法的选择 在 NIRs 建模过程中,常需选择一定的预处理方法,以消除光谱测量过程中由外界因素带来的干扰和噪声。比较研究了 OPUS 软件中附带的多种光谱预处理方法,不同预处理方法对模型效果的影响见表 2。通过比较,可得知,原始光谱经“一阶导数 + 矢量归一化”处理后所得的 R^2 和 RMSECV 最理想,分别为 99.86, 0.023 9, 可以很好地反映样品信息。因此本实验采用对原始光谱进行“一阶导数 + 矢量归一化”预处理方法。

3.3 主因子数(维数)的选择 在相同波长范围和相同数据预处理方法条件下,不同维数的选择会对相关系数 R^2 和内部交叉验证均方差 RMSECV 造成直接的影响。本实验采用留一交互验证法得到 RMSECV 选择维数。从图 3 可以看出,随着维数的增加,校正模型性能在逐渐地提高, RMSECV 不断减小。但是当维数大于 6 时,模型性能没有更好地提

高,同时过多的数据维数也会降低模型的计算效率,因此,本文采用软件推荐的维数为 6。

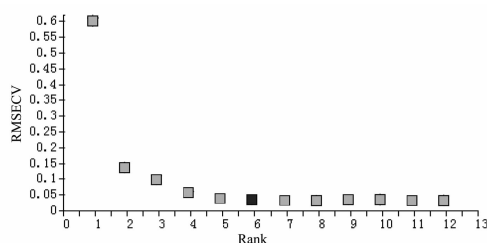


图 3 RMSECV 与维数的关系

3.4 模型的建立与验证

3.4.1 白芍提取过程芍药苷定量模型的建立 用校正样品集进行内部交叉验证,得到最终优化的模型结果为:原始图谱在波长范围 8 987 ~ 7 301.5 cm^{-1} 和 6 283.2 ~ 5 457.8 cm^{-1} 内,经“一阶导数 + 矢量归一化”处理后,选定维数为 6,建立白芍提取过程芍药苷的 PLS 定量校正模型,结果如图 4 所示, $R^2 = 99.86$, RMSECV = 0.023 9。结果表明样品光谱与其中芍药苷含量之间存在良好的相关性。

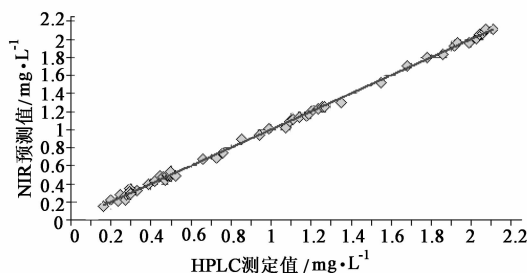


图 4 NIR 预测值与 HPLC 测定值之间的相关性

3.4.2 NIRs 模型的预测 采用上述所建模型来对 A 组预测集样品进行预测,所得结果见表 4。其相关系数 $r = 0.998 9$, 预测均方差 RMSEP = 0.080 8, 预测相对标准偏差 RSEP = 7.33%。

由表 4 可以看出,虽然 B, C 两组样品中芍药苷的浓度范围覆盖了 A 组样品,但考虑到实验所用白芍分别为 3 个不同批次,从而致使 B, C 两组白芍样品难以充分代表 A 组样品的组成信息,而这也是导致上述数据的相对偏差和 RSEP 值偏大的原因所在,因此分别从每组 29 个样品中挑选部分样品组成验证集,确保所选验证样品集更具代表性,且含量范围处于校正样品集含量范围之内。为避免最高值(C-16)和最低值(B-17)被选入验证集,对输入 OPUS 软件中的样品光谱进行数据类型设置,按 3:1 的比例划分校正集和验证集,设定第一测试样品从 A-3 开始,数据长度(测试样品)为 1,间距(校正样

表 4 模型 1 的 NIR 预测

样品号	HPLC 测定值 /g·L ⁻¹	NIR 测得值 /g·L ⁻¹	偏差	相对偏差 /%
A-1	0.533	0.521	0.012	2.32
A-2	0.780	0.744	0.035	4.55
A-3	0.992	0.914	0.078	7.85
A-4	1.130	1.013	0.118	10.39
A-5	1.242	1.133	0.110	8.81
A-6	1.355	1.249	0.106	7.84
A-7	1.407	1.300	0.107	7.57
A-8	1.452	1.380	0.072	4.98
A-9	1.539	1.446	0.093	6.07
A-10	1.577	1.482	0.096	6.07
A-11	1.622	1.481	0.140	8.66
A-12	1.656	1.527	0.129	7.79
A-13	1.690	1.573	0.117	6.93
A-14	1.744	1.608	0.136	7.79
A-15	1.757	1.647	0.111	6.29
A-16	1.767	1.669	0.098	5.52
A-17	0.252	0.200	0.052	20.62
A-18	0.314	0.254	0.061	19.36
A-19	0.355	0.317	0.038	10.60
A-20	0.400	0.354	0.046	11.62
A-21	0.427	0.377	0.050	11.62
A-22	0.437	0.391	0.045	10.38
A-23	0.426	0.414	0.013	2.93
A-24	0.461	0.396	0.064	13.95
A-25	0.431	0.433	-0.002	-0.48
A-26	0.474	0.463	0.011	2.23
A-27	0.451	0.442	0.010	2.13
A-28	0.441	0.470	-0.029	-6.52
A-29	0.452	0.437	0.014	3.13

表 5 模型 2 的 NIR 预测

样品号	HPLC 测定值 /g·L ⁻¹	NIR 测得值 /g·L ⁻¹	偏差	相对偏差 /%
A-3	0.992	0.972	0.020	2.03
A-6	1.355	1.339	0.016	1.15
A-9	1.539	1.549	-0.010	-0.64
A-12	1.656	1.670	-0.014	-0.86
A-15	1.757	1.758	-0.001	-0.06
A-18	0.314	0.304	0.010	3.23
A-21	0.427	0.390	0.037	8.57
A-24	0.461	0.420	0.041	8.82
A-27	0.451	0.444	0.007	1.58
B-1	0.328	0.368	-0.040	-12.30
B-4	0.759	0.741	0.018	2.37
B-7	0.987	1.023	-0.036	-3.62
B-10	1.141	1.136	0.005	0.41
B-13	1.211	1.208	0.003	0.26
B-16	1.268	1.257	0.011	0.89
B-19	0.233	0.218	0.015	6.51
B-22	0.273	0.233	0.040	14.79
B-25	0.293	0.260	0.033	11.12
B-28	0.294	0.297	-0.003	-0.88
C-2	1.092	1.091	0.001	0.06
C-5	1.678	1.697	-0.019	-1.14
C-8	1.915	1.918	-0.002	-0.13
C-11	2.023	2.001	0.023	1.11
C-14	2.065	2.073	-0.007	-0.36
C-17	0.306	0.316	-0.010	-3.24
C-20	0.447	0.480	-0.033	-7.37
C-23	0.483	0.494	-0.011	-2.33
C-26	0.501	0.492	0.009	1.77
C-29	0.499	0.507	-0.008	-1.70

品)为2,按照上述建模方法建立校正模型,记为模型2,其预测结果见表5所示,此时其相关系数 $r = 0.9994$,预测均方差 $RMSEP = 0.021$,预测集样品的预测相对标准偏差 $RSEP = 1.91%$,远小于上述模型7.33%的预测相对标准偏差,模型性能大大改善,预测精度大大提高。这说明如果实验中获取的样本量足够多,代表性足够强的话,将能使所建模型的适用性能增强,提高预测结果的准确性。

4 讨论

代表性的建模样本的选择是建立优秀的 NIRS

定量校正模型的关键,所谓“代表性”就是校正样品集必须包括今后待测样品的全部背景信息^[8]。本实验对3个不同批次的白芍提取过程进行NIRS分析,通过划分不同校正集和验证集,建立模型1和模型2。其中模型2对验证样品预测的 $RSEP = 1.91%$,远小于模型1的7.33%的预测相对标准偏差。因为在NIRS分析时,必须用足够多的光谱数据点才能区别光谱中的重叠信息。对于中药等复杂体系,在其提取过程中变化的不仅仅只有目标成分,还有其他

许多未知成分等,虽然光谱波长范围筛选可以提高目标成分与 NIRS 的相关性,但其他成分的变化始终会对目标成分的分析带来一定的干扰。因此 NIRS 反映的不仅仅是单个成分的吸收特性,还有样品中整体化学成分的组成信息。本实验所用白芍为 3 个不同批次,不同批次白芍在芍药苷和其他化学成分的组成背景信息上也存在着差异,这就使得不同批次样品的 NIRS 或多或少存在着偏差。模型 2 分别从每个批次白芍提取样品中选择一定样品组成验证集,相比之模型 1,模型 2 中校正集样品已经包含了 3 个批次白芍的化学组成信息,更能充分代表验证集样品的背景信息,所以模型 2 的稳定性和适用性较模型 1 更好。虽然本文用于建模的样本数偏少,但所建立的 NIRS 校正模型预测性能结果令人满意,展现出了良好的应用前景,具有较强的实用性。通过今后进一步充实不同产地和不同批次白芍药材信息,提高建模用样品的代表性,进一步优化模型,建立更加合理、准确、有效的校正模型,NIRS 可以满足实际白芍提取过程芍药苷含量快速检测的需要。

本文采用 NIRS 技术,对白芍水提过程中芍药苷含量变化进行快速分析研究,结果显示样品光谱在“8 987 ~ 7 301. 5, 6 283. 2 ~ 5 457. 8 cm^{-1} ”进行“一阶导数 + 矢量归一化”处理后,白芍水提过程中芍药苷的吸收特性可以得到很好地体现,运用 PLS 建立起满意的芍药苷定量分析模型。相比于 HPLC 分析方法,NIRS 能立即得出有效成分的含量,准确反映有效成分的变化趋势,让我们更好地理解制药工艺

过程,是一种极有前途的在线检测技术。研究表明,NIRS 作为一种快速、简便、无污染的分析方法,可用于白芍提取过程的快速分析。同时芍药苷作为白芍中的主要成分之一,在白芍制药工艺过程中常被用来作为质控指标^[8],因此本文研究结果也为白芍产业化提取过程的在线检测奠定了一定的基础。

[参考文献]

- [1] 陆婉珍. 现代近红外光谱分析技术[M]. 北京:中国石化出版社,2006:9.
- [2] 严诗楷,罗国安,王义明,等. 栀子药材提取工艺的红外光谱实时控制方法研究[J]. 光谱学与光谱分析, 2006,26(6):1026.
- [3] 杨南林,程翼宇,瞿海斌. 一种用于中药纯化过程的近红外光谱分析新方法[J]. 化学学报, 2003, 61(5):742.
- [4] 瞿海斌,李斌,刘雪松,等. 红参醇提液浓缩过程近红外光谱在线分析方法[J]. 中国药学杂志, 2005, 40(24):1897.
- [5] 章顺楠,杨海雷,刘占强,等. 近红外光谱法在线监测复方丹参滴丸料液中有效成分含量[J]. 药物分析杂志,2009,29(2):192.
- [6] 中国药典. 一部[S]. 2010:97.
- [7] 郭兰萍,黄璐琦. 近红外光谱技术及其在中药道地性研究中的应用[J]. 中国中药杂志. 2009,34(14):1755.
- [8] 李衫,王亚楠,刘小宇,等. 白芍总苷提取和纯化工艺的研究[J]. 食品科学. 2007,28(5):173.

[责任编辑 蔡仲德]