

# 药渣粒径对中药提取过程在线检测取样系统影响的模拟研究

隋丞琳 林兆洲 吴志生 徐冰 李洋 史新元\* 乔延江\*

(北京中医药大学 中药信息工程研究中心,北京 100102)

**摘要** 以栀子提取液为研究载体,采用离线方式采集经不同孔径滤网过滤后样品的近红外光谱,并以高效液相测定值为参考值,采用偏最小二乘算法建立栀子苷定量校正模型,探讨基于中药提取过程的在线检测旁路预处理系统可靠性及预处理系统中滤器孔径大小对所建立定量校正模型的影响。结果表明,提取液经筛分过滤后样品与原提取液样品经配对 t 检验,验证同一样品经过滤后所含栀子苷浓度无显著性差异,可以代表原液样品进行数据分析,进一步证明旁路预处理系统可用于在线检测的可行性。经 300 目标准筛过滤后的样品以原始光谱建立的定量模型预测能力最好(校正集的交叉验证均方差 RMSECV = 0.1962,预测均方差 RMSEP = 0.1867,系统偏差 Bias = 0.0256)。通过实践明确取样系统的可行性,合理选择预处理系统滤器孔径,能够有效改善并提高所建模型预测能力,解决中药提取在线过程分析应用技术难题。

**关键词** 中药; 近红外光谱; 提取过程; 在线检测; 取样系统

## 1 引言

为保障中药制剂质量稳定,安全有效,中药行业结合自身特色研发、应用、建立适合中药制剂特点的生产过程在线检测技术与快速质量分析方法,是增强中药产品竞争力的发展要求。

近红外光谱(Near infrared spectroscopy, NIR)分析技术具有操作简便、分析速度快、样本处理量大等优点,符合在线检测要求<sup>[1,2]</sup>,是中药生产过程质量检测的有效手段。目前,中药提取环节的在线近红外实时检测多为旁路采样方式<sup>[3~5]</sup>,在检测之前通常需进行样品预处理,以消除取样条件对分析结果的影响,提高在线分析准确度<sup>[6,7]</sup>。中药提取过程常含有药渣颗粒或杂质粉末,这些杂质粒径大小不一,若进入光路将导致不同程度光谱质量下降<sup>[8]</sup>。此外,中药成分复杂,有效组分间化学信息相互干扰,降低了在线 NIR 分析的准确度。

为解决上述因素造成的在线分析光谱信号稳定性差等普遍存在的应用技术难题,确保在线过程分析中样本代表性和预测模型准确性,本研究采用离线方式模拟制备了含不同粒径药渣颗粒的样本,考察了所采集样本的代表性,在此基础上探讨了药渣粒径对近红外光谱建模结果的影响,以期为验证中药提取过程在线检测旁路预处理系统的可靠性及其优化设计提供指导。

## 2 实验部分

### 2.1 仪器与试剂

XDS Rapid Liquid Analyzer 近红外光谱仪(美国 FOSS 公司); RW 20 数显机械搅拌器(德国 IKA 公司); 高效液相色谱仪(美国 Waters 公司),配备 Waters 2695 泵, Waters 2998 检测器, Empower 色谱工作站; 标准分样筛(100, 200, 300 和 500 目)

栀子药材(由亚宝北中大药厂提供),经北京中医药大学李卫东副研究员鉴定为茜草科植物栀子(*Gardenia jasminoides* Ellis)的干燥成熟果实; 对照品栀子苷(中国药品生物制品检定所,批号 110749-200714),乙腈(色谱纯,美国 Fisher 公司); 水为自制高纯水(美国 Millipore 公司)。

### 2.2 栀子提取液制备

取经破壳处理后的栀子药材 350 g 于 5000 mL 烧杯中,加高纯水 12 倍,加热至提取液沸腾,保持微

2013-01-28 收稿; 2013-06-05 接受

本文系国家“重大新药创制”科技重大专项(No. 2010ZX09502-002)资助

\* E-mail: yjqiao@263.net; shixinyuan01@163.com

沸状态 60 min, 全过程持续搅拌(转速 450 r/min)。为防止提取溶剂挥发, 烧杯口用塑料薄膜封严。自提取液沸腾开始, 每隔 5 min 吸取药液 25 mL 于 50 mL 磨口锥形瓶中, 密封, 待用, 所取样品依次标号, 整个提取过程共获得样品 12 个。每次取样位置不变, 取样结束后, 立即向烧杯中补充相同体积高纯水。相同工艺及取样过程平行操作 3 个批次(A, B, C), 共获得样品 36 个。

### 2.3 待测样品制备

将锥形瓶内栀子提取液混匀, 移取适量于近红外样品管中, 作为原液样品。剩余药液全部缓慢倾注于 100 目分样筛, 收集所有滤液, 将其混匀, 移取适量于近红外样品管中, 得到经 100 目筛滤过后样品(简称 100 目样品); 剩余滤液过 200 目筛, 收集滤液, 混匀, 移取适量, 得 200 目样品; 同法, 得 300 目和 500 目样品。

### 2.4 栀子苷含量测定

对照品溶液的制备: 按《中华人民共和国药典》一部(2010 年版) 栀子药材项下的高效液相色谱法测定栀子苷含量的方法配制。

供试品溶液的制备: 待 2.3 项下全部样品的近红外光谱采集完毕, 精密移取各样品等量, 置 25 mL 容量瓶中, 加 50% 甲醇稀释, 超声处理 20 min, 以 50% 甲醇定容, 滤过, 取续滤液, 既得(每 1 mL 中约含栀子苷 30~40  $\mu\text{g}$ ), 即精确测定 2.3 项下各样品浓度, 作为 NIR 的参考值。

色谱条件: 采用  $\text{C}_{18}$  色谱柱(250 mm  $\times$  4.6 mm, 5  $\mu\text{m}$ , Waters 公司) 色谱柱, 流动相: 乙腈-水(15: 85, V/V), 流速 1 mL/min, 柱温 30  $^{\circ}\text{C}$ , 检测波长 238 nm, 进样量 10  $\mu\text{L}$ 。理论塔板数以栀子苷峰计算, 应不低于 1500。

### 2.5 近红外光谱数据采集

采集近红外光谱前, 反复震摇样品杯内药液至药渣均匀分散。透射方式采集光谱, 以仪器内置背景作为参比。光谱采集范围: 400~2500 nm, 分辨率: 0.5 nm, 扫描次数 32 次。所得近红外原始光谱见图 1。

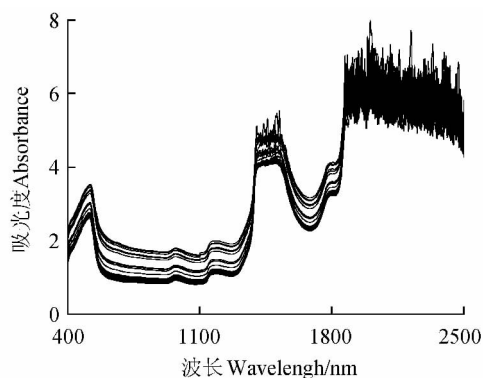


图 1 近红外原始光谱

Fig. 1 Original NIR spectra

## 3 结果与讨论

### 3.1 样本集的划分

Kennard-Stone (K-S) 算法是一种有效的、应用广泛的校正集选取方法。K-S 算法以光谱间欧氏距离为基础, 选择代表性强, 分布范围广的样本作为校正集样本<sup>[9]</sup>。由于本研究中栀子苷含量分布范围较窄, 不同过滤筛目数下样本性质差异较小, 仅考虑样品间的光谱差异, 因此, 采用 K-S 法作为划分校正集和验证集的方法。选取 24 个样品作为校正集, 用于建立校正模型, 剩余样品作为验证集, 对模型进行外部验证。不同样品类别中校正集和验证集样品中栀子苷浓度分布情况, 见表 1。

### 3.2 预处理系统可行性考察

为避免旁路预处理系统对过程分析结果的影响, 降低过程分析手段实现的风险, 需考察原样品和经预处理样品的一致性程度, 即经预处理的样品能否代表原样品进行后续分析。图 2 为高效液相色谱法测得的栀子苷浓度随提取过程的变化。由图 2 可知, 不同目数样品所反映的变化趋势基本一致。为判断同一样品经过滤后与原液样品是否存在显著性差异, 选取显著水平  $\alpha = 0.05$ , 将不同目数样品分别与原液样品进行配对 t 检验, 所得 P 均大于 0.05, 表明同

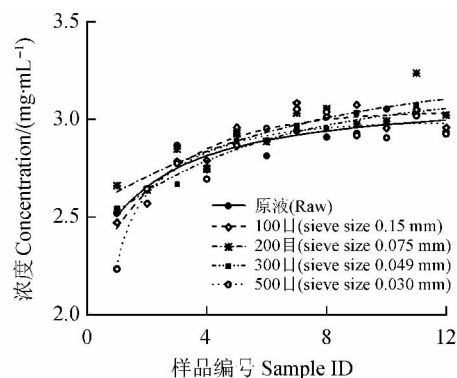


图 2 原液与不同目数样品中栀子苷含量变化趋势图

Fig. 2 Trend of geniposide concentration in original extract and extracts with different sizes of particles

表 1 校正集和验证集样品中梔子苷浓度分布情况表( g/L)

Table 1 Distributon of geniposide concentration in calibration and validation sets

样品名称 Sample name	样品集 Sample sets	样本数 Number	浓度范围 Concentration range ( g/L)	均值 ± 标准偏差 Mean ± SD ( g/L)
原液 Raw extract	校正集 Calibration	24	2.22 ~ 3.17	2.85 ± 0.24
	验证集 Validation	12	2.35 ~ 3.12	2.86 ± 0.23
100 目样品 ( Sieve size 0.15 mm)	校正集 Calibration	24	1.98 ~ 3.24	2.76 ± 0.30
	验证集 Validation	12	2.45 ~ 3.09	2.83 ± 0.23
200 目样品 ( Sieve size 0.075 mm)	校正集 Calibration	24	1.99 ~ 3.16	2.83 ± 0.26
	验证集 Validation	12	2.28 ~ 3.24	2.85 ± 0.26
300 目样品 ( Sieve size 0.049 mm)	校正集 Calibration	24	2.01 ~ 3.11	2.81 ± 0.27
	验证集 Validation	12	2.36 ~ 3.14	2.86 ± 0.23
500 目样品 ( Sieve size 0.030 mm)	校正集 Calibration	24	2.21 ~ 3.32	2.88 ± 0.27
	验证集 Validation	12	2.24 ~ 3.09	2.85 ± 0.26

Notes: Sieve size stands for extracts filtrated by sieve size of 0.15 mm ,0.075 mm ,0.049 mm and 0.030 mm , respectively.

一样品筛分前后药液中所含梔子苷浓度均无显著性差异,即过滤后样品可以代表原液样品进行数据分析。该结果进一步说明,将旁路系统用作过程分析手段,对中药水提过程实现在线实时检测切实可行。

### 3.3 校正模型的建立和光谱预处理方法筛选

运用 Vision 5.0( Foss NIR Systems , Silver Spring , MD , USA) 光谱分析软件,以偏最小二乘法( PLS) 建立梔子苷含量的校正模型,并用验证集样品进行外部验证。

由图 1 可知,原始光谱在 1800 ~ 2500 nm 和 1400 ~ 1520 nm 存在波动。1800 ~ 2500 nm 波段下所有样本光谱波动均较为剧烈,建模时将其剔除的原因是本研究所采用仪器采集该波段下光谱波动剧烈以及组合频区在该波段下存在饱和吸收特征<sup>[10]</sup>,用于建模会导致模型性能剧烈下降。而波段 1400 ~ 1520 nm 的波动可能包含不同粒径大小药渣因素,为分析不同粒径大小药渣对建模结果的影响,将其保留。因此,本研究选用 800 ~ 1800 nm 作为梔子苷含量的建模波段。

比较了原始光谱、一阶导数、二阶导数、Savitzky-Golay( 9: 2) 平滑法( 简称 SG 平滑)、多元散射校正( Multiplicative signal correction, MSC)、标准正态变换( Standard normal variate, SNV)<sup>[11,12]</sup> 等光谱预处理方法对模型性能的影响。利用四折内部交叉验证法,以潜变量因子数对预测残差平方和( Predict of error square sum, PRESS) 作图( 图 3),选择最优预处理方法。由图 3 可知,对于梔子提取液样品,采用原始光谱、MSC、SNV 建模结果较好,其中原始光谱建模 PRESS 值最小,因此,本研究采用原始光谱建立数学模型。

### 3.4 不同目数样品所建模型预测性能分析

本研究对 3 个批次梔子提取液的实验样本进行数据分析,采用偏最小二乘回归模型,内部留一交叉验证( Cross-validation) 将光谱数据与高效液相色谱法分析结果进行关联建立模型,利用外部样本集验证模型预测性能。

模型的内部稳健性和拟合效果主要以校正集相关系数  $R_{cal}$ 、校正误差均方根( Root mean square error of calibration, RMSEC) 和校正集的交叉验证均方差( Root mean square error of cross validation, RMSECV) 为评价指标。 $R_{cal}$  越接近于 1, RMSEC 和 RMSECV 越小,则所建模型的拟合效果及模型稳定性越好。所建模型对相应验证集样品的预测结果以验证集相关系数  $R_{val}$  和预测误差均方根( Root mean square error

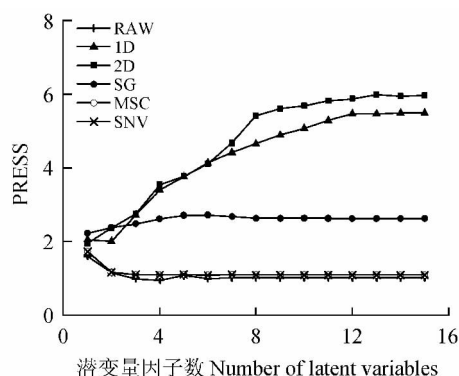


图 3 不同预处理方法预测残差平方和图

Fig. 3 Predictive residual of error square sum ( PRESS) values by different spectral preprocessing methods

Raw: 原始光谱; 1 D: 一阶导数; 2 D: 二阶导数; SG: SG 平滑; MSC: 多元散射校正; SNV: 标准正态变换。

Raw: raw spectra; 1 D: first derivative; 2 D: second derivative; SG: Savitzky-Golay Smoothing; MSC: Multiplicative Signal Correction; SNV: Standard Normal Variate.

of prediction, RMSEP) 为评价指标。 $R_{val}$  越接近于 1, RMSEP 越小, RMSECV 与 RMSEP 越接近于 1, 则模型的预测性能越好。为方便比较不同样品间的预测结果, 本研究还选用系统偏差 Bias<sup>[13~15]</sup> 判断模型间差异的一个参数。Bias 为外部验证残差的均值, Bias 值越趋近于零, 所建模型的预测性能越好, 预测精度越高。

原液和经过滤样品建模的模型预测结果见表 3。为方便比较不同样本所建模型的性能, 图 4 简要介绍比较方法。

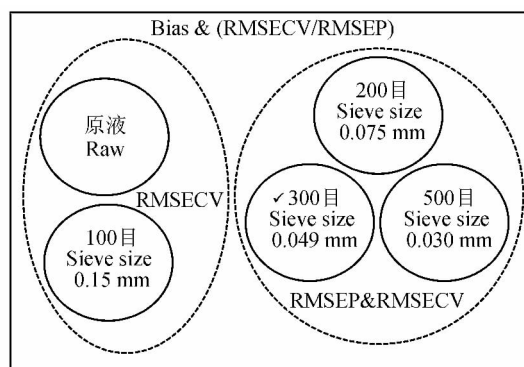


图 4 不同模型预测性能比较示意图

Fig. 4 Schematic of comparison of prediction performance among different models

Bias, RMSECV, RMSEP, RMSECV/RMSEP 是比较模型预测性能的依据: 以 Bias 和 RMSECV/RMSEP 将所有样品所建模型分为两类, 用虚线圆表示; 以 RMSECV 区分原液和 100 目样品所建模型; 以 RMSEP 和 RMSECV 区分 200 目, 300 目和 500 目样品 “&”代表“且”

Bias, root mean square error of cross validation, root mean square error of prediction (RMSECV), RMSEP and (RMSECV/RMSEP) were the basis of comparison of model prediction performance: all models were divided into two groups by combination use of Bias and RMSECV/RMSEP, and each group was circled by dashed line. RMSECV was used to distinguish between raw extract and sieve size 0.15 mm. The difference among sieve size of 0.15 mm, 0.075 mm, 0.049 mm and 0.030 mm were evaluated by RMSEP and RMSECV

结合表 3 中模型预测参数可知, 原液样品和 100 目样品所建模型的预测性能较差, 其 Bias 分别为 -0.0713 和 -0.1097, RMSECV/RMSEP 分别为 1.29 和 0.75, 明显低于其它目数样本所建模型。同时, 原液样本建模 RMSECV 为 0.2469, 高于经过滤样品的相应参数。此结果表明, 未经过滤的原药液样品中含不溶性固体药渣或杂质, 其进入光路引起光传递路径和能量等的改变使得模型的预测结果存在随机性和不确定性, 因此需对样品进行过滤以改善模型预测能力。100 目样品所建模型 RMSECV 为 0.1720, 校正模型稳定性较原液样品建模明显提高, 但与其它目数模型相比预测性能并不理想 (RMSEP 为 0.2279)。这可能是由于药材颗粒大小、性状及表面粗糙程度等不同引起的非线性光散射。原液及 100 目样品内固体颗粒种类较为多样, 光能量被吸收或发生散射 (Rayleigh 散射和 Lorenz-Mie 散射) 造成光谱质量下降。

表 3 原液与经不同孔径筛分样品建模结果

Table 3 Comparison of calibration models using extracts filtrated by different sieve meshes

样品名 Sample ID	潜变量数 Factors	$R_{cal}$	RMSEC	RMSECV	$R_{val}$	RMSEP	RMSECV/ RMSEP	Bias
原液样品 Raw extract	4	0.9276	0.0885	0.2469	0.6349	0.1921	1.29	-0.0713
100 目样品 Sievesize 0.15 mm	4	0.8936	0.1325	0.1720	0.6380	0.2279	0.75	-0.1097
200 目样品 Sievesize 0.075 mm	4	0.9453	0.0836	0.2035	0.6919	0.2066	0.98	0.0386
300 目样品 Sievesize 0.049 mm	3	0.8886	0.1212	0.1962	0.8636	0.1867	1.05	0.0256
500 目样品 Sievesize 0.030 mm	5	0.9805	0.0522	0.2273	0.8172	0.2146	0.98	0.0578

对其它目数样品的建模结果进行比较,发现 300 目样品所建模型的 RMSECV 为 0.1962, RMSEP 为 0.1867, Bias 为 0.0256, 明显优于其它目数样品所建模型。其中, 200 目样品和 500 目样品所建模型 (Bias 分别为 0.0386 和 0.0578, RMSECV/RMSEP 均为 0.98) 优于原液样品和 100 目样品所建预测模型结果, 却比 300 目样品所建模型结果稍差。造成 500 目样品所建模型结果较差的可能原因是伴随溶剂作用的存在光谱在小颗粒杂质中传播也较为复杂等因素, 导致建模准确性下降<sup>[16]</sup>。

上述结果表明, 在线采集原药液近红外光谱不仅较难实施, 而且所采集光谱受到药渣干扰较严重所建模型结果不理想, 有必要在采样前对原药液进行预处理, 除去部分药渣。对于过滤装置而言, 选择适宜研究载体的滤器孔径大小, 能够提高近红外光谱在线检测的准确性。本研究所涉 4 种规格标准筛旨在缩小在应用中过滤器滤网孔径选择范围, 并得出采用 300 目滤网过滤后药液所建模型预测性能优良, 可选为栀子提取过程在线近红外光谱检测旁路预处理系统过滤装置滤网孔径。

采用离线模拟手段可为在线预处理系统功能组成模块的选配提供指导, 降低在线应用风险及成本。此外, 中药提取过程的在线近红外光谱检测设备旁路预处理系统的过滤装置孔径应视实际研究载体而选定, 应借助离线方法考察不同中药材、提取溶剂、生产工艺等条件下, 药渣或固体杂质颗粒大小对近红外光谱所建数学模型的影响。同时, 考虑到在线预处理系统装置的工作效率和工程实施难度, 建议在能够满足分析目的的情况下, 尽量选取滤网孔径较大的滤器, 若为保障分析精度选择小孔径滤器, 可使用多级过滤方式。

合理设计在线近红外光谱样品过滤装置, 能够在保证所得数据客观反映物料真实动态的同时, 降低分析手段自身引入的如采样环境、条件等干扰因素对预测模型结果的影响, 提升所建模型的预测准确性。

## References

- 1 Luypaert J, Massart D L, Vander Heyden Y. *Talanta*, **2007**, 72(3): 865–883
- 2 Roggo Y, Chalus P, Maurer L, Lema M, Carmen E, Aurélie J, Nadine. *J. Pharm. Biomed. Anal.*, **2007**, 44(3): 683–700
- 3 Wu Y J, Jin Y, Ding H Y, Luan L J, Chen Y, Liu X S. *Spectrochim. Acta Part A.*, **2011**, 79(5): 934–939
- 4 Wu Y J, Jin Y, Li Y R, Sun D, Liu X S, Chen Y. *Vib. Spectrosc.*, **2012**, 58: 109–118
- 5 JIANG Xue, ZHANG Jing-Shuo, QIU Min, XIE Hong-Ping, GUO Li-Ping. *Chinese J. Anal. Chem.*, **2006**, 34(9): 171–174  
蒋雪, 张经硕, 邱敏, 谢洪平, 郭丽萍. *分析化学*, **2006**, 34(9): 171–174
- 6 CHU Xiao-Li, YUAN Hong-Fu, LU Wan-Zhen. *Mod. Sci. Instrum.*, **2004**, (2): 3–17  
褚小立, 袁洪福, 陆婉珍. *现代科学仪器*, **2004**, (2): 3–17
- 7 FU Qing-Ling. *Contr. Instru. Chem. Ind.*, **2002**, 29(1): 59–66  
符青灵. *化工自动化及仪表*, **2002**, 29(1): 59–66
- 8 Rinnan Å, Berg F V D, Engelsen S B. *TrAC-Trend Anal. Chem.*, **2009**, 28(10): 1201–1222
- 9 LI Hua, WANG Ju-Xiang, XING Zhi-Na, SHEN Gang. *Spectroscopy and Spectral Analysis*, **2011**, 31(2): 362–365  
李华, 王菊香, 邢志娜, 申刚. *光谱学与光谱分析*, **2011**, 31(2): 362–365
- 10 Wu Z S, Du M, Xu B, Lin Z Z, Shi X Y, Qiao Y J. *J. Mol. Struct.*, **2012**, 1019: 97–102
- 11 Wu Z S, Du M, Sui C L, Shi X Y, Qiao Y J. *Anal. Methods*, **2012**, 4(4): 1084–1088
- 12 Wu Z S, Xu B, Du M, Sui C L, Shi X Y, Qiao Y J. *J. Pharm. Biomed. Anal.*, **2012**, 62: 1–6
- 13 DU Min, WU Zhi-Sheng, LIN Zhao-Zhou, SUI Cheng-Lin, SHI Xin-Yuan, QIAO Yan-Jiang. *Chinese J. Pharm. Anal.*, **2012**, 32(10): 1796–1800  
杜敏, 吴志生, 林兆洲, 隋丞琳, 史新元, 乔延江. *药物分析杂志*, **2012**, 32(10): 1796–1800
- 14 Bellon M V, Fernandez A E, Palagos B, Roger J M, McBratney A. *TrAC-Trend Anal. Chem.*, **2010**, 29(9): 1073–1081
- 15 HUANG Cai-Jin, LIU Xian, YANG Zeng-Ling, HAN Lu-Jia. *Spectroscopy and Spectral Analysis*, **2009**, 29(5): 1264–1267  
皇才进, 刘贤, 杨增玲, 韩鲁佳. *光谱学与光谱分析*, **2009**, 29(5): 1264–1267
- 16 Andrew J O, Roger D J, Anthony C M. *Analyst*, **1998**, 123: 2297–2302

