

# 近红外漫反射光谱法快速测定积雪草总苷中积雪草苷的含量

赵娜 吴志生 袁瑞娟 孙杨 詹雪艳 乔延江

(北京中医药大学中药学院,北京,100102)

**摘要** 目的:建立近红外漫反射光谱法快速测定积雪草总苷中积雪草苷的含量方法。方法:采用高效液相色谱法测定市场上流通的66批次积雪草总苷中积雪草苷的含量,同时采集近红外漫反射光谱,对原始光谱进行多元散射校正(MSC)、二阶导数(2D)和Norris导数(ND)平滑等光谱预处理方法,采用偏最小二乘回归法(PLSR)建立近红外定量校正模型,实现积雪草总苷中指标性成分积雪草苷含量的快速测定。结果:所建立的积雪草苷近红外定量校正模型,验证集相对误差RMSEP、标准偏差与标准误差的比值RPD分别为0.0352和6.47,表明所建近红外模型预测准确度高。同时,校正集均方根误差RMSEC和验证集均方根误差RMSEP接近,分别为1.17和1.25,交叉验证均方根误差RMSECV为3.43,表明所建近红外模型稳健性好。结论:所建立的积雪草苷近红外定量校正模型具有很好的预测准确度和稳健性,为市场上积雪草总苷质量的快速评价提供方法。

**关键词** 近红外光谱;积雪草总苷;积雪草苷;漫反射

## Rapid Determination the Content of Asiaticoside in Centella Total Glucosides by Near Infrared Reflectance Spectroscopy

Zhao Na, Wu Zhisheng, Yuan Ruijuan, Sun Yang, Zhan Xueyan, QIAO Yanjiang

(School of Chinese Pharmacy, Beijing University of Chinese Medicine, Beijing 100102, China)

**Abstract Objective:** To establish near - infrared spectrometry quantitative models for the determination of asiaticoside in Centella Total Glucosides. **Methods:** The contents of asiaticoside of 66 samples of Centella Total Glucosides were determined by high - performance liquid chromatography. And the near - infrared spectrogram of 66 samples of Centella Total Glucosides was collected at the same time. After smoothing pretreatment by Multiple scattering correction (MSC), the second derivative (2D) and Norris derivative (ND), partial least square regression (PLS) was used to establish near - infrared spectrometry quantitative models for the determination of asiaticoside. **Results:** In validation set, the relative error (RSEP), and the ratio of standard deviation and standard error of prediction were 0.0352 and 6.47, respectively, which indicated that the models had good accuracy. The root mean square error (RMSEC) in calibration set and the root mean square error (RMSEP) in validation set were 1.17 and 1.25, respectively, and the root mean square error of cross validation (RMSECV) was 3.43, which showed that the near infrared model was robust. **Conclusion:** The established models have good accuracy and robustness, which provide an efficient approach for the fast determination of asiaticoside in Centella Total Glucosides.

**Key Words** Near - infrared spectrometry (NIR); Centella Total Glucosides; Asiaticoside; Diffuse reflectance

doi:10.3969/j.issn.1673-7202.2013.11.005

积雪草总苷(Centella Total Glucosides)是伞形科植物积雪草*Centella asiatica*(L.)Urb.的全草经加工制成的总皂苷,主要成分为羟基积雪草苷、积雪草苷、积雪草苷B等皂苷化合物,具有促进创面愈合、抑制瘢痕形成、抗抑郁、改善记忆、抗氧化、保护神经元等方面药理作用。目前,积雪草总苷的质量控制采用HPLC指纹图谱与指标性成分含量相结合来评价积雪草总苷的质量<sup>[1]</sup>,其指标性成分积雪草苷的含量采用高效液相色谱法测定,操作繁琐费时,难以实现质量的快速评价和实时监控。近红外光谱(Near - Infrared Spectroscopy, NIR)分析技术是一种典型的快速无损分析技术,

在中药材<sup>[2]</sup>、中药制剂<sup>[3]</sup>、中药生产过程<sup>[4-6]</sup>以及药物<sup>[7]</sup>的质量快速评价中已得到广泛的应用。本研究以积雪草总苷为研究载体,采集66份11个厂家不同批次的积雪草总苷样品的近红外漫反射光谱,对原始光谱进行多元散射校正(MSC)、二阶导数(2D)和Norris导数(ND)平滑的光谱预处理,采用偏最小二乘法(PLS)建立近红外定量校正模型,实现积雪草总苷中指标性成分积雪草苷含量的快速测定以及市场上积雪草总苷质量的快速评价和质量监控。

## 1 仪器与试剂

仪器: Antaris 傅里叶变换近红外光谱仪(美国 Thermo Nicolet 公司)配有 InGaAs 检测器、积分球漫反射采样系统、Result 操作软件和 TQ Analyst 光谱分析软件;色谱柱为 Agilent TC - C<sub>18</sub>(4.6mm × 250mm, 5 μm); Shimadzu LC - 20AT 高效液相色谱仪(日本岛津公司)配有 LC - 20AT 高压泵、DGU - 20A5 在线脱气机、SIL - 20A 自动进样器、CTO - 10AS 柱温箱、SPD - 20A 紫外检测器;电子天平 BS 110S(北京赛多利斯仪器系统有限公司)。

试剂:积雪草苷(中国食品药品检定研究院), β - 环糊精(天津光复化工研究所), 色谱纯乙腈(Fisher 赛默飞世尔科技有限公司), 娃哈哈纯净水。

积雪草总苷样本来自市场上流通的广西、陕西、江苏三省 11 个积雪草总苷提取物厂家的 66 个不同批次样品, 各个批次的样品均为粒径小于 80 目的干燥粉末, 密封包装。

## 2 实验部分

### 2.1 积雪草苷的含量测定

2.1.1 积雪草苷对照品溶液的制备 精密称取积雪草苷对照品 4.3mg 至 10mL 容量瓶中, 乙腈 - 水(1:3)定容, 摇匀, 称定质量, 超声溶解 20min, 冷却后再称定质量, 用乙腈 - 水(1:3)补足重量, 摇匀, 微孔滤膜过滤(0.45 μm), 即得积雪草苷对照品溶液。

2.1.2 供试品溶液的制备 精密称取积雪草总苷样品 7.5mg, 至 10mL 容量瓶中, 乙腈 - 水(1:3)定容, 摇匀, 称定质量, 超声溶解 20min, 冷却后再称定质量, 用乙腈 - 水(1:3)补足, 摇匀, 微孔滤膜过滤(0.45 μm), 即得。

2.1.3 色谱条件 Agilent TC - C<sub>18</sub>(4.6 mm × 250 mm, 5 μm) 色谱柱; 乙腈 - 2mmol/L β - 环糊精溶液(25:75)为流动相, 流速为 1mL/min; 柱温为 25℃; 紫外检测波长 205nm, 进样量 10 μL。

2.1.4 样品的测定 按照 2.1.3 液相色谱条件采集供试品溶液的色谱图, 采用外标标准曲线法测定积雪草总苷中积雪草苷的含量。

2.2 近红外漫反射光谱的采集 积雪草总苷样品: 以内置背景为参照, 积分球漫反射方式采集近红外光谱, 扫描范围 4000 ~ 10000cm<sup>-1</sup>, 分辨率 8 cm<sup>-1</sup>, 扫描次数 32 次, 增益 2, 每个样品重复测定 3 次, 以其平均光谱作为积雪草总苷各样本的近红外光谱。积雪草总苷的原始近红外光谱见图 1, 其中横坐标为波数值, 纵坐标为吸光度。

2.3 校正集样本选择和光谱预处理 66 个积雪草总

苷样本, 采用 SPXY 法<sup>[8]</sup>选择 2/3 作为校正集样本, 1/3 为验证集样本, SPXY 算法原代码来自 PLS - Toolbox 2.1, 在 MATLAB(Mathwork Inc)环境中实现。

利用 TQ Analyst 8 软件, 采用多元散射校正(multivariate scatter correction, MSC)和标准正则变换(standard normal variate, SNV)等散射校正法、一阶导数(1st derivative, 1D)和二阶导数(2nd derivative, 2D)等光谱导数法及 Savitzky - Golay 平滑(SG)和 Norris 导数(Norris derivative, ND)平滑等光谱数据预处理方法, 提高近红外分析信号的信噪比, 筛选光谱波段, 并确定合适的潜变量因子数, 采用偏最小二乘回归(Partial Least Square Regression, PLSR)建立积雪草苷近红外定量校正模型。

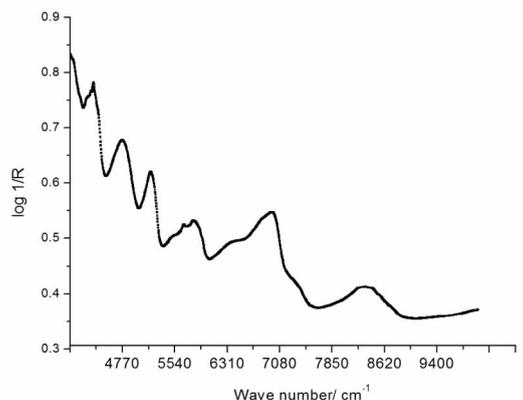


图 1 积雪草总苷的近红外漫反射光谱

2.4 近红外定量模型的评价指标 近红外定量模型通常用样本的预测值和化学测量值相关系数(correlation coefficient, R)、交叉验证均方根误差(Root Mean Squared Error of Cross - Validation, RMSECV)、校正集均方根误差(Root Mean Squared Error of Calibration, RMSEC)、验证集均方根误差(Root Mean Squared Error of Prediction, RMSEP)、验证集相对误差(Relative Standard Error of Prediction, RSEP)<sup>[9-10]</sup>及验证集标准偏差与标准误差的比值(Ratio of standard deviation and standard error of prediction, RPD)<sup>[10-11]</sup>等指标来评价所建近红外定量模型的性能, 优化建模过程中参数。预测集中各样本模型预测值和化学测量值相关系数 R<sub>p</sub>和 RPD 越大, RMSECV、RMSEC、RMSEP 和 RSEP 的值越小, 表明模型的预测性能越好。其中, RPD = SD/RMSEP, 其中 SD 为验证集化学值的标准差, 当 RPD 大于 3 时, 表明模型具有很好的预测能力<sup>[12]</sup>。

## 3 结果与讨论

3.1 校正集和验证集样本的划分 66 份积雪草总苷样品中积雪草苷含量范围为 13.73% ~ 69.82%, 平均含量为 31.70%, 标准偏差为 10.97%。SPXY 法将 66

个样本划分为校正集和验证集,两个样本集中积雪草苷含量的分布如表 1,校正集中积雪草苷含量的数值范围大于验证集中积雪草苷含量的数值范围,而且校正集中积雪草苷含量的数值范围与全部样本的积雪草苷含量的数值范围相当,表明所选取校正集样本能代表原 66 份样品的全部信息,而且其化学测量值的范围覆盖了验证集的化学测量值的范围。

表 1 校正集和验证集中积雪草苷的含量分布

样本集	样本数	积雪草苷含量范围%	平均值%	标准偏差%
calibration set	44	13.73 ~ 69.82	29.43	11.51
validation set	22	14.77 ~ 42.76	36.23	8.28

3.2 光谱预处理方法的选择 建立积雪草苷偏最小二乘回归模型时,以 RMSECV、RMSEC、RMSEP 或 RSEP、Rp 或 RPD 等模型评价指标为指导选择合适的 NIR 光谱预处理方法,剔除仪器噪音、基线不稳等因素带来干扰信号,提取与目标成分测量值的相关信息,提高模型的稳健性和预测的准确性<sup>[13]</sup>。本研究采用 TQ Analyst 8 软件,改变光谱预处理方法并建立相应的最小二乘回归模型,以所建模型的 RMSECV、RMSEC、RMSEP 或 RSEP 指标数值小,Rp 或 RPD 指标值较大为目标选择合适的光谱预处理方法,采用多元散射校正(MSC)消除积雪草总苷固体颗粒不均匀和表面散射等光散射信号的干扰,利用二阶导数(2D)扣消除基线漂移和仪器背景信息,Norris 导数光谱平滑来消除仪器噪音的干扰。

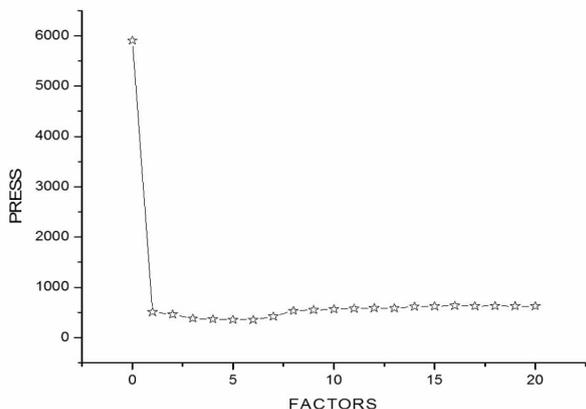


图 2 积雪草苷定量校正模型 PRESS - FACTORS 间的关系

3.3 潜变量因子数的确定 在确定校正集样本的情况下,采用偏最小二乘法建立多元校正模型,潜变量因子数的选择对模型的预测能力有很大的影响。如果选择的潜变量因子数过少,光谱中一些有效信息未被选入,建模时因信息利用不全而导致模型预测能力差,出现欠拟合现象。相反,如果选择的潜变量因子数过多,一些代表噪音的因子被带到模型中,模型预测能力下降,出现过拟合现象。采用内部交叉验证法,预测残差

平方和(PRESS)值随潜变量因子数(FACTORS)的变化见图 2,PRESS 值越小,模型的预测性能越好,即最小的 PRESS 值对应的最佳潜变量因子数,最佳潜变量因子数为 6。

3.4 模型的建立与评价 采用 SPXY 法从 66 个样本中选择 2/3 校正集样本,对其原始 NIR 光谱进行多元散射校正,二阶求导和 Norris 导数平滑后,选取潜变量因子数为 6,建立偏最小二乘定量模型。模型的 RMSEC、RMSEP 和 RSEP 分别为 1.17、1.25 和 0.0353,均比较小,积雪草苷含量的近红外预测值和高效液相实测值相关性见图 3,校正集相关系数为 0.9947,验证集的相关系数为 0.9880,而且验证集标准偏差与标准误差的比值 RPD 为 6.47,大于 RPD 的阈值 3.0,表明所建模型的预测准确度高<sup>[12]</sup>。所建模型 RMSEC 与 RMSEP 接近,表明所建模型对校正集样本数据有小的预测误差,好的预测准确度,而且所建模型适用于验证集样本,该模型具有较好的拓展性和稳健性。同时,所建模型 RMSECV 值为 3.43,数值较小,表明所建模型的稳健性好。

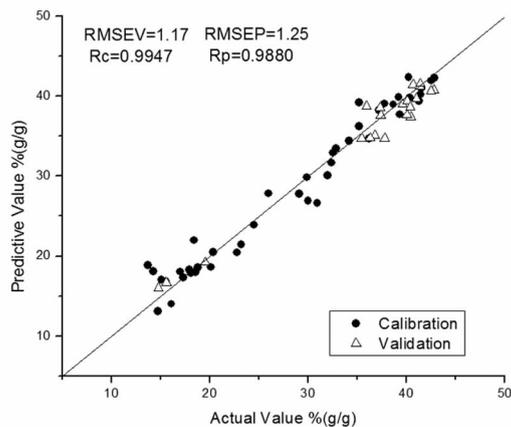


图 3 积雪草苷含量实测值和预测值相关图

#### 4 结论

本研究以积雪草总苷为研究载体,采集市场上流通的 66 个批次积雪草总苷样本的近红外漫反射光谱,采用偏最小二乘回归建立指标性成分积雪草苷定量校正模型。所建立的近红外定量校正模型的交叉验证均方根误差 RMSECV 值小,校正集均方根误差 RMSEC 和验证集均方根误差 RMSEP 接近,表明所建模型具有很好的稳健性,同时模型的 RSEP 和 RPD 分别为 0.0352 和 6.47,表明所建模型预测准确度高,可实现积雪草总苷中积雪草苷含量的快速测定。该研究为市场流通的积雪草总苷产品质量的快速评价提供方法,对积雪草总苷产品的质量的控制具有重要意义。

(下接第 1286 页)

结果如表 6 所示。

表 5 验证实验点的选取

实验	乙醇浓度	乙醇倍量	煎煮时间	实验点的说明
YZ1	80	10	2	Y1 < 70% 且 Y2 < 50%
YZ2	80	6	1.8	Y1 < 70% 且 Y2 < 50%
YZ3	100	6	0.5	Y1 > 70%, Y2 > 50%, 95% 置信区间外
YZ4	100	6.5	0.9	Y1 > 70%, Y2 > 50%, Y1 95% 置信区间外, Y2 95% 置信区间内
YZ5	100	8	0.5	Y1 > 70%, Y2 > 50%, 95% 置信区间内

表 6 设计空间验证实验结果

实验	Y1 (%)	Y2 (%)
YZ 1	27.33	23.08
YZ 2	26.25	21.37
YZ 3	67.18	58.06
YZ 4	61.91	50.52
YZ 5	70.65	62.56

由验证实验结果可知, YZ1、YZ2 和 YZ5 均符合预设要求。YZ3 和 YZ4 中 Y1 < 70%, Y2 > 50%, Y2 均符合预设要求而 Y1 却没有达到, 这表明由于 95% 置信区间的加入, 相比未加入 95% 置信区间, 设计空间边界的不确定性得到了较好的印证。

#### 4 结论

本文采用中心点复合设计对丹参醇提工艺进行优化, 并构建了丹参醇提工艺的设计空间。加入 95% 置信区间后, 设计空间的范围有所缩小, 提高了设计空间的可靠性。实验结果表明, 基于 QbD 理念建立丹参醇提设计空间可提高工艺过程的稳健性和可靠性。

#### 参考文献

- [1] International Conference on Harmonisation, ICH Q8 (R2); Pharmaceutical Development, 2009.
- [2] Pisano R, Fissore D, Barresi A A, et al. Quality by design: optimization of a freeze-drying cycle via design space in case of heterogeneous drying behavior and influence of the freezing protocol[J]. *Pharmaceutical Development and Technology*, 2013, 18(1): 280-295.
- [3] Martinez L, Antonio P, Lorenz L, et al. Use of near-infrared spectroscopy to quantify drug content on a continuous blending process: Influence of mass flow and rotation speed variations[J]. *European Journal of Pharmaceutics*, 2013, 84(3): 606-615.
- [4] Defne Kayrak - Talay, Dale Steven, Wassgren Carl, et al. Quality by design for wet granulation in pharmaceutical processing: Assessing models for a priori design and scaling[J]. *Powder Technology*, 2013, 240: 7-18.
- [5] Teckoe Jason, Mascaro Tracey, Farrell Thomas P, et al. Process Optimization of a Novel Immediate Release Film Coating System using QbD Principles[J]. *AAPS PharmSciTech*, 2013, 14(2): 531-540.
- [6] Streefland M, Van Herpen PFG, Van de Waterbeemd B, et al. A practical approach for exploration and modeling of the design space of a bacterial vaccine cultivation process[J]. *Biotechnol Bioeng*, 2009, 104(3): 492-504.
- [7] 徐冰, 史新元, 乔延江, 等. 中药制剂生产工艺设计空间的建立[J]. *中国中药杂志*, 2013, 38(6): 924-929.
- [8] 徐冰, 罗赣, 林兆洲, 等. 基于过程分析技术和设计空间的金银花醇沉加醇过程终点检测[J]. *高等学校化学学报*, 2013, 34(10): 2284-2289.
- [9] 翟学佳. 丹参药材热不稳定性考察及提取工艺优化[J]. *中国医院药学杂志*, 2009, 29(19): 1619-1622.

(2013-10-30 收稿)

(上接第 1282 页)

#### 参考文献

- [1] 国家药典委员会. 中华人民共和国药典(2010 年版一部)[S]. 北京: 中国医药科技出版社, 2010: 389.
- [2] 李国辉, 张录达, 杨建文, 等. 栽培和野生中药材灯盏花的近红外光谱鉴别模型[J]. *光谱学与光谱分析*, 2007, 27(10): 1959-1961.
- [3] 蒋受军, 刘丽娜, 朱斌, 等. 近红外光谱法测定注射用丹参(冻干)中丹参素、原儿茶醛及总酚含量的初步研究[J]. *药物分析杂志*, 2008, 28(7): 1094-1098.
- [4] 金叶, 丁海樱, 吴永江, 等. 近红外光谱技术用于血必净注射液提取过程的在线检测研究[J]. *药物分析杂志*, 2012, 32(7): 1214-1221, 1234.
- [5] Zhisheng Wu, Bing Xu, Min du, et al. Validation of a NIR quantification method for the determination of chlorogenic acid in *Lonicera japonica* solution in ethanol precipitation process[J]. *Journal of Pharmaceutical and Biomedical Analysis*, 2012, 6: 1-6.
- [6] 于宝珠, 辛明, 施朝晨, 等. 微型近红外分析仪用于药物制剂混合中的在线检测[J]. *中国药事*, 2008, 22(7): 560-563.
- [7] 侯少瑞, 冯艳春, 胡昌勤. 近红外光谱法快速分析注射用头孢曲松钠

及其水分的含量[J]. *药物分析杂志*, 2008, 28(6): 936-941.

- [8] Galvo R. K, Araujo M. C., José G. E., et al. A method for calibration and validation subset partitioning[J]. *Talanta*, 2005, 67(4): 736-740.
- [9] Blanco M, Peguero A. Influence of physical factors on the accuracy of calibration models for NIR spectroscopy[J]. *J Pharm Biomed Anal*, 2010, 52(1): 59-65.
- [10] 戴传云, 高晓燕, 杨波, 等. 近红外光谱法测定双黄连口服液液中绿原酸和连翘苷的含量[J]. *光谱学与光谱分析*, 2010, 30(2): 358-362.
- [11] Joubert E, Manley M, Botha M. Use of NIRS for Quantification of Mangiferin and Hesperidin Contents of Dried Green Honeybush (*Cyclopia genistoides*) Plant Material[J]. *Agricultural And Food Chemistry*, 2006, 54(15): 5279-5283.
- [12] Fei Shen, Xiaoying Niu, Danting Yang, et al. Determination of Amino Acids in Chinese Rice Wine by Fourier Transform Near-Infrared Spectroscopy[J]. *Agricultural And Food Chemistry*, 2010, 58(17): 9809-9816.
- [13] 褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. *化学进展*, 2004, 16(4): 528-542.

(2013-10-30 收稿)