

基于四种 NIR 仪器类型的清开灵注射液中 黄芩苷成分的多变量检测限研究

彭严芳, 史新元, 周璐薇, 裴艳玲, 李洋, 吴志生*, 乔延江*

北京中医药大学中药信息工程中心, 北京 100102

摘要 选取 a, b, c 和 d 四种类型近红外光谱仪, 分别采用透射模式测定清开灵注射液近红外光谱, 以高效液相色谱法测定值作为参考值, 建立不同仪器类型清开灵注射液中黄芩苷偏最小二乘(PLS)和间隔偏最小二乘(iPLS)定量模型, 并计算模型的多变量检测限(MDL)。四种仪器的 PLS 模型相关系数(r_{pre})和预测均方差(SEP) 分别为 0.976 2 和 $230.4 \mu\text{g} \cdot \text{mL}^{-1}$ (a), 0.956 1 和 $246.4 \mu\text{g} \cdot \text{mL}^{-1}$ (b), 0.966 2 和 $264.4 \mu\text{g} \cdot \text{mL}^{-1}$ (c), 0.998 5 和 $71.5 \mu\text{g} \cdot \text{mL}^{-1}$ (d), 其中 d 型仪器较其他三种类型能获得更好的模型性能。经 iPLS 变量筛选后, a 和 b 两种类型仪器得到的 iPLS 模型 R_{pre}^2 和 SEP 分别为 0.977 1 和 $218.4 \mu\text{g} \cdot \text{mL}^{-1}$, 0.975 4 和 $219.4 \mu\text{g} \cdot \text{mL}^{-1}$, 相较其 PLS 模型预测性能未见明显提高; c 和 d 未筛选出变量。不同仪器的 MDL ($\Delta_{0.05, 0.05}$) 均低于 $250 \mu\text{g} \cdot \text{mL}^{-1}$, 其中 c 和 d 型 MDL 分别低至 $58 \mu\text{g} \cdot \text{mL}^{-1}$ 和 $2.9 \mu\text{g} \cdot \text{mL}^{-1}$ 。结果表明不同类型仪器定量预测性能和 MDL 不同。本文创新性采用多变量检测限理论探讨了不同类型近红外仪器的检测性能, 结果表明这一方法具有可行性。在实际应用中应根据研究载体的特征选择合适的仪器类型, 以确保定量准确性。

关键词 近红外; 清开灵注射液; 偏最小二乘; 间隔偏最小二乘; 多变量检测限

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2013)09-00-06

引言

近红外光谱法(near infrared spectroscopy, NIRS)具有操作简便、分析快速、不破坏分析样品、无污染, 特别适宜于在线分析等优点^[1], 已在国外制药生产过程分析中得到成功的应用^[2-4]。国内大量学者也已开始致力于近红外在中药过程分析领域的应用研究^[5-9]。由于中药生产过程生产工艺复杂, 药物组成复杂, 有效成分含量低, 变化范围大, 而不同类型近红外仪器具有不同的检测性能和适用范围, 所以近红外在中药过程分析中的应用尚存在许多问题。不同类型近红外仪器获得的光谱信息受外界环境影响的不一致和光谱响应值的不一致, 使不同 NIR 仪器对同一载体具有不同的定性和定量性能。如何根据特定载体, 选择适宜类型的近红外光谱仪用于中药过程分析是当前急需解决的关键问题之一。NIRS 技术类型较多, 根据其分光原理, 可分为全息光栅型(GT)、傅里叶型(FT)、声光可调型(AOTF)、微阵列型

(MEMS)等, 不同类型仪器具有不同的检测性能和适用范围。以清开灵注射液为载体, 对 GT, FT, AOTF 和 MEMS 四种不同类型近红外光谱技术用于中药体系的多变量检测限(MDL)进行研究, 以探讨不同类型近红外光谱仪的检测性能, 为不同需求下近红外仪器类型的选择提供参考。

1 实验部分

1.1 仪器和试剂

GT(a)型近红外光谱仪(公司保护), FT(b)型近红外光谱仪(公司保护), AOTF(c)型近红外光谱仪(公司保护), MEMS(d)型近红外光谱仪(公司保护)。1100 型高效液相色谱仪包括四元泵、真空脱气机、自动进样器、柱温箱、二极管阵列检测器(DAD)及 HP 数据处理工作站(美国 Agilent 公司)。色谱级甲醇(美国 Tedia 公司), 磷酸(天津大学试剂厂, 分析级), 纯净水(杭州娃哈哈集团有限公司)。黄芩苷对照品由中国食品药品检定研究院提供(批号: 110777-

收稿日期: 2012-00-00, 修订日期: 2012-00-00

基金项目: 国家科技部十一五重大新药创制专项(2010ZX09502-002, 2011ZX09201-201-24)资助

作者简介: 彭严芳, 女, 1987 年生, 北京中医药大学中药信息工程中心硕士研究生 e-mail: pengyf0816@163.com

* 通讯联系人 e-mail: yjqiao@263.net; WZS@bucm.edu.cn

201005), 6 批清开灵注射液样本由指定药厂提供。

1.2 光谱条件

a 型仪器光谱: 采用近红外透射模式, 以仪器内部的空气为背景, 分辨率为 0.5 nm, 光谱扫描范围 400~2 500 nm, 扫描次数 32 次, 每个样品平行测定 3 次, 取平均光谱。

b 型仪器光谱: 采用近红外透射模式, 以仪器内部的空气为背景, 分辨率为 8 cm⁻¹, 光谱扫描范围 4 000~10 000 cm⁻¹, 扫描次数 32 次, 每个样品平行测定 3 次, 取平均光谱。

c 型仪器光谱: 采用光纤模式, 分辨率为 2 nm, 光谱扫描范围 1 100~2 300 nm, 扫描次数 500 次, 每个样品平行测定 3 次, 取平均光谱。

d 型仪器光谱: 采用光纤透射模式, 分辨率为 1 nm, 光谱扫描范围 1 350~1 800 nm, 扫描次数 16 次, 每个样品平行测定 3 次, 取平均光谱。

1.3 色谱条件

SunFire C18 色谱柱 (4.6×150 mm, 5 μm, Waters); 流动相: 甲醇-水-磷酸(47:53:0.2); 检测波长: 276 nm, 流速: 1 mL·min⁻¹, 柱温: 30 °C, 进样量: 10 μL。

1.4 样品制备

取 6 批清开灵注射液(黄芩苷含量按 2010 年版《中华人民共和国药典》清开灵注射液项下高效液相色谱法测定), 每批样品均用纯净水稀释成系列浓度。共 118 个样本, 浓度范围为 1~6 mg·mL⁻¹, 分别取其中的 3 批样本作为训练集, 另 3 批样本作为预测集。

1.5 数据处理

运用 Unscrambler 7.0(挪威 CAMO 软件公司)软件对光谱进行预处理和模型计算。间隔偏最小二乘算法工具包由 Nørgaard 等提供的网络共享(<http://www.models.kvl.dk/source/iToolbox/>), 其余各计算程序均自行编写, 采用 MATLAB 软件工具(美国 Mathwork 公司)计算。

1.6 多变量检测限原理^[10-14]

本文采用基于两类误差统计学理论和误差传递理论得到的逆矩阵模型多变量检测限(multivariate detection limit, MDL)计算公式(1)来计算不同仪器的 MDL。该公式同时考虑了一类误差(假阳性)和二类误差(假阴性), 能够全面反映分析方法在一定置信限下被检出的最低浓度。

$$MDL = \Delta_{p,q} [(1+h)MSEC - \sigma_e^2] \quad (1)$$

其中, $\Delta_{p,q}$ 为在自由度 V 下, 非中心 t 分布在置信限为 p, q 时的分位数。 h 为未知样本的杠杆值, 可通过偏最小二乘(PLS), 主成分分析(PCA)等模型求得, $MSEC$ 为校正集均方差。为被分析物参考值测量误差。

2 结果与讨论

2.1 光谱预处理方法筛选

图 1 为清开灵注射液近红外原始光谱。由图 1 可知, 不同类型近红外仪器的光谱在同一波长点处的吸收值不一致, 谱图差异较大。随着清开灵注射液中间体中黄芩苷浓度的改变, d 型近红外光谱仪的光谱吸收值较其余三种近红外光谱仪变化大。四种不同类型仪器所得近红外光谱受光谱重叠、基线漂移和随机噪音等因素影响较大, 需对光谱进行预处

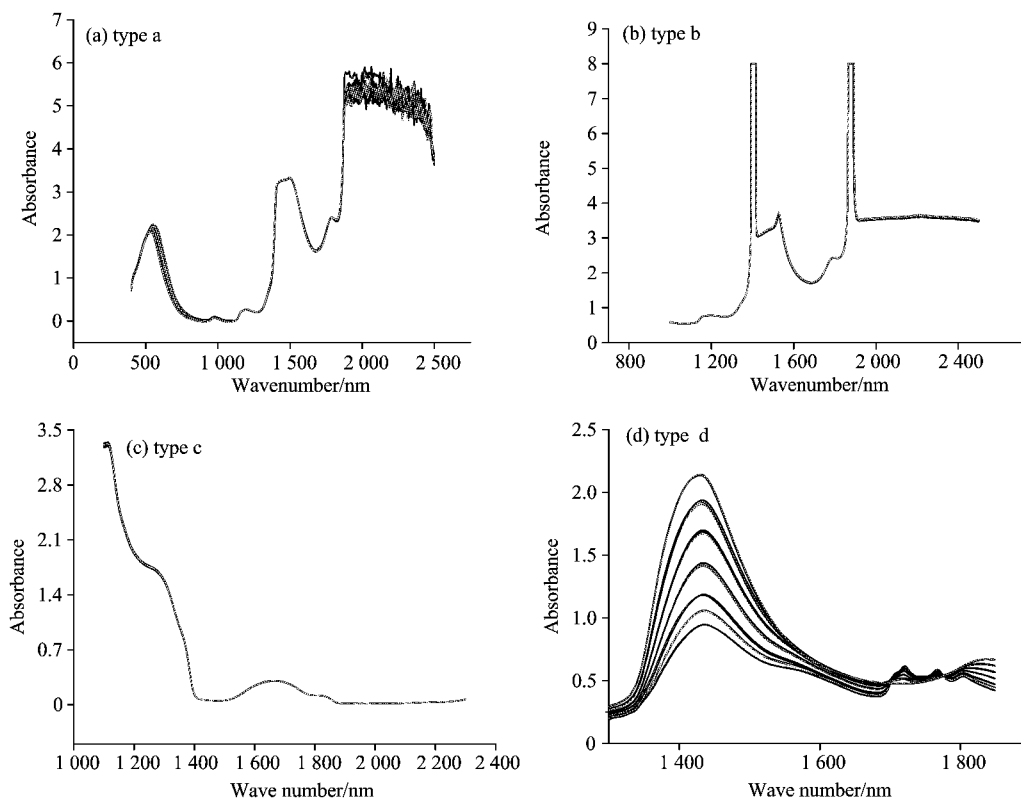


Fig. 1 Original NIR spectra

Table 1 Results of Qingkailing injection PLS models based on different pretreatment methods (a, b, c, d $\mu\text{g} \cdot \text{mL}^{-1}$)

Pretreatment	Latent factors	Calibration set		Validation set	
		r^2	RMSEC	r	RMSECV
a					
Raw	1	0.965 3	268.2	0.965 7	271.6
1D	4	0.967 6	259.2	0.960 1	289.1
2D	20	0.977 6	215.2	0.917 0	410.3
2D+SG	18	0.973 5	234.1	0.937 0	362.2
SG	2	0.965 4	268.2	0.965 0	272.2
SNV	1	0.965 6	267.1	0.965 3	268.3
1D+SG	3	0.959 9	288.7	0.9548	310.1
b					
Raw	12	0.985 2	177.5	0.968 9	261.8
1D	12	0.981 0	201.2	0.930 1	388.1
2D	3	0.810 5	636.2	0.801 3	655.3
1D+SG	11	0.980 5	204.1	0.908 3	445.2
SG	4	0.927 8	392.2	0.915 0	442.2
SNV	11	0.987 4	163.1	0.971 3	253.3
SNV+SG	9	0.985 6	175.7	0.976 9	233.1
c					
Raw	5	0.992 8	108.7	0.990 4	126.7
1D	4	0.989 8	125.2	0.982 1	176.1
2D	4	0.977 0	194.2	0.954 4	278.3
SNV+SG	3	0.988 3	138.8	0.984 2	165.4
Baseline	4	0.992 4	111.2	0.988 0	140.2
SNV	4	0.992 4	112.1	0.988 3	138.8
d					
Raw	1	0.998 2	77.5	0.997 7	92.6
1D	1	0.997 1	101.5	0.996 7	107.7
2D	1	0.996 2	104.7	0.996 3	116.7
SG	1	0.998 2	77.5	0.998 0	84.8
Baseline	1	0.998 4	74.6	0.998 2	81.4
SNV	1	0.997 2	99.1	0.996 4	119.9

理。本文考察了原始光谱(Raw)、一阶导数(1D)、二阶导数(2D)、Savitzky-Golay 平滑(SG)、标准正则变换(SNV)、基线校正(Baseline)等预处理方法及各预处理方法的组合对近红外 PLS 定量模型的影响。

四种类型光谱经不同预处理后所建偏最小二乘(PLS)模型的相关系数 r 和交叉验证均方根误差 RMSECV 值见表 1。由表 1 可知, a, b, c 和 d 四种仪器所得光谱分别采用 SNV, SNV+SG, RAW, Baseline 预处理方法时, 模型的 r^2 和 RMSECV 值最小。表明不同类型仪器的最佳预处理方式不同, 实际应用中应进行预处理方法的考察, 获得最佳预处理方法以确保模型的准确性。

2.2 偏最小二乘回归模型的建立

采用 2.1 节中最佳预处理方法分别建立不同类型仪器光谱的清开灵注射液中黄芩苷的 PLS 定量模型, 采用外部验证集评价模型预测性能。模型预测值与 HPLC 测定值的相关关系见图 2。由图 2 可知, a, b, c 和 d 四种仪器所得黄芩苷 PLS 模型的预测集相关系数 r_{pre} 和预测均方差 SEP 分别为

0.976 2 和 230.4 $\mu\text{g} \cdot \text{mL}^{-1}$; 0.956 1 和 246.4 $\mu\text{g} \cdot \text{mL}^{-1}$; 0.966 2 和 264.4 $\mu\text{g} \cdot \text{mL}^{-1}$; 0.998 5 和 71.5 $\mu\text{g} \cdot \text{mL}^{-1}$ 。以上结果表明四种技术类型所建 PLS 模型均具有良好的预测性能。说明了这四种类型近红外光谱仪均可用于清开灵注射液中黄芩苷含量的检测。d 型近红外光谱技术所建 PLS 模型能够获得比 a, b, c 更好的预测性能。

2.3 间隔偏最小二乘回归(iPLS)模型的建立

运用 iPLS 算法筛选近红外光谱特征波段, 比较不同区间划分对模型预测性能的影响, 结果如表 2 所示。表 2 中 a 型仪器光谱全谱模型潜变量因子为 10, iPLS 模型确定的最佳区间数为 22。模型的 r_{pre} 和 SEP 分别 0.977 1 和 218.4 $\mu\text{g} \cdot \text{mL}^{-1}$, 与 2.2 节中全谱 PLS 模型相比, iPLS 提高了模型性能; b 型仪器光谱全谱模型潜变量因子为 5, iPLS 模型确定的最佳区间数是 20。模型的 r_{pre} 和 SEP 分别为 0.975 4 和 219.4 $\mu\text{g} \cdot \text{mL}^{-1}$, 同样提高了模型性能。c 和 d 型仪器光谱采用 iPLS 方法未能筛选出最佳建模区间, 可能受仪器本身检测波段范围以及算法本身的局限的影响。iPLS 算法旨在将整个光谱以不同区间大小进行划分, 通过对每个区间建立 PLS 模型与全谱建模结果进行比较, 以筛选出最佳建模区间^[15]。而 c 和 d 两仪器光谱全谱原谱和基线校正后所建模型就有较好的模型性能, 划分区间可能导致有用信息丢失, 从而筛选不出最佳的建模区间。具体原因还需进一步研究。

2.4 不同类型近红外光谱技术的多变量检测限(MDL)研究

采用式(1)计算清开灵注射液中黄芩苷在不同类型近红外仪器的 MDL(式中 σ_c^2 参考值的采样误差忽略), 各类型近红外仪器在不同置信限下的 MDL 估计值如表 3 所示。当 p, q 均取 5% 时(原料药取 2%, 成药取 5%, 生物样品分析取 15%), 四种类型近红外光谱仪的 MDL 都低于 240 $\mu\text{g} \cdot \text{mL}^{-1}$, 能够满足清开灵注射液中间体快速定量分析的要求。由表 3 可知, iPLS 模型的 MDL 均低于相应 PLS 模型, 表明通过筛选波段可以得到比全谱建模更低的检测限, 从而提高模型预测性能。c 和 d 两种新类型 NIR 仪器 PLS 模型的多变量检测限估计值分别为 58 和 2.9 $\mu\text{g} \cdot \text{mL}^{-1}$, 远低于 a 和 b 两种类型仪器。表明技术装备的改进一定程度上克服了 NIR 技术检测限高的问题, 提高了 NIR 技术用于中药过程分析的可行性。近红外光谱仪分光系统从传统的滤光片型到全息光栅型和傅里叶变换型的转变, 使近红外光谱仪分光系统波长范围、波长准确性和扫描速度得到了很大的提高, 但可移动部件的存在使仪器带宽易受环境影响。而声光可调型分光器无可移动部件, 体积小, 更适于现场分析, 但易受温度和电磁波干扰^[16]。随着半导体集成技术的发展, 微机电系统型近红外光谱仪兼具重量轻、体积小、探测速度快、寿命长、可集成化、可批量制造以及成本低廉, 精度高等优点^[17], 适于在线分析。

3 结论

以清开灵注射液为载体, 考察了不同类型近红外仪器的 PLS 模型定量性能。其中 d 型近红外光谱仪获得了最佳模型预测性能, r_{pre} 和 SEP 分别为 0.998 5 和 71.5 $\mu\text{g} \cdot \text{mL}^{-1}$, 可

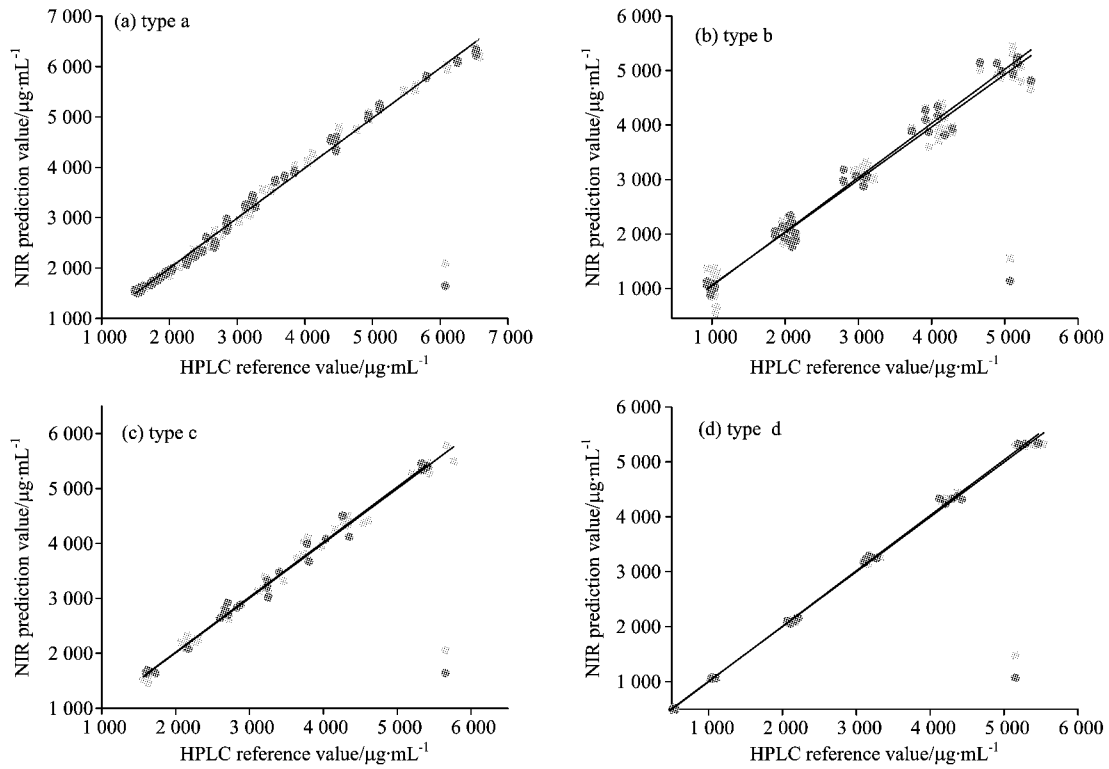


Fig. 2 Correlation graph of Near-infrared predictive value of the different types of NIR technology with HPLC reference value

Table 2 Predict ability of iPLS model of different NIR instrument in different Optimization interval ($\mu\text{g} \cdot \text{mL}^{-1}$)

Instrument a				Instrument b			
Interval number	Selected interval	Latent factors	RMSECV	Interval number	Selected interval	Latent factors	RMSECV
20	3	9	100.1	20	7	5	120.1
22	3	10	96.2	22	13	5	180.2
24	4	10	114.1	24	8	5	164.1
26	4	7	103.3	26	9	5	128.3
28	4	8	102.3	28	21	5	170.3
30	4	10	120.5	30	10	5	160.5
32	5	9	116.7	32	11	5	176.7
34	5	7	123.2	34	11	4	183.2
36	5	10	128.9	36	12	5	168.9
38	8	6	120.5	38	13	5	220.5
40	6	7	118.1	40	13	4	228.1

Table 3 The MDL of Baicalin in qingkailing injection from different NIR equipment type under different $\Delta_{p, q}$ ($\mu\text{g} \cdot \text{mL}^{-1}$)

Equipment type	Method	$\Delta_{0.1, 0.1}$	$\Delta_{0.1, 0.05}$	$\Delta_{0.1, 0.01}$	$\Delta_{0.05, 0.1}$	$\Delta_{0.05, 0.05}$	$\Delta_{0.05, 0.01}$	$\Delta_{0.01, 0.1}$	$\Delta_{0.01, 0.05}$	$\Delta_{0.01, 0.01}$
a	PLS	190	210	260	220	240	290	270	290	350
	iPLS	100	110	140	110	130	150	140	150	180
b	PLS	140	160	200	160	190	220	210	220	260
	iPLS	90	110	130	110	120	140	130	140	170
c	PLS	45	51	63	52	58	70	64	71	83
d	PLS	2.3	2.6	3.2	2.6	2.9	3.5	3.2	3.6	4.2

能因为 d 型仪器光谱所包含信息更加丰富,更能反映出目标成分含量变化对近红外光谱的影响,这从 d 型仪器所采集到的原始光谱中得到确证。

四种不同类型的近红外光谱技术在不同建模方法下的多变量检测限估计值均低于 $250 \mu\text{g} \cdot \text{mL}^{-1}$ 。其中, d 型近红外光谱技术多变量检测限估计值低至 $2.9 \mu\text{g} \cdot \text{mL}^{-1}$, 表现出

对复杂中药体系良好的适用性。同时, MDL 值的大小是近红外仪器性能评价的适宜指标。

不同类型仪器检测性能不同, 实际应用中, 应结合自身需求和经济性原则, 选择合适的仪器类型。本文所得到的结论能够对其它中药研究载体的研究提供一定的参考。作者声明不涉及商业用途。

References

- [1] OZAKI Y. *Analytical Sciences*, 2012, 28(6): 545.
- [2] Genkawa Takuma, Watari Masahiro, Nishii Takashi, et al. *Applied Spectroscopy*, 2012, 66(7): 773.
- [3] Krapf Lutz, Gronauer Andreas, Schmidhalter Urs, et al. *Journal of Near Infrared Spectroscopy*, 2012, 19(6): 479.
- [4] Xiong Haoshu, Gong Xingchu, Qu Haibin. *Journal of Pharmaceutical and Biomedical Analysis*, 2012, 70(11): 178.
- [5] JIANG Bo-hai, WANG Qing, WANG Shi-sheng, et al(姜博海, 汪晴, 王世盛, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2013, 33(1): 20.
- [6] CHEN Chen, LI Wen-long, QU Hai-bin, et al(陈 晨, 李文龙, 瞿海斌, 等). *Chinese Herbal Medicines(中草药)*, 2013, 44(001): 47.
- [7] Wu Z, Peng Y, Cheng W, et al. *Bioresource Technology*, 2013, 137: 394.
- [8] Xu Bing, Wu Zhisheng, Lin Zhaozhou, et al. *Analytica Chimica Acta*, 2012, 720: 22.
- [9] XU Bing, SHI Xin-yuan, QIAO Yan-jiang, et al(徐 冰, 史新元, 乔延江, 等). *China Journal of Traditional Chinese Medicine and Pharmacy(中华中医药杂志)*, 2012, 4: 3.
- [10] Alcalá Manel, León J Joshua, Roperó Jorge, et al. *Journal of Pharmaceutical Sciences*, 2008, 97(12): 5318.
- [11] Blanco Marcel, Castillo Miguel, Peinado Antonio, et al. *Analytica Chimica Acta*, 2007, 581(2): 318.
- [12] Boqué R, Larrechí M S, Rius F X. *Chemometrics and Intelligent Laboratory Systems*, 1999, 45(1): 397.
- [13] Clayton C A, Hines J W, Elkins P D. *Analytical Chemistry*, 1987, 59(20): 2506.
- [14] Wu Zhisheng, Sui Chenglin, Xu Bing, et al. *Journal of Pharmaceutical and Biomedical Analysis*, 2013, 77(4): 16.
- [15] Pereira A F C, Pontes M J C, Neto F F G, et al. *Food Research International*, 2008, 41(4): 341.
- [16] JI Hai-Yan(吉海彦). *Modern Scientific Instruments(现代科学仪器)*, 2001, 6: 25.
- [17] CHU Xiao-li, LU Wan-zhen(褚小立, 陆婉珍). *Analytical Instrumentation(分析仪器)*, 2008, 1: 3.

Multivariate Detection Limits of Baicalin in Qingkailing Injection Based on Four NIR Spectrometer Type

PENG Yan-fang, SHI Xin-yuan, ZHOU Lu-wei, PEI Yan-ling, LI Yang, WU Zhi-sheng*, QIAO Yan-jiang*
Research Center of Traditional Chinese Medicine Information Engineering, Beijing 100102, China

Abstract Multivariate detection limits (MDLs) of different types of near-infrared instruments have been investigated to guide the selection of device type for TCM NIR analysis. In this paper, Near-infrared spectroscopy of QingKaiLing injection was performed in transmission mode on four near-infrared spectrometers named a, b, c and d, respectively. High performance liquid chromatography (HPLC) was used as the reference method to determine the content of baicalin in QingKaiLing Injection. Partial least squares (PLS) and interval partial least squares (iPLS) quantitative models of baicalin in Qing Kai Ling Injection were established and MDLs of quantitative models based on different types of instruments were calculated. The determination coefficient of prediction (r_{pre}) and the root mean square errors of prediction (SEP) of PLS models of four different near-infrared spectrometers are 0.976 2 and $230.4 \mu\text{g} \cdot \text{mL}^{-1}$ (a), 0.956 1 and $246.4 \mu\text{g} \cdot \text{mL}^{-1}$ (b), 0.966 2 and $264.4 \mu\text{g} \cdot \text{mL}^{-1}$ (c), 0.998 5 and $71.5 \mu\text{g} \cdot \text{mL}^{-1}$ (d). And the model of instrument d shows a better prediction performance than the other three types. There are no remarkable superiorities in predictability in iPLS models of instruments a and b after variable selection, since the r_{pre} and SEP values for instruments a and b are 0.977 1 and $218.4 \mu\text{g} \cdot \text{mL}^{-1}$, 0.975 4 and $219.4 \mu\text{g} \cdot \text{mL}^{-1}$, respectively. Models c and d show no results of variable selection. MDLs ($\Delta_{0.05, 0.05}$) of different instruments are all less than $250 \mu\text{g} \cdot \text{mL}^{-1}$, and the MDLs of instruments c and d reach to 58 and $2.9 \mu\text{g} \cdot \text{mL}^{-1}$. The results reveal that the predictability of models and corresponding MDLs are different for different detection equipments. This paper innovatively used the theory of MDL to investigate the detection performance of different types of NIR instruments. The results demonstrated the feasibility of the approach. And it is expected

ted that in actual applications, choosing the right type of instrument should be based on the characteristics of the study carrier to ensure quantitative accuracy.

Keywords Near-infrared; Qingkailing injection ; PLS; iPLS; Multivariate detection limits

(Received 2012; accepted 2012)

* Corresponding author