Evaluation of the value of near infrared (NIR) spectromicroscopy for the analysis of glycyrrizhic acid in licorice

ARTICLE · APRIL 2015

DOI: 10.1016/S1875-5364(15)30022-4 · Source: PubMed

5 AUTHORS, INCLUDING:

Zhisheng Wu

Beijing University of Chinese Medicine and …

50 PUBLICATIONS **108** CITATIONS

SEE PROFILE

Available online at www.sciencedirect.com
ScienceDirect
Chinese Journal of Natural Medicines 2015, **13**(1): 0001–0005
doi: 10.3724/SP.J.1009.2015.00000
Chinese Journal of Natural Medicines

Evaluation of NIR spectromicroscopy for the analysis of glycyrrizhic acid in licorice

WU Zhi-Sheng, ZHOU Lu-Wei, DAI Sheng-Yun, SHI Xin-Yuan*, QIAO Yan-Jiang*

Beijing University of Chinese Medicine, Beijing 100102, China

Available online 20 Mar. 2015

[ABSTRACT] Previous work has established that hyperspectral data could be employed to qualitatively elucidate the spatial composition of Chinese medicinal plant tablets. To gain more insight into this technology, the quantitative profile provided by near infrared (NIR) spectromicroscopy was further studied by determining the glycyrrhizic acid content in licorice, *Glycyrrhiza uralensis*. Thirty-nine samples from twenty-four different origins were analyzed using NIR spectromicroscopy. Partial least squares, interval partial least square (iPLS), and least squares support vector regression (LS-SVR) method were performed to develop linear and non-linear calibration models. Optimal calibration parameters (number of interval numbers, kernel parameter, etc.) were explored. The root mean square error of prediction (RMSEP) and the coefficient of determination (R^2) of the iPLS model were 0.7177% and 0.9361 in the prediction set. The RMSEP and R^2 of LS-SVR model were 0.5155% and 0.9514 in the prediction set. These results demonstrated that the glycyrrhizic acid content in licorice could barely be analyzed by NIR spectromicroscopy. The conclusions revealed that good quality quantitative data was difficult to obtain from microscopic NIR spectra for a complicated Chinese medicinal plant material.

[KEY WORDS] NIR hyperspectral imaging; *Glycyrrhiza uralensis*; Partial least squares; Least squares support vector regression

[CLC Number] R917 **[Document code]** A **[Article ID]** 2095-6975(2015)03-0000-05

Introduction

As a new technique, near-infrared chemical imaging (NIR-CI) can acquire spatial distribution information in addition to the spectral information by combining conventional NIR spectroscopy with digital imaging [1-2]. As the saying goes, “a picture is worth a thousand words”. NIR-CI is a very attractive technique to obtain a great amount of spectral and spatial information of pharmaceutical products. However, data quality is especially critical for NIR-CI to be a useful and robust analytical method for Chinese materia medica (CMM).

If the quality of the hyperspectral data is poor, an inaccurate chemical image will be generated, which will lead to erroneous conclusions. Previous studies established that hyperspectral data could be effectively employed to qualitatively elucidate the spatial composition of Compound licorice tablets and Rukuaixiao tablets [3-4]. Further research on NIR-CI should extend to the quantification of an active ingredient, which is challenging in the CMM field. If NIR-CI could detect a low-dose active pharmaceutical ingredient (API) with excellent hyperspectral information (data), the application of NIR hyperspectral data will be greatly extended in the CMM system.

Though spatial information is of value, which is not the focus of this paper. To elucidate the quality of the NIR hyperspectral information, microscopic spectral data is the first and most important step [5-6]. Ravn *et al*. reported the comparison of common quantitative approaches on pharmaceutical solid dosage forms by NIR-CI. It was noted that three key points for a complete hyperspectral image analysis were important, i.e. the data quality and preprocess method [7]. Burger *et al* reported a comparison between microscopic spectral data and spectrometer data for quantitative information about organic and biological samples [8]. The results demonstrated that the prediction error of microscopic spectral data was between that of the two spectrometers data.

In this paper, the tasks were to obtain data sets of licorice from different origins using NIR spectromicroscopy,

[Received on] 27-Apr.-2014

[Research funding] This project was supported by the National Natural Science Foundation of China (No. 81303218), the Doctoral Fund of the Ministry of Education of China (No. 20130013120006), and the Beijing University of Chinese Medicine Special Subject of Outstanding Young Teachers and Innovation Team Foundation.

[*Corresponding author] E-mail: yjqiao@263.net (QIAO Yan-Jiang); xyshi@126.com (SHI Xin-Yuan)

These authors have no any conflict of interest to declare.

and to determine the glycyrrhizic acid content coupled with linear and non-linear calibration techniques [9]. Partial least squares (PLS) and least squares support vector regression (LS-SVR) were proposed to determine the glycyrrhizic acid content in licorice using NIR spectromicroscopy.

Materials and Methods

Materials

Samples of licorice were provided by the Institute of Medicinal Plant Development (Chinese Academy of Medical Science & Peking Union Medical College, Beijing, China). These materials were collected from twenty-four different geographical regions of China, and were identified as the root and rhizome from *Glycyrrhizia uralensis* Fisch. ex DC. by Prof. Wang Wenquan. Standards of glycyrrhizic acid were purchased from the National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China). The purity was above 98 %. HPLC-grade acetonitrile was purchased from Merck (Darmstadt, Germany). Deionized water was purified by a Milli-Q water system (Millipore Corp., Billerica, MA, USA). Other reagents were of analytical grade.

Preparation of solutions

A standard solution at a concentration of 0.1 mg/mL was prepared by dissolving accurately weighed glycyrrhizic acid and diluting with methanol. Water content and powder size factors should be avoided for sample analysis. All the samples were dried for 24 hours at 50 ℃ and ground in a blender, and the powder was sieved (60 mesh). Each sample (100 mg) of licorice powder was analyzed using NIR spectromicroscopy before HPLC analysis.

Sample solutions were prepared by dissolving licorice powder (100 mg) in a 100 mL amber volumetric flask and adding 70% ethanol (50 mL) The mixture was sonicated with a KQ-500GDV ultrasonic bath (250 W, 40 kHz) for 30 min (Kunshan Ultrasonic, Jiangsu, China). Then it was diluted to volume with 70% ethanol. These solutions were filtered through filter paper. The final solution was further filtered using a 0.45 μm filter directly into HPLC vials for immediate HPLC analysis.

Microscopic spectra acquisition and processing

The Kennard–Stone (KS) algorithm was used to divide the spectral data set into calibration and validation (50/50). The details about the glycyrrhizic acid content are as follows: range, 5.23−105.42 mg g^{-1}; mean, 38.63 mg g^{-1}; standard deviation, 22.86. Each sample of licorice powder (100 mg) was placed on a sample stage with a glass fragment, and manually centered and focused. Image cubes were acquired with a NIR line mapping system (Spectrum Spotlight 400 FT-IR and 400N FT-NIR Imaging Systems, PerkinElmer, UK). The spectrum was collected with an average of eight scans over the wavelength region 7 800−4 000 cm^{-1} using 16 cm^{-1} resolution. The field of view was set to 2.5 mm × 2.5 mm using pixel size 25 μm × 25 μm [4]. Each image cube containing 10 000 full NIR spectra was collected in 5 min.

Microscopic spectral data were analyzed and processed with ISys 5.0 Software (Malvern Instruments, Malvern, UK). Other data analysis was performed by Spotlight 400 software, Hyper View software (PerkinElmer) and home-made routines programmed in MATLAB code (MATLAB, The Mathworks, MA, USA).

HPLC apparatus

The same samples were analyzed after NIR-CI measurement using an Agilent 1100 HPLC System, comprised of a quaternary solvent delivery system, an on-line degasser, an auto-sampler, a column temperature controller and a diode-array detector (DAD) coupled with an analytical workstation (Agilent Technologies, Palo Alto, CA, USA). Samples were primarily separated on a Sunfire-C18 column (150 mm × 4.6 mm; 5 μm particles, Waters Co., USA).

Chromatographic conditions

The mobile phase was a mixture of acetonitrile (mobile phase A) and water containing 0.05% phosphoric acid (mobile phase B). The eluted gradient was used according to the following profile: the initial composition was 20% of mobile phase A; after 5 min this was changed, with a linear gradient, over 30 min, to 50% of mobile phase A, then reached 100% in 3 min; after 38 min this proportion of mobile phase A was returned to 80% in 3 min, and in 41 – 45 min increased to 20% A. The mobile phase flow rate was 1.0 mL per min. The DAD detector was operated at 248 nm. The column temperature was set at 25˚C. The injection volume was 10 μL. Chromatographic peaks were identified by comparing the retention time with standard material.

Results and Discussion

Quantitative analysis of glycyrrhizic acid by an HPLC method

The HPLC method was adopted to determine glycyrrhizic acid content in licorice, as guided by ChP [10]. Glycyrrhizic acid content was calculated by percentage (glycyrrhizic acid content/each total weighing). Figure 1 shows the typical HPLC chromatograms of the extraction solution and the glycyrrhizic acid standard solution. The retention time of glycyrrhizic acid in the sample solution was the same as the reference standard solution. The calibration curve of the HPLC method was investigated before the real sample analysis. The calibration curve exhibited good linearity ($Y = 5 \times 10^9 X - 25\,383$, $R^2 = 0.9990$) within the content range (0.019 6 – 0.147 μg).

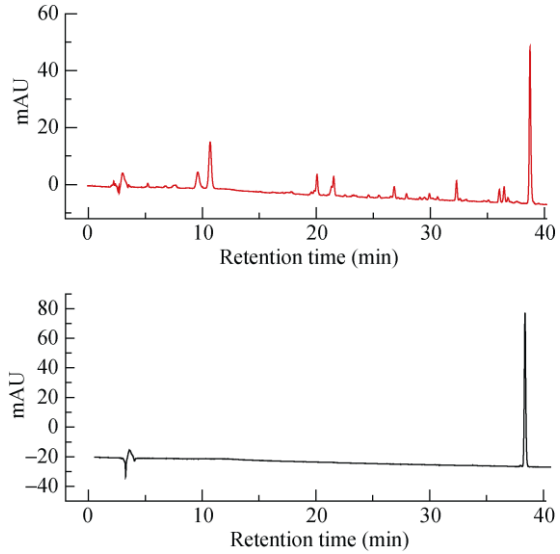

Fig. 1 Chromatograms of the extraction solution (a) and glycyrrhizic acid standard solution (b)

Spectra pre-treatments

Each pure NIR image was obtained with two spatial co-ordinates and one spectral coordinate. Therefore, a crucial step was unfolding the 3-D hyper cube. When hypercube (X×Y×λ) is unfolded, the two-way matrix (XY×λ) will be obtained. The mean spectra of each two-way matrix were obtained, as described in Figure 2. The trends of the mean spectra with different concentration levels were similar, and there were no characteristic peaks and valleys. However, there was some noise at the beginning and end parts of mean spectra. Hence, it required different pre-treatments for optimal spectra information.

As can be seen in Table 1, PLS models were established based on different preprocessing treatments. Relative good results were obtained using second derivative and nine latent factors. The calibration gave RMSE and R^2 values of 0.75% and 0.919 9. In the validation process, the RMSE and R^2 were 5.61% and 0.988 9. Second derivative was better because that it could reduce shifts and slope changes of the baseline and also improve the resolution of overlapping bands. But the calibration model was a little 'over-fitting' because the RMSE ratio value between validation set and calibration set was above 1.2. In addition, interference information was included in the full-spectral model, which leads to poor prediction results for the samples. The wavelength selection for the spectra denoted as iPLS model was investigated.

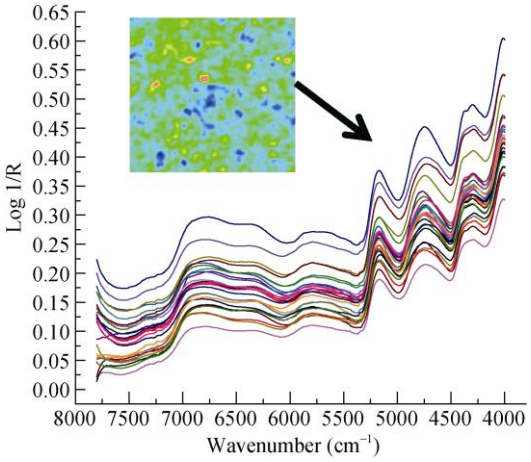

Fig. 2 Mean spectra of each RGB image

Linear model performance

The iPLS model was developed on spectral subintervals of equal width, and the prediction performance of these local models were compared with the global (full-spectrum) model. The comparison was mainly based on the RMSECV (root mean squared error of cross-validation, a segment size of five) parameter. The full-spectrum data-set was split into different intervals. Figure 3 shows that the performance of several interval models surpassed the full-spectrum model, and the interval number 18 showed the best results.

Table 1 The results for the PLS model with different pre-processing methods

Pretreatments	Latent factors	Calibration set		Validation set	
		R^2	RMSE	R^2	RMSE
Raw	2	0.5009	0.6137	0.2390	0.7788
1D	5	0.6878	0.4853	0.3062	0.7499
2D	9	0.9199	0.0750	0.6022	0.5610
SG	1	NOD	NOD	NOD	NOD
MSC	3	0.6735	0.4836	0.3275	0.7171
MSC+SG	1	NOD	NOD	NOD	NOD

*Raw: raw spectra, 1D: First derivative, 2D: Second derivative, MSC: Multiplicative signal correction, SG: Savitzky- Golay. NOD: Not detected

Figure 4 shows the iPLS result of 18 equidistant subintervals, where the bar plots of RMSECV values in each variable range are shown. Models were developed with one to nine latent variables. It could be observed that interval number, 1, 3, 4, 10, with 6, 5, 9, 9 latent variables, respectively, produced models with better performance than the full-spectrum model with RMSECV = 0.597%. The best interval was number 4, corresponding to 4 636.3–4848.2 cm^{-1}

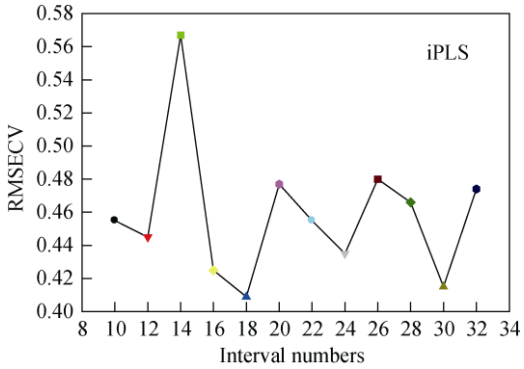

Fig. 3 Efficient interval numbers by iPLS model

with RMSECV = 0.409 %. Figure 5a illustrates the calibration and prediction results of the iPLS model. Implementation of variable selection showed improvement of model performance as the RMSE value decreased from 0.6137% to 0.2194%. The RMSEP and R^2 parameters were 0.7177% and 0.9361.

Nonlinear model performance

To obtain a robust LS-SVR model, two important parameters, i.e., ε and σ (kernel parameter the RBF kernel) need to be optimized. The parameter ε regulates the radius of the ε tube around the regression function and thus the number of support vectors is finally selected to construct the regression function (leading to a sparse solution). Too large a ε value results in less support vectors (more data points will fit in the ε tube) and, consequently, produce a smoother (less complex) regression function. The hyper-parameter values were selected by performing a simplex optimization based on 10-fold cross-validation for support vector-based methods. Finally, the LS-SVR model with ε =54.4 and σ^2 =4966.3, gave the RMSEP and R^2 parameters of 0.5155 % and 0.9514, respectively, as shown in Fig. 5b.

According to the analysis of Fig. 5, the overall results indicated that the glycyrrhizic acid content of licorice could barely be analyzed by NIR spectromicroscopy. There is a huge dispersion of the points, especially for the predictive data. The result demonstrated that good quality hyperspectral data for this CMM was not obtained for quantitative analysis by NIR spectromicroscopy. Additional research should be conducted in the future to thoroughly evaluate this novel technology in CMM.

Conclusions

In the present study, the overall results showed that glycyrrhizic acid content in licorice can barely be analyzed by

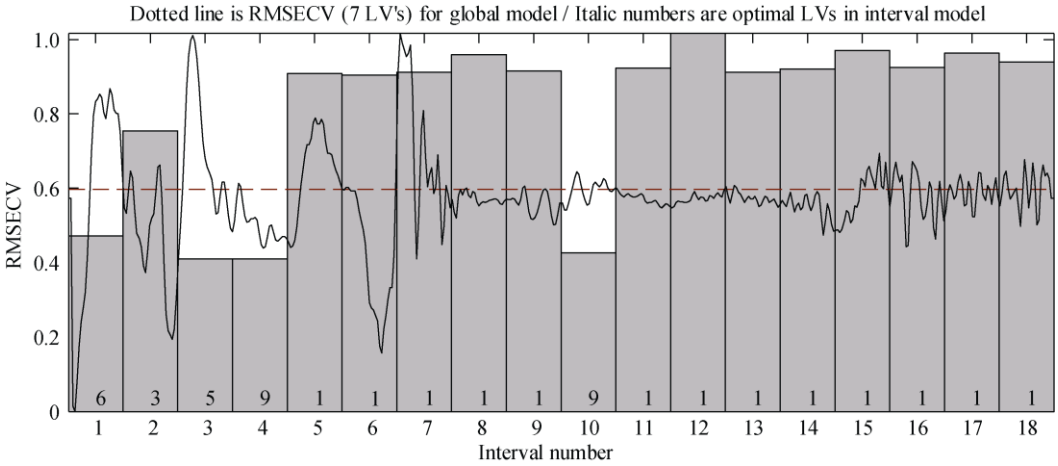

Fig. 4 Cross-validated prediction errors for 18 interval models (bars) and full-spectrum model (red dotted line) versus interval number for 1–9 latent variables in localized models and 7 latent variables for global model

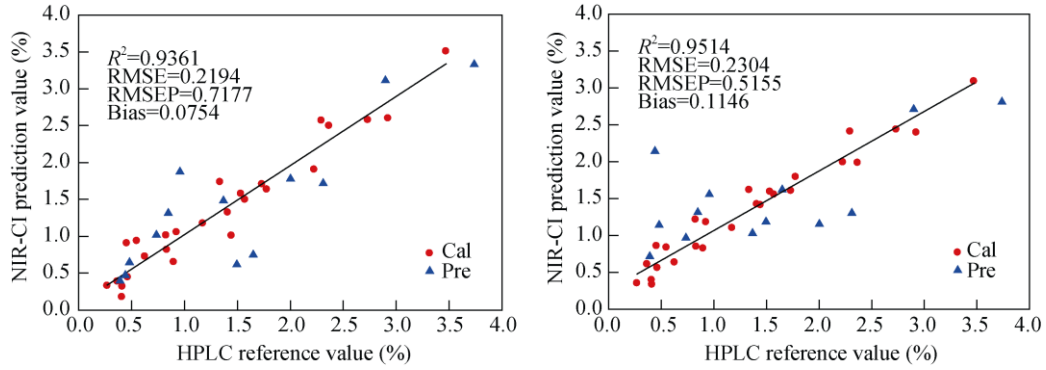

Fig. 5 Reference versus NIR spectromicroscopy predicted by iPLS model (a) and LS-SVR model (b)

linear and non-linear calibration techniques. Good data from spectromicroscopy are difficult to obtain in a complex matrix system. Therefore, additional research should be conducted in the future to thoroughly evaluate this novel technology, i.e., novel chemometric methods are needed to extract accurate quantitative information for use in CMM analysis.

References

[1] Gowen AA, O'Donnell CP, Cullen PJ, *et al*. Recent applications of chemical imaging to pharmaceutical process monitoring and quality control [J]. *Eur J Pharm Biopharm*, 2008, **69** (1): 10-22.

[2] Amigo JM, Cruz J, Bautista M, *et al*. Study of pharmaceutical samples by NIR chemical-image and multivariate analysis [J]. *Anal Chem*, 2008, **27** (8): 696-713.

[3] Wu ZS, Tao O, Cheng W, *et al*. Visualizing excipient composition and homogeneity of Compound Liquorice tablets by near-infrared chemical imaging [J]. *Spectrochim Acta A*, 2012, **86**: 631-636.

[4] Wu ZS, Tao O, Cheng W, *et al*. Visualizing excipient composition and homogeneity of Compound Liquorice tablets by near-infrared chemical imaging [J]. *Chin J Anal Chem*, 2011, **39** (5): 628-634.

[5] Gendrin C, Roggo Y, Collet C. Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review [J]. *J Pharm Biomed*, 2008, **48** (3): 533-553.

[6] Reich G. Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications [J]. *Adv Drug Deliver Rev*, 2005, **57** (8): 1109-1143.

[7] Ravn C, Skibsted E, Bro R. Near-infrared chemical imaging (NIR-CI) on pharmaceutical solid dosage forms-comparing common calibration approaches [J]. *J Pharm Biomed*, 2008, **48** (3): 554-561.

[8] Burger J, Geladi P. Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples [J]. *Analyst*, 2006, **131**: 1152-1160.

[9] Wang D, Ma ZH, Liu Z, *et al*. Comparative research on the NIR and MIR micro-imaging of two similar plastic materials [J]. *Spectrosc Spect Anal*, 2011, **31** (9): 2377-2382.

[10] *Pharmacopeia of People's Republic of China* [S]. Beijing, China: China Medical Science Press, 2010.