A Rapid Method to Determine the Harpagoside and Cinnamic Acid in *Radix Scrophulariae* by NIR Spectroscopy

Qun ZHOU,1 Shaqing CAI,2 Jianhua WANG,2 Suqin SUN1

1 Department of Chemistry, Tsinghua University, Beijing 100084, China
(E-mail: sunsq@chem.tsinghua.edu.cn)
2 Department of Pharmacognosy, School of Pharmaceutical Sciences, Peking University, Beijing, 100083, China

Quality control is utterly important to ensure efficient and safe use of medicinal herbs. The most important is the method should be fast, easy to use and non-destructively. 20 kinds of Xuanshen (*Radix Scrophulariae*), a common used medicinal herb in China, had been analyzed by Near-Infrared (NIR) Reflectance Spectroscopy. The contents of harpagoside and cinnamic acid, effective components of Xuanshen samples, were determined by High Performance Liquid Chromatography (HPLC). Also, chemometrics were served to process the HPLC data. Using partial least squares (PLS) regression model and pretreatment such as derivative and Standard Normal Variate (SNV), high correlation coefficient values (R²) and low standard errors of prediction (SEP) values were obtained. The correlation coefficient values of harpagoside and cinnamic acid were 0.998 and 0.999. Their SEP values were 0.00163 and 0.000488, respectively. This provides a rapid, non-solvent effects and non-destructive method to quantitative analyses some medicinal herbs.

(Received on August 9, 2001, Accepted on September 30, 2001)

Medicinal herbs, especially Chinese medicinal herbs, is a mystery to the world, for both the eastern and western countries. How to separate and determine the components, especially the effective components, is a key point in this area. This will involve the quality control of medicinal herbs. Xuanshen (*Radix Scrophulariae*) is a common used medicinal herb in China, which can nourish body's essential fluid and has detoxicating effect. It can cure the pyrexia, constipation, faecities and so on. According to the Chinese pharmacopeia, Xuanshen is the root of the Scrophularia nigrinae Hemsl., but the method of quality control has not been described. It is known that Harpagoside is a high content characteristic ingredient in Xuanshen. It is also reported that the Harpagoside has many efficacies such as reduce the blood pressure, ease pain, treat Hepatitis B Virus (HBV) and promote immunity. Cinnamic acid has laxative effect, it seems could explain that why Xuanshen could cure constipation. In this paper, an accurate method was established to determine the contents of Harpagoside and Cinnamic acid by High Performance Liquid Chromatography (HPLC). To avoid the solvent effects and time consuming, Near infrared spectroscopy (NIR) was served to process the HPLC data via chemometrics.

NIR is a powerful tool in the quantitative analysis. It has been routinely used in pharmaceutical industry and agricultural products for several decades. As the developing of NIR technology, many new ideas inspired and many interests aroused in pharmaceutical industry. Different from traditional quantitative method, such as HPLC, NIR could access various components in sample at the same time. This will be help to save time and decrease the consumer of the solution. Due to the non-destructive and no conventional preparation in the quality control, Near infrared reflectance analysis (NIRA) has been employed in the study of particles and tablets of medicines. Some medicinal herbs such as Ginseng and American Ginseng were also classified and quantified. Recently, Fourier Transform NIR (FT-NIR) made a great progress with its advantages; good resolution, high throughput and precision of wavelength and intensity. As a promising instrument, it can replace the conventional dispersive NIR completely. So the FT-NIR was chosen as the optimal instrument in our study.

In NIR, there are several ways to do the multi-content quantitative analysis, such as Multiple linear regression (MLR), Principal component analysis (PCA), Principal component regression (PCR) and Partial least squares (PLS). The Spectrum Quant+ (a software made by Perkin-Elmer Ltd.) calibration procedure is based on either a modified form of PCR or on PLS fit for one or more properties. PLS seeks to express the variation in the property information by correlating it with the spectral information, whereas PCR is a combination of PCA and MLR. In the PCA stage, PCR only seeks to account for variation in the spectral data and then in the MLR stage PCR correlates this with the property data. There are two algorithms in PLS: PLS1 and PLS2. For the PLS1 algorithm, each property is analyzed individually with respect to the spectral data. This means that if there is a high degree of correlation between properties, it is more efficient to use the PCR algorithms. Hence, only spectral information that relates to property variation will be accounted for in PLS. To model and predict the contents, a calibration set and a predict set were selected from the samples in PLS1 algorithms. Also in PLS1, the spectra are modeled by a different set of factors for each property and the contents are modeled by the respective factors. For PLS2 algorithm, all regression models are calculated simultaneously. PLS2 is used when significant amount of noise associated with the property values for multiple properties, as it will tend to average out. Since PLS2 could operate on all property data simultaneously, so correlation that is poor with one of the properties may affect other properties, resulting in larger than usual standard error of prediction. So the PLS1 algorithm was chosen to analyze the harpagoside and cinnamic acid contents in this paper.
Experimental

Sample preparation
The 20 kinds of commercial Xuanshen samples were purchased from 15 states in China. To avoid the factor of particle size, all the samples were grinded into powder in agate mortar, and screened through a 100-mesh sieve. To avoid the factor of sample cup oriental and particle size distribution, all the samples were put into a self-made quartz sample cup at the same direction.

Data Collection
The Xuanshen powders were determined by Perkin-Elmer System 2000 FT-IR system (Perkin-Elmer Limited, Beaconsfield Bucks., England) with a diffused reflection accessory (Beijing Second Optical Instrument Factory, Beijing, China). NIR source. The scanning range is from 3900 to 7000 cm\(^{-1}\), resolution 16cm\(^{-1}\), quartz beam splitter and MCT detector. The spectra were collected at 1 cm\(^{-1}\) interval by averaging 128 scans.

Data Analysis
All the Xuanshen samples were analyzed by Quant+ software Version 5.0. It is a multi-component analysis software which developed by Perkin-Elmer Ltd.. Some pretreatment was chosen to process the data, include derivative, smooth, Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV).

High Performance Liquid Chromatography (HPLC)
The Xuanshen powder (0.5g) was dissolved in 50ml methanol-water (3:7). Soak for 1 hour to extract Harpagoside and Cinnamic acid by using ultrasonic vibration in 30 minutes. After filtrated, the samples were ready to the HPLC test. The chromatographic peaks were identified by comparing retention times against harpagoside and cinnamic acid standards.

A Beckman System Gold Liquid chromatography was used to determine the Harpagoside and Cinnamic acid contents at the room temperature. The HPLC was equipped with a 126 pump and a 167 UV detector. The experimental conditions were as follows:
- Sample: 20 μL
- Column: UTRASPhere 5μm ODS quark-column (4.6mm × 45mm) and analytical column (4.6mm × 250mm) both from Beckman company
- Mobile phase: acetonitrile (LC grade, Fisher company) - water (containing 1.0% acetic acid) (20:80→50:50) (20minutes)
- Flow-rate: 1 mL/min
- The detection wavelength: 278 nm.

Result and Discussion

Method development and calibration
NIR spectra of 20 Xuanshen samples were collected (shown in Fig.1).

Then the Quant+ software was used to construct a method. 15 samples were selected to form the calibration set that covered the full range of the sample content. Other 5 samples were used to set up a validation (or prediction set) which could assess the prediction capability of the PLS1 model constructed from the calibration set. The HPLC results of Harpagoside and Cinnamic acid were shown in Table 1. The range of the harpagoside is from 0.0551% to 0.2442% and the mean±SD is 0.1182±0.0487 %; whereas the cinnamic acid is from 0.01170% to 0.05289% and the mean±SD is 0.03298±0.01177 %.

<table>
<thead>
<tr>
<th>No.</th>
<th>Sample code</th>
<th>Purchase Time (%)</th>
<th>Harpagoside (%)</th>
<th>Cinnamic acid (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1214</td>
<td>1997.8</td>
<td>0.1145</td>
<td>0.03108</td>
</tr>
<tr>
<td>2</td>
<td>1215</td>
<td>1997.7</td>
<td>0.1374</td>
<td>0.03358</td>
</tr>
<tr>
<td>3</td>
<td>1222</td>
<td>1997.9</td>
<td>0.1049</td>
<td>0.04718</td>
</tr>
<tr>
<td>4</td>
<td>1219</td>
<td>1998.2</td>
<td>0.1003</td>
<td>0.04113</td>
</tr>
<tr>
<td>5</td>
<td>1223</td>
<td>1997.8</td>
<td>0.2074</td>
<td>0.04184</td>
</tr>
<tr>
<td>6</td>
<td>1231</td>
<td>1998.2</td>
<td>0.0721</td>
<td>0.01291</td>
</tr>
<tr>
<td>7</td>
<td>1217</td>
<td>1997.12</td>
<td>0.1545</td>
<td>0.05518</td>
</tr>
<tr>
<td>8</td>
<td>1218</td>
<td>1997.12</td>
<td>0.1119</td>
<td>0.03872</td>
</tr>
<tr>
<td>9</td>
<td>1226</td>
<td>1998.2</td>
<td>0.1153</td>
<td>0.05247</td>
</tr>
<tr>
<td>10</td>
<td>1230</td>
<td>1998.2</td>
<td>0.0668</td>
<td>0.01804</td>
</tr>
<tr>
<td>11</td>
<td>1216</td>
<td>1998.2</td>
<td>0.1743</td>
<td>0.03098</td>
</tr>
<tr>
<td>12</td>
<td>1220</td>
<td>1998.2</td>
<td>0.0700</td>
<td>0.02542</td>
</tr>
<tr>
<td>13</td>
<td>1228</td>
<td>1998.1</td>
<td>0.1270</td>
<td>0.03781</td>
</tr>
<tr>
<td>14</td>
<td>1229</td>
<td>1998.2</td>
<td>0.0788</td>
<td>0.02854</td>
</tr>
<tr>
<td>15</td>
<td>1221</td>
<td>1997.8</td>
<td>0.2442</td>
<td>0.03048</td>
</tr>
<tr>
<td>16</td>
<td>1232</td>
<td>1998.2</td>
<td>0.1433</td>
<td>0.05289</td>
</tr>
<tr>
<td>17</td>
<td>1225</td>
<td>1998.2</td>
<td>0.0869</td>
<td>0.02947</td>
</tr>
<tr>
<td>18</td>
<td>1227</td>
<td>1998.2</td>
<td>0.0551</td>
<td>0.01170</td>
</tr>
<tr>
<td>19</td>
<td>1224</td>
<td>1998.2</td>
<td>0.0785</td>
<td>0.01894</td>
</tr>
<tr>
<td>20</td>
<td>1233</td>
<td>1998.8</td>
<td>0.1212</td>
<td>0.04126</td>
</tr>
</tbody>
</table>

To get better results, some pretreatment were used to analysis the samples.

1) Baseline correction. Baseline correction could offset removes flat deviations from zero absorbance at baseline position. Derivative is a useful way for baseline correction. As everyone knows, derivative spectra usually have sharper features than the original spectra. In quantitative analysis, they are used to reduce the effects of overlapping bands. The first derivative removes any baseline offset and the second derivative also removes any linear slope at the expense of adding noise. Some people believe that derivative even helps “resolve” spectral features. Since second-derivative spectra have sharp minima where there are maxima in the original spectrum and so can be used to identify band positions in complex regions.

2) Normalization. Normalization is ordinate multiplication of the spectrum to correct for differences in pathlength, or other linear changes with absorbance). Multiplicative Scatter Correction (MSC) is a normalization option which can compensate for wavelength-dependent light scattering variations encountered during reflectance spectroscopy. These variations can be seen as both an offset and a slope within the spectrum. With MSC, the scatter of each standard is estimated relative to the mean standard and normalized such that each
The Standard Error of Estimate (SEE) for the regression gives an indication of the quality of fit of the regression. It can be described by following equation:

$$ SEE = \sqrt{ \frac{RSS}{n-k-1} } $$

The SEP estimate (SEP Est.) gives an estimate of the standard error of prediction (SEP), that is, the magnitude of the error expected when independent samples are predicted using the model. In effect, a standard is removed and a model built using the other ns-1 standards. The removed standard is then predicted using this model. This is done for each standard in the calibration set. The SEP Est. is calculated from:

$$ SEP_{\text{estimate}} = \sqrt{ \frac{\sum (\hat{P}_i - P_i)^2}{n-1} } $$

This formula provides a very fast way of estimating the error to be found in the property values when predicting.

Obviously, the best condition of pretreatment could be selected from table 2 and table 3. Due to the excellent signal-to-noise ratio of FT-NIR and more than 100 scans, smooth is not an effective way to optimize the model. So smooth is usually not necessary in FT-NIR. To emphasize the normalization of the spectra, MSC and SNV were used. But these two ways were also could not improve the correlation coefficient. Hence the size of Xuanshen particles was a very important factor. Since it could not be homogeneous, the scattering must be occurred in NIR reflection spectra. To depress the multiplicative interference of scatter, the derivative, both first and second, with (or without) SNV were good choices. By using those pretreatments, the correlation coefficient values of harragoside and cinamnic acid were up to 0.998 and 0.999. Their SEP values were down to 0.00163 and 0.0000488, respectively. The final results of NIR predicted versus HPLC determined of the two components were shown in Fig.2 and Fig 3.

![Fig. 2 The plot of Estimated (NIR predicted) versus Specified (HPLC Determined) of harragoside content, which multiple correlation is 0.9993.](image-url)
Xuanshen, one of medicine herbs in China, was quantitatively analyzed by HPLC and NIR. This is only a tentative experiment in Chinese medicinal herbs. HPLC provided a new way to separate the Xuanshen samples, and contents of two effective component harpagoside and cinnamic acid were determined. By processing HPLC data, it can be known that NIR spectroscopy is also available in the quantitative analysis of Xuanshen. Compare to the routine method HPLC, NIR spectroscopy also presents some advantages: fast (get the result of two components at the same time), non-destructive (need not sample prepare), and accuracy (no solution effects). NIR will be a powerful and promising tool in the quantitative control of medicinal herbs. Hence, the medicine herbs could be used routinely and safely by the people all over the world someday in the near future.

References
4. Chang, Il Moo, Ger Offen DE 4322567(CIA61K31/70) 20 Jan 1994; CA, 1994, 120: P208577x