COMBINATION UV-VIS SPECTROSCOPY AND PARTIAL LEAST SQUARE FOR DETECTING ADULTERATION PARACETAMOL AND PIROXICAM IN TRADITIONAL MEDICINES

Pretty Falena Atmanda Kambira, Dion Notario*, Untung Gunawan, Sherina Dhamayanti, Raymunanda W.K. Ningrum, Sondang G.O. Ambarita, Grafty Polin

Department of Pharmacy, Faculty of Medicine and Health Sciences, Atma Jaya Catholic University of Indonesia, Jalan Pluit Raya no 2, Jakarta Utara, 14440, Indonesia

Received January 11, 2020; Accepted March 17, 2020

ABSTRACT

An analytical method based on combination UV-Vis spectroscopy and chemometric was developed for detecting commonly listed adulterants such as paracetamol and piroxicam simultaneously in traditional medicines. No complex sample preparation and separation are required except grinding, dissolving, and filtering. The spectral interferences were resolved by multivariate techniques. Wavelengths selection and number of components optimization were performed by a combination of Genetic Algorithm and Partial Least Square (GA-PLS) followed by backward elimination through Jack-Knife Partial Least Square Regression (JK-PLSR). The capability PLSR model for quantitative analysis was assessed from the coefficient of determination (R²) and root mean square of error prediction/cross-validation (RMSEP/RMSECV) dan predicted residual sum of square (PRESS). Classification performance of PLS Discriminant Analysis (PLS-DA) was evaluated from the area under the receiver operating characteristic curve (AUROCC). For ensuring the sensitivity of the method, the detection limits from the two pseudo-univariate lines were estimated. The R², RMSEP, RMSECV, AUROCC, and detection limit obtained from the selected models of paracetamol and piroxicam were >0.999, <0.25 mg/L, <0.15 mg/L, 100%, and <0.4 mg/L respectively. Therefore, the proposed method is suitable for the rapid screening of adulterated herbal medicine.

Keywords: adulteration; paracetamol; partial least square; piroxicam; UV-vis spectroscopy.

INTRODUCTION

Traditional herbal medicines have been used by the Indonesian population for a long time on any occasions whether to boost immune systems or improve health or to treat diseases. The abundant source of Indonesian medicinal plants makes traditional herbal medicine still commercially available and consumed today. However, to enhance its efficacy and boost sales, chemical adulterants are often added by some manufacturers to their herbal products. Indonesian National Agency of Drug and Food Control (NADFC) found an illegal traditional medicine production facility in Tangerang, West Java, in June 2019. The manufacturer added paracetamol as a chemical adulterant in their products (BPOM, 2019). Paracetamol and piroxicam are listed as common chemical adulterants in traditional herbs for treating pain, gout, and rheumatic (Ching et al., 2018).

Consumption of adulterated herb leads to therapeutic error and may cause the patient to worsen dependent on what kinds of adulterant(s). Paracetamol may lead to hepatotoxicity caused by its reactive metabolite N-acetyl-para-benzo-quinone imine. Piroxicam, a non-steroid anti-inflammatory drug (NSAID), induces peptic ulcer and gastrointestinal hemorrhages (Drini, 2017; Tittarelli et al., 2017). Some adverse events associated with NSAIDs were reported in
patients who took adulterated herbal medicines, including Cushing Syndrome (Xu et al., 2018). Most of the chemical adulterants are mixed homogeneously in traditional herb preparation, so it is difficult to distinguish an adulterated powdered herb sample without modern analytical techniques. Some analytical methods for detecting paracetamol and piroxicam in traditional herb preparation were developed, such as thin-layer chromatography (Mustariche et al., 2017) and liquid chromatography coupled with tandem mass spectrometry (Ching et al., 2018; Kim et al., 2014). However, the previous methods were time-consuming due to complicated sample preparation steps, not eco-friendly because they used toxic organic solvent, and not practical for initial screening since they required complex instrumentations. For this reason, a new simple analytical method is necessary to provide rapid detection of these adulterants.

Ultraviolet-visible (UV-Vis) spectroscopy provides a rapid and sensitive detection based on light-absorbing properties of chemical adulterants in the UV-Vis region. Lack of specificity due to spectral interferences can be resolved by multivariate techniques (Contreras et al., 2010). The strength of combination spectroscopy and chemometric techniques are no separation processes required, more rapid processing provided, and non-toxic solvent used so that moving forward to the green chemistry. However, the appropriate wavelength selection is essential to generate an accurate model.

The visual observations by comparing multispectral data are practical and commonly used to perform variables selection (Rahayu et al., 2018; Rohman et al., 2017), but the holistic changes in spectra cannot be elucidated clearly by visual observation only (Wang et al., 2017). The stepwise variable elimination/addition according to Jack-Knife Partial Least Square Regression (JK-PLSR) was recommended as the best variable selection technique based on the significance test (Westad and Martens, 2000). Nevertheless, the selection of optimal variables with stepwise elimination/addition is time-consuming due to the variables elimination/addition that must be done each in order. The limitation of stepwise variable elimination/addition can be overcome by performing Genetic-Algorithm-PLS (GA-PLS). Based on natural selection, the variables with a higher likelihood to “persist” (fit to the data) would be selected (Mehmood et al., 2012). GA-PLS would be able to screen a large number of variables, then stepwise variable elimination/addition by JK-PLSR would be performed more efficiently. In this study, we apply GA-PLS and JK-PLSR to the UV-Vis spectroscopy data for screening and eliminating non-significant variables at once to build the optimum quantitative models. The absorbances of the selected wavelengths were re-processed with PLS Discriminant Analysis (PLS-DA) for identifying adulterated and non-adulterated samples (qualitative analysis). As far as we know, the wavelength selection strategy through combination GA-PLS and JK-PLSR has not yet been implemented in UV-Vis spectroscopy to analyze a mixture of paracetamol and piroxicam in herbal medicine. Therefore, we applied these chemometric techniques in combination with UV-Vis spectroscopy to develop a fast method for the detection of paracetamol and piroxicam in traditional medicine based on the pattern recognition of the UV-Vis multispectra data.

**METHODS**

Chemicals NaOH (Merck, Germany), piroxicam (98.5%, Putra Bakti Niaga, Indonesia), paracetamol (99.9%, Putra Bakti Niaga, Indonesia), double distilled water, drug-free herbs, and commercial samples obtained from local markets in Jakarta and Tangerang in August-September 2019.

**Instrumentation**

The absorbances of calibration and sample solutions were measured by Shimadzu UV-1800 UV/VIS spectrophotometer with 1 cm path length quartz cell under scan-mode from 200 – 500 nm with an interval of 1 nm. Data acquisition was conducted by UVProbe.
Combination UV-Vis Spectroscopy And Partial Least Square...

2.52 (Shimadzu) software and saved in .spc file format until the next analysis step.

**Calibration solutions**

Five hundred mg piroxicam and paracetamol reference standards were weighed at the operating range of an analytical balance, dissolved in 0.1 N NaOH, and diluted with double distilled water in an appropriate volumetric flask to obtain concentration about 0.25 mg/mL accurately. Meanwhile, a blank sample solution, extracted from drug-free herbs powder, was prepared in the same way as the standard solutions. Standard solutions and blank solutions were mixed in certain proportions and diluted to the desired concentration previously designed by central composite design (CCD) with additional sampling points (Table 1).

**Table 1.** Calibration curve design by central composite design and additional sampling points

<table>
<thead>
<tr>
<th>run.order</th>
<th>x1.coded</th>
<th>x2.coded</th>
<th>x3.coded</th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>1.414</td>
<td>15</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>-1.000</td>
<td>-1.000</td>
<td>-1.000</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
<td>0.000</td>
<td>-1.414</td>
<td>15</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>1.000</td>
<td>1.000</td>
<td>-1.000</td>
<td>20</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>-1.414</td>
<td>0.000</td>
<td>0.000</td>
<td>8</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>1.000</td>
<td>-1.000</td>
<td>-1.000</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>-1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>10</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>-1.000</td>
<td>1.000</td>
<td>-1.000</td>
<td>10</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>0.000</td>
<td>1.414</td>
<td>0.000</td>
<td>15</td>
<td>22</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>1.000</td>
<td>-1.000</td>
<td>1.000</td>
<td>20</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>0.000</td>
<td>-1.414</td>
<td>0.000</td>
<td>15</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>14</td>
<td>-1.000</td>
<td>-1.000</td>
<td>1.000</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>1.414</td>
<td>0.000</td>
<td>0.000</td>
<td>22</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

**Additional sampling points**

| 16 | NA | NA | NA | 20 | 0  | 22 |
| 17 | NA | NA | NA | 20 | 0  | 22 |
| 18 | NA | NA | NA | 20 | 0  | 22 |
| 19 | NA | NA | NA | 20 | 23 | 0  |
| 20 | NA | NA | NA | 20 | 23 | 0  |
| 21 | NA | NA | NA | 20 | 23 | 0  |
| 22 | NA | NA | NA | 20 | 0  | 0  |

x1 = approximate concentration of drug-free herbs in working solution (mg/L); x2 = concentration of piroxicam in working solution (mg/L); x3 = concentration of paracetamol in working solution (mg/L)
Sample preparation
Samples were homogenized and approximately 500 mg homogenates were weighed in the operating range of an analytical balance, dissolved in 0.1 N NaOH, and diluted with double distilled water to obtain concentration about 10-20 mg/L. Each sample was analyzed in triplicates.

Data preparation and statistical analysis
Multispectral data in .spc format were converted to .csv and exported to R software for further analysis. Calibration data were divided into two groups labeled as data train and data test. Data train (n=17) were used for generating a regression model by PLSR technique and data test (n=5) for performing external validation. Wavelength selection was conducted by GA-PLS and optimized with backward elimination after performed a t-test for each wavelength by JK-PLSR. Root mean square error of cross-validation (RMSECV) and prediction (RMSEP) was calculated by applying eq. 1 to leave-one-out crossvalidation (LOO-CV) of data train and back-calculation from data test respectively. The lowest number of the components and wavelength resulting coefficient of determination (R²) above 0.999, smaller RMSECV/RMSEP, and lowest predicted residual sum of square (PRESS) was selected (Rohman et al., 2016). For discriminating between adulterated and non-adulterated samples, qualitative analysis based on PLS-DA was performed and visualized in biplot of two first principal components. The sensitivity of the proposed method was assessed by calculating the limit of detection from pseudo-univariate linear regression (LODpu) according to eq. 2 (Allegrini and Olivieri, 2014; Oleneva et al., 2019). All of the statistical operations including variable selection, PLSR, and PLS-DA calculation performed by plsVarSel, pls and mixOmics package installed in R 3.6.0 (Mehmood et al., 2012; Mevik and Wehrens, 2007; R Core Team, 2019; Rohart et al., 2017).

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_{predicted} - y_{actual})^2}{N}} \quad \text{.........1}
\]

RMSE is the root mean square error (mg/L); \(y_{predicted}\) and \(y_{actual}\) are predicted and calculated drug concentration in spiked standard solution (mg/L); \(N\) is the number of calibration samples.

\[
LOD_{pu} = 3.3s_{pu}^{-1}\left(1 + h_{0\text{min}} + 1/I\right)\text{var}_{pu}^{1/2} \quad \text{.........2}
\]

The terms \(LOD_{pu}\), \(s_{pu}\), and \(\text{var}_{pu}\) are limit of detection, slope, and residual variance calculated from the pseudo-univariate line between \(y_{predicted}\) and \(y_{actual}\) respectively; \(h_{0\text{min}}\) is the lowest projected leverage; \(I\) is the number of the calibration points.
RESULTS AND DISCUSSION

A traditional medicine preparation adulterated with pharmaceutical substances such as paracetamol and piroxicam is challenging to detect just by relying on visual observation. Nevertheless, the light-absorbing properties of paracetamol and piroxicam in aqueous solution can be used to identify them in herbal medicine. The molecular structure of paracetamol and piroxicam contain conjugated double bonds that π electrons are in. Conjugated double bonds are essential for absorbing light in the UV-Vis region (chromophore) due to π→π* electronic transition (Rafi et al., 2018). Besides, the presence of the auxochrome group (-OH) in the chromophore system may strengthen the light-absorbing properties of paracetamol and piroxicam. Differences in a chromophore-auxochrome system of paracetamol, piroxicam, and endogenous compound in herbs cause variation of spectral patterns (Figure 1), from which adulterated and non-adulterated herbal medicine are discriminated.

In the case of adulteration of traditional edicines with chemicals, the composition of natural ingredients with counterfeiting materials correlates with each other. When chemicals are added in a higher proportion, then the ratio of natural component is lower and vice versa. This correlation can lead to prediction errors due to two dependent variables correlated with each other. The use of an appropriate experimental design aims to prevent the effects of undesirable correlations within dependent variables due to the components are arranged in a structured and statistically unrelated way (Tomuta et al., 2017). At present, some types of experimental designs that were used initially to optimize responses such as CCD are widely utilized to design calibration curves in multivariate analysis (Hassaninejad-Darzi et al., 2016; Pinto et al., 2019).

The relationship between the concentration of paracetamol and piroxicam with absorbance in multiple wavelengths were modeled by PLSR. The first model of paracetamol and piroxicam which all wavelengths included showed a good fit to data; however, the accuracy of prediction could be improved by optimizing wavelength selection (Table 2). Wavelength selection through GA-
PLS was able to extract more than one hundred variables to be less than 20 variables (Mehmood et al., 2012) so that it was more efficient to re-optimize variables selection through backward elimination after jack-knifing (Mevik and Wehrens, 2007). The number of components also optimized by assessing the plot between RMSECV/RMSEP and PRESS against the number of components. The number of components resulting in minimum RMSECV/RMSEP and PRESS was selected (Table 2).

Table 2. Validation parameters of PLSR

<table>
<thead>
<tr>
<th>Model</th>
<th>Compound</th>
<th>Wavelengths (nm)</th>
<th>Ncomp</th>
<th>RMSECV</th>
<th>RMSEP</th>
<th>R² (CV)</th>
<th>R² (EV)</th>
<th>PRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Piroxicam</td>
<td>all wavelengths</td>
<td>4</td>
<td>0.541</td>
<td>0.326</td>
<td>0.9952</td>
<td>0.9984</td>
<td>4.98</td>
</tr>
<tr>
<td>2</td>
<td>Piroxicam</td>
<td>213*, 298*, 351, 353*, 365*, 369*, 393*, 398*, 404*, 414*, 417*, 476</td>
<td>4</td>
<td>0.141</td>
<td>0.208</td>
<td>0.9997</td>
<td>0.9993</td>
<td>0.34</td>
</tr>
<tr>
<td>3</td>
<td>Piroxicam</td>
<td>213*, 298*, 351, 353*, 365*, 369*, 393*, 398*, 404*, 414*, 417*</td>
<td>4</td>
<td>0.124</td>
<td>0.217</td>
<td>0.9997</td>
<td>0.9993</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td>Piroxicam</td>
<td>213*, 298*, 353*, 365*, 369*, 393*, 398*, 404*, 414*, 417*</td>
<td>4</td>
<td>0.125</td>
<td>0.212</td>
<td>0.9997</td>
<td>0.9993</td>
<td>0.26</td>
</tr>
<tr>
<td>5</td>
<td>Paracetamol</td>
<td>all wavelengths</td>
<td>3</td>
<td>0.225</td>
<td>0.362</td>
<td>0.9991</td>
<td>0.9979</td>
<td>0.86</td>
</tr>
<tr>
<td>6</td>
<td>Paracetamol</td>
<td>239*, 253*, 263*, 335, 351*, 378, 381, 455, 500</td>
<td>4</td>
<td>0.090</td>
<td>0.128</td>
<td>0.9999</td>
<td>0.9997</td>
<td>0.14</td>
</tr>
<tr>
<td>7</td>
<td>Paracetamol</td>
<td>239*, 253*, 263*, 335*, 351*, 381, 455, 500</td>
<td>4</td>
<td>0.090</td>
<td>0.128</td>
<td>0.9999</td>
<td>0.9997</td>
<td>0.14</td>
</tr>
<tr>
<td>8</td>
<td>Paracetamol</td>
<td>239*, 253*, 263*, 335*, 351*, 381, 455</td>
<td>4</td>
<td>0.096</td>
<td>0.128</td>
<td>0.9998</td>
<td>0.9997</td>
<td>0.16</td>
</tr>
<tr>
<td>9</td>
<td>Paracetamol</td>
<td>239*, 253*, 263*, 335*, 351*, 381*</td>
<td>3</td>
<td>0.087</td>
<td>0.157</td>
<td>0.9999</td>
<td>0.9996</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*Variable is significant at 95% confidence interval (p<0.05); CV = cross validation; EV = external validation; Ncomp = number of components; RMSECV = root mean square error of cross validation; RMSEP = root mean square error of prediction; PRESS = predictive residual sum of square; Selected models were typed in bold letter.
Combination UV-Vis Spectroscopy And Partial Least Square...

The goodness of fit of the multivariate regression is assessed by considering $R^2$, RMSEP, RMSECV, and PRESS values. Generally, the acceptance criteria of $R^2$ and RMSEP/RMSECV are above 0.999 and less than 0.4, respectively (Rohman and Man, 2010). Internal and external validation parameters showed a high linear correlation from the selected models, and no overfitting was found due to the significance of all variables in the selected model (Table 2 & Figure 2). Variable wavelength selection through GA-PLS followed by jack-knifing and backward elimination proved capable of generating a more accurate calibration model than using all of the wavelengths without selection and overcoming visual wavelength selection limitation, which very subjective wavelength selection is performed. The optimum wavelength selection resulted in a calibration model with high accuracy so that no data pre-processing such derivative curve, smoothing, or mean centering required.

The sensitivity of the models was assessed by calculating the detection limit from the pseudo-univariate line. The detection limit values of paracetamol and piroxicam were found in a low concentration of 0.23 and 0.36 mg/L, respectively. A low detection limit showed that the proposed analytical method is considered sensitive for detecting paracetamol and piroxicam commonly added to herbal preparation in medium to high proportion.

Classification of multispectral data conducted by PLS-DA at the selected wavelengths showed an excellent performance. The high discriminant capacity of the PLS-DA model was confirmed by the receiver operating characteristic (ROC) curve located closer to the top left corner and quantitatively represented by the area under the ROC curve (AUROCC) (Hajian-Tilaki, 2013). Based on AUROCC values (Figure 3), the chances of the PLS-DA model to discriminate spectral patterns between one counterfeit to the others were about 77 – 83%, while non-adulterated with adulterated samples was distinguished correctly (AUROCC = 1). Therefore, this method can be considered selective for determining adulterated and non-adulterated samples. The two first principal components carried out about 98% variances, which means almost all variance of the data can be explained by PLS-DA. There was shown that adulterated and non-
adulterated herbs could be distinguished with good separation on the biplot of two first principal components. In addition, based on our findings, there was no adulteration found in the commercial samples (Figure 4).

![Graph showing ROC curve](image)

**Figure 3.** Receiver operating characteristic (ROC) curve of the PLS-DA model. Blank samples (BLANK), spiked samples with paracetamol (ADD01), piroxicam (ADD02) and both paracetamol and piroxicam (ADD12)

![Graph showing multispectral data classification](image)

**Figure 4.** Classification of multispectral data obtained from blank samples (BLANK); unknown commercial samples (UNK); spiked sample with paracetamol (ADD01), piroxicam (ADD02) and both paracetamol and piroxicam (ADD12)
CONCLUSION
A selective, sensitive, rapid, and accurate analytical method based on UV-Vis spectroscopy and chemometric for detecting paracetamol and piroxicam in herbal medicine is successfully developed. Variable selection enhances the performance of the calibration model so that data pre-processing was not necessary. The limitation of this method is restricted only for paracetamol and piroxicam quantities. Therefore, enclosing other commonly chemical adulterants would be better for creating a vast database for future research.

ACKNOWLEDGEMENT
The authors would like to thank the Faculty of Medicine and Health Sciences Atma Jaya Catholic University of Indonesia for funding this research.

REFERENCES


