Implementation of chemometrics as a Solution to detecting and preventing falsification of herbal medicines in Southeast Asia: A review

Arif Nur Ikhsan1, Fella Syifa1, Mabrurotul Mustafidah1, Abdul Rohman2,3*

1Postgraduate Student, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.
2Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.
3Center of Excellence, Institute for Halal Industry and Systems (PUI-PT IHIS UGM), Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

ARTICLE INFO
Received on: 17/01/2021
Accepted on: 22/06/2021
Available Online: 05/09/2021

Key words: Chemometrics, authentication, identification, narrative review, Southeast Asia.

ABSTRACT
Herbal medicine has been known for centuries as one of the sophisticated traditional medicine systems and is widely used in Southeast Asia. Based on their beneficial and economical values, herbal medicines are sometimes found falsified. Chemometrics is a potential tool be developed to detect and prevent falsification. This review aimed to explore the application of chemometrics using certain chemical and biological responses for detecting the adulteration of herbal medicines in Southeast Asia. The study was carried out as a narrative review to evaluate published papers between 2000 and 2020. A total of 34 papers met the inclusion criteria and were reviewed and evaluated. Reviewed articles reveal that chemometrics successfully detected the falsification of herbal medicine with great accuracy and flexibility. Chemometrics was advantageous, practical, and cost-effective to detect adulteration and prevent falsification of herbal medicines. Nonetheless, the data published still diminish compared to the number of herbal medicines in Southeast Asia. Updated, verified, and integrated data among Southeast Asia researchers may be collected to sharpen the chemometric method. The implementation of chemometrics in Southeast Asia to detect and prevent herbal medicines was fundamental and increased the potency of herbal medicines.

INTRODUCTION
Herbal medicine has been known for centuries as one of the sophisticated traditional medicine systems in preventing or treating disease. It is one of the traditional medicines used for many generations and is acknowledged as an alternative medicine (Ahmad, 2002). People in Southeast Asian countries have a long history of using plants and herbs for their healing properties. The rich-in-flora biodiversity has influenced people to use plants as remedies besides food supplies (Granato et al., 2018). Some of them are turmeric (Curcuma longa) and ginger (Zingiber cassumunar) which are widely cultivated in Southeast Asian countries such as Indonesia (Rohaeti et al., 2015).

Since there are various factors that influence the chemical composition of the plants, it is important to make sure that the plant has an active component and it is there in a sufficient amount (Gad et al., 2013). The efficiencies of herbal medicine depend on several factors in order to ensure the reliability and repeatability of the pharmacological effect (Rohman et al., 2014). The problem is that the material used in herbal medicine by the communities and drug industry is sometimes found to be adulterated (Sultana et al., 2011). For this reason, authenticity is an important aspect to be considered as one of the quality controls. The other problems related to herbal medicine’s authenticity are incorrect identification, substitution, adulterations, dilution with lower grade materials, and labeling the products of different origins (Rohman et al., 2014).

There was much research applied and developed regarding the use of analytical methods to authenticate herbal medicine such as high-performance liquid chromatography.
The duplicated articles had to be discarded manually and 84 journal articles were collected. All papers were initially reviewed by title, abstract, and methods to be selected before the final assessment. The inclusion criteria were papers (1) containing a pure or mixture extract of herbal medicines which originate from Southeast Asia, (2) using chemometrics for analysis, (3) published between January 2000 and November 2020, (4) in the English language, and (5) containing any research designs, while the exclusion criteria included those papers where (1) all studies have been cited and (2) the review paper was not included. The assessment was based on inclusion and exclusion criteria. There were eight more qualified journals that were added to the pooled journals to accomplish the review. A total of 34 papers were selected for a more detailed review. The research strategy used in the review is illustrated in Figure 1.

RESULTS AND DISCUSSION

Results

Authentication of Herbal Medicines

Consumption of herbal medicines and products has been increasing nowadays for primary healthcare needs. The uncontrollable quality in herbal productions can have an effect on the efficacy and safety of the starting material, preparations, and finished products. Herbal materials contain complex mixtures of chemical constituents which possess variability due to many factors such as climate, harvest seasons, plant origins, drying, storage, and the extraction procedure. The adulteration might lead to variant efficacy and safety. As for herbal medicines,
identification, determination, and authentication techniques are commonly performed to fulfill the requirement of quality control (Azemin et al., 2018). The adulteration practice, either intentionally or unintentionally, usually includes partial or full substitution of original crude drugs with other substances which are cheaper and either free from or inferior in therapeutic properties.

The application of chemometrics can significantly improve the quality of the fingerprint obtained from complex chromatographic or spectroscopic profiles. Chemometrics is basically classified into two main categories: pattern recognition methods and multivariate calibration for quantitative purposes. Pattern recognition by means of multivariate statistical analysis can be divided into two categories: unsupervised and supervised. Unsupervised pattern recognition is utilized for data visualization by observing the relationship between samples and variables with no predetermined class. On the other hand, supervised pattern recognition methods have been used intensively for the classification of datasets. The most popular techniques for the classification of herbal products include linear discriminant analysis (LDA), k-nearest neighbor, soft independent modeling of class analogy (SIMCA), ANN, partial least square-discriminant analysis (PLS-DA), and orthogonal projections to latent structures-discriminant analysis (OPLS-DA) (Gad and Bouzabata, 2017). Table 1 explains several analytical methods that have been reported which applied the fingerprinting analytical approach using chromatography [Thin-layer chromatography (TLC), GC, and HPLC/ultra-performance liquid chromatography] and spectrometry (UV–Vis, Infrared, near magnetic resonance (NMR), and MS) or both in tandem like gas chromatography-mass spectrometer (GC-MS) and liquid chromatography-mass spectrometer that could be used for fingerprint analysis and authentication of some herbal medicines (Rohman et al., 2020b; Septyanti et al., 2016).

The combination of chromatography fingerprint and multivariate analysis [Partial least square regression (PLSR), PCA, and discriminant analysis (DA)] has been extensively used for species identification, discrimination, and authentication of medicinal plant such as authentication of fresh ginger (Zingiber officinale) from another country (Yudhavarosit et al., 2014), authentication of KG (Kaempferia galanga) from related plant species (Septyanti et al., 2016), and discrimination of red and white rice bran (Sabir et al., 2017). Other methods of chemometrics, namely, PLSR, PCR, and PCA, were combined with fourier transform infrared spectroscopy (FTIR) spectroscopy to be successfully applied for the authentication of Eurycoma longifolia from different places in Kalimantan (Triyasmono et al., 2020) and authentication of Nigella sativa oil adulterated with grapeseed oil (Nurulhidayah et al., 2011). PCA and DA were successfully used for the quantification of rice bran oil with extra virgin olive oil based on peak intensities at FTIR spectra (Rohman and Man, 2012). Rapid authentication using UV-Vis spectra was used in combination with chemometrics for discrimination in four species of Curcuma (Curcuma longa, Curcuma xanthorrhiza, Curcuma aeruginosa, and Curcuma manga). PCA and DA were used for the classification of the four species (M. Rafi et al., 2018).

FTIR and HPLC fingerprints of phytochemicals using chromatographic and spectroscopic techniques may provide valuable information about qualitative and quantitative analysis of medicinal herbs in which pattern recognition can be achieved using chemometrics including PCA and hierarchical cluster analysis (HCA). The application of combining both methods was used for the authentication of different extracts of Orthosiphon stamineus leaves combined with chemometrics (PCA and HCA) for quality control (Saidan et al., 2015). The combination of chemometrics (PCA, HCA, and DA) with GC-MS was also found to be able to discriminate and classify the untargeted volatile compounds in some varieties of Ficus deltoidea based on untargeted volatile compounds which could be used for their quality control (Azemin et al., 2018).

The authentication methodology was developed using hyperspectral imaging using chemometrics (PCA, PLS-DA, and ANN) applied for exploring the data and constructing and authenticating the retail samples for nutmeg powder authentication from other species (Kiani et al., 2019). Proton nuclear magnetic resonance (1H-NMR) spectroscopy gained its popularity because of its capability of detection, identification, and authentication between groups of samples with relatively simple sample preparation (Wijayanti et al., 2019). Proton NMR-based metabolite fingerprinting combined with chemometric methods of PCA and PLS-DA has been used for the authentication of C. xanthorrhiza adulterated with Z. cassumunar (Wijayanti et al., 2019). Another analytical method also can be considered such as TLC. TLC has been developed for a long time for the analysis of drugs and medicinal plants. TLC has many advantages for analysis such as simple preparation, use of only a small volume of organic solvent, and the fact that it can be used for qualitative and semiquantitative analysis.

By using the developed methods using 1H-NMR spectroscopy and TLC methods combined chemometrics of PCA, OPLS-DA was successfully used for classifying between pure and adulterated samples such as Curcuma species, especially C. longa adulterated with C. manga (Windarsh et al., 2018), and java turmeric (Rohman et al., 2020b) which can be observed in Figure 2. Other methods for analytical techniques such as polarography, X-ray fluorescence spectrometry, inductively coupled plasma-optical emission spectroscopy (ICP-OES), and inductively coupled plasma-mass spectrometry (ICP-MS) were combined with multivariate PCA and LDA for the classification and authentication of varieties of Piper betle (Atikul Islam et al., 2020). Species Classification

The chemometrics analysis, namely, PCA and HCA, combined with HPLC successfully differentiated three Panax species which is important to the locals in Southeast Asia (Table 1) (Xia et al., 2016). Although it is possible to identify and discriminate visually by the chromatograms, minor variations among the chromatograms might not be seen. The help from chemometrics could assist in evaluating these minor differences. The result of PCA and HCA was comparable and the three principal components accounted for >80.5%. The other studies conducted similar analysis and concluded that chemometrics PCA and HCA allowed the discrimination between the varieties based on metabolite fingerprint or volatile compounds (Azemin et al., 2018; Maulidiani et al., 2012; Srisonp et al., 2019; Wirasuta et al., 2017). Meanwhile, Rafi et al. (2018) identified and discriminated Curcuma species. All of the four species are closely related and used as herbal drinks and food supplements in Indonesia. UV–
Table 1. Implementation of chemometrics for herbal medicine authentication.

<table>
<thead>
<tr>
<th>No</th>
<th>Subject</th>
<th>Analytical method</th>
<th>Chemometric method</th>
<th>Aim of research</th>
<th>Main findings</th>
<th>Reference</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arabica coffee</td>
<td>UV-Vis spectroscopy</td>
<td>PLS and DA</td>
<td>Classifying arabica coffee from different regions</td>
<td>Understanding the use of PLS and DA to classify the different qualities of same subject</td>
<td>(Suhandy and Yulia, 2018)</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Ananas comosus L.</td>
<td>SPE-LC</td>
<td>CA and PCA</td>
<td>Identification of Ananas comosus L. bioactive compounds and discrimination from three varieties</td>
<td>Understanding the use of PCA and CA to discriminate each species</td>
<td>(Khan et al., 2020)</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Centella asiatica, Hydrocotyle bonariensis, and Hydrocotyle sibthorpioides</td>
<td>¹H-NMR</td>
<td>PCA and HCA</td>
<td>Discrimination and identification of the markers between three varieties of pega</td>
<td>Understanding the use of ¹H-NMR as instrument for authentication in similar samples</td>
<td>(Maulidiani et al., 2012)</td>
<td>2 &amp; 3</td>
</tr>
<tr>
<td>4</td>
<td>C. longa and C. xanthorrhiza</td>
<td>HPLC and FTIR</td>
<td>PLS</td>
<td>Determining curcumin in Curcuma species</td>
<td>The use of HPLC combined with PLS for determining each content from two similar species</td>
<td>(Rohman et al., 2015)</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>C. longa, C. xanthorrhiza, and Z. cassumunar</td>
<td>FTIR</td>
<td>PCA and CVA</td>
<td>Discriminating three closely related Curcuma species</td>
<td>The use of FTIR combined with PCA and CVA for determining each component from two related species</td>
<td>(Rohacti et al., 2015)</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>C. longa</td>
<td>FTIR</td>
<td>PCA, PLS-DA, and PLSR</td>
<td>Authentication of turmeric powder from Curcuma zedoaria and C. xanthorrhiza as adulterant</td>
<td>Providing information about authentication in powder form</td>
<td>(Angeline et al., 2019)</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>C. xanthorrhiza</td>
<td>¹H-NMR</td>
<td>PCA and PLS-DA</td>
<td>Authentication of C. xanthorrhiza adulterated with Z. cassumunar</td>
<td>Understanding the use of ¹H-NMR as instrument for authentication</td>
<td>(Wijayanti et al., 2019)</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>C. longa</td>
<td>¹H-NMR and TLC</td>
<td>PCA, PLS-DA, and OPLS-DA</td>
<td>Authentication of C. longa from adulterated powder</td>
<td>Giving information related to TLC as instrument for authentication</td>
<td>(Windarsih et al., 2018)</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>C. longa, C. xanthorrhiza, C. aeruginosa, and C. manga.</td>
<td>UV-Vis spectroscopy</td>
<td>PCA and DA</td>
<td>Discriminating four Curcuma species</td>
<td>The use of UV-Vis spectroscopy for authentication</td>
<td>(M. Rafi et al., 2018)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>10</td>
<td>E. longifolia</td>
<td>FTIR</td>
<td>PCA and PLSR</td>
<td>Authenticating E. longifolia from Z. officinale as adulterant</td>
<td>Providing information about falsifying high-economical-value extract with low-economical-value extract in powder form</td>
<td>(Triyasmono et al., 2020)</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Flavonoid in the medicinal plant extract</td>
<td>FTIR and UV-Vis spectroscopy</td>
<td>PLS and LDA</td>
<td>Measuring significance of flavonoid content in many medicinal plant extracts</td>
<td>Understanding the use of FTIR for determining flavonoid contents</td>
<td>(Wulandari et al., 2016)</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>F. deltoidea Jack</td>
<td>GC-MS</td>
<td>PCA, HCA, and DA</td>
<td>Authenticating F. deltoidea Jack among eight similar varieties</td>
<td>Providing information about volatile compounds which are used for authenticating each variety</td>
<td>(Azem et al., 2018)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>No</td>
<td>Subject</td>
<td>Analytical method</td>
<td>Chemometric method</td>
<td>Aim of research</td>
<td>Main findings</td>
<td>Reference</td>
<td>Classification</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------------</td>
</tr>
<tr>
<td>13</td>
<td><em>Garcinia mangostana</em> L.</td>
<td>HPLC and FTIR</td>
<td>PLS and PCA</td>
<td>Quantitative analysis of mangosteen using FTIR spectra</td>
<td>Application of FTIR spectra and chemometrics for predicting mangosteen contents and classifying by its origin</td>
<td>(Rohman et al., 2020a)</td>
<td>1 &amp; 3</td>
</tr>
<tr>
<td>14</td>
<td>Ground roasted robusta coffee</td>
<td>UV-Vis spectroscopy</td>
<td>PCA and SIMCA</td>
<td>Classifying fresh and expired ground roasted robusta coffee</td>
<td>The use of chemometrics for analyzing quality of products</td>
<td>(Yulia and Suhandy, 2018)</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Java turmeric</td>
<td>TLC and ³H-NMR</td>
<td>PCA and OPLS-DA</td>
<td>Authenticating <em>C. xanthorrhiza</em> from <em>C. aeruginosa</em> as adulterant</td>
<td>The use of OPLS-DA for authentication</td>
<td>(Rohman et al., 2020b)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>16</td>
<td><em>K. galangal</em></td>
<td>HPLC</td>
<td>PCA and DA</td>
<td>Identification and discriminating of <em>K. galangal</em> from another related species</td>
<td>The use of HPLC fingerprints for authenticating oilplant from its related species</td>
<td>(Septyanti et al., 2016)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>17</td>
<td><em>M. fragrans</em> Houttt.</td>
<td>Hyperspectral imaging</td>
<td>PCA, PLSDA, and ANN</td>
<td>Authenticating <em>M. fragrans</em> Houttt. from adulterated powder</td>
<td>Understanding the use of hyperspectral imaging for authentication</td>
<td>(Kiani et al., 2019)</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td><em>M. fragrans</em></td>
<td>PTR-MS and FT-ESI-MS</td>
<td>PCA</td>
<td>Discriminating low-quality nutmeg and high-quality nutmeg</td>
<td>Chemometrics can be used for classifying different qualities of extract</td>
<td>(Van Ruth et al., 2019)</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>Mangosteen pericarp</td>
<td>HPLC and FTIR</td>
<td>PLS and PCR</td>
<td>Determining the contents of alpha-mangostin, gamma-mangostin, and gartanin from <em>G. mangostana</em> L.</td>
<td>Quantifying three different contents in extract simultaneously</td>
<td>(Tejamukti et al., 2020)</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td><em>Mangifera indica</em> L., with three species: Harumanis, Tong Dam, and Susu mangoes</td>
<td>GC</td>
<td>CA and DA</td>
<td>Identification and discrimination of volatile compounds for different <em>Mangifera</em> species</td>
<td>Volatile compound can be used for profiling and discriminating each related species</td>
<td>(Zakaria et al., 2018)</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td><em>Mitragyna speciosa</em></td>
<td>ICP-MS</td>
<td>DFA</td>
<td>Discriminating kratom from its origin, suborigin, and strain</td>
<td>Providing information about discriminating based on the origin, suborigin, and strain</td>
<td>(Braley and Hondrogiannis, 2020)</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td><em>N. sativa</em> L. seed oil</td>
<td>FTIR</td>
<td>PCA, PLSD, and PCR</td>
<td>Authentication <em>N. sativa</em> L. seed oil from grapeseed oil as adulterant</td>
<td>Chemometrics can be used for authentication in oil form</td>
<td>(Nurrulhidayah et al., 2011)</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td><em>O. stamineus</em></td>
<td>FTIR</td>
<td>PCA and SIMCA</td>
<td>For authentication of <em>Orthosiphon</em> from variety of herbal samples</td>
<td>The use of SIMCA for authentication from different variety of samples</td>
<td>(Sim et al., 2004)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>24</td>
<td><em>O. stamineus</em></td>
<td>TLC</td>
<td>PCA</td>
<td>Discriminating <em>O. stamineus</em> from different regions</td>
<td>The use of TLC as an instrument for authentication of species from different regions</td>
<td>(Kartini et al., 2020)</td>
<td>3</td>
</tr>
<tr>
<td>25</td>
<td><em>O. stamineus</em></td>
<td>UV-Vis spectroscopy, FTIR, and HPLC</td>
<td>PCA and HCA</td>
<td>Classifying and discriminating <em>O. stamineus</em> from nine related species</td>
<td>Comparing authentication instrument from related research</td>
<td>(Saidan et al., 2015)</td>
<td>1 &amp; 4</td>
</tr>
<tr>
<td>No</td>
<td>Subject</td>
<td>Analytical method</td>
<td>Chemometric method</td>
<td>Aim of research</td>
<td>Main findings</td>
<td>Reference</td>
<td>Classification</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------------</td>
</tr>
<tr>
<td>26</td>
<td>Panax notoginseng, Panax vietnamensis, and Panax stipuleanatus</td>
<td>HPLC</td>
<td>PCA and HCA</td>
<td>Distinction and identifying of Panax species</td>
<td>The use of characteristic chromatogram peaks for tdistinction of each species</td>
<td>(Xia et al., 2016)</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>Phyllanthus niruri</td>
<td>FTIR</td>
<td>PCA, LDA, and SIMCA</td>
<td>Discriminating six different species of Phyllanthus sp.</td>
<td>Application of LDA for discriminating sample from different species</td>
<td>(Dharmaraj et al., 2006)</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>P. nigrum L.</td>
<td>GC-MS</td>
<td>PCA</td>
<td>Discriminating P. nigrum L. from different geographical origins</td>
<td>Volatile compound may be used for discriminating samples from different origins</td>
<td>(Mercer et al., 2019)</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>P. betle L.</td>
<td>TLC</td>
<td>PCA and HCA</td>
<td>Authenticating P. betle L. and predicting its antifungal activity</td>
<td>Predicting antifungal activity based on TLC profile</td>
<td>(Wirasuta et al., 2017)</td>
<td>2 &amp; 4</td>
</tr>
<tr>
<td>30</td>
<td>P. betle L.</td>
<td>ICP-oes and ICP-MS</td>
<td>PCA and LDA</td>
<td>Authenticating and identifying P. betle L. varieties</td>
<td>Applying ICP-oes and ICP-MS as instruments for authentication</td>
<td>(Atikul Islam et al., 2020)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>31</td>
<td>Red rice bran and white rice bran</td>
<td>HPLC</td>
<td>PCA and DA</td>
<td>Identifying and discriminating between red rice bran and white rice bran</td>
<td>Authenticating samples from their specific compounds</td>
<td>(Sabir et al., 2017)</td>
<td>1 &amp; 2</td>
</tr>
<tr>
<td>32</td>
<td>Rice bran oil</td>
<td>FTIR</td>
<td>PCA, DA, PLS, and PCR</td>
<td>Authenticating rice bran oil from extra virgin oil as adulterant</td>
<td>Application of FTIR for analyzing adulterated oils</td>
<td>(Rohman and Man, 2012)</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>Sandalwood</td>
<td>GC-MS</td>
<td>SA, CA, and PCA</td>
<td>Identifying sandalwood crude drugs from traditional drugstores</td>
<td>Analyzing specific volatile compound from crude drugs</td>
<td>(Srisopon et al., 2019)</td>
<td>2</td>
</tr>
<tr>
<td>34</td>
<td>Z. officinale</td>
<td>HPLC</td>
<td>HCA, PCA, and LDA</td>
<td>Identification of Z. officinale authentication from adulterant</td>
<td>The use of HCA for authenticating samples from adulterated samples</td>
<td>(Yuthhavorasit et al., 2014)</td>
<td>1</td>
</tr>
</tbody>
</table>

Vis spectroscopy in combination with PCA and DA was used for the classification of the four species. The results showed that species classification using DA gave better separation with 95.5% of the samples classified correctly into their groups by leave-one-out cross-validation. Another study conducted by Rohman et al. (2020b) ingsuccessfully discriminated pure and adulterated C. xanthorrhiza with C. aeruginosa using TLC and proton 1H-NMR coupled with PCA and OPLS-DA as multivariate analysis. OPLS-DA is supervised pattern recognition which uses orthogonal X and Y variables for classification. The result shows that both PCA and OPLS-DA successfully differentiated the two Curcuma species. The study’s findings were also similar to other studies (Atikul Islam et al., 2020; Azemin et al., 2018; Sabir et al., 2017; Septyanti et al., 2016).

Canonical variate analysis (CVA) was also one of the multivariate analyses used for species classifications. Rohaeti et al. (2015) used FTIR spectroscopy coupled with PCA and CVA to classify three Curcuma species (Curcuma longa, Curcuma xanthorrhiza, and Z. cassumunmar). CVA is one of the supervised pattern recognition techniques with the goal of finding a linear combination of variables that exhibit the maximum among group variations to within-groups variations and work effectively when the samples are more than the variables. PCA and CVA could discriminate the three closely related species more accurately than by just visual analysis of the FTIR spectra, in which CVA gave clearer classification based on species. Dharmaraj et al. (2011) showed the application of PCA-LDA, genetic algorithm-based linear discriminant analysis (GA-LDA), SIMCA, and neural networks together with FTIR spectroscopy to distinguish six different species of Phyllanthus. The best result was obtained with GA-LDA in which the similar FTIR spectra from the six species could easily be distinguished compared to any other method.

**Origin Determination**

Different chemometric studies have been carried out for classifying samples from their different environments which imply different characteristics (Table 1). Some samples produce some unique metabolites or profiles which can be used as bases for classifying their origin. The environmental differences like light intensity may affect the production of some plant metabolite contents. The difference in light intensity may be observed using research intended by Maulidiani et al. (2012). Using 1H-NMR PCA and HCA, the difference in metabolite contents from different light intensities was observed. The effects of the growth-lighting condition on metabolite contents was investigated. This approach successfully discriminates samples which grow in high-
and low-light-intensity environments based on their metabolites (triterpenoids, flavonoids, and phenolic compounds). Based on the study, light exposure may accumulate triterpenoids, flavonoids, and phenolic compounds which possess antioxidant activity (Maulidiani et al., 2012). However, samples exposed to lower light intensity produced lower triterpenoids, flavonoids, and phenolic compounds. These compounds will help plants to reduce UV radiation damage from light. The extracts grown with full-day light exposure exhibited a stronger antioxidant activity and contained more metabolites as compared to plants grown in 50% shade.

TLC combined with PCA was also successful in classifying *O. stamineus* from different origins. This classification based on the Rf value, height, and area of each peak obtained from a videodensitogram had significant differences. The results of the PCA score plot of PC1 and PC2 clearly distinguished three clusters of samples with the RF values of 0.0–0.1; 0.1–0.2; 0.2–0.3, and 0.9–1.0 which are the most important compounds for clustering of samples (Kartini et al., 2020). Volatile organic compounds can also be used for classifying the samples’ origins. Using high-resolution-gas chromatography-mass spectrometry, over 200 different compounds have been found. Some signature peaks can be used for the classification of each sample’s origin. Using PCA and fold change analysis, the researcher could differentiate *Piper nigrum* L. from its origin (Malaysian and Indian black pepper) by means of 11 unique non-polar compounds present in the pepper samples among others: 4(10)-thujene; bicyclo(3.1.0) hexan-2-ol; alpha-copaene; artemisinin; (2e,4e,14e)-n-isobutylicosa-2,4,14-trienamide; elemene isomer (e-elemene) andrographolide; (e)-1-(piperidin-1-yl) dodec-2-en-1-one; (2E,4E,6E)-7-(benzo[d][1,3]dioxol- 5-yl)-1-piperidin-1-yl) pent-2-en-1-one; and gibberellic acid.

A research was conducted for classifying three different origins, coffee Arabica Gayo, Arabica Kintamani, and Arabica Wamena, based on their unique taste. The research was carried out using UV-Vis spectroscopy at 190–1,100 nm wavelength combined with PLSR-DA and PCA to classify 296 samples from Gayo, Kintamani, and Wamena. By using this method, all of the samples were 100% correctly classified from their origin (Suhandy and Yulia, 2018). Another similar method conducted by Rohman et al. (2020a) classified mangosteen pericarps from different locations using FTIR spectra. Using PCA at 1,000–800 cm⁻¹, the researcher could classify mangosteen pericarp from different regions. Using ICP-MS combined with chemometrics provides another perspective in plant origin classification. The study was carried out by using the 13 elements found in kratom samples. Discriminant function analysis (DFA) using the leave-one-out model successfully classified kratom from its origins (100%), suborigins (100%), and strain (86%) (Braley and Hondrogiannis, 2020).

**Activity and Quality Determination**

Flavonoid contents in several herbal medicine extracts were reported to have antioxidant, anti-inflammatory, and anticancer activity (Miller, 1996). The biological activities are closely related to the medicine plant extracts’ quality. Using near-infrared spectroscopy combined with LDA, SIMCA, and support vector machines successfully determined the flavonoid content in each medicinal plant extract. Using this method, the quality of different medicinal plant extracts can be accurately identified with SIMCA as the best chemometric method with 91.2% accuracy (Wulandari et al., 2016).

The study conducted by Tejamukti et al. (2020) reveals that *xanthone* derivates can be determined using FTIR combined multivariate calibration of PLSR and PCR. Like flavonoids, a xanthone derivate also has various biological activities among others (Rivero and Garibay, 2019), such as antioxidant, analgesic, anti-inflammatory, antifungal, and anthelmintic activities (El-Kenawy et al., 2018). Using FTIR combined PLSR, the xanthone derivate contents can be predicted precisely and accurately with the coefficient of determination ($R^2$) 0.9573 followed by an RMSEC value of 0.0487% and RMSEP value of 0.12% (Tejamukti et al., 2020).

Another study was conducted to evaluate the antifungal activity of *Piper betle* L. This research using numerical chromatographic parameters of the HPLC peak markers combined HCA and PCA. Using PCA, antifungal activity can be determined from the first component’s PCA, which linearly correlated to the biomarker used. This method proved it can be used to predict the antifungal activity from *P. betle* L. (Wirasuta et al., 2017).

Moreover, determining the high quality and low quality of nutmeg (*Myristica fragrans*) has been carried out based on...
volatile and nonvolatile compounds using FI-ESI-MS and PTR-MS combined with PCA (Van Ruth et al., 2019). Intensities of volatiles (trymyrinstin and essential oil) and non-volatiles are highly correlated but can diminish gradually and even be reserveds with rising molecular masses of the non-volatiles. By using this approach, nearly 100% correct prediction was obtained and high-quality and low-quality nutmeg were successfully discriminated.

Discriminating quality has been carried out for ground roasted robusta coffee. UV-Vis spectroscopy combined with PCA and SIMCA is used to evaluate fresh and expired samples based on caffeine, trigonelline, and chlorogenic acid contents. The spectral data were pre-treated using standard normal variate (SNV) with thea spectral range spectral of 230–400 nm. The results of this research show that this approach can be used to discriminate each sample successfully with a 100% correct classification rate (Yulia and Suhandy, 2018).

Comprehensive research carried out by Saidan et al. (2015) correlated HPLC and FTIR combined with chemometrics with different biological activities. This approach shows that chemometrics can be used as a tool to classify and discriminate distinct features of extracts that can be correlated with their biological activities (Saidan et al., 2015).

DISCUSSION

The Benefit of Using Chemometrics for Herbal Medicines for Authentication and further Developments

A review was carried out to determine the implementation of chemometrics for the authentication of herbal medicine which grows and is cultivated or traded in Southeast Asian countries. Based on this review, chemometrics successfully simplifies and discriminates herbal medicines using various data. Processed data can be used for the authentication of various herbal medicines in Southeast Asia. From 34 selected journals, the aim, instruments, and chemometric method were evaluated. Based on Table 1, chemometrics can be used for various analytical purposes. However, in this review, we discriminate and group each function among others: authentication of herbal medicines, species classification, origin determination, and activity related to quality determination.

This review reveals that chemometrics was a high potential method to be applied to authenticate various herbal medicines in Southeast Asia because of the low cost and simplicity. Using chemometrics, analyzing herbal medicine will become easier and faster. Therefore, herbal medicine industries, brokers, and distributors can authenticate their herbal medicines and prevent herbal medicine adulteration. From the chemometrics based on the data obtained from different samples, the data play an important role in this technique. Numerous data needed should be collected, verified, and integrated among researchers in Southeast Asia. Comprehensive data among various herbal medicines, origins, varieties, and potencies will improve the quality of the chemometric technique. Integrated data that can be accessed by researchers can prevent the falsification of herbal medicines in the future.

Limitations

In early 1970, Weiner et al. (1970), Jurs et al. (1969), and Pestemer (1974) were publishing manuscripts that now we recognize as chemometrics. In the last three decades, chemometrics was widely used in various fields, such as chemical analysis (Bereton, 2014; Honold et al., 2016; Shafii et al., 2019). In the present day, chemometrics has been widely developed and used as a tool to help researchers simplify and interpret complex data. Implementation of chemometrics for authentication of herbal medicines in Southeast Asia performed well for the authentication of herbal medicines (Nurrulhidayah et al., 2011; Triyasmono et al., 2020), species classification (Khan et al., 2020; Rohaeti et al., 2015), origin determination (Braley and Hondrogiannis, 2020; Mercer et al., 2019), and activity related to quality determination (Kiani et al., 2019; Saidan et al., 2015; Wirasuta et al., 2017). The limitation of chemometrics was caused mainly by the limited data obtained. There are still diminishing data and research related to the authentication of herbal medicine compared to the number of herbal medicines in Southeast Asia. Therefore, the data should be updated and verified among Southeast Asian researchers to sharpen the authentication method.

CONCLUSION

The application of chemometrics for authentication of herbal medicines in Southeast Asia was advantageous and could be developed as a new approach. However, this review exhibits various researches related to the application of chemometrics which support the development of chemometrics for the analysis of herbal medicines in the future. Collaboration and integration data among countries in Southeast Asia have become an important issue to be discussed to maximize this new approach’s potency.

ABBREVIATIONS

1H-NMR : Proton Nuclear Magnetic Resonant
ANN : Artificial Neural Network
CVA : Canonical Variate Analysis
DA : Discriminant Analysis
DFA : Discriminant Function Analysis
FTIR : Fourier Transform Infrared Spectroscopy
GA-LDA : Genetic Algorithm based Linear Discriminant Analysis
GC : Gas Chromatography
GC-MS : Gas Chromatography-Mass Spectrometer
HCA : Hierarchical Cluster Analysis
HPLC : High Performance Liquid Chromatography
HR-GC-MS : High Resolution Gas Chromatography Mass Spectrometry
ICP-OES : Inductively Coupled Plasma-Optical Emission Spectroscopy
LC-MS : Liquid Chromatography-Mass Spectrometer
LDA : Linear Discriminant Analysis
OPLS-DA : Orthogonal Projections To Latent Structures-Discriminant Analysis
PCA : Principal Component Analysis
PLS-DA : Partial Least Square - Discriminant Analysis
PLSR : Partial Least Square Regression
SIMCA : Soft Independent Modeling Of Class Analogy
TLC : Thin Layer Chromatography

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for
important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work. All the authors are eligible to be an author as per the international committee of medical journal editors (ICMJE) requirements/guidelines.

**FUNDING**

There is no funding to report.

**CONFLICTS OF INTEREST**

The authors report no financial or any other conflicts of interest in this work.

**ETHICAL APPROVALS**

This study does not involve experiments on animals or human subjects.

**PUBLISHER’S NOTE**

This journal remains neutral with regard to jurisdictional claims in published institutional affiliation.

**REFERENCES**


Honold PJ, Nouard ML, Jacobsen C. Fish oil extracted from fish-fillet by-products is weakly linked to the extraction temperatures but strongly linked to the omega-3 content of the raw material. Eur J Lipid Sci Tech, 2016; 118(6):874–84; doi:10.1002/ejlt.201500343


Mercer ZJA, Chua HS, Mahon P, Hwang SS, Ng SM. Authentication of geographical growth origin of black pepper (P. nigrum L.) based on volatile organic compounds profile: a case study for Malaysia and India black peppers. Conference: 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), 2019,p.3


Authentication of herbal medicine neem (Azadirachta indica) using thin layer chromatography and 'H-NMR based-metabolite fingerprinting coupled with multivariate analysis. Molecules, 2020b; 25(17):3928; doi:10.3390/molecules25173928


How to cite this article: