Using Fourier Transform Infrared (FTIR) Spectroscopy Coupled with Multivariate Analysis to Diagnose Polycystic Ovary Syndrome (PCOS)

Afrah Mohammed Hassan Salman1, M. Mohammed A. Al-Zubaidi2, Maysaa Ali Abdul khaleq2

1Department of Clinical Laboratory Sciences, College of Pharmacy, Mustansiriya University, Baghdad, Iraq
2Al-Rashed University College/Department of Dentistry, Baghdad, Iraq

Corresponding author: Afrah Mohammed Hassan Salman, Email: afrahbassil@yahoo.com

ABSTRACT

The biochemical diagnostic test is important in the identification of patients with polycystic ovary syndrome (PCOS). Fourier Transform Infrared (FTIR) measurement which analyzes the overall metabolic profiling of serum may provide more biochemical information. Here we applied the FTIR spectroscopy with multivariate analysis for a PCOS diagnosis. Sixty patients and thirty one healthy subjects were recruited in the present study. FTIR measurement was carried out on the subjects’ serum. The most significant variables obtained from FTIR data were nominated by variable importance in the projection (VIP) value following an OPLS-DA model creating from three different spectral regions: first region/ 900-1200 cm\(^{-1}\), second region/ 1500-1700 cm\(^{-1}\), and third region/ 2800-3100 cm\(^{-1}\). An OPLS-DA of the FTIR data showed variances in the components of patients’ serum. All OPLS-DA models showed validation (P-value < 0.05) based on CV-ANOVA, 900-1200 cm\(^{-1}\) region (R2Y(cum)= 0.926; Q2(cum)= 0.57), 1500-1700 cm\(^{-1}\) (R2Y(cum)= 0.828; Q2(cum)= 0.594), and 2800-3100 cm\(^{-1}\) (R2Y(cum)= 0.927; Q2(cum)= 0.717). These results showed that spectral biomarkers separated patients’ serum. In summary, our work proves that vibrational spectroscopy (FTIR) combined with multivariate analysis (OPLS-DA) showing the possibility to become a chemical reagent-free approach for biochemical assessment of bio-fluid.

Keywords: Blood serum, FTIR spectroscopy, Multivariate analysis, polycystic ovary syndrome

INTRODUCTION

Polycystic ovary syndrome (PCOS) is one of the most common endocrinological disorders that affects not only the reproductive function, but also the overall body metabolic function of approximately 5% to 20% of the reproductive-age women worldwide [1-2]. The PCOS is clinically characterized by hyperandrogenism, menstrual disorder, chronic amenorrhea, insulin resistance, chronic anovulation, and/or poly cystic ovarian morphology [2]. The complex disorder of PCOS was reported might be caused by the interaction of genetic and environmental factors that lead to diverse phenotypes [3-4]. Vibrational spectroscopic techniques including infrared absorption have expanded across the last two decades by introducing their techniques for different applications and providing both specific chemical knowledge and molecular structure for analyzed samples with no need to use an extrinsic labels [5-6]. Vibrational spectroscopy has some privilege because it is an applicable technique for reagent-free tests that makes it available for a non-destructive diagnosis in comparison with conventional diagnostic and screening problems linked to bio-fluid. This technique is relatively simple, reproducible, require a minimum samples preparation and without large sample amount requirements [7-8]. The basic idea of FTIR technique is when any biochemical compounds expose to infrared radiation, and when the energy of IR match the energy of a specific molecular vibration, the absorption happens. FTIR spectroscopy can offer a significant contribution to identify unknown materials, define the sample’s quality, and assess the quantity of chemical components in blend samples [6]. Several spectroscopic studies have been conducted on FTIR spectroscopy in trying to employ it for medical diagnosis, using tissue samples [9-10] and recently using bio-fluids e.g. serum [11-13], plasma [14-15] and tears [16]. Despite its potentiality, FTIR spectroscopy is not employed in clinical diagnosis, because most of the published studies were experiment of concept research studies [17]. The progress of using vibrational spectroscopic in medical diagnosis involves two-stage plan: the first stage is analysis/classification of biomarkers and the next stage is identification of these biomarkers [18]. The obtained metabolic fingerprint from FTIR spectra allows for identification a broad biochemical alterations caused by disease. Spectral data analysis and biomarkers identification are need to decrease the total number of variables in the spectrum from hundreds to around a dozen, in an effort to reduce the influence of no longer useful variables and noise [19]. This is possible to perform by the combined of multivariate analysis to FTIR spectra [20]. The aim of our work is to assess the capability of FTIR spectroscopy, yielding a comprehensive metabolic profile valuable for diagnosis of patients with polycystic ovary syndrome (PCOS). For that, the first step will focus on multivariate analysis approach for differentiating PCOS patients, whereas the second step will tentatively assign the chosen FTIR spectral markers.

MATERIALS AND METHODS

Subjects and preparing serum sample

The blood was collected in a plain tube. After blood has been taken from 60 patients and 31 healthy subjects, the collected blood was centrifuged at 3000 rpm for 10 min to separate and collect serum before storage at - 40 °C. At the stage of analysis, the serum was melted at room temperature. All subjects were completely informed and provided the consent forms before starting this study; regarding each PCOS patient, a basic medical history was taken. Furthermore, clinical examination and diagnosis were achieved by Gynecologist using abdominal ultrasound.
examination and Rotterdam 2003 criteria for PCOS in the High Institute for Infertility Diagnosis and Assisted Reproductive Technologies/Baghdad-Iraq during December 2018 to July 2019. Ethics Board in the College of Pharmacy/Mustansiriya University/Iraq was approved this study.

Hormonal measurement
The serum level of Luteinizing Hormone (LH) and Follicle Stimulating Hormone (FSH) were measured using commercial Enzyme-Linked Immunosorbent Assay (ELISA) kits (LH ELISA kit; Cat. No. ENZ-KIT107-0001 and FSH ELISA kit; Cat. No ENZ-KIT108-0001) according to the manufacturer’s instructions. The absorbance was recorded at 405 nm by BioTek Synergy H1, for estimation of LH and FSH. Microsoft Excel 2010 software was used for data analysis. Data were represented in the result table as mean ± standard deviation. Student’s t-test was used to compare patient group with healthy control group. Probability (P-value) <0.05 was considered the difference statistically significant.

FTIR spectra acquisition
The FTIR spectra were obtained by using FTIR-8400S, SHIMADZU. Absorption mode was used to analyze the samples within the region 400-4000 cm⁻¹, with 45 scans and at a resolution of 4 cm⁻¹. The laboratory temperature and humidity levels were kept during spectral acquisition at 23 °C and ± 39 %, respectively. After melting the frozen serum samples, 10 µl of serum was transferred to a 3 cm diameter of potassium bromide cell and spread on it. Potassium bromide cell was cleaned by using Chloroform (stab./Amylene) HPLC (Biosolve; cat no. 03080602) between each serum sample to prevent the interfering between samples. The samples analyses were achieved in a triplicate by each subject.

FTIR spectral pre-processing
At first, the spectral dimension was lowered 400-4000 cm⁻¹ to three specific bio-organic component regions 900-1200 cm⁻¹, 1500-1700 cm⁻¹, and 2800-3100 cm⁻¹, which represent the features of carbohydrates and nucleic acids, proteins, and lipids vibrations, respectively. Then, second derivative was calculated, and 13-point Savitzky-Golay algorithm was used for smoothing. Finally, vector normalization throughout the entire spectral region was applied for spectra normalization before conversion the spectra into ASCII format and then the data gathered in a single worksheet of Microsoft Excel (2010). The original FTIR spectra were obtained and pre-processed by using Shimadzu IR solution 1.60 software.

Multivariate and univariate analysis of FTIR spectra
After collection the normalized second derivative FTIR spectra of PCOS patient and healthy control groups, the average spectral data of each group was imported into Spectrograph software v1.2.10 to draw the average spectra, whereas the whole spectral data of each region was introduced to SIMCA 14.1 software (MKS Umetrics AB, Umeå, Sweden) for achieving principal component analysis (PCA) to detect and remove the outlier spectra that might not fit the model and after that achieving orthogonal partial least squares discriminant analysis (OPLS-DA) with unit variance (UV) scaling method to separate between the studied groups. The models were described by the criteria of both R2 (reflect the goodness of fit) and Q2 (reflect the goodness of prediction). When the variable importance in the projection (VIP) values exceeded 1.5, the variables were chosen as the most important variable for differences in the model. Furthermore, CV-ANOVA was used to assess the significance of each OPLS-DA model. Univariate analysis (Student’s t-test) was performed between patients with PCOS and healthy control groups by using Microsoft Excel 2010 software. P-value <0.05 is considered significant.

RESULTS
Patient and healthy subjects
Overall ninety-one participants were involved, 60 patients with PCOS and 31 healthy control. The variances in clinical and biochemical characteristics between the studied groups are reviewed in Table 1. The mean patients’ age was 25.18 ± 4.86 years and the mean age of the healthy subjects was 25.67 ± 5.37 years. There was statistically significant elevation of the mean LH level (P<0.01) and no statistically significant elevation of the mean FSH level in a PCOS patient group when compared to control group.

Table 1: Basic Clinical and biochemical characteristics of patients with polycystic ovary syndrome and control subjects

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>PCOS (n=60) Mean ± SD</th>
<th>Control (n=31) Mean ± SD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>25.18 ± 4.86</td>
<td>25.67 ± 5.37</td>
<td>65 x 10⁻²</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>24.26 ± 0.39</td>
<td>24.80 ± 0.51</td>
<td>42 x 10⁻³</td>
</tr>
<tr>
<td>LH (mIU/ml)</td>
<td>11.67 ± 2.31</td>
<td>5.33 ± 1.27</td>
<td>1.58 x 10⁻⁴</td>
</tr>
<tr>
<td>FSH (mIU/ml)</td>
<td>7.36 ± 1.56</td>
<td>6.50 ± 1.48</td>
<td>1.37 x 10⁻²</td>
</tr>
</tbody>
</table>

PCOS: Polycystic Ovary Syndrome; BMI: Body Mass Index; LH: Luteinizing Hormone; FSH: Follicle Stimulating Hormone

FTIR spectral overview
To better distinguish singular spectral component, second derivative together with 13 smoothing points Savitzky-Golay algorithm for the serum spectra of polycystic ovary syndrome (PCOS) and healthy control were normalized, averaged and displayed in Figure 1. At first view, the primary data exhibited that the mean serum spectra of polycystic ovary syndrome (PCOS) (blue lines) were distinctive when they compared with the mean serum spectra healthy control (red lines), the differences are
noticed in the three spectral regions (900-1200, 1500-1700 and 2800-3100 cm⁻¹), indicative of the metabolic and genomic, lipidomic and proteomic variation between polycystic ovary syndrome (PCOS) and healthy control subjects.

Figure 1: Mean of pre-processed FTIR spectra within the recorded regions obtained from three replicates of 60 PCOS serum samples and 31 healthy control serum samples

A) region 900-1200 cm⁻¹
B) region 1500-1700 cm⁻¹
C) region 2800-3100 cm⁻¹. X-axes correspond to wave number (cm⁻¹) and Y-axes correspond to absorbance. Blue color lines represent the patients with polycystic ovary syndrome (PCOS) and red color lines represent the healthy control spectra.

Classification of FTIR spectra and assignment of the spectral band

The biochemical features of bio-fluid spectra are extremely complicated and, therefore, it is better to analyze FTIR data by multivariate methods. In this study, for a comprehensible evaluation of spectral variance of each region between the studied groups we successfully used orthogonal partial least squares discriminant analysis (OPLS-DA) to explore the spectral profile differences between serum of patients with polycystic ovary syndrome and serum of healthy control, after employing principal component analysis (PCA). It was observed that carbohydrates and nucleic acids (spectral region 900-1200 cm⁻¹), proteins (spectral region 1500-1700 cm⁻¹), and lipids (spectral region 2800-3100 cm⁻¹) have valuable meaning for discrimination of PCOS sera. The spectral features are examined below in detail to clarify the mechanism of changes and obtain further understanding into the pathophysiology of polycystic ovary syndrome. In
addition to its classification capability, the OPLS-DA model offers an opportunity to separate the variables which are substantially related to discrimination. Student’s t-test (p<0.05) focused on the variables with VIP>1.5 taken from the created OPLS-DA model.

### Carbohydrates and nucleic acids region

Principal component analysis (PCA) for carbohydrate and nucleic acid feature (900-1200 cm⁻¹ region) is first performed to build up a picture of serum sample variances and assess the outlier serum samples, from which three outlier serum samples from PCOS and one outlier serum sample from healthy control groups falling outside the Hotelling’s T² ellipse (95% confidence region) (Figure 2A). The score plot of PCA represented as the first two PCs (PC1 (13.2%) and PC2 (11.5%)) and the model expressed the goodness of fit, R²X (cum) = 0.82, and predictability, Q² (cum) = 0.413. Then, the PCA model is utilized for OPLS-DA method. In the OPLS-DA model, after excluding the outlier serum samples from PCOS and healthy control groups, all spectral serum samples lay inside the Hotelling T² ellipse (95% confidence region) of the score plot (Figure 2B). The score plot represents a good separation between healthy control and polycystic ovary syndrome (PCOS) groups, and the serum samples allocated in each group’s region. OPLS-DA model represents a good fit and predictability with R²Y=0.926 and Q²=0.57, respectively; the validation assessment of the model for predictability was significant (P<0.05) for cross validated ANOVA (CV-ANOVA). These findings intimated variations in the pathophysiology of polycystic ovary syndrome (PCOS).

![Figure 2: Multivariate data analysis of serum carbohydrates and nucleic acids](image)

A) PCA score plot of the pre-processed FTIR spectra, component 1 (X-axis) explaining 13.2% and component 2 (Y-axis) explaining 11.5% of the variation

B) OPLS-DA score plot, X-axis shows the first principal component and Y-axis shows the first orthogonal component. Ellipses are the 95% confidence interval. Circles represent polycystic ovary syndrome (PCOS) group (n=60) for PCA and (n=57) for OPLS-DA score plot, squares represent healthy control group (n=31) for PCA and (n=30) for OPLS-DA.

The differential variables were selected after applying the following two criteria: 1) variable importance in projection (VIP) value exceeds 1.5 in OPLS-DA model; 2) P-value of Student’s t-test lower than 0.05. Table 2 represents the major assignments of the selected wave numbers with their absorption values in the 900-1200 cm⁻¹ region, that dominated by the vibration of PO₃ of nucleic acid, deoxyribose, and C-O of sugar ring of nucleic acid at different wave numbers including 900, 1048, 1070, 1072, 1080, 1082, 1094, and 1096 cm⁻¹.

<table>
<thead>
<tr>
<th>Wave number (cm⁻¹)</th>
<th>Absorbance</th>
<th>P-value</th>
<th>VIP value</th>
<th>Literature Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1070</td>
<td>0.659224</td>
<td>0.521274</td>
<td>9.26x10⁻²</td>
<td>1.83258</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Symmetric phosphate (PO₃) stretching [21]</td>
</tr>
<tr>
<td>900</td>
<td>0.106877</td>
<td>0.236334</td>
<td>2.7x10⁻²</td>
<td>1.82489</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Deoxyribose ring vibration [22]</td>
</tr>
<tr>
<td>1072</td>
<td>0.621799</td>
<td>0.488437</td>
<td>1.52x10⁻⁵</td>
<td>1.73553</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(PO₃) of nucleic acid vibration [23]</td>
</tr>
<tr>
<td>1048</td>
<td>0.614482</td>
<td>0.489394</td>
<td>2.12x10⁻⁵</td>
<td>1.61017</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C-O stretching of sugar ring of nucleic acid [6]</td>
</tr>
<tr>
<td>1096</td>
<td>0.6317</td>
<td>0.477421</td>
<td>2.76x10⁻⁵</td>
<td>1.59522</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Symmetric phosphate (PO₃) stretching [24]</td>
</tr>
<tr>
<td>1080</td>
<td>0.392146</td>
<td>0.500194</td>
<td>9.22x10⁻⁵</td>
<td>1.58041</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Symmetric phosphate (PO₃) stretching [25]</td>
</tr>
<tr>
<td>1082</td>
<td>0.384169</td>
<td>0.497025</td>
<td>2.70x10⁻⁴</td>
<td>1.57709</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Symmetric phosphate (PO₃) stretching [26]</td>
</tr>
</tbody>
</table>
Proteins region
Principal component analysis (PCA) for protein feature (1500-1700 cm\(^{-1}\) region) is first performed to build up a picture of sample variances and assess the outlier serum samples, from which two outlier serum samples from PCOS and four outlier serum samples from healthy control groups falling outside the Hotelling’s T2 ellipse (95% confidence region) (Figure 3A). The score plot of PCA represented as the first two PCs (PC1 (53.7%) and PC2 (15.1%)) and the model expressed the goodness of fit, R\(^2\) (cum) = 0.977, and predictability, Q\(^2\) (cum) = 0.911. Then, the PCA model is utilized for OPLS-DA method. In the OPLS-DA model, after excluding the outlier serum samples from PCOS and healthy control groups, all spectral serum samples lay inside the Hotelling T2 ellipse (95% confidence region) of the score plot (Figure 3B). The score plot represents definite healthy control and polycystic ovary syndrome (PCOS) grouping with a small overlap. OPLS-DA model represents a good fit and predictability with R\(^2\)Y=0.828 and Q\(^2\)=0.594, respectively; the validation assessment of the model for predictability was significant (P<0.05) for cross validated ANOVA (CV-ANOVA). These findings intimated variations in the pathophysiology of polycystic ovary syndrome (PCOS).

Table 3: FTIR band assignments of the polycystic ovary syndrome (PCOS) and healthy control groups in the protein spectral region (1500-1700 cm\(^{-1}\))

<table>
<thead>
<tr>
<th>Wave (cm(^{-1}))</th>
<th>Absorbance</th>
<th>P-value</th>
<th>VIP value</th>
<th>Literature Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCOS</td>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1648</td>
<td>0.201187</td>
<td>0.117069</td>
<td>3.67x10^{-5}</td>
<td>C=O different conformation of amide I [28] Amide I [29]</td>
</tr>
<tr>
<td>1646</td>
<td>0.16449</td>
<td>0.091304</td>
<td>2.26x10^{-6}</td>
<td>C=O stretching of amide I [30]</td>
</tr>
<tr>
<td>1650</td>
<td>0.218399</td>
<td>0.116621</td>
<td>2.43x10^{-7}</td>
<td>C=O stretching of amide I [25]</td>
</tr>
<tr>
<td>1588</td>
<td>0.839009</td>
<td>0.783913</td>
<td>1.46x10^{-6}</td>
<td>N-H bending and C-N stretching of Amide II [6]</td>
</tr>
<tr>
<td>1586</td>
<td>0.857661</td>
<td>0.803493</td>
<td>1.79x10^{-8}</td>
<td>N-H bend and C-N stretch of amide II [31]</td>
</tr>
<tr>
<td>1584</td>
<td>0.877991</td>
<td>0.830988</td>
<td>2.19x10^{-7}</td>
<td>(\alpha)-helical structures of amide II [32]</td>
</tr>
</tbody>
</table>
Lipids region
Principal component analysis (PCA) for lipid feature (2800-3100 cm⁻¹ region) is first performed to build up a picture of serum sample variances and assess the outlier serum samples, from which two outlier serum samples from PCOS and three serum samples from healthy control groups falling outside the Hotelling’s T2 ellipse (95% confidence region) (Figure 4A). The score plot of PCA represented as the first two PCs (PC1 (18.2%) and PC2 (14.8)) and the model expressed the goodness of fit, R²X (cum) = 0.834, and predictability, Q2(cum) = 0.512. Then, the PCA model is utilized for OPLS-DA method. In the OPLS-DA model, after excluding the outlier serum samples from PCOS and healthy control groups, all spectral serum samples lay inside the Hotelling T2 ellipse (95% confidence region) of the score plot (Figure 4B). The score plot represents distinct separation between healthy control and polycystic ovary syndrome (PCOS) groups and the serum samples allocated in each group’s region. OPLS-DA model shows a good fit and predictability with R²Y = 0.927 and Q² = 0.717, respectively; the validation assessment of the model for predictability was significant (P<0.05) for cross validated ANOVA (CV-ANOVA). These findings intimated variations in the pathophysiology of polycystic ovary syndrome (PCOS).

**Figure 4:** Multivariate data analysis of serum lipids
A) PCA score plot of second derivative FTIR spectra, component 1 (X-axis) explaining 18.2% and component 2 (Y-axis) explaining 14.8% of the variation
B) OPLS-DA score plot, X-axis shows the first principal component and Y-axis shows the first orthogonal component. Ellipses are the 95% confidence intervals. Circles represent polycystic ovary syndrome (PCOS) group (n=60) for PCA and (n=58) for OPLS-DA score plot, squares represent healthy control group (n=31) for PCA and (n=28) for OPLS-DA.

The differential wave numbers were selected after applying the following two criteria: 1) variable importance in projection (VIP) value exceeds 1.5 in OPLS-DA model; 2) P-value of Student’s t-test lower than 0.05. Table 4 represents the major assignments of the selected wave numbers with their absorption values in the 2800-3100 cm⁻¹ region, that dominated by the vibration of CH₂ and C-H at different wave numbers including 2846, 2848, 2850, 2852, 2854, and 2916 cm⁻¹.

**Table 4:** FTIR band assignments of the polycystic ovary syndrome (PCOS) and healthy control groups in the lipid spectral region (2800-3100 cm⁻¹)

<table>
<thead>
<tr>
<th>Wave number (cm⁻¹)</th>
<th>Absorbance</th>
<th>P-value</th>
<th>VIP value</th>
<th>Literature Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCOS</td>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2854</td>
<td>0.537966</td>
<td>0.274077</td>
<td>2.1x10⁻⁴</td>
<td>1.87993</td>
</tr>
<tr>
<td>2848</td>
<td>0.620214</td>
<td>0.343104</td>
<td>4.6x10⁻⁴</td>
<td>1.86827</td>
</tr>
<tr>
<td>2850</td>
<td>0.584703</td>
<td>0.286858</td>
<td>7.3x10⁻⁵</td>
<td>1.84117</td>
</tr>
<tr>
<td>2852</td>
<td>0.553621</td>
<td>0.259517</td>
<td>2.0x10⁻⁵</td>
<td>1.83217</td>
</tr>
<tr>
<td>2846</td>
<td>0.612739</td>
<td>0.439416</td>
<td>3.3x10⁻⁸</td>
<td>1.62166</td>
</tr>
</tbody>
</table>

Afrah Mohammed Hassan Salman et al / Using Fourier Transform Infrared (FTIR) Spectroscopy Coupled with Multivariate Analysis to Diagnose Polycystic Ovary Syndrome (PCOS)

Systematic Review Pharmacy Vol 11, Issue 4, 2020
DISCUSSION
The common characteristic feature of women with PCOS is metabolic abnormalities. However, it remains controversy about whether such metabolic abnormalities are a function of PCOS itself or of accompanying a well-known risk factors for metabolic disease [1,2]. The pathogenesis of PCOS is still unclear because of the etiological complex as well as the multiple endocrine and metabolic factors. Furthermore, there is no single marker for PCOS diagnosis because of the phenotypic heterogeneity and multi-symptomatic character of the syndrome [2]. Following Rotterdam criteria for PCOS diagnosis, two other diagnostic criteria were modified causing controversies [2]. The heterogeneity in various criteria for diagnosis of PCOS imposes difficult efforts in gathering results and making solid conclusions based on all studies [2]. In this research, we have attempted to explore the combination of FTIR spectroscopy and multivariate analysis for finding potential spectral biomarkers in serum of polycystic ovary syndrome and the opportunity of using this method as a complementary tool for Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR). The first attempt to use FTIR for serum-based diagnosis of human polycystic ovary syndrome is showed here. The presented OPLS-DA results as illustrated above demonstrated that the metabolic profiling based on FTIR technique can distinguish and identify functional groups of metabolites that exist in sera. Those metabolites were considered in the absorption regions of carbohydrates and nucleic acids (900-1200 cm\(^{-1}\)), proteins (1500-1700 cm\(^{-1}\)), and lipids (2800-3100 cm\(^{-1}\)). The FTIR spectroscopy was very sensitive for studying of carbohydrates and nucleic acids, proteins, and lipids vibration. The absorbance spectrum of the first region (complex region) at 900-1200 cm\(^{-1}\) is ascribed to the vibration of carbohydrates, phosphates and nucleic acids [40]. According to earlier published studies by Fabian et al. [41], and Wood et. al. [25], the 1000-1150 cm\(^{-1}\) spectral region is typical for carbohydrate, whereas the 900-1300 cm\(^{-1}\) region is mostly attributed to nucleic acids. Our results show that the FTIR absorption of PCOS patients' serum decreases at 900, 1080, and 1082 cm\(^{-1}\) and increased absorbance 1048, 1070, 1072, 1094, and 1096 cm\(^{-1}\) in the PCOS patient's serum. Increases in the proteins spectral region, the second region (1500-1700 cm\(^{-1}\)), of patients with PCOS at 1584, 1586, 1588, 1590, 1944, 1646, 1648, and 1650 cm\(^{-1}\) may be associated with increase protein biosynthesis when compared with control subjects. These observed changes in the main characteristics of the protein feature reflect the changes in protein secondary structure or protein folding/unfolding events may have occurred due to the PCOS. The symmetry of secondary structure appears in the IR spectrum as a change in the amide I bands, whereas the interaction or linking between hydration water and protein residuals appear as a change in the amide II bands. The alteration of both amides I and amide II bands in serum PCOS patients revealed in our result, and thus the protein content was affected. Furthermore, albumin account for 65% of total plasma protein [44], and recent study found high concentration of albumin, which correlate positively with androgen concentrations, in patients with PCOS [45]. Hence, the increases in serum protein contents may be as a result of increase serum albumin concentration. In patients with PCOS increases were observed in the lipid spectral region, the third region (2800-3100 cm\(^{-1}\)), at 2846, 2848, 2850, 2852, 2854, and 2916 cm\(^{-1}\) when compared with control subjects suggest a change in the lipids content.

As women with PCOS tend to have excessive adiposity in the abdominal cavity, the release of fatty acid by lipid catabolic process, lipolysis, is more obvious in these women [46]. It was reported by Li et al. study, which conducted to investigate serum lipid profile in patients with PCOS, that long chain saturated fatty level increased in patient with PCOS when compared with control [47]. This study demonstrates the biomedical potential of FTIR spectroscopic technique which, although not providing a clear picture of the metabolic pathway alteration, solely depends on the data of FTIR spectra and subsequent multivariate analysis shows acceptable method for PCOS diagnosis.

CONCLUSION
The cost-effective method of FTIR spectroscopy might be used as an influential technique in a clinical laboratory to distinguish the biochemical profiling of PCOS patients' serum. FTIR spectroscopy in combination with multivariate analysis offered a method for PCOS diagnosis. The most discriminatory power of variables (22 data-points) was extracted to recognize the differences between PCOS patients' serum and healthy subjects' serum by applying an OPLS-DA analysis on second derivative spectra. The characteristic patterns of these peaks provide understanding serum-based bio-molecules in patients with PCOS and healthy subjects at different FTIR absorption regions. This approach could serve as a useful laboratory method in clinical diagnosis of patients after further validation of this method in depth.
REFERENCES


