*Sys Rev Pharm 2020; 11(4): 552 – 560*
A multifaceted review journal in the field of pharmacy
E-ISSN 0976-2779 P-ISSN 0975-8453

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

Afrah Mohammed Hassan Salman[1*], Mohammed A. Al-Zubaidi[1], Maysaa Ali Abdul khaleq[2]
[1]Department of Clinical Laboratory Sciences, College of Pharmacy, Mustansiriyah University, Baghdad, Iraq
[2]Al-Rasheed University College/Department of Dentistry, Baghdad, Iraq
**Corresponding author**: Afrah Mohammed Hassan Salman, Email: afrahbassil@yahoo.com

*Article History:* *Submitted: 23.01.2020* *Revised: 26.03.2020* *Accepted: 30.04.2020*

## ABSTRACT
The biochemical diagnostic test is important in the identification of patients with polycystic ovary syndrome (PCOS). Fourier Transfer 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

**Corresponding author**:
Afrah Mohammed Hassan Salman
Department of Clinical Laboratory Sciences, College of Pharmacy, Mustansiriyah University, Baghdad, Iraq
Email: afrahbassil@yahoo.com

**DOI:** 10.31838/srp.2020.4.82



## 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 $^{o}$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/Mustansiriyah 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 $^{o}C$ 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

| Characteristics | PCOS (n=60) Mean ± SD | Control (n=31) Mean ± SD | P-value |
|---|---|---|---|
| Age (years) | 25.18 ± 4.86 | 25.67 ± 5.37 | $65 \times 10^{-2}$ |
| BMI ($kg/m^2$) | 24.26 ± 0.39 | 24.80 ± 0.51 | $42 \times 10^{-2}$ |
| LH (mIU/ml) | 11.67 ± 2.31 | 5.33 ± 1.27 | $1.58 \times 10^{-24}$ |
| FSH (mIU/ml) | 7.36 ± 1.56 | 6.50 ± 1.48 | $1.37 \times 10^{-2}$ |

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.

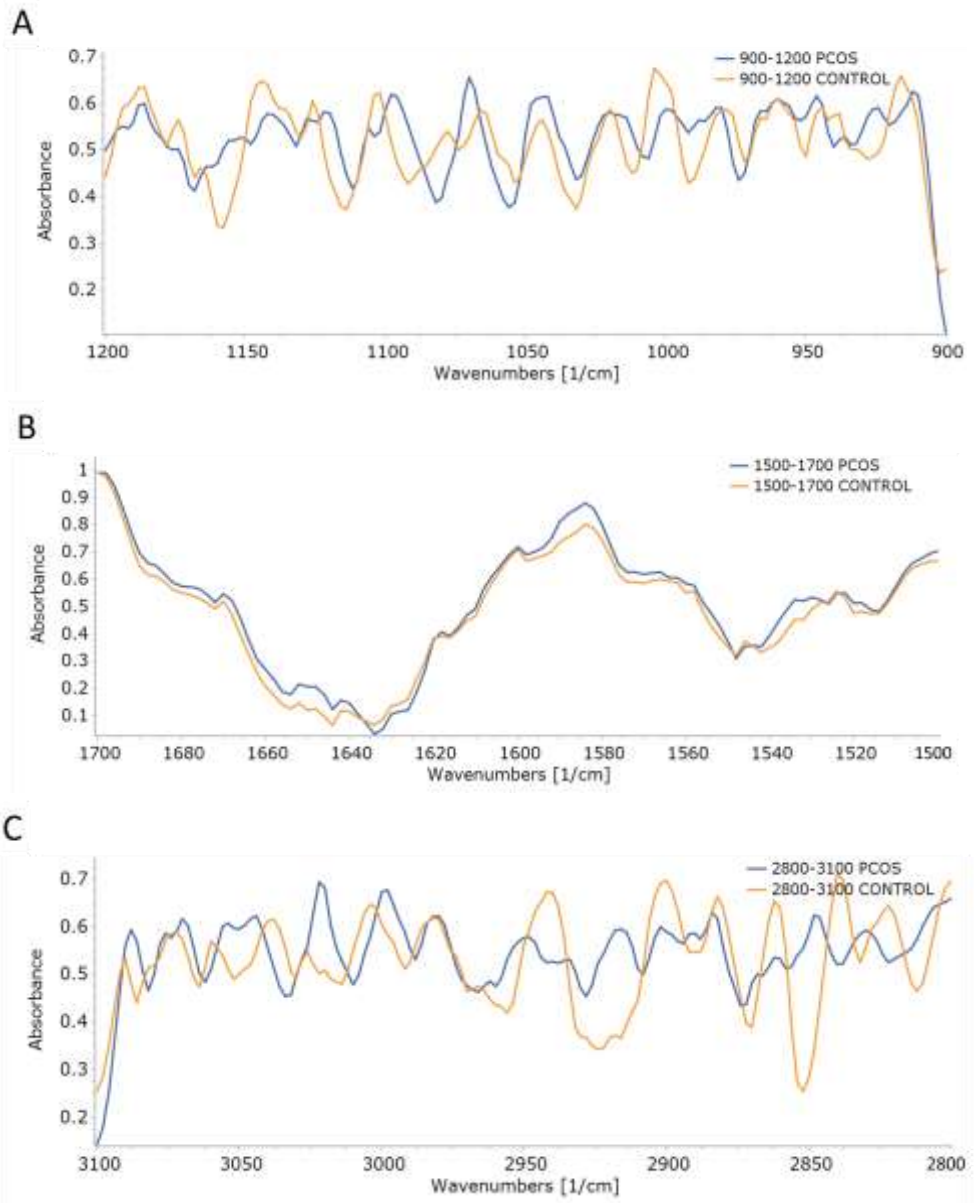


Figure 1: Mean of pre-processed FTIR spectra within the recorded regions obtained from three replicates of 60 PCOS serum samples and 31 healthy control serum samples

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).

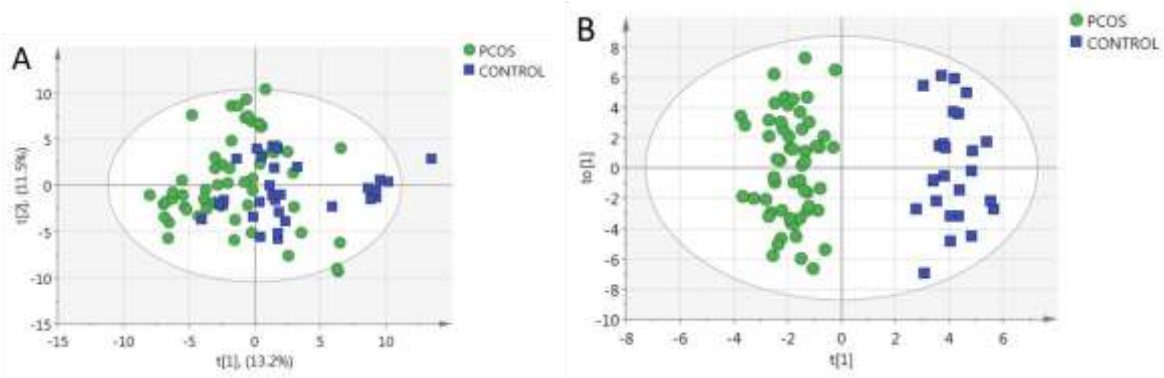


Figure 2: Multivariate data analysis of serum carbohydrates and nucleic acids

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 $PO_2^-$ 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 2: FTIR band assignments of the polycystic ovary syndrome (PCOS) and healthy control groups in the carbohydrate and nucleic acid spectral region (900-1200 $cm^{-1}$).

| Wave number ($cm^{-1}$) | Absorbance | | P-value | VIP value | Literature Assignment |
|---|---|---|---|---|---|
| | PCOS | Control | | | |
| 1070 | 0.659224 | 0.521274 | $9.26x10^{-7}$ | 1.83258 | Symmetric phosphate ($PO_2^-$) stretching [21] |
| 900 | 0.106877 | 0.236634 | $2.7x10^{-7}$ | 1.82489 | Deoxyribose ring vibration [22] |
| 1072 | 0.621799 | 0.488437 | $1.52x10^{-5}$ | 1.73553 | ($PO_2^-$) of nucleic acid vibration [23] |
| 1048 | 0.614482 | 0.489394 | $2.12x10^{-5}$ | 1.61017 | C-O stretching of sugar ring of nucleic acid [6] |
| 1096 | 0.6317 | 0.477421 | $2.76x10^{-5}$ | 1.59522 | Symmetric phosphate ($PO_2^-$) stretching [24] |
| 1080 | 0.392146 | 0.500194 | $9.22x10^{-5}$ | 1.58041 | Symmetric phosphate ($PO_2$-) stretching [25] |
| 1082 | 0.384169 | 0.497025 | $2.70x10^{-4}$ | 1.57709 | Symmetric phosphate ($PO_2^-$) stretching [26] |

| | | | | | |
|---|---|---|---|---|---|
| 1094 | 0.608843 | 0.45762 | $8.54x10^{-5}$ | 1.55896 | Symmetric phosphate ($PO_2^-$) stretching [27] |

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, R2X (cum) = 0.977, and predictability, Q2 (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 R2Y=0.828 and Q2=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).

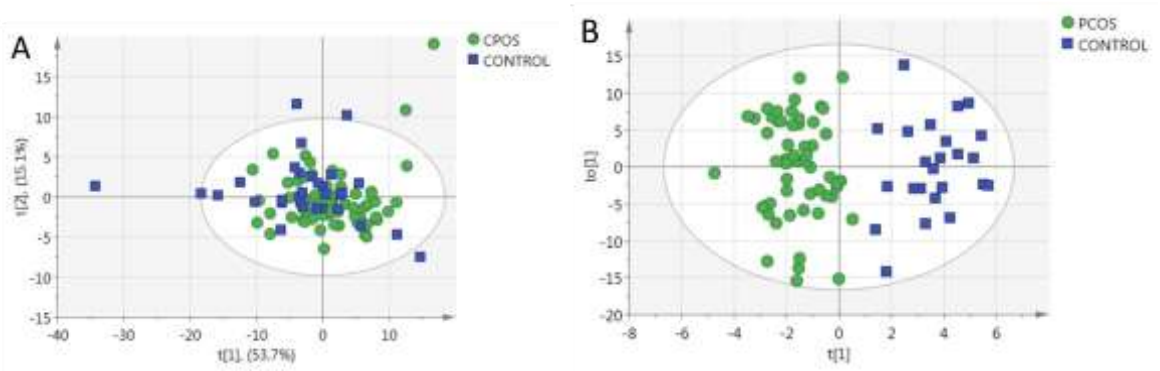


Figure 3: Multivariate data analysis of serum proteins

A) PCA score plot of second derivative FTIR spectra, component 1 (X-axis) explaining 53.7% and component 2 (Y-axis) explaining 15.1% 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=27) 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 3 represents the major assignments of the selected wave numbers with their absorption values in the 1500-1700 $cm^{-1}$ region, that dominated by the vibration of C=O, N-H, and C-N in amide I and II at different wave numbers including 1584, 1586, 1588, 1590, 1644, 1646, 1648, and 1650 $cm^{-1}$

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

| Wave number ($cm^{-1}$) | Absorbance | | P-value | VIP value | Literature Assignment |
|---|---|---|---|---|---|
| | PCOS | Control | | | |
| 1648 | 0.201187 | 0.117069 | $3.67x10^{-8}$ | 1.96534 | C=O different conformation of amide I [28]<br>Amide I [29] |
| 1646 | 0.16449 | 0.091304 | $2.26x10^{-6}$ | 1.89095 | C=O stretching of amide I [30] |
| 1650 | 0.218399 | 0.116621 | $2.43x10^{-7}$ | 1.80128 | C=O stretching of amide I [25] |
| 1588 | 0.839009 | 0.783913 | $1.46x10^{-6}$ | 1.7473 | N-H bending and C-N stretching of Amide II [6] |
| 1586 | 0.857661 | 0.803493 | $1.79x10^{-6}$ | 1.71194 | N-H bend and C-N stretch of amide II [31] |
| 1584 | 0.877991 | 0.830988 | $2.19x10^{-5}$ | 1.57807 | α-helical structures of amide II [32] |

| | | | | | |
|---|---|---|---|---|---|
| 1644 | 0.111162 | 0.055303 | $1.01x10^{-3}$ | 1.57413 | C=O stretching of amide I [33] |
| 1590 | 0.809107 | 0.765447 | $1.01x10^{-4}$ | 1.51125 | N-H bend and C-N stretch of amide II [31] |

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).

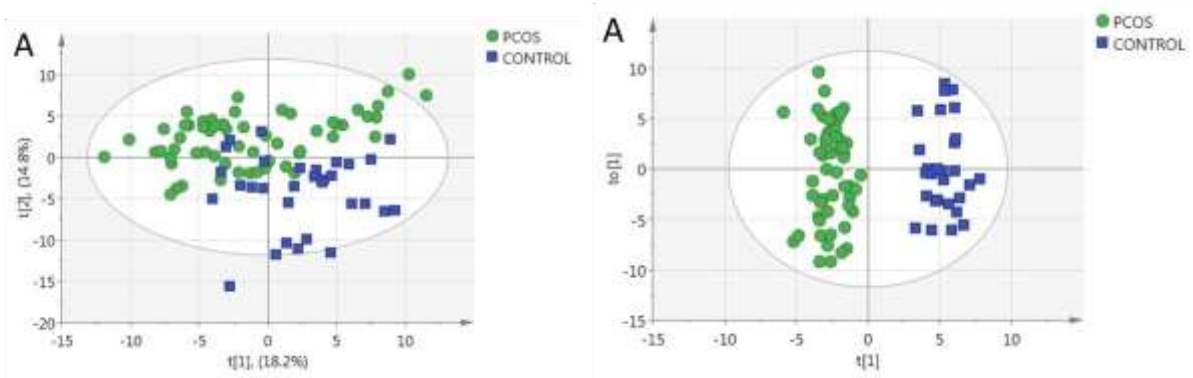


Figure 4: Multivariate data analysis of serum lipids

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 $CH_2$ 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 $cm^{-1}$)

| Wave number ($cm^{-1}$) | Absorbance | | P-value | VIP value | Literature Assignment |
|---|---|---|---|---|---|
| | PCOS | Control | | | |
| 2854 | 0.537966 | 0.274077 | $2.1x10^{-14}$ | 1.87993 | Symmetric stretching vibration of $CH_2$ [34] |
| 2848 | 0.620214 | 0.343104 | $4.6x10^{-14}$ | 1.86827 | Symmetric stretching vibration of $CH_2$ [35] |
| 2850 | 0.584703 | 0.286858 | $7.3x10^{-13}$ | 1.84117 | Symmetric stretching vibration of $CH_2$ [36] |
| 2852 | 0.553621 | 0.259517 | $2.0x10^{-12}$ | 1.83217 | Symmetric stretching vibration of $CH_2$ [37] |
| 2846 | 0.612739 | 0.439416 | $3.3x10^{-9}$ | 1.62166 | Symmetric stretching of $CH_2$ [38] |

| | | | | | |
|---|---|---|---|---|---|
| 2916 | 0.590437 | 0.387453 | $3.55x10^{-8}$ | 1.52038 | Stretching vibration of C-H [39] |

## 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 is 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 $cm^{-1}$, while increases at 1048, 1070, 1072, 1094, and 1096 $cm^{-1}$. The observed changes would offer a signature in this region of FTIR spectra to characterize PCOS. However, it was reported that the environmental factor affect the IR absorption of nucleic acids and their absorption obviously appear in the absence of glycogen [42]. Hence, any change in glycogen content, which overlaps with nucleic acid absorbance in this region, can influence the nucleic acid signals and lead to difficult identification and measurement the change in the value of any peak wave number precisely in such alterations [43]. Since the IR absorbance is related to the concentration of various bio-molecules in the sample, change of any bio-molecules concentration with respect to each other would distinguish themselves as alteration in the spectra. This is may be the reason for the decreased absorbance 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


1. Couto Alves A, Valcarcel B, Mäkinen V-P, Morin-Papunen L, Sebert S, Kangas AJ, et al. Metabolic profiling of polycystic ovary syndrome reveals interactions with abdominal obesity. Int J Obes. 2017;41(9):1331–40.
2. RoyChoudhury S, Mishra BP, Khan T, Chattopadhayay R, Lodh I, Datta Ray C, et al. Serum metabolomics of Indian women with polycystic ovary syndrome using 1 H NMR coupled with a pattern recognition approach. Mol BioSyst. 2016;12(11):3407–16.
3. Chen Y-X, Zhang X-J, Huang J, Zhou S-J, Liu F, Jiang L-L, et al. UHPLC/Q-TOFMS-based plasma metabolomics of polycystic ovary syndrome patients with and without insulin resistance. Journal of Pharmaceutical and Biomedical Analysis. 2016;121:141–50.
4. Murri M, Insenser M, Escobar-Morreale HF. Metabolomics in polycystic ovary syndrome. Clinica Chimica Acta. 2014;429:181–8.
5. Parachalil DR, Bruno C, Bonnier F, Blasco H, Chourpa I, Baker MJ, et al. Analysis of bodily fluids using vibrational spectroscopy: a direct comparison of Raman scattering and infrared absorption techniques for the case of glucose in blood serum. Analyst. 2019;144(10):3334–46.
6. iBiMED University of Aveiro, Nunes A. FTIR Spectroscopy - A Potential Tool to Identify Metabolic Changes in Dementia Patients. AND. 2016;2(2):1–9.
7. Movasaghi Z, Rehman S, ur Rehman DrI. Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues. Applied Spectroscopy Reviews. 2008;43(2):134–79.
8. Mitchell AL, Gajjar KB, Theophilou G, Martin FL, Martin-Hirsch PL. Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting: Vibrational spectroscopy of biofluids: laboratory to clinical setting. J Biophoton. 2014;7(3–4):153–65.
9. Rehman S, Movasaghi Z, Darr JA, Rehman IU. Fourier Transform Infrared Spectroscopic Analysis of Breast Cancer Tissues; Identifying Differences between Normal Breast, Invasive Ductal Carcinoma, and Ductal Carcinoma In Situ of the Breast. Applied Spectroscopy Reviews. 2010;45(5):355–68.
10. Kallenbach-Thieltges A, Großerüschkamp F, Mosig A, Diem M, Tannapfel A, Gerwert K. Immunohistochemistry, histopathology and infrared spectral histopathology of colon cancer tissue sections. J Biophoton. 2013;6(1):88–100.
11. Backhaus J, Mueller R, Formanski N, Szlama N, Meerpohl H-G, Eidt M, et al. Diagnosis of breast cancer with infrared spectroscopy from serum samples. Vibrational Spectroscopy. 2010;52(2):173–7.
12. Ollesch J, Heinze M, Heise HM, Behrens T, Brüning T, Gerwert K. It's in your blood: spectral biomarker candidates for urinary bladder cancer from automated FTIR spectroscopy: Spectral cancer biomarkers from high-throughput FTIR spectroscopy. J Biophoton. 2014;7(3–4):210–21.
13. Hands JR, Dorling KM, Abel P, Ashton KM, Brodbelt A, Davis C, et al. Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectral discrimination of brain tumour severity from serum samples: Serum spectroscopy gliomas. J Biophoton. 2014;7(3–4):189–99.
14. Barlev E, Zelig U, Bar O, Segev C, Mordechai S, Kapelushnik J, et al. A novel method for screening colorectal cancer by infrared spectroscopy of peripheral blood mononuclear cells and plasma. J Gastroenterol. 2016;51(3):214–21.
15. Peuchant E, Richard-Harston S, Bourdel-Marchasson I, Dartigues J-F, Letenneur L, Barberger-Gateau P, et al. Infrared spectroscopy: a reagent-free method to distinguish Alzheimer's disease patients from normal-aging subjects. Translational Research. 2008;152(3):103–12.
16. Travo A, Paya C, Déléris G, Colin J, Mortemousque B, Forfar I. Potential of FTIR spectroscopy for analysis of tears for diagnosis purposes. Anal Bioanal Chem. 2014;406(9–10):2367–76.
17. Baker MJ, Hussain SR, Lovergne L, Untereiner V, Hughes C, Lukaszewski RA, et al. Developing and understanding biofluid vibrational spectroscopy: a critical review. Chem Soc Rev. 2016;45(7):1803–18.
18. Trevisan J, Park J, Angelov PP, Ahmadzai AA, Gajjar K, Scott AD, et al. Measuring similarity and improving stability in biomarker identification methods applied to Fourier-transform infrared (FTIR) spectroscopy: Stability and similarity in biomarker identification in FTIR spectroscopy. J Biophoton. 2014;7(3–4):254–65.
19. Le Corvec M, Jezequel C, Monbet V, Fatih N, Charpentier F, Tariel H, et al. Mid-infrared spectroscopy of serum, a promising non-invasive method to assess prognosis in patients with ascites and cirrhosis. Strnad P, editor. PLoS ONE. 2017;12(10):e0185997.
20. Worley B, Powers R. PCA as a Practical Indicator of OPLS-DA Model Reliability. CMB. 2016;4(2):97–103.
21. Liu Y, Xu Y, Liu Y, Zhang Y, Wang D, Xiu D, et al. Detection of cervical metastatic lymph nodes in papillary thyroid carcinoma by Fourier transform infrared spectroscopy. Br J Surg. 2011;98(3):380–4.
22. Mello MLS, Vidal BC. Changes in the Infrared Microspectroscopic Characteristics of DNA Caused by Cationic Elements, Different Base Richness and Single-Stranded Form. Tajmir-Riahi H-A, editor. PLoS ONE. 2012;7(8):e43169.
23. Zlotogorski-Hurvitz A, Dekel BZ, Malonek D, Yahalom R, Vered M. FTIR-based spectrum of salivary exosomes coupled with computational-aided discriminating analysis in the diagnosis of oral cancer. J Cancer Res Clin Oncol. 2019;145(3):685–94.
24. Olszty S, Szymborska-Ma K. Spectroscopic techniques in the study of human tissues and their components. Part I: IR spectroscopy. Acta of Bioengineering and Biomechanics. 2012;14(3):101–15.

25. Wood BR, Quinn MA, Tait B, Ashdown M, Hislop T, Romeo M, et al. FTIR microspectroscopic study of cell types and potential confounding variables in screening for cervical malignancies. Biospectroscopy. 1998;4(2):75–91.
26. Fung MFK, Senterman MK, Mikhael NZ, Lacelle S. Pressure-tuning fourier transform infrared spectroscopic study of carcinogenesis in human endometrium. Biospectroscpoy. 1996;2(3):155–65.
27. Dovbeshko G. FTIR spectroscopy studies of nucleic acid damage. Talanta. 2000;53(1):233–46.
28. Jiang W, Saxena A, Song B, Ward BB, Beveridge TJ, Myneni SCB. Elucidation of Functional Groups on Gram-Positive and Gram-Negative Bacterial Surfaces Using Infrared Spectroscopy. Langmuir. 2004;20:11433–42.
29. Šťovíčková L, Tatarkovič M, Logerová H, Vavřinec J, Setnička V. Identification of spectral biomarkers for type 1 diabetes mellitus using the combination of chiroptical and vibrational spectroscopy. Analyst. 2015;140(7):2266–72.
30. Di Giambattista L, Grimaldi P, Gaudenzi S, Pozzi D, Grandi M, Morrone S, et al. UVB radiation induced effects on cells studied by FTIR spectroscopy. Eur Biophys J. 2010;39(6):929–34.
31. Seredin P, Goloshchapov D, Ippolitov Y, Vongsvivut J. Spectroscopic signature of the pathological processes of carious dentine based on FTIR investigations of the oral biological fluids. Biomed Opt Express. 2019;10(8):4050.
32. Böcker U, Ofstad R, Wu Z, Bertram HC, Sockalingum GD, Manfait M, et al. Revealing Covariance Structures in Fourier Transform Infrared and Raman Microspectroscopy Spectra: A Study on Pork Muscle Fiber Tissue Subjected to Different Processing Parameters. Appl Spectrosc. 2007;61(10):1032–9.
33. Bartošová A, Blinová L, Gerulová K. Characterisation Of Polysacharides And Lipids From Selected Green Algae Species By FTIR-ATR Spectroscopy. Research Papers Faculty of Materials Science and Technology Slovak University of Technology. 2015;23(36):97–102.
34. Ostrowska KM, Garcia A, Meade AD, Malkin A, Okewumi I, O'Leary JJ, et al. Correlation of p16INK4A expression and HPV copy number with cellular FTIR spectroscopic signatures of cervical cancer cells. Analyst. 2011;136(7):1365.
35. Chaber R, Arthur CJ, Depciuch J, Łach K, Raciborska A, Michalak E, et al. Distinguishing Ewing sarcoma and osteomyelitis using FTIR spectroscopy. Sci Rep. 2018;8(1):15081.
36. Zhang D, Lan L, Bai Y, Majeed H, Kandel ME, Popescu G, et al. Bond-selective transient phase imaging via sensing of the infrared photothermal effect. Light Sci Appl. 2019;8(1):116.
37. Ami D, Mereghetti P, Leri M, Giorgetti S, Natalello A, Doglia SM, et al. A FTIR microspectroscopy study of the structural and biochemical perturbations induced by natively folded and aggregated transthyretin in HL-1 cardiomyocytes. Sci Rep. 2018;8(1):12508.
38. Sherif MS, Mervat AA, Eman AM. Infrared spectroscopic investigation of erythrocyte membrane-smoke interactions due to chronic cigarette smoking. gpb. 2017;36(03):273–80.
39. Caetano Júnior PC, Strixino JF, Raniero L. Analysis of saliva by Fourier transform infrared spectroscopy for diagnosis of physiological stress in athletes. Res Biomed Eng. 2015;31(2):116–24.
40. Lin H, Deng K, Zhang J, Wang L, Zhang Z, Luo Y, et al. Biochemical detection of fatal hypothermia and hyperthermia in affected rat hypothalamus tissues by Fourier transform infrared spectroscopy. Bioscience Reports. 2019;39(3):BSR20181633.
41. Fabian H, Jackson M, Murphy L, Watson PH, Fichtner I, Mantsch HH. A comparative infrared spectroscopic study of human breast tumors and breast tumor cell xenografts. Biospectroscopy. 1995;1(1):37–45.
42. Lasch P, Boese M, Pacifico A, Diem M. FT-IR spectroscopic investigations of single cells on the subcellular level. Vibrational Spectroscopy. 2002;28(1):147–57.
43. Sahu RK, Argov S, Salman A, Huleihel M, Grossman N, Hammody Z, et al. Characteristic Absorbance of Nucleic Acids in the Mid-IR Region as Possible Common Biomarkers for Diagnosis of Malignancy. Technol Cancer Res Treat. 2004;3(6):629–38.
44. Petibois C, Cazorla G, Cassaigne A, Déléris G. Plasma Protein Contents Determined by Fourier-Transform Infrared Spectrometry. Clinical Chemistry. 2001;47(4):730–8.
45. Caglar GS, Oztas E, Karadag D, Pabuccu R, Eren AA. The association of urinary albumin excretion and metabolic complications in polycystic ovary syndrome. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2011;154(1):57–61.
46. Carvalho LML, dos Reis FM, Candido AL, Nunes FFC, Ferreira CN, Gomes KB. Polycystic Ovary Syndrome as a systemic disease with multiple molecular pathways: a narrative review. Endocrine Regulations. 2018;52(4):208–21.
47. Goyal M, Dawood A. Debates regarding lean patients with polycystic ovary syndrome: A narrative review. J Hum Reprod Sci. 2017;10(3):154.