

# UPLC/Q-TOF-MS based plasma metabolomics and clinical characteristics of polycystic ovarian syndrome

XUEMEI FAN<sup>1\*</sup>, JIANFA JIANG<sup>2\*</sup>, ZHIQIANG HUANG<sup>1</sup>, JUMEI GONG<sup>1</sup>, YIMING WANG<sup>1</sup>, WEI XUE<sup>3</sup>,  
YAN DENG<sup>3</sup>, YANFANG WANG<sup>3</sup>, TINGPING ZHENG<sup>3</sup>, AIJUN SUN<sup>3</sup> and GUOAN LUO<sup>1</sup>

<sup>1</sup>Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology (Ministry of Education), Department of Chemistry, Tsinghua University, Beijing 100084; <sup>2</sup>Department of Obstetrics and Gynecology, Third Xiangya Hospital, Central South University, Changsha, Hunan 410013; <sup>3</sup>Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing 100730, P.R. China

Received July 5, 2017; Accepted September 18, 2018

DOI: 10.3892/mmr.2018.9643

**Abstract.** The present study aimed to develop novel diagnostic methods for polycystic ovarian syndrome (PCOS) by screening and identifying specific PCOS-associated metabolic markers using plasma metabolomics. Ultra-performance liquid chromatography/quadrupole-time of flight-mass spectrometry was adopted to establish the plasma metabolic fingerprint of 49 patients and 50 normal controls, in order to screen the potential metabolic markers. In addition, these markers were integrated with the clinical indexes, followed by focused analysis to obtain diagnostic markers. The present results demonstrated that not only was the concentration of palmitoyl sphingomyelin in plasma of patients with PCOS significantly increased; however, a statistically significant difference between the two PCOS subgroups was additionally demonstrated. At the same time, the concentrations of cyclic guanosine monophosphate (cGMP) and dehydroepiandrosterone sulphate in the plasma of patients of the subgroup 1 were significantly elevated. These markers were additionally integrated with the clinical index number of follicles in the left ovary and high-density lipoprotein (HDL-C), followed by

receiver operating characteristic curve analysis, which demonstrated a diagnostic accuracy of ~90% in the control and the two subgroups. The integrated marker system consisting of palmitoyl sphingomyelin, cGMP and androsterone sulfate, as well as the number of left follicles and HDL-C may be used for the accurate diagnosis and classification of PCOS. These results confirmed that the abnormalities in hormone metabolism and lipid metabolism disorder were primarily involved in the onset of PCOS.

## Introduction

Polycystic ovarian syndrome (PCOS) is an endocrine syndrome manifesting as dysgonesis and metabolic disorder, with a prevalence rate accounting for 5-10% of the women of childbearing age and 50-70% with anovulatory infertility (1,2). The majority of previous studies speculated that PCOS is a polygenic inheritance disease (3-5). Peroxisome proliferators-activated receptor gene and tumor necrosis factor (6) and its receptor genes may affect the occurrence of the disease. The pathogenesis may be associated with insulin resistance (7), metabolism (8), inflammation and an immune response (3). However, no genes have been demonstrated to be associated directly with the pathogenesis of PCOS. Previous studies proposed that the pathological characteristic of PCOS is a hypothalamic-pituitary-ovarian axis dysfunction (9-11). However, a number of patients additionally present with hyperinsulinemia and hyperandrogenism, which may directly damage the great vessels, leading to sympathetic activity and causing glucose and lipid metabolism disorders (12), thereby increasing blood viscosity and resulting in obesity. In recent years, the mechanism of metabolic disorders in patients with PCOS has been the focus of intensive research. Previous studies adopted nuclear magnetic resonance (13,14), gas chromatography-mass spectrometry (8), liquid chromatography-mass spectrometry (7,15) and other methods to investigate the metabolite disorder and the possible pathogenesis in patients with PCOS. The results of these studies demonstrated that patients with PCOS suffered from an amino acid metabolism disorder, in which the valine and leucine levels were increased,

*Correspondence to:* Professor Yiming Wang, Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology (Ministry of Education), Department of Chemistry, Tsinghua University, 1 Qinghuayuan, Beijing 100084, P.R. China  
E-mail: wangyim@tsinghua.edu.cn

Professor Aijun Sun, Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, 1 Shuaifuyuan, Beijing 100730, P.R. China  
E-mail: 2249256116@qq.com

\*Contributed equally

**Key words:** polycystic ovarian syndrome, metabolomics, potential metabolic marker, integrated marker system, accurate diagnosis

whereas, the glycine level was decreased, which may enhance insulin sensitivity. Additionally, patients with PCOS demonstrated abnormal plasma lipid metabolism, in which the lactic acid, long-chain fatty acid, triglycerides and low-density lipoprotein levels were elevated. Furthermore, an elevated level of acetylglycoprotein may predict a mild chronic inflammation in patients with PCOS.

Due to the heterogeneity and polymorphism of clinical manifestations, the clinical diagnosis and treatment of PCOS is difficult. Additionally, the studies mentioned above identified that alterations in the *in vivo* metabolites serve a vital role in the pathogenesis of PCOS. Therefore, in the present study, ultra-performance liquid chromatography/quadrupole-time of flight-mass spectrometry (UPLC/Q-TOF-MS) technology was utilized to establish the metabolic fingerprints of patients with PCOS and healthy controls, in order to screen the specific biomarkers associated with PCOS and evaluate their accuracy in the diagnosis and classification of the disease.

## Materials and methods

**Instruments and reagents.** The following instruments were used: Waters ACQUITY UPLC (Waters Corp., Milford, MA, USA) utilized the sub-micron particle technology; Waters Xevo G2 Q-TOF (Waters Corp.) high-resolution MS with electron spray ionization (ESI) source; MIKRO 22R tabletop high-speed low-temperature centrifuge (Hettich GmbH & Co., KG, Tuttlingen, Germany); and VORTEX-5 turbine mixer (Haimen Kylin-Bell Lab Instruments Co., Ltd., Jiangsu, China).

The reagents utilized were: LC/MS grade methanol and acetonitrile (Merck KGaA, Darmstadt, Germany); chromatographic grade formic acid (Fluka, Buchs, Switzerland); leucine enkephalin (Sigma-Aldrich; Merck KGaA); and ultrapure water prepared using Milli-Q ultrapure water system (18.2 M $\Omega$ ; EMD Millipore, Billerica, MA, USA). All the other reagents used were either of analytical grade or of the best grade available.

**Subjects.** The present study was approved by the Ethics Committee of Peking Union Medical College Hospital (approval no. S-733). Written informed consent was obtained from all participants.

A total of 49 patients diagnosed with PCOS at the Department of Obstetrics and Gynecology, Peking Union Medical College Hospital (Beijing, China), between July 2014 and March 2015, were included and not administered any drug treatment prior to sampling. Inclusion criteria: Age 18–40 years, diagnosed with PCOS according to the 2003 Rotterdam consensus (16,17): i) Oligo-ovulation or anovulation; ii) clinical manifestations of excessive androgen and (or) hyperandrogenism; iii) polycystic alterations in the ovary: Ultrasound demonstrated  $\geq 12$  follicles with a diameter of 2–9 mm at unilateral or bilateral ovaries, and (or) ovarian volume  $\geq 10$  ml; and iv) patients with two out of the above three features with exclusion of other causes for excessive androgen. Simultaneously, 50 healthy volunteers were recruited with the following inclusion criteria: Age 18–40 years and a normal menstrual cycle with the absence of PCOS, diabetes and other

endocrine diseases, in addition to a lack of a family history of PCOS.

**Collection of samples and clinical data.** A total of 8 ml morning fasted elbow vein blood was collected from healthy volunteers at the 2nd–4th day of menstruation and from patients at the 2nd–4th day of menstruation or  $\geq 30$  days following menopause with the absence of dominant follicles in ultrasound as assessed by professional medical staff. The sample in the EDTA anticoagulant tube was centrifuged at 4°C and 870 x g for 15 min, from which the plasma was collected, dispensed in 1.5 ml tubes and preserved at -80°C.

Concurrently, all the subjects underwent evaluation of general indexes and various biochemical markers. The general indexes included: Body height (m), body weight (kg), waist circumference (cm), and hip circumference (cm), without shoes and in an unlined garment, based on which the body mass index [BMI=body weight/height<sup>2</sup> (kg/m<sup>2</sup>)] and waist:hip ratio (WHR=waist circumference/hip circumference) were calculated. Subsequently, at 3–7 days following menstruation or  $\geq 30$  days following menopause, the subjects underwent an ultrasound examination to record the number of right follicles, the size of right follicles (cm), the number of left follicles and the size of left follicles (cm).

In addition, the biochemical markers included: Follicle stimulating hormone (FSH), luteinizing hormone (LH), prolactin (PRL), estradiol (E2), testosterone (T), total serum cholesterol (TC), triglycerides (TG), high-density lipoprotein (HDL-C), low-density lipoprotein (LDL-C), apolipoprotein A1 (ApoA1), apolipoprotein B (ApoB), high-sensitivity C-reactive protein (hsCRP), glucose (GLU), insulin (INS) and glycosylated hemoglobin (HbA1c). Plasma FSH, LH, PRL, E2, T and INS were measured by chemiluminescence using DxI800 analysers (Beckman Coulter, Brea, CA, USA). Plasma TC, TG, HDL-C, LDL-C, ApoA1, ApoB and hsCRP were measured by biochemical assays using Beckman AU5800 analyser (Beckman Coulter). HbA1c was performed in an automated chemistry analyzer Variant II Turbo (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

**Preparation of analytical samples.** The preserved plasma sample was thawed at 4°C and mixed by oscillation. Subsequently, 100  $\mu$ l sample was collected, added to 400  $\mu$ l methanol, followed by vortex agitation for 2 min to allow protein precipitation. The mixture was centrifuged (4°C at 17,700 x g for 20 min), from which 400  $\mu$ l supernatant was transferred to a 1.5 ml tube, and the solvent volatilized using N<sub>2</sub>. Subsequently, the sample was mixed with 120  $\mu$ l (3-fold of the amount of plasma) reconstitution solvent acetonitrile-water (1:4 v/v) for resolubilization, followed by vortex oscillation for 2 min and centrifugation (4°C at 17,700 x g for 30 min). In total, 10  $\mu$ l supernatant was placed into the autosampler vial for analysis.

**UPLC/Q-TOF-MS analysis.** A Boston Green ODS BEH-C18 column (100x2.1 mm, i.d. 1.7  $\mu$ m; Waters Corp.) was used for chromatography, with a flow rate of 0.4 ml/min and a 10  $\mu$ l injection volume, and the column temperature was maintained at 40°C. The chromatographic mobile phase was composed of A and B phases: The A phase referred to 0.1% (v/v) formic acid-water and

B phase referred to 0.1% (v/v) formic acid-acetonitrile, wherein the gradient elution was 0-6.5 min, 5-50% B; 6.5-10.5 min, 50-60% B; 10.5-13.5 min, 60-70% B; 13.5-17.5 min, 70-95% B; 17.5-20 min, 95-95% B; 20-23 min, 95-5% B.

ESI-Q-TOF-MS was collected in a negative ion mode; capillary voltage 3,000 V, cone voltage 30 V, desolvation gas temperature 350°C, ion source temperature 110°C, desolvation gas flow 600 l/h and cone gas flow 50 l/h; acquisition time 0-23 min, scanning range m/z 100-1,500, each scan process lasting for 0.1 sec with an interval of 0.02 sec. In the test, leucine enkephalin (200 ng/ml) was used as the tuned liquid while the mass axis was corrected in real time to ensure the accuracy and reproducibility of the instrument.

The solution for quality control (QC) was obtained by mixing an equivalent amount (10  $\mu$ l) of each plasma sample that assessed the stability of the instrument. Prior to the experiment, 10 QC samples were successively tested for balancing the columns, and subsequently, the QC of the sample was performed one time following every five random samples that were analyzed. The stability of the instrument was first evaluated by the overlap of the chromatographic peaks of the QC samples; subsequently, eight ion peaks were randomly extracted from the chromatographic peaks for stability and reproducibility analyses.

**Statistical analysis.** The raw mass spectra data were generated following chromatographic separation and mass spectrum acquired from the plasma samples, followed by filtering the noise peaks testing and peak-matching using MassLynx 4.0 (Waters Corp.) software. A three-dimensional data matrix was obtained, including the retention time, mass-to-charge ratio and peak area, which would be used for subsequent multivariate and focused analyses for characteristic indexes.

The multivariate statistical analyses were performed using SIMCA-P 12.0 (Umetrics, Umea, Sweden) software. Data patterns were recognized using partial least-squares discriminant analysis (PLS-DA) and orthogonal partial least-squares discriminant (OPLS-DA) methods. Additionally, differential variables with significant contribution were selected according to the variable important in projection (VIP) obtained by OPLS-DA model. An independent samples t-test was performed to obtain the metabolites list containing the retention time, mass-to-charge ratio, response intensity and other information.

Identification of metabolites was primarily based on the accurate molecular mass, retention time and ion mode. In the present study, the potential metabolic markers were identified using the following methods: i) Mass deviation from the actual chemical formula (expressed as mDa or ppm), reasonable double bond number, and i-FIT value were used to judge the possibility of the chemical formula, followed by calculation of the potential equation; ii) mass spectrum data were matched with those in the MassBank (<http://www.massbank.jp/>) and Metlin library (<http://metlin.scripps.edu/>); iii) data were corroborated with the fragment peaks reported in the literature; and iv) the molecular formula was imported into the Human Metabolome Database (<http://www.hmdb.ca/>) and Kyoto Encyclopedia of Genes and Genomes Ligand ([www.genome.jp/kegg/ligand.html](http://www.genome.jp/kegg/ligand.html)) for searching and identifying the structural information,

in order to obtain the PCOS disease-associated potential biomarkers.

All the statistical analyses were performed using SPSS 21.0 software (IBM Corp., Armonk, NY, USA). Kolmogorov-Smirnov was used to evaluate and screen the normality of the general indexes and biochemical markers. Continuous variables in the normal distribution are presented as the mean  $\pm$  standard deviation. Differential characteristic indexes were screened using an independent samples t-test. The focused metabolic markers data were analyzed using a one-way analysis of variance followed by Tukey's Honest Significant Difference test.  $P < 0.05$  was considered to indicate a statistically significant difference.

**Focused analysis of characteristic indexes.** Since there were a number of clinical indexes and potential metabolic markers obtained by metabolomics analysis to determine the characteristic indexes for PCOS, artificial neural network (ANN) was adopted to establish the NeuroFuzzy logic model. The variables with the most significant contribution to the model were demonstrated, thereby, generating the focused variable sets and screening the smallest specific index set of PCOS, known as the focused integrated marker system. The simple procedures included were as follows: Firstly, the markers underwent a focused analysis for screening the potential optimal characteristic index set, followed by an evaluation of the model. The 99 samples were randomly divided into the training set (83 samples; 80%) and test set (16 samples; 20%) for the ANN model training. The data of the test set were used to predict and evaluate the training results, thereby ultimately obtaining the  $R^2$  value of the fitted line of the real and predicted value with the prediction accuracy. Subsequently, the receiver-operating characteristic (ROC) was adopted to analyze the sensitivity and specificity of the key markers obtained by focused analysis in the diagnosis of PCOS, where the area under the curve was used to determine the diagnostic value. All the ROC analyses were performed using SPSS 21.0 software.

## Results

**Clinical data.** The ages of the control and PCOS groups were  $25.08 \pm 3.59$  years and  $27.32 \pm 4.90$  years, respectively, without a statistically significant difference. Therefore, 19 crucial clinical indexes in the two groups were analyzed (Table I), and the statistical analyses demonstrated that PRL, ApoA1 and HbA1c did not differ significantly between the two groups. However, the BMI ( $P = 7.21 \times 10^{-6}$ ) and WHR ( $P = 1.65 \times 10^{-4}$ ) were significantly elevated, and dyslipidemia was demonstrated in the PCOS group. The HDL-C level was decreased, whereas, the LDL-C, TC, TG and ApoB levels were elevated in the PCOS group compared with the control. Simultaneously, the sex hormones (T, E2, LH), INS, GLU and inflammatory factors were significantly increased in the PCOS group ( $P < 0.05$ ), which prompted abnormalities in the *in vivo* endocrine metabolism in the patients with PCOS.

**Methodological examination of UPLC/Q-TOF-MS.** The UPLC/Q-TOF-MS plasma analysis was performed and examined for the methodology, wherein the QC samples were used

Table I. Clinical indexes in the control and PCOS groups.

No.	Index	Control group, n=50	PCOS group, n=49	P-value
1	T, ng/ml	0.44±0.14	0.63±0.19 <sup>b</sup>	1.00x10 <sup>-02</sup>
2	E2, pg/ml	35.21±10.84	63.88±51.53 <sup>b</sup>	3.67x10 <sup>-04</sup>
3	PRL, ng/ml	14.91±7.40	12.49±9.86	1.69x10 <sup>-01</sup>
4	LH, IU/l	5.64±5.95	13.57±8.36 <sup>b</sup>	5.23x10 <sup>-07</sup>
5	FSH, IU/l	6.83±2.03	6.20±1.55 <sup>a</sup>	3.86x10 <sup>-02</sup>
6	hsCRP, mg/l	0.80±2.11	2.13±2.70 <sup>b</sup>	8.09x10 <sup>-03</sup>
7	HDL-C, mmol/l	1.47±0.25	1.38±0.35 <sup>a</sup>	1.59x10 <sup>-02</sup>
8	LDL-C, mmol/l	2.28±0.46	2.71±0.81 <sup>b</sup>	2.16x10 <sup>-03</sup>
9	TC, mmol/l	4.03±0.52	4.48±0.87 <sup>b</sup>	2.38x10 <sup>-03</sup>
10	TG, mmol/l	0.76±0.32	1.12±0.94 <sup>a</sup>	1.48x10 <sup>-02</sup>
11	ApoA1, g/l	1.40±0.13	1.38±0.20	6.22x10 <sup>-01</sup>
12	ApoB, g/l	0.68±0.13	0.83±0.20 <sup>b</sup>	2.43x10 <sup>-05</sup>
13	INS, uIU/ml	8.91±9.16	18.99±25.75 <sup>a</sup>	1.22x10 <sup>-02</sup>
14	GLU, mmol/l	4.78±0.49	5.00±0.40 <sup>a</sup>	1.51x10 <sup>-02</sup>
15	HbA1c, %	5.17±0.48	5.16±0.26	9.55x10 <sup>-01</sup>
16	Waist:hip ratio	0.77±0.04	0.82±0.07 <sup>a</sup>	1.65x10 <sup>-04</sup>
17	Number of right follicles	3.51±3.15	9.69±2.61 <sup>b</sup>	7.70x10 <sup>-13</sup>
18	Number of left follicles	2.19±2.52	10.35±1.77 <sup>b</sup>	8.58x10 <sup>-21</sup>
19	BMI	20.40±1.72	24.38±5.39 <sup>b</sup>	7.21x10 <sup>-06</sup>

<sup>a</sup>P<0.05 vs. the control group; <sup>b</sup>P<0.01 vs. the control group. PCOS, polycystic ovarian syndrome; FSH, Follicle stimulating hormone; LH, luteinizing hormone; PRL, prolactin; E2, estradiol; T, testosterone; TC, total serum cholesterol; TG, triglycerides; HDL-C, high-density lipoprotein; LDL-C, low-density lipoprotein; ApoA1, apolipoprotein A1; ApoB, apolipoprotein B; hsCRP, high-sensitivity C-reactive protein; GLU, glucose; INS, insulin; HbA1c, glycosylated hemoglobin; BMI, body mass index.

to evaluate the stability and reproducibility of the method. A total of 10 peaks from the metabolic fingerprints were randomly selected from which, the ion peaks were extracted and analyzed by 10 injections. The results demonstrated that the relative standard deviation of the peak areas were <10%, which suggested that this method demonstrated adequate stability and reproducibility, and was able to meet the requirements of the study.

The 99 plasma samples of the control and PCOS groups were analyzed by UPLC/Q-TOF-MS using the preferred method, in order to construct the characteristic metabolic fingerprints. Fig. 1 demonstrated the base-peak total ion current (TIC) of the plasma metabolic fingerprints in the control and PCOS groups in the negative-ion mode; a marked difference between these two TIC plots was observed, which suggested significant metabolic abnormalities in patients with PCOS compared with the controls.

**Multivariate statistical analysis of UPLC/Q-TOF-MS data and pattern recognition.** In the present study, multivariate statistical analysis and pattern recognition methods, PLS-DA and OPLS-DA, were used to analyze the metabolomics data.

Fig. 2 demonstrated the score plot and response permutation test results obtained by the PLS-DA method. The model-associated parameter Q<sup>2</sup> was 0.791 (R<sup>2</sup>X=0.289, R<sup>2</sup>Y=0.833) and the response permutation test returned R<sup>2</sup>=0.426 and Q<sup>2</sup>=-0.0513, which suggested successful

modeling and absence of overfitting. The PLS-DA model may be used to differentiate the control and the PCOS groups. Notably, the PCOS group may be divided into two subgroups according to their metabolic status, subgroups 1 and 2. Furthermore, 2/50 controls were categorized as subgroup 1 (N49 and N50).

Fig. 3 demonstrates the OPLS-DA results of metabolomics data of patients with PCOS and controls; the Q<sup>2</sup> was 0.824 (R<sup>2</sup>X=0.375, R<sup>2</sup>Y=0.962). These OPLS-DA results were consistent with that of PLS-DA, demonstrating two subgroups with significant metabolic differences in the patients with PCOS. In addition, the two control samples, N49 and N50, deviated from the control group and approached the subgroup 1.

#### *Identification and screening of potential metabolic markers.*

As presented in the OPLS-DA load schematic, the variables distal from the origin served a crucial role in the classification of each group. According to the VIP values of markers in the model as well as S-plot analysis, samples were differentiated by combining the mass spectrometric isotopes corresponding to the accessible database. Ultimately, 34 potential metabolic markers serving a decisive role in sample classification, as well as alterations in the contents of each group, were identified (Table II). These may be used as potential biomarkers of PCOS-induced variations in the human metabolic state, including phospholipids, fatty acids, pyrimidines, sterols and other small molecular compounds.



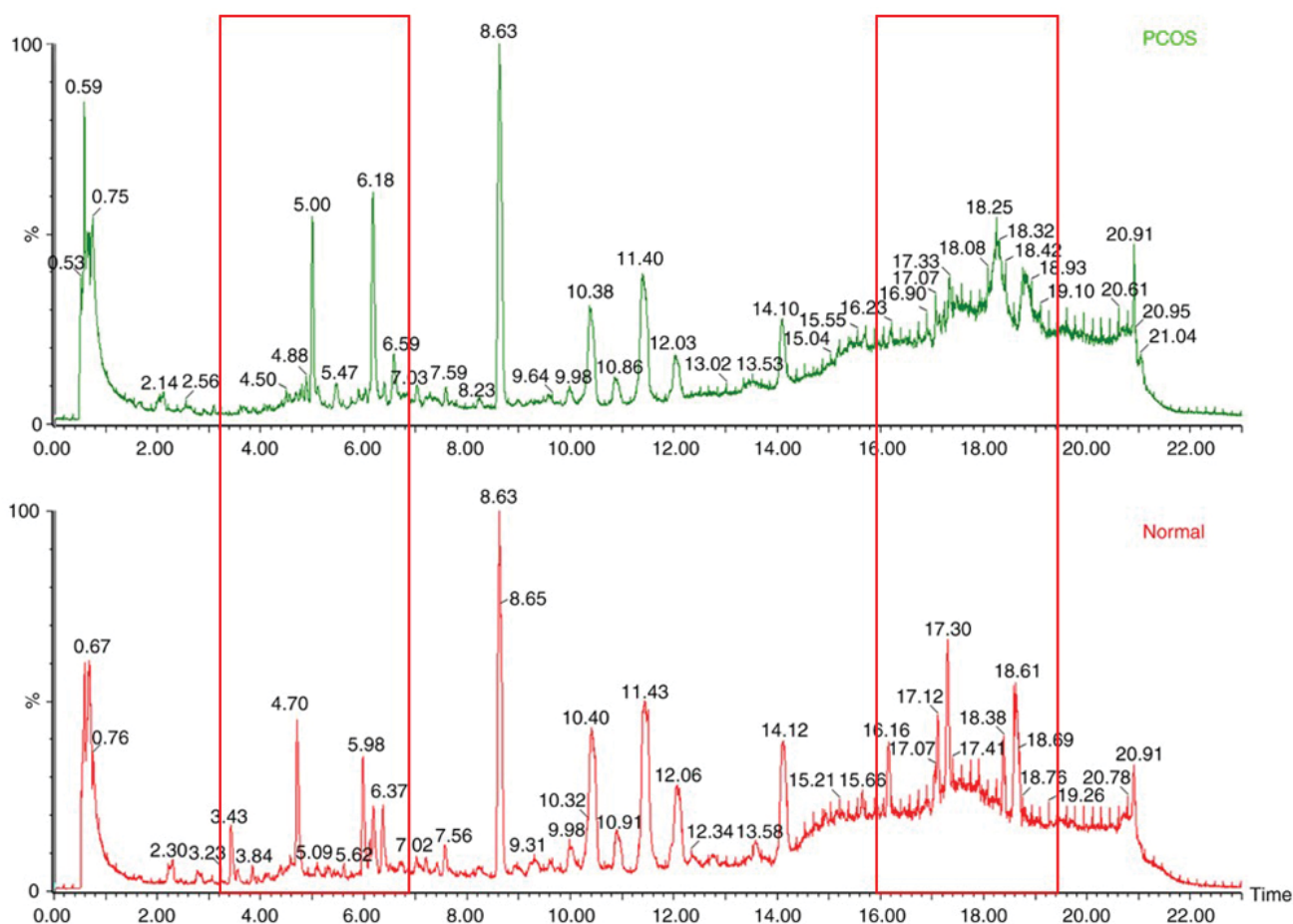


Figure 1. Base-peak ion current of plasma metabolic fingerprints of the control (left) and PCOS (right) groups. The red box indicates areas that are markedly different between the two total ion current plots. PCOS, polycystic ovarian syndrome.

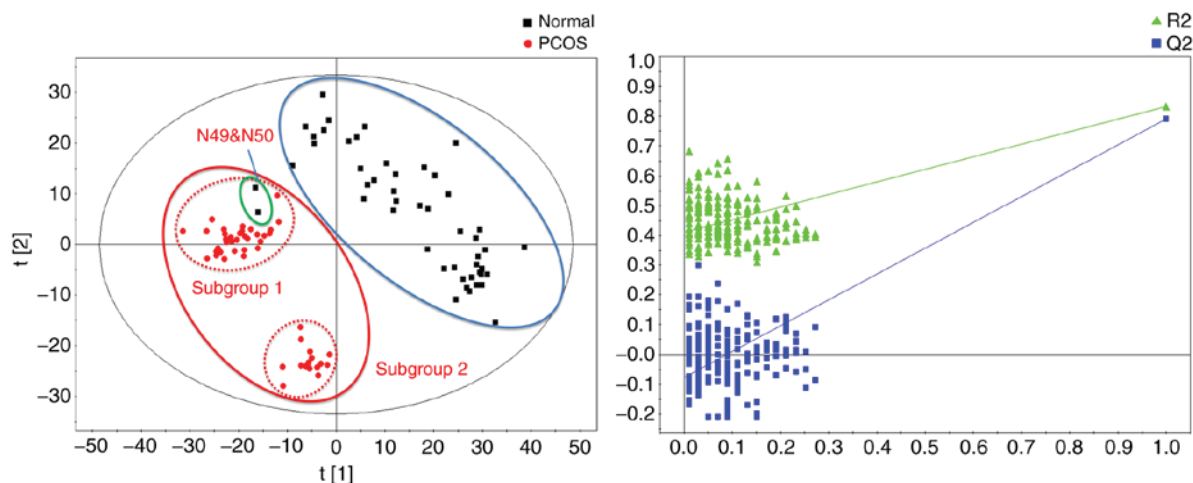


Figure 2. Partial least-squares discriminant analysis and response permutation test results of metabolomics data of patients with PCOS and controls. PCOS, polycystic ovarian syndrome.

Based on the PLS-DA and OPLS-DA results, the patients with PCOS may be divided into two metabolic subgroups, subgroup 1 and 2, according to the *in vivo* metabolic profile. Therefore, the alterations in the 34 potential metabolic markers in the two subgroups were analyzed. Table III demonstrated significant differences in the plasma levels of the 10 potential markers between the two subgroups, including

DG (18:1n9/0:0/20:4n3), SM (d18:1/16:0), PG (18:0/16:0), D-Glyceraldehyde 3-phosphate, androsterone sulfate and cysteine-S-sulfate ( $P < 0.05$ ).

*Focused analysis results of characteristic indexes.* A total of 19 differential clinical indexes and 34 potential metabolic markers were performed as focused analysis of the

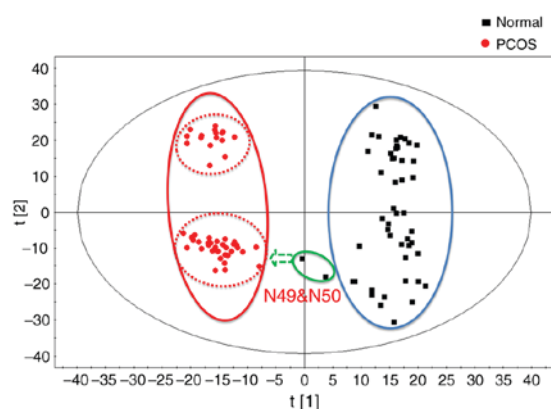


Figure 3. Orthogonal partial least-squares discriminant analysis results of metabolomics data of patients with PCOS and controls. PCOS, polycystic ovarian syndrome.

characteristic indexes. The NeuroFuzzy logic-based focused analysis results were presented in Table IV.

As presented in Table IV, the three groups of focus variables contributed >80% to the control group, and the two PCOS subgroups, particularly the predictive accuracy of the focused, integrated markers reached ≥95% or above, which was superior to the predictive accuracy of all the potential metabolic markers and clinical indexes. This suggests that each focus variable may characterize the information represented by all the variables. However, their potential usage in clinical disease diagnosis requires further evaluation by ROC analysis.

**ROC analysis of focused variables.** The five clinical indexes, three metabolic markers and five integrated markers obtained by ANN focused analysis were analyzed independently, and the following multifactor equations were constructed:

Focused clinical indexes:

$$Y = 1.431A + 3.56B + 15.138C + 6.616D + 0.752E - 58.902;$$

Where A is the BMI, B is the number of left follicles, C refers to HDL-C, D is LDL-C and E refers to hsCRP.

Focused metabolic markers:

$$Y = 0.011X_1 + 0.408X_2 + 0.017X_3 - 16.535;$$

Where  $X_1$ ,  $X_2$  and  $X_3$  refer to palmitoylphosphatidylcholine, cyclic guanosine monophosphate (cGMP) and dehydroepiandrosterone sulphate (DHEAS), respectively.

Focused integrated markers:

$$Y = 2.906B + 15.775C + 0.034X_1 + 0.402X_2 + 0.009X_3 - 42.579;$$

Where B is the number of left follicles, C is HDL-C, and  $X_1$ ,  $X_2$  and  $X_3$  refer to palmitoylphosphatidylcholine, cGMP and DHEAS, respectively.

The obtained dependent variables, Y, were analyzed using ROC curves to evaluate the accuracy, specificity and sensitivity for focused clinical indexes, focused metabolic markers and focused integrated markers in the diagnosis of PCOS, and classification of the two subgroups, as presented in Tables V-VII.

ROC analysis results demonstrated that for the control and PCOS groups, the diagnostic capacity of the indexes was in the order of as follows: Focused clinical indexes > focused

integrated markers > focused metabolic markers. However, in the subgroup 1 and 2 of PCOS, the diagnostic capacity of the indexes was ordered as focused metabolic markers > focused integrated markers > focused clinical indexes. The alterations in the plasma contents of the focused metabolic markers, including palmitoyl sphingomyelin, cGMP and androsterone sulfate were identified (Fig. 4). Comprehensively considering the diagnostic capacity of the focused indexes with respect to the control group, the PCOS group, and the two PCOS subgroups, the focused integrated markers demonstrated excellent performance, which was able to obtain a diagnostic accuracy >90% for each group (Table VII). These results suggested that the focused integrated markers may be considered as a marker system for the clinical diagnosis and classification of PCOS.

## Discussion

In the present study, the plasma metabolic fingerprints of the patients with PCOS (n=49) and normal controls (n=50) were established and demonstrated, to the best of our knowledge, for the first time that patients with PCOS may be divided into two groups, subgroup 1 and subgroup 2, according to their metabolic profiles. Therefore, focused analyses were performed on the characteristic markers of the PCOS group and the two subgroups and demonstrated an integrated marker system consisting of two clinical indexes and three metabolic markers that may be used for accurate diagnosis and classification of patients with PCOS.

For the establishment of a chromatographic fingerprint, liquid chromatography conditions including flow rate, injection volume and column temperature were optimized to minimize their effects on sample separation. Conversely, positive and negative ion modes, desolution gas flow, temperature and other factors were analyzed by mass spectrometry. Furthermore, analysis of the plasma samples provided additional information of the compounds employing the negative ion mode compared with the positive ion mode. Therefore, the negative ion mode was adopted for the establishment of the chromatographic fingerprint for the samples in the present study.

The multivariate statistical analysis and pattern recognition of plasma metabolic fingerprints of the patients with PCOS and controls demonstrated the occurrence of significant metabolic abnormalities in patients with PCOS compared with the controls. Additionally, two metabolic subtypes were demonstrated in the patients with PCOS, wherein 33/49 cases belonged to the PCOS subgroup 1 and 16 belonged to PCOS subgroup 2. These results suggested that the metabolic abnormalities were not consistent in the patients with PCOS, and there were at least two types of principal metabolic abnormalities. The analyses of the 34 identified potential metabolic markers identified that the alterations in the fat metabolism-associated markers were marked in subgroup 2 compared with subgroup 1.

Furthermore, PLS-DA and OPLS-DA of the metabolic data demonstrated that two samples (N49 and N50) of the control group exhibited metabolic abnormalities. Further analyses identified that although the clinical indexes, including blood glucose and glycosylated hemoglobin were normal in these two samples, the plasma contents of the 34 potential metabolic markers were already within the range in patients with PCOS (Table VIII). These results suggested a certain degree of metabolic disorders

Table II. Alterations in 34 potential metabolic markers in the control and PCOS groups.

No.	m/z	Common name	Change	Control group, n=50	PCOS group, n=49	Associated metabolic pathway
1	750.5448	PE [O-18:1(1Z)/20:4(5Z,8Z,11Z,14Z)]	↓	16.38±10.05	0.20±1.43 <sup>a</sup>	Glycerophospholipid metabolism
2	534.3507	LysoPE [0:0/22:1(13Z)]	↓	6.36±2.92	1.01±2.56 <sup>a</sup>	Glycerophospholipid metabolism
3	494.3535	LysoPC (16:0)	↓	5.24±2.83	0.72±2.16 <sup>a</sup>	Lecithin metabolism
4	746.5124	PE [O-16:1(1Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z)]	↓	73.58±51.44	11.12±12.02 <sup>a</sup>	Glycerophospholipid metabolism
5	592.3552	LysoPC [20:2(11Z,14Z)]	↓	5.54±2.63	0.44±1.78 <sup>a</sup>	Lecithin metabolism
6	766.5379	PE [22:4(7Z,10Z,13Z,16Z)/16:0]	↓	85.10±63.52	11.84±18.27 <sup>a</sup>	Glycerophospholipid metabolism
7	687.5446	DG [18:1n9/0:0/20:4n3]	↑	1.98±6.40	22.65±13.19 <sup>a</sup>	Diacylglycerol metabolism
8	856.6024	PC [16:1(9Z)/22:2(13Z,16Z)]	↓	26.67±16.46	1.36±7.03 <sup>a</sup>	Glycerophospholipid metabolism
9	748.5655	SM (d18:1/16:0)	↑	2.16±7.06	37.22±25.27 <sup>a</sup>	Sphingomyelin metabolism
10	749.5689	PG(18:0/16:0)	↑	0.52±2.22	10.86±7.78 <sup>a</sup>	Glycerophospholipid metabolism
11	215.0242	D-Glyceraldehyde 3-phosphate	↑	11.99±2.72	19.12±5.72 <sup>a</sup>	ATP metabolism
12	291.082	5,6-Dihydrouridine	↑	147.07±43.38	228.59±39.63 <sup>a</sup>	Pyrimidine metabolic cycle
13	369.1722	Androsterone sulfate	↑	13.28±35.74	113.90±116.75 <sup>a</sup>	Hormone metabolism
14	395.4668	25-Methyl-1-hexacosanol	↓	3.44±2.20	0.61±1.56 <sup>a</sup>	
15	417.2988	11'-Carboxy- $\alpha$ -chromanol	↑	13.00±4.31	20.88±8.06 <sup>a</sup>	
16	508.4052	LysoPC (O-18:0)	↓	7.20±1.52	2.10±3.90 <sup>a</sup>	Lecithin metabolism
17	431.3093	3,7-Dihydroxy-5-cholestenoic acid	↑	11.43±3.59	16.52±6.26 <sup>a</sup>	Bile acid metabolism
18	489.3537	3- $\beta$ -Hydroxy-4- $\beta$ -methyl-5- $\alpha$ -cholest-7-ene-4- $\alpha$ -carboxylate	↑	16.96±4.03	24.47±7.52 <sup>a</sup>	Bile acid metabolism
19	539.4314	(9-cis,9'-cis)-7',8,8'-Tetrahydro-y,y'-Carotene	↑	13.23±12.34	29.33±17.38 <sup>a</sup>	
20	747.5625	Palmitoylsphingomyelin	↑	4.22±14.24	75.06±51.95 <sup>a</sup>	Sphingomyelin metabolism
21	775.5975	PG (18:1(9Z)/18:0)	↓	29.18±13.37	4.19±7.09 <sup>a</sup>	Glycerophospholipid metabolism
22	855.5933	Thyroxine sulfate	↓	15.02±10.58	0.00±0.00 <sup>a</sup>	ATP metabolism
23	518.3935	Clupanodonylcarmitine	↑	9.44±2.60	14.58±5.06 <sup>a</sup>	Fat metabolism
24	640.2944	S-(PGI2)-glutathione	↑	8.35±4.27	13.19±4.45 <sup>a</sup>	Immune modulation
25	397.201	Galbanic acid	↑	3.79±7.54	19.7±16.84 <sup>a</sup>	
26	391.2837	Chenodeoxycholic acid	↑	14.41±8.43	24.98±11.45 <sup>a</sup>	Lipid metabolism, cholesterol metabolism
27	398.1299	S-Adenosylmethionine	↓	7.96±5.01	0.80±2.43 <sup>a</sup>	Thiol amino acid metabolic cycle
28	589.3353	2-Arachidonoylglycerophosphocholine	↑	12.40±2.54	16.43±3.97 <sup>a</sup>	Glycerophospholipid metabolism
29	453.2927	Cholic acid	↓	9.92±3.10	3.69±4.61 <sup>a</sup>	Lipid metabolism, cholesterol metabolism
30	647.4328	Oryzanol A	↓	12.21±6.88	2.22±4.36 <sup>a</sup>	Endocrine modulation
31	343.994	Cyclic GMP	↑	29.93±16.10	54.94±19.23 <sup>a</sup>	Purine metabolism
32	747.5151	PG [18:1(9Z)/16:0]	↓	38.51±26.40	3.87±7.80 <sup>a</sup>	Glycerophospholipid metabolism
33	441.3515	4a-Methylzymosterol-4-carboxylic acid	↑	10.37±3.30	17.74±8.13 <sup>a</sup>	Citric acid cycle
34	199.9709	Cysteine-S-sulfate	↑	14.71±9.85	30.10±13.28 <sup>a</sup>	

<sup>a</sup>P<0.05 vs. the control group; <sup>b</sup>P<0.05 vs. the subgroup 1. PCOS, polycystic ovarian syndrome; GMP, guanosine monophosphate.

Table III. Alterations in 34 potential metabolic markers in the control and PCOS subgroups.

No.	Common name	Change	Control group, n=50	PCOS subgroup 1, n=33	PCOS subgroup 2, n=16
1	PE [O-18:1(1Z)/20:4(5Z,8Z,11Z,14Z)]	↓	16.38±10.05	0.30±1.75 <sup>a</sup>	0.00±0.00 <sup>a</sup>
2	LysoPE [0:0/22:1(13Z)]	↓	6.36±2.92	0.76±2.12 <sup>a</sup>	1.51±3.32 <sup>a</sup>
3	LysoPC (16:0)	↓	5.24±2.83	0.62±1.99 <sup>a</sup>	0.93±2.54 <sup>a</sup>
4	PE [O-16:1(1Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z)]	↓	73.58±51.44	10.99±14.27 <sup>a</sup>	11.39±5.29 <sup>a</sup>
5	LysoPC [20:2(11Z,14Z)]	↓	5.54±2.63	0.47±1.89 <sup>a</sup>	0.39±1.58 <sup>a</sup>
6	PE [22:4(7Z,10Z,13Z,16Z)/16:0]	↓	85.10±63.52	11.78±21.88 <sup>a</sup>	11.97±6.80 <sup>a</sup>
7	DG (18:1n9/0:0/20:4n3)	↑	1.98±6.40	19.49±12.55 <sup>a</sup>	29.17±12.37 <sup>a,b</sup>
8	PC [16:1(9Z)/22:2(13Z,16Z)]	↓	26.67±16.46	2.03±8.53 <sup>a</sup>	0.00±0.00 <sup>a</sup>
9	SM (d18:1/16:0)	↑	2.16±7.06	29.55±24.16 <sup>a</sup>	53.03±20.06 <sup>a,b</sup>
10	PG (18:0/16:0)	↑	0.52±2.22	8.89±7.47 <sup>a</sup>	14.94±6.95 <sup>a,b</sup>
11	D-Glyceraldehyde-3-phosphate	↑	11.99±2.72	17.62±5.95 <sup>a</sup>	22.22±3.74 <sup>a,b</sup>
12	5,6-Dihydrouridine	↑	147.07±43.38	225.47±34.05 <sup>a</sup>	235.02±49.87 <sup>a</sup>
13	Androsterone sulfate	↑	13.28±35.74	161.48±115.02 <sup>a</sup>	15.76±11.79 <sup>b</sup>
14	25-Methyl-1-hexacosanol	↓	3.44±2.20	0.56±1.35 <sup>a</sup>	0.70±1.97 <sup>a</sup>
15	11'-Carboxy-a-chromanol	↑	13.00±4.31	19.94±7.18 <sup>a</sup>	22.82±9.58 <sup>a</sup>
16	LysoPC (O-18:0)	↓	7.20±1.52	2.63±4.28 <sup>a</sup>	1.01±2.78 <sup>a</sup>
17	3,7-Dihydroxy-5-cholestenoic acid	↑	11.43±3.59	16.05±5.90 <sup>a</sup>	17.49±7.05 <sup>a</sup>
18	3-b-Hydroxy-4-b-methyl-5-a-cholest-7-ene-4-a-carboxylate	↑	16.96±4.03	23.44±6.85 <sup>a</sup>	26.59±8.58 <sup>a</sup>
19	(9-cis,9'-cis)-7,7',8,8'-Tetrahydro-γ,γ-Carotene	↑	13.23±12.34	29.52±18.84 <sup>a</sup>	28.95±14.46 <sup>a</sup>
20	Palmitoylsphingomyelin	↑	4.22±14.24	58.99±49.95 <sup>a</sup>	108.20±39.64 <sup>a,b</sup>
21	PG [18:1(9Z)/18:0]	↓	29.18±13.37	4.29±7.08 <sup>a</sup>	3.97±7.33 <sup>a</sup>
22	Thyroxine sulfate	↓	15.02±10.58	0.00±0.00 <sup>a</sup>	0.00±0.00 <sup>a</sup>
23	Clupanodonylcarnitine	↑	9.44±2.60	13.52±4.74 <sup>a</sup>	16.76±5.15 <sup>a</sup>
24	S-(PGJ2)-glutathione	↑	8.35±4.27	15.40±2.79 <sup>a</sup>	8.64±3.75 <sup>b</sup>
25	Galbanic acid	↑	3.79±7.54	28.25±13.71 <sup>a</sup>	2.07±3.84 <sup>b</sup>
26	Chenodeoxycholic acid	↑	14.41±8.43	24.21±11.81 <sup>a</sup>	26.58±10.87 <sup>a</sup>
27	S-Adenosylmethionine	↓	7.96±5.01	0.68±2.22 <sup>a</sup>	1.04±2.87 <sup>a</sup>
28	2-Arachidonoylglycerophosphocholine	↑	12.40±2.54	16.32±3.88 <sup>a</sup>	16.65±4.27 <sup>a</sup>
29	Cholic acid	↓	9.92±3.10	3.72±4.40 <sup>a</sup>	3.63±5.16 <sup>a</sup>
30	Oryzanol A	↓	12.21±6.88	1.76±3.55 <sup>a</sup>	3.15±5.69 <sup>a</sup>
31	Cyclic GMP	↑	29.93±16.1	67.37±7.42 <sup>a</sup>	29.3±4.98 <sup>b</sup>
32	PG [18:1(9Z)/16:0]	↓	38.51±26.40	4.62±8.75 <sup>a</sup>	2.31±5.28 <sup>a</sup>
33	4a-Methylzymosterol-4-carboxylic acid	↑	10.37±3.30	18.32±7.98 <sup>a</sup>	16.54±8.57 <sup>a</sup>
34	Cysteine-S-sulfate	↑	14.71±9.85	25.82±11.97 <sup>a</sup>	38.92±11.63 <sup>a,b</sup>

<sup>a</sup>P<0.05 vs. the control group; <sup>b</sup>P<0.05 vs. the subgroup 1. PCOS, polycystic ovarian syndrome.Table IV. NeuroFuzzy logic-based focused analysis results and prediction accuracy (R<sup>2</sup>).

Type	Focus variable	R <sup>2</sup>
Focused clinical indexes	BMI, number of left follicles, HDL-C, LDL-C, hsCRP	0.8284
Focused metabolic markers	Palmitoyl sphingomyelin, cGMP, DHEAS	0.9027
Focused integrated markers	Number of left follicles, HDL-C, Palmitoyl sphingomyelin, cGMP, DHEAS	0.9574
All clinical indexes	19 clinical indexes	0.8256
All metabolic markers	34 potential metabolic markers	0.9392
All markers	34 potential metabolic markers and 18 clinical indexes	0.9377

BMI, body mass index; cGMP, cyclic guanosine monophosphate; HDL-C, high density lipoprotein; DHEAS, dehydroepiandrosterone sulphate; LDL-C, low density lipoprotein; hsCRP, high sensitivity C-reactive protein.



Table V. Receiver-operating characteristic analysis results of focused clinical indexes.

Type	Control vs. PCOS	Control vs. subgroup 1	Control vs. subgroup 2	Subgroup 1 vs. subgroup 2
AUC	0.994	0.993	0.995	0.621
Sensitivity, %	98.0	100.0	100.0	62.5
Specificity, %	96.0	94.0	98.0	72.7
95% CI	0.983-1.000	0.982-1.000	0.984-1.000	0.447-0.795

CI, confidence interval; AUC, area under the curve; PCOS, polycystic ovarian syndrome.

Table VI. Receiver-operating characteristic analysis results of focused metabolic markers.

Type	Control vs. PCOS	Control vs. subgroup 1	Control vs. subgroup 2	Subgroup 1 vs. subgroup 2
AUC	0.876	0.958	0.708	1.000
Sensitivity, %	95.9	93.9	93.8	100.0
Specificity, %	68.0	94.0	64.0	100.0
95% CI	0.807-0.945	0.910-1.000	0.586-0.829	1.000-1.000

CI, confidence interval; AUC, area under the curve; PCOS, polycystic ovarian syndrome.

Table VII. Receiver-operating characteristic analysis results of focused, integrated markers.

Type	Control vs. PCOS	Control vs. subgroup 1	Control vs. subgroup 2	Subgroup 1 vs. subgroup 2
AUC	0.986	0.998	0.963	0.928
Sensitivity, %	100.0	96.7	100.0	90.9
Specificity, %	86.0	100.0	86.2	87.5
95% CI	0.970-1.000	0.992-1.000	0.921-1.000	0.851-1.000

CI, confidence interval; AUC, area under the curve; PCOS, polycystic ovarian syndrome.

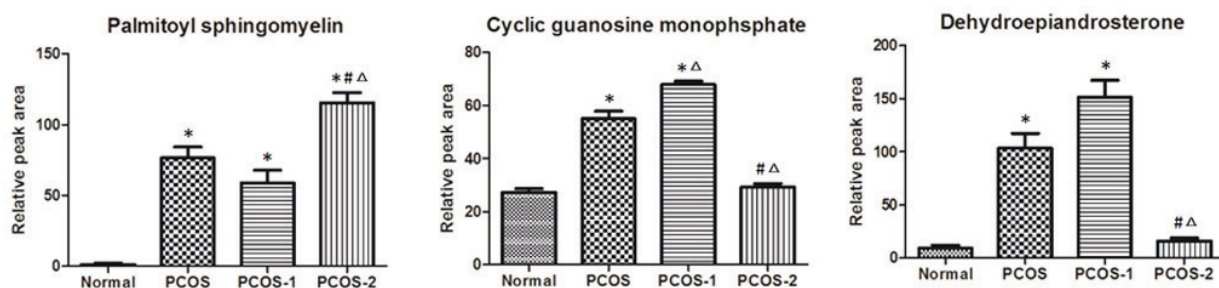


Figure 4. Alterations of differential focused metabolic markers. PCOS-1, the subgroup 1; PCOS-2, the subgroup 2. \*P<0.05 vs. the control group, #P<0.05 vs. PCOS-1, ΔP<0.05 vs. the PCOS group. PCOS, polycystic ovarian syndrome.

in these two controls, which may be attributed to the early stage of PCOS, wherein the alterations of clinical indexes are not as marked yet. Therefore, an accurate diagnosis is unlikely to be achieved relying only on the clinical indexes. This additionally suggests that alterations in metabolites are more sensitive when compared with the relatively stable clinical indexes.

Focused analyses of the characteristic indexes demonstrated that the focused clinical indices, including BMI, the

number of left follicles, HDL-C, LDL-C and hsCRP may be used for accurate diagnosis of patients with PCOS. These parameters demonstrated marked abnormalities in the PCOS group, suggesting that fat metabolism, hormone metabolism, inflammatory response and other *in vivo* processes were involved in PCOS. Lipid metabolism is associated with PCOS, and the plasma HDL level is decreased in patients with PCOS with INS resistance (18). Yilmaz *et al* (19) demonstrated

Table VIII. Contents of 34 potential metabolic markers in the controls N49 and N50.

No.	m/z	Control group	N49	N50	PCOS group
1	750.5448	17.06±9.66	0.00 <sup>a</sup>	0.00 <sup>a</sup>	0.00±0.00
2	534.3507	6.44±2.81	0.00 <sup>a</sup>	8.97 <sup>b</sup>	1.03±2.58
3	494.3535	5.29±2.76	0.00 <sup>a</sup>	8.11 <sup>b</sup>	0.60±2.03
4	746.5124	76.14±50.88	18.05 <sup>a</sup>	5.87 <sup>a</sup>	10.73±11.83
5	592.3552	5.63±2.56	0.00 <sup>a</sup>	6.38 <sup>b</sup>	0.45±1.79
6	766.5379	88.40±62.68	5.88 <sup>a</sup>	5.76 <sup>a</sup>	9.51±8.18
7	687.5446	0.96±3.59	35.24 <sup>a</sup>	17.64 <sup>a</sup>	23.12±12.90
8	856.6024	27.78±15.83	0.00 <sup>a</sup>	0.00 <sup>a</sup>	1.17±6.97
9	748.5655	0.87±3.09	35.31 <sup>a</sup>	30.68 <sup>a</sup>	37.99±24.94
10	749.5689	0.09±0.63	11.68 <sup>a</sup>	9.93 <sup>a</sup>	11.09±7.69
11	215.0242	11.82±2.65	15.26	16.55	19.19±5.76
12	291.082	144.07±41.61	222.40 <sup>a</sup>	215.62 <sup>a</sup>	229.88±38.99
13	369.1722	7.65±18.18	215.30 <sup>a</sup>	81.87 <sup>a</sup>	115.79±117.22
14	395.4668	3.59±2.12	0.00 <sup>a</sup>	0.00 <sup>a</sup>	0.62±1.57
15	417.2988	12.88±4.3	14.82	16.90	20.83±8.14
16	508.4052	7.36±1.13	0.00 <sup>a</sup>	6.84 <sup>b</sup>	1.95±3.79
17	431.3093	11.41±3.65	13.15	10.54	16.54±6.33
18	489.3537	16.76±3.99	21.08	22.02	24.6±7.53
19	539.4314	12.54±12.11	29.638 <sup>a</sup>	30.0624 <sup>a</sup>	29.66±17.41
20	747.5625	1.45±3.79	76.03 <sup>a</sup>	65.34 <sup>a</sup>	76.37±51.66
21	775.5975	29.99±13.01	12.61 <sup>a</sup>	6.88 <sup>a</sup>	3.72±6.36
22	855.5933	15.64±10.33	0.00 <sup>a</sup>	0.00 <sup>a</sup>	0.00±0.00
23	518.3935	9.14±2.16	15.37	17.97	14.64±5.10
24	640.2944	8.08±4.14	14.99	14.52	13.05±4.37
25	397.201	2.92±6.28	28.09 <sup>a</sup>	21.36 <sup>a</sup>	19.89±16.96
26	391.2837	13.82±8.07	28.79 <sup>a</sup>	28.49 <sup>a</sup>	25.18±11.48
27	398.1299	8.15±4.97	0.00 <sup>a</sup>	6.70 <sup>b</sup>	0.68±2.29
28	589.3353	12.46±2.57	10.54	11.16	16.43±4.00
29	453.2927	10.17±2.79	8.11 <sup>b</sup>	0.00 <sup>a</sup>	3.68±4.66
30	647.4328	12.72±6.54	0.00 <sup>a</sup>	0.00 <sup>a</sup>	2.26±4.39
31	343.994	27.92±12.92	81.11 <sup>a</sup>	75.18 <sup>a</sup>	55.06±19.42
32	747.5151	39.93±25.97	9.14 <sup>a</sup>	0.00 <sup>a</sup>	3.61±7.68
33	441.3515	10.25±3.26	16.31 <sup>a</sup>	10.06 <sup>b</sup>	17.79±8.21
34	199.9709	14.11±9.59	29.91 <sup>a</sup>	28.16 <sup>a</sup>	30.29±13.34

<sup>a</sup>Within the diseased range; <sup>b</sup>Within the normal range.

that the LDL level was increased in obese patients with PCOS compared with the non-obese patients with PCOS. This phenomenon may be attributed to the weakened inhibitory effect of INS resistance on the free fatty acids, which leads to the increased levels of free fatty acids in the liver, thereby resulting in increased LDL synthesis. Additionally, an abnormal increase in the number of follicles is an obvious manifestation of patients with a polycysticovary and a critical representative of disruption in hormone metabolism (20). Previous studies (21,22) observed that the INS resistance in patients with PCOS was associated with chronic inflammation; the hsCRP level in the PCOS patients was significantly increased compared with the normal individuals and positively correlated to BMI, WHR and LDL (23). The results of

the present study demonstrated that the BMI and number of left follicles were notably increased, and the plasma LDL-C and hsCRP levels were elevated in patients with PCOS. This suggested the occurrence of INS resistance, abnormal lipid metabolism and inflammatory response in patients with PCOS in addition to abnormal hormone metabolism. Furthermore, the analyses of the PCOS subgroups 1 and 2 (Table IX) demonstrated that patients of the two subgroups exhibited metabolic hormone abnormalities. Furthermore, the abnormal lipid metabolism was prominent in the subgroup 2 compared with subgroup 1, whereas, BMI, HDL-C and LDL-C were elevated in comparison with the control group. Therefore, it may be hypothesized that in addition to the hormone metabolism disorders, patients with PCOS in the subgroup 2

Table IX. Analysis of clinical, biological indexes in the control and PCOS subgroups.

No.	Clinical index	Control group, n=50	PCOS subgroup 1, n=33	PCOS subgroup 2, n=16
1	T, ng/ml	0.44±0.14	0.79±0.99	0.69±0.25 <sup>a</sup>
2	E2, pg/ml	35.21±10.85	58.49±56.53 <sup>a</sup>	74.99±38.53 <sup>a</sup>
3	PRL, ng/ml	14.92±7.41	13.81±11.66	9.75±3.05 <sup>a</sup>
4	LH, IU/l	5.64±5.95	14.02±9.37 <sup>a</sup>	12.63±5.93 <sup>a</sup>
5	FSH, IU/l	6.83±2.03	6.42±1.62	5.74±1.33 <sup>a</sup>
6	hsCRP, mg/l	0.81±2.11	1.41±2.06	3.62±3.29 <sup>a,b</sup>
7	HDL-C, mmol/l	1.47±0.25	1.44±0.33	1.25±0.37 <sup>a</sup>
8	LDL-C, mmol/l	2.29±0.46	2.53±0.79	3.06±0.74 <sup>a,b</sup>
9	TC, mmol/l	4.03±0.52	4.42±0.83 <sup>a</sup>	4.60±0.95 <sup>a</sup>
10	TG, mmol/l	0.76±0.32	1.15±1.06 <sup>a</sup>	1.05±0.63
11	ApoA1, g/l	1.40±0.13	1.42±0.18	1.29±0.21 <sup>b</sup>
12	ApoB, g/l	0.68±0.13	0.79±0.19 <sup>a</sup>	0.91±0.18 <sup>a,b</sup>
13	INS, uIU/ml	8.91±9.16	14.72±11.43 <sup>a</sup>	27.80±41.47
14	GLU, mmol/l	4.78±0.49	5.02±0.40 <sup>a</sup>	4.98±0.41
15	HbA1c, %	5.17±0.48	5.12±0.29	5.25±0.20
16	Waist-hip ratio	0.77±0.04	0.82±0.07 <sup>a</sup>	0.81±0.07 <sup>a</sup>
17	Number of right follicles	3.51±3.14	9.19±3.05 <sup>a</sup>	10.59±1.13 <sup>a,b</sup>
18	Number of left follicles	2.19±2.52	10.23±2.06 <sup>a</sup>	10.56±1.11 <sup>a</sup>
19	BMI	20.40±1.72	23.97±5.28 <sup>a</sup>	25.23±5.68 <sup>a</sup>

<sup>a</sup>P<0.05 vs. the control group; <sup>b</sup>P<0.05 vs. the subgroup 1. PCOS, polycystic ovarian syndrome; FSH, Follicle stimulating hormone; LH, luteinizing hormone; PRL, prolactin; E2, estradiol; T, testosterone; TC, total serum cholesterol; TG, triglycerides; HDL-C, high-density lipoprotein; LDL-C, low-density lipoprotein; ApoA1, apolipoprotein A1; ApoB, apolipoprotein B; hsCRP, high-sensitivity C-reactive protein; GLU, glucose; INS, insulin; HbA1c, glycosylated hemoglobin; BMI, body mass index.

additionally demonstrated abnormalities in INS resistance and lipid metabolism.

In comparison with the control group, the plasma contents of the focused metabolites were increased in the PCOS group, which suggested that the alterations in the content were associated with PCOS. Furthermore, comparing the alterations of the three metabolites between the two subgroups identified that palmitoyl sphingomyelin was elevated in the two subgroups, and the plasma level in the subgroup 2 was increased compared with in subgroup 1. Additionally, the plasma contents of the hormone metabolism-associated cGMP and DHEAS were increased in the subgroup 1. These results suggested that PCOS involves hormone and lipid metabolism disorders, and the two subgroups have different metabolic phenotypes, which are consistent with the results from the clinical indices.

Palmitoyl sphingomyelin is a lipid with sphingosine as the skeleton. A recent study (24) on PCOS metabolism demonstrated that PCOS was associated with the decreased oxidation capacity of fatty acids, thereby leading to accumulation of excessive fatty acids in patients with PCOS in the form of triglycerides. This results in significantly increased octadecenoic acid, palmitoleic acid and other fatty acids. The present study demonstrated that the content of phospholipid metabolites was increased in patients with PCOS, and the increasing trend of the PCOS subgroup 2 was increased compared with that of the PCOS subgroup 1, which indicated abnormal lipid metabolism in the patients of subgroup 2.

cGMP, which serves the role of acetylcholine, is widely distributed in various tissues. Previous studies (25,26) demonstrated that the cGMP is positively correlated to DHEAS. The present study additionally demonstrated that the cGMP level in the PCOS subgroup 1 was elevated compared with the control group. However, the difference between the PCOS subgroup 2 and the control group was not significant, which was in agreement with the results of the DHEAS content.

DHEAS is a sex hormone precursor. Previously, specific studies demonstrated that the total T, free T and DHEAS levels in patients with PCOS were increased by 50-75% compared with the healthy population (27,28). The increased DHEAS may induce follicular atresia and cystic expansion, promote the synthesis of T, elevate serum T levels, as well as improve polycystic ovaries and hyperandrogenism. Metabolomics analysis demonstrated that the plasma DHEAS level in the PCOS subgroup 2 was similar to that in the control group, whereas the plasma DHEAS level in the PCOS subgroup was elevated compared with the control group. Concurrently, the analysis of the content of the clinical, biochemical index T additionally demonstrated that the T level in the PCOS subgroup 1 was greater compared with the subgroup 2, thereby designating an abnormal hormone level in the patients of PCOS subgroup 1.

Combining the results of the focused clinical indexes and metabolic markers, it was demonstrated that the lipid metabolism and hormone metabolism abnormalities were

most notable in patients with PCOS. In addition, the hormone metabolism abnormality was dominant in the subgroup 1, whereas, the hormone and lipid metabolism disorders occurred in the subgroup 2.

ANN and ROC analyses demonstrated that the focused clinical indexes may be used to accurately differentiate the controls and patients with PCOS, whereas, the focused metabolic markers may be used to accurately distinguish the two PCOS subgroups. Furthermore, the screened focused integrated marker system (including two clinical indexes and three metabolites) may be used for the diagnosis of PCOS patients, as well as the classification of the two PCOS subgroups with a predictive diagnosis rate >90% between the various groups. These results indicated that although the clinical indexes and metabolic markers exhibited advantages, the focused, integrated marker system combines their advantages and may accurately diagnose the occurrence and classification of the disease.

PCOS is a heterogeneous disease with various clinical manifestations. It is a syndrome exhibited by the combined action of multiple factors, and its potential pathogenesis is excessive androgen secretion and lipid metabolism abnormality. The present study not only confirmed that the hormone metabolism disorders and lipid metabolism abnormalities were involved in the occurrence of PCOS; however, additionally demonstrated two metabolic subtypes in the patients with PCOS for the first time to the best of our knowledge. A focused analysis of the clinical indexes and potential metabolic biomarkers to establish a focused, integrated marker system was additionally conducted, which may be used for accurate diagnosis and classification of PCOS. Therefore, the present study provides a novel index system for the clinical diagnosis of PCOS, in addition to a scientific basis for individualized diagnosis and treatment of patients with PCOS.

### Acknowledgements

Not applicable.

### Funding

The present study was supported by the National Natural Science Foundation of China (grant no. 81302731) and the Major National Science and Technology Project (grant no. 2017ZX09301032).

### Availability of data and materials

The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

### Authors' contributions

XF and JJ wrote the manuscript and analyzed the data. ZH, JG, WX, YD, YaW and TZ collected the clinical samples and analyzed the data. YiW and AS designed the study, and reviewed and edited the manuscript. GL contributed to the interpretation of the metabolomics data, and reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

### Ethics approval and consent to participate

The present study was approved by the Ethics Committee of Peking Union Medical College Hospital (approval no. S-733; Beijing, China). Written informed consent was obtained from all participants.

### Patient consent for publication

Written informed consent was obtained from all participants.

### Competing interests

The authors declare they have no competing interests.

### References

1. Glintborg D and Andersen M: An update on the pathogenesis, inflammation, and metabolism in hirsutism and polycystic ovary syndrome. *Gynecol Endocrinol* 26: 281-296, 2010.
2. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE, *et al*: Positions statement: Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: An androgen excess society guideline. *J Clin Endocrinol Metab* 91: 4237-4245, 2006.
3. Knebel B, Janssen OE, Hahn S, Jacob S, Gleich J, Kotzka J and Muller-Wieland D: Increased low grade inflammatory serum markers in patients with polycystic ovary syndrome (PCOS) and their relationship to PPARgamma gene variants. *Exp Clin Endocrinol Diabet* 116: 481-486, 2008.
4. Lerchbaum E, Schwetz V, Giuliani A and Obermayer-Pietsch B: Influence of a positive family history of both type 2 diabetes and PCOS on metabolic and endocrine parameters in a large cohort of PCOS women. *Eur J Endocrinol* 170: 727-739, 2014.
5. Zhao S, Tian Y, Gao X, Zhang X, Liu H, You L, Cao Y, Su S, Chan WY, Sun Y, *et al*: Family-based analysis of eight susceptibility loci in polycystic ovary syndrome. *Sci Rep* 5: 12619, 2015.
6. Kume T, Calan M, Yilmaz O, Kocabas GU, Yesil P, Temur M, Bicer M and Calan OG: A possible connection between tumor necrosis factor alpha and adipon levels in polycystic ovary syndrome. *J Endocrinol Invest* 39: 747-754, 2016.
7. Chen YX, Zhang XZ, Huang J, Zhou SJ, Liu F, Jiang LL, Chen M, Wan JB and Yang DZ: UHPLC/Q-TOFMS-based plasma metabolomics of polycystic ovary syndrome patients with and without insulin resistance. *J Pharm Biomed Anal* 121: 141-150, 2016.
8. Zhao Y, Fu L, Li R, Wang LN, Yang Y, Liu NN, Zhang CM, Wang Y, Liu P, Tu BB, *et al*: Metabolic profiles characterizing different phenotypes of polycystic ovary syndrome: Plasma metabolomics analysis. *BMC Med* 10: 153, 2012.
9. Rautio K, Tapanainen JS, Ruokonen A and Morin-Papunen LC: Rosiglitazone treatment alleviates inflammation and improves liver function in overweight women with polycystic ovary syndrome: A randomized placebo controlled study. *Fertil Steril* 87: 202-226, 2007.
10. Baskind NE and Balen AH: Hypothalamic-pituitary, ovarian and adrenal contributions to polycystic ovary syndrome. *Best Pract Res Clin Obstet Gynaecol* 37: 80-97, 2016.
11. Lian Y, Zhao F and Wang W: Central leptin resistance and hypothalamic inflammation are involved in letrozole-induced polycystic ovary syndrome rats. *Biochem Biophys Res Commun* 476: 306-312, 2016.
12. Cheng F, Zhao L, Wu Y, Huang T, Yang G, Zhang Z, Wu Y, Jia F, Wu J, Chen C and Liu D: Serum vascular endothelial growth factor B is elevated in women with polycystic ovary syndrome and can be decreased with metformin treatment. *Clin Endocrinol (Oxf)* 84: 386-393, 2016.
13. Atiomo W and Daykin CA: Metabolomic biomarkers in women with polycystic ovary syndrome: A pilot study. *Mol Hum Reprod* 18: 546-553, 2012.
14. Sun LY, Hu WH, Liu Q, Hao Q, Sun B, Zhang Q, Mao S, Qiao J and Yan X: Metabonomics reveals plasma metabolic changes and inflammatory marker in polycystic ovary syndrome patients. *J Proteome Res* 11: 2937-2946, 2012.



15. Zhao XJ, Xu F, Qi B, Hao S, Li Y, Li Y, Zou L, Lu C, Xu G and Hou L: Serum metabolomics study of polycystic ovary syndrome based on liquid chromatography-mass spectrometry. *J Proteome Res* 13: 1101-1111, 2014.
16. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group: Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 81: 19-25, 2004.
17. Lujan ME, Chizen DR and Pierson RA: Diagnostic criteria for polycystic ovary syndrome: Pitfalls and controversies. *J Obstet Gynaecol Can* 30: 671-679, 2008.
18. Xue YJ, Xie H, Sun ML *et al*: Characteristics of abnormal lipid metabolism in PCOS patients with insulin resistance. *Chin J Mod Med* 24: 94-98, 2014 (In Chinese).
19. Yilmaz M, Biri A, Bukan N, Karakoç A, Sancak B, Törüner F and Paşaoğlu H: Levels of lipoprotein and homocysteine in non-obese and obese patients with polycystic ovary syndrome. *Gynecol Endocrinol* 20: 258-263, 2005.
20. Padmapriya B and Kesavamurthy T: Detection of follicles in polycystic ovarian syndrome in ultrasound images using morphological operations. *J Med Imaging Health Inform* 6: 240-243, 2016.
21. Walch K, Grimm C, Zeillinger R, Huber JC, Nagele F and Heffler LA: A common interleukin-6 gene promoter polymorphism influences the clinical characteristics of women with polycystic ovary syndrome. *Fertil Steril* 81: 1638-1641, 2004.
22. Vgontzas AN, Trakada G, Bixler EO, Lin HM, Pejovic S, Zoumakis E, Chrousos GP and Legro RS: Plasma interleukin 6 levels are elevated in polycystic ovary syndrome independently of obesity or sleep apnea. *Metabolism* 55: 1076-1082, 2006.
23. Verit FF: High sensitive serum C-reactive protein and its relation-ship with other cardiovascular risk factors in normoinsulinemic polycystic ovary patients without metabolic syndrome. *Arch Gynecol Obstet* 281: 1009-1014, 2010.
24. Dong F, Deng D, Chen H, Cheng W, Li Q, Luo R and Ding S: Serum metabolomics study of polycystic ovary syndrome based on UPLC-QTOF-MS coupled with a pattern recognition approach. *Anal Bioanal Chem* 407: 4683-4695, 2015.
25. Martina V, Benso A, Gigliardi VR, Masha A, Origlia C, Granata R and Ghigo E: Short-term dehydroepiandrosterone treatment increases platelet cGMP production in elderly male subjects. *Clin Endocrinol (Oxf)* 64: 260-264, 2006.
26. Martina V, Origlia C, Bruno GA, Messina M, Ferri M and Pescarmona GP: Serum DHEAS levels correlate with platelet cGMP in normal women. *J Endocrinol Invest* 24: RC28-RC30, 2001.
27. Holte J, Bergh T, Gennarelli G and Wide L: The independent effects of polycystic ovary syndrome and obesity on serum concentrations of gonadotrophins and sex steroids in premenopausal women. *Clin Endocrinol (Oxf)* 41: 473-481, 1994.
28. Huang A, Brennan K and Azziz R: Prevalence of hyperandrogenemia in the polycystic ovary syndrome diagnosed by the National Institutes of Health 1990 criteria. *Fertil Steril* 93: 1938-1941, 2010.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.