Steroids 150 (2019) 108432

Contents lists available at ScienceDirect

Steroids

journal homepage: www.elsevier.com/locate/steroids

Individual and cyclic estrogenic profile in women: Structure and variability of the data

Cristina Bozzolino[a], Sara Vaglio[a], Eleonora Amante[a,b], Eugenio Alladio[a,b,*], Enrico Gerace[b], Alberto Salomone[a,b], Marco Vincenti[a,b]

[a] Dipartimento di Chimica, Università degli Studi di Torino, via P. Giuria 7, 10125 Torino, Italy
[b] Centro Regionale Antidoping e di Tossicologia "A. Bertinaria", regione Gonzole 10/1, 10043 Orbassano, TO, Italy

ARTICLE INFO

Keywords:
Estrogens
Menstrual cycle
PARAllel FACtor analysis (PARAFAC)
Multivariate statistics
Gas chromatography-mass spectrometry (GC-MS)

ABSTRACT

The concentration of estrogens in the body fluids of women is highly variable, due to the menstrual cycle, circadian oscillations, and other physiological and pathological causes. To date, only the cyclic fluctuations of the principal estrogens (estradiol and estrone) have been studied, with limited outcome of general significance. Aim of the present study was to examine in detail the cyclic variability of a wide estrogens' panel and to interpret it by multivariate statistics.

Four estrogens (17α-estradiol, 17β-estradiol, estrone, estriol) and eleven of their metabolites (4-methoxyestrone, 2-methoxyestrone, 16α-hydroxyestrone, 4-hydroxyestrone, 2-hydroxyestrone, 4-methoxyestradiol, 2-methoxyestradiol, 4-hydroxyestradiol, 2-hydroxyestradiol, estriol, 16-epiestriol, and 17-epiestriol) were determined in urine by a gas chromatography – mass spectrometry method, which was developed by design of experiments and fully validated according to ISO 17025 requirements. Then, urine samples collected every morning for a complete menstrual cycle from 9 female volunteers aged 24–35 years (1 parous) were analysed.

The resulting three-dimensional data (subjects × days × estrogens) were interpreted using several statistical tools. Parallel Factor Analysis compared the estrogen profiles in order to explore the cyclic and inter-individual variability of each analyte. Principal Component Analysis (PCA) provided clear separation of the sampling days along the cycle, allowing discrimination among the luteal, ovulation, and follicular phases. The scores obtained from PCA were used to build a Linear Discriminant Analysis classification model which enhanced the recognition of the three cycle's phases, yielding an overall classification non-error rate equal to 90%. These statistical models may find prospective application in fertility studies and the investigation of endocrinology disorders and other hormone-dependent diseases.

1. Introduction

Estrogens play a variety of crucial roles in the menstrual cycle and throughout the entire life of women. The menstrual cycle is the cyclic orderly sloughing of the uterine lining, in response to the interactions of hormones produced by the hypothalamus, pituitary and ovaries [1]. The duration of a complete menstrual cycle spans from 21 to 35 days, with an average of 28 days. The menstrual cycle is usually divided into the follicular and the luteal phases. The follicular phase begins from the first day of menses until ovulation, which typically occurs around the 14th day. After ovulation, the luteal phase starts and lasts 14 further days, on average [1–4]. Lifestyle factors, such as smoking, physical activity and alcohol consumption may affect the phases of the menstrual cycle [5]. Abnormally high and low values of body mass index (BMI) are frequently associated to menstrual dysfunctions, due to the correlation of the estrogens metabolism with the nutrition and dietary composition and the role of adipose tissue in aromatase conversion [6]. The natural rhythmic fluctuations of the estrogens that control the menstrual cycle influence the fertility [7–11] and various physical and psychological conditions [3,12–15].

An important methodological issue with the study of estrogens data is how to align the cycles of the different women to allow comparisons [9]. In the Nurses' Health Study II, this issue was overcome by sampling all the women during their luteal phase [16]. The main problem for this approach is the difficult recognition of the menstrual phase in women with irregular periods. To date, only the variation of estrone or estradiol levels were evaluated across complete menstrual cycles, possibly because these are the main estrogens circulating in the human body,

* Corresponding author.
E-mail address: eugenio.alladio@unito.it (E. Alladio).

https://doi.org/10.1016/j.steroids.2019.108432
Received 22 January 2019; Received in revised form 11 June 2019; Accepted 18 June 2019
Available online 04 July 2019
0039-128X/

together with estriol [17]. A comprehensive evaluation of an extended estrogenic profile was previously proposed with the purposes of detecting any possible correlations between estrogens and breast cancer risk: the urinary estrogenic profile of 15 free and conjugated estrogens was collected from a large cohort of pre-menopausal women and retrospectively interpreted, on the basis of their clinical history [16] Extended estrogenic profiles were also correlated with terminal duct lobular unit involution, a marker of increased breast cancer risk [18]. In parallel studies on post-menopausal women, the determination of blood estrogens and metabolites revealed a lower risk of breast cancer for the subjects with high levels of the hydroxylated 2-pathway metabolites [19,20].

The estrogen determinations most frequently reported in the literature are conducted on either urine or oral fluid, using radioimmunoassay (RIA), enzyme immunoassay (EIA), or enzyme-linked immunosorbent assay (ELISA) [22]. While these immunoassay methods provide high throughput, efficiency, ease of use, fast turnaround time and low cost, they frequently do not have the necessary specificity and sensitivity to accurately measure low estrogen concentrations, due to cross-reactivity with structurally similar substances [21,22]. This limits the chance of estrogen profiling during the follicular and late luteal phases, when their concentration level is particularly low. In contrast, the hyphenation of chromatographic and mass spectrometric (MS) techniques provides the simultaneous dosage of both parent estrogens and their metabolites ensuring at the same time extremely low detection limits [21–23]. Liquid chromatography (LC) and gas chromatography (GC) coupled with MS are consistently used in multi-analyte profiling, with LC-MS increasingly favored for its straightforward applicability, even if GC-MS has traditionally dominated the analysis of estrogens and other endogenous steroids for years. Actually, GC-MS provides broad steroids profiles after a single derivatization step, achieving high specificity, good sensitivity, and limited matrix effects [23]. In general, the advantage of high-resolution separation is increasingly valued in targeted and untargeted metabolomics to obtain complete urinary endogenous steroid profiles that include estrogens, androgens, corticoids, and progesterone [24–26]. Multi-residual GC-MS methods for the detection of wide estrogen profiles have occasionally been developed in the past [27–30], even if the laborious sample preparation steps somehow contributed to the progressive decline of GC-MS procedures in favor of LC-MS.

In the present study, 15 estrogens were monitored in nine women along one menstrual period using an optimized and fully validated GC-MS method. The collected data were used to build a preliminary multivariate model shaping the menstrual cycle, which may represent a valuable tool in the study of fertility issues, as well as in the screening and evaluation of various pathological conditions, including endocrinology disorders and hormone-dependent cancers.

2. Material and methods

2.1. Chemicals and reagents

4-methoxyestrone, 4-methoxyestradiol, 2-methoxyestrone, 16α-hydroxyestrone, 2-methoxyestradiol, 2-hydroxyestradiol, 4-hydroxyestrone, 4-hydroxyestradiol, 17-epiestriol and 16-epiestriol were purchased from Toronto Research Chemical Inc. (Toronto, ON, Canada). 17α-estradiol, 17β-estradiol, hexane, methanol, ethyl acetate, ascorbic acid, ammonium iodide, *tert*-butyl methyl ether (TBME), dithioerythritol and *N*-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA), β-glucuronidase/arylsulfatase (from *Helix pomatia*) mixture, were provided by Sigma-Aldrich (Milan, Italy). Estrone, 2-hydroxyestrone, estrone 3-(β-D-glucuronide) sodium salt, estriol, estrone-d_4 and 17β-estradiol-d_4 were supplied by LGC Promochem SRL (Milan, Italy). β-glucuronidase from *Escherichia coli* was purchased from Roche Life Science (Indianapolis, IN, USA). Ultra-pure water was obtained from a Milli-Q® UF-Plus apparatus (Millipore, Bedford, MA, USA). C-18 endcapped Solid-Phase Extraction (SPE) cartridges were provided by UCT Technologies (Bristol, PA, USA) and estrone 3-sulfate sodium salt was supplied by Steraloids Inc. (Newport, RI, USA).

All stock standard solutions were prepared in methanol at 1 mg/mL and stored at –20 °C until use. Working solutions containing a mixture of all analytes were prepared at the final concentrations of 20 μg/mL and 1 μg/mL by appropriate dilution with methanol. Estrone-d_4 and 17β-estradiol-d_4 were used as isotopically labelled internal standards for quantitation and were added from separate methanol working solutions at the final concentrations of 100 μg/mL and 50 μg/mL, respectively.

2.2. Sample preparation

The sample preparation conditions were optimized after design of experiments [31], described elsewhere [32]. The urine sample (6 mL) was fortified with both 17β-estradiol-d_4 and estrone-d_4 internal standard solutions at the final concentrations of 50 ng/mL and 25 ng/mL. After that, the pH was checked and, if necessary, some drops of HCl were added to attain a final pH of 5.5. 2 mL acetate buffer 1.1 M (pH 5.5) and 50 μL ascorbic acid 1 M were added, too. Ascorbic acid was necessary to protect the labile catechol groups and prevent their degradation [27,29]. A deconjugation step, useful to transform the glucuronide and sulphate conjugated estrogens [2,17,33] into the free form, was executed by adding 20 μL of β-glucuronidase/arylsulfatase mixture to the urine samples, which were then incubated at 37 °C overnight. The next morning, 100 μL β-glucuronidase from *Escherichia Coli* was added, together with 50 μL of ascorbic acid solution and the final enzymatic deconjugation of the remaining glucuronide estrogens was carried out for 1 h at 58 °C. Once the hydrolysis was completed, the mixture was cooled to room temperature and 2 mL of 0.1 M carbonate buffer (pH 9) with some drops of NaOH 1 M were added, to obtain a final pH > 9. Liquid-liquid extraction (LLE) was performed by adding 5 mL of ethyl acetate and hexane (2:3 v/v) mixture to the sample, which was subsequently shaken in a vortex multimixer (Tecnovetro, Monza, Italy) for 5 min and subjected to centrifugation (model Megafuge 1.0 Heraeus; ASHI, Milan, Italy) at 4000 rpm for 5 min. The extraction process was repeated twice, and the two combined organic phases were transferred into a vial and evaporated to dryness under a gentle stream of nitrogen at 40 °C using a Techne Sample Concentrator (Barloworld Scientific, Stone, UK). The dried residue was reconstituted and derivatized for 1 h at 70 °C by adding 50 μL of MSTFA/NH_4I/dithioerythritol (1.000:2:4 v/w/w) solution. A 2 μL aliquot was injected into the GC/MS system in the splitless mode.

2.3. GC-MS analysis

All analyses were conducted on an Agilent 6890 N Network GC System interfaced to a 5975 inert XL Mass Selective Detector (Agilent Technologies, Milan, Italy). The GC was equipped with a J&W Scientific HP-1 17.0 m × 200 μm (i.d.) × 0.11 μm (f.t.) capillary column. The helium gas carrier was employed at a constant pressure of 23.25 psi and 1.1 mL/min initial flow. The GC oven temperature was initially set at 200 °C, held for 2 min, then was raised to 225 °C with an 8 °C/min ramp. Then, the temperature was increased to 234 °C with a 3 °C/min heating rate, held for 3 min and raised again to 245 °C with a 3 °C/min ramp. The final oven temperature of 315 °C was reached with a 40 °C/min heating rate and held for 3 min. The total run time was 19.54 min. The GC injector and transfer line were maintained at the temperature of 280 °C. Trimethylsilyl derivatives of the analytes were ionized by electron ionization (EI) at 70 eV. Data were acquired in the selected ion monitoring (SIM) mode at a dwell time of 20 ms [32,34].

2.4. Method validation

The analytical method was validated according to the Eurachem

criteria and recommendations [35]: linearity range, selectivity, specificity, limit of detection (LOD), limit of quantitation (LOQ), intra-assay precision and accuracy, repeatability, matrix effect, extraction recovery, and carry-over were determined. A pool of urines collected from healthy male volunteers (laboratory personnel), was negativized by solid-phase extraction using a C-18 end-capped cartridge previously conditioned with 2-propanol and ultra-pure water. The absence of any detectable trace of estrogens was verified. The resulting sample was used as blank urine and spiked with the standard solutions within the validation procedure. Full details about the validation of the analytical method are reported in a dedicated publication [32].

2.4.1. Linearity, LOD, LOQ

The calibration was performed by internal standardization using the least squares regression method from five replicate analyses for each data-point at six concentration levels in the range 1–50 ng/mL. Linearity was evaluated by lack-of-fit test, analysis of variance (ANOVA) test, Mandel's test, and relative standard deviation (RSD) of the slope, according to the approach described by Desharnais et al. [36]. Moreover, the residual plots and the deviation from back-calculated concentrations were examined. When heteroscedastic distribution of data-points was observed, a weighting factor of x^{-1} or x^{-2} was employed, depending on the rate of the variance increase with the concentration (linear or quadratic).

LOD and LOQ were estimated for all the target analytes using the Hubaux-Vos' algorithm at a significant level of 95% [37] from the 30 data-points collected to build the calibration lines. To confirm the correct estimation further, the calculated LOD and LOQ values were experimentally verified with blank samples spiked at concentrations close to the detectable and quantifiable values, respectively. In the operational practice, LOQ values were assumed at the lower level of the calibration curves.

2.4.2. Repeatability and accuracy

The retention time repeatability was verified on the chromatographic peak of the target analytes recorded in the 30 overall analyses used to build the five calibration curves (see above). Deviations below 1% from calibrators and controls were considered satisfactory. The repeatability of the relative ion abundance was evaluated on the selected ion chromatograms for each target analyte. The variations were considered acceptable within ± 20%, with respect to the controls.

For all analytes, intra-day repeatability and accuracy were evaluated on 10 blank urine samples spiked with all the target analytes at three concentration levels (1.0 ng/mL, 5.0 ng/mL and 25 ng/mL). Precision and accuracy were estimated from the percent variation coefficient (CV%) and percent bias (bias%), respectively. Precision was considered satisfactory when the CV% values were below 15% for the low calibration level and below 10% for the other levels. Satisfactory accuracy was achieved when the experimentally determined average concentration lied within ± 10% from the expected value.

2.4.3. Matrix effect, extraction recovery, enzyme performance, carry-over

The matrix effect was evaluated at the three concentration levels defined above by comparing the experimental results obtained from blank urine samples (mean value from five replicates) and blank deionized water solution both spiked after the extraction step at the same concentration. The matrix effect for each target analyte was expressed as the percentage ratio between the two measured concentrations. Extraction recovery was calculated at the same concentration levels by comparing the experimental results from blank urine samples spiked respectively before and after the extraction step (5 replicates each) and expressed as the percentage ratio between the two data.

The efficiency of β-glucuronidase and arylsulfatase to achieve exhaustive hydrolysis of the conjugated metabolites was tested at three concentration levels (1.0 ng/mL, 5.0 ng/mL and 25 ng/mL) by measuring the percentage ratio between the recovered concentration of estrone glucuronide (and sulfate) spiked into a blank sample and that of free estrone spiked to another blank sample at the same molar concentration. All the analyses were performed in duplicate.

The carry-over effect was evaluated by injecting in alternate sequence five blank urine samples spiked with all the analytes at the highest concentration and five blank deionized water solutions. Moreover, the carry-over effect was considered negligible if the S/N ratio was lower than 3 at the analytes' retention time for each monitored ion chromatogram obtained from the latter solutions.

2.5. Real urine sample collection

First morning urine samples were collected every day during a complete menstrual cycle (28 days) from 9 female volunteers aged 24–35 years, average 27.6 y ± 3.4 (1 parous). The 252 total samples were maintained at –20 °C and randomly analysed once the monthly collection was completed. All the women were healthy. None of them was taking any pharmaceutical drug or combined oral contraceptive pills in the period of sample collection. For all urine samples, the analytical determinations were normalized against their creatinine concentration to compensate for the physiological urinary dilution [2,4]. Creatinine was determined by the alkaline picrate photometric method using the dedicated kit on Architect C800 instrumentation (Abbott srl, Rome). In order to follow privacy regulations, an anonymous code was attributed to each participant subject who, anyway, voluntarily donated samples to the present project.

2.6. Chemometrics

Multivariate data analysis was carried out using Matlab® (The MathWorks, MA, USA) version 9.0.0 with PCA Toolbox version 1.2 [38], *N*-way Toolbox version 2.10 [39] and Classification Toolbox version 5.0 [40].

Data were arranged into a three-dimensional array (3-way), labelled as $\underline{\mathbf{X}}$, with dimensions ($I \times J \times K$) chosen as follows: (I) 9 subjects (representing the female volunteers), (J) 28 days (representing the menstrual cycle duration), (K) 15 variables (representing the studied estrogens). To analyse the three-dimensional data, a Parallel Factor Analysis (PARAFAC) model [41–43] was applied. The Alternating Least Squares (ALS) algorithm basically decomposes the $\underline{\mathbf{X}}$ 3-way array into three two-dimensional matrices , namely **A** ($I \times L$), **B** ($J \times L$) and **C** ($K \times L$), where the former variables (I, J, K) are expressed as a function of a new multivariate parameter (L) representing the loadings [41–43]. The **B** and **C** matrices show the natural fluctuation of each estrogen concentration throughout the 28-day menstrual cycle.

In order to separate the different phases of the menstrual cycle (i.e., follicular phase, ovulation and luteal phase), the Principal Component Analysis (PCA) [44] was carried out, as an exploratory method for multivariate data analysis. Since PCA works on two dimensional data matrices, the 3-way matrix $\underline{\mathbf{X}}$ was unfolded in a J × IK matrix (i.e., 28 × 135), after autoscaling. The PCA model was built employingfollowing a venetian blinds cross-validation procedure, with a number of data splits equals to 5. The optimal number of principal components (PCs) was chosen from the predicted residual sums of squares (PRESS), root mean squared error of cross-validation (RMSECV) and the scree plot. Further parameters, including eigenvalues, percentage variance captured by each PC (Var%) and percentage cumulative variance captured by the model (CumVar%) were also evaluated [44].

Lastly, a linear discriminant analysis (LDA) model was built to verify the classification power of the multivariate estrogenic profile with respect to the phase of the menstrual cycle (e.g., luteal phase, ovulation and follicular phase). The variables used to build the LDA model were the first 10 PCs scores, obtained as linear combinations of the original estrogen concentrations. This approach has the advantage of removing noise from the dataset and improving the classification performances. The data multi-normality was verified and again a cross-

validation procedure was performed by applying the venetian blinds design technique with 5 data splits. The classification criterion based on the Bayes' rule assigned each sampled day to the category showing the highest probability [40].

3. Results and discussion

3.1. Method optimization and validation

The DoE optimization of sample preparation [32] was aimed to achieve simpler and faster extraction conditions than those used in previous studies [27–30]. The best combination of drying temperature (found at 40 °C) and extraction solvent was found with the ethyl acetate + hexane (2:3 v/v) mixture as it corresponded to higher resolution and intensity of the chromatographic peaks with respect to TBME [32].

Optimal chromatographic separation among the estrogens of similar chemical structure (for example, 2-methoxyestrone and 16α-hydroxyestrone) was obtained by using a slow increase of the oven temperature (3 °C/min) between 225 °C and 245 °C interrupted by a hold time at 234 °C for 3 min. Nevertheless, the full chromatographic run was completed in less than 20 min and the retention times of the target analytes lied between 6.58 min (17α-estradiol) and 10.50 min (16-epiestriol).

3.1.1. Linearity, LOD, LOQ

Full validation data are reported elsewhere [32]. The linearity of the calibration curves was tested in the concentration range of 1.0–50 ng/mL for all the analytes. Lack of fit's, ANOVA, RSD slope and Back-calculation tests proved to yield calculated results below the respective critical values for all the target analytes. Among the target analytes, only 17α-estradiol, 4-hydroxyestradiol and 2-methoxyestradiol were characterized by a quadratic calibration model. Most of the estrogens' models used an x^{-2} weighting correction, except 17α-estradiol, 2-methoxyestradiol, estrone, 4-methoxyestrone and estriol. From the residual plots, the calibration linearity was confirmed by the presence of random residuals patterns along the concentration ranges for all the analytes.

LOD values ranged between 0.2 ng/mL and 0.4 ng/mL. The LOQ values, estimated below 1.0 ng/mL for all target analytes, were verified experimentally. The first point (1.0 ng/mL) of each calibration range was successfully tested for precision and accuracy, as reported below, and was subsequently used as operational LOQ.

3.1.2. Repeatability and accuracy

Ion abundance and retention time repeatability proved experimentally appropriate. Intra-assay precision and accuracy satisfied the target criteria, as the CV% are lower than 15% for all the analytes at all tested concentration levels, while the percent bias (bias%) lied between −8.2% (2-hydroxyestrone) and +12% (2-hydroxyestradiol) at 1.0 ng/mL, −11% (4-hydroxyestradiol) and + 6.8% (4-hydroxyestrone) at 5.0 ng/mL and −6.2% (17α-estradiol) and +5.6% (2-hydroxyestradiol) at 25 ng/mL [32].

3.1.3. Matrix effect, extraction recovery, enzyme performance, carry-over

The matrix effect values ranged from −12% for 4-methoxyestrone to +15% for 16α-hydroxyestrone at low level, from −9.2% for 4-hydroxyestrone to +12% for 17-epiestriol at medium level and from −5.6% for 2-hydroxyestradiol to +6.3% for 16α-hydroxyestrone at high level. These scattered values are close to the experimental bias and do not evidence any significant matrix effect. The average recovery efficiency was 99%, with minima and maxima ranging from 89% for 4-hydroxyestrone to 108% for 4-hydroxyestradiol and 17-epiestriol at 1.0 ng/mL; from 87% for 17α-estradiol to 107% for 4-hydroxyestrone at 5.0 ng/mL; from 94% for 17α-estradiol to 110% for 2-hydroxyestradiol at 25 ng/mL. Again, the extraction recovery was virtually complete at all concentration levels allowing a correct estimation of the target analytes' concentration.

The percent hydrolysis achieved by both β-glucuronidase and arylsulfatase on estrone glucuronide and estrone sulfate at all concentration levels was close to 100%, supporting the claim that the deconjugation efficiency on phase II metabolites could be considered complete. No carry-over effect was observed.

3.2. PARAFAC model

The PARAFAC approach is commonly employed in environmental data analysis, when repeated chronological monitoring of sampling sites yields three-dimensional data structures. The same statistical tool is suitable for our chronological monitoring of estrogen profiles [39,41–43]. A PARAFAC model was built to extract the concentration profile for each estrogen along the 28-day menstrual cycle, by smoothing the large individual variability of estrogenic profiles, that proved significant for the 9 investigated women. Due to the different duration of the menstrual cycles, spanning between 28 and 30 days, the ovulation peak occurred at different days, from the 13th to the 17th day, in agreement with the literature [17,45–47]. To comply with this source of variability, the extreme sampling days were removed from the series collected from the women with a menstrual cycle longer than 28 days. Actually, the extreme days (i.e. the first and the last of the menstrual cycle) exhibited comparable results with the subsequent and preceding samples, respectively. The final PARAFAC processing allowed the equalize each menstrual cycle within a unique scale so as to evaluate and compare the natural variation of the estrogenic levels. The number of significant factors for the PARAFAC model was two, that explain a CumVar% of 86.98%, relative to Var% = 75.17% and Var % = 10.81% for factor 1 and factor 2, respectively.

All the extrapolated estrogenic profiles are reported in Fig. 1 and exhibit several remarkable features. In particular, 17β-estradiol (2a) and estrone (2b) show two peaks, the first occurring close to the ovulation with a time-shift of 3–4 days between the two hormones, while the second smoother peak appears in the period around the 20th-25th day of the cycle. These profiles are comparable to those reported in the literature [7–9,12,48]. In contrast, no peak is observed for 17α-estradiol (2a) in the central part of the cycle and only a faint increase of its level is detectable in the luteal phase of the cycle. The lack of correlation between 17α- and 17β-estradiol profiles may explain the scarce specificity of the immunoassays methods used for their quantification.

Several metabolite profiles are characterized by the occurrence of a single concentration peak around the ovulation, namely 2-hydroxyestradiol (1c), 4-hydroxyestradiol (1c), and 2-methoxyestrone (1f) at the 15th day, but 4-hydroxyestrone (1e) and 2-hydroxyestrone (1e) together with estrone (1b) at the 17th day. Surprisingly, 2-methoxyestrone shows a chronological correlation with hydroxyestradiol isomers instead of hydroxyestrone isomers, as it would be expected. On the other hand, 4-methoxyestrone (1f) show a sharp peak in the follicular phase of the cycle, that is not observed for the isomer 2-methoxyestrone. The different behaviour observed for 2- and 4-methoxyestrone isomers contrasts with those recorded for the analogous hydroxyestrone (1e) and hydroxyestradiol (1c) isomers. All these observations add details on the complex regulating system of the estrogen biochemistry active during the ovulation phase which can not be explained by straightforward and progressive metabolic pathways [14].

4-methoxyestradiol (1d), 16-epiestriol, and 17-epiestriol (1 g) display a profile in which the concentration increases around the ovulation and remains quite stable for the subsequent 10 days, whereas 2-methoxyestradiol (1d) shows a constant decrease along the cycle.

Barrett et al. [7] and Venners et al. [9] determined the urinary concentration of estrone alongside the entire menstrual cycle by immunoassay: the resulting profiles showed the same pattern that we observed in the PARAFAC profile, even if the analytical methods were different. A comparable profile was also observed by Baird et al. [8]

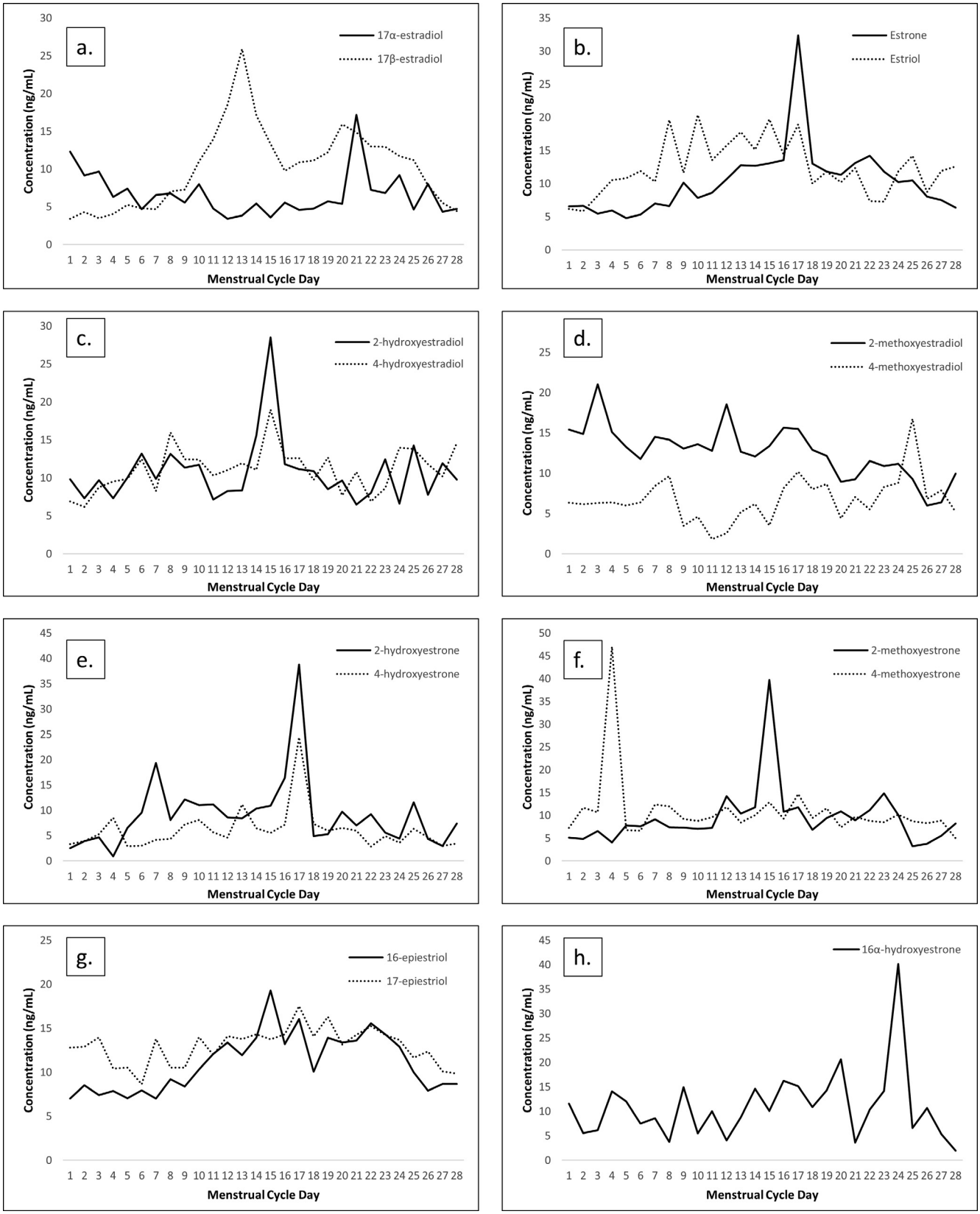

Fig. 1. Concentration profile (normalized for the creatinine value) of the target analytes along the 28-day menstrual cycle achieved by applying PARAFAC approach: (a) 17α-estradiol and 17β-estradiol, (b) estrone and estriol, (c) 2-hydroxyestradiol and 4-hydroxyestradiol, (d) 2-methoxyestradiol and 4-methoxyestradiol, (e) 2-hydroxyestrone and 4-hydroxyestrone, (f) 2-methoxyestrone and 4-methoxyestrone, (g) 16α-hydroxyestrone and 16-epiestriol, and (h) 17-epiestriol.

who used a radio-immunoassay method. Likewise, the profiles of 17α- and 17β-estradiol that we observed substantially overlaps the ones reported by Roney et al. [12] and Barrett et al. [48], although in these studies the concentrations were measured by immunoassay in the oral fluid. Basically, two peaks are observed, the first one just before the ovulation and the second during the luteal phase. Noteworthy, the first peak has not been observed in our study for 17α-estradiol.

The correspondence of our data with literature profiles and the agreement between oral fluid and urine data, and between GC-MS and immunoassay methods represent further confirmation of the reliability of the present approach to gain general information about the relative concentration of the circulating hormones. The multi-residual GC-MS method proposed in this study proved to represent a fast, cheap, practical, and reliable analytical tool for the monitoring of an extended estrogenic profile in young women (24–35 years), overcoming the lack of specificity typical of immunoassay methods.

3.3. PCA results and LDA model

Principal component analysis (PCA) was performed on the complete 28 × 135 data matrix with the purpose of discovering any underlying structure in the data The optimal number of principal components (PC) to be considered was two representing a CumVar% of 28.22% and a RMSECV% of 16.13%. The limited percentage of total variance explained by PC1 + PC2 (only 28.22%) is coherent with the large variability of the data. In practice, the new PC variables, as linear combination of the old ones (concentration of estrogens), emphasize the information content present in the data while reducing the contribution of their random fluctuation. The scores plot of PC1 (Var% 14.91%) vs PC2 (Var% 13.31%) is reported in Fig. 2A and shows the occurrence of three broad clusters corresponding to the three phases of the menstrual cycle: follicular, ovulation, and luteal. The follicular and ovulation phase data are separated along PC2, while the ovulation and luteal phase data along PC1.

By plotting PC1 and PC2 as a function of the menstrual cycle day in two separate diagrams (Fig. 2B, C), the phase transitions become visible and the starting point of both the ovulation and luteal phase can be clearly identified.

A preliminary LDA model was built using the information extracted by the PCA scores. While the PARAFAC technique demonstrated that the original data were affected by large internal variability which prevented the construction of a reliable and stable classification model based on them, the PC scores are free from correlation and noisy pattern. Therefore, the PCA scores were used instead of the original estrogen data to build the LDA model.

The multi-normality of the PC scores was successfully checked (Fig. 3A) and then a cross-validating procedure was applied to the 28 × 10 matrix (10 PCs were considered). A cross-validated non-error rate of 90% was achieved, together with an accuracy equal of 93%, as is shown in the confusion matrix reported in Table 1. Only two data-points were misclassified, namely the 17th day, which was classified in class 3 (luteal phase) instead of class 2 (ovulation), and the 18th day, which was classified in class 1 (follicular phase) instead of class 3. The misclassification of the 17th and 18th days was not surprising since both data-points belong to the transition period from the ovulation to the luteal phase and correspond to a sudden drop of the estrone, 2- and 4-hydroxyestrone concentrations (Fig. 1b, e). The accurate classification of all days belonging to the transition from the follicular to the ovulation phase is explained by the smoother concentration increment observed for 17β-estradiol, 2- and 4-hydroyestradiol from the 11th to the 15th day of the cycle (Fig. 1a, 1c).

The scores plot reported in Fig. 3B shows the good partition of the days in three well-defined classes corresponding to the follicular phase, the ovulation and the luteal phase. The loadings plot, representing the PC variables in the space of the LDA canonical variables (Fig. 3C), indicates the correspondence between class discrimination and PCs. In

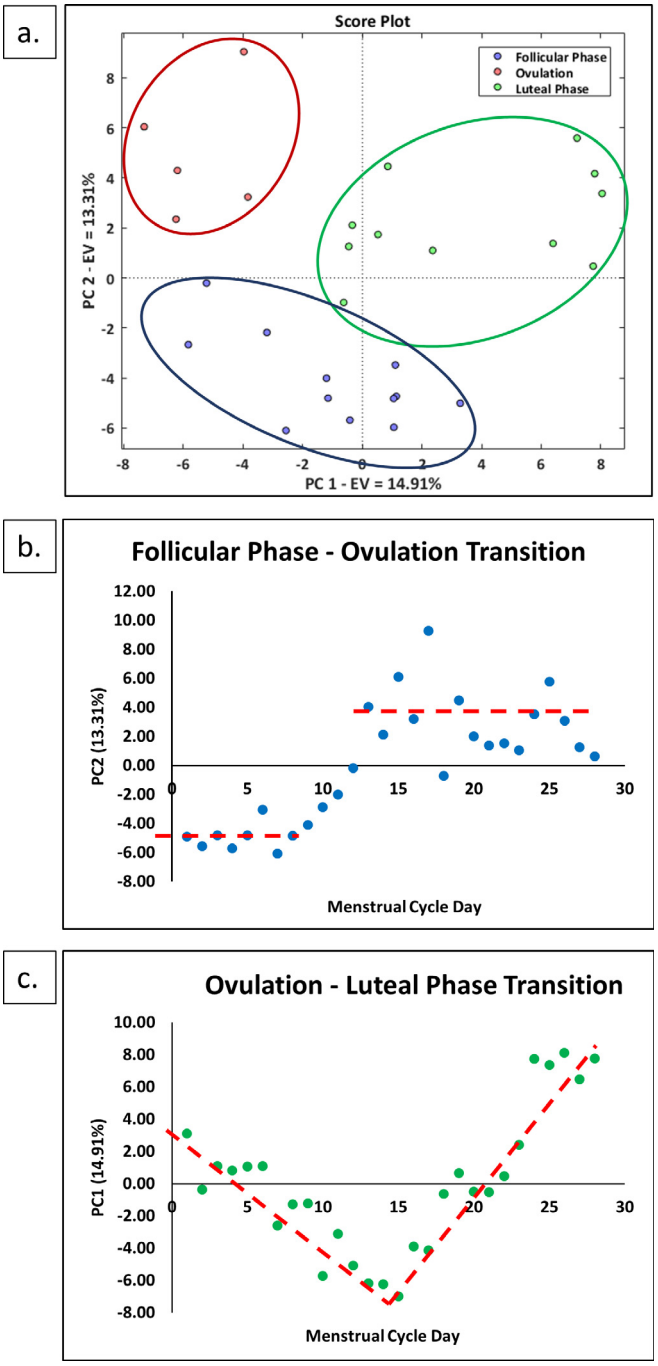

Fig. 2. Results provided by the PCA model: (a) Score plot relevant to PC1 (Var. = 14.91%) vs PC2 (Var. = 13.31%) showing the occurrence of three different clusters corresponding to the three phases of the menstrual cycle, i.e. follicular phase (blue dots), ovulation (red dots) and luteal phase (green dots). (b) PC1 vs menstrual cycle day, representing the transition from the follicular phase to the ovulation phase. (c) PC2 vs menstrual cycle day, representing the transition from the ovulation phase to the luteal phase.

particular, PC2 is high during the ovulation and low during the follicular and luteal phases. Hence, it is able to identify the ovulation period from the other phases of the cycle. On the other hand, the luteal phase is characterized by elevated values of PC1 (and PC7), which is low in the follicular phase and especially low during the ovulation, distinguished also by a high value of PC8.

Studies that use the menstrual cycle phase as a proxy for directly measured ovarian hormone levels typically fail to capture their inherent variability. The lack of reliable methods to divide the menstrual cycle

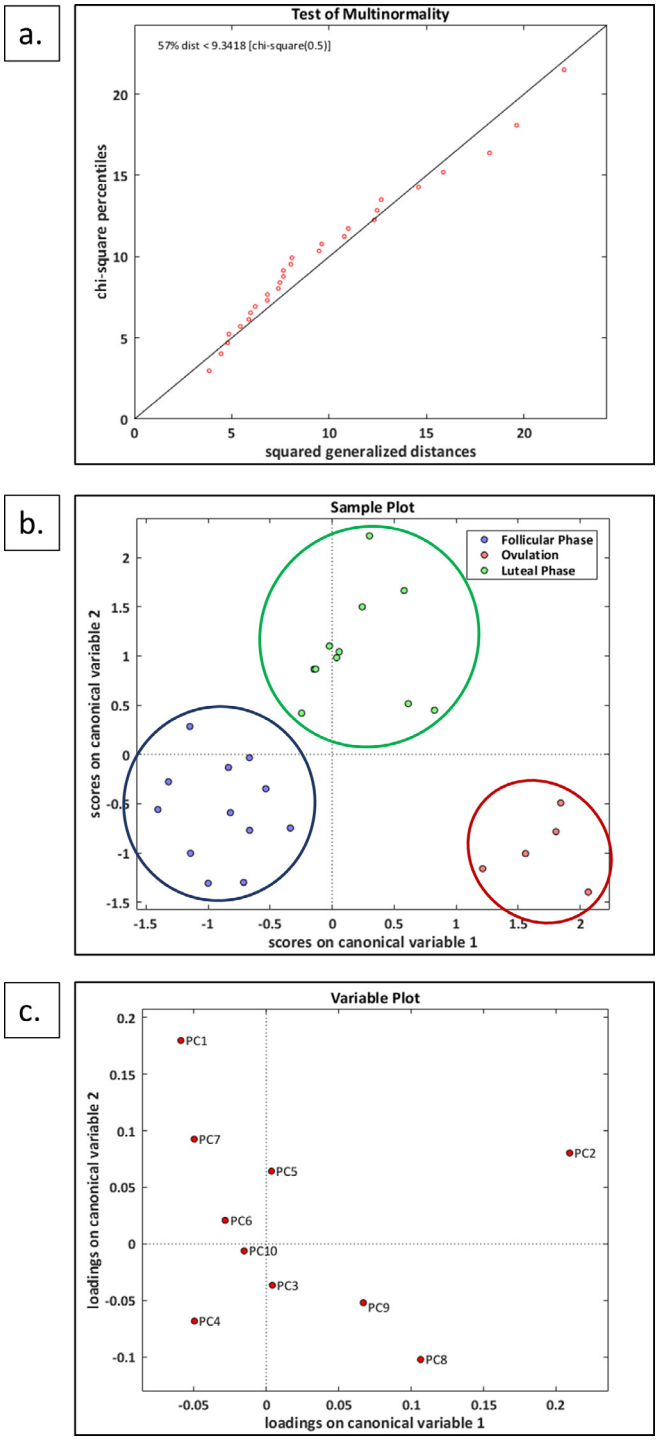

Fig. 3. Results achieved by building the LDA model. (a) Multinormality test graph. (b) Score Plot relevant to the first two latent variables, showing the partition of the data in three well-defined classes: follicular phase (blue dots), ovulation (red dots) and luteal phase (green dots). (c) Loading Plot relevant to the first two latent variables, showing the PCs that mainly characterize the three classes of samples.

into its component phases was proved, as divergent outcomes may be produced by using different methods [3,10]. The application of the present multivariate statistical model to GC-MS data is expected to overcome this limit and allow a correct definition of the phases of the menstrual cycle, an important issue in the study of fertility. For example, Barrett et al. [7] who determined the concentration of urinary estrone alongside a complete menstrual cycle established the difference

Table 1
Confusion matrix provided by the LDA model. The rows represent the real classes, while the columns represent the predicted ones; the correctly classified samples are reported on the diagonal. Overall non-error rate is reported, too.

Confusion matrix	Follicular phase	Ovulation	Luteal phase
Follicular Phase	12	0	0
Ovulation	0	4	0
Luteal phase	1	0	10
Non-error rate	90%		

in ovarian function between nulliparous and parous women. In general, the menstrual cycle features represent important indicators of the reproductive health and endocrine function. For example, Small et al. [11] found a connection between the menstrual cycle variability and the likelihood of pregnancy, Venners et al. [9] discovered that higher estrogen concentrations were associated with the occurrence of clinical pregnancy, and Baird et al. [8] studied the hormonal pattern most appropriate for pre-implantation. All the areas of interest linked to reproduction could benefit from a multivariate interpretation of a wide estrogen profile, such as the one proposed in the present study, which may find application also in the investigation of a variety of physical and psychological disorders.

4. Conclusions

In the present study, a GC-MS method is proposed for the simultaneous detection of 15 estrogens in the urine of a group of young women, that involves easy sample pretreatment, overcoming some of the limitations of previously published GC-MS protocols [28,29]. The reliability of the procedure was validated following a rigorous protocol and good performances were obtained, particularly in terms of efficient extraction recovery and adequate sensitivity, making the GC-MS approach competitive with the more demanding LC-MS/MS technique. In case that the concentration of one specific estrogen has to be determined with high accuracy, the method can be further improved by using a dedicated isotopically-labelled homolog as the internal standard.

Despite the large variability of the experimental data, the use of multivariate statistics on urine sample sequences collected from nine women – in particular the application of the PARAFAC approach – proved capable to extract the typical concentration profile for each analyte along the menstrual cycle, including estriol and the eleven metabolites not previously investigated in women. As a matter of fact, most of the existing literature only reports the variations of estrone and estradiol concentrations across the complete menstrual cycle, whereas in the present case a generalized picture for a broad urinary estrogen panel along the whole menstrual period has been described for the first time.

The advantages of using multivariate data analysis was made evident by the application of PCA, which yielded an easier visualization and efficient partition of the data into three groups, corresponding to the three phases of the menstrual cycle, namely the follicular phase, ovulation, and the luteal phase, together with the transitions between the phases.

The preliminary LDA model built on the PCA scores produced a reliable classification of each day along the cycle series, with a satisfactory cross-validated non-error rate of 90%. Therefore, the multivariate comparison of the estrogen profile collected from a single urine sample with the proposed model is likely to provide a trustworthy classification of this sample in terms of phase of the menstrual cycle (follicular, ovulation, luteal). Possible applications of the model include the detection of the fertile days along the cycle, the screening of pathological conditions, and the identification of particular stressing or psychological conditions of the investigated subjects. Further refinement of the present classification model is underway, as its full

validation will require a much larger training and test sets than the one used in this proof-of-concept contribution based the on the recruitment of nine volunteers.

References

[1] B.G. Reed, B.R. Carr, The Normal Menstrual Cycle and the Control of Ovulation, MDText.com, Inc., 2000.

[2] A. Bellem, S. Meiyappan, S. Romans, G. Einstein, Measuring estrogens and progestagens in humans: an overview of methods, Gend. Med. 8 (2011) 283–299, https://doi.org/10.1016/j.genm.2011.07.001.

[3] P.C. Regan, Rhythms of desire: the association between menstrual cycle phases and female sexual desire, Can. J. Hum. Sex. 5 (1996) 145–156.

[4] H.S. Driver, E. Werth, D.-J. Dijk, A.A. Borbély, The menstrual cycle effects on sleep, Sleep Med. Clin. 3 (2008) 1–11, https://doi.org/10.1016/j.jsmc.2007.10.003.

[5] Y. Liu, E.B. Gold, B.L. Lasley, W.O. Johnson, Factors affecting menstrual cycle characteristics, Am. J. Epidemiol. 160 (2004) 131–140, https://doi.org/10.1093/aje/kwh188.

[6] P. Montero, C. Bernis, V. Fernandez, S. Castro, Influence of body mass index and slimming habits on menstrual pain and cycle irregularity, J. Biosoc. Sci. 315–323 (1996).

[7] E.S. Barrett, L.E. Parlett, G.C. Windham, S.H. Swan, Differences in ovarian hormones in relation to parity and time since last birth, Fertil. Steril. 101 (2014) 1773–1780.e1, https://doi.org/10.1016/j.fertnstert.2014.02.047.

[8] D.D. Baird, A.J. Wilcox, C.R. Weinberg, F. Kamel, D.R. McConnaughey, P.I. Musey, D.C. Collins, Preimplantation hormonal differences between the conception and non-conception menstrual cycles of 32 normal women, Hum. Reprod. 12 (1997) 2607–2613, https://doi.org/10.1093/humrep/12.12.2607.

[9] S.A. Venners, X. Liu, M.J. Perry, S.A. Korrick, Z. Li, F. Yang, J. Yang, B.L. Lasley, X. Xu, X. Wang, Urinary estrogen and progesterone metabolite concentrations in menstrual cycles of fertile women with non-conception, early pregnancy loss or clinical pregnancy, Hum. Reprod. 21 (2006) 2272–2280, https://doi.org/10.1093/humrep/del187.

[10] J.B. Brown, Types of ovarian activity in women and their significance: the continuum (a reinterpretation of early findings), Hum. Reprod. 17 (2011) 141–158, https://doi.org/10.1093/humupd/dmq040.

[11] C.M. Small, A.K. Manatunga, M. Klein, C.E. Dominguez, H.S. Feigelson, R. McChesney, M. Marcus, Menstrual cycle variability and the likelihood of achieving pregnancy., Rev. Environ. Health. 25 (n.d.) 369–78.

[12] J.R. Roney, Z.L. Simmons, Hormonal predictors of sexual motivation in natural menstrual cycles, Horm. Behav. 63 (2013) 636–645, https://doi.org/10.1016/j.yhbeh.2013.02.013.

[13] D.M. Campagne, G. Campagne, The premenstrual syndrome revisited, Eur. J. Obstet. Gynecol. Reprod. Biol. 130 (2007) 4–17, https://doi.org/10.1016/j.ejogrb.2006.06.020.

[14] A. Shechter, D.B. Boivin, Sleep, hormones, and circadian rhythms throughout the menstrual cycle in healthy women and women with premenstrual dysphoric disorder, Int. J. Endocrinol. 2010 (2010) 1–17, https://doi.org/10.1155/2010/259345.

[15] D.H. Schwartz, S.E. Romans, S. Meiyappan, M.J. De Souza, G. Einstein, The role of ovarian steroid hormones in mood, Horm. Behav. 62 (2012) 448–454, https://doi.org/10.1016/j.yhbeh.2012.08.001.

[16] A.H. Eliassen, D. Spiegelman, X. Xu, L.K. Keefer, T.D. Veenstra, R.L. Barbieri, W.C. Willett, S.E. Hankinson, R.G. Ziegler, Urinary estrogens and estrogen metabolites and subsequent risk of breast cancer among premenopausal women, Cancer Res. 72 (2012) 696–706, https://doi.org/10.1158/0008-5472.CAN-11-2507.

[17] H. Samavat, M.S. Kurzer, Estrogen metabolism and breast cancer, Cancer Lett. 356 (2015) 231–243, https://doi.org/10.1016/j.canlet.2014.04.018.

[18] H. Oh, Z.G. Khodr, M.E. Sherman, M. Palakal, R.M. Pfeiffer, L. Linville, B.M. Geller, P.M. Vacek, D.L. Weaver, R.E. Chicoine, R.T. Falk, H.N. Horne, D. Papathomas, D.A. Patel, J. Xiang, X. Xu, T. Veenstra, S.M. Hewitt, J.A. Shepherd, L.A. Brinton, J.D. Figueroa, G.L. Gierach, Relation of serum estrogen metabolites with terminal duct lobular unit involution among women undergoing diagnostic image-guided breast biopsy, Horm. Cancer 7 (2016) 305–315, https://doi.org/10.1007/s12672-016-0265-2.

[19] R.T. Falk, L.A. Brinton, J.F. Dorgan, B.J. Fuhrman, T.D. Veenstra, X. Xu, G.L. Gierach, Relationship of serum estrogens and estrogen metabolites to postmenopausal breast cancer risk: a nested case-control study, Breast Cancer Res. 15 (2013), https://doi.org/10.1186/bcr3416.

[20] J.N. Sampson, R.T. Falk, C. Schairer, S.C. Moore, B.J. Fuhrman, C.M. Dallal, D.C. Bauer, J.F. Dorgan, X.-O. Shu, W. Zheng, L.A. Brinton, M.H. Gail, R.G. Ziegler, X. Xu, R.N. Hoover, G.L. Gierach, Association of estrogen metabolism with breast cancer risk in different cohorts of postmenopausal women, Cancer Res. 77 (2017) 918–925, https://doi.org/10.1158/0008-5472.CAN-16-1717.

[21] X. Xu, T.D. Veenstra, S.D. Fox, J.M. Roman, H.J. Issaq, R. Falk, J.E. Saavedra, L.K. Keefer, R.G. Ziegler, Measuring fifteen endogenous estrogens simultaneously in human urine by high-performance liquid chromatography-mass spectrometry, Anal. Chem. 77 (2005) 6646–6654, https://doi.org/10.1021/ac050697c.

[22] J.M. Faupel-Badger, B.J. Fuhrman, X. Xu, R.T. Falk, L.K. Keefer, T.D. Veenstra, R.N. Hoover, R.G. Ziegler, Comparison of liquid chromatography-tandem mass spectrometry, RIA, and ELISA methods for measurement of urinary estrogens, Cancer Epidemiol. Biomarkers Prev. 19 (2010) 292–300, https://doi.org/10.1158/1055-9965.EPI-09-0643.

[23] A.A. Franke, L.J. Custer, Y. Morimoto, F.J. Nordt, G. Maskarinec, Analysis of urinary estrogens, their oxidized metabolites, and other endogenous steroids by benchtop orbitrap LCMS versus traditional quadrupole GCMS, Anal. Bioanal. Chem. 401 (2011) 1319–1330, https://doi.org/10.1007/s00216-011-5164-3.

[24] S.H. Lee, Y.J. Yang, K.M. Kim, B.C. Chung, Altered urinary profiles of polyamines and endogenous steroids in patients with benign cervical disease and cervical cancer, Cancer Lett. 201 (2003) 121–131.

[25] M.H. Choi, J.-Y. Moon, S.-H. Cho, B.C. Chung, E.J. Lee, Metabolic alteration of urinary steroids in pre- and post-menopausal women, and men with papillary thyroid carcinoma, BMC Cancer 11 (2011) 342, https://doi.org/10.1186/1471-2407-11-342.

[26] J. Lee, H.M. Woo, G. Kong, S.J. Nam, B.C. Chung, Discovery of urinary biomarkers in patients with breast cancer based on metabolomics, Mass Spectrom. Lett. 4 (2013) 59–66, https://doi.org/10.5478/MSL.2013.4.4.59.

[27] U. Knust, T. Strowitzki, B. Spiegelhalder, H. Bartsch, R.W. Owen, Optimization of an isotope dilution gas chromatography/mass spectrometry method for the detection of endogenous estrogen metabolites in urine samples, Rapid Commun. Mass Spectrom. 21 (2007) 2245–2254, https://doi.org/10.1016/j.jcrs.2012.09.017.

[28] X.Y. Xiao, D. McCalley, Quantitative analysis of estrogens in human urine using gas chromatography/negative chemical ionisation mass spectrometry, Rapid Commun. Mass Spectrom. 14 (2000) 1991–2001, https://doi.org/10.1002/1097-0231(20001115)14:21<1991::AID-RCM125>3.0.CO;2-H.

[29] J.-Y. Moon, K.J. Kim, M.H. Moon, B.C. Chung, M.H. Choi, A novel GC-MS method in urinary estrogen analysis from postmenopausal women with osteoporosis, J. Lipid Res. 52 (2011) 1595–1603, https://doi.org/10.1194/jlr.D016113.

[30] P. Hoffmann, M.F. Hartmann, T. Remer, K.P. Zimmer, S.A. Wudy, Profiling oestrogens and testosterone in human urine by stable isotope dilution/benchtop gas chromatography-mass spectrometry, Steroids 75 (2010) 1067–1074, https://doi.org/10.1016/j.steroids.2010.06.014.

[31] R. Leardi, Experimental design in chemistry: a tutorial, Anal. Chim. Acta 652 (2009) 161–172, https://doi.org/10.1016/j.aca.2009.06.015.

[32] C. Bozzolino, S. Vaglio, E. Amante, E. Alladio, E. Gerace, A. Salomone, M. Vincenti, Optimization and validation of a GC-MS quantitative method for the determination of an extended estrogenic profile in human urine, MethodsX (submitted contextually).

[33] R.G. Ziegler, B.J. Fuhrman, S.C. Moore, C.E. Matthews, Epidemiologic studies of estrogen metabolism and breast cancer, Steroids 99 (2015) 67–75, https://doi.org/10.1016/j.steroids.2015.02.015.

[34] E. Alladio, R. Caruso, E. Gerace, E. Amante, A. Salomone, M. Vincenti, Application of multivariate statistics to the steroidal module of the athlete biological passport: a proof of concept study, Anal. Chim. Acta 922 (2016), https://doi.org/10.1016/j.aca.2016.03.051.

[35] Eurachem, The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, 2014. doi: 10.1016/S0014-2999(99)00500-2.

[36] B. Desharnais, F. Camirand-Lemyre, P. Mireault, C.D. Skinner, Procedure for the selection and validation of a calibration model I-description and application, J. Anal. Toxicol. 41 (2017) 261–268, https://doi.org/10.1093/jat/bkx001.

[37] A. Hubaux, G. Vos, Decision and detection limits for linear calibration curves, Anal. Chem. 42 (1970) 849–855, https://doi.org/10.1021/ac60290a013.

[38] D. Ballabio, A MATLAB toolbox for principal component analysis and unsupervised exploration of data structure, Chemom. Intell. Lab. Syst. 149 (2015) 1–9, https://doi.org/10.1016/j.chemolab.2015.10.003.

[39] C.A. Andersson, R. Bro, The N-way toolbox for MATLAB, Chemom. Intell. Lab. Syst. 52 (2000) 1–4, https://doi.org/10.1016/S0169-7439(00)00071-X.

[40] Davide Ballabio, Viviana Consonni, Classification tools in chemistry. Part 1: linear models, PLS-DA Anal. Methods 5 (16) (2013) 3790, https://doi.org/10.1039/c3ay40582f.

[41] R. Bro, PARAFAC, Tutorial applications, Chemom. Intell. Lab. Syst. 38 (1997) 149–171, https://doi.org/10.1016/S0169-7439(97)00032-4.

[42] P. Geladi, Analysis of multi-way (multi-mode) data, Chemom. Intell. Lab. Syst. 7 (1989) 11–30, https://doi.org/10.1016/0169-7439(89)80108-X.

[43] A.K. Smilde, R. Bro, P. Geladi, Multi-way Analysis with Applications in the Chemical Sciences, 1st ed., John Wiley & Sons Ltd, Chichester, UK, 2004.

[44] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831, https://doi.org/10.1039/c3ay41907j.

[45] S.B. Brown, S.E. Hankinson, Endogenous estrogens and the risk of breast, endometrial, and ovarian cancers, Steroids 99 (2015) 8–10, https://doi.org/10.1016/j.steroids.2014.12.013.

[46] L.G.A. de Chuffa, L.A. Lupi-Júnior, A.B. Costa, J.P.A. de Amorim, F.R.F. Seiva, The role of sex hormones and steroid receptors on female reproductive cancers, Steroids 118 (2017) 93–108, https://doi.org/10.1016/j.steroids.2016.12.011.

[47] E. Folkerd, M. Dowsett, Sex hormones and breast cancer risk and prognosis, Breast 22 (2013) S38–S43, https://doi.org/10.1016/j.breast.2013.07.007.

[48] E.S. Barrett, I. Thune, S.F. Lipson, A.-S. Furberg, P.T. Ellison, A factor analysis approach to examining relationships among ovarian steroid concentrations, gonadotrophin concentrations and menstrual cycle length characteristics in healthy, cycling women, Hum. Reprod. 28 (2013) 801–811, https://doi.org/10.1093/humrep/des429.