Optimization and validation of a GC–MS quantitative method for the determination of an extended estrogenic profile in human urine: Variability intervals in a population of healthy women

Title: Optimization and validation of a GC-MS quantitative method for the determination of an extended estrogenic profile in human urine. Variability intervals in a population of healthy women

Short title: GC-MS method to determine the women urinary estrogenic profile

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

Affiliations:

[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*

***corresponding author:** eleonora.amante@unito.it

Keywords: *estrogens; urine; gas chromatography – mass spectrometry; Design of Experiment; validation;*

Abstract

An analytical method based on gas chromatography-mass spectrometry (GC-MS) was developed for the determination of a wide panel of urinary estrogens, together with their principal metabolites. Due to the low concentration of estrogens in urine, an efficient sample pretreatment was optimized by a design of experiment (DoE) procedure to achieve satisfactory sensitivity. A second DoE was built for the optimization of the chromatographic run, with the purpose of reaching the most efficient separation of analytes with potentially interfering ions and similar chromatographic properties. The method was fully-validated using a rigorous calibration strategy: from several replicate analyses of blank urine samples spiked with the analytes, calibration models were built with particular attention to the study of heteroscedasticity and quadraticity. Other validation parameters, including the limit of detection, intra-assay precision and accuracy, repeatability, selectivity, specificity and carry-over were obtained using the same set of data. Further experiments were performed to evaluate matrix effect and extraction recovery. Then the urinary estrogen profiles of 138 post-menopausal healthy women were determined. These profiles provide a representation of physiological concentration ranges, which, in forthcoming studies, will be matched on the base of multivariate statistics with the urinary estrogenic profile of women with breast or ovarian cancer.

1. Introduction

The estrogens biosynthesis is regulated by several enzymes, most of which belonging to the class of cytochromes (CYP). CYP17 converts pregnenolone and progesterone to precursors of androgens. The aromatase CYP19 converts testosterone and androstenedione to 17β-estradiol and estrone, respectively. 17β-estradiol is then transformed into (i) estrone, by means of a dehydrogenase enzyme; (ii) estriol, by CYP1A1; (iii) 2-OH-estradiol and 4-OH-estradiol, by CYP1B1. A methyl group is subsequently transferred to the latter estrogens (by catechol-O-methyltransferase – COMT) to form 2- and 4-MeO-estradiol, which are then excreted (Greenlee et al., 2007). In turn, Estrone can be transformed into (i) 17α- and 17β-Estradiol by means of an oxidoreductase enzyme; (ii) 2- and 4-OH-estrone, by CYP1A1 (further metabolized to the corresponding oxy-methylated metabolites, similarly to the above-cited hydroxy-estradiol; (iii) 16α-OH-estrone, by the action of CYP3A. The latter can be further metabolized to estriol and 16-epiestriol (Mason, 2002; Zhu & Lee, 2005). A comprehensive scheme of the metabolic pathway is reported in Figure 1.

The enzymes involved in the estrogenic metabolic pathway are differently expressed in women during their fertile and post-menopausal ages. The variation is particularly strong for aromatase (Bulun et al., 2007; Cui, Shen, & Li, 2013). In pre-menopausal women, estrogens are produced primarily in the ovaries, corpus luteum and placenta (Cui et al., 2013) whereas in the post-menopausal period, the most important organs or their biosynthesis are the liver, heart, skin, brain, and adipose tissue, the last becoming the prevalent aromatase expressing body site (Bulun et al., 2007; Cui et al., 2013; Simpson et al., 1999). Estradiol is the main biosynthetic product and also the most potent estrogen during the pre-menopausal period, while estrone plays a predominant role after menopause (Cui et al., 2013). In general, a substantial decline of circulating estrogens is observed after the menopause, with several consequences on the women health, including bone decalcification and increased cardiovascular disease (Simpson et al., 1999). Furthermore, the metabolic changes occurring after the

fertile period are associated with a several forms of cancer (Moon, Kim, Moon, Chung, & Choi, 2011), even if the specific role of estrogen metabolites is still not clear.

The role of estrogens in breast cancer evolution is well established (Berrino et al., 1996; Lippert, Seeger, & Mueck, 2000), even though a few contradictory results can be found in the literature, especially concerning the proposal of new biomarkers for its diagnosis. For example, Bradlow *et al.* observed an increased urinary 16α-hydroxyestrone/2-hydroxyestrone ratio in women affected by several forms of breast cancer and recommended to use this concentration ratio as a predictive biomarker for breast cancer (Bradlow, Davis, Lin, Sepkovic, & Tiwari, 1995). In contrast, a review from 2011 Obi *et al.* rejected this proposal and suggested to focus on the estrogen metabolites profiles instead (Obi, Vrieling, Heinz, & Chang-Claude, 2011). In a recent work, our research group developed a multivariate interpretation of an extended urinary androgenic profile as an effective approach for the diagnosis of prostate carcinoma (De Luca et al., 2020). A similar strategy will be tested for the diagnosis and prognostic monitoring of breast, ovarian and uterus cancer. In the past , the quantification of urinary estrogens had been predominantly completed by radioimmunoassay, enzyme immunoassay and enzyme-linked immunosorbent assay (Faupel-Badger et al., 2010). The limitations of these techniques are linked to risk of cross-reactivity, which affects the test's specificity, and the low sensitivity, limiting their applicability. Chromatography coupled to mass spectrometry offers the combined advantages of lower detection limits and concurrent multi-analyte determination. Gas chromatography – mass spectrometry methods (GC-MS), in particular, have been reported for the detection of estrogen profiles (Hoffmann, Hartmann, Remer, Zimmer, & Wudy, 2010; Knust, Strowitzki, Spiegelhalder, Bartsch, & Owen, 2007; Moon et al., 2011; Xiao & McCalley, 2000).

In this paper, we describe the development of a GC-MS method for the detection of 14 urinary estrogens, including Estrone, 17α-Estradiol, 17β-Estradiol, and their main metabolites Estriol, 2-Hydroxyestradiol, 4-Hydroxyestradiol, 2-Methoxyestradiol, 4-Methoxyestradiol, 2-Hydroxyestrone,

4-Hydroxyestrone, 16α-Hydroxyestrone, 2-Methoxyestrone, 4-Methoxyestrone, and 16-Epiestriol. The analytical protocol and the chromatographic programming were optimized using a multivariate Design of Experiment (DoE) approach (Leardi, 2009) and the protocol was fully validated.

In the post-menopausal age, the probability of developing ovarian, breast, and uterine cancers increases substantially (Surakasula, Nagarjunapu, & Raghavaiah, 2014). For most of the estrogens monitored in the present study, the existing literature does not report the range of their physiological concentrations. Therefore, the new method was initially applied to define the urinary estrogens' concentration range for a population of 138 post-menopausal healthy women, recruited within a planned screening program. The results reported in this study will be exploited in future investigations aimed at the multivariate characterization of the urinary estrogenic profile of women with breast or ovarian cancer.

2. Experimental

2.1. Chemicals and Reagents

17α-/17β-Estradiol, ascorbic acid, methanol, ethyl acetate, hexane, ammonium iodide, dithioerythritol, *tert*-butyl methyl ether (TBME), *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA), and β-glucuronidase/arylsulfatase (from *Helix pomatia*) mixture were purchased from Sigma-Aldrich (Milan, Italy). Estrone, Estrone 3-(β-D-glucuronide) sodium salt, 2-Hydroxyestrone, Estriol, Estrone-d_4 and 17β-Estradiol-d_4 were purchased from LGC Promochem SRL (Milan, Italy). The sulfate sodium salt of Estrone was supplied by Steraloids Inc. (Newport, RI, USA). All the other standards were purchased from Toronto Research Chemical Inc. (Toronto, ON, Canada). β-Glucuronidase from *Escherichia coli* was supplied by Roche Life Science (Indianapolis, IN, USA). A Milli-Q® UF-Plus apparatus (Millipore, Bedford, MA, USA) was used to obtain Ultra-pure water. C-18 endcapped Solid-Phase Extraction (SPE) cartridges were obtained from UCT Technologies (Bristol, PA, USA).

All stock standard solutions (for both the analytes and the internal standards) were prepared in methanol at 1 mg/mL and stored at 20 °C until use. Two working solutions, at the concentrations of 20 µg/mL and 1 µg/mL, containing the 14 analytes, were obtained by appropriate dilution with methanol. The isotopically labelled Estrone-d_4 and 17β-estradiol-d_4 were added from separate methanol working solutions at the final concentrations of 100 µg/mL and 50 µg/mL, respectively.

2.2.Analytical protocol

The analytical protocol was optimized starting from a method previously developed for the detection of an extended urinary androgenic profile (Alladio et al., 2016; Amante et al., 2018). The Design of Experiment (Leardi, 2009) study was performed using the Chemometric Agile Tool (CAT) developed by Leardi et al. (Gruppo di Chemiometria Italiana, n.d.) in R environment ("R: The R Project for Statistical Computing," 2018).

2.3.Optimization of the analytical protocol

2.3.1. Optimization of the sample preparation

Two critical factors were investigated: (i) the composition of the liquid-liquid extraction (LLE) solvent and (ii) the drying temperature of the extract. The LLE procedure involves two consecutive extractions, with 5 mL solvent each. The following four combinations of solvents were tested: (i) hexane / hexane; (ii) *terz*-butyl methyl ether (TBME) / hexane; (iii) TBME / TBME; (iv) ethyl acetate + hexane (2:3 v/v) mixture (twice), which were coded as -1.5, -0.5, +0.5, +1.5, respectively. With regard to the drying temperature of the extracts, the levels of the experimental domain were set at 40°C, 50°C, and 60°C, coded as -1, 0, +1, respectively. A total of 36 experiments (4×3×3, i.e. the levels of the full-factorial design, with each experiment evaluated in triplicate) were performed in random order. A model described by the following equation was calculated:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2$$

where y represents the intensity of the peaks defined by means of the Total Ion Current (TIC) collected by the MS detector, b_i represent the regression coefficients for the corresponding factors x_i (0 = intercept, 1 = extraction solvent, 2 = drying temperature). The interaction (x_1x_2) and quadratic terms for both LLE solvent (x_1^2) and drying temperature (x_2^2) were also evaluated.

2.3.2 Optimization of the chromatographic run

The optimization of the oven temperature programming was performed by applying a Design of Experiment strategy involving a face-centered central composite design (FCCD). The levels for all the investigated factors were coded from -1 to 1 (including the the central value 0,0,0). The following conditions, relative to the central oven temperature ramp, were varied, as follows: (i) initial temperature (range: 225–235 °C); (ii) temperature ramp (range: 3–7 °C/min); (iii) final temperature (range: 250–260 °C). The middle levels., coded as 0, were set at the central point of the ranges, i.e. 230 °C for the initial temperature, 5 °C/min for the temperature ramp, and 255°C for the final temperature. Every experiment was executed in triplicate randomly, resulting in 45 total experiments. A model analogous to the one described above was optimized. The DoE results are detailed in the Results section. The optimal starting and final temperatures turned to be equal to 225°C and 250°C, respectively, with a ramp of 3°C/min. Since all the analytes were eluted before the final temperature was reached, the latter was reduced to 245 °C. Moreover, a 3 min holding was added to the oven program at 234 °C in order to achieve better separation of 2-methoxyestrone and 16α-Hydroxyestrone. The optimized chromatographic run was completed in less than 20 minutes and the retention times of the target analytes lied between 6.58 min (17α-Estradiol) and 10.50 min (16-Epiestriol). Figure 2 shows the SIM chromatogram obtained from a blank urine sample fortified with all the target analytes at the concentration of 50 ng/mL.

2.4. Analytical protocol

In the optimized protocol, 6 mL of urine were fortified with both 17β-estradiol-d_4 and estrone-d_4 at the final concentrations of 50 ng/mL and 25 ng/mL, respectively. Then, 2 mL of acetate buffer 1.1 M (pH 5.5) was added, together with some drops of HCl 10 N, if necessary, to reach the final pH of 5.5. Subsequently, 50 µL ascorbic acid 1 M was added. This step is aimed to prevent the degradation of the labile catechol groups (Kim, Lee, Chung, Pyo, & Lee, 2014; Knust et al., 2007; Moon et al., 2011). A deconjugation step, useful to transform the conjugated estrogens into the free form, was executed by adding 20 µL of β-glucuronidase/arylsulfatase mixture, followed by the incubation of the urine samples at 37 °C overnight (Bellem, Meiyappan, Romans, & Einstein, 2011; Samavat & Kurzer, 2015; Ziegler, Fuhrman, Moore, & Matthews, 2015). The next morning, a second deconjugation step was performed using 100 µL β-glucuronidase from *Escherichia Coli* together with 50 µL of ascorbic acid solution. The reaction was performed at 58 °C for 1 hour. Once the hydrolysis was completed, the mixture was cooled to room temperature. The extraction was performed at basic pH (above the value of 9), reached by adding 2 mL of 0.1 M carbonate buffer (pH 9) and some drops of NaOH 1 M. The liquid-liquid extraction (LLE) was performed twice for each sample: 5 mL of ethyl acetate / hexane mixture (2:3 v/v) were added, the tubes were shaken in a vortex multimixer (Tecnovetro, Monza, Italy) for 5 minutes, and centrifugated (Megafuge 1.0 Heraeus; ASHI, Milan, Italy) at 4000 rpm for 5 min. The two combined organic phases were evaporated to dryness under a gentle stream of nitrogen at 40 °C (Techne Sample Concentrator, Barloworld Scientific, Stone, UK). The dried residue was reconstituted using 50 µL of MSTFA/NH_4I/dithioerythritol (1.000:2:4 v/w/w) derivatizing solution. The derivatization was performed at 70 °C for 1 hour. An aliquot of 2 µL was injected into the GC/MS system in the splitless mode. The oven temperature was programmed as follows: the starting temperature of 200°C was held for 2 minutes; then, the temperature was raised to 225°C with a ramp of 8 °C/min; a slower ramp of 3 °C/min was applied to reach the temperature of 234°C, which was held for 3 minutes. The final temperature of 315°C was reached with a ramp of 40°C/min and held for further 3 minutes. The overall run time was equal to 16.15 minutes.

The retention times and the target ions of the analytes and internal standards are reported in Table 1. The chromatogram from a female urine sample is reported in Supplementary Materials, Figure S1.

2.5.Validation of the analytical method

The analytical method was validated according to the ISO/IEC 17025:2017 criteria and recommendations: in particular, 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 parameters were evaluated. Blank urine samples were collected from healthy male volunteers (laboratory personnel), then pooled, and finally loaded on solid phase extraction (SPE) cartridges to remove all the possible estrogen traces, while keeping most of the non-steroidal matrix components. These blank samples were spiked with standard solutions of the target analytes to obtain the desired concentrations.

2.5.1. Linearity

For linearity study, a total of 30 experiments were performed, consisting in five replicates of the six calibration points, in the range 1–50 ng/mL (i.e., 1, 2, 5, 10, 25, 50 ng/mL). Statistical tests were executed to evaluate the linearity of data-point distributions, including analysis of variance (ANOVA) test, lack-of-fit test, Mandel's test, and evaluation of the relative standard deviation (RSD) of the slope, according to the approach described elsewhere (Desharnais, Camirand-lemyre, Mireault, & Skinner, 2017; Desharnais, Camirand-Lemyre, Mireault, & Skinner, 2017). Furthermore, the residual plots and the deviation from back-calculated concentrations were examined. The ANOVA, lack-of-fit, and Mandel tests were considered as passed when the calculated values (F_{exp}) were lower than the corresponding critical value at $\alpha = 0.05$ significance level. The specific critical values are reported in Table 2. For the RSD slope and back calculation tests, the threshold values of 5% and 20% were respectively adopted. The analysis of the residuals plot revealed the structure of the data, if homoscedastic or heteroscedastic. In the case of heteroscedastic distribution of data-points, a

weighting factor of 1/x or $1/x^2$ was employed, depending on the rate of the variance increase with the concentration (linear or quadratic). For the choice of the best weight and for the computation of the calibration model, an R-routine developed by Desharnais et al. was employed (Desharnais, Camirand-lemyre, et al., 2017; Desharnais, Camirand-Lemyre, et al., 2017).

2.5.2. LOD and LOQ

The Hubaux-Vos algorithm was used to estimate the limit of detection (LOD) and of quantitation (LOQ) for all the analytes (Hubaux & Vos, 1970). For this computation, the same 30 data-points collected to build the calibration curves were used. To experimentally confirm the correct estimation, blank samples were spiked at concentrations values corresponding to the computed LOD and LOQ values. In the operational practice, LOQ values were assumed at the lower level of the calibration curves.

2.5.3. Precision and accuracy

Intra-day precision and accuracy were evaluated on 10 blank urine samples spiked with all the target analytes at three concentration levels (i.e., 1.0 ng/mL, 5.0 ng/mL and 25 ng/mL). Accuracy and precision were estimated from the percent bias (bias%) and the percent variation coefficient (CV%), respectively.

2.5.4. Matrix effect and extraction recovery

The matrix effect was evaluated at the same three concentration levels selected for precision and accuracy. The experimental results obtained from blank urine samples (mean value from five replicates) were compared to those obtained from blank deionized water solution. Both the blank urine and the blank deionized water were spiked after the extraction step. The matrix effect for each target analyte was expressed as the percentage ratio between the two measured concentrations. The extraction recovery was calculated (at the same concentration levels, with five replicates) by

comparing the results obtained from blank urine samples spiked respectively before and after the extraction step and expressed as the percentage ratio between the two data.

2.5.5. Enzyme performance

The deconjugation efficiency of β-glucuronidase and arylsulfatase was tested at three concentration levels (1.0 ng/mL, 5.0 ng/mL and 25 ng/mL). The percentage ratio between the recovered concentrations of Estrone glucuronide and Estrone sulfate spiked into a blank sample and those of the free estrone spiked to another blank sample at the same concentration was calculated. All the analyses were performed in duplicate.

2.5.6. Repeatability, specificity and selectivity

The retention time repeatability was verified on the 30 experiments performed for the linearity study, Deviations below 1% were considered satisfactory. The relative ion abundance repeatability was evaluated on the selected ion chromatograms for each target analyte. The variations were considered acceptable within ±20%, with respect to the controls.

Ten blank urine samples were analyzed and the signal–to–noise ratio (S/N) was measured on the selected ion chromatograms at the retention times of all the analytes of interest and a S/N < 3 was considered satisfactory to verify the method's specificity. Furthermore, the presence of possible interfering compounds from endogenous substances or derivatization by-products was tested at the retention times of the target analytes.

2.5.7. Carry-over

The injection of a blank non-spiked sample after the highest calibration level was used to evaluate the carry-over effect. It was considered negligible if the S/N ratio was lower than 3 at the analytes retention time.

2.6.Subjects recruitment

The urine samples were collected from healthy post-menopausal volunteer women. Every woman included in the project was informed verbally and in writing and signed an informed consent. The information reported were: hormone replacement therapy, familiarity for uterus and/or mammalian cancer, fertility, smoke habits and pharmacological therapies.

3. Results and discussion

3.1. Optimization of the analytical procedure

3.1.1. Optimization of the sample preparation protocol

The values of b_i coefficients and their statistical significance, as determined by t-tests, are reported in Figure 3a. The extraction solvent proved to be the only significant factor (i.e. p-value < 0.05), while the drying temperature, interaction, and quadratic terms of the regression model did not show any significant influence. The two-and three-dimensional response surfaces reported in Figure 3b-3c indicate that the maximum response value corresponds to $x_1 < 1.5$ and $x_2 = -1$. Consequently, the use of TBME or ethyl acetate + hexane mixture (2:3 v/v) for both extraction steps provided comparable results. In the end, the best combination of drying temperature (found at 40°C, i.e. $x_2 = -1$) and extraction solvent involved the use of ethyl acetate + hexane mixture (2:3 v/v) corresponding to higher resolution and intensity of the chromatographic peaks with respect to TBME.

3.1.2. *Optimization of the chromatographic run*

Figure 3d depicts the histogram for the b_i coefficients and their statistical significance. The only significant factor appears to be the temperature ramp (p-value < 0.05) with a negative trend (i.e., the lower the ramp, the higher the response). The two and three-dimensional response surfaces confirmed this trend. The maximum response value lied at the point encoded as [-1, -1, 0] indicating optimal conditions for the following experimental values: initial and final temperature 225 °C and 250°C,

respectively, while the best temperature ramp was 3 °C/min. The final temperature was subsequently reduced to 245 °C, while a 3-min step at constant 234 °C was introduced in the oven program.

Thanks to the dedicated DoE model for the oven temperature programming, excellent chromatographic separation among the fourteen analytes was achieved (Figure 2), despite their similar structures and physico-chemical properties, among which very close chromatographic properties. Effective separation was particularly important for the accurate quantification of each analyte, since their mass spectra share several signals (both among and beyond the characteristic ions reported in Table 1), producing a variety of interferences and possibly resulting in inaccurate determinations.

3.2. Method validation

Within the 1-50 ng/mL calibration range, all data distributions turned out heteroskedastic, making the use of weighting highly recommendable. Depending on the extent of heteroskedasticity, the estrogens models either used an x^{-2} (for most of the analytes) or a x^{-1} weighting correction (17α-Estradiol, 2-Methoxyestradiol, Estrone, 4-Methoxyestrone, and Estriol). Linear response with concentration was recorded for 7 out of 14 of the analytes, with the exclusion of 17α-Estradiol, 17β-Estradiol, 2-Hydroxyestradiol, 2-Methoxyestradiol, 4-Methoxyestradiol, 2-Hydroxyestrone and 16-Epiestriol for which a quadratic calibration model was adopted. Mandel's test confirmed these outcomes. Furthermore, the lack-of-fit test was not passed for 17α-Estradiol, 17β-Estradiol and 16-Epiestriol, thus confirming that the variance explained by the linear calibration is larger than the residual variance, so that a higher (i.e. quadratic) term should be introduced to fit the data better than traditional linear approach. Moreover, the residual plots were evaluated, and the calibration models were confirmed by the presence of random residuals patterns along the concentration ranges for all the analytes. The variability of the back-calculated concentrations was evaluated too, aiming to compare directly the back-calculated concentrations with the theoretical values of the calibration standards. The results turned lower than 20% for all the target analytes, thus showing the goodness

of our methodology to provide accurate measurements in the future for real unknown samples. Furthermore, RSD slope values turned to be lower than the critical 5% limit for all the tested analytes. This result confirmed once again the occurrence of a small dispersion of the experimental data around the calibration curves. All the results are reported in Table 2 together with the calculated limit of detection (LOD) and limit of quantitation (LOQ) values. The mass spectra of all the analytes at the lower limit of quantitation (LLQ) and at the upper limit of quantitation (ULQ), plus the mass spectra of the internal standards are reported in Supplementary Materials, Figures S2-S16.

Precision data were considered satisfactory, since the CV% values lied below 25% for the low calibration level and below 15% for the other levels. Satisfactory accuracy results were also achieved, with experimental average concentrations lying within ±15% from the expected value (Table 3).

Likewise, matrix effect and the extraction recovery results were fully adequate, as their values were uniformly within ± 10% from the expected 100% and none of them exceed the planned threshold of ± 15% (Table 3).

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.

The repeatability, selectivity and sensitivity turned to be good for all the analytes. Finally, no carry-over effect was noticed, since a S/N lower than 3 was observed in the blank sample injected after the highest concentrate calibration point for all the target compounds.

3.3. Variability interval in a population of healthy women

A total of 138 samples from post-menopausal women were collected and analyzed. The age of the recruited women was 61.4±5.5 years and the body mass index was 26.6±6.0. For all the analytes, the median value is reported, together with the first decile, the first quartile, the third quartile and the ninth decile values for each target compound (Table 4), representing the main reference values of the distributions of the investigated population. A visual representation (e.g. boxplots) of the distribution

of concentrations is reported in Supplementary Material, Figure S17. The median concentration for all the estrogens appear to be quite low with respect to the mean values recorded for the women in the fertile age, even considering - for the latter - the large variability associated with the menstrual cycle (Bozzolino et al., 2019). In fact, the medians detected for post-menopausal women were close to the lower calibration level, namely between 1 ng/mL and 10 ng/mL). The 2-hydroxy and 2-methoxy derivatives of 17β-estradiol and estrone resulted homogeneously more concentrated than the equivalent 4-metabolites. This is in accordance with the predominance of the hydroxylation in correspondence of C2 with respect to the equivalent reaction in C4 position (Moon et al., 2011; Zhu & Lee, 2005). Moreover, estradiol metabolites were consistently less concentrated than the estrone equivalents, coherently with the greater role of estrogens in postmenopausal age (Cui et al., 2013) and in contrast with what is observed in the pre-menopausal age (Bozzolino et al., 2019).

4. Conclusions

An analytical method based on GC-MS was developed for the simultaneous quantification of 14 urinary estrogens, thanks to two design of experiments models used to optimize the sample pretreatment protocol and the chromatographic method. In particular, optimal separation was achieved, overcoming the potential interferences among analytes sharing very similar mass spectra and chromatographic properties. The method was fully validated following a rigorous protocol that allows the statistical evaluation of several parameters with a relatively restricted set of experiments, including homoscedasticity, linearity of the calibration range, LOD, LOQ, accuracy, and precision. Further experiments were executed to evaluate the potential matrix effect, extraction recovery and the deconjugation enzyme efficiency.

The method was applied to the determination of the estrogenic profile in urine samples collected from 138 allegedly healthy post-menopausal women during a screening protocol. The results, substantially in accordance with literature data (Cui et al., 2013; Moon et al., 2011; Zhu & Lee, 2005), form a

preliminary data-set useful to establish a physiological “normality range” to be matched with possible “pathological” profiles, using multivariate statistical strategies.

Acknowledgments

Financial support from the Italian “Ministero dell’Istruzione, dell’Università e della Ricerca” within a PRIN 2017 call for bids (Research Projects of Relevant National Interest—grant 2017Y2PAB8) is gratefully acknowledged.

References

Alladio, E., Caruso, R., Gerace, E., Amante, E., Salomone, A., & Vincenti, M. (2016). Application of multivariate statistics to the Steroidal Module of the Athlete Biological Passport: A proof of concept study. *Analytica Chimica Acta*, *922*, 19–29. https://doi.org/10.1016/j.aca.2016.03.051

Amante, E., Alladio, E., Salomone, A., Vincenti, M., Marini, F., Alleva, G., … Porpiglia, F. (2018). Correlation between chronological and physiological age of males from their multivariate urinary endogenous steroid profile and prostatic carcinoma-induced deviation. *Steroids*, *139*, 10–17. https://doi.org/10.1016/j.steroids.2018.09.007

Bellem, A., Meiyappan, S., Romans, S., & Einstein, G. (2011, October). Measuring estrogens and progestagens in humans: An overview of methods. *Gender Medicine*. Elsevier. https://doi.org/10.1016/j.genm.2011.07.001

Berrino, F., Muti, P., Micheli, A., Krogh, V., Pisani, P., & Secreto, G. (1996). Serum Sex Hormone Levels After Menopause and Subsequent Breast Cancer Results The characteristics of the breast cancer. *Journal of the National Cancer Institute*, *88*(5), 291–296.

Bozzolino, C., Vaglio, S., Amante, E., Alladio, E., Gerace, E., Salomone, A., & Vincenti, M. (2019). Individual and cycling estrogenic profile in women: Structure and variability of the data. *Steroids*, *150*, 1–10.

Bradlow, H. L., Davis, D. L., Lin, G., Sepkovic, D., & Tiwari, R. (1995). Effects of pesticides on the ratio of 16α/2-hydroxyestrone: A biologic marker of breast cancer risk. *Environmental Health Perspectives*, *103*(SUPPL. 7), 147–150. https://doi.org/10.2307/3432525

Bulun, S. E., Chen, D., Lu, M., Zhao, H., Cheng, Y., Demura, M., … Lin, Z. (2007). Aromatase excess in cancers of breast, endometrium and ovary. *Journal of Steroid Biochemistry and Molecular Biology*, *106*(1–5), 81–96. https://doi.org/10.1016/j.jsbmb.2007.05.027

Cui, J., Shen, Y., & Li, R. (2013). Estrogen synthesis and signaling pathways during ageing. *Trends in Molecular Medicine*, *19*(3), 197–209. https://doi.org/10.1016/j.molmed.2012.12.007.Estrogen

De Luca, S., Amante, E., Fiori, C., Alleva, G., Alladio, E., Marini, F., … Porpiglia, F. (2020). Prospective evaluation of urinary steroids and prostate carcinoma-induced deviation: preliminary results. *Minerva Urologica e Nefrologica*. https://doi.org/10.23736/S0393-2249.19.03529-X

Desharnais, B., Camirand-lemyre, F., Mireault, P., & Skinner, C. D. (2017). Procedure for the Selection and Validation of a Calibration Model I — Description and Application. *Journal of Analytical Toxicology*, *41*, 261–268. https://doi.org/10.1093/jat/bkx001

Desharnais, B., Camirand-Lemyre, F., Mireault, P., & Skinner, C. D. (2017). Procedure for the selection and validation of a calibration model II-theoretical basis. *Journal of Analytical Toxicology*, *41*(4), 269–276. https://doi.org/10.1093/jat/bkx002

Faupel-Badger, J. M., Fuhrman, B. J., Xu, X., Falk, R. T., Keefer, L. K., Veenstra, T. D., … Ziegler, R. G. (2010). Comparison of liquid chromatography-tandem mass spectrometry, RIA, and ELISA methods for measurement of urinary estrogens. *Cancer Epidemiology Biomarkers and Prevention*, *19*(1), 292–300. https://doi.org/10.1158/1055-9965.EPI-09-0643

Greenlee, H., Chen, Y., Kabat, G. C., Wang, Q., Kibriya, M. G., Gurvich, I., … Ahsan, H. (2007). Variants in estrogen metabolism and biosynthesis genes and urinary estrogen metabolites in women with a family history of breast cancer. *Breast Cancer Research and Treatment*, *102*(1), 111–117. https://doi.org/10.1007/s10549-006-9308-7

Gruppo di Chemiometria Italiana. (n.d.). http://gruppochemiometria.it/index.php/software.

Hoffmann, P., Hartmann, M. F., Remer, T., Zimmer, K. P., & Wudy, S. A. (2010). Profiling

oestrogens and testosterone in human urine by stable isotope dilution/benchtop gas chromatography-mass spectrometry. *Steroids*, *75*(13–14), 1067–1074. https://doi.org/10.1016/j.steroids.2010.06.014

Hubaux, A., & Vos, G. (1970). Decision and Detection limits for linear Calibration Curves. *Analytical Chemistry*, *42*(8), 849–855. https://doi.org/10.1021/ac60290a013

Kim, E. J., Lee, D., Chung, B. C., Pyo, H., & Lee, J. (2014). Association between urinary levels of bisphenol-A and estrogen metabolism in Korean adults. *Science of the Total Environment*, *470–471*, 1401–1407. https://doi.org/10.1016/j.scitotenv.2013.07.040

Knust, U., Strowitzki, T., Spiegelhalder, B., Bartsch, H., & Owen, R. W. (2007). Optimization of an isotope dilution gas chromatography/mass spectrometry method for the detection of endogenous estrogen metabolites in urine samples. *Rapid Communications in Mass Spectrometry*, *21*(14), 2245–2254. https://doi.org/10.1016/j.jcrs.2012.09.017

Leardi, R. (2009). Experimental design in chemistry: A tutorial. *Analytica Chimica Acta*, *652*(1–2), 161–172. https://doi.org/10.1016/j.aca.2009.06.015

Lippert, T. H., Seeger, H., & Mueck, A. O. (2000). The impact of endogenous estradiol metabolites on carcinogenesis. *Steroids*, *65*(7), 357–369. https://doi.org/10.1016/S0039-128X(00)00101-X

Mason, J. I. (2002). *Genetics of steroid biosynthesis and function. Modern genetics* (Vol. 6). New York: Taylor & Francis.

Moon, J.-Y., Kim, K. J., Moon, M. H., Chung, B. C., & Choi, M. H. (2011). A novel GC-MS method in urinary estrogen analysis from postmenopausal women with osteoporosis. *Journal of Lipid Research*, *52*(8), 1595–1603. https://doi.org/10.1194/jlr.D016113

Obi, N., Vrieling, A., Heinz, J., & Chang-Claude, J. (2011). Estrogen metabolite ratio: Is the 2-hydroxyestrone to 16α-hydroxyestrone ratio predictive for breast cancer? *International Journal*

of Women's Health, *3*(1), 37–51. https://doi.org/10.2147/IJWH.S7595

R: The R Project for Statistical Computing. (2018).

Samavat, H., & Kurzer, M. S. (2015). Estrogen metabolism and breast cancer. *Cancer Letters*. https://doi.org/10.1016/j.canlet.2014.04.018

Simpson, E., Rubin, G., Clyne, C., Robertson, K., O'Donnell, L., Davis, S., & Jones, M. (1999). Local estrogen biosynthesis in males and females. *Endocrine-Related Cancer*, *6*(2), 131–137. https://doi.org/10.1677/erc.0.0060131

Surakasula, A., Nagarjunapu, G., & Raghavaiah, K. (2014). A comparative study of pre- and post-menopausal breast cancer: Risk factors, presentation, characteristics and management. *Journal of Research in Pharmacy Practice*, *3*(1), 12–18. https://doi.org/10.4103/2279-042x.132704

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

Zhu, B. T., & Lee, A. J. (2005). NADPH-dependent metabolism of 17β-estradiol and estrone to polar and nonpolar metabolites by human tissues and cytochrome P450 isoforms. *Steroids*, *70*(01), 225–244. https://doi.org/10.1016/j.steroids.2005.01.002

Ziegler, R. G., Fuhrman, B. J., Moore, S. C., & Matthews, C. E. (2015, July). Epidemiologic studies of estrogen metabolism and breast cancer. *Steroids*. Elsevier. https://doi.org/10.1016/j.steroids.2015.02.015

Table 1. Retention time (t_R), monitored characteristic ions and their relative abundances (in brackets) of the target compounds and the internal standards.

Analyte	Internal Standard	t_R (min)	Quantifier (m/z)	Qualifier 1 (m/z)	Qualifier 2 (m/z)
17α-Estradiol	**17β-Estradiol-d_4**	6.58	416 (100%)	285 (96%)	326 (13%)
Estrone	**Estrone-d_4**	6.73	414 (100%)	399 (79%)	309 (28%)
17β-Estradiol	**17β-Estradiol-d_4**	6.97	416 (100%)	285 (69%)	232 (54%)
4-Methoxyestrone	**Estrone-d_4**	7.76	444 (100%)	429 (25%)	261 (32%)
4-Methoxyestradiol	**17β-Estradiol-d_4**	8.02	446 (100%)	315 (47%)	416 (9%)
16α-Hydroxyestrone	**Estrone-d_4**	8.27	286 (100%)	430 (13%)	415 (6%)
2-Methoxyestrone	**17β-Estradiol-d_4**	8.29	444 (100%)	261 (13%)	429 (11%)
2-Methoxyestradiol	**17β-Estradiol-d_4**	8.56	315 (100%)	416 (18%)	431 (10%)
2-Hydroxyestrone	**Estrone-d_4**	8.73	502 (100%)	293 (22%)	306 (12%)
2-Hydroxyestradiol	**17β-estradiol-d_4**	8.98	504 (100%)	373 (8%)	413 (2%)
4-Hydroxyestrone	**Estrone-d_4**	9.16	502 (100%)	293 (14%)	306 (10%)
4-Hydroxyestradiol	**17β-Estradiol-d_4**	9.49	504 (100%)	373 (20%)	489 (5%)
Estriol	**17β-Estradiol-d_4**	10.02	345 (100%)	504 (54%)	386 (50%)
16-Epiestriol	**17β-Estradiol-d_4**	10.50	345 (100%)	504 (52%)	386 (39%)

Internal Standard	t_R (min)	Quantifier (m/z)	Qualifier 1 (m/z)	Qualifier 2 (m/z)
Estrone-d_4	6.72	417 (100%)	402 (78%)	-
17β-Estradiol-d_4	6.94	420 (100%)	287 (95%)	-

Table 2. Validation parameters and results relative to the evaluation of the calibration curves for all the target analytes in the dynamic range of calibration (e.g., 1.0 – 50.0 ng/mL), as follows: coefficient of determination (R^2), limit of detection (LOD), limit of quantitation (LOQ), lack-of-fit, ANOVA and RSD slope tests, back-calculation results, Mandel's test experimental values, type of model, relative weights, and calibration model equation. The critical values of the significance tests are reported, too.

Analyte	Correlation coefficient (R^2)	LOD (LOQ) (ng/mL)	Lack-of-fit test (F_{exp})[a]	ANOVA (F_{exp})[b]	RSD slope test (%)[c]	Back calculation test (%)[d]	Mandel's test (F_{exp})[e]	Model	Weight	Calibration model equation
17α-Estradiol	0.9977	0.33 (0.67)	5.31	1.62	2.37	17	18.81	Quadratic	x^{-1}	$0.0964x^2 + 0.9312x + 0.0521$
17β-Estradiol	0.9985	0.27 (0.55)	3.73	1.39	1.94	15	10.30	Quadratic	x^{-2}	$0.0615x^2 + 0.8662x + 0.0102$
2-Hydroxyestradiol	0.9985	0.40 (0.80)	1.15	1.02	1.97	16	4.67	Quadratic	x^{-2}	$-0.1266x^2 + 2.4425x - 0.1070$
4-Hydroxyestradiol	0.9976	0.35 (0.69)	2.26	1.18	2.45	19	1.17	Linear	x^{-2}	$1.8495x - 0.1117$
2-Methoxyestradiol	0.9985	0.28 (0.55)	1.12	1.02	1.97	16	4.67	Quadratic	x^{-1}	$0.1075x^2 + 1.3862x + 0.0535$
4-Methoxyestradiol	0.9988	0.25 (0.49)	1.23	1.03	1.75	13	4.63	Quadratic	x^{-2}	$0.0696x^2 + 1.0797x + 0.0191$
Estrone	0.9980	0.31 (0.63)	0.77	0.97	2.23	14	1.15	Linear	x^{-1}	$1.6611X + 0.0289$
2-Hydroxyestrone	0.9985	0.27 (0.54)	1.07	1.01	1.92	15	3.30	Quadratic	x^{-2}	$-0.3604x^2 + 3.7179x + 0.0358$
4-Hydroxyestrone	0.9966	0.41 (0.83)	1.37	1.05	2.94	19	2.29	Linear	x^{-2}	$0.4676x + 0.0073$

16α-Hydroxyestrone	0.9988	0.24 (0.49)	0.17	0.88	1.74	8	1.36	Linear	x^{-2}	5.9750x + 0.0918
2-Methoxyestrone	0.9993	0.19 (0.38)	0.10	0.87	1.34	16	1.01	Linear	x^{-2}	2.8365x + 0.0542
4-Methoxyestrone	0.9993	0.19 (0.38)	0.29	0.90	1.36	15	1.16	Linear	x^{-1}	1.9278x + 0.0501
Estriol	0.9990	0.22 (0.44)	0.20	0.89	1.56	11	1.48	Linear	x^{-1}	0.3282x + 0.0126
16-Epiestriol	0.9983	0.29 (0.58)	3.09	1.30	2.04	15	11.00	Quadratic	x^{-2}	$0.0387x^2$ + 0.4659x + 0.0101

[a] F_{crit} = 2.776 (n_1 = 4 and n_2 = 24 degrees of freedom)
[b] F_{tab} = 3.842 (n_1 = 1 and n_2 = 28 degrees of freedom)
[c] %RSD threshold = 5.00%
[d] % threshold = 20%
[e] F_{crit} = 2.572 (n_1 = 1 and n_2 = 27 degrees of freedom)

Table 3. Intra-day precision (CV%), accuracy (bias%), matrix effect and recovery for each analyte tested.

Analyte	Precision (CV%)			Accuracy (bias%)			Matrix effect (%)			Recovery (%)		
	1 ng/mL	5 ng/mL	25 ng/mL	1 ng/mL	5 ng/mL	25 ng/mL	1 ng/mL	5 ng/mL	25 ng/mL	1 ng/mL	5 ng/mL	25 ng/mL
17α-Estradiol	13	9.7	7.3	+7.9	+2.7	-6.2	+8.1	+10	+4.4	106	87	94
17β-Estradiol	7.5	1.4	4.1	+7.6	+5.3	-1.1	+3.1	+1.8	+4.4	101	98	100
2-Hydroxyestradiol	13	9.4	3.5	+12	-2.6	+5.6	-6.4	-2.1	-5.6	105	97	110
4-Hydroxyestradiol	7.3	8.0	1.4	+6.5	-11	+2.3	-1.8	-1.1	-0.3	108	90	101
2-Methoxyestradiol	8.1	6.6	1.5	+10	+6.1	-0.5	+2.2	+0.1	+2.0	104	99	98
4-Methoxyestradiol	4.9	0.4	1.4	+2.0	+3.5	-0.05	+1.8	+0.6	+3.7	100	99	98
Estrone	13	5.6	1.2	+7.3	+1.8	+3.6	-3.9	-7.3	-2.1	107	97	102
2-Hydroxyestrone	12	3.4	1.7	-8.2	+5.0	-2.1	-9.9	-8.6	-1.2	93	100	99
4-Hydroxyestrone	9.7	4.1	3.0	-7.8	+6.8	-1.0	-5.0	-9.2	-4.8	89	107	99
16α-Hydroxyestrone	13	4.4	5.0	-6.6	-2.0	-2.2	+15	+8.1	+6.3	95	99	101
2-Methoxyestrone	11	7.0	3.8	+8.9	-2.1	+1.7	+8.5	+10	+3.1	93	100	101
4-Methoxyestrone	12	9.3	7.1	+3.6	+4.8	-4.1	-12	-2.4	-0.1	99	104	100
Estriol	13	9.2	2.9	+9.6	+4.6	+1.5	-9.5	-0.6	-1.5	95	102	96
16-Epiestriol	1.3	1.9	2.0	+3.7	+5.1	-0.5	-0.2	-1.1	-3.7	102	95	97

Table 4. Values of the first decile, first quartile, median, third quartile and ninth decile of the distributions of the 14 monitored estrogens.

Target analyte	1st decile (ng/mL)	1st quartile (ng/mL)	Median (ng/mL)	3rd quartile (ng/mL)	9th decile (ng/mL)
17α-Estradiol	1.74	1.75	1.83	2.00	2.36
17β-Estradiol	2.00	2.05	2.12	2.27	3.49
2-Hydroxyestradiol	4.24	4.24	4.35	4.71	5.22
4-Hydroxyestradiol	2.18	2.21	2.26	2.47	2.96
2-Methoxyestradiol	6.25	6.30	6.39	6.69	7.54
4-Methoxyestradiol	1.82	1.84	1.94	2.07	2.73
Estrone	1.32	1.48	1.79	2.56	3.95
2-Hydroxyestrone	1.32	5.14	10.55	19.86	67.04
4-Hydroxyestrone	5.36	5.72	5.83	10.26	19.55
16α-Hydroxyestrone	1.17	1.29	1.39	1.67	2.61
2-Methoxyestrone	6.03	6.35	9.57	32.64	137.35
4-Methoxyestrone	3.34	3.96	4.43	8.13	73.28
Estriol	1.94	2.18	2.64	3.55	4.76
16-Epiestriol	3.83	3.83	3.92	5.32	8.17