ENDOCRINOLOGY

# Reproducibility over time of measurements of androgens, estrogens and hydroxy estrogens in urine samples from post-menopausal women

S. Rinaldi[1], C.N. Moret[2], R. Kaaks[1], C. Biessy[1], M.S. Kurzer[3], H. Déchaud[4], P.H.M. Peeters[2] & P.A.H. van Noord[2]

[1]*International Agency for Research on Cancer, Hormones and Cancer Group, 69372 Lyon, France;* [2]*Julius Center for General Practice and Patient Oriented Research, University Medical Center, 3508 GA Utrecht , The Netherlands;* [3]*Department of Food Science and Nutrition, University of Minnesota, St. Paul, 55108 Minnesota, USA;* [4]*Service de Radioanalyse et Radiopharmacie, Centre de Médecine Nucléaire, Hôpital Neurocardiologique, 69003 Lyon, France*

**Abstract.** Sex steroid concentrations in urine samples from post-menopausal women have been associated with risk of various chronic diseases. The basic requirement for the assessment of risk in such large-scale epidemiological studies is that subjects be ranked accurately by their average, long-term hormone levels. We examined the reproducibility over time of measurements of urinary testosterone (T), 5α-androstane-3α, 17β-diol (ADIOL), estrone ($E_1$), estradiol ($E_2$), 2-hydroxy estrone and 2-hydroxy estradiol, (2(OH)-E), 16α-hydroxyestrone (16α(OH)-$E_1$) and the ratio of 2(OH)-E and 16α(OH)-$E_1$, in a representative sub-sample of post-menopausal women (n = 43) participating in an ongoing prospective cohort study. Women collected three first morning urine voids on different occasions, with average time difference between the first and the third urine sample of 5.1 years. T, ADIOL, $E_1$ and $E_2$ were measured by radio immunoassay after enzymatic hydrolysis, solid-phase extraction and HPLC purification of the samples, while 2(OH)-E and 16α(OH)-$E_1$ were assayed by solid-phase enzyme immunoassay after enzymatic hydrolysis. Intra-class correlation co-efficients (ICCs) over time were very good for T ($r = 0.85$), acceptable for $E_2$, $E_1$ and ADIOL ($r > 0.55$), but low for 2(OH)-E, 16α(OH)-$E_1$ and their ratio ($r < 0.46$). The adjustment for creatinine concentrations did not increase these correlations.

**Key words:** Androgens, Estrogens, Hydroxy estrogens, Post-menopausal women, Reproducibility, Urine

**Abbreviations:** ADIOL = 5α-androstane-3α, 17β-diol; $E_2$ = estradiol; $E_1$ = estrone; 2(OH)-E = 2-hydroxy estrone and 2-hydroxy estradiol; 16α(OH)-$E_1$ = 16α-hydroxyestrone; T = testosterone

## Introduction

Sex steroid concentrations in blood or urine of post-menopausal women have been associated with an increase in risk of various chronic diseases, including cancer [1, 2] and osteoporosis [3, 4]. Although blood samples have been used in many studies, urine samples may be collected in large-scale epidemiological studies for a number of reasons. First, contrary to blood withdrawal, urine collection is non-invasive and can be done directly by the study subjects, without medical assistance. Second, the collection of overnight or 24-hour urine samples may minimise variability of hormones that show a pulsatile secretion or diurnal variability (e.g., cortisol, growth hormone) [5, 6]. Finally, certain hormones can only be measured accurately in urine, because plasma levels are very low (e.g., 2/16α-hydroxy estrogen metabolites) [7, 8].

A difficulty inherent in epidemiological studies of hormones is that random errors in the measurement of individuals' hormone levels attenuate relative risk estimates and decrease the power of statistical tests for hormone–disease association [9]. The basic requirement for the accurate assessment of relative risk in epidemiological studies is that subjects be ranked accurately by their average, long-term hormone levels. While several studies have been published on the reproducibility and validity of hormone measurements in blood over time [10–13], few have addressed the reproducibility and validity of these measurements in urine samples [14, 15].

In preparation for a prospective study on breast cancer risk and sex-steroid concentrations in first morning voids from post-menopausal women in a Dutch prospective cohort, we examined the reproducibility over time of measurements of urinary testosterone (T), 5α-androstane-3α, 17β-diol (ADIOL), estrone ($E_1$), estradiol ($E_2$), 2-hydroxy estrone and 2-hydroxy estradiol (2(OH)-E), 16α-hydroxyestrone (16α(OH)-$E_1$) and the ratio between 2(OH)-E and 16α(OH)-$E_1$.

## Material and methods

### Subjects and urine collections

All women born between 1911 and 1945 living in Utrecht and surroundings were invited to participate in a large population based screening programme for early detection of breast cancer (the DOM-cohort) between 1975 and 1986 [16]. Women who responded to the first invitation to participate in this study were then invited for regular screening examinations. At each visit the women filled in a lifestyle questionnaire containing questions regarding breast cancer risk factors, medical history, exogenous hormone use and last day of the cycle. Furthermore, anthropometric measurements (height, weight, waist and hip circumferences) were taken for all participants, and the women were asked to bring a first morning urine sample on the day of their examination. Upon their arrival at the laboratory, urine samples were stored at $-20\,°C$ in 250-ml plastic polypropylene jars with a screw cap, without preserving agents.

Forty-three women were randomly selected from the 27,718 women participating in the DOM-cohort. The women were post-menopausal at recruitment (natural menopause defined as no menstruation during the previous 12 months and no surgical interventions), did not use exogenous hormones, and had provided at least three different urine samples.

The average time difference between the collection of the first (t1) and the second (t2) urine sample was 1.1 year (minimum 0.9 years, maximum 4.4 years), while it was 4.0 years between the second and the third (t3) sample (minimum 1.3 years, maximum 7.7 years). Urine samples were shipped on dry ice to the Hormone and Cancer Group, International Agency for Research on Cancer (IARC) laboratory for hormone analyses. On their arrival, samples were re-aliquotted in 3 ml-aliquots and stored at $-80\,°C$ until analysis. One aliquot was then used for the simultaneous measurements of T, ADIOL, $E_1$ and $E_2$, and a second one for the measurement of 2(OH)-E and $16\alpha(OH)$-$E_1$.

### Sex-steroid assays

#### T, ADIOL, $E_1$ and $E_2$
T, ADIOL, $E_1$ and $E_2$ were measured by radioimmunoassay (RIA) after enzymatic hydrolysis, solid phase extraction and high-performance liquid chromatography (HPLC) purification of the urinary samples.

The method for hydrolysis and purification of urinary sex-steroids has been described previously in detail [17, 18]. In brief, 1 ml of acetate buffer (pH = 5.2) and 5.6 mg of Helix Pomatia powder (Sigma-Aldrich Chimie, Lyon, France) were added to a 2-ml sample of urine, and the samples were incubated at $45\,°C$ for 22 hour. After hydrolysis, the samples were extracted by a C18 cartridge (LiChrolut, 500 mg) coupled to a $NH_2$ cartridge (LiChrolut, 200 mg, Merck, Darmstadt, Germany). Sex-steroids were eluted with 5 ml of acetonitrile through the column assembly. The eluent was then dried at $37\,°C$, under a gentle stream of dry nitrogen. The dried extract was then redissolved into 400 µl of 50% ethanol solution, passed through a syringe filter and then injected into an HPLC (100 µl) (1100 HPLC Hewlett Packard model, equipped with a Spherisorb S5 ODS2 column (Waters, Milford, USA), mobile phase: linear gradient 90% water/10% acetonitrile, to 100% acetonitrile in 25 min). Three different fractions of the eluent were collected: between 15.5 and 16.2 min for $E_2$, between 16.3 and 17.2 min for $E_1$ and T (which were eluted in the same fraction), and between 17.3 and 18.3 for ADIOL. All the fractions were dried under a gentle stream of nitrogen at $37\,°C$, and then redissolved in known quantities of albumin buffer at pH 7.4 (albumin fract.5, Merck).

Quantification of T, ADIOL, $E_1$ and $E_2$ concentrations was performed by RIA. Specific antibodies for $E_2$ and T were made by our laboratory (Dr Déchaud, Service de Radioanalyse et Radiopharmacie, Hôpital Neurocardiologique, Lyon, France), while for $E_1$ and ADIOL antibodies were purchased from Sigma and P.A.R.I.S. (Compiegne, France). [³H]-labelled tracers of T, $E_2$ and $E_1$ were purchased from Amersham Biosciences Europe (Orsay, France), while for ADIOL they were purchased from Perkin-Elmer (Courtaboeuf, France). Since $E_1$ and T were eluted in the same fractions, cross-reactivities of the respective antibodies were tested, but were found to be negligible (0.022% for anti-$E_1$ vs. T, and of 0.0055% for anti-T vs. $E_1$).

All measurements were done in duplicate, including hydrolysis, solid phase extraction, HPLC and RIA steps, and all samples collected from each subject were measured randomly in the same HPLC and RIA analytical batch. Laboratory technicians were blind to the time of sampling of the samples (t1, t2, or t3). For quality control, two control samples containing known amount of steroids were run through all the analysis (from hydrolysis to RIA) in each analytical batch. Intra-batch coefficients of variation were 6.9% for $E_1$ (at a concentration of 1900 ng/l), 4.7% for $E_2$ (at a concentration of 530 ng/l), 9.3% for ADIOL (at a concentration of 14,500 ng/l) and 6.7% for T (at a concentration of 2900 ng/l). Inter-batch coefficients of variation were 16% for $E_1$, 17% for $E_2$, 15% for ADIOL and 10% for T at the same concentrations. The detection limits of the method were 2.0 ng/100 ml for $E_1$ and $E_2$, 7.8 ng/100 ml for ADIOL and 1.6 ng/100 ml for T.

Results were expressed as concentrations, in ng analyte per litre. To correct for the variability in urine dilution, creatinine was measured in each sample by kinetic Jaffé reaction (Hitachi 717, Roche, Central

Laboratory for Biochemistry, Hôpital de l'Antiquaille, Lyon, France), and results were also expressed per mg creatinine.

Recovery and surcharge tests for hormone concentrations within the post-menopausal range were performed in duplicate for each hormone on three different urine samples. The dilution tests were done by diluting samples in cascade with urine from which all sex steroids had first been removed by stripping by solid phase extraction. For the surcharge recovery tests, a fixed volume of stripped urine was loaded with increasing, known quantities of each of the four sex steroids. Mean recovery for surcharge tests were 101% for $E_1$, 108% for $E_2$, 100% for ADIOL and 104% for T. Mean recovery for dilution tests were 94% for $E_1$, 104% for $E_2$, 119% for ADIOL and 109% for T.

## 2(OH)-E and 16α(OH)-E₁

2(OH)-E and 16α(OH)-$E_1$ were measured by solid-phase enzyme immunoassay (Estramet, Immuna Care Corporation, Bethlehem, USA), and the analyses were performed according to the protocols from the manufacturer. In brief, each urine sample was hydrolysed by β-glucuronidase/aryl sulfatase in acidic milieu, and was pipetted into microtiter plates with high-affinity murine monoclonal antibody. The enzyme alkaline phosphatase was linked to each estrogen metabolite, which was used for the competition with the standard or the analyte in the assay. Plates were read at end point at 405 nm (paranitrophenyl phosphate as chromogen) by an automated plate reader (Spectracount, Packard, Meriden, USA). All the samples were assayed in triplicate, and samples collected from the same subject were measured in the same analytical batch. The ratio between 2(OH)-E and 16α(OH)-$E_1$ was calculated from the averages of the triplicate measurements for each sample.

Validity, reproducibility and detection limits of the method have been published previously [19]. In our study, intra-batch coefficients of variation were 10.9% for 2(OH)-E and 12.7% for 16α(OH)-$E_1$ for concentrations of 3500 ng/l, while inter-batch coefficients of variation were 17.3 and 21.2% for 2(OH)-E and 16α(OH)-$E_1$, respectively. For the ratio 2(OH)-E/ 16α(OH)-$E_1$, intra-batch coefficients of variation were of 16.3% for a ratio of 1, while inter-batch coefficients of variation were of 17.9%.

All hormone measurements were performed at the laboratory for hormone analyses, Hormones and Cancer Group, IARC (Lyon, France).

Immunoassays were performed using an automated liquid handling system (Multiprobe II, Packard) with computer connections to a liquid scintillation analyser (TRI-CARB 1900CA, Packard) and an UV-reader (Spectracount, Packard).

## Statistical analyses

Statistical analyses were performed using the SAS statistical software package (SAS Institute Inc., Cary, NC, USA). The natural logarithm was used for all variables, to normalise their frequency distributions. Statistical analyses included the calculation of geometric means and confidence intervals, and calculation of correlation coefficients. Pearson correlations were used to examine the correlations over time of T, ADIOL, $E_1$, $E_2$, 2(OH)-E, 16α(OH)-$E_1$ and the ratio 2(OH)-E/16α(OH)-$E_1$, and were calculated by computing average values from the three different points in time (t1, t2 and t3) for each hormone and for each subject. Within- and between-subject variations were evaluated by computing intra-class correlation coefficients (ICCs), from the following model:

$$y = \text{between-subject} + \text{within-subject} + \text{error} \quad (1)$$

where $y$ denotes the value obtained by the urinary measurements of the hormones, and the terms on the right hand side are random effects. This corresponds to the variance decomposition:

$$\text{Var}[y] = \text{Var}[\text{between-subject}] + \text{Var}[\text{within-subject}] + \text{Var}[\text{error}]$$

Here Var[$y$] represents the total variance of the measurements, and Var[between-subject], Var[within-subject] and Var[error] are the variances attributable to the between-subjects variation, to within-subject variation, the variation between the three different points in time, and to random errors between the replicates within the same subject and the same sample (laboratory error), respectively. These variances were estimated by maximising the likelihood of the random effect model of Equation (1), as represented by the SAS procedure mixed.

The within and between-sample ICCs were calculated as

$$(\text{Var}[\text{between-subject}] + \text{Var}[\text{within-subject}])/ (\text{Var}[\text{between-subject}] + \text{Var}[\text{within-subject}] + \text{Var}[\text{error}])$$

and

$$\text{Var}[\text{between-subject}]/(\text{Var}[\text{between-subject}] + \text{Var}[\text{within-subject}] + \text{Var}[\text{error}]),$$

respectively.

For creatinine analyses, between-subject variations were evaluated by computing ICCs, from the following model:

$$y = \text{between-subject} + \text{error} \quad (2)$$

This corresponds to the variance decomposition:

$$\text{Var}[y] = \text{Var}[\text{between-subject}] + \text{Var}[\text{error}]$$

where Var[$y$] represents the total variance of the measurements, Var[between-subject] is the variance attributable to the between-subject variations, and Var[error] is the variation within a given subject in the three different points in time.

These variances were estimated by maximising the likelihood of the random effect model of Equation (2), as represented by the SAS procedure mixed.

The between-sample ICC was defined as Var-[within-subject]/(Var[within-subject] + Var [error]).

Calculations of hormone concentrations corrected by creatinine were obtained by the ratio of hormone concentration (expressed in ng/l) and of creatinine concentrations (mg/l). The statistical significance of mean changes in the hormone levels over time was assessed by analysis of variance (ANOVA) (using log-transformed variables). As the numbers of observations within the various blocks were not equal, all analyses of variance were done by generalised linear models, using the SAS statistical software package.

All statistical tests and corresponding $p$ values were two-sided with 95% confidence intervals.

## Results

Geometric means and confidence intervals for age, body mass index (BMI), T, ADIOL, $E_1$, $E_2$, 2(OH)-E, 16α(OH)-$E_1$, 2(OH)-E/16α(OH)-$E_1$ ratios and creatinine are presented in Table 1. BMI did not change significantly over time. Measurements for all hormones were well above the detection limit of the assay method, and absolute levels were found to be in the range of those previously reported [2, 13, 19, 20]. A trend towards decreased levels from t1 to t3 was observed for all hormones, and for $E_1$ and $E_2$ these trends were statistically significant ($p < 0.05$). The ratio of 2(OH)-E/16α(OH)-$E_1$ did not show any significant variation over time. The correction for creatinine values reduced the trend towards lower levels over time, although the trends remained substantial for $E_1$ and $E_2$ only ($p < 0.05$). Very similar results were obtained by adjusting hormone levels for creatinine concentrations by multivariate regression analysis (data not shown).

Within-subject variance, error (laboratory replicate) variance, and intra-class (ICCs) and Pearson's correlations over time are shown in Table 2. ICCs for laboratory replicates were all above 0.93, indicating good analytical reproducibility. Within-subject ICCs ranged from 0.55 to 0.85 for T, ADIOL, $E_1$ and $E_2$, but they were much lower for 2(OH)-E, 16α(OH)-$E_1$ and their ratio (ICCs < 0.46). The adjustment of results for creatinine levels did not substantially change the ICCs for T, ADIOL, $E_1$ and $E_2$, but slightly increased the ICCs for 2(OH)-E from 0.20 to 0.32, and decreased the ICCs observed for 16α(OH)-

**Table 1.** Age, BMI, and urinary hormones over time (geometric means and 95% confidence intervals, CI)[a]

|  | t1[b] Mean (CI) | t2 Mean (CI) | t3 Mean (CI) | Annual change of hormone levels (%) (CI) |
|---|---|---|---|---|
| Age (years)[c] | 57.6 (49.2–65.7) | 58.7 (50.3–67.0) | 63.0 (53.4–73.5) | |
| BMI[c] (kg/m$^2$) | 25.6 (18.3–34.5) | 25.5 (19.1–33.7) | 26.2 (19.8–35.9) | |
| Hormones (ng/l) | | | | |
| T | 1657.2 (1280.4–2144.8) | 1546.7 (1209.8–1977.5) | 1244.4 (973.3–1590.9) | −5.59 (−9.27 to −1.6) |
| ADIOL | 11,164.7 (8770.1–14,213.3) | 10,205.2 (8099.5–12858.5) | 9974.9 (8182.0–11,870.6) | −2.37 (−7.05 to 2.54) |
| $E_2$[d] | 316.3 (225.1–444.5) | 225.3 (175.8–288.8) | 187.9 (138.8–254.3) | −9.31 (−14.51 to −3.79) |
| $E_1$[d] | 1478.3 (1136.2–1923.3) | 1002.0 (796.7–1260.3) | 939.7 (734.9–1201.5) | −8.16 (−13.65 to −2.32) |
| 2(OH)-E | 4907.1 (4309.1–5588.0) | 4633.0 (3940.7–5447.0) | 4392.7 (3776.9–5109.1) | −5.46 (−9.95 to −0.75) |
| 16α(OH)-$E_1$ | 3526.1 (3075.2–4043.2) | 2699.9 (2289.2–3184.2) | 2775.9 (2442.0–3155.4) | −5.34 (−8.30 to 2.29) |
| Ratio 2(OH)-E/ 16α(OH)-E | 1.56 (1.39–1.74) | 1.75 (1.57–1.96) | 1.58 (1.38–1.82) | 0.18 (−4.52 to 5.11) |
| Creatinine[e] | 654.5 (564.3–759.2) | 578.2 (485.2–689.1) | 551.2 (479.2–633.9) | −4.80 (−8.48 to −0.98) |
| Hormones corrected for creatinine (ng hormone/mg creatinie) | | | | |
| T | 2.6 (2.1–3.1) | 2.7 (2.1–3.3) | 2.33 (1.9–2.9) | −0.82 (−3.51 to 1.93) |
| ADIOL | 17.3 (13.9–21.6) | 17.65 (13.9–22.4) | 18.74 (15.8–22.3) | 2.29 (−3.22 to 8.10) |
| $E_2$[d] | 0.5 (0.4–0.6) | 0.4 (0.3–0.5) | 0.3 (0.3–0.4) | −4.73 (−9.31 to 0.07) |
| $E_2$[d] | 2.3 (1.9–2.8) | 1.7 (1.5–2.0) | 1.7 (1.4–2.1) | −3.62 (−7.52 to 0.45) |
| 2(OH)-E | 8.3 (7.3–9.6) | 8.4 (7.5–9.6) | 8.0 (7.1–9.1) | −1.37 (−5.76 to 3.22) |
| 16α(OH)-$E_1$ | 5.5 (4.5–6.1) | 4.8 (4.4–5.2) | 5.1 (4.6–5.7) | −0.35 (−3.97 to 2.78) |

[a] n = 43 post-menopausal women; [b] t1, t2, t3 = timepoint 1, timepoint 2, timepoint 3; [c] means, minimum and maximum based on arithmetic means; [d] statistically signifiant between t1 and t3 ($p < 0.05$); [e] mg/l.

**Table 2.** Variance components, and intraclass and Pearson correlation coefficients (with 95% confidence intervals, CI) of urinary sex steroids (values in italics are creatinine adjusted)

| Hormone (ng/l) | Var 'subject' | Var 'within-subject' | Var 'error' | Lab-Replicate ICC (CI) | Within-subjects ICC (CI) | Within-subjects Pearson correlations between t1 and t3 (CI) |
|---|---|---|---|---|---|---|
| T | 0.59 | 0.09 | 0.01 | 0.98 (0.97–0.99) | 0.85 (0.79–0.82) | 0.84 (0.73–0.91) |
| | *0.42* | *0.06* | *0.01* | *0.98 (0.97–0.99)* | *0.86 (0.80–0.93)* | *0.89 (0.80–0.94)* |
| ADIOL | 0.32 | 0.22 | 0.04 | 0.93 (0.91–0.95) | 0.55 (0.42–0.78) | 0.37 (0.08–0.60) |
| | *0.32* | *0.22* | *0.04* | *0.93 (0.91–0.95)* | *0.55 (0.42–0.79)* | *0.41 (0.12–0.64)* |
| $E_2$ | 0.66 | 0.37 | 0.03 | 0.97 (0.96–0.98) | 0.62 (0.50–0.83) | 0.56 (0.30–0.74) |
| | *0.43* | *0.29* | *0.03* | *0.95 (0.93–0.97)* | *0.57 (0.44–0.81)* | *0.45 (0.18–0.66)* |
| $E_1$ | 0.42 | 0.26 | 0.02 | 0.96 (0.95–0.97) | 0.59 (0.46–0.81) | 0.54 (0.28–0.73) |
| | *0.23* | *0.13* | *0.02* | *0.94 (0.92–0.96)* | *0.60 (0.48–0.81)* | *0.49 (0.22–0.69)* |
| 2(OH)-E | 0.04 | 0.16 | 0.03 | 0.96 (0.95–0.97) | 0.20 (0.10–1.00) | 0.20 (−0.15–0.51) |
| | *0.05* | *0.09* | *0.03* | *0.95 (0.94–0.96)* | *0.32 (0.19–0.91)* | *0.20 (−0.15–0.51)* |
| $16\alpha(OH)-E_1$ | 0.10 | 0.11 | 0.03 | 0.96 (0.95–0.97) | 0.46 (0.33–0.79) | 0.56 (0.29–0.74) |
| | *0.03* | *0.07* | *0.03* | *0.92 (0.90–0.94)* | *0.26 (0.14–1.00)* | *0.35 (0.04–0.60)* |
| Ratio 2(OH)-E/ $16\alpha(OH)-E_1$ | 0.03 | 0.11 | 0.03 | 0.93 (0.91–0.95) | 0.18 (0.08–1.00) | 0.09 (−0.26–0.41) |
| Creatinine | 0.12 | | 0.14 | | 0.46 | 0.57 (0.32–0.74) |

$E_1$, from 0.46 to 0.26. The ICC over time for creatinine was 0.46. For all the hormones, Pearson's correlation coefficients and ICCs were comparable. Since Pearson's correlation coefficients reflect the agreement of reproducibility of hormone levels over time, and ICCs reflects this tendency also on absolute scale, this finding suggested that ICCs over time were not underestimated because of the decline of hormone concentrations due to ageing.

The cross-classification of subjects by tertiles of hormone levels over time is presented in Table 3. To evaluate the accuracy of subject classification by hormone levels on the basis of only one biological sample collection, subjects were classified into tertiles (due to the small sample size). The values on the diagonal line represent the concordance between the ranking of subjects according to the sample collected in t1, and the ranking of subjects according to the mean of samples collected in t2 and t3 (longer exposure over time). For T (within-subject ICCs of 0.85), 81% of the subjects (34 on 42) were well classified (assuming the mean levels of t2 and t3 as reference concentration). For the ratio of 2(OH)-E/$16\alpha(OH)-E_1$ (ICCs = 0.18), only 42% of the subjects (14 on 33) were correctly classified.

## Discussion

We evaluated the reproducibility of urinary sex-steroid concentrations over time in first-morning voids from post-menopausal women, measuring T, ADIOL, $E_1$ and $E_2$ by RIA after enzymatic hydrolysis, solid-phase extraction and HPLC purification of the

**Table 3.** Cross classification of subjects due to their urinary hormone levels: first urine sample by the mean of the second and the third

**Testosterone (n = 42)**

| | | Tertile for mean of sample 2 and 3 | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| Tertile for sample 1 | 1 | 11 | 3 | 0 | 14 |
| | 2 | 3 | 10 | 1 | 14 |
| | 3 | 0 | 1 | 13 | 14 |
| | | 14 | 14 | 14 | |

**Ratio 2(OH)-E/$16\alpha(OH)-E_1$ (n = 33)[a]**

| | | Tertile for mean of sample 2 and 3 | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| Tertile for sample 1 | 1 | 5 | 3 | 2 | 10 |
| | 2 | 2 | 4 | 4 | 10 |
| | 3 | 4 | 4 | 5 | 13 |
| | | 11 | 11 | 11 | |

Values on the diagonal line represent concordance between sample in t1 and the mean of samples in t2 and t3.
[a] Nine women were missing one value of either 2(OH)-E or $16\alpha(OH)-E_1$ and were therefore excluded from the analyses.

samples, and 2(OH)-E and 16α(OH)-E$_1$ by solid-phase enzyme immunoassay. Reproducibility over time was found to be good for T, reasonably high for ADIOL, E$_1$ and E$_2$, but low for 2(OH)-E, 16α(OH)-E$_1$ and the ratio of 2(OH)-E/16α(OH)-E$_1$.

For all hormones, there was a trend toward lower concentrations with time, possibly reflecting decreased production with age [21]. In fact, the lowest concentrations of hormones were found at t3 (between 2.2 and 8.8 years after t1), the time point at which women were the oldest, and further in time from the start of menopausal status. The decrease in hormone levels was higher for E$_1$ and E$_2$ than for the other hormones, while the ratio between 2(OH)-E and 16α(OH)-E$_1$ remained constant over time. The decrease in estrogens was far greater than the decrease in creatinine, and persisted after adjustment for creatinine. This suggests that this hormone decrease was not due to changes in urine volumes.

The trend toward lower concentrations over time (before and after the adjustment for creatinine values, for almost all hormones) suggested that there were probably no major problems of sex steroid degradation with storage, since the samples that were stored the longest had the highest concentrations. This confirms previous observations for T, E$_1$ and E$_2$ [22, 23]. Our study also showed that 2(OH)-E and 16α(OH)-E$_1$ were stable enough to be measured in urine samples stored at −20 °C for more 25 years without any added preserving agents. Absolute values in the frozen samples were not different from those in fresh urine samples from post-menopausal women (average data within "normal" post-menopausal range on fresh urine samples were provided by the kit manufacturer), or in urine samples that were stored for much shorter periods of time, measured with the same method [8, 24–26]. The apparent stability of hormones during storage time, even at a temperature of −20 °C, may be an important issue for prospective epidemiological studies in which biological samples are stored frozen over many years.

To our knowledge, only one study has been published so far on the reproducibility of urinary sex steroid measurements over time in post-menopausal women [15]. In this previous study, 60 women who participated as controls in a case–control study were invited to give a second 24-hour urine sample approximately 1 year after the first sample. For post-menopausal women, Pearson's correlation coefficients were found to be 0.41 for T and 0.57 for ADIOL. In our study we found within-subject ICCs of 0.85 for T and 0.55 for ADIOL for measurements over time, despite the fact that our measurements were performed on first-morning voids, rather than on 24-hour urine samples, and despite the longer time intervals between samples. Two other studies have reported the reproducibility of urinary hormones over time, but they focused only on estrogen metabolites in pre-menopausal women [14, 27]. Over a 1-year interval, ICCs in the luteal phase were found to range from 0.56 for 16(OH)-E$_1$ to 0.71 for the ratio between the two metabolites [14], and over a 2-month interval, correlations were found to be even higher [27]. In our study, ICCs over-time for estrogen metabolites in post-menopausal women were not satisfactory. This could have different explanations. First, measurements of estrogens metabolites in urine samples from post-menopausal women could be problematic, since absolute levels are lower compared to levels found in pre-menopausal women [19, 26]. Second, the time interval between the urine samples in our study was longer than in the previous studies (between 0.9 and 7.7 years). Third, the ratio between 2(OH)-E and 16α(OH)-E$_1$ could be influenced by changes in diet [28, 29] and smoking habit [30], factors that could easily vary over a long period of time. Furthermore, the width of the confidence intervals of ICCs and Pearson's correlation coefficients for 2(OH)-E, 16α(OH)-E$_1$ and their ratio could suggest that the power of the study is rather limited for those hormones.

Urine collection over 24-hours is known to be the best option to accurately determine a woman's daily hormone excretion [31], but in practice this kind of collection represents a heavy burden for study participants. First-morning voids or overnight collections can be obtained more easily, but may decrease somewhat the reliability of measurements. Most of the studies published on hormone measurements in overnight urine samples or first morning voids use creatinine measurements to adjust for dilution errors [32, 33]. Some papers have questioned whether such creatinine adjustments will actually reduce, or rather increase, measurement errors [34–36]. Creatinine excretion can vary depending on muscle mass, physical exercise, diet and age [35, 37] of the subjects. In addition, creatinine measurements are subjected to methodological errors as are any other measurements. In our study, the adjustment for creatinine did not substantially change the ICCs of hormones over time. Creatinine ICC over time was only 0.46. Since creatinine is an index of lean body mass [38], this low ICC could be due to variation in the loss of lean body tissue during aging.

In conclusion, the classification of T, ADIOL, E$_1$ and E$_2$ concentrations in first morning urine samples from post-menopausal women appeared to be quite reproducible over time. Reproducibility of 2(OH)-E and 16α(OH)-E$_1$ and the ratio of 2(OH)-E/16α(OH)-E$_1$, on the other hand, appeared to be very low over a long period of time. Thus, urinary measurements of sex-steroids by RIA after sample purification appears to be an appropriate alternative to serum measurements in epidemiological studies, as a non-invasive method easy-to-apply on a large scale. The high variability of 2(OH)-E and 16α(OH)-E$_1$, however, suggests that one urine sample may not be representative of these metabolites over time.

## References

1. Thomas HV, Reeves GK, Key TJ. Endogenous estrogen and postmenopausal breast cancer: A quantitative review. Cancer Causes Control 1997; 8: 922–928.
2. Secreto G, Toniolo P, Berrino F, et al. Serum and urinary androgens and risk of breast cancer in postmenopausal women. Cancer Res 1991; 51: 2572–2576.
3. Ettinger B, Pressman ASP, Bauer DC, Cauley JA, Cummings SR. Associations between low levels of serum estradiol, bone density, and fractures among elderly women: The study of osteoporotic fractures. 1998 (report).
4. Garnero P, Sornay-Rendu E, Claustrat B FAU, Delmas PD. Biochemical markers of bone turnover, endogenous hormones and the risk of fractures in postmenopausal women: The OFELY study. 2000 (report).
5. Moreira-Andres MN, Canizo FJ, Hawkins F. Is there a place for urinary growth hormone measurement? Acta Endocrinol (Copenh) 1993; 128: 197–201.
6. Brook CG, Jacobs HS, Stanhope R, Adams J, Hindmarsh P. Pulsatility of reproductive hormones: Applications to the understanding of puberty and to the treatment of infertility. Baillieres Clin Endocrinol Metab 1987; 1: 23–41.
7. Ikegawa S, Lahita R, Fishman J. Concentration of 16 alpha-hydroxyestrone in human plasma as measured by a specific RIA. J Steroid Biochem 1983; 18: 329–332.
8. Falk RT, Rossi SC, Fears TR, et al. A new ELISA kit for measuring urinary 2-hydroxyestrone, 16alpha-hydroxyestrone, and their ratio: Reproducibility, validity, and assay performance after freeze-thaw cycling and preservation by boric acid. Cancer Epidemiol Biomarkers Prev 2000; 9: 81–87.
9. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Oxford University Press, New York, 1992.
10. Cauley JA, Gutai JP, Kuller LH, Powell JG. Reliability and interrelations among serum sex hormones in postmenopausal women. Am J Epidemiol 1991; 133: 50–57.
11. Toniolo P, Koenig KL, Pasternack BS, et al. Reliability of measurements of total, protein-bound, and unbound estradiol in serum. Cancer Epidemiol Biomarkers Prev 1994; 3: 47–50.
12. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer FE. Reproducibility of plasma hormone levels in postmenopausal women over a 2–3-year period. Cancer Epidemiol Biomarkers Prev 1995; 4: 649–654.
13. Muti P, Trevisan M, Micheli A, et al. Reliability of serum hormones in premenopausal and postmenopausal women over a one-year period. Cancer Epidemiol Biomarkers Prev 1996; 5: 917–922.
14. Michaud DS, Manson JE, Spiegelman D, et al. Reproducibility of plasma and urinary sex hormone levels in premenopausal women over a one-year period. Cancer Epidemiol Biomarkers Prev 1999; 8: 1059–1064.
15. Micheli A, Muti P, Pisani P, et al. Repeated serum and urinary androgen measurements in premenopausal and postmenopausal women. J Clin Epidemiol 1991; 44: 1055–1061.
16. de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 1984; 37: 1–44.
17. Venturelli E, Manzari A, Cavalleri A, Benzo M, Secreto G, Marubini E. Urinary testosterone measurement by gas chromatography after solid-phase extraction and high-performance liquid chromatography. J Chromatogr 1992; 582: 7–12.
18. Minut GJ, Cambie M, Lanzoni M, Marubini E, Secreto G. Urinary 5 alpha-androstanediol and 5 beta-androstanediol measurement by gas chromatography after solid-phase extraction and high-performance liquid chromatography. Int J Biol Markers 1999; 14: 154–159.
19. Falk RT, Gail MH, Fears TR, et al. Reproducibility and validity of radioimmunoassays for urinary hormones and metabolites in pre- and postmenopausal women. Cancer Epidemiol Biomarkers Prev 1999; 8: 567–577.
20. Key TJ, Wang DY, Brown JB, et al. A prospective study of urinary oestrogen excretion and breast cancer risk. Br J Cancer 1996; 73: 1615–1619.
21. Vermeulen A. Sex hormone status of the postmenopausal woman. Maturitas 1980; 2: 81–89.
22. Bolelli G, Muti P, Micheli A, et al. Validity for epidemiological studies of long-term cryoconservation of steroid and protein hormones in serum and plasma. Cancer Epidemiol Biomarkers Prev 1995; 4: 509–513.
23. Kley HK, Schlaghecke R, Kruskemper HL. Stability of steroids in plasma over a 10-year period. J Clin Chem Clin Biochem 1985; 23: 875–878.
24. Meilahn EN, De Stavola B, Allen DS, et al. Do urinary oestrogen metabolites predict breast cancer? Guernsey III cohort follow-up. Br J Cancer 1998; 78: 1250–1255.
25. Ursin G, London S, Stanczyk FZ, et al. Urinary 2-hydroxyestrone/16alpha-hydroxyestrone ratio and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1999; 91: 1067–1072.
26. Ziegler RG, Rossi SC, Fears TR, et al. Quantifying estrogen metabolism: An evaluation of the reproducibility and validity of enzyme immunoassays for 2-hydroxyestrone and 16alpha-hydroxyestrone in urine. Environ Health Perspect 1997; 105(Suppl 3): 607–614.
27. Chen C, Malone KE, Prunty J, Daling JR. Measurement of urinary estrogen metabolites using a monoclonal enzyme-linked immunoassay kit: Assay performance and feasibility for epidemiological studies. Cancer Epidemiol Biomarkers Prev 1996; 5: 727–732.

28. Michnovicz JJ, Adlercreutz H, Bradlow HL. Changes in levels of urinary estrogen metabolites after oral indole-3-carbinol treatment in humans. J Natl Cancer Inst 1997; 89: 718–723.

29. Fowke JH, Longcope C, Hebert JR. Brassica vegetable consumption shifts estrogen metabolism in healthy postmenopausal women. Cancer Epidemiol Biomarkers Prev 2000; 9: 773–779.

30. Michnovicz JJ, Naganuma H, Hershcopf RJ, Bradlow HL, Fishman J. Increased urinary catechol estrogen excretion in female smokers. Steroids 1988; 52: 69–83.

31. Muti P, Deutsch A, Freudenheim J, Bolelli G, Hill L, Trevisan M. Why and how to measure urinary sex steroid metabolites in epidemiological studies in women. Nutr Metab Cardiovasc Dis 2000; 10: 85–91.

32. Gorgels WJ, Graaf Y, Blankenstein MA, Collette HJ, Erkelens DW, Banga JD. Urinary sex hormone excretions in premenopausal women and coronary heart disease risk: A nested case-referent study in the DOM-cohort. J Clin Epidemiol 1997; 50: 275–281.

33. Zheng W, Dai Q, Custer LJ, et al. Urinary excretion of isoflavonoids and the risk of breast cancer. Cancer Epidemiol Biomarkers Prev 1999; 8: 35–40.

34. Zacur H, Kaufman SC, Smith B, et al. Does creatinine adjustment of urinary pregnanediol glucuronide reduce or introduce measurement error? Gynecol Endocrinol 1997; 11: 29–33.

35. Boeniger MF, Lowry LK, Rosenberg J. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: A review. Am Ind Hyg Assoc J 1993; 54: 615–627.

36. Alessio L, Berlin A, Dell'Orto A, Toffoletto F, Ghezzi I. Reliability of urinary creatinine as a parameter used to adjust values of urinary biological indicators. Int Arch Occup Environ Health 1985; 55: 99–106.

37. Walser M. Creatinine excretion as a measure of protein nutrition in adults of varying age. JPEN J Parenter. Enteral Nutr 1987; 11: 73S–78S.

38. Welle S, Thornton C, Totterman S, Forbes G. Utility of creatinine excretion in body-composition studies of healthy men and women older than 60 y. Am J Clin Nutr 1996; 63: 151–156.

*Address for correspondence*: Dr Sabina Rinaldi, International Agency for Research on Cancer, Hormones and Cancer Group, 150 Cours Albert Thomas, 69372 Lyon, France
Phone: +33-4-72738328; Fax: +33-4-72738361
E-mail: rinaldi@iarc.fr