

# Monitoring the menstrual cycle: Comparison of urinary and serum reproductive hormones referenced to true ovulation

Judith Roos\*, Sarah Johnson†, Sarah Weddell†, Erhard Godehardt‡, Julia Schiffner§, Günter Freund\* and Christian Gnoth\*

\*Endocrinology- and IVF-Centre, Grevenbroich, Germany, and Department of Gynaecology and Obstetrics, University of Cologne, Cologne, Germany, †SPD Development Company Ltd, Bedford, UK, ‡Medical Faculty (Biostatistics), Heinrich-Heine-University Düsseldorf, HELLenthal, Germany, and §Mathematical Institute, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

**ABSTRACT** **Objective** The aim of the study was to examine relationships and interindividual variations in urinary and serum reproductive hormone levels relative to ultrasound-observed ovulation in menstrual cycles of apparently normally menstruating women.

**Methods** This was a prospective study of normally menstruating women (no known subfertility), aged 18–40 years ( $n = 40$ ), who collected daily urine samples and attended the study centre for blood samples and transvaginal ultrasound during one complete menstrual cycle. Serum luteinising hormone (LH), progesterone, estradiol, urinary LH, pregnanediol-3-glucuronide (P3G) and estrone-3-glucuronide were measured. Ultrasound was conducted by two physicians and interpreted by central expert review.

**Results** Menstrual cycle length varied from 22 to 37 days (median 27 days). Ovulation by ultrasound ranged from day 8 to day 26 (median day 15). Serum and urinary hormone profiles showed excellent agreement. Estrogen and LH hormone peaks in urine and serum showed a range of signal characteristics across the study group before and after ovulation. The rise in estrogen and LH always occurred before ovulation; the progesterone rise from baseline always occurred after ovulation.

**Conclusions** Urinary and serum reproductive hormones showed excellent agreement and may be used interchangeably. The beginning of the surge in serum and urinary LH was an excellent predictor of ovulation. The rise in progesterone and P3G above baseline was a consistent marker of luteinisation confirming ovulation. Both LH and progesterone surges delivered clear, sharp signals in all volunteers, allowing reliable detection and confirmation of ovulation.

**KEY WORDS** Estradiol; Estrone-3-glucuronide (E3G); Luteinising hormone (LH); Ovulation; Pregnanediol-3-glucuronide (P3G); Progesterone; Reproductive hormones; Transvaginal ultrasound

Trial registration number: NCT01802060

Correspondence: Christian Gnoth, green-ivf, Grevenbroich Endocrinology and IVF Center, Rheydter Str. 143, 41515 Grevenbroich, Germany. Tel: + 49 2181 491513. Fax: + 49 2181 491534. E-mail: dr.christian.gnoth@green-ivf.de

## INTRODUCTION

Menstrual cycle length and hormonal levels in premenopausal women are relevant indicators of female reproductive health and show high variability among women of early, mid and late reproductive age<sup>1</sup>. There is a fluent shift from physiological variability to pathological cycle irregularity, with ovulatory dysfunction being associated with subfertility<sup>2</sup>. Due to changing reproductive behaviour with long periods of hormonal contraception use, sometimes for up to or more than 20 years, knowledge of natural menstrual cycles has become less certain<sup>3,4</sup>.

The wide extent of intra- and interindividual menstrual cycle variability is well described in the literature<sup>5-7</sup>. Reports state that the mean individual cycle length variability is 6.7 days<sup>6</sup>; 46% of women have cycles that vary in length by  $\geq 7$  days; and anovulatory cycles occur in up to 12% of apparently normally menstruating women<sup>7</sup>. Therefore, it is sometimes difficult to anticipate when to measure reproductive hormones and time clinical visits in order to detect key events in the menstrual cycle, such as the estrogen rise, the luteinising hormone (LH) surge and ovulation<sup>8</sup>, and to distinguish between physiological variation and pathological dysfunction.

Much of our knowledge concerning the natural menstrual cycle and its variations derives from studies published several decades ago<sup>9</sup>. Cycle monitoring, however, is still one of the most important diagnostic procedures for evaluating subfertile couples and timing reproductive interventions and certain natural contraceptive methods<sup>10,11</sup>. Only 55% of women's self-estimated ovulation days, however, fall within their actual fertile window<sup>12</sup>. Detecting the day of ovulation is important because one single episode of intercourse on the day prior to ovulation delivers nearly the same chance of resulting in pregnancy as multiple episodes of intercourse across the cycle<sup>13</sup>. Therefore, reliable cycle monitoring is necessary to optimise natural conception potential, effective non-hormonal contraception and early diagnosis of cycle disturbances or subfertility<sup>7</sup>.

Due to the importance of cycle monitoring, the Menstrual Cycle Monitoring (MeMo) study (trial registration number NCT01802060) strove to update knowledge of menstrual cycles<sup>14</sup>. The study used tools such as modern ultrasound equipment for daily follicular monitoring, which provides high-resolution,

reliable follicle counts of even small antral follicles, and modern/new laboratory assays for measurement of serum and urinary reproductive hormones. The goal of the study was to create reproductive hormone ranges referenced to an independently determined ovulation day and to evaluate the benefits of urinary hormone measurements to replace expensive, invasive and labour-intensive endocrine monitoring of serum hormones. This manuscript reports the reference ranges for serum hormones and their corresponding relationship with urinary hormone levels referenced to ultrasound-detected ovulation.

## METHODS

The MeMo study was a prospective study carried out to create ranges for serum reproductive hormones (serum LH, estradiol [E<sub>2</sub>] and progesterone) and their urinary equivalents (urinary LH, estrone-3-glucuronide [E3G], pregnanediol-3-glucuronide [P3G]) referenced to objective ovulation detected by transvaginal ultrasound.

The study was conducted at green-ivf, Grevenbroich, Germany, a private centre for gynaecological endocrinology and reproductive medicine, in close cooperation with Swiss Precision Diagnostics (SPD) Development Company Ltd, Bedford, UK.

### Study population and design

The study was conducted between August 2012 and April 2013. Fifty-one women were recruited who fulfilled the following inclusion criteria: age 18 to 40 years; menstrual cycle length 23 to 35 days; two or more natural cycles prior to study participation; use of non-hormonal contraception to avoid pregnancy or abstention from intercourse during the observational study phase; signed informed consent and data protection forms. Exclusion criteria included: menopausal symptoms (e.g., night sweats, hot flushes); use of hormonal treatments; known subfertility; evidence of endocrine disorders (polycystic ovary syndrome, untreated thyroid dysfunction); taking of drugs that interfere with steroid hormone metabolism (e.g., antibiotics); having impaired liver or kidney function; obesity; having previously undergone major pelvic surgery (e.g., ovarian cyst, endometriosis).

## Study visits and data collection

Volunteer recruitment targeted women who reported having normal menstrual cycles. The women were recruited by word-of-mouth advertising and on a 'first come, first served' basis. The study lasted for one menstrual cycle, starting on day 1 of the cycle and concluding with the onset of the next menses. The volunteers were required to collect daily urine samples to measure the corresponding urinary hormones. Transvaginal ultrasonography was conducted every 2 days, from cycle day 5 to 11, and daily once a follicle diameter had reached 16 mm until ovulation occurred and the dominant follicle disappeared. Individual visit schedules in between were timed and determined by the examiners based on follicle growth up to 16 mm. Blood samples were taken whenever transvaginal ultrasound was performed. All ultrasound examinations and hormonal blood tests were performed at the study site in the morning (most visits) or in the late afternoon (some visits). After ultrasound-observed ovulation, the monitoring was finalised by blood samples taken 7 and 9 days after ovulation. Volunteers were reimbursed €350 for their participation.

## Laboratory analysis

Daily morning urine samples were collected in 200 ml pots containing sodium azide to prevent degradation. Volunteers kept one pot for each day at + 4°C until returning it to the study site at the next visit, where the pots were stored at - 20°C for later analysis. Urinary LH, E3G and P3G were measured using SPD Development Company Ltd proprietary in-house assays on the AutoDELFI A analysis platform (PerkinElmer, Waltham, MA, USA). Intra- and interassay percentage confidence values were < 5% for all assays used; validation, together with further details of these assays, has been previously reported elsewhere<sup>14,15</sup>. The assay used to detect urinary LH consists of immobilised biotinylated antibody (antibody no. 2119; SPD Development Company Ltd) that recognises the  $\alpha$  LH subunit bound to streptavidin plates, and a second, europium-labelled antibody that recognises the  $\beta$  subunit (antibody no. 2301; SPD Development Company Ltd); thus, it is only able to measure intact LH. Urine volume and dilution effects were checked by validated creatinine adjustment producing smoother hormone profiles<sup>15</sup>. Creatinine was measured by ABX Pentra XL

80 (Horiba, Irvine, CA, USA). Serum samples were analysed for E<sub>2</sub>, progesterone and LH from blood collection tubes on the day of the visit. This was carried out using the ADVIA Centaur XP Immunoassay System (Siemens, Erlangen, Germany) at the study site in order to receive test results on the same morning and assist with appropriate timing of the next visit. Centaur and AutoDELFI A analysis platforms are fully automatic, high-throughput immunoassay systems. All assays were performed strictly according to the manufacturers' instructions.

## Transvaginal ultrasound examination

All transvaginal ultrasounds were performed by the same skilled examiners (CG and JR) using a Siemens G50 ultrasound monitor equipped with a 3.7–7.5 MHz vaginal transducer. Follicles were measured in two representative diameters orthogonally from the follicles' inner boundary. All antral follicles were counted in classes from 3 mm onwards. Endometrial thickness and appearance were also assessed. Digital pictures of the follicles were saved on the clinic data terminal. At every visit, all measurements, the return of urine samples and any special or adverse events were recorded in each volunteer's daily documenting diary and the next visit was scheduled. Assays were conducted in duplicate, and panel image review was performed to reduce errors.

## Data management and analysis

Medical and laboratory data were recorded using the clinic's digital management program MEDISTAR (Hannover, Germany) and Excel (Microsoft Corporation, Redmond, WA, USA). Data were analysed using SAS software (version 9.2; SAS Institute, Cary, NC, USA) and the statistics program R, a free software environment for statistical computing and graphics (www.r-project.org; R Core Team, 2014), to create ranges referenced to the day of objective ovulation and logarithmic scatter plots for comparison of urine and serum hormones.

The day of objective ovulation was determined as the day before disappearance of the dominant follicle. Daily ultrasound prior to ovulation was not possible in all participants; therefore, for four volunteers, in whom ovulation occurred between two scan days, the day of ovulation was considered as the day following

the last day when the dominant follicle was present. Peak hormonal levels were defined as the highest of all measurements taken across the whole cycle. The hormonal rises (E3G, E<sub>2</sub>, P3G and progesterone) and surge in urinary and serum LH were assigned retrospectively and defined as the first rise above the best-fit line of previous measurements.

Descriptive statistics were calculated for demographic characteristics and basic cycle data (median, range, mean and standard deviation [SD]). Percentiles were calculated to create hormonal ranges, and scatter plots to assess correlation, agreement and interchangeability of urinary and serum hormones. The prediction line in the scatter plots was calculated using a variance component model (R package MethComp)<sup>16</sup>. This model takes the specific data structure into account, since there were a number of volunteers with repeated measurements per individual cycle.

### Ethics approval

The study was conducted in accordance with the Declaration of Helsinki. All patients gave their signed informed consent after reading the information material. The study was approved by the ethics committee of the Chamber of Physicians (Düsseldorf, Germany) and the Federal Institute for Drugs and Medical Devices (Bonn, Germany).

## RESULTS

Initially, 51 volunteers were recruited to the MeMo study. All reported natural, regular menstrual cycles and no known subfertility. Three women withdrew for personal reasons and seven were withdrawn from the study for the following reasons: ovarian cysts at the time of their first scan, of which they were previously unaware ( $n = 2$ ); pre-trial pregnancy ( $n = 2$ ); and non-compliance with the study protocol ( $n = 3$ ). A total of 41 women completed one menstrual cycle with continuous monitoring of urinary and serum reproductive hormones and transvaginal ultrasound. One woman was found to have an anovular cycle and was therefore excluded from the analysis.

In 36 of the remaining 40 volunteers, the day of ovulation was determined unambiguously by transvaginal ultrasound (90%), as was expected from the

literature<sup>17</sup>. In four women, ovulation was estimated to have occurred between two scan days. All antral follicles were counted in classes from 3 mm onwards but, unlike in previous reports, waves of antral follicle development were not observed<sup>18</sup>. Three women conceived during cycle monitoring. Table 1 describes the volunteers' menstrual cycle characteristics (excluding cycle length and luteal phase data for the three pregnant volunteers for whom this information was not available). The mean age of the volunteers was 28.9 years (SD 4.76 years, median 29.5 years, range 18 to 37 years). Seventeen of the volunteers had had a previous pregnancy and 12 had previously had a live birth. Ninety-five percent of the volunteers were white Caucasian and 5% were Asian.

The cycle lengths varied from 22 to 37 days (median 27 days). The day of ovulation ranged from day 8 to day 26 (median day 14.5). The luteal phase length was relatively constant at  $12.26 \pm 2.49$  days.

The mean serum hormone levels of E<sub>2</sub>, progesterone and LH, relative to the day of ovulation, are given in Table 2. Details of the corresponding urinary hormone ranges obtained in this study have recently been published elsewhere<sup>14</sup>. The ranges for reproductive hormones derived from daily urinary and selected serum hormone analysis (median and 10th–90th percentile range) relative to the day of true ovulation observed by ultrasound and scatter plots are shown in Figures 1–3.

Figure 1A shows the ranges of serum and urinary LH levels, referenced to the day of objective ovulation determined by ultrasound (median and 10th–90th percentile range). Figure 1B is a scatter plot, with prediction lines and 95% prediction limits, of urinary LH against serum LH (coefficient of correlation  $r = 0.72$ ). If measurements in urine are compared with serum measurements from the day before, the correlation is slightly decreased ( $r = 0.69$ ). Figure 1C shows a scatter plot of urinary LH not creatinine-corrected against urinary LH creatinine-corrected, to assess measurement errors in urine due to dilution and volume effects (coefficient of correlation  $r = 0.87$ ). The scatter plots demonstrate that urinary LH (creatinine-corrected and not corrected) and urinary and serum LH are highly correlated and show good agreement. Thus, serum and urine levels of LH may be used interchangeably.

Figure 2A shows the ranges (median and 10th–90th percentile range) observed for E<sub>2</sub> and E3G referenced

**Table 1** Basic cycle characteristics of women with self-reported regular menses in the MeMo study.

Cycle characteristic	n	Median	Minimum	Maximum	5th percentile	25th percentile	75th percentile	95th percentile	Mean	SD	Lower 95% CI mean	Upper 95% CI mean
Cycle length, days	37*	27.00	22.00	37.00	22.00	26.00	29.00	34.00	27.70	3.37	26.58	28.82
Day of ovulation	40	14.50	8.50	26.50	11.00	12.75	17.00	23.50	15.35	3.84	14.12	16.58
Length of luteal phase, days	37*	12.50	3.00	15.50	6.50	11.50	13.50	15.50	12.26	2.49	11.43	13.09

\*Excludes three women who became pregnant during the course of the study.  
CI, confidence interval; SD, standard deviation.

to the day of objective ovulation, as determined by ultrasound. Figure 2B is a scatter plot of urinary E3G against serum E<sub>2</sub> (coefficient of correlation  $r = 0.51$ ). If measurements in urine are compared with serum measurements from the day before, the correlation is improved ( $r = 0.54$ ). Figure 2C is a scatter plot of urinary E3G not creatinine-corrected against urinary E3G creatinine-corrected, to assess measurement errors in urine due to dilution or volume effects ( $r = 0.84$ ). Urinary E3G (creatinine-corrected and not corrected) and serum E<sub>2</sub> are highly correlated and show a good agreement. Thus, serum E<sub>2</sub> and urinary E3G may also be used interchangeably.

Figure 3A shows the reference ranges for progesterone and P3G referenced to the day of objective ovulation, as determined by ultrasound (median and 10th–90th percentile range). Figure 3B is a scatter plot of urinary P3G against serum progesterone (coefficient of correlation  $r = 0.81$ ). If measurements in urine are compared with serum measurements from the day before, the correlation is improved ( $r = 0.82$ ). Figure 3C is a scatter plot of urinary P3G not creatinine-corrected against urinary P3G creatinine-corrected, to assess measurement errors in urine due to dilution or volume effects ( $r = 0.91$ ). Urinary P3G (creatinine-corrected and not corrected) and serum progesterone are highly correlated and show good agreement. Thus, serum progesterone and urinary P3G may be used interchangeably.

Figure 4 shows the differences in days between hormone peaks (highest levels measured across the cycle) and objective ovulation (median and 5th–95th percentile).

Urinary LH peak preceded ovulation in many women (mean time from peak to objective ovulation  $-0.35 \pm 0.62$  days), but in 23% of cycles the LH peak was detected after ovulation (ranging from  $-2.5$  days before to  $0.5$  days after ovulation). Serum LH peak also preceded ovulation in many women (mean time from peak to objective ovulation  $-0.53 \pm 1.15$  days), but in more than 25% of women the serum LH peak was detected after ovulation (ranging from  $-2.5$  days before to  $3$  days after ovulation). Urinary follicle stimulating hormone (FSH) peak was measured with a mean time to objective ovulation of  $-2.35 \pm 4.32$  days (ranging from  $-3.5$  days before to  $0.5$  days after ovulation). Serum FSH was only measured once on the occasion of the first visit on day 5. Mean serum FSH was  $7.6 \pm 2.2$  mIU/ml.

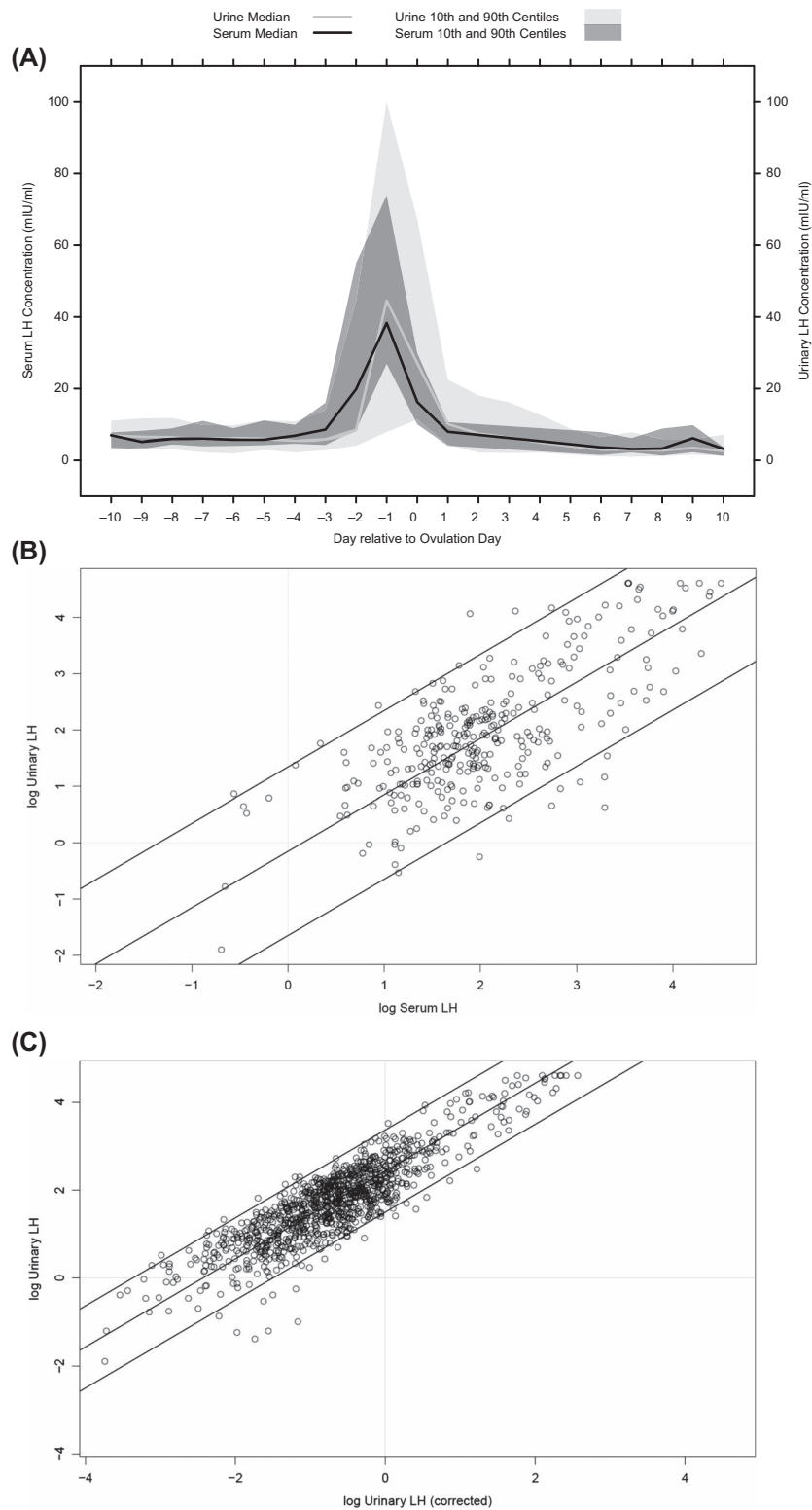
**Table 2** Mean serum hormone levels of each hormone measured, relative to the day of ovulation.\*

Day relative to ovulation*	n	Total LH (mIU/ml) Median (10th–90th percentile)	E2 (pmol/l) Median (10th–90th percentile)	Progesterone (nmol/l) Median (10th–90th percentile)
– 16	4	5.2 (4.4–8.5)	161.9 (106.7–177.5)	2.4 (1.8–2.7)
– 15	1	5.5 (5.5–5.5)	83.7 (83.7–83.7)	1.9 (1.9–1.9)
– 14	3	6.4 (5.2–12.5)	164.8 (156.3–166.3)	1.0 (0.8–2.4)
– 13	5	7.5 (5.6–9.8)	154.5 (128.0–165.2)	1.6 (0.9–1.9)
– 12	4	6.7 (3.6–14.0)	153.1 (143.4–168.1)	2.1 (1.7–2.2)
– 11	8	5.0 (3.9–7.5)	162.1 (81.9–218.1)	1.5 (0.3–2.2)
– 10	11	7.0 (3.5–7.8)	215.9 (135.1–258.8)	1.5 (0.5–2.0)
– 9	11	5.1 (3.1–8.2)	175.1 (136.2–269.5)	1.4 (0.3–2.5)
– 8	19	6.0 (4.4–8.9)	238.6 (138.5–351.9)	1.0 (0.4–2.6)
– 7	14	6.0 (3.9–11.0)	259.6 (157.6–334.6)	0.8 (0.5–1.6)
– 6	15	5.8 (4.0–9.0)	309.8 (199.3–608.1)	1.3 (0.5–2.6)
– 5	18	5.8 (4.2–11.1)	516.3 (262.8–626.9)	1.4 (0.5–2.1)
– 4	18	6.9 (4.6–9.9)	446.9 (337.3–809.1)	0.9 (0.4–2.4)
– 3	27	8.6 (4.3–16.1)	779.7 (458.4–1144.5)	1.3 (0.6–2.3)
– 2	20	19.8 (8.5–55.0)	954.6 (580.9–1359.4)	1.9 (0.8–3.8)
– 1	30	38.3 (26.9–74.0)	853.1 (489.3–1220.8)	3.1 (2.0–5.0)
0	26	16.3 (10.1–29.9)	301.6 (177.9–494.1)	4.6 (2.9–9.7)
1	6	8.0 (4.1–10.6)	314.6 (248.3–448.2)	10.3 (7.2–15.6)
2	4	5.5 (1.6–14.5)	478.7 (268.0–595.1)	20.7 (8.0–32.9)
3	1	5.8 (5.8–5.8)	266.1 (266.1–266.1)	22.7 (22.7–22.7)
4	1	4.5 (4.5–4.5)	720.6 (720.6–720.6)	33.4 (33.4–33.4)
5	1	3.1 (3.1–3.1)	687.2 (687.2–687.2)	30.8 (30.8–30.8)
6	18	3.6 (1.4–7.9)	394.6 (274.4–675.8)	36.3 (20.2–63.9)
7	14	3.1 (2.2–6.2)	409.3 (266.1–681.5)	32.0 (21.8–63.1)
8	18	3.3 (1.3–8.9)	521.1 (355.1–698.9)	32.7 (21.4–56.7)
9	13	6.2 (2.3–9.8)	461.8 (286.0–672.9)	33.1 (17.1–63.4)
10	6	3.2 (1.2–3.7)	373.3 (215.3–468.1)	21.3 (11.0–29.7)
11	2	1.9 (1.8–2.1)	302.1 (200.5–403.7)	17.1 (9.8–24.4)

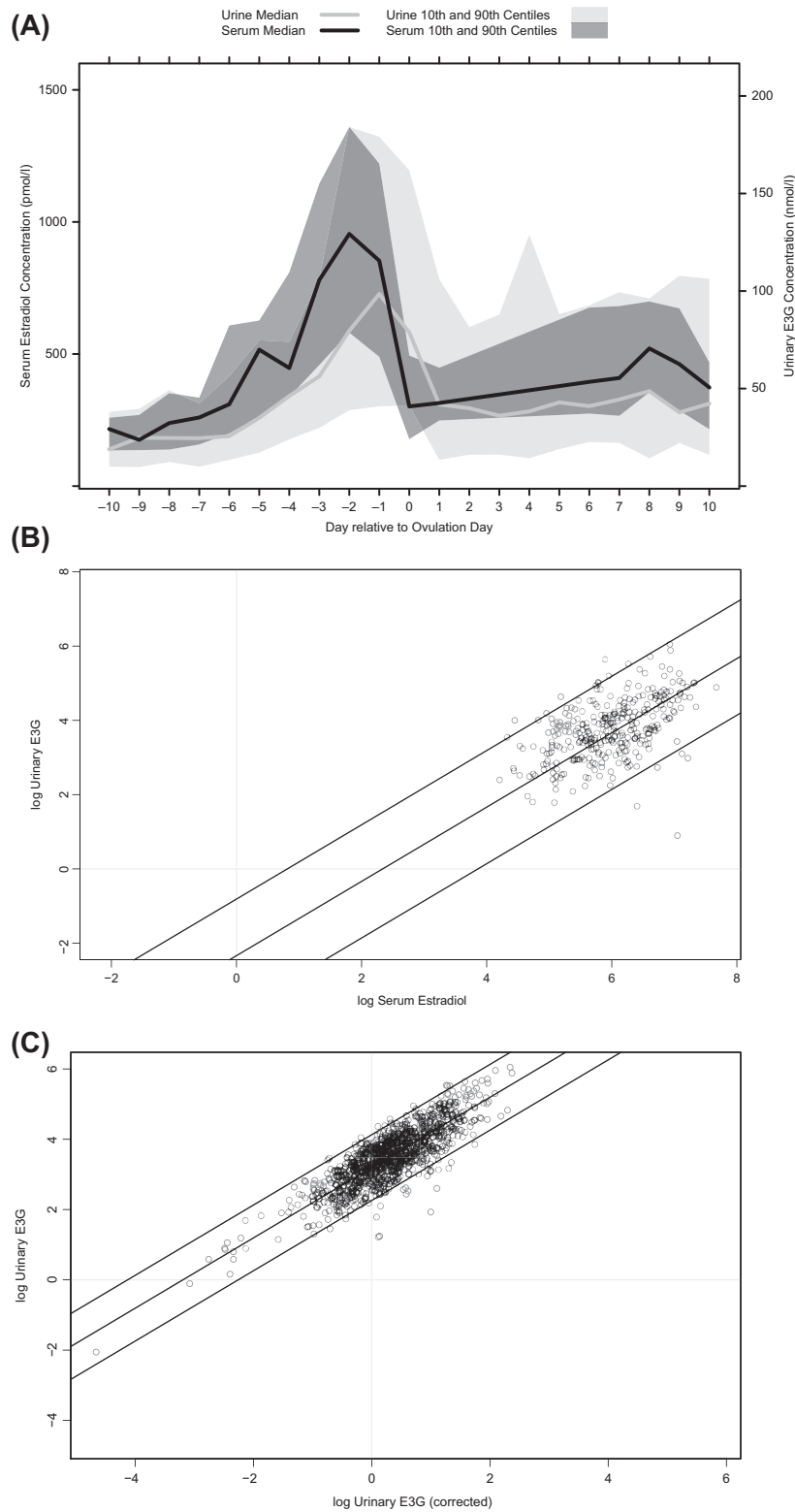
\*As determined by ultrasound, days rounded down to whole days. *Note:* Daily taking of blood samples and daily ultrasound were not possible for all volunteers, resulting in some days having < 10 serum samples on which to base the analysis. The high agreement of urinary and serum hormone profiles proves that the visits were timed correctly.

Physiologically, E3G and E<sub>2</sub> show at least two peaks, one in the follicular phase and one in the luteal phase of the menstrual cycle. The follicular phase peaks of E<sub>2</sub> and E3G were always higher than the luteal phase peaks of E<sub>2</sub> and E3G and were detected before and after ovulation. The mean time for the E3G peak was  $+0.98 \pm 4.46$  days after ovulation, with a range for the follicular E3G peak to objective ovulation from  $-5.5$  days to 19.5 days. The mean time for the E<sub>2</sub> peak was  $-0.63 \pm 2.70$  days before/after objective ovulation (ranging from  $-2.5$  before to 8 days after ovulation). Urinary P3G peaked  $7.65 \pm 2.35$  days after objective ovulation, ranging from 2 to 25.5 days after ovulation.

The differences in days between the first rise in reproductive hormones above baseline (definition: surge) and objective ovulation (5th–95th percentile, median) are shown in Figure 5. The hormonal surges in serum LH, urinary LH, E<sub>2</sub> and E3G were always preovulatory. The urinary LH surge occurred  $0.81 \pm 0.89$  days prior to ovulation with a range across the study group of only 4 days: a sharp and clear signal. The rise in urinary E3G occurred  $5.25 \pm 1.84$  days prior to ovulation, delivering a broad signal of 10 days across the study group, due to waves of rising estrogen in some cycles. For the serum hormones we detected an LH surge  $2.20 \pm 0.92$  days prior to ovulation with

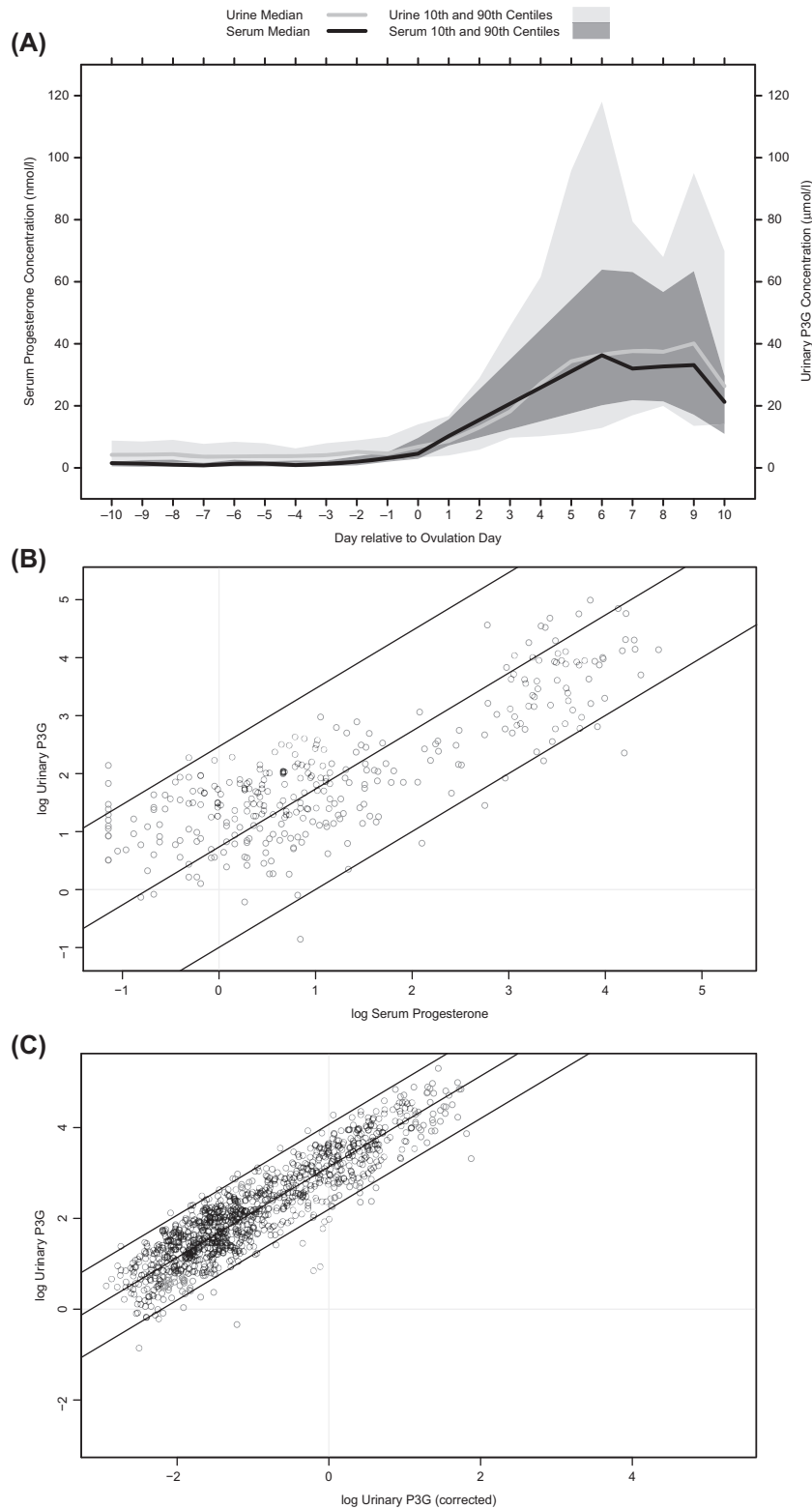


**Figure 1** (A) Serum (dark grey, mIU/ml and urinary (light grey, mIU/ml) LH ranges referenced to the day of objective ovulation (median and 10th–90th percentile range). (B) Scatter plot, including prediction lines and 95% prediction limits, of urinary LH against serum LH. (C) Scatter plot, including prediction lines and 95% prediction limits, of urinary LH not creatinine-corrected against urinary LH creatinine-corrected.

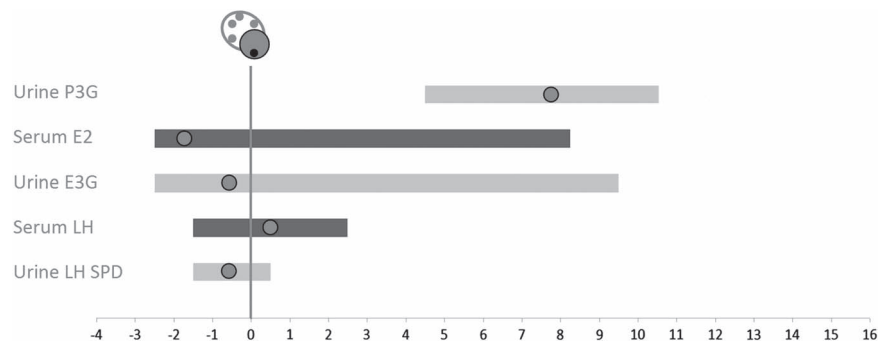


**Figure 2** (A)  $E_2$  (dark grey, pmol/l) and E3G (light grey, nmol/l) nomogram referenced to the day of objective ovulation (median and 10th–90th percentile range). (B) Scatter plot, including prediction lines and 95% prediction limits, of urinary E3G against serum  $E_2$ . (C) Scatter plot, including prediction lines and 95% prediction limits, of urinary E3G not creatinine-corrected against urinary E3G creatinine-corrected.





**Figure 3** (A) Progesterone (dark grey, nmol/l) and P3G (light grey, µmol/l) ranges referenced to the day of objective ovulation (median and 10th–90th percentile range). (B) Scatter plot, including prediction lines and 95% prediction limits, of urinary P3G against serum progesterone. (C) Scatter plot, including prediction lines and 95% prediction limits, of urinary P3G not creatinine-corrected against urinary P3G creatinine-corrected.



**Figure 4** Peaks of reproductive hormones in relation to objective ovulation. Median (dark dot); bars: 5th–95th percentile. For serum peaks there might be a small imprecision: in 11 volunteers, a serum sample was not collected on the day following peak LH level, so the LH peak may not be exact to the nearest day. For peak E<sub>2</sub> levels, a sample was not collected on the following day in nine volunteers.

a range of 4 days. The rise in serum E<sub>2</sub> occurred  $5.75 \pm 1.84$  days prior to ovulation with a range of 9 days. The difference in time between the LH surge following the estrogen rise was relatively constant at  $3.05 \pm 2.13$  days in urine (5th–95th percentile: 0–5.1 days, range 13 days) and  $3.59 \pm 1.96$  days in serum (5th–95th percentile: 0.8–6 days, range 11 days). The rise in urinary P3G occurred  $3.20 \pm 1.07$  days after ovulation, showing a relatively sharp, clear signal with a range of only 5 days. After ovulation, serum progesterone was only measured twice on days +7 and +9.

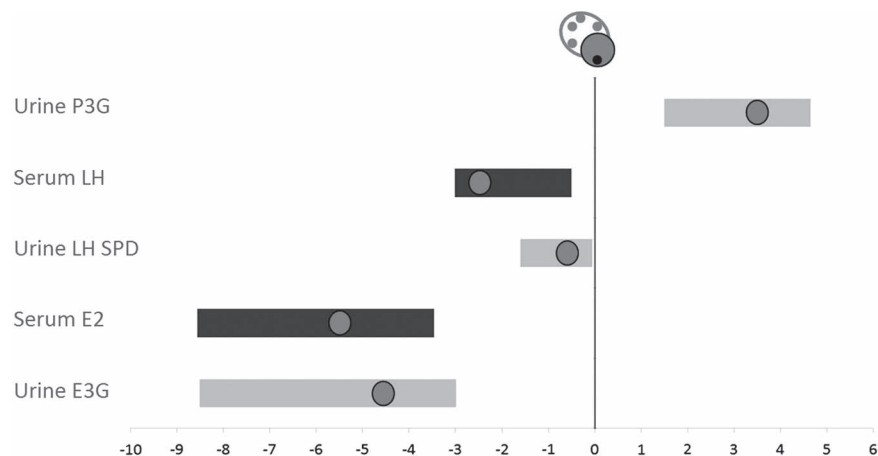
DISCUSSION

**Findings and interpretation**

The study found that serum and urinary hormone profiles showed excellent agreement, with a short

delay for the urinary signal; they can thus be used interchangeably. Estrogen and LH hormone peaks in urine and serum showed a range of signal characteristics across the study group and may occur after ovulation. The estrogen and LH hormone surges were always preovulatory, with the rise in estrogen occurring around 5 days before ovulation, followed by the LH surge 3–4 days later, delivering a very sharp signal. The LH surge started 1 day before, and never after, objective ovulation. Objective ovulation was followed by a rise in progesterone around 3 days later. Both the LH and progesterone surges delivered sharp signals across the whole study group and allowed reliable detection and confirmation of the key event, ovulation.

Figures 1–3 clearly show that urinary and serum reproductive hormones are highly correlated and show



**Figure 5** Reproductive hormone surges in relation to objective ovulation. Median (dark dot); bars: 5th–95th percentile. The hormonal surges in E3G, E<sub>2</sub>, P3G and progesterone and in urinary and serum LH were defined as the first rise above the best-fit line of previous measurements.

excellent agreement (Figures 1A, 1B; 2A, 2B; 3A, 3B) and may therefore be used interchangeably. Creatinine correction was not considered necessary to exclude volume or dilution effects if the first morning urine was used for analysis (Figures 1C, 2C, 3C). This is partly due to the fact that the changes in fertility hormone levels during the menstrual cycle are so dynamic that the volume and concentration of urine does not impact on the ability to observe these changes. Furthermore, the use of creatinine correction has previously been shown to be unnecessary when evaluating a trend analysis rather than absolute values<sup>15</sup>.

If measurements in urine are compared with serum measurements from the day before, due to an expected delay in urine signals, the correlation is slightly decreased for LH, revealing that there is no relevant delay. For E3G and P3G the coefficients of correlation were slightly increased, demonstrating the expected delay by steroid metabolism, which is less than 1 day.

This study confirms that considerable interindividual variations exist in hormonal profiles and the day of ovulation in women with apparently normal cycles. Having a 'normal' menstrual cycle history does not guarantee the cycle is ovulatory, and even very short or long cycles can be fertile. A menstrual history alone is not enough to diagnose regular ovulatory function and neither normal nor abnormal menstrual history can be relied on for assessment of fertility status<sup>7</sup>. The MeMo study delivers new ranges for reproductive hormones and compares urinary and serum reproductive hormone profiles. Established and widely used reference values in serum for different phases of the menstrual cycle have serious limitations because they were based on only 20 volunteers and the day of ovulation was referenced to the LH peak and not to independently determined objective ovulation<sup>19</sup>.

Analysis of urinary hormone profiles obtained by the use of different assays reveals the dependency on the assay system<sup>20</sup>. We have recently reported that the profile of the LH surge is highly assay-dependent<sup>14</sup>. Assays measuring intact LH, such as the SPD urinary LH assay reported here, provide physiologically relevant information, whereas those that measure total LH (including the beta-core LH derivative) provide profiles that do not accurately reflect menstrual cycle physiology.

### Relevance of the findings

The MeMo study provides evidence that urinary and serum hormone monitoring are both useful and

reliable means of predicting and confirming ovulation. The LH surge, 3 to 4 days after the estrogen rise, is the most effective and prospective marker of impending ovulation in  $0.81 \pm 0.89$  days time. LH peaks do not always precede ovulation and are therefore unreliable predictors of ovulation. The progesterone rise provides an excellent marker for the end of the fertile period and confirms ovulation. In this respect, the MeMo study underlines that monitoring of menstrual cycles using urinary metabolites of reproductive hormones is a very promising method of ovulation detection and confirmation, either for contraceptive purposes or to achieve pregnancy and to help in scheduling clinic visits<sup>8,21,22</sup>. The chronology of serum hormonal surges in menstrual cycles was analysed approximately three decades ago<sup>23–25</sup>. Surges were defined as a two- to three-fold increase in hormonal levels, so do not truly represent the surge. Ovulation was determined using less reliable methods such as laparoscopy. Therefore, these early results with considerably shorter time spans between the rise in estrogen and LH and ovulation are unreliable.

The recently published BioCycle study, which did not provide serum and laboratory-performed urinary hormone analysis<sup>26</sup> with confirmation of ovulation based on luteal progesterone measurements, reported similar levels of variability of menstrual cycles. However, the MeMo study goes a stage further, as objective evidence of ovulation is presented; thus, it provides the best contemporary data to date regarding hormone levels in relation to true ovulation, with findings regarding the temporal relationship of serum LH and E<sub>2</sub> to ovulation, comparable to the results of the study by Behre *et al.*<sup>27</sup>, which did not report hormonal ranges or examine urine concentration.

### Strengths and weaknesses of the study

The key strength of the MeMo study was the use of modern ultrasound equipment for daily follicular monitoring, which provided high-resolution, reliable follicle counts of even small antral follicles. It also used modern/new laboratory assays for measurement of serum and urinary reproductive hormones to examine the precise relationships and variations in hormone levels in relation to ultrasound-observed objective ovulation.

This study also had some limitations. Daily ultrasound prior to ovulation was not possible for

all volunteers and therefore for some volunteers the day of ovulation was estimated, introducing an error of  $\pm 1$  day. Daily taking of blood samples was also not possible for all volunteers, resulting in some days having  $< 10$  samples on which to base the analysis. Nevertheless, the agreement of urinary and serum hormone profiles was high and presumably would have been even higher with more frequent testing and more samples, especially in the luteal phase. On the other hand, MeMo proves that two measurements of progesterone on day 7 and day 9 after ovulation are sufficient to assess luteal function.

The volunteer recruitment targeted women with self-reported normal menstrual cycles; therefore, the study might have underrepresented the extent of population variability and frequency of cycle disturbances, e.g., anovulation. Another weakness is that 'only' 40 women (volunteers) could be included, due to the difficulty in recruiting women who are willing to undergo invasive study procedures, such as repeated transvaginal ultrasound and the frequent taking of blood samples. This is why an extension of the MeMo study to subsequent cycles to assess intraindividual variations was not possible.

### Future research

Urinary hormone measurements have been found to be a reliable tool for endocrine cycle monitoring. An analysis of multiple, sequential cycles from 'normally' menstruating women or women with cycle disorders would be of great interest if the monitoring could be performed simply by analysing daily morning urine samples. This would enable the

development of algorithms that could be used to prospectively identify hormonal surges. Accurate urinary monitoring may one day replace blood tests, and direct data transfer from live monitoring of menstrual cycles would also be of considerable benefit to clinics.

### CONCLUSIONS

The results reported here demonstrate that urinary and serum reproductive hormone levels show excellent agreement and provide reliable cycle monitoring and scheduling of clinic visits. The MeMo study also provides new ranges for urinary and serum reproductive hormones<sup>14</sup>, validating the principle of endocrine monitoring of menstrual cycles for contraception, achievement of pregnancy or examination of cycle abnormalities.

### ACKNOWLEDGEMENTS

The authors wish to thank all team members of green-ivf, especially Silvia Heil and Jennifer Neye, for helping to recruit volunteers, providing volunteer support during the study, and carrying out data recording and documentation work.

**Declaration of interest:** SPD Development Company Ltd and green-ivf provided financial funding for the whole study. Beckman-Coulter GmbH (Sinsheim, Germany) provided support with free AMH Generation II assays.

### REFERENCES

- Mihm M, Gangooly S, Muttukrishna S. The normal menstrual cycle in women. *Anim Reprod Sci* 2011; 124:229–36.
- Dasharathy SS, Mumford SL, Pollack AZ, et al. Menstrual bleeding patterns among regularly menstruating women. *Am J Epidemiol* 2012;175:536–45.
- Gnoth C, Frank-Herrmann P, Schmoll A, et al. Cycle characteristics after discontinuation of oral contraceptives. *Gynecol Endocrinol* 2002;16:307–17.
- Gnoth C, Godehardt E, Frank-Herrmann P, et al. Definition and prevalence of subfertility and infertility. *Hum Reprod* 2005;20:1144–7.
- Cole LA, Ladner DG, Byrn FW. The normal variabilities of the menstrual cycle. *Fertil Steril* 2009; 91:522–7.
- Johnson SR, Miro F, Barrett S, Ellis JE. Levels of urinary human chorionic gonadotrophin (hCG) following conception and variability of menstrual cycle length in a

- cohort of women attempting to conceive. *Curr Med Res Opin* 2009;25:741–8.
7. Creinin MD, Keverline S, Meyn LA. How regular is regular? An analysis of menstrual cycle regularity. *Contraception* 2004;70:289–92.
  8. Howards PP, Schisterman EF, Wactawski-Wende J, et al. Timing clinic visits to phases of the menstrual cycle by using a fertility monitor: the BioCycle Study. *Am J Epidemiol* 2009;169:105–12.
  9. Treloar AE, Boynton RE, Behn BG, Brown BW. Variation of the human menstrual cycle through reproductive life. *Int J Fertil* 1967;12:77–126.
  10. Blockeel C, Knez J, Polyzos NP, et al. Should an intrauterine insemination with donor semen be performed 1 or 2 days after the spontaneous LH rise? A prospective RCT. *Hum Reprod* 2014;29:697–703.
  11. Frank-Herrmann P, Heil J, Gnoth C, et al. The effectiveness of a fertility awareness based method to avoid pregnancy in relation to a couple's sexual behaviour during the fertile time: a prospective longitudinal study. *Hum Reprod* 2007;22:1310–19.
  12. Zinaman M, Johnson S, Ellis J, Ledger W. Accuracy of perception of ovulation day in women trying to conceive. *Curr Med Res Opin* 2012;28:749–54.
  13. Bilian X, Heng Z, Shang-chun W, et al. Conception probabilities at different days of menstrual cycle in Chinese women. *Fertil Steril* 2010;94:1208–11.
  14. Johnson S, Weddell S, Godbert S, et al. Development of the first urinary reproductive hormone ranges referenced to independently determined ovulation day. *Clin Chem Lab Med* 2015. E-Pub. ahead of print doi 10.1515/cclm-2014-1087.
  15. Miro F, Coley J, Gani MM, et al. Comparison between creatinine and pregnanediol adjustments in the retrospective analysis of urinary hormone profiles during the human menstrual cycle. *Clin Chem Lab Med* 2004;42:1043–50.
  16. Carstensen B, Gurrin L, Ekstrom C, Figurski M. MethComp: Functions for analysis of agreement in method comparison studies. R package version 1.22. 2013. Available at <http://CRAN.R-project.org/package=MethComp>
  17. Ecochard R, Marret H, Rabilloud M, et al. Sensitivity and specificity of ultrasound indices of ovulation in spontaneous cycles. *Eur J Obstet Gynecol Reprod Biol* 2000;91:59–64.
  18. Baerwald AR, Adams GP, Pierson RA. Ovarian antral folliculogenesis during the human menstrual cycle: A review. *Hum Reprod Update* 2012;18:73–91.
  19. Stricker R, Eberhart R, Chevailler M-C, et al. Establishment of detailed reference values for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the Abbott ARCHITECT analyzer. *Clin Chem Lab Med* 2006;44:883–7.
  20. Direito A, Bailly S, Mariani A, Ecochard R. Relationships between the luteinizing hormone surge and other characteristics of the menstrual cycle in normally ovulating women. *Fertil Steril* 2013;99:279–85.
  21. Blackwell LF, Brown JB, Cooke D. Definition of the potentially fertile period from urinary steroid excretion rates. Part II. A threshold value for pregnanediol glucuronide as a marker for the end of the potentially fertile period in the human menstrual cycle. *Steroids* 1998;63:5–13.
  22. Blackwell LF, Vigil P, Cooke DG, et al. Monitoring of ovarian activity by daily measurement of urinary excretion rates of oestrone glucuronide and pregnanediol glucuronide using the Ovarian Monitor, Part III: variability of normal menstrual cycle profiles. *Hum Reprod* 2013;28:3306–15.
  23. World Health Organization, Task Force on Methods for the Determination of the Fertile Period, Special Programme of Research, Development and Research Training in Human Reproduction. Temporal relationships between ovulation and defined changes in the concentration of plasma estradiol-17 beta, luteinizing hormone, follicle-stimulating hormone, and progesterone. I. Probit analysis. *Am J Obstet Gynecol* 1980;138:383–90.
  24. Pauerstein CJ, Eddy CA, Croxatto HD, et al. Temporal relationships of estrogen, progesterone, and luteinizing hormone levels to ovulation in women and infrahuman primates. *Am J Obstet Gynecol* 1978;130:876–86.
  25. Kerin J. Ovulation detection in the human. *Clin Reprod Fertil* 1982;1:27–54.
  26. Hambridge HL, Mumford SL, Mattison DR, et al. The influence of sporadic anovulation on hormone levels in ovulatory cycles. *Hum Reprod* 2013;28:1687–94.
  27. Behre HM, Kuhlage J, Gassner C, et al. Prediction of ovulation by urinary hormone measurements with the home use ClearPlan Fertility Monitor: Comparison with transvaginal ultrasound scans and serum hormone measurements. *Hum Reprod* 2000;15:2478–482.