Intercycle variability of ovarian reserve tests: results of a prospective randomized study

J.Kwee1,2, R.Schats1, J.McDonnell1, C.B.Lambalk1 and J.Schoemaker1

1Division of Reproductive Endocrinology and Fertility and the IVF Centre, Department of Obstetrics and Gynaecology, Vrije Universiteit Medical Centre, PO Box 7057, 1007 MB Amsterdam, The Netherlands
2To whom correspondence should be addressed. E-mail: j.kwee@vumc.nl

BACKGROUND: This study was designed to assess prospectively the intercycle variability (ICV) of basal FSH (bFSH), clomiphene citrate challenge test (CCCT) (analysis of the CCCT was performed by the parameter: ∑bFSH + sFSH) and exogenous FSH ovarian reserve test (EFORT) (analysis of the EFORT included the following parameters: estradiol (E2) increment and inhibin B increment 24 h after administration of FSH), and secondarily to assess the influence of the variability of these ovarian reserve tests. METHODS: Eighty-five regularly menstruating patients, aged 18–39 years, participated in this prospective study, randomized, by a computer-designed four-blocks system into two groups. Forty-three patients underwent a CCCT, and 42 patients underwent an EFORT. Each test was performed 1–4 times in subsequent cycles, one test per cycle. During the first three cycles, patients were treated with intrauterine insemination (IUI). Follicle number and oocyte yield during IVF ovarian stimulation in the fourth cycle were taken as measures for ovarian reserve. RESULTS: The per cycle variance of bFSH ranged from 1.8 to 4.4 (maximum to minimum ratio of 2.44, \(P < 0.0001\)), while that of CCCT ranged from 21.3 to 70.6 (3.31, \(P < 0.0001\)). No significant change in per cycle variance was found for the E2 increment (1.25, \(P > 0.2\)) and inhibin B increment (1.31, \(P > 0.2\)), which were the EFORT parameters. A large ICV of CCCT and bFSH test results was strongly associated with lower ovarian reserve. CONCLUSIONS: Our study shows that the ICV of the inhibin B increment and the E2 increment in the EFORT is stable in consecutive cycles, which indicates that this reproducible test is a more reliable tool for determination of ovarian reserve than bFSH and CCCT. Women with limited ovarian reserve show a strong ICV of bFSH and FSH response to clomiphene citrate.

Key words: bFSH/CCCT/EFORT/intercycle variability/ovarian reserve

Introduction
Basal FSH (bFSH) (Scott et al., 1989), clomiphene citrate challenge tests (CCCTs) (Navot et al., 1987) and the exogenous FSH ovarian reserve test (EFORT) (Fanchin et al., Kwee et al., 2003) have been shown to be of predictive value for the ovarian reserve with respect to stimulation and pregnancy rates in patients undergoing IVF (Sharara et al., 1998). Kwee et al. (2003) compared the predictive value of bFSH, CCCT and the EFORT on the outcome of ovarian stimulation in IVF treatment, and concluded that the EFORT was the best endocrine test for the prediction of ovarian reserve. A small number of studies have documented the intercycle variability of bFSH and CCCT in the same patient. Scott et al. (1990) documented the intercycle variability of FSH and found it to vary from patient to patient. Hannoun et al. (1998) documented a high degree of intercycle variability of the CCCT when performed in the same patient.

Knowledge of the intercycle variability of an ovarian reserve test is important for correct interpretation of test results. No data at all are available about the intercycle variability of the EFORT. The purpose of our study was to assess prospectively the intercycle variability of bFSH, CCCT and EFORT, and secondarily to assess the influence of the variability of these tests on the prediction of ovarian reserve (ovarian capacity), defined by us as the maximal number of follicles which can be stimulated under maximal ovarian stimulation with FSH.

Materials and methods
Study population
This study is part of a prospective randomized study of regularly menstruating patients on the determination of ovarian capacity, called the DOC study. From June 1997 to May 1999, 85 patients aged 18–39 years who were eligible for intrauterine insemination (IUI) entered the study. Their infertility was either idiopathic for >3 years and/or due to a male factor and/or cervical hostility. Patients had to have regular menstrual cycles, two ovaries and at least one patent Fallopian tube. Excluded were patients with either polycystic ovary syndrome or a severe male factor, subsequently treated by ICSI and defined as (i) <1 × 10^6 motile spermatozoa after Percoll centrifugation (gradient 40/90); and/or (ii) >20% antibodies present on the spermatozoa after
processing with Percoll centrifugation (gradient 40/90); and/or (iii) >50% of the spermatozoa without an acrosome. Other exclusion criteria were untreated or insufficiently corrected endocrinopathies, clinically relevant systemic diseases or a body mass index >28 kg/m².

The study protocol was approved by the Committee on ethics of research involving human subjects of the Vrije Universiteit Medical Centre, Amsterdam, The Netherlands. Informed consent was signed by all the couples participating in the study.

**Treatment protocol**

Patients were randomized by a computer-designed four-blocks system into two groups of 43 patients, for the study of the CCCT, and 42 patients for the study of the EFORT. Each test was performed once per cycle for up to four cycles. During the first three cycles, patients were treated with IU; in the fourth cycle, IVF with maximal ovarian stimulation was performed. The IVF cycle had to be initiated within a year from the first cycle.

It seems that in IVF ovarian stimulation, the maximal effect is reached with FSH dosages up to 225 IU per day (Out et al., 2000, 2001; Latin-American Puregon IVF Study Group, 2001). Using these results, we concluded that an initial stimulation by three ampoules of 75 IU of FSH under a long (GnRH agonist suppressed) protocol probably gives a maximal IVF stimulation, the outcome of which, in terms of number of follicles and oocytes, could be used as the gold standard for the cohort size.

**Clomiphene citrate challenge test (CCCT).** Starting on the fifth day of the menstrual cycle (CD 1 = day of onset of menses), 100 mg of clomiphene citrate (Serophene®, Serono, Geneva, Switzerland) was administered for 5 days. In this study, on CD 2 or 3 (basal values) and on CD 10 (stimulated values), the serum FSH (sFSH) was determined. Analysis of the CCCT (Kwee et al., 2003) was performed by the parameter: ΣbFSH + sFSH.

**Exogenous FSH ovarian reserve test (EFORT).** On CD 3, 300 IU of recFSH (Gonal-F®, Serono) were administered s.c. In this study, blood samples for the determination of FSH, estradiol (E₂) and inhibin B were drawn: just before (basal values) and 24 h after (stimulated values) the administration of FSH. Analysis of the EFORT (Kwee et al., 2003) included the following parameters: E₂ increment and inhibin B increment 24 h after administration of FSH.

The bFSH level was determined as an integral part of all CCCTs and EFORTs. All tests (CCCT and EFORT) during the first three test cycles were followed by regular IU cycles. Each cycle was monitored with serial transvaginal sonography (TVS) to evaluate if clomiphene citrate caused multifollicular growth and what the effect of a single injection with 300 IU of recFSH in the early follicular phase was. When the leading follicle reached a diameter of 18–20 mm (measured in two perpendicular directions), 10 000 IU of HCG (Profasi®, Serono) was administered to induce final follicular maturation. The IU incremental rise and inhibin B increment 24 h after administration of FSH.

The bFSH level was determined as an integral part of all CCCTs and EFORTs. All tests (CCCT and EFORT) during the first three test cycles were followed by regular IU cycles. Each cycle was monitored with serial transvaginal sonography (TVS) to evaluate if clomiphene citrate caused multifollicular growth and what the effect of a single injection with 300 IU of recFSH in the early follicular phase was. When the leading follicle reached a diameter of 18–20 mm (measured in two perpendicular directions), 10 000 IU of HCG (Profasi®, Serono) was administered to induce final follicular maturation. The IU incremental rise and inhibin B increment 24 h after administration of FSH.

The bFSH level was determined as an integral part of all CCCTs and EFORTs. All tests (CCCT and EFORT) during the first three test cycles were followed by regular IU cycles. Each cycle was monitored with serial transvaginal sonography (TVS) to evaluate if clomiphene citrate caused multifollicular growth and what the effect of a single injection with 300 IU of recFSH in the early follicular phase was. When the leading follicle reached a diameter of 18–20 mm (measured in two perpendicular directions), 10 000 IU of HCG (Profasi®, Serono) was administered to induce final follicular maturation. The IU incremental rise and inhibin B increment 24 h after administration of FSH.

The bFSH level was determined as an integral part of all CCCTs and EFORTs. All tests (CCCT and EFORT) during the first three test cycles were followed by regular IU cycles. Each cycle was monitored with serial transvaginal sonography (TVS) to evaluate if clomiphene citrate caused multifollicular growth and what the effect of a single injection with 300 IU of recFSH in the early follicular phase was. When the leading follicle reached a diameter of 18–20 mm (measured in two perpendicular directions), 10 000 IU of HCG (Profasi®, Serono) was administered to induce final follicular maturation. The IU incremental rise and inhibin B increment 24 h after administration of FSH.

The bFSH level was determined as an integral part of all CCCTs and EFORTs. All tests (CCCT and EFORT) during the first three test cycles were followed by regular IU cycles. Each cycle was monitored with serial transvaginal sonography (TVS) to evaluate if clomiphene citrate caused multifollicular growth and what the effect of a single injection with 300 IU of recFSH in the early follicular phase was. When the leading follicle reached a diameter of 18–20 mm (measured in two perpendicular directions), 10 000 IU of HCG (Profasi®, Serono) was administered to induce final follicular maturation. The IU incremental rise and inhibin B increment 24 h after administration of FSH.

**IVF treatment.** The ovarian stimulation protocol was performed according to a long GnRH agonist protocol starting in the mid-luteal phase. On CD 3 of the first cycle, the CCCT or the EFORT was performed as described above. In the subsequent midluteal phase, 7 days after ovulation, daily s.c. injections with triptoreline acetate (Decapeptyl®, 0.1 mg/day; Ferring, Hoofddorp, The Netherlands) were started. Because of the administration of the GnRH agonist, patients were advised to use a barrier type of contraception during this cycle. On CD 3 of the next cycle, ovarian stimulation was started with daily s.c. injections of a fixed dose of 225 IU of uFSH (Metrodin HP®, 75 IU/ampoule; Serono), because this dosage probably gives a maximal effect in follicle stimulation. Standard procedures were followed including TVS (Alok SSD-1700, 5.0 MHz probe) on CD 2 or 3 and on CD 9 or 10. Daily TVS was performed from the moment when the leading follicle reached a diameter of 16 mm. Ovarian stimulation was continued until the largest follicle reached a diameter of at least 18 mm. The maximum duration of uFSH administration allowed was 16 days. If these criteria were met, Metrodin HP® and Decapeptyl® were discontinued and 10 000 IU of HCG (Profasi®, 10 000 IU/ampoule; Serono) were administered. On the day of HCG, TVS was performed to count the result of ovarian stimulation (all follicles ≥10 mm) expressed as the total number of follicles.

**Serum assay**

Serum E₂ and FSH were determined by commercially available immunometric assays (Amerlite, Amersham, UK). For E₂, the inter- and intra-assay coefficient of variation (CV) was 11% at 250 pmoI/l and 8% at 8000 pmoI/l, the intra-assay CV was 13% at 350 pmoI/l, 9% at 1100 pmoI/l and 9% at 5000 pmoI/l. The lower limit of detection for E₂ was 90 pmoI/l. In the EFORT and CCCT, we measured E₂ by a sensitive radioimmunoassay (Sorin, Biomedica, Saluggia, Italy). For measurement of E₂ (abbreviated as EE), the inter-assay CV was 11% at 60 pmoI/l, 8% at 200 pmoI/l, 11% at 550 pmoI/l and 9% at 900 pmoI/l. The intra-assay CV was 4% at 110 pmoI/l and 5% at 1000 pmoI/l. The lower limit of detection for EE was 18 pmoI/l. For FSH, the inter-assay CV was 9% at 3 IU/l and 5% at 35 IU/l, the intra-assay CV was 9% at 5 IU/l, 8% at 15 IU/l and 6% at 40 IU/l. The lower limit of detection for FSH was 0.95 IU/l. Inhibin B was determined immunochemically by a commercially available assay (Serotec Ltd, Oxford, UK). For inhibin B, the inter-assay CV was 17% at 25 ng/l, 14% at 55 ng/l and 9% at 120 ng/l and the intra-assay CV was 8% up to 40 ng/l and 5% at >40 ng/l. The lower limit of detection for inhibin B was 13 ng/l.

Half way through the study, the Amerlite assay used to assess FSH was suddenly withdrawn from the market and had to be replaced by another commercially available assay (Delfia, Wallac, Finland). The two assays have been compared and showed excellent linear correlation, although a shift in the values took place [Delfia FSH = 1.28 × Amerlite FSH + 0.01 (r = 0.9964)]. For the Delfia FSH, the inter-assay CV was 5% at 3.5 IU/l and 3% at 15 IU/l. All FSH determinations have been recalculated and are expressed according to the Delfia assay. Values below the detection limit of an assay were assigned a value equal to the detection limit of that assay.

**Statistical analysis**

Descriptive statistics and univariate tests were carried out using SPSS for Windows. Data on intercycle variability were analysed using random coefficient models which are a generalization of ordinary regression models. Such models allow us to remove variation due to other factors, such as age. A brief description of random coefficient models and details of the models fitted are given in the Appendix.

The measure of intercycle variation used was the within-patient SD of each test variable over time. As an SD can only be defined when at least two measurements have been done, patients with only one measurement point were excluded from the analysis.

Secondly, we examined the relationship between intercycle variation and ovarian reserve by calculating the correlation coefficient between variability and ovarian reserve. For the ovarian reserve, we used the result of ovarian stimulation expressed as the total number of follicles.
retrieved oocytes as gold standard. That means that only patients with four tests were included in this part of the analysis.

**Results**

The characteristics of the two groups are given as means ± SD in Table I. No significant differences were noted between the groups in baseline characteristics, CD 3 measurements or outcome parameters. In the CCCT group, 69.8% had a primary infertility and 30.2% a secondary infertility. The cause of infertility was an idiopathic factor for 62.8%, a male factor for 30.2% and a cervical factor for 7.0%. In the EFORT group, 66.7% had a primary infertility and 33.3% a secondary infertility. The cause of infertility was an idiopathic factor for 63.4%, a male factor for 31.7% and a cervical factor for 4.9%.

As shown in Table II, in the CCCT group, six patients received one test, 10 received two tests, one received three tests and 26 received four tests. Six patients conceived after the first test, seven after the second test and one after the third. Three patients dropped out of the study for unknown reasons after two tests.

In the EFORT group, four patients became pregnant spontaneously after randomization and before starting with IUI treatment. Two patients received one test, one received two tests, two received three tests and 33 received four tests. After the first test, one patient conceived; another conceived following the second and two after the third test. One patient became pregnant in the first month of IVF treatment. One patient dropped out of the study for unknown reasons after one test.

To exclude bias in analysis of the patients groups due to pregnancy and drop out rates, we performed statistical analysis on the characteristics. No significant differences were noted between the two study groups during the study.

The regression analysis revealed no systematic change in mean level of the clinical variables over time (Table III). However, significant intercycle variation was seen in two of the four variables. The per cycle variance of bFSH ranged from 1.8 to 4.4 (maximum to minimum ratio of 2.44, \( P < 0.0001 \)), while that of CCCT ranged from 21.3 to 70.6 (3.31, \( P < 0.0001 \)). No significant change in per cycle variance was found for the E_2 increment (1.25, \( P > 0.2 \)) and inhibin B increment (1.31, \( P > 0.2 \)).

Table IV shows the correlation between ovarian reserve and intercycle variability. The intercycle variability of both CCCT and bFSH (Figure 1A and B) was strongly and negatively linked with ovarian reserve. This negative correlation indicates that women who have high intercycle variation in these factors tend to have a lower ovarian reserve. The intercycle variability of the E_2 increment showed no apparent correlation with ovarian reserve, while the intercycle variability of the inhibin B increment was positively correlated with ovarian reserve and was significant at the 6% level.

**Discussion**

Recently, we showed that the inhibin B increment and E_2 increment in the EFORT were the best predictors of the total

---

**Table I. Characteristics of the groups**

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>CCCT group ((n = 43))</th>
<th>EFORT group ((n = 42))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>33.0 ± 3.3</td>
<td>32.5 ± 3.7</td>
</tr>
<tr>
<td>Duration of infertility (years)</td>
<td>3.8 ± 1.3</td>
<td>4.1 ± 1.5</td>
</tr>
<tr>
<td>Cycle 1 day 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSH (IU/l)</td>
<td>7.0 ± 2.3</td>
<td>7.7 ± 2.2</td>
</tr>
<tr>
<td>E_2 (pmol/l)</td>
<td>130.7 ± 59.2</td>
<td>121.2 ± 85.8</td>
</tr>
<tr>
<td>Inhibin B (ng/l)</td>
<td>108.0 ± 45.8</td>
<td>92.3 ± 54.2</td>
</tr>
<tr>
<td>Treatment results cycle 4</td>
<td>( n = 26 )</td>
<td>( n = 33 )</td>
</tr>
<tr>
<td>No. of ampoules of FSH</td>
<td>32.0 ± 4.7</td>
<td>32.8 ± 7.9</td>
</tr>
<tr>
<td>E_2 level on the day of hCG (pmol/l)</td>
<td>15 116.5 ± 24 798.2</td>
<td>13 848.2 ± 22 688.9</td>
</tr>
<tr>
<td>Values are means ± SD. No significant differences.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table II. Breakdown of number of patients that became pregnant or dropped out during the subsequent IUI and IVF treatments divided over the two test groups**

<table>
<thead>
<tr>
<th>Total ( n = 85 )</th>
<th>Cycle 0</th>
<th>Cycle 1, IUI</th>
<th>Cycle 2, IUI</th>
<th>Cycle 3, IUI</th>
<th>Cycle 4, IVF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pregnant</td>
<td>Drop outs</td>
<td>Pregnant</td>
<td>Drop outs</td>
<td>Pregnant</td>
</tr>
<tr>
<td>CCCT ((n = 43))</td>
<td>0</td>
<td>43</td>
<td>6</td>
<td>37</td>
<td>7</td>
</tr>
<tr>
<td>EFORT ((n = 42))</td>
<td>4</td>
<td>38</td>
<td>1</td>
<td>36</td>
<td>1</td>
</tr>
</tbody>
</table>

*Pregnant after randomization but before treatment.

Pregnant = ongoing pregnancy defined as positive heartbeat by a gestational age of 12 weeks.
number of follicles obtained after maximal ovarian stimulation in an IVF treatment, i.e. cohort size (Kwee et al., 2003). Age, basal values of FSH, E_2 and inhibin B and the outcome of the CCCT in this respect, each, and in combination, showed a much lower performance. In order to allow correct interpretation of these tests, we here estimated their intercycle variability.

This study shows that the intercycle variability of the inhibin B and E_2 increment in the EFORT is not significantly different in consecutive cycles. This is in contrast to bFSH and CCCT which vary significantly from cycle to cycle. This outcome indicates that the EFORT potentially is a highly reproducible predictor for ovarian reserve.

The issue of intercycle variability of ovarian reserve tests as a marker for ovarian reserve has been studied in the past (Scott et al., 1990; Brown et al., 1995; Hannoun et al., 1998). Scott et al. (1990) evaluated and documented the intercycle variation of two cycles of bFSH. They found that two bFSH values that are in agreement might be used to counsel patients regarding their performance during ovarian stimulation. Women with a normal bFSH had a small range in the intercycle variation, in contrast to women with an elevated bFSH showing a much greater variation. If the patient had wide fluctuations in her bFSH values, she was more likely to respond poorly to stimulation. They therefore suggested serial screening of bFSH because they found the diagnostic and predictive value of a normal value of a single determination of bFSH limited. Brown et al. (1995) evaluated the intercycle variation of bFSH in a group of normally cycling women, unselected for fertility. They concluded that a single day 3 FSH level <20 IU/l is highly predictive of all subsequent values within a year in women under 40. Women over 40 years with a ‘normal’ day 3 level had a 50% chance of having an elevated day 3 FSH level in the subsequent cycle. In this group, the determination of more than one day 3 FSH level was likely to have prognostic significance.

Hannoun et al. (1998) documented the variation of the results of the CCCT performed in the same patient from cycle to cycle. They showed a high degree of intercycle variability. However, they did not test the influence of this variability on the ovarian reserve.

Our study confirms that the intercycle variation of bFSH and CCCT was strongly negatively correlated with the outcome of IVF. A small intercycle variability of bFSH and CCCT was associated with a ‘normal’ ovarian reserve. Patients with large fluctuating bFSH and CCCT results in consecutive cycles had fewer follicles after controlled stimulation. This confirms that high intercycle variability of bFSH and CCCT could act as a marker for low response as results from limited ovarian reserve. However, from our study, it appeared that intercycle variability of bFSH and CCCT outcome has no added value to a single measurement for prediction of ovarian reserve. This is probably because elevated bFSH and abnormal outcome of CCCT are coupled to the phenomenon of high intercycle variability. Women with one single ‘elevated’ bFSH show a high intercycle variability and have a high chance of a ‘poor’ response after ovarian stimulation, and women with one single ‘normal’ bFSH show a small intercycle variability and have a high chance for an ‘adequate’ response after ovarian stimulation (Scott et al., 1990). Comparing the prediction of intercycle variability of bFSH and CCCT with one single screening of bFSH and CCCT (Kwee et al., 2003), there was no added value of the intercycle variability for the prediction of ovarian reserve. Therefore, it seems that there is no need for evaluation of intercycle variability of bFSH and CCCT as a marker for ovarian reserve.

From a pathophysiological point of view, large intercycle variability in bFSH and CCCT, particularly in those patients with limited ovarian reserve, remains an intriguing phenomenon. Early follicular phase fluctuations in FSH are a reflection of the balance between ovarian steroid and peptide inhibition.

### Table III. Per cycle variance for each of the four variables

<table>
<thead>
<tr>
<th>Cycle</th>
<th>bFSH (IU/l)^a</th>
<th>CCCT (IU/l)^a</th>
<th>EFORT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inhibin B increment (ng/l)^b</td>
<td>E_2 increment (pmol/l)^b</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.1</td>
<td>21.3</td>
<td>2498</td>
</tr>
<tr>
<td>2</td>
<td>4.4</td>
<td>70.6</td>
<td>3003</td>
</tr>
<tr>
<td>3</td>
<td>3.0</td>
<td>40.6</td>
<td>2441</td>
</tr>
<tr>
<td>4</td>
<td>1.8</td>
<td>40.6</td>
<td>2296</td>
</tr>
</tbody>
</table>

^aP < 0.0001.

^bNot significant.

### Table IV. Correlation between ovarian capacity and intercycle variance

<table>
<thead>
<tr>
<th></th>
<th>bFSH (IU/l)</th>
<th>CCCT (IU/l)</th>
<th>EFORT</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>–0.52</td>
<td>–0.54</td>
<td>0.337</td>
</tr>
<tr>
<td>P-value</td>
<td>0.008</td>
<td>0.000</td>
<td>0.060</td>
</tr>
</tbody>
</table>
and the hypothalamic–pituitary drive during the period of follicular recruitment at the moment just before the selection of the dominant follicle. It signifies the amount of inhibin and/or E2 produced by the cohort of follicles, responsible for the negative feedback on the FSH secretion. The bFSH increases when ovarian reserve diminishes (Lenton et al., 1988), supposedly because the small antral follicles produce less inhibin B and possibly E2.

De Koning et al. (2000) found that elevated bFSH results from a pituitary more sensitive to GnRH, leading to higher FSH and LH pulses. Thus, this means that FSH in the early follicular phase as well as after stimulation with clomiphene is influenced in many ways and is potentially susceptible to large variations particularly with a small follicle cohort. Therefore, bFSH and CCCT can only act as a very indirect measure for the actual cohort size.

Conversely, why is the intercycle variability of the EFORT very small? EFORT is a very accurate predictor of ovarian response. Therefore, a likely explanation would be that the monthly available ovarian follicle cohort is surprisingly constant, even in women with a small cohort. In ~70 days of follicular development, the primary follicles reach the early antral stage. They become sensitive to FSH (Gougeon, 1996). At the end of the first 70 days (Schoemaker et al., 1993; Van der Meer et al., 1994; Scheele and Schoemaker, 1996), in the early follicular phase, a cohort of antral follicles, probably ~20 in number, each with a different sensitivity to FSH, is present in the ovaries, ready to continue their development under the stimulation of FSH. This is the stage in which we perform the EFORT. The EFORT is a direct test, which demonstrates the ability of the ovaries to initiate aromatase activity in response to a fixed dose of exogenously administered FSH (300 IU of FSH). The aromatase activity results in increased follicular concentrations of E2 since the aromatase substrate, androstenedione, is already available in abundance. Inhibin B is produced by the granulosa cells of the developing cohort of follicles and therefore directly reflects the ovarian reserve (Groome et al., 1996; Welt et al., 1997). Apparently the month to month variability of the cohort of follicles measuring 2–5 mm, present very early in the follicular phase of the cycle (Chang et al., 1998), is very small.

Indeed, Scheffer et al. (1999), using sonography, found a limited intercycle variability of antral follicle count (AFC) in regularly cycling women. So far, no other studies have been published that describe intercycle variability of AFC. Similarly, no studies have been published to date on intercycle variability of other available tests for ovarian reserve such as GnRH agonist stimulation test (GAST) (Padilla et al., 1990; Ravhon et al., 2000) and the measurement of anti-Müllerian hormone (AMH) (Vet et al., 2002). Ideally, a test for ovarian reserve should be short, simple, accurate and precise. We conclude that the EFORT in comparison with bFSH and CCCT has superior stable characteristics in terms of intercycle variability. This high reproducibility and its optimal performance to predict ovarian response (Kwee et al., 2003) make the EFORT a useful endocrine test to predict ovarian reserve.

Acknowledgements
The authors acknowledge the help of Dr Corry Popp-Snijders and her staff, particularly for the endocrine laboratory work, and the staff of the IVF centre for assistance during the execution of the protocol. This study was financially supported by an unrestricted grant from Serono, Geneva, Switzerland.

Appendix
Random coefficient models are a generalization of ordinary regression models. Consider the following regression model

\[ y = \alpha + \beta x + \varepsilon \]

where \( y \) is the dependent variable, \( x \) is an explanatory variable, \( \alpha \) and \( \beta \) are unknown coefficients and \( \varepsilon \) is the random error. Usually, the coefficients \( \alpha \) and \( \beta \) are considered as being fixed.
parameters, taking the same value for all individuals. However, in some situations, one or other of the coefficients (typically the coefficient associated with the explanatory variable) may vary from individual to individual. We may therefore choose to view $\beta$ as being a random variable with a given distribution (usually normal) with a fixed mean $\beta_0$ and variance $\sigma^2$, i.e. $\beta \sim N(\beta_0, \sigma^2)$.

Alternatively, we can write $\beta = \beta_0 + \beta_1$ where $\beta_0$ is fixed and $\beta_1 \sim N(0, \sigma^2)$.

Entering this expression into the regression model gives

$$y = \alpha + (\beta_0 + \beta_1)x + \varepsilon = \alpha + \beta_0x + (\varepsilon + \beta_1x)$$

Since $\beta_1$ is random, the term in parentheses is random and represents the random variation in the data. Usually, $\beta_1$ is assumed to be uncorrelated to $\varepsilon$.

The differences between the models are 2-fold. First, the observed variation is split into two components. Secondly, the amount of variation is to some degree dependent on the covariate $x$.

A hypothesis test of no $x$ effect on the variation (i.e. $H_0$: $\sigma^2 = 0$) can be carried out on the values of the log-likelihood functions of two models, one containing the extra term, the other without it.

In our analyses, we defined three binary variables, $T_2$, $T_3$ and $T_4$. $T_2$ took the value 1 if an observation corresponded to the second cycle and 0 if it did not. The other variables were defined analogously. Our model was therefore

$$y = \alpha + \Sigma (\beta_0 + \beta_1)T + \gamma (\text{age}) + (\varepsilon + \Sigma \beta_1 T)$$

where the summation is over the $T$ variables. Note that we are not really interested in the coefficients ($\beta_0$) but we have included them in the analyses.

References


Submitted on August 26, 2003; accepted on November 11, 203