Predictive Value and Clinical Impact of Basal Follicle-Stimulating Hormone in Subfertile, Ovulatory Women

Jan W. van der Steeg, Pieternel Steures, Marinus J. C. Eijkemans, J. Dik F. Habbema, Peter G. A. Hompes, Frank J. Broekmans, Peter X. J. M. Bouckaert, Patrick M. M. Bossuyt, Fulco van der Veen, and Ben W. J. Mol, on behalf of the Collaborative Effort for Clinical Evaluation in Reproductive Medicine (CECERM) Study Group*.

Context: Basal FSH is a marker for ovarian reserve.

Objectives: The objective of the study was to investigate the predictive value of basal FSH on spontaneous ongoing pregnancy in subfertile ovulatory women.

Design: This was a prospective cohort study.

Setting: The study was conducted in 19 fertility centers in The Netherlands.

Participants: Subfertile ovulatory women without two-sided tubal pathology and in whom the man had normal sperm parameters (total motile count > 3 x 10^8) participated in the study.

Interventions: Interventions included a fertility work-up, including a basal FSH measurement on cycle d 3.

Main Outcome Measures: Spontaneous ongoing pregnancy was measured.

Results: We included 3519 consecutive couples of which 562 (16%) had a spontaneous ongoing pregnancy within 1 yr. Basal FSH levels of 8 IU/liter or higher were associated with a decreased probability of spontaneous ongoing pregnancy [hazard ratio (HR) 0.93/IU liter (95% confidence interval [CI] 0.87–0.99)]. In a multivariable analysis, female age (HR 0.97/yr, 95% CI 0.95–0.99), cycle length (HR 0.96/d, 95% CI 0.93–1.0), and FSH levels 8 IU/liter or greater (HR 0.93/IU liter, 95% CI 0.87–0.99) were strong negative predictors for spontaneous ongoing pregnancy. Addition of FSH to a prediction model based on female age, duration of subfertility, previous pregnancy, referral status, and semen analysis changed the probability to conceive spontaneously from 30% or greater to less than 30% in 97 of 3219 couples (3.0%).

Conclusions: In ovulatory women, a basal FSH level of 8 IU/liter or higher is associated with decreasing fecundity, independent of female age and cycle length. Because the number of couples in whom the FSH level alters management decisions is low, we do not recommend routine testing of basal FSH in subfertile couples. (J Clin Endocrinol Metab 92: 2163–2168, 2007)
In view of this issue, we performed a prospective cohort study in a general subfertile population to assess the predictive value of basal FSH for spontaneous ongoing pregnancy in ovariary women.

**Subjects and Methods**

Between January 2002 and February 2004, we included consecutive subfertile couples in a prospective cohort study, performed in 19 hospitals in The Netherlands. The local ethics committee of each participating center gave institutional review board approval for this study. All couples underwent a basic fertility work-up, consisting of a fertility history, semen analysis, assessment of ovulation, basal FSH level testing on cycle d 3, postcoital test, and assessment of the fallopian tubes according to the guidelines of the Dutch Society of Obstetrics and Gynecology (18, 19). The menstrual cycle was considered regular if the duration of the cycle was between 23 and 35 d, with an intercycle variation of less than 8 d. Ovulation was assessed by means of a basal body temperature chart, a midluteal serum progesterone, or sonographic monitoring of the cycle. Women with an irregular menstrual cycle or documented anovulation according to the guidelines of the Dutch Society of Obstetrics and Gynecology were excluded.

Basal FSH concentrations were measured at least once on cycle d 2–4 in each participating center with commercially available immunometric assays (Table 1).

Seminal analysis was performed at least once according to the World Health Organization manual (20). The total motile sperm count (TMC) was calculated from semen volume, semen concentration, and percentage progressive motility. Men with a prewash TMC less than 3 × 10⁹ were not included in the study (21).

Tubal pathology was assessed by a chlamydia antibody test (CAT) or directly assessed by hysterosalpingography or laparoscopy. In case of a positive CAT, the tubal status was subsequently evaluated with hysterosalpingography or laparoscopy, whereas in cases of a negative CAT, tubal pathology was considered absent (22). The CAT could be tested with immune-fluorescence technique or enzyme immune assays (BioMerieux, Paris, France; Medac GmbH, Wedel, Germany; Savyon Diagnostics Ltd., France). The CAT was considered to be positive for immune-fluorescence technique if the titer was greater than 1:16 and for ELISA if the level was greater than 1.1. Women with two-sided tubal pathology were excluded from the study.

After completion of the fertility work-up, the chance of spontaneous pregnancy within 1 yr, leading to live birth, was calculated with a validated prediction model (http://www.freya.nl/probability.php) (23, 24). Couples in whom the probability of spontaneous pregnancy was 40% or greater within 12 months were counseled for expectant management for a period of at least 6 months. After 6 months of expectant management, it was up to the couples to decide to start treatment or to wait for a longer period. Couples with a probability less than 30% were counseled for treatment according to the national fertility guidelines, whereas those with a probability between 30 and 40% were asked to participate in a randomized clinical trial comparing intrauterine insemination with ovarian hyperstimulation to expectant management (25–27).

Follow-up started at completion of the basic fertility work-up and ended after 12 months. The primary end point was spontaneous conception resulting in ongoing pregnancy. The first day of that menstrual cycle was considered to mark the end of time until spontaneous pregnancy. Spontaneous ongoing pregnancy was defined as the presence of fetal cardiac activity at transvaginal sonography at a gestational age of at least 12 wk, resulting from treatment-independent conception.

Time to pregnancy was considered to be censored when treatment started or at the last day of contact during follow-up and ongoing pregnancy had not occurred. For all couples that were lost to follow-up, the general practitioner was sent a questionnaire on the last known reproductive status of the couple.

**Analysis**

Female age, menstrual cycle length, and basal FSH level are all likely to be related to ovarian aging (28). Therefore, we decided to analyze these variables together.

First, we assessed the linearity of these three variables in relation to time to spontaneous ongoing pregnancy using spline functions (29). Variables with nonlinear associations were redefined, based on these spline functions.

We then analyzed the predictive value of female age, cycle length, and basal FSH level using Cox proportional hazard modeling of the time to spontaneous ongoing pregnancy. The results were expressed as a hazard ratio (HR). Thereafter we checked whether the results of the analysis could be confirmed in a multivariable Cox proportional hazard model that included other prognostic factors.

Second, we aimed to estimate the clinical impact if FSH were to be used in a prediction model for pregnancy. For this purpose, it was considered appropriate that couples with a probability of less than 30% on a spontaneous pregnancy were counseled for treatment, whereas couples with a probability of 30% or less were counseled for expectant management. This policy is supported by the findings of a recently published randomized clinical trial, which showed that above a probability of 30% fertility treatment (intrauterine insemination with mild ovarian stimulation) did not improve pregnancy rates, compared with expectant management (27).

To estimate the added value of basal FSH, we calculated two probabilities of spontaneous pregnancy for each couple. The first probability was calculated with a validated prediction model prediction for pregnancy (model Human4, revised 2003). This model incorporates the variables female age, duration of subfertility, primary or secondary subfertility, referral status, and semen analysis. The second probability was calculated with the same model, but also including basal FSH as prognostic factor.

### Table 1. FSH assays with their analytical variations used in the 19 hospitals

<table>
<thead>
<tr>
<th>FSH assay</th>
<th>Firm</th>
<th>Analytical variation</th>
<th>No. of hospitals</th>
<th>Women, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access 2</td>
<td>Beckman Coulter Access (Fullerton, CA)</td>
<td>5.6% at 9.95 IU/liter; 5.4% at 15.45 IU/liter; 4.3% at 36.40 IU/liter</td>
<td>2</td>
<td>298 (8.5)</td>
</tr>
<tr>
<td>Advia Centaur</td>
<td>Bayer Diagnostics B.V. (Tarrytown, NY)</td>
<td>3.9% at 6.9 IU/liter; 2.4% at 23.4 IU/liter</td>
<td>4</td>
<td>1191 (34)</td>
</tr>
<tr>
<td>Architect platform</td>
<td>Abbott Laboratories (Abbott Park, IL)</td>
<td>4.5% at 4.2 IU/liter; 3.8% at 16 IU/liter</td>
<td>1</td>
<td>228 (6.5)</td>
</tr>
<tr>
<td>AutoDelfia</td>
<td>Wallac Oy (Turku, Finland)</td>
<td>3.1% at 2.56 IU/liter; 2.1% at 11.4 IU/liter; 2.6% at 44.5 IU/liter</td>
<td>2</td>
<td>367 (10)</td>
</tr>
<tr>
<td>AxSym FSH</td>
<td>Abbott Laboratories</td>
<td>5.8% at 3.0 IU/liter; 6.9% at 54 IU/liter</td>
<td>1</td>
<td>217 (6.2)</td>
</tr>
<tr>
<td>Elecsys 2010</td>
<td>Roche (Copenhagen, Denmark)</td>
<td>2.6% at 6.1 IU/liter; 3.5% at 45.0 IU/liter</td>
<td>3</td>
<td>284 (8.1)</td>
</tr>
<tr>
<td>Immulite 2000</td>
<td>Diagnostic Products Corp. (Los Angeles, CA)</td>
<td>3.1% at 5.8 IU/liter; 3.2% at 16.8 IU/liter; 2.7% at 40.9 IU/liter</td>
<td>5</td>
<td>833 (24)</td>
</tr>
<tr>
<td>Modular E170</td>
<td>Roche</td>
<td>4.4% at 3.5 IU/liter; 4.4% at 12.5 IU/liter</td>
<td>1</td>
<td>101 (2.9)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>19</strong></td>
<td><strong>3519</strong></td>
</tr>
</tbody>
</table>
factor (FSH model). We next identified the couples for which the probability of less than 30% changed to a probability of 30% or greater (or otherwise) with this new model. In other words, we assessed how many couples would change from treatment to expectant management (or otherwise).

Calculations and analyses were performed with SPSS 11.0 (SPSS Inc., Chicago, IL) and S-plus 6.0 (MathSoft Inc., Seattle, WA) programs.

Results

The study profile is shown in Fig. 1. We registered 7108 subfertile couples in 19 hospitals. Of these couples, 6072 (85%) completed their fertility work-up. In 5380 women (89%), basal FSH was measured. In 288 couples the duration of subfertility was less than 1 yr, 636 had an ovulation disorder, 685 had a severe male factor, whereas 251 had double-sided tubal pathology, leaving 3519 couples (58%) that fulfilled the inclusion criteria. Baseline characteristics are presented in Table 2. Median basal FSH level was 6.6 IU/liter (fifth to 95th percentile: 3.5–12). A basal FSH less than 8 IU/liter was found in 72% of all women, between 8 and 10 IU/liter in 17%, between 10 and 15 IU/liter in 8.9%, and 15 IU/liter or greater in 2.5% of all women.

The follow-up status of all patients at 12 months is shown in Fig. 1. We completed follow-up for 3270 couples (93%). Of all couples, 562 (16%) had a spontaneous ongoing pregnancy within 1 yr, including four (0.7%) multiple pregnancies. Unsuccessful pregnancies occurred in 61 couples (1.8%), of which 59 (1.7%) miscarried and two (0.1%) resulted in an ectopic pregnancy, i.e. 9.4 and 0.3% of all pregnancies, respectively. Within 12 months 1373 of all couples (39%) started treatment, whereas 1274 (36%) had neither started

---

**TABLE 2. Baseline characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Mean/median&lt;sup&gt;a&lt;/sup&gt;</th>
<th>5th to 95th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female age (yr)</td>
<td>32.6</td>
<td>25–40</td>
</tr>
<tr>
<td>Male age (yr)</td>
<td>35.3</td>
<td>27–45</td>
</tr>
<tr>
<td>Duration of subfertility (yr)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.5</td>
<td>1.0–4.1</td>
</tr>
<tr>
<td>Subfertility, primary (%)</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Semen analysis, TMC (10&lt;sup&gt;6&lt;/sup&gt;)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>54.0</td>
<td>5.6–304</td>
</tr>
<tr>
<td>Cycle length (d)</td>
<td>28.1</td>
<td>23–33</td>
</tr>
<tr>
<td>FSH (IU/liter)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.6</td>
<td>3.5–12</td>
</tr>
</tbody>
</table>

<sup>a</sup> For these variables, medians are provided instead of means because of a nonnormal distribution.

---

van der Steeg et al. • FSH in Subfertile Women
J Clin Endocrinol Metab, June 2007, 92(6):2163–2168

---

Fig. 1. Study profile.
treatment nor become pregnant. We did not complete follow-up for 12 months for 249 couples (7.1%).

For female age and cycle length, we found a linear association with the probability of spontaneous ongoing pregnancy. In contrast, a nonlinear association was observed for basal FSH level (Fig. 2). At less than 8 IU/liter, FSH level was not associated with the probability of spontaneous ongoing pregnancy, whereas above 8 IU/liter, the probability of a spontaneous ongoing pregnancy decreased. We therefore divided basal FSH level into two continuous variables, the first ranging from 0 to 8 and the second starting at a basal FSH level of 8 IU/liter.

Female age, cycle length, and FSH concentration, starting at a level of 8 IU/liter, were each associated with a decreased probability of spontaneous ongoing pregnancy (Table 3). In the multivariable analysis, they remained significant predictors (Table 3). At a per-unit increase of FSH over 8 IU/liter, the probability of a spontaneous pregnancy decreased with 7%. This translates to a 40% reduction of the probability to conceive in women with a basal FSH of 15 IU/liter and a reduction of 58% in women with a FSH of 20 IU/liter, in comparison with women with a basal FSH less than 8 IU/liter.

To assess possible change in treatment policy, we calculated in 3219 couples a probability for the occurrence of spontaneous pregnancy within 1 yr based on the model of Hunault (model Hunault). The probability was compared with the probability calculated with the same model including basal FSH as prognostic factor (FSH model). Probabilities are plotted in a scatterplot (Fig. 3). With the model of Hunault, 1242 couples (39%) had a probability of 30% or greater, whereas the remaining 1977 couples (61%) had a probability less than 30%. When probabilities were calculated with the FSH model (including FSH), 1255 couples (39%) had a probability of 30% or greater, whereas the remaining 1964 couples (61%) had a probability less than 30%. In a total of 97 of the 3219 couples (3.0%), the probability changed to 30% or greater to less than 30% [42 (1.3%)] or the other way around [55 (1.7%)]. As a consequence, with a probability of less than 30% as threshold for counseling treatment, 42 couples (1.3%) would be counseled for fertility treatment instead of expectant management, and 55 couples (1.7%) would be counseled for expectant management instead of fertility treatment. So in total in less than 5% of all couples, FSH altered treatment management.

Another approach was the calculation of the fraction of couples in which the probability of a spontaneous pregnancy decreased more than 5% and more 10% due to their FSH level. In 21 of 3219 couples (0.6%), the probability decreased by 10% or more with the prediction model including FSH, whereas in 73 of 3219 couples (2.3%), the probability decreased 5% or more.

Discussion

This study assessed the predictive value and clinical impact of basal FSH in subfertile ovulatory women. Our results show that increased basal FSH levels are associated with a decreased fecundity, even when corrected for cofactors female age and cycle length. Female age, cycle length, and basal FSH were all independent predictors for spontaneous pregnancy. Yet addition of basal FSH to the Hunault prediction model for spontaneous pregnancy did not lead to a large number of clinically relevant changes in pregnancy chances.

Many studies reporting on basal FSH have introduced cutoff values, i.e., 10, 15, or 20 IU/liter (13, 14). In contrast, we analyzed basal FSH as a continuous variable. In fact, pregnancy rates in relation to FSH level appear to follow a continuum, with subtle declines noted from as low as 8 IU/liter, thus not justifying a single cutoff value basal FSH.

A possible limitation of our study is that we assessed the basal FSH level only once for each woman. A review reported that intercycle variation and hourly variations may result in disparate FSH (30). Maybe a repeat of measurements would alter the accuracy that we report, although it was shown in another study focused on IVF treatment that repeated FSH measurements did not provide a substantial benefit in predicting poor ovarian response in IVF (31).

Another potential limitation is the fact that different FSH assays were used in the 19 participating clinics. Addition of the type of FSH assay to the Cox regression analysis, however, did not alter the results (data not shown).

A receiver operating characteristic curve analysis was not performed for two reasons; first, receiver operating characteristic analysis does not take time to event, e.g., time to spontaneous pregnancy, into account; second, the area under the curve is a measure of discrimination, or the ability to separate two groups, such as case-patients and controls. The sensitivity and specificity have no direct diagnostic meaning for the patient, and the issue is not the risk for having a positive test result but the risk for getting pregnant. The predicted probability, given the risk factors (basal FSH, female age, cycle length) or the posttest probability, can be

![Fig. 2. Spline function of basal FSH level. This figure expresses the association between basal FSH levels on cycle d 3 vs. the relative hazard (fecundity ratio) of a spontaneous ongoing pregnancy. Dotted lines represent 95% CI.](image-url)
TABLE 3. Results of the univariable and multivariable Cox proportional hazard analysis

<table>
<thead>
<tr>
<th>Univariable analysis</th>
<th>Multivariable analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female age (yr)</td>
<td>HR 0.96 (0.94–0.97)</td>
</tr>
<tr>
<td>Cycle length (per day shorter)</td>
<td>HR 0.95 (0.92–0.99)</td>
</tr>
<tr>
<td>Basal FSH level until 8 (per IU/liter)</td>
<td>1.0 (0.96–1.1)</td>
</tr>
<tr>
<td>Basal FSH level above 8 (per IU/liter)</td>
<td>0.93 (0.87–0.98)</td>
</tr>
</tbody>
</table>

more useful clinically in assessing future chance to conceive spontaneously than sensitivity or specificity (32).

Day 3 estradiol was advocated but not obligatory in our study protocol, but 13 of the 19 participating hospitals assessed d 3 estradiol routinely. Data on 1991 estradiol measurements of the 3519 women (57%) showed estradiol level not to be related to time to spontaneous ongoing pregnancy [HR 0.75; 95% confidence interval (CI) 0.32–1.7].

Whereas many data exist on the clinical value of basal FSH in relation to IVF, only two retrospective studies and a small prospective study have reported on basal FSH in relation to spontaneous pregnancy. The first study, performed in a case-control setting, failed to demonstrate that basal FSH had predictive capacity (13). In contrast to our study, the primary end point in that study was not specifically spontaneous pregnancy, but pregnancy also including pregnancies after fertility treatment. The same holds true for the only existing prospective study on this subject (15). A methodological issue of that prospective study was that they did not perform any regression analysis; they just compared the mean values of basal FSH in the pregnant and nonpregnant group of patients. The second retrospective study reported FSH to be associated with a lower fecundity in women with basal FSH levels above 20 IU/liter (14), but the study was underpowered to demonstrate an effect over a wider range of FSH values. Because of the large number of subfertile couples in our study, we were able to show a significant decline in fecundity with increasing basal FSH levels.

One reason for subfertility of these patients may be diminished ovarian reserve (33). Our study focused on female age, cycle length, and basal FSH as markers for ovarian aging. Apart from these variables, other tests for ovarian aging have been proposed, such as the clomiphene citrate challenge test, GnRH-agonist stimulation test, measurement of antimullerian hormone and basal inhibin B, and antral follicle count (AFC) (34–38). Of all these tests, the AFC is the only one that has been found to be superior over basal FSH in the prediction of ovarian response in IVF, although the difference is limited (11, 12, 39–41). Whether the AFC should be part of the basic fertility work-up remains subject of further research.

This study showed that the addition of basal FSH to the Hunault prediction model for spontaneous pregnancy did lead to a change in treatment policy in less than 5% of all couples.

However, the test may still be useful in a more selected group of subfertile women. Such a selection could be based on the prognostic profile of the couples. The fraction of couples in which treatment altered changed from 3.0 to 8.0% when we focused on only those couples that had probability between 25 and 35% chance to conceive, i.e. close to the cutoff probability of 30%, used to decide for treatment decisions. The largest change in management occurred in the group of women aged over 38 yr. However, because most of these women are usually advised to start treatment, FSH assessment will, generally speaking, not influence management. In women in the age category between 30 and 38 yr, management would change in 3%. Cost-effective analysis is needed to decide whether routine FSH assessment in this group should be recommended.

In summary, our results show that basal FSH levels above 8 IU/liter or greater have predictive value for spontaneous ongoing pregnancy in a general subfertile population, even after taking female age and cycle length into account. Because the number of couples in whom the FSH level alters management decisions is low, we do not recommend routine testing of basal FSH in subfertile couples.

Acknowledgments

CECERM Study Group investigators and participating centers in The Netherlands: P. F. M. van der Heijden, Almelo, Twenteborg Ziekenhuis; W. A. Schols, Amersfoort, Meander Medisch Centrum; M. H. Mocttar, Amsterdam, Academisch Medisch Centrum; H. R. Verhoeve, Amsterdam, Onze Lieve Vrouwe Gasthuis; P. G. A. Hompes, Amsterdam, Vrije Universiteit Medisch Centrum; J. M. Burggraaff, Emmen, Schepere Ziekenhuis; G. J. E. Oosterhuis, Enschede, Medisch Spectrum Twente; H.
References

1. Evers JL. 2002 Female subfertility. Lancet 360:151–159
14. van Rooij IA, de Jong E, Broekmans FJ, Looman CW, Habbema JDF, te Velde ER 2004 High follicle-stimulating hormone levels should not necessarily lead to the exclusion of subfertile patients from treatment. Fertil Steril 81:1478–1485