

Combining cycle day 7 follicle count with the basal antral follicle count improves the prediction of ovarian response

Fatih Durmusoglu, M.D., Koray Elter, M.D., Pinar Yoruk, M.D., and Mithat Erenus, M.D.

Center for ART, Department of Obstetrics and Gynecology, Marmara University School of Medicine, Istanbul, Turkey

Objective: To analyze the predictive value of cycle day 7 follicle count (CD7-FC) for poor ovarian response during IVF in women down-regulated with a luteal start of GnRH analogue (long protocol).

Design: A retrospective analysis.

Setting: University hospital.

Patient(s): Ninety-one consecutive IVF cycles of 82 subjects.

Intervention(s): Basal levels of FSH and E₂ were determined in the spontaneous cycle before the IVF cycle. During the IVF cycle, the number of basal and CD7 follicles and CD7 endometrial thickness were determined by ultrasound, and CD7 serum E₂ levels were measured.

Main Outcome Measure(s): Ovarian response determined according to the number of mature oocytes retrieved.

Result(s): On receiver operating characteristic analysis, CD7-FC had the highest combination of sensitivity and specificity to detect women with poor ovarian response when compared with the basal ovarian reserve tests. When a combined basal antral FC and CD7-FC evaluation was used with the optimum cutoff values of 6.5 and 7.5, respectively, sensitivity and specificity improved to 85% and 90%, respectively.

Conclusion(s): Cycle day 7 follicle count during a long IVF protocol is helpful in predicting ovarian response in combination with the antral FC. This combination has high positive and negative predictive values. This may help clinicians and women to cancel cycles earlier and decrease the psychological, financial, and medical burden of a later cancellation. (Fertil Steril® 2004;81:1073–8. ©2004 by American Society for Reproductive Medicine.)

Key Words: Ovarian reserve, antral follicle count, cycle day 7 follicle count, poor response, IVF

Received April 29, 2003;
revised and accepted
August 29, 2003.

Presented at the 19th
Annual Meeting of the
European Society for
Human Reproduction and
Embryology, Madrid,
Spain, June 29–July 2,
2003.

Reprint requests: Fatih
Durmusoglu, M.D.,
Marmara Universitesi
Hastanesi, Kadin
Hastaliklari ve Dogum AD,
Tophanelioglu cd. No. 13-
15, Altunizade, Istanbul,
Turkey (FAX: 90-216-326-
68-09; E-mail:
fatihdurmusoglu@
superonline.com).

0015-0282/04/\$30.00
doi:10.1016/j.fertnstert.2003.
08.044

It is known that reproductive aging is related to chronological aging and that both the quantity and quality of primordial follicles diminish as women age. However, chronological age alone is of limited value for the prediction of individual ovarian response (1–3). Therefore, various tests have been developed to assess ovarian reserve and to predict the response to ovarian stimulation (4). Even so, cancellation rates due to poor ovarian response remain high, between 5% and 30% in IVF-embryo transfer (ET) treatment with a down-regulation regimen, and the chance for success with the IVF treatment is still not predictable for some women (5, 6). The result is a cancellation or a decreased success rate due to a poor ovarian response.

Therefore, ovarian response to stimulation is the ultimate test. Earlier cancellations may help to decrease the psychological, financial, and medical burden of ovarian stimulation. To our knowledge, none of the cycle characteristics during ovarian stimulation has been compared with each other as a predictor of poor ovarian response. Cutoff values for the cycle characteristics have not been suggested either. Therefore, in this retrospective study, we analyzed the predictive value of cycle day 7 follicle count (CD7-FC) for poor ovarian response during IVF in women down-regulated with a luteal start of GnRH analogue (long protocol) and compared its predictive value with CD7 serum E₂ levels and endometrial thickness and also with basal ovarian reserve tests.

MATERIALS AND METHODS

Subjects

This study is a retrospective analysis of 91 intracytoplasmic sperm injection (ICSI) cycles of 82 patients with an indication for IVF/ICSI between January 2002 and December 2002. Women older than 42 years old or with a high (>13 mIU/mL) basal serum FSH level are not offered IVF/ICSI treatment in our center. All women underwent an ICSI cycle with a long protocol. The cycles of all patients whose spontaneous cycle before the treatment cycle had been followed and who had two ovaries were included in this study. The complete data for the parameters presented in this study were available for analysis. Institutional Review Board approval was obtained from the Marmara University School of Medicine.

Treatment Protocol

In our center, treatment cycles were performed after a spontaneous cycle, during which a workup including basal ovarian reserve tests and office hysteroscopy was done. Basal serum FSH and E₂ levels were determined in this spontaneous cycle before the treatment cycle. For all patients, pituitary desensitization was performed with leuprolide acetate SC daily (Lucrin; Abbott, Istanbul, Turkey) starting 1 week before the expected menses. Luteal administration of 1.0 mg/day was given during the luteal phase of the cycle before treatment until the second day of the onset of menses and decreased to 0.5 mg/day from day 2 until the day of hCG injection.

After down-regulation was achieved (serum E₂ <40 pg/mL), ovarian stimulation was commenced at day 3 with a daily dose of 150–300 IU recombinant FSH (rFSH) IM (Gonal-F; Serono, Istanbul, Turkey; or Puregon; Organon, Istanbul, Turkey) in combination with 300–600 IU of hMG (Humegon; Organon or Menogon, Ferring, Istanbul, Turkey; both hMG contain 75 IU FSH and 75 IU LH). On the same day, basal antral follicles, which were between 2 and 10 mm in diameter, were counted. The cycles of subjects with a basal follicle of >10 mm or evidence of an ovarian pathology were canceled. The starting dose was adjusted according to the patient's age and basal serum FSH and E₂ values at the preceding cycle and basal antral follicle count (AFC).

The subjects returned on day 7 of stimulation for an assessment of follicular recruitment and growth by transvaginal ultrasound. Follicles were measured, and those of ≥2 mm were counted and recorded. The gonadotropin dose and timing of subsequent scans were determined by the subject's response to controlled ovarian stimulation. When there were at least three follicles that were ≥16 mm in diameter, hCG was administered and transvaginal oocyte retrieval was performed 36 hours later. All subjects received 10,000 IU of hCG. Subjects who did not have at least one follicle of >10 mm after 9 days of gonadotropin stimulation and those who had one follicle of ≥16 mm but no follicles between 10 and

16 mm at any day had their cycles canceled before oocyte retrieval.

Assessment of meiotic maturity was performed after stripping the cumulus cells. All metaphase II oocytes were injected. Embryo transfer was performed 72 hours after oocyte retrieval. All patients received supplemental P in oil, 50 mg IM, each day until a pregnancy test was performed 10 days after ET.

Ovarian Response

Women were divided into two groups according to the number of metaphase II oocytes retrieved: group A (poor responders), ≤3; and group B, >3. Ovum pickup (OPU) cancellations because of impaired follicular growth in response to exogenous gonadotropins were also included in the group of poor responders.

Assays and Ultrasound Measurements

Serum FSH and E₂ concentrations were determined using the Immulite immunoassay system (Diagnostic Products Corporation, Los Angeles, CA). This assay is standardized to the World Health Organization Second International Reference Preparation 78/549. The interassay and intra-assay coefficients of variation were 6.6% and 5.4% for FSH and 5.4% and 4.4% for E₂, respectively.

Transvaginal ultrasound was performed on day 3 of the stimulation cycle by the two senior authors (F.D. and M.E.), each of whom has more than 5 years of IVF experience using a GE Logiq 200 Pro (GE Medical Systems, Milwaukee, WI) with a 6.5-MHz vaginal transducer. All ovarian follicles measuring 2–10 mm on both ovaries were counted on CD3, and those measuring ≥2 mm were counted on CD7. The total number of follicles per patient was used for calculations.

Statistical Analysis

Groups were compared using Student's *t*-test and χ^2 test where appropriate. Receiver operating characteristic (ROC) analysis was performed to determine the efficacy of different variables in detecting women with poor ovarian response. Diagnostic sensitivity and specificity were calculated, and the ROC curve was constructed by plotting the sensitivity against the false-positive rate (1-specificity) of various cutoff values for predicting poor response. Area under the ROC curve (AUC) was calculated for the significantly different parameters between groups.

The value with the optimal combination of sensitivity and specificity was chosen as the cutoff value. The ideal screening test is one that approaches or reaches the upper left corner of the graph (100% sensitivity and 100% specificity). The cutoff point for each test that has the best combination of sensitivity and specificity is located at the "knee" of the graph and is labeled for each parameter. A test that approximates a coin flip is the diagonal from the lower left to the upper right corner of the graph, which has an AUC value of

TABLE 1

Characteristics of normal and poor responders.

	Group A (poor responders)	Group B (normal responders)	P
Cycles (n)	24	67	—
Age (y)	34.54 ± 4.84	30.15 ± 5.52	.001
BMI (kg/m ²)	24.73 ± 3.03	24.18 ± 4.15	NS
Duration of infertility (y)	8.63 ± 5.62	5.82 ± 4.83	.02
Rate of previously pregnant women (%)	25.0	20.9	NS
Rate of women who had previous IVF failure (%)	33.3	31.3	NS
Basal serum FSH level (in the previous spontaneous cycle, IU/mL)	7.18 ± 3.27	6.89 ± 1.87	NS
Basal serum E ₂ level (in the previous spontaneous cycle, pg/mL)	57.11 ± 41.90	40.81 ± 24.96	NS
Basal AFC (in the treatment cycle)	4.55 ± 3.00	8.55 ± 3.22	<.001
Mean daily dose of rFSH on days 3–6 (IU)	210.2 ± 61.9	160.9 ± 96.0	.02
Mean daily dose of hMG on days 3–6 (IU)	403.1 ± 132.4	362.3 ± 169.9	NS
CD7 FC	4.58 ± 2.90	10.64 ± 3.73	<.001
CD7 E ₂ (pg/mL)	165.28 ± 145.28	448.87 ± 430.66	<.001
CD7 endometrial thickness (mm)	4.96 ± 1.31	6.71 ± 1.55	<.001
OPU cancellation rate (%)	33.3 (8/24)	0 (0/67)	<.001
Pregnancy rate (/cycle, %)	4.2 (1/24)	34.3 (23/67)	.004

Note: Values are means ± SD. AFC = antral follicle count; BMI = body mass index; CD7 = cycle day 7; FC = follicle count; OPU = ovum pickup.

Durmusoglu. Day 7 follicle count and poor response. *Fertil Steril* 2004.

0.5. Positive and negative predictive values were also calculated.

Binary combinations of the parameters, which were found to be significantly different between groups, were also analyzed for their predictive role. SPSS version 10.0 (SPSS, Inc., Chicago, IL) was used for the statistical analysis, and $P < .05$ was considered statistically significant.

RESULTS

In the group of poor responders, age, duration of infertility, and mean daily dose of rFSH used from day 3 to day 6 were significantly higher, and basal AFC, CD7-FC, CD7 serum E₂ level, and endometrial thickness were significantly lower than in the group of normal responders

(Table 1). Body mass index and basal serum FSH and E₂ levels in the previous spontaneous cycle were comparable between groups (Table 1). The number of previously pregnant women and women with previous IVF failure were also comparable between groups (Table 1). Ovum pickup cancellation and pregnancy rates were significantly different between groups (Table 1).

On ROC analysis, CD7-FC had the highest AUC value compared with the other parameters (Table 2 and Fig. 1). The optimum cutoff value of 7.5, with a sensitivity of 83% and specificity of 79%, was observed for the CD7-FC. The cutoff value for the AFC, which had the next highest AUC value, was 6.5, with a sensitivity of 85% and a specificity of 74%.

TABLE 2

Results of the receiver operating characteristic analysis for different parameters.

	AUC ± SE	P ^a	95% CI	Optimum cutoff	Sn (%)	Sp (%)	PPV (%)	NPV (%)
CD7-FC	0.91 ± 0.03	<.001	0.85–0.97	7.5	83.3	79.1	58.8	93.0
Basal AFC	0.82 ± 0.05	<.001	0.71–0.92	6.5	85.0	74.1	53.1	93.5
CD7 endometrial thickness	0.81 ± 0.05	<.001	0.72–0.90	6.75 mm	95.5	55.4	42.0	97.3
CD7 serum E ₂ level	0.79 ± 0.07	.001	0.66–0.93	158.5 pg/mL	69.2	82.0	50.0	91.1
Age	0.72 ± 0.06	.002	0.60–0.83	32.5 years	70.8	64.2	41.5	86.0
Duration of infertility	0.66 ± 0.07	.02	0.54–0.79	6.75 years	65.2	68.2	40.5	84.6

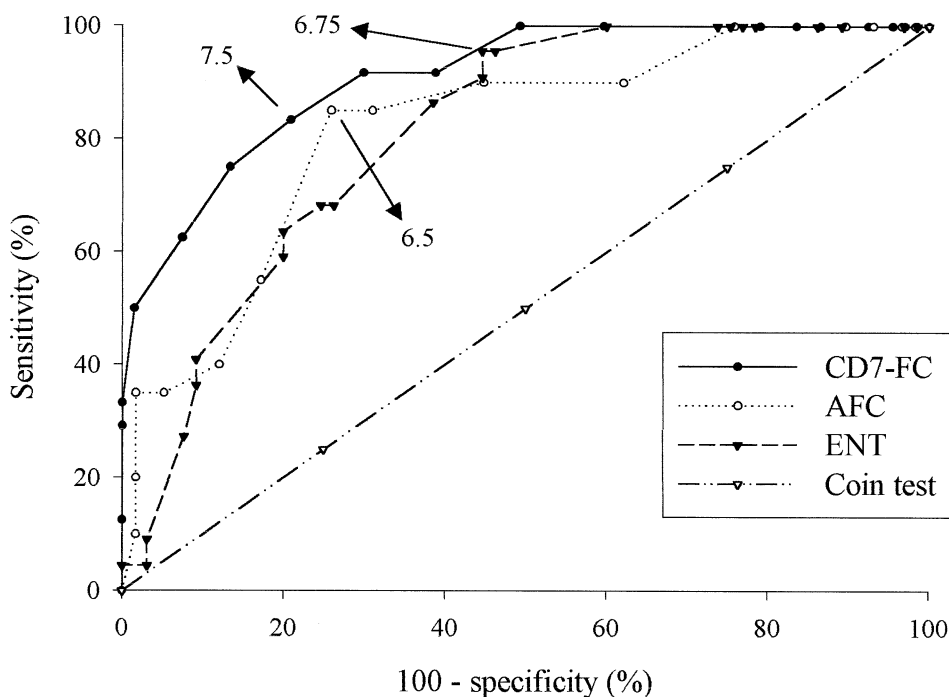
Note: AFC = antral follicle count; AUC = area under the curve; CD7 = cycle day 7; CI = confidence interval; FC = follicle count; NPV = negative predictive value; PPV = positive predictive value; SE = standard error; Sn = sensitivity; Sp = specificity.

^a Significance of the difference from a test that approximates a coin flip, which has an AUC of 0.5.

Durmusoglu. Day 7 follicle count and poor response. *Fertil Steril* 2004.

FIGURE 1

Receiver operating characteristic curves for CD7-FC, AFC, and CD7 endometrial thickness (ENT), which have significantly higher AUC values than a coin test, i.e., a test that approximates a coin flip, which is shown by a diagonal from the lower left to the upper right corner of the graph and also signifies an AUC of 0.5. Arrows indicate the optimum cutoff values, which are the closest points to the left upper corner of the graph.



Durmusoglu. Day 7 follicle count and poor response. *Fertil Steril* 2004.

The predictive roles of different combinations are shown in Table 3. When a combined AFC and CD7-FC evaluation was used with the optimum cutoff values of 6.5 and 7.5, respectively, sensitivity and specificity improved to 85% and

90%, respectively (Table 3). Positive and negative predictive values for the combined use of AFC and CD7-FC to detect women with poor ovarian response were 74% and 95%, respectively (Table 3). The likelihood ratio for a poor ovarian response was 39.1 ($P < .001$).

TABLE 3

The predictive roles of different combinations.^a

	Sn (%)	Sp (%)	PPV (%)	NPV (%)
Age and B-AFC	52.4	87.5	57.9	84.8
B-AFC and CD7-FC	85.0	89.7	73.9	94.5
Age and CD7-FC	58.3	89.6	66.7	85.7
CD7-FC and CD7-E ₂	53.3	92.3	61.5	89.6
CD7-FC and CD7-ENT	81.8	88.1	69.2	93.7

Note: B-AFC = basal antral follicle count; CD7 = cycle day 7; ENT = endometrial thickness; FC = follicle count; NPV = negative predictive value; PPV = positive predictive value; Sn = sensitivity; Sp = specificity.
^a Statistical analysis for the combinations was performed by comparing cycles in which both of the individual variables were below their corresponding cutoff values with those in which any of the variables was above its corresponding cutoff value.

Durmusoglu. Day 7 follicle count and poor response. *Fertil Steril* 2004.

DISCUSSION

In this retrospective analysis, we observed that hormonal and ultrasound parameters on CD7 were successful predictors of poor ovarian response. Among the basal ovarian reserve tests, age and AFC were also found to be effective. Serum FSH was comparable between groups. This may be due to the selection bias in the present study, since only subjects with FSH levels ≤ 13 mIU/mL were accepted for IVF treatment in our center. Among these individual parameters, ultrasound parameters, i.e., CD7-FC and AFC, had the highest combination of sensitivity and specificity.

Upon analysis of different combinations, we observed that the AFC and CD7-FC combination had the highest power to predict poor ovarian response. Women with <6.5 antral follicles on CD3 and <7.5 follicles on CD7 during an IVF cycle with a long protocol were 39 times more likely to

have a poor ovarian response than those who had higher basal or day 7 FCs. We found also that CD7-FC was a better predictor of poor ovarian response than basal AFC.

Various reports have been published recently on the usefulness of ovarian ultrasound in predicting impaired ovarian response (7–10). The basal AFC is one of these ultrasound markers, which has been shown to be indicative of poor response in assisted reproduction (7, 10). Recently, Bancsi et al. (11) have suggested that AFC, as a single predictor, had the best discriminative potential for poor response, expressed by the largest area under the ROC curve, when compared with the other basal ovarian reserve tests, i.e., total ovarian volume and basal serum levels of FSH, E₂, and inhibin B on CD3. They concluded that the AFC was the most powerful predictor of poor response (11). However, in that study, AFCs were obtained on CD3 in a spontaneous cycle.

We found that the basal AFC on day 3 of the stimulation cycle, after ovarian down-regulation, also was the single best predictor for the poor ovarian response when compared with the basal hormonal ovarian reserve tests. To our knowledge, the present study is the first one that compares the AFC after complete pituitary suppression with the hormonal ovarian reserve tests during a spontaneous cycle, although there are studies that have suggested that AFCs after complete pituitary suppression are indicative of poor ovarian response (7, 10).

Although several studies, which have been the subject of recent reviews, have attempted to develop tests for predicting ovarian response, these tests have limited value, and doubts remain about their accuracy and interpretation (4, 12–14). Therefore, the clinician is sometimes left puzzled by discordant responses, and ovarian response to stimulation remains the ultimate test. To our knowledge, none of the cycle characteristics during ovarian stimulation in a long protocol has been analyzed, and threshold values for these parameters have not been suggested as predictors for poor ovarian response.

Identification of poor responders early during stimulation is important for decreasing the side effects, cost, and risks of ovarian stimulation. During controlled ovarian stimulation, women are usually assessed for follicular recruitment and growth on CD7, and that day is usually the earliest day after a basal evaluation that ovarian response can be evaluated.

We observed that CD7-FC was an even better predictor for poor ovarian response than the basal AFC. Combined use of the basal AFC and CD7-FC had positive and negative predictive values of 74% and 95%, respectively. This means that women who have <6.5 antral follicles on CD3 and <7.5 follicles on CD7 during an IVF cycle with a long protocol have a 26% chance of getting >3 metaphase II oocytes upon retrieval. It should be emphasized that cutoff values in the present study can be used only if one uses the endpoints selected in this study. Therefore, these cutoffs apply only for

the prediction of poor ovarian response, which was defined according to the number of metaphase II oocytes retrieved (≤ 3) in the present study.

Basal AFC was a better predictor than the hormonal ovarian reserve tests, which were evaluated in this study. Unfortunately, ovarian volume has not been routinely measured in our center since it requires additional time. In the present study, CD7-FC also was a better predictor for poor ovarian response than serum E₂ level on CD7. The superiority of these ultrasound measurements over hormonal measurements is important when the ease and availability of ultrasound, which is a prerequisite for IVF, and cost, time consumption, and burden of blood samples are considered.

In a review of the literature, we found only a few reports that evaluated the association of AFC with IVF (7, 8, 10). However, only one of these articles used cutoff values for antral follicle number that were based on a true threshold value (8). In that study, Frattarelli et al. (8) suggested an optimum cutoff value of 10 for the total number of antral follicles per patient on day 3 of the stimulation cycle, after ovarian down-regulation and before beginning treatment with gonadotropins. We arrived at 6.5 for the corresponding value. This difference may be due to the difference in the statistical methods. They did not use the ROC analysis. Frattarelli et al. (8) observed the positive and negative predictive values of 12% and 97%, respectively, for their threshold value. The difference between these two values calls into question whether the optimum threshold value with the best combination of sensitivity and specificity has been determined.

All the significantly different parameters, except the gonadotropin dose, were analyzed by the ROC curve in the present study. Their independence was also evaluated by analyzing their predictive role in different combinations. Although the results of the prospective studies evaluating the use of higher doses of gonadotropins in poor responders have shown either minimal or no benefit (15), the difference in the gonadotropin dose may have biased the results in the present study. This bias is due to the retrospective nature of our study and would probably accentuate our results if it could be eliminated in a prospective design.

In conclusion, CD7-FC during a long IVF protocol is helpful in predicting ovarian response in combination with the AFC. This combination has high positive and negative predictive values. This may help clinicians and women to cancel cycles earlier and decrease the psychological, financial, and medical burden of a later cancellation. The superiority of these ultrasound parameters over hormonal parameters for both the third and seventh days of stimulation may decrease the necessity of blood samples. However, if maximum accuracy in counseling is warranted, endocrine testing, especially of FSH and E₂, should not be abandoned.

References

1. Check JH, Lurie D, Callan C, Baker A, Benfer K. Comparison of the cumulative probability of pregnancy after in vitro fertilization-embryo transfer by infertility factor and age. *Fertil Steril* 1994;61:257-61.
2. Padilla SL, Garcia JE. Effect of maternal age and number of in vitro fertilization procedures on pregnancy outcome. *Fertil Steril* 1989;52:270-3.
3. Toner JP, Philput CB, Jones GS, Muasher SJ. Basal follicle stimulating hormone level is a better predictor of in vitro fertilization performance than age. *Fertil Steril* 1991;55:784-91.
4. Sharara FI, Scott RT Jr, Seifer DB. The detection of diminished ovarian reserve in infertile women. *Am J Obstet Gynecol* 1998;179(3 Pt 1):804-12.
5. Yong PY, Brett S, Baird DT, Thong KJ. A prospective randomized clinical trial comparing 150 IU and 225 IU of recombinant follicle-stimulating hormone (Gonal-F[®]) in a fixed-dose regimen for controlled ovarian stimulation in in vitro fertilization treatment. *Fertil Steril* 2003;79:308-15.
6. Hohmann FP, Macklon NS, Fauser BC. A randomized comparison of two ovarian stimulation protocols with gonadotropin-releasing hormone (GnRH) antagonist cotreatment for in vitro fertilization commencing recombinant follicle-stimulating hormone on cycle day 2 or 5 with the standard long GnRH agonist protocol. *J Clin Endocrinol Metab* 2003;88:166-73.
7. Chang MY, Chiang CH, Hsieh TT, Soong YK, Hsu KH. Use of the antral follicle count to predict the outcome of assisted reproductive technologies. *Fertil Steril* 1998;69:505-10.
8. Frattarelli JL, Lauria-Costab DF, Miller BT, Bergh PA, Scott RT. Basal antral follicle number and mean ovarian diameter predict cycle cancellation and ovarian responsiveness in assisted reproductive technology cycles. *Fertil Steril* 2000;74:512-7.
9. Lass A, Skull J, McVeigh E, Margara R, Winston RM. Measurement of ovarian volume by transvaginal sonography before ovulation induction with human menopausal gonadotrophin for in-vitro fertilization can predict poor response. *Hum Reprod* 1997;12:294-7.
10. Tomas C, Nuojua-Huttunen S, Martikainen H. Pretreatment transvaginal ultrasound examination predicts ovarian responsiveness to gonadotrophins in in-vitro fertilization. *Hum Reprod* 1997;12:220-3.
11. Bancsi LF, Broekmans FJ, Eijkemans MJ, de Jong FH, Habbema JD, te Velde ER. Predictors of poor ovarian response in in vitro fertilization: a prospective study comparing basal markers of ovarian reserve. *Fertil Steril* 2002;77:328-36.
12. Broekmans FJ, Scheffer GJ, Bancsi LF, Dorland M, Blankenstein MA, te Velde ER. Ovarian reserve tests in infertility practice and normal fertile women. *Maturitas* 1998;30:205-14.
13. Galtier-Dereure F, De Bouard V, Picto MC, Vergnes C, Humeau C, Bringer J, et al. Ovarian reserve test with the gonadotrophin-releasing hormone agonist buserelin: correlation with in-vitro fertilization outcome. *Hum Reprod* 1996;11:1393-8.
14. Barnhart K, Osheroff J. We are overinterpreting the predictive value of serum follicle-stimulating hormone levels. *Fertil Steril* 1999;72:8-9.
15. Tarlatzis BC, Zepiridis L, Grimbizis G, Bontis J. Clinical management of low ovarian response to stimulation for IVF: a systematic review. *Hum Reprod Update* 2003;9:61-76.