Journal of Assisted Reproduction and Genetics, Vol. 21, No. 3, March 2004 (© 2004)

Assisted Reproduction

Use of Stimulated Serum Estradiol Measurements for the Prediction of Hyperresponse to Ovarian Stimulation in in Vitro Fertilization (IVF)[1]

D. J. Hendriks,[2,4] E. R. Klinkert,[2] L. F. J. M. M. Bancsi,[2] C. W. N. Looman,[3] J. D. F. Habbema,[3] E. R. te Velde,[2] and F. J. Broekmans[2]

Submitted July 18, 2003; accepted February 23, 2004

***Purpose*:** In ovarian stimulation an exaggerated ovarian response is often seen and is related to medical complications, such as ovarian hyperstimulation syndrome (OHSS), and increased patient discomfort. If it were possible to identify hyperresponders at an early stage of the stimulation phase, adaptation of the stimulation protocol would become feasible to minimize potential complications. Therefore, we studied the usefulness of measuring stimulated serum estradiol (E_2) levels in predicting ovarian hyperresponse.

***Methods*:** A total of 109 patients undergoing their first IVF treatment cycle using a long protocol with GnRH agonist was prospectively included. The E_2 level was evaluated on day 3 and 5 of the stimulation phase. Two outcome measures were defined. The first was ovarian **hyperresponse** (collection of $\geq$15 oocytes at retrieval and/or peak $E_2 > 10000$ pmol/L, or cancellation due to $\geq$30 follicles growing and/or peak $E_2 > 15000$ pmol/L, or OHSS developed). The second outcome measure comprised a subgroup representing the more severe hyperresponders, named **extreme-response** (cancellation or OHSS developed).

***Results*:** The data of 108 patients were analyzed. The predictive accuracy of E_2 measured on stimulation day 3 towards ovarian **hyperresponse** was clearly lower than that of E_2 measured on stimulation day 5 (area under the receiver operating characteristic curve (ROC_{AUC}) 0.75 and 0.81, respectively). For **extreme-response** the predictive accuracy of E_2 measured on stimulation day 3 or 5 was comparable (ROC_{AUC} 0.81 and 0.82, respectively). For both outcome measures the stimulated E_2 tests yielded only acceptable specificity with moderate sensitivity at higher cutoff levels. Prediction of extreme-response seemed slightly more effective due to a lower error rate.

***Conclusions*:** There is a significant predictive association between E_2 levels measured on stimulation day 3 and 5 and both ovarian hyperresponse and extreme-response in IVF. However, the clinical value of stimulated E_2 levels for the prediction of hyperresponse is low because of the modest sensitivity and the high false positive rate. For the prediction of extreme-response the clinical value of stimulated E_2 levels is moderate.

KEY WORDS: Estradiol; IVF; ovarian hyperresponse.

INTRODUCTION

Hyperresponder patients in IVF are characterized by exaggerated response to exogenous gonadotropin therapy manifested by recruitment of large numbers of follicles, rapid estradiol (E_2) response, and a significant cycle cancellation rate in order to prevent

[1] Presented as an oral presentation at the 18th annual meeting of the ESHRE in Vienna, June 30–July 3, 2002.

[2] Department of Reproductive Medicine, University Medical Center Utrecht, Utrecht, Heidelberglaan, Utrecht, The Netherlands.

[3] Department of Public Health, Faculty of Medicine and Health Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands.

[4] To whom correspondence should be addressed; e-mail: d.hendriks@azu.nl.

1058-0468/04/0300-0065/0

ovarian hyperstimulation syndrome (OHSS). Shortly after the first pregnancy following IVF was reported, evidence arose that overall pregnancy chances per IVF attempt increased significantly when more than one embryo was transferred. Subsequently, ovarian stimulation protocols were developed aiming at the ongoing growth of many follicles to obtain multiple oocytes for fertilization in vitro. These stimulation protocols often take more than 2 weeks, are expensive, and not without danger. Patients are at a substantially higher risk for developing OHSS and associated complications (1,2). OHSS is a serious and potentially life-threatening complication of IVF (3,4).

In several studies it was demonstrated that women who exhibit a high response to ovarian stimulation tend to have higher success rates than normal responders (5,6). In some studies however, an appreciably lower fertilization, implantation, and ongoing pregnancy rate was seen for high responders who progress to oocyte retrieval and embryo transfer. Possibly, embryonic viability, endometrial receptivity, or both are reduced in these high responder patients undergoing standard stimulation protocols (7–9).

For the reasons described in the aforementioned, it would be useful to identify women at risk for a hyperresponse and subsequently tailor the stimulation protocol to obtain a moderate ovarian response and thereby minimize patient discomfort and health risks. A rise in serum vascular endothelial growth factor (VEGF) between day of hCG administration and day of oocyte collection was shown to be a good predictor of the OHSS, especially in conjunction with the number of follicles stimulated and the level of E_2 at the day of hCG. However, the serum VEGF measured early in the stimulation phase appeared to have no predictive value toward the risk of OHSS (10). Eldar-Geva et al. (11) measured serum inhibin B levels early during FSH administration for IVF and found it of value in predicting ovarian response in IVF. Other studies confirmed that inhibin B levels measured early in the course of ovarian hyperstimulation for IVF are a useful tool in the prediction and monitoring of overresponse (12,13). One of the disadvantages, however, of inhibin B measurements is that to date they cannot be obtained on a daily basis.

Many ovarian stimulation protocols include the monitoring of serial serum E_2 levels. According to a recent report, E_2 levels obtained early after initiation of ovarian stimulation directly reflect follicular activity and ovarian responsiveness to exogenous gonadotropins (14). Use of serum E_2 levels obtained shortly after the initiation of gonadotropin stimulation could provide an early marker of ovarian responsiveness, and might in case of predicting a high ovarian response assist in deciding whether to cancel the cycle or to adapt the stimulation-dose by decreasing the amount of gonadotropins. Ultimately, this will decrease patient discomfort and the cost of continued monitoring and medication in cycles, which most likely will be cancelled due to exaggerated ovarian response. This study was designed to assess whether serum E_2 levels measured early in the course of ovarian stimulation could adequately identify patients with a high risk of an ovarian hyperresponse, even before the response is visible on ultrasound.

MATERIALS & METHODS

Patients

Patients who were prepared for IVF treatment were asked to participate in this study, independent of the cause or the duration of the infertility. There were no particular exclusion criteria. A total of 109 patients were included and written informed consent was obtained. The study was approved by the Institutional Review Board. Reasons for infertility were divided into four groups, based on the etiology of their infertility: (a) immunologic and sperm abnormalities, combined as a male infertility group (56%); (b) tuboperitoneal disease, encompassing both tubal disease and endometriosis (27%): (c) unexplained, if no cause of infertility was found (9%): (d) patients with polycystic ovarian syndrome (PCOS) (8%). In the majority of patients conventional IVF was planned ($n = 97$), whereas 12 patients were scheduled for intracytoplamic sperm injection (ICSI, $n = 12$).

Hormone Assays

During ovarian stimulation for IVF, E_2 concentrations were determined from venous blood obtained on the morning of the third, and fifth day of gonadotropin therapy. Concentrations of E_2 were assayed with a monoclonal enzyme immunoassay from Abbot Laboratories (Abbot Park, IL, USA), which was performed on the semiautomatic Imx analyser. Between run coefficients of variation were 10.1%, 7.0%, and 6.9% at 533, 1354, and 4197 pmol/L, respectively ($n = 49$, 49 and 30).

Treatment Protocol

A description of the IVF treatment protocol was published previously (15). Briefly, patients started with leuprolide acetate (Lucrin; Abbott, Hoofddorp, The Netherlands) in the midluteal phase to achieve pituitary desensitization. After the onset of menses, ovarian stimulation was started with 225 IU/d of hMG (Pergonal; Serono Benelux BV, The Hague, The Netherlands or Humegon; Organon, Oss, The Netherlands), human FSH (Metrodin HP; Serono Benelux, The Hague, The Netherlands) or a combination of both. In PCOS patients (oligoamenorrhea, elevated serum LH level, and/or elevated androgen levels) and in women under the age of 30 the starting dose was decreased. Likewise, women with elevated basal FSH or above the age of 40 received a raised starting dose. As starting dose adaptation is far from perfect in the prevention of ovarian hyperresponse, the tests under study aimed at identifying hyperresponders on the basis of information obtained during the actual stimulation cycle.

Ovarian stimulation was monitored by transvaginal ultrasonography and E_2 measurements from day 8 of gonadotropin therapy according to our standard treatment protocol. After 8 days the stimulation dose was increased if the E_2 level was under 200 pmol/L and/or no dominant follicle growth was observed. Serum E_2 levels were measured on day 3 and day 5 of gonadotropin therapy, but were not used as monitoring variable in the actual cycle. When at least three follicles attained a diameter of 18–20 mm, 10000 IU hCG (Profasi, Serono Benelux, or Pregnyl, Organon) was administered to induce oocyte maturation and 34–36 h later oocyte retrieval was performed. At a serum E_2 level >15000 pmol/L and/or growth of a large number (≥ 30) of follicles the cycle was cancelled, even before the hCG criteria were met. Coasting or dose reduction was not performed in these cycles. A maximum of two embryos was transferred in women <38 years of age. Above this age a maximum of three embryos was transferred. The luteal phase was supported either by hCG (Profasi; Serono Benelux) or micronized progesterone (Progestan; Nourypharma BV, Oss, The Netherlands).

Outcome Measures

For the statistical analysis we were interested in two outcome measures. The first one was ovarian **hyperresponse**, defined as the collection of ≥ 15 oocytes at retrieval and/or $E_2 > 10000$ pmol/L at the day of hCG, *or* cycle cancellation on the day of hCG because of a high risk of ovarian hyperstimulation syndrome (OHSS) (≥ 30 follicles and/or peak $E_2 > 15000$ pmol/L; criteria based on retrospective analysis of OHSS patients in our institute), *or* actual development of an OHSS. OHSS was classified according to the classification of OHSS of the Royal College of Obstetricians and Gynecologists (16). Only cases with a moderate or severe type were classified as OHSS. The second outcome measure comprised a subgroup representing the more severe forms of ovarian hyperresponse: cancellations for OHSS risk or actual development of OHSS. This outcome was called **extreme-response.** This approach was used as we anticipated that a shift from using all hyperresponders as endpoint toward the use of only the subgroup extreme-responders would change the predictability by the tests under study.

Methods of Analysis

Data were analyzed with the Statistical Program for Social Sciences (SPSS Inc., Chicago, IL, USA). Values are presented as median and range. To compare hyperresponders with their controls the Mann–Whitney test or the χ^2 test was performed whenever appropriate. This was also done for comparison of the extreme-responders with their controls. Logistic regression analysis was used to study the association between age of the patient and the two prognostic variables (E_2 on stimulation day 3 and E_2 on stimulation day 5) and the outcome measure ovarian hyperresponse. This procedure was also performed for the outcome measure extreme-response. The area under the receiver operating characteristic curve (ROC_{AUC}) (17), which can yield values from 0.5 (no predictive power) to 1 (perfect predictive power), was computed to assess the overall predictive accuracy of the logistic models. In order to judge upon the clinical value, logistic models were analyzed in terms of sensitivity, specificity, positive and negative predictive values, and error rates (proportion of patients who were falsely predicted of achieving a normal or hyperresponse). The probability of an event occurring as calculated from the logistic model was used as the test outcome and considered normal or abnormal based on various cutoff levels. A p value of < 0.05 was considered statistically significant.

RESULTS

Of the 109 patients entered into the study 108 became eligible for analysis (96 patients undergoing

conventional IVF and 12 patients undergoing ICSI). One patient was excluded from analysis because E_2 levels were not measured on day 3 and 5 of gonadotropin therapy. A total of 49 patients were diagnosed as ovarian hyperresponders. Of these hyperresponders 17 were classified as extreme-responders. Of this last subgroup 12 cycles were cancelled due to a high risk of developing OHSS and 5 patients underwent a follicle aspiration and embryo transfer but developed an OHSS.

Median values and 10–90 percentile ranges of the various clinical and endocrine data for the outcome measure ovarian **hyperresponse** are presented in Table I. Patients with a hyperresponse were somewhat younger compared to patients with a normal response (= controls). Indications for assisted reproductive treatment were almost identical. However, the proportion of patients with PCOS was higher in the hyperresponse group. For the hyperresponders there was a significant difference regarding amount and duration of gonadotropin stimulation, with lower dosages used during shorter time periods. E_2 concentrations measured on stimulation day 3 and 5 were significantly higher in the patients with an ovarian hyperresponse compared to normal responders. As expected, E_2 peak concentration and the number of oocytes recruited were significantly higher in patients with an ovarian hyperresponse. In Table II. median values and 10–90 percentile ranges of the patient characteristics and IVF outcome for the outcome measure **extreme-response** are presented. The control group ($N = 91$) combines normal responders and patients with ≥15 oocytes at oocyte retrieval and/or $E_2 > 10000$ pmol/L. There was no difference in age between patients with an extreme-response and those in the control group. For the extreme-responders there was a significant difference regarding the amount of gonadotropin stimulation. However, the duration of gonadotropin stimulation was not significantly different. E_2 concentrations measured on stimulation day 3, 5, and on the day of hCG administration were significantly higher in the patients with an extreme-response compared to the controls.

Table III depicts odds ratios, statistical significance, and ROC_{AUC}'s of the logistic regression analysis of three variables (age, E_2sd3 and E_2sd5) for both outcome measures. For prediction of ovarian hyperresponse, the E_2 level measured on stimulation day 5 appeared to have the best discriminative potential as a single predictor (ROC_{AUC} 0.81). For prediction of extreme-response, there was no clear difference in discriminative potential between E_2 levels measured on stimulation day 3 or 5.

To compare the performance of the E_2 levels measured on stimulation day 3 and 5 in a more clinical fashion we calculated classic test characteristics. In a logistic model we applied three arbitrary cutoff levels for the probability of ovarian hyperresponse or extreme-response of 0.33, 0.50, and 0.67. Test characteristics are given for each cutoff level for the probability of hyperresponse (Table IV) and extreme-response (Table V) together with the corresponding E_2 level. For both outcome measures (ovarian hyperresponse and extreme-response), increasing the

Table I. Patient Characteristics, IVF Outcome and E_2 Test Results in Ovarian *Hyperresponse*

Variables	Controls[a] ($N = 59$)	Hyperresponse ($N = 49$)	p value
Age (y)	34 (27–39)	31 (27–36)	0.05[b]
Primary infertility (n) (%)	41 (69.5)	36 (73.5)	NS (0.68)
Diagnosis of infertility			
Tubal pathology (%)	16 (27.1)	13 (26.5)	NS (0.46)
Male factor (%)	33 (55.9)	27 (55.1)	
Unexplained (%)	7 (11.9)	3 (6.1)	
PCOS (%)	3 (5.1)	6 (12.3)	
Number of ampules (n)	33 (18–60)	30 (19–36)	0.01[b]
Duration of stimulation (d)	11 (9–15)	10 (8–12)	<0.001[b]
E_2sd3 (pmol/L)	200 (200–401)	329 (200–648)	<0.001[b]
E_2sd5 (pmol/L)	390 (200–823)	921 (351–1755)	<0.001[b]
Peak E_2 level (pmol/L)	5470 (1719–8290)	11900 (6480–17060)	<0.001[b]
Number of oocytes (n)	8 (4–13)[c]	17 (10–28)[d]	<0.001[b]

Note. Values are presented as median and range (10–90%) or as number (%).
[a] Defined as Normal Response.
[b] Mann-Whitney test. NS = not significant on χ^2 test.
[c] $n = 56$ (3 cycles cancelled due to poor ovarian response).
[d] $n = 37$ (12 cycles cancelled due to exaggerated response).

Table II. Patient Characteristics, IVF Outcome and E_2 Test Results in the Subgroup: *Extreme-response*[a]

Variables	Controls[b] ($N = 91$)	Extreme-response ($N = 17$)	p value
Age (y)	33 (27–38)	33 (28–37)	NS (0.59)[c]
Primary infertility (n) (%)	67 (73.6)	10 (58.8)	NS (0.93)
Diagnosis of infertility			
Tubal pathology (%)	24 (26.4)	5 (29.4)	NS (0.39)
Male factor (%)	53 (58.2)	7 (41.2)	
Unexplained (%)	8 (8.8)	2 (11.8)	
PCOS (%)	6 (6.6)	3 (17.6)	
Number of ampules (n)	30 (19–53)	27 (16–39)	0.05[c]
Duration of stimulation (d)	11 (8–14)	10 (8–13)	NS (0.14)[c]
E_2sd3 (pmol/L)	202 (200–466)	437 (200–931)	<0.001[c]
E_2sd5 (pmol/L)	504 (200–1200)	1504 (307–2968)	<0.001[c]
Peak E_2 level (pmol/L)	7002 (2279–12316)	15981 (4863–20023)	<0.001[c]
Number of oocytes (n)	11 (4–20)[d]	13 (7–26)[e]	N/A

Note. Values are presented as median and range (10–90%) or as number (%). N/A = Not applicable.
[a] Defined as 1) cancellation due to exaggerated ovarian response, or 2) development of OHSS.
[b] Defined as 1) normal response, and 2) >15 oocytes at retrieval and/or $E_2 > 10000$ pmol/L.
[c] Mann-Whitney test. NS = Not significant on χ^2 test.
[d] $n = 88$ (3 cycles cancelled due to poor ovarian response).
[e] $n = 5$ (12 cycles cancelled due to exaggerated response).

cutoff for an abnormal E_2 level measured on stimulation day 3 or 5 demonstrated a shift from relatively high sensitivity and low specificity, to a low sensitivity along with a good specificity and a higher positive predictive value. The number of patients with an abnormal test result became lower at higher cutoff levels. For the prediction of the outcome ovarian hyperresponse we found a very high error rate (∼30%) (Table IV). This compared unfavourably with the prediction of extreme-response (∼14%) (Table V). For both ovarian hyperresponse and extreme-response we identified the false positive patients to see if they could become at risk for a poor ovarian response by decreasing the amount of gonadotropins. For E_2 measured on stimulation day 3 and 5 for the prediction of ovarian hyperresponse, there were 4 patients with a false positive test result at a cutoff level of the probability of hyperresponse of 0.67. The median peak E_2 level for these 4 patients was 4985 pmol/L (range 3650–8660). The median amount of oocytes collected at retrieval was 7 (range 4–13). For extreme-response, there were also 4 patients with a false positive test result at a cutoff level of the probability of extreme-response of 0.50. The median peak E_2 level for these 4 patients was 10450 pmol/L (range 4050–11800). The median amount of oocytes collected at retrieval was 11 (range 4–21).

DISCUSSION

In this study we prospectively investigated whether serum E_2 levels measured early in the course of

Table III. Univariate Logistic Regression Analysis with Odds Ratio and Areas Under the Receiver Operating Characteristic Curve (ROC_{AUC}) of Age, E_2sd3, and E_2sd5 for the prediction of ovarian *Hyperresponse* (*Upper Panel*) and for the Prediction of *Extreme-Response*[a] (*Lower Panel*)

	Odds ratio (95% CI) for Hyperresponse ($n = 49$)	p value	ROC_{AUC} (95% CI)
Age (per year)	0.92 (0.83–1.01)	0.08	0.61 (0.50–0.72)
E_2sd3 (per 100 pmol/L)	2.05 (1.42–2.97)	<0.001	0.75 (0.65–0.84)
E_2sd5 (per 100 pmol/L)	1.37 (1.20–1.56)	<0.001	0.81 (0.73–0.89)
	Odds ratio (95% CI) for Extreme-response ($n = 17$)		
Age (per year)	0.96 (0.84–1.10)	0.57	0.54 (0.40–0.69)
E_2sd3 (per 100 pmol/L)	1.99 (1.41–2.82)	<0.001	0.81 (0.69–0.93)
E_2sd5 (per 100 pmol/L)	1.29 (1.15–1.45)	<0.001	0.82 (0.69–0.94)

[a] Subgroup of Hyperresponse defined as 1) cancellation due to exaggerated ovarian response, or 2) development of OHSS.

Table IV. Predictive Performance of the Logistic Model for E_2sd3 and E_2sd5 at Several Cutoff Levels for the Probability of *Hyperresponse* and the Corresponding E_2 Levels

	E_2 (pmol/L)	Patients (%)	Sens	Spec	NPV	PPV	Error rate (*n*)	FP rate (*n*)	FN
E_2sd3									
Probability[a]:									
0.33	230	55 (51%)	0.76	0.69	0.77	0.67	28% (30)	33% (18)	12
0.50	325	33 (31%)	0.51	0.86	0.68	0.76	30% (32)	24% (8)	24
0.67	440	21 (19%)	0.35	0.93	0.63	0.81	33% (36)	19% (4)	32
E_2sd5									
Probability[a]:									
0.33	500	60 (56%)	0.78	0.62	0.77	0.63	31% (33)	37% (22)	11
0.50	730	39 (36%)	0.59	0.83	0.71	0.74	28% (30)	26% (10)	20
0.67	970	27 (25%)	0.47	0.93	0.68	0.85	28% (30)	15% (4)	26

Note. For each probability cutoff for the occurrence of ovarian Hyperresponse the corresponding E_2 level and the number (percentage) of patients with an abnormal test result is given. Sens = sensitivity, Spec = specificity, NPV = negative predictive value, PPV = positive predictive value, FP = number of false positives, FN = number of false negatives.
[a] Cutoff level of probability for occurrence of ovarian Hyperresponse.

ovarian hyperstimulation in IVF can reliably predict ovarian hyperresponse. We observed that stimulated E_2 levels only moderately predict the outcome ovarian hyperresponse and the subgroup extreme-response as evidenced by the values for the area under the ROC. However, the ROC_{AUC} as an overall measure of predictive accuracy, does not provide full insight into the clinical applicability. Therefore we also calculated classical test characteristics for E_2sd3 and E_2sd5 at various cutoff levels. The choice of a cutoff point, however, depends on the appreciation of false positive and false negative results and on the consequences drawn by the clinician from an abnormal test. If a predicted ovarian hyperresponse or extreme-response will lead to adjustment of the stimulation dose or to cancellation of the treatment, then false positive results may have unwanted consequences. In such scenarios maximum specificity is a strong prerequisite. Consequently, if the threshold for an abnormal test is set at a high level, sensitivity and the number of patients to whom the test applies will become low. On the other hand, a false negative result will leave the patient not to have the opportunity to possibly benefit from dose adjustment and consequently remain or become at risk for cancel due to high ovarian response or the occurrence of OHSS. Clearly, this would be nothing less than if no test was applied and therefore false positives may be seen as a much graver misclassification.

The classical test characteristics show us that the E_2 level measured on stimulation day 3 or 5 is from a clinical point of view not a useful test for predicting ovarian hyperresponse. As shown in Table IV, even by increasing the cutoff point for an abnormal test

Table V. Predictive Performance of the Logistic Model for E_2sd3 and E_2sd5 at Several Cutoff Levels for the Probability of *Extreme-response*[a] and the Corresponding E_2 Levels

	E_2 (pmol/L)	Patients (%)	Sens	Spec	NPV	PPV	Error rate (*n*)	FP rate (*n*)	FN
E_2sd3									
Probability[b]:									
0.33	500	14 (13%)	0.41	0.92	0.89	0.50	16% (17)	50% (7)	10
0.50	600	7 (6%)	0.29	0.98	0.88	0.71	13% (14)	29% (2)	12
0.67	700	4 (4%)	0.16	0.99	0.85	0.75	15% (17)	25% (1)	16
E_2sd5									
Probability[b]:									
0.33	1400	17 (16%)	0.59	0.92	0.92	0.59	13% (14)	41% (7)	7
0.50	1600	10 (9%)	0.47	0.98	0.90	0.80	10% (11)	20% (2)	9
0.67	1800	6 (5%)	0.26	0.99	0.87	0.83	14% (15)	17% (1)	14

Note. For each probability cutoff for the occurrence of Extreme-response the corresponding E_2 level and the number (%) of patients with an abnormal test result is given. Sens = sensitivity, Spec = specificity, NPV = negative predictive value, PPV = positive predictive value, FP = number of false positives, FN = number of false negatives.
[a] Subgroup of Hyperresponse defined as 1) cancellation due to exaggerated ovarian response, or 2) development of OHSS.
[b] Cutoff level of probability for occurrence of Extreme-response.

result there still remains a high error rate (~30%) with a high false positive rate (~20%). Looking at the false positive patients, it may be possible that by decreasing the amount of gonadotropins used, some patients could become at risk for developing a poor ovarian response. For the prediction of extreme-response classical test characteristics are slightly better. There is a lower error rate (~14%), even when the cutoff point for an abnormal test is increased. The false positive rate (~20%) does not become more favorable in comparison to the prediction of ovarian hyperresponse. This, however, could be less serious as false positive patients seem less at risk for developing poor ovarian response when decreasing the gonadotropin dose in comparison to the prediction of ovarian hyperresponse.

If an adequate test would be present to predict the patient with an exaggerated ovarian response the question arises whether it is clinically relevant to obtain early information on expected ovarian response. Considering the test characteristics shown in this study, an E_2 level measured on stimulation day 5 for instance could provide us with sufficient information to decide whether to leave the stimulation unaltered or to adjust the gonadotropin dose in trying to prevent extreme ovarian response. From the literature little is known about decreasing the gonadotropin dose early in the course of ovarian hyperstimulation during an IVF treatment. There is one study using high responders, whose endometrial receptivity is considered to be decreased significantly, in which the use of the step-down principle was investigated. The data suggest that impaired endometrial receptivity in patients with high levels of E_2 can be improved by decreasing E_2 levels, albeit in the next cycle (8). Also, the application of minimal stimulation protocols, in which gentle manipulation of the FSH window is applied, may result in a reduced degree of stimulation of the ovaries (18). However, in these minimal stimulation protocols still a large number of oocytes can be collected at retrieval and the E_2 levels at day of hCG still can be high, while the risk of poor response seems to be increased (19). This implies that prediction of outcome during the actual stimulation remains necessary, and studies on dose adaptation in that actual cycle are mandatory.

Some other indicators have been studied to help in predicting exaggerated ovarian response before the start of the IVF cycle. Tomas *et al.* examined the predictive value of early-follicular-phase antral follicle count (AFC) detected by ultrasound (20). They found that the number of baseline follicles strongly correlated with the number of recovered oocytes. This was also found in a recent study by van Rooij *et al.* (21). In another study a significant correlation was found between baseline ovarian volume and development of OHSS (22). Anti-Müllerian hormone (AMH) measured in early follicular phase has also recently been shown to be predictive for high ovarian response (21). Finally, women with polycystic ovary syndrome (PCOS) and young, regularly cycling, women with the characteristics of PCOS at ultrasound are considered to be at risk for exaggerated ovarian response (23). Pre-IVF tests as listed above may indicate situations in which the management in terms of starting dose can be adjusted accordingly. To date no clear information exists as to the value of measuring stimulated E_2 early in the running cycle in patients that are considered having a normal risk of exaggerated response on the basis of pre-IVF tests.

In conclusion, it seems not to be useful from a clinical point of view to use E_2 levels measured on stimulation day 3 or 5 in predicting ovarian hyperresponse. Predicting extreme-response could be useful, as E_2 measured on stimulation day 5 as a test is accurate enough. However, it has to be investigated whether dose adjustment during gonadotropin stimulation in the running cycle would provide a clear reduction in the occurrence of extreme ovarian response. If dose adaptation yields no clear prevention of extreme-response, the measurement of E_2 early in the course of ovarian hyperstimulation then may be considered as not useful for the management of extreme ovarian response.

REFERENCES

1. Morris RS, Paulson RJ, Sauer MV, Lobo RA: Predictive value of serum estradiol concentrations and oocyte number in severe ovarian hyperstimulation syndrome. Hum Reprod 1995;10:811–814
2. Delvigne A, Vandromme J, Barlow P, Lejeune B, Leroy F: Are there predictive criteria of complicated ovarian hyperstimulation in IVF? Hum Reprod 1991;6:959–962
3. Navot D, Bergh PA, Laufer N: Ovarian hyperstimulation syndrome in novel reproductive technologies: Prevention and treatment. Fertil Steril 1992;58:249–261
4. Brinsden PR, Wada I, Tan SL, Balen A, Jacobs HS: Diagnosis, prevention and management of ovarian hyperstimulation syndrome. Br J Obstet Gynaecol 1995;102:767–772
5. Kligman I, Rosenwaks Z: Differentiating clinical profiles: Predicting good responders, poor responders, and hyperresponders. Fertil Steril 2001;76:1185–1190
6. Pena JE, Chang PL, Chan LK, Zeitoun K, Thornton MH, Sauer MV: Supraphysiological estradiol levels do not affect oocyte and embryo quality in oocyte donation cycles. Hum Reprod 2002;17:83–87

7. Tarin JJ, Sampaio MC, Calatayud C, Castellvi RM, Bonilla-Musoles F, Pellicer A: Relativity of the concept 'high responder to gonadotrophins.' Hum Reprod 1992;7:19–22
8. Simon C, Garcia Velasco JJ, Valbuena D, Peinado JA, Moreno C, Remohi J, Pellicer A: Increasing uterine receptivity by decreasing estradiol levels during the preimplantation period in high responders with the use of a follicle-stimulating hormone step-down regimen. Fertil Steril 1998;70:234–239
9. Ertzeid G, Storeng R: The impact of ovarian stimulation on implantation and fetal development in mice. Hum Reprod 2001;16:221–225
10. Agrawal R, Tan SL, Wild S, Sladkevicius P, Engmann L, Payne N, Bekir J, Campbell S, Conway G, Jacobs H: Serum vascular endothelial growth factor concentrations in in vitro fertilization cycles predict the risk of ovarian hyperstimulation syndrome. Fertil Steril 1999;71:287–293
11. Eldar-Geva T, Margalioth EJ, Ben Chetrit A, Gal M, Robertson DM, Healy DL, Diamant YZ, Spitz IM: Serum inhibin B levels measured early during FSH administration for IVF may be of value in predicting the number of oocytes to be retrieved in normal and low responders. Hum Reprod 2002;17:2331–2337
12. Enskog A, Nilsson L, Brannstrom M: Peripheral blood concentrations of inhibin B are elevated during gonadotrophin stimulation in patients who later develop ovarian OHSS and inhibin A concentrations are elevated after OHSS onset. Hum Reprod 2000;15:532–538
13. Fawzy M, Lambert A, Harrison RF, Knight PG, Groome N, Hennelly B, Robertson WR: Day 5 inhibin B levels in a treatment cycle are predictive of IVF outcome. Hum Reprod 2002;17:1535–1543
14. Phelps JY, Levine AS, Hickman TN, Zacur HA, Wallach EE, Hinton EL: Day 4 estradiol levels predict pregnancy success in women undergoing controlled ovarian hyperstimulation for IVF. Fertil Steril 1998;69:1015–1019
15. van Kooij RJ, Looman CW, Habbema JD, Dorland M, te Velde ER: Age-dependent decrease in embryo implantation rate after in vitro fertilization. Fertil Steril 1996;66:769–775
16. Navot D, Bergh PA: Ovarian hyperstimulation syndrome: A practical approach. *In* Ovarian Endocrinopathies. The proceedings of the 8th Reinier de Graaf Symposium, Amsterdam 2–4 September 1993, pp. 215–225
17. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36
18. Fauser BC, Devroey P, Yen SS, Gosden R, Crowley WF Jr, Baird DT, Bouchard P: Minimal ovarian stimulation for IVF: Appraisal of potential benefits and drawbacks. Hum Reprod 1999;14:2681–2686
19. 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;88:166–173
20. Tomas C, Nuojua-Huttunen S, Martikainen H: Pretreatment transvaginal ultrasound examination predicts ovarian responsiveness to gonadotrophins in in-vitro fertilization. Hum Reprod 1997;12:220–223
21. Van Rooij IA, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, Jong FH, Themmen AP: Serum anti-Mullerian hormone levels: A novel measure of ovarian reserve. Hum Reprod 2002;17:3065–3071
22. Danninger B, Brunner M, Obruca A, Feichtinger W: Prediction of ovarian hyperstimulation syndrome by ultrasound volumetric assessment of baseline ovarian volume prior to stimulation. Hum Reprod 1996;11:1597–1599
23. Dor J, Shulman A, Levran D, Ben Rafael Z, Rudak E, Mashiach S: The treatment of patients with polycystic ovarian syndrome by in-vitro fertilization and embryo transfer: A comparison of results with those of patients with tubal infertility. Hum Reprod 1990;5:816–818