FERTILITY AND STERILITY®
VOL. 81, NO. 1, JANUARY 2004

Published by Elsevier Inc.
Printed on acid-free paper in U.S.A.

Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization

László F. J. M. M. Bancsi, Ph.D.,[a] *Frank J. M. Broekmans, Ph.D.,*[a] *Caspar W. N. Looman, M.Sc.,*[b] *J. Dik F. Habbema, Ph.D.,*[b] *and Egbert R. te Velde, Ph.D.*[a]

University Medical Center Utrecht, Utrecht, and Erasmus University Rotterdam, Rotterdam, The Netherlands

Objective: To study the value of a single antral follicle count and the additional value of repeated counts in different cycles for the prediction of poor ovarian response in IVF.

Design: Prospective.

Setting: Tertiary fertility center.

Patient(s): One hundred twenty women undergoing their first IVF cycle.

Intervention(s): Measurement of the number of antral follicles on cycle day 3 in two spontaneous cycles.

Main Outcome Measure(s): Ovarian response.

Result(s): A single antral follicle count is clearly predictive of poor ovarian response and there is good agreement between repeated measurements in subsequent cycles (area under the receiver operating characteristic curve [ROC_{AUC}]; cycle 1: 0.87, cycle 2: 0.85). In a logistic regression analysis, information obtained after the second cycle contributed significantly to the prediction of poor response by the antral follicle count of the first cycle. The predictive accuracy of the highest of two counts (ROC_{AUC} 0.89) was slightly better than that of each single count. The predictive model with the highest count yielded slightly higher values of specificity and positive predictive value. Sensitivity, negative predictive value, and error rates were slightly lower.

Conclusion(s): A single antral follicle count is a good predictor of poor ovarian response in IVF. Although the impact of a second antral follicle count on ovarian response predictions in IVF is statistically significant, clinical relevance is very limited. Repeating an antral follicle count in a subsequent cycle is not recommended. (Fertil Steril® 2004;81:35–41. ©2004 by American Society for Reproductive Medicine.)

Key Words: Antral follicle count, IVF, ovarian reserve, poor response, repeated measurements

Received January 30, 2003; revised and accepted June 4, 2003.

Reprint requests: László F.J.M.M. Bancsi, Ph.D., Department of Reproductive Medicine, Division of Obstetrics, Neonatology and Gynecology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands (FAX: +31 302 505 433; E-mail: lbancsi@azu.nl).

[a] Department of Reproductive Medicine, Division of Obstetrics, Neonatology and Gynecology, University Medical Center Utrecht.

[b] Department of Public Health, Faculty of Medicine and Health Sciences, Erasmus University Rotterdam.

0015-0282/04/$30.00
doi:10.1016/j.fertnstert.2003.06.011

Delayed childbearing warrants a continuous focus on the phenomenon of reproductive aging. Aging of the ovary seems to play a key role in this process, which comprises both primordial follicle reduction and oocyte quality deterioration. The remaining follicle pool and its quality is usually referred to as ovarian reserve.

In an IVF program, ovarian aging is characterized by lowered pregnancy rates and decreased ovarian responsiveness to gonadotropin administration. The so-called poor responders are likely to suffer from diminished ovarian reserve and consequently have impaired pregnancy prospects due to loss of oocyte quality. Early identification of patients who will react insufficiently to ovarian stimulation is, therefore, of great value for adequate counseling and perhaps also for the choice of the best stimulation regimen.

As a prognosticator of individual ovarian potential, chronological age is of limited value because women of the same age can be at different stages in the process of follicular depletion. This feature is also related to the wide range of age at onset of menopause, which marks "total" follicular depletion (1). Basal FSH has been shown to be a better marker of individual ovarian reserve than age (2), and is to date commonly used in many infertility centers. As a predictor of ovarian response in IVF, basal FSH generally demonstrates a fair per-

formance. However, with regard to subsequent pregnancies its predictive ability is quite limited (3, 4). An adequate test for individual reproductive potential is yet to be established.

Several other endocrine ovarian reserve tests were proposed in the past decade. Also from the field of quantitative ultrasonography achievements in the development of tests that assess ovarian reserve have been reported on. Syrop et al. (5, 6) introduced ovarian volume as a prognosticator for IVF outcome. After the finding of Ruess et al. (7), that the number of antral follicles in healthy volunteers declined with increasing age, several studies on the relation between the number of small antral follicles and response to ovarian stimulation in patients undergoing assisted reproduction have been performed (8, 9). Low numbers of small antral follicles are associated with decreased ovarian response during controlled ovarian hyperstimulation (COH) before IVF, supporting the concept of a reduced primordial follicle pool delivering a small antral follicle cohort. The study by Chang et al. (8) also showed a trend toward lower pregnancy rates in women with few antral follicles. However, this trend lacked significance, probably due to the limited size of the population under study.

Scheffer et al. (10) addressed the issue of intercycle variation in the number of antral follicles in fertile volunteers and found a rather wide 95% coverage interval around the mean difference between repeated measurements. If the same were the case in infertility patients, this might influence the predictive performance of the number of antral follicles toward ovarian response in IVF. Therefore, the aim of this study was to investigate the predictive accuracy of single and repeated antral follicle counts in the prediction of poor ovarian response and to assess the degree of agreement between antral follicle counts in subsequent cycles.

MATERIALS AND METHODS

Patients

In this study 130 patients, who were referred to our IVF clinic for their first IVF cycle in this pregnancy attempt, were included prospectively. All patients met the following criteria: [1] regular spontaneous menstrual cycle (25–35 days); [2] presence of both ovaries; [3] no evidence of endocrine disorders (normal levels of TSH, T, androstenedione (A), and PRL); and [4] written informed consent.

The institutional review board approved of the study. Patients were classified according to the main cause of infertility (tubal, male, or unexplained). In the majority of patients conventional IVF was planned (n = 112), whereas the remaining 18 patients were scheduled for intracytoplasmatic sperm injection (ICSI).

The currently presented study is part of a larger investigation on the relationship between various ovarian reserve markers and IVF outcome. Data on the predictive performance of multiple basal markers of ovarian reserve toward IVF outcome were published previously (11). Baseline data of patient characteristics and treatment variables were drawn from this study and reprinted by permission from the American Society for Reproductive Medicine.

Transvaginal Ultrasonography

On day 3 of a spontaneous cycle, and again on day 3 of the second subsequent cycle, ovarian ultrasonography was performed using a Toshiba Capasee SSA-220A (Toshiba Medical Systems Europe BV, Zoetermeer, The Netherlands) with a 7.5-MHz transvaginal probe. Round or oval echo-free structures were regarded as follicles and all follicles up to 5 mm were included in the data analysis as the antral follicle count.

One investigator (LB) performed all ultrasound measurements.

Hormone Assay

During ovarian stimulation for IVF, plasma concentrations of E_2 were assayed with a monoclonal enzyme immunoassay from Abbott Laboratories (Abbott Park, IL) on the semiautomated IMx analyzer. Between run coefficients of variation were 10.1%, 7.0%, and 6.9% at 533, 1,354, and 4,197 pmol/L, respectively (n = 49, 49, and 30).

Treatment Protocol

The IVF was conducted within 3 months from the antral follicle counts. A detailed description of the protocol was published previously (12). In short, in all patients a long protocol of down-regulation with 1 mg of leuprolide acetate daily (Lucrin; Abbott, Hoofddorp, The Netherlands) was applied from the midluteal phase onward, followed by ovarian stimulation with follitropin alpha (Gonal-F; Serono Benelux BV, 's Gravenhage, The Netherlands) in a fixed daily dose of 150 IU. Stimulation was monitored by transvaginal ultrasonography and E_2 measurements. After 7 days the gonadotropin dose was adjusted individually.

After the development of at least three leading follicles hCG (Profasi; Serono Benelux BV) was administered and a transvaginal, ultrasound-guided oocyte retrieval was performed 36 hours later. The maximum number of embryos at transfer was either two (women $<$38 years) or three (women $>$38 years). The luteal phase was supported by hCG (Profasi) if the peak E_2 level was $<$10,000 pmol/L and if $<$15 oocytes were harvested. In all other cases micronized P (Progestan; Nourypharma BV, Oss, The Netherlands) was applied.

Ovarian Response

To date, there is no uniform definition of poor response (4, 13). In a recent study, when changing the cutoff level of the number of ultrasonographically visible follicles at oocyte retrieval to define poor response from 2 to 5, the results of comparisons of patient characteristics and treatment outcome between poor and normal responders were similar (14).

At a mean fertilization rate of 50%–60% in IVF, retrieval of four oocytes would be needed to reach an average of two embryos available for transfer. Thus, in this study we defined poor response as: [1] collection of fewer than four oocytes at retrieval or [2] cycle cancellation due to impaired follicular reaction (<3 follicles) in response to exogenous gonadotropins.

High response was defined as the collection of >20 oocytes at retrieval. Patients whose cycles were canceled because they were considered at risk of ovarian hyperstimulation syndrome (OHSS) due to exaggerated follicle growth (>30 follicles or peak E_2 >15,000 pmol/L), were also defined as high responders. In all other cases ovarian response was classified as normal. For purpose of analysis normal and high responders were considered as one group.

Pregnancy

Clinical and ongoing pregnancies were defined as the presence of fetal cardiac activity beyond 6 and 12 weeks of gestation, respectively. Multiple pregnancies were dealt with in the same manner as singleton pregnancies.

Eligibility for Analysis

Data of all patients who underwent ovarian stimulation were analyzed. Patients whose cycles were canceled due to either risk of OHSS or poor response to hormone stimulation (<3 follicles) were included in ovarian response analyses but not in pregnancy rate calculations, because it cannot be excluded that such patients would have become pregnant if IVF were performed. On the other hand, patients with complete absence of follicle growth and E_2 level below the detection limit of the assay (<200 pmol/L) were considered to have a zero chance of pregnancy. These patients were included into pregnancy rate calculations and counted as nonpregnant.

Methods of Analysis

Data were analyzed with the SPSS (SPSS Inc., Chicago, IL) and General Linear Interactive Modelling packages (GLIM, NAG, Oxford, United Kingdom) (15). A two-sided *P* value <.05 was considered to indicate statistical significance. We used the Mann Whitney *U* test, the χ^2 test, and Fisher's exact test for comparison of clinical data and IVF treatment variables between poor responders and normal responders. A Wilcoxon signed rank test and Pearson's correlates were used to compare antral follicle counts from the first and second cycle.

Reproducibility of the antral follicle count was analyzed by calculating the 95% coverage interval of the difference between two counts (16, 17). The 95% coverage interval is defined as 1.96*SD above or below the mean difference of two measurements. Reproducibility was also studied assuming individual-specific average levels and Poisson distributed variability between individual-specific measurements. The coverage interval for the difference between two Poisson distributed counts is accordingly estimated as 1.96 times the square root of two times the Poisson mean.

Logistic regression was applied to study the value of antral follicle counts, age, and infertility diagnosis for the prediction of poor ovarian response. We also analyzed several possible prognostic variables derived from the combination of the two antral follicle counts: mean, highest, and lowest antral follicle count.

Multiple logistic regression analysis with forward selection of parameters was applied, with *P*<.10 for entry. To assess the predictive power of logistic models, areas under the receiver operating characteristic curve (ROC_{AUC}) were computed. Values can range from 0.5 (no predictive power) to 1 (perfect prediction). The ROC_{AUC} reflects the proportion of all discordant pairs of patients, consisting of one poor responder and one normal responder, in which the model predicts a higher probability of poor response for the poor responder patient than for the normal responder. The predictive performances of the logistic models were further analyzed in terms of sensitivity, specificity, positive and negative predictive values, and error rates (total proportion of patients in whom the occurrence of poor or normal ovarian response was predicted falsely).

TABLE 1

Comparison of basic clinical data of normal responders (NR) and poor responders (PR).

Variable	NR (n = 84)	PR (n = 36)	*P* value
Mean (± SD) age (y)	34.4 ± 4.7	36.2 ± 5.4	.06
Mean (± SD) duration of infertility (mo)	33 ± 16	39 ± 39	.94
Mean (± SD) body mass index (kg/m^2)	22.9 ± 3.7	24.4 ± 4.2	.10
No. (%) of patients with indicated type of infertility			
Primary	48 (57)	18 (50)	.47
Secondary	36 (43)	18 (50)	
No. (%) of patients with indicated infertility diagnosis			
Tubal	17 (20)	6 (17)	<.001
Male	50 (60)	9 (25)	<.001
Unexplained	17 (20)	21 (58)	
Mean (± SD) cycle 1 antral follicle count (n)	11.6 ± 6.2	4.2 ± 3.3	<.001
Mean (± SD) cycle 2 antral follicle count (n)	11.4 ± 6.3	4.6 ± 3.3	<.001
Mean (± SD) mean antral follicle count (n)	11.5 ± 5.9	4.4 ± 3.2	<.001
Mean (± SD) highest antral follicle count (n)	13.2 ± 6.3	5.1 ± 3.4	<.001
Mean (± SD) lowest antral follicle count (n)	9.9 ± 5.7	3.7 ± 3.1	<.001

Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

TABLE 2

IVF treatment results in normal responders (NR) and poor responders (PR).

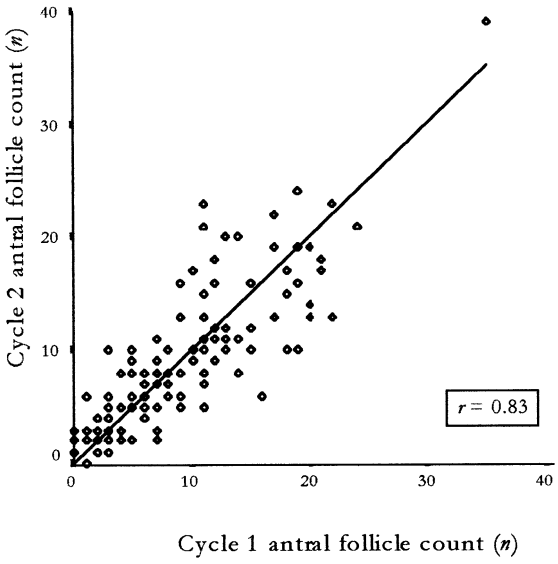
Variable	NR (n = 84)	PR (n = 36)	*P* value
Mean (± SD) no. of ampules	25.4 ± 8.9	29.7 ± 10.8	.022
Mean (± SD) duration of stimulation (d)	12.9 ± 2.9	13.6 ± 3.2	.18
Rate of gonadotropin dose increase (%)	18 (15)	42 (15)	.006
Mean (± SD) peak E_2 level (pmol/L)	7594 ± 3779	1696 ± 1661	<.001
Mean (± SD) no. of oocytes[a]	9.5 ± 4.8	2.1 ± 0.8	<.001
Mean (± SD) fertilization rate (%)	55 ± 32	75 ± 33	.006
Mean (± SD) no. of embryos per ET[b]	2.1 ± 0.5	1.7 ± 0.6	.003
Implantation rate per embryo (%)	30 (41/137)	10 (3/30)	.024
Clinical pregnancy rate (%)[c]	40 (31/77)	10 (3/30)	.003
Ongoing pregnancy rate (%)[c]	31 (24/77)	10 (3/30)	.024

[a] Data for oocyte retrieval cycles (overall n = 97, normal responders n = 77, and poor responders n = 20).
[b] Data for embryo transfer cycles (overall n = 83, normal responders n = 65, and poor responders n = 18).
[c] Data for oocyte retrieval cycles or cycle cancellation due to complete absence of follicular response (overall n = 107, normal responders n = 77, and poor responders n = 30).

Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

FIGURE 1

Plot of the first antral follicle count against the second antral follicle count. The *solid line* represents equality of the two counts.



Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

RESULTS

Of the 130 included patients, 120 became eligible for analysis (n = 102 for IVF and n = 18 for ICSI). Six patients conceived spontaneously while on the waiting list for IVF, two patients dropped out due to intercurrent disease, and two patients withdrew their consent.

A total of 36 patients were diagnosed as poor responders. In 20 patients follicular aspiration yielded less than four oocytes and 16 cycles were canceled due to either complete absence of ovarian response (n = 10; included in pregnancy calculations) or to development of only one or two follicles during hormone stimulation (n = 6; not included in pregnancy calculations). Ten patients were diagnosed as high responders, after either cycle cancellation because of risk of OHSS (n = 7; not included in pregnancy calculations) or retrieval of more than 20 oocytes (n = 3; included in pregnancy calculations). Thus, of the 120 patients with data for ovarian response analysis, 107 were included into pregnancy calculations.

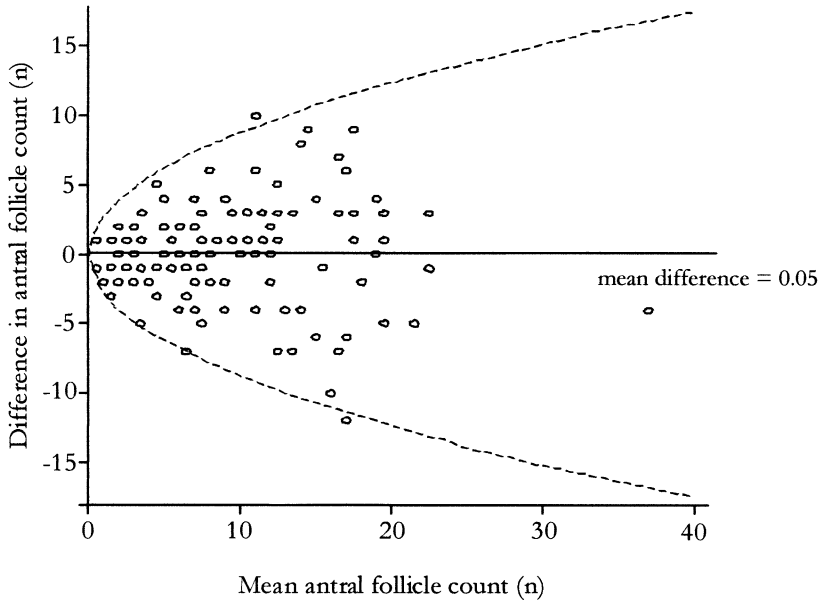
Table 1 presents basic clinical data for poor responders and normal responders. Antral follicle counts from cycle 1 and cycle 2 were similar (P=.81). All antral follicle variables were significantly lower in poor responders. The IVF treatment results are shown in Table 2. Poor responders showed significantly lower pregnancy rates. The reduced implantation rate per embryo is more striking, considering the fact that embryo quality was similar in both groups (P=.28).

There was a good correlation between antral follicle counts of cycle 1 and cycle 2 (r = 0.83). At higher values, the points deviated more from the line of equality (Fig. 1). Reproducibility was assessed by plotting mean antral follicle count against difference in antral follicle counts, as shown in Figure 2. The mean difference between two measurements was 0.05. The overall 95% coverage limits of the mean difference were −7.26 and 7.36. However, Figure 2 shows increasing variability around the mean difference when the mean count increases. This pattern of variability is consistent with the assumption that the two measurements are independent Poisson counts (P=.14). The corresponding coverage interval is indicated in Figure 2.

Univariate logistic regression analysis revealed a good predictive power of all antral follicle count variables for the prediction of poor response (Table 3). Age and infertility diagnosis appeared to be far less potent predictors of poor response. When applying stepwise forward selection on all variables presented in Table 3, the highest antral follicle

FIGURE 2

Plot of mean antral follicle count against difference in antral follicle counts (cycle 1 antral follicle count − cycle 2 antral follicle count). The *solid line* represents the mean difference. The *dotted line* represents the upper and lower limit of the 95% coverage interval of the mean difference.



Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

count was selected in the first step and no other variable contributed significantly thereafter (ROC_{AUC} = 0.89, $P<.001$). To assess whether a second antral follicle count added significantly to a first count, we performed forced entry of cycle 1 or cycle 2 antral follicle count before applying stepwise forward selection. Again, the highest antral follicle count was selected (ROC_{AUC} = 0.89, $P=.03$ and ROC_{AUC} = 0.89, $P=.005$, respectively), indicating that the highest of two counts significantly contributes to the prediction of poor response by each single count. Forced entry of age and infertility diagnosis did not improve predictive power.

The clinical performance of the best single predictor after one measurement (cycle 1 antral follicle count) and the strongest predictor based on two measurements (highest antral follicle count) was compared on classic test properties. To elucidate the influence of changing the cutoff point, we calculated the test characteristics for several cutoff values of the number of antral follicles (Table 4). At lower cutoff points sensitivity and negative predictive value decrease and specificity and positive predictive value increase. The highest antral follicle count displayed better specificity and positive predictive value, whereas at the same cutoff levels sensitivity was lower than for cycle 1 antral follicle count. The error rate is about 2% lower for the highest antral follicle count.

DISCUSSION

With regard to the relationship between the number of small antral follicles and ovarian reserve, Gougeon (18) demonstrated a clear association between the number of growing follicles and the number of resting follicles in the follicle pool. Moreover, Pellicer et al. (19) found a high correlation when comparing numbers of selectable follicles (2–5 mm) measured by three-dimensional ultrasonography with the numbers of selectable follicles in histological slices. Therefore, it is plausible that the number of antral follicles, originating from the cohort of growing follicles, reflects the size of the pool of resting follicles, and thus ovarian reserve.

In this study, a single investigator performed all antral follicle counts. There is evidence that counts of antral follicles are subject to limited interobserver variation only (20). In addition, we studied interobserver variation for both real-time two-dimensional transvaginal ultrasonography and stored three-dimensional transvaginal ultrasound images and between method variation comparing two- and three-dimensional transvaginal ultrasonography (21). The data indicated adequate two- and three-dimensional interobserver reproducibility as well as acceptable between method variation. Therefore, ultrasound-based antral follicle counts seem to be a reliable tool for the assessment of ovarian reserve and

TABLE 3

Logistic regression analysis for the prediction of poor response, expressed as odds ratios and area under the receiver operating characteristic curve (ROC_{AUC}) for age, infertility diagnosis, and various antral follicle count variables.

Variable	Odds ratio (95% CI)	*P* value	ROC_{AUC}
Univariate			
Age (per year)	1.08 (0.99–1.17)	.07	0.61
Infertility diagnosis	na	<.001	0.71
Male	1.00	na	na
Tubal	1.96 (0.61–6.32)	.26	na
Unexplained	6.86 (2.64–17.83)	<.001	na
Cycle 1 antral follicle count (n)	0.70 (0.60–0.81)	<.001	0.87
Cycle 2 antral follicle count (n)	0.71 (0.61–0.82)	<.001	0.85
Mean antral follicle count (n)	0.67 (0.58–0.81)	<.001	0.87
Highest antral follicle count (n)	0.68 (0.58–0.79)	<.001	0.89
Lowest antral follicle count (n)	0.70 (0.60–0.81)	<.001	0.87
Multivariate			
Highest antral follicle count (n)	0.68 (0.58–0.79)	<.001	0.89

Note: CI = confidence intervals; na = not applicable.

Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

studies on the prediction of poor ovarian response to exogenous gonadotropins.

Our data confirm that a single antral follicle count is a good indicator of the occurrence of poor or normal response in IVF (8, 22). If poor response is adequately predictable, then an antral follicle count before IVF could be useful for patient counseling and perhaps treatment management. Moreover, there is a significant relation between the nature of ovarian response and pregnancy rates (Table 2). Nevertheless, antral follicle counts cannot be applied to predict pregnancy (cycle 1 antral follicle count: $P=.53$; cycle 2 antral follicle count: $P=.88$).

The present study demonstrated that the number of antral follicles on cycle day 3 varies from cycle to cycle according to expected fluctuations, especially with higher number of antral follicles. This does not exclude the possibility of biologic variation in cohort sizes between cycles. The predictive accuracy with regard to poor ovarian response in IVF demonstrated little intercycle variation. Starting off with either the first or the second antral follicle count, however, it appeared that the highest count from two cycles gave a statistically significant improvement of the prediction of poor response. Does this finding imply clinical relevance of counting antral follicles in two different cycles instead of only in one cycle?

Before answering this question, one should consider how the predictive information might influence the clinical decision on whether or not to treat, or, if treatment is to be applied, on whether or not to differentiate the treatment protocol. Poor response may result in cycle cancellation during the course of COH. If oocyte pick-up is performed, poor response is clearly related to low pregnancy rates (Table 2). Therefore, the knowledge that a patient will probably have a poor response may lead to counsel this patient on her meager prospects. In addition, the treatment protocol may be altered by applying higher doses of gonadotropins in false negatives.

On the other hand, if a normal response is predicted, the decision on whether or not to perform IVF is not influenced. Thus, the prediction of poor response may have profound consequences, whereas the prediction of normal response does not. Therefore, a false-positive prediction (normal response in case of a predicted poor response) followed by negative counseling, denial of treatment, or increased daily dosages of gonadotropins and consequently a higher risk of OHSS, might be considered more serious than a false-negative prediction (poor response in case of a predicted normal response). Consequently, in our clinical setting high specificity and a high positive predictive value are more important than high sensitivity and a high negative predictive value. This requirement is only satisfied at low antral follicle counts (4 follicles or less). The data in Table 4 demonstrate that specificity and positive predictive value of the highest follicle count are just slightly higher than that of cycle 1 antral follicle count at low antral follicle counts. This might suggest that the highest antral follicle count is more powerful, at the expense, however, of substantially decreased sensitivity. Moreover, to obtain this subtle advantage the whole population needs to be tested twice. We conclude, therefore, that performing antral follicle counts in subsequent cycles to predict poor response in IVF is not useful.

The analysis thus far is based on the scenario that either one or two antral follicle counts are performed in all patients in subsequent cycles. A clinically useful alternative could be only to perform a second antral follicle count in patients with an intermediate cycle 1 antral follicle count, who still may be at risk of poor response. We investigated this sequential diagnostic scenario in our data and it appeared that in case of only patients with less than seven antral follicles in the first cycle would have been subjected to a cycle 2 antral follicle count (40% of all patients), the same overall predictive performance would have been obtained. Note that these results are based on a limited number of patients and should be tested prospectively.

In conclusion, in this study it appeared that a single antral follicle count is an adequate predictor of poor ovarian response in IVF. The impact of a second antral follicle count in a subsequent cycle on ovarian response predictions is statistically significant. Clinical relevance, however, in terms of changes in specificity and sensitivity is very limited. Thus, the application of an antral follicle count in different cycles before IVF is not advocated.

TABLE 4

Performance of two antral follicle count variables in the prediction of poor response in IVF.

Follicles (n)	Patients[a]	Sens	NPV	Spec	PPV	Error rate	FP	FN
				Cycle 1 antral follicle count				
≤10	70 (58)	0.94	0.96	0.57	0.49	32%	36	2
≤8	61 (51)	0.92	0.95	0.67	0.54	26%	28	3
≤6	48 (40)	0.81	0.90	0.77	0.60	22%	19	7
≤5	41 (34)	0.72	0.88	0.82	0.63	20%	15	9
≤4	32 (27)	0.61	0.84	0.88	0.69	20%	10	14
≤3	23 (19)	0.47	0.80	0.93	0.74	21%	6	19
≤2	14 (12)	0.33	0.77	0.98	0.86	22%	2	24
				Highest antral follicle count				
≤10	66 (55)	0.94	0.96	0.62	0.52	30%	32	2
≤8	54 (45)	0.89	0.94	0.74	0.59	22%	22	4
≤6	39 (33)	0.75	0.89	0.86	0.69	18%	12	9
≤5	30 (25)	0.61	0.84	0.90	0.73	18%	8	14
≤4	23 (19)	0.53	0.82	0.95	0.83	18%	4	17
≤3	17 (14)	0.39	0.79	0.96	0.82	21%	3	22
≤2	7 (6)	0.17	0.73	0.99	0.86	26%	1	30

Note: For each cutoff number of antral follicles the number of patients with an antral follicle count equal to or below the cutoff value is given. Sens = sensitivity; NPV = negative predictive value; Spec = specificity; PPV = positive predictive value; FP = number of false positives; FN = number of false negatives.

[a] Values in parentheses are percentages.

Bancsi. Repeated antral follicle counts. Fertil Steril 2004.

References

1. Treloar AE. Menstrual cyclicity and the pre-menopause. Maturitas 1981;3:249–64.
2. 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.
3. Bancsi LFJMM, Huijs AM, den Ouden CT, Broekmans FJM, Looman CWN, Blankenstein MA, et al. Basal follicle-stimulating hormone levels are of limited value in predicting ongoing pregnancy rates after in vitro fertilization. Fertil Steril 2000;73:552–7.
4. Bancsi LFJMM, Broekmans FJM, Mol BWJ, Habbema JDF, te Velde ER. Performance of basal follicle-stimulating hormone in the prediction of poor ovarian response and failure to become pregnant after in vitro fertilization: a meta-analysis. Fertil Steril 2003;79:1091–100.
5. Syrop CH, Dawson JD, Husman KJ, Sparks AE, Van Voorhis BJ. Ovarian volume may predict assisted reproductive outcomes better than follicle stimulating hormone concentration on day 3 [In Process Citation]. Hum Reprod 1999;14:1752–6.
6. Syrop CH, Willhoite A, Van Voorhis BJ. Ovarian volume: a novel outcome predictor for assisted reproduction. Fertil Steril 1995;64:1167–71.
7. Ruess ML, Kline J, Santos R, Levin B, Timor-Tritsch I. Age and the ovarian follicle pool assessed with transvaginal ultrasonography. Am J Obstet Gynecol 1996;174:624–7.
8. 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.
9. 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.
10. Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, te Velde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril 1999;72:845–51.
11. Bancsi LFJMM, Broekmans FJM, 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. 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;66:769–75.
13. Lashen H, Ledger W, Lopez-Bernal A, Barlow D. Poor responders to ovulation induction: is proceeding to in-vitro fertilization worthwhile? Hum Reprod 1999;14:964–9.
14. van Rooij IAJ, Bancsi LFJMM, Broekmans FJM, Looman CWN, Habbema JDF, te Velde ER. Women older than 40 years of age and those with elevated follicle-stimulating hormone levels differ in poor response rate and embryo quality in in vitro fertilization. Fertil Steril 2003;79:482–8.
15. Aitken MD, Anderson SO. AnonymousStatistical modelling in GLIM. Oxford: Clarendon Press, 1989.
16. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995; 346:1085–7.
17. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
18. Gougeon A. Caracteres qualitatifs et quantitatifs de la population folliculaire dans l'ovaire humain adulte. Contracept Fertil Sex 1984;12: 527–35.
19. Pellicer A, Ardiles G, Neuspiller F, Remohi J, Simon C, Bonilla-Musoles F. Evaluation of the ovarian reserve in young low responders with normal basal levels of follicle-stimulating hormone using three-dimensional ultrasonography. Fertil Steril 1998;70:671–5.
20. Pache TD, Wladimiroff JW, de Jong FH, Hop WC, Fauser BC. Growth patterns of nondominant ovarian follicles during the normal menstrual cycle. Fertil Steril 1990;54:638–42.
21. Scheffer GJ, Broekmans FJM, Bancsi LFJMM, Habbema JDF, Looman CWN, te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: reproducibility of antral follicle counts. Ultrasound Obstet Gynecol 2002;20:270–5.
22. Frattarelli JL, Lauria-Costa DF, Miller BT, Bergh PA, Scott RT Jr. 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.