Antral Follicle Assessment as a Tool for Predicting Outcome in IVF—Is it a Better Predictor than Age and FSH?

RAVIT NAHUM,¹ JAN L. SHIFREN,¹ YUCHIAO CHANG,¹ LUCY LEYKIN,¹ KEITH ISAACSON,¹ and THOMAS L. TOTH,^{1,2}

Submitted: July 28, 2000

Accepted: October 13, 2000

Purpose: The purpose of this study is to determine if baseline antral follicle assessment may serve as additional information in predicting in vitro fertilization outcome.

Methods: Prospective, descriptive preliminary study of in vitro fertilization outcome. From July 1998 to July 1999, 224 patients underwent antral follicle assessment (follicle 2–6 mm in diameter) on baseline of the planned, stimulated in vitro fertilization cycle. The outcomes were analyzed with respect to antral follicle assessment (≤ 6 or > 6), basal cycle day 3 follicle stimulated hormone (≤ 10 or > 10 IU/L) and maternal age (≤ 35 or > 35 years).

Results: The clinical pregnancy rate was significantly higher in the group with baseline antral follicle > 6 compared to that in the group with antral follicle ≤ 6 (51% vs. 19%, respectively). Controlling for patient age, and basal follicle stimulated hormone, the pregnancy rate was significantly higher in the group with antral follicle > 6 compared to that in the group with antral follicle ≤ 6 . The cancellation rate was significantly increased with advancing maternal age, elevated basal follicle stimulated hormone levels, and baseline antral follicle ≤ 6 . The cancellation rate was significantly higher in the group with antral follicle ≤ 6 compared to that in the group with antral follicle ≥ 6 (33% vs. 1%, respectively).

Conclusions: In vitro fertilization outcome is strongly correlated with both maternal ages, basal cycle, day 3 follicle, stimulated hormone, and antral follicle assessment. Antral follicle assessment was a better predictor of in vitro fertilization outcome than were age or follicle stimulated hormone. Antral follicle assessment may provide a marker for ovar-

ian age that is distinct from chronological age or hormonal markers.

KEY WORDS: Antral follicle; IVF outcome; pregnancy rate; cancellation rate.

INTRODUCTION

The outcome of in vitro fertilization (IVF) treatment is highly dependent on ovarian response to hormonal stimulation. Poor ovarian response is associated with a higher cancellation rate, reduced number of retrieved oocytes, and lower pregnancy rate (1). A variety of screening tests have been developed to provide a reliable assessment of the ovarian reserve and to predict response to gonadotropin stimulation. Factors that may help to determine ovarian response are valuable prognosticators of IVF outcome.

Previous studies in infertile women have shown that advanced maternal age and elevated follicle stimulating hormone (FSH) level on cycle Day 3 are associated with reduced ovarian response to gonadotropin stimulation and decreased IVF outcome (2,3). Maternal age alone is a weak predictor for evaluation of IVF outcome. Toner *et al.* (4) found that pregnancy rates decreased markedly as age increased, but basal FSH levels provided better predictive values than did age alone for pregnancy rate. Recent data show that pregnancy rates with basal FSH ≤ 10 IU/L were statistically significantly higher than those with FSH > 10 IU/L (5).

Aging is associated with progressive follicular depletion and diminished oocyte quality, which is accompanied by a reduction in the size or activity of the cohort of follicles available to respond to gonadotropin stimulation (6). The rate of oocyte depletion with age

¹ The Vincent IVF Unit, Department of Obstetrics and Gynecology, and Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.

² To whom correspondence should be addressed at Vincent IVF unit, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston, Massachusetts 02114. e-mail: toth.thomas@harvard.edu.

is highly variable, and this variability manifests itself in IVF patients who respond unexpectedly poorly to ovarian stimulation despite favorable age, whereas an older woman may respond surprisingly well. Several studies evaluated additional prognosticators of IVF outcome. Andolf *et al.* (7) found a nearly linear decrease in mean ovarian volume with age. Other studies demonstrated a strong correlation between follicular density, ovarian volume, and follicular response in IVF (8,9). Sharara *et al.* (10) concluded that decreased ovarian volume may be a better predictor of reduced ovarian reserve than is basal FSH. Recently, Reuss *et al.* (11) were able to detect an age-related decrease in the number of antral follicles in healthy volunteers. Tomas *et al.* (12) showed that in IVF patients, the number of antral follicles after ovarian suppression correlated with the number of recovered oocytes and was more accurate than ovarian volume as a predictor for follicular response. Others (13) reported an association between AF count, patient age, basal FSH, basal E2, and the number of oocytes retrieved, but no correlation with pregnancy rate was seen.

There is little information available on the relative predictive value of AF assessment on important IVF outcomes as pregnancy and cancellation rates. Therefore, the purpose of this study is to determine the predictive value of AF assessment on IVF outcome as an adjunct to basal FSH and maternal age.

MATERIALS AND METHODS

Patients and Study Protocol

This study is a descriptive, prospective analysis of AF assessment as determinants of IVF outcome in 272 consecutive IVF cycles of 224 patients cared for at the Massachusetts General Hospital, Vincent Obstetrics and Gynecology Service, IVF Unit from July 1998 to July 1999. The study was approved by the Institutional Review Board (No. 99-7347).

All women who were treated at the Vincent IVF Unit during the study period had basal FSH and E2 levels on cycle Day 3 of a spontaneous cycle determined 2–3 months prior to participation in IVF ovarian stimulation protocols. The IVF protocol for induction of multiple follicular development utilized pituitary down regulation with daily SC administration of GnRH analogue (Leuprolide acetate, LA; Lupron, TAP Pharmaceuticals, Deerfield, IL) commencing in the midluteal phase of the preceding cycle

and continued until the day of hCG or in the follicular phase of the stimulated cycle. Controlled ovarian hyperstimulation with gonadotropins (rFSH; Gonal-F; Serono Laboratories, Randolph, MA, or Follistim, Organon, West Orange, NJ) was initiated following confirmation of adequate pituitary suppression in an individualized step-down protocol. Follicular development was monitored by serial estradiol (E2) measurements and transvaginal ultrasound examination. Ovulation induction was initiated with 10,000 IU human chorionic gonadotropin (hCG) (Profasi; Serono Laboratories, Randolph, MA) when at least three ovarian follicles reached ≥ 18 mm in diameter as determined by transvaginal ultrasound. Follicular aspiration was carried out under transvaginal ultrasound guidance 34–36 h after the hCG administration. Transcervical intrauterine embryo transfer was performed 3–5 days after retrieval, followed by luteal phase support with progesterone 50 mg/day IM from 2 days after retrieval until determination of clinical pregnancy. A clinical pregnancy was defined by the presence of a gestational sac on ultrasound.

All patients underwent AF assessment using transvaginal ultrasound (General Electric, RT-X200, Milwaukee, USA), equipped with a 6.5 MHz vaginal probe. The number of antral follicles (2–6 mm in diameter) was measured on cycle Day 3 of the planned IVF cycle (not on spontaneous cycle) after ovarian suppression with GnRH analogue and prior to administration of gonadotropins. In 46 patients, a flare protocol was applied for ovarian stimulation, and AF assessment was measured on cycle Day 2 prior to administration of GnRH analogue and gonadotropins. Evidence suggests that the number of antral follicles is independent of hormonal stimulation for its growth (14). Patients were divided into two groups based on AF assessment of ≤ 6 or > 6 based on the study of Tomas *et al.* (12) that defined inactive ovaries as AF assessment ≤ 5.4 . Cancellation criteria included less than three dominant mature follicles on the day of hCG administration.

Serum FSH levels were measured by using MEIA (Abbott Laboratories, Abbott Park, IL) immunoassay kit. The interassay and the intraassay coefficients of variation were 3.48% and 4.52%, respectively. Serum E2 levels were measured also by using the Abbott Imx Estradiol assay (Abbott Laboratories, Abbott Park, IL). The interassay and the intraassay coefficients of variation were 7.8% and 5.8%, respectively.

Cycle cancellation rate and pregnancy rate were analyzed with respect to baseline AF assessment (≤ 6 or > 6 AF), basal cycle Day 3 FSH (≤ 10 or > 10 IU/L), and patient age (≤ 35 or > 35 years).

Statistical Analysis

Fisher's exact tests were used to compare pregnancy rates and cancellation rates between different groups. Multiple logistic regression models were used to examine the effect of age, basal FSH levels, and number of AF assessment on pregnancy rates simultaneously. Continuous data were expressed as mean $\pm$ SD. Student's tests were used to compare the mean age, FSH, E2 levels, and number of oocytes retrieved between the different groups. P -value $< .05$ was accepted as statistically significant.

RESULTS

A total of 272 consecutive IVF cycles of 224 consecutive patients enrolled in treatment cycles for IVF were included in the study. Among these cycles, 107 (40%) patients had baseline AF assessment ≤ 6 and 165 (60%) had baseline AF assessment > 6 . Among the patients, 139 (51%) were ≤ 35 years of age and 133 (49%) were > 35 years of age. The mean age in the group with AF ≤ 6 was significantly higher compared to that in the group with AF > 6 (37 ± 3.5 vs. 34 ± 4.2 years, respectively; $P < 0.001$). The group with AF ≤ 6 have a significantly higher basal FSH compared to the group with AF > 6 (8.3 ± 2.7 IU/L vs. 6.5 ± 1.5 IU/L, respectively; $P < .001$). When analyzing the response to gonadotropin stimulation, the group with AF ≤ 6 utilized more ampules of gonadotropin compared to the group with AF > 6 (49.3 ± 13 vs. 31.2 ± 13 ampules, respectively; $P < .001$). Women in the group with AF ≤ 6 had significantly lower mean peak E2 compared to those in the group with AF > 6 (1400 ± 607 vs. 2120 ± 684 pg/ml, respectively; $P < .001$). Furthermore, the mean number of retrieved oocytes was significantly lower in the group with AF ≤ 6 compared to that in the group with AF > 6 (4.8 ± 2.1 vs. 11.9 ± 4.0 , respectively; $P < .001$). There was a statistically significant correlation between baseline AF assessment and the number of retrieved oocytes ($P = .007$).

As shown in Table I, the clinical pregnancy rate was statistically significantly higher in the group with baseline AF > 6 compared to the group with AF ≤ 6 (51% vs. 20%, respectively; $P < .001$). The same number of embryos (2,3) was transferred in both group of

Table I. Pregnancy Rates in Women with AF Assessment ≤ 6 or > 6

Group	AF ≤ 6	AF > 6	P value
All patients	19% (13/70)	51% (84/165)	.001
Age ≤ 35 years	22% (5/23)	53% (56/106)	.014
Age > 35 years	17% (8/47)	47% (28/59)	.001
FSH ≤ 10 IU/L	18% (10/57)	50% (81/162)	.001
FSH > 10 IU/L	22% (2/13)	66% (2/3)	.036
Age ≤ 35 , FSH ≤ 10	18% (3/17)	52% (54/104)	.009
Age ≤ 35 , FSH > 10	33% (2/6)	100% (2/2)	.43 ^a
Age > 35 , FSH ≤ 10	18% (7/40)	47% (27/58)	.005
Age > 35 , FSH > 10	14% (1/7)	100% (1/1)	.25 ^a

^a The number of subjects was too small to reach any statistical significance.

patients to reduce the chance of multiple pregnancies. Stratified by patient age, the clinical pregnancy rate for patients ≤ 35 years was significantly higher in the group with AF > 6 compared to the group with AF ≤ 6 (53% vs. 22%, respectively; $P = .014$). The clinical pregnancy rate for patients > 35 years was significantly higher in the group with AF > 6 compared to the group with AF ≤ 6 (47% vs. 17%, respectively; $P = .001$). Controlling for basal FSH level, the pregnancy rate for women with FSH ≤ 10 IU/L was significantly higher in the group with AF > 6 compared to the group with AF ≤ 6 (50% vs. 18%, respectively; $P < .001$), in the group with FSH > 10 and AF > 6 , the number of subjects was too small to reach any statistical evaluation. When stratified by both age and basal FSH, the pregnancy rate for patients with favorable FSH (≤ 10 IU/L) was still significantly higher in the group with AF > 6 compared to the group with AF ≤ 6 regardless of patients age; 52% vs. 18% ($P = .001$) for patients ≤ 35 years, and 47% vs. 18% ($P = .005$) for patients > 35 years.

Using a multiple logistic regression model to simultaneously examine the predictive value of these determinants of IVF outcome (age, basal FSH, and AF), AF was the found to be the most significant predictor of pregnancy ($P < .05$).

The cancellation rate was significantly increased with advancing maternal age, elevated basal FSH levels, and baseline AF ≤ 6 (Table II). The overall cancellation rate was significantly lower in the group with AF > 6 compared to the group with AF ≤ 6 (1% vs. 33%, respectively; $P = .002$). Controlling for both patient age and basal FSH, the cancellation rate for patients ≤ 35 years of age with a favorable FSH ≤ 10 IU/L was significantly lower in the group with AF > 6 compared to the group with AF ≤ 6 (32% vs. 1%, respectively; $P < .001$). For patients > 35 years of age with FSH > 10 IU/L, the cancellation rate increased

Table II. Cancellation Rates in Women with AF Assessment ≤ 6 or AF > 6

Group	AF ≤ 6	AF > 6	P value
All patients	33% (35/105)	51% (2/167)	.002
Age ≤ 35 years	28% (9/32)	1% (1/107)	.001
Age > 35 years	36% (26/73)	2% (1/60)	.001
FSH ≤ 10 IU/L	32% (27/84)	1% (2/164)	.001
FSH > 10 IU/L	38% (8/21)	0% (0/3)	.53 ^a
Age ≤ 35 , FSH ≤ 10	1% (1/105)	32% (8/25)	.001
Age ≤ 35 , FSH > 10	14% (1/7)	0% (0/2)	1.00 ^a
Age > 35 , FSH ≤ 10	32% (19/59)	2% (1/59)	.001
Age > 35 , FSH > 10	50% (7/14)	0% (0/1)	1.00 ^a

^a The number of subjects was too small to reach any statistical significance.

to 50% in the group with AF ≤ 6 ($P < .001$). Even though the cancellation rate was still higher among those in the group with AF > 6 , age > 35 years, and FSH > 10 IU/L, the number of subjects was too small to reach any statistical evaluation.

DISCUSSION

This study suggests that AF assessment was a better predictor of IVF outcome than were maternal age and FSH. AF assessment contributes to the predictive value of established and widely used markers of IVF outcome, including maternal age and basal FSH level. AF assessment was a good predictor of the actual number of retrieved oocytes, pregnancy rate, and cancellation rate, independent of patient age and FSH levels.

The provision of primordial germ cells, the ovarian reserve, is a critical limiting factor in determining the reproductive life span of women and the potential chances of success of fertility treatments. The recruitment of a sufficient number of follicles during ovarian stimulation is a crucial factor in the success of assisted reproductive techniques. Poor ovarian response is major determinant of reduced IVF outcome. Quantitative assessments of ovarian reserve could be used clinically to estimate the potential ovarian response before considering IVF treatment. The availability of additional screening tests to identify patients who may respond poorly to ovarian stimulation would assist clinicians in fertility treatment counseling.

Several tests for predicting ovarian reserve have been reported. It is well documented that advanced maternal age and elevated basal FSH levels are highly accurate predictors of IVF outcome. They are widely used as screening tests of decreased ovarian reserve,

although a significant percentage of patients may have a limited response to ovarian stimulation despite normal hormonal screening parameters (15). Reuss *et al.* (11) showed in healthy volunteers an age-related decrease in the number of AF between 22 and 42 years of age by 60%. Other studies demonstrated in infertile women a negative correlation between age, follicular density, and ovarian volume (8,9). It has been suggested that ovarian volume measurement may be used as a prognosticator for ovarian responsiveness; however, it is of limited value in cases of (a) PCOS patients with increased ovarian volume due to large stroma, or (b) patients with large endometriomas, which may compromise ovarian reserve but still have a large volume, or (c) patients who had ovarian surgery, which may decrease ovarian volume but not always reduce follicular pool. Pellicer *et al.* (16) demonstrated a decrease in AF count without a difference in ovarian volume in younger low responders with normal basal FSH levels compared with normal responders. In an effort to find an additional tool as a prognostic predictor value that will help us counseling patients, we assessed the value baseline AF assessment in predicting ovarian response and IVF outcome. Several studies have looked at AF count as a predictor of IVF outcome, but no correlation with pregnancy and cancellation rates were seen (12,13). This study showed that the group with AF ≤ 6 had a reduced number of oocytes retrieved, lower pregnancy rate and higher cancellation rate, regardless of age or basal FSH levels when compared to the group with AF ≥ 6 . As the number of follicles ready to respond to gonadotropins is already limited, for patients with AF ≤ 6 , despite using an aggressive ovarian stimulation protocol it may not improve PR or reduce cancellation rates.

This study suggests that IVF outcome is strongly correlated with both patient ages, basal FSH levels, and cycle day 3 AF assessment. AF assessment was done on cycle day 3 of the planned IVF cycle and not prior to IVF, as this study was a descriptive study. A new practical study that maybe will yield more information when counseling the patients prior to enrollment in IVF program will include assessment of the number of antral follicles on a spontaneous cycle before starting IVF treatment.

AF assessment was a better predictor of IVF outcome than were age and FSH. AF assessment is a reliable, noninvasive, simple tool that can predict IVF outcome, including pregnancy and cancellation rate. This suggests that AF assessment may provide a marker of ovarian age that is distinct from chronological age or hormonal markers.

Poor ovarian response remains a major limitation to IVF outcome. Although valuable, none of the tests for prediction of ovarian response have perfect sensitivity or specificity. Therefore, the treatment strategies should not be determined by a single test but rather by using all data available. The availability of additional reliable screening tests of ovarian reserve should help identify patients who will respond poorly to ovarian stimulation, and would assist clinicians in counseling patients and selecting optimal treatment.

REFERENCES

1. Pellicer A, Lightman A, Diamond MP, *et al.*: Outcome of in vitro fertilization in women with low response to ovarian stimulation. *Fertil Steril* 1987;47:812–815
2. Scott RT, Toner JP, Muasher SJ, Oehninger SC, Robinson S, Rosenwaks Z: Follicle stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome. *Fertil Steril* 1989;51:651–654
3. Muasher SJ, Oehninger S, Simonetti S, Matta J, Ellis LM, Liu HC, Jones GS, Rosenwaks Z: The level of basal and/or stimulated serum gonadotropin levels in prediction of stimulation response and in vitro fertilization outcome. *Fertil Steril* 1988;50:298–307
4. Toner JP, Philput CB, Jones GS, *et al.*: Basal follicle-stimulating hormone level is a better predictor of in vitro fertilization performance than age. *Fertil Steril* 1991;55:784–791
5. Nahum RT, Shifren JL, Chang Y-C, Leykin L, Isaacson K, Toth TL: Age, follicle-stimulating hormone and estradiol—which is a better predictor for ovarian response in IVF. *Assist Reprod* 1999;9:176–180
6. Scott RT, Hofmann GE: Prognostic assessment of ovarian reserve. *Fertil Steril* 1995;63:1–11
7. Andolf E, Jorgensen C, Svalenius E, Sundén B: Ultrasound measurement of the ovarian volume. *Acta Obstet Gynecol Scand* 1987;66:387–389
8. Syrop CH, Willhoite A, Van Voorhis BJ: Ovarian Volume: A novel predictor for assisted reproduction. *Fertil Steril* 1995;64:1167–1171
9. Lass A, Skull J, McVeigh E, Margara R, Winston RM: Measurement of ovarian volume by transvaginal sonography before ovulation induction with human menopausal gonadotropin for in-vitro fertilization can predict poor response. *Hum Reprod* 1997;12:294–297
10. Sharara FI, McClanrock HD: The effect of aging on ovarian volume measurements in infertile women. *Obstet Gynecol* 1999;94:57–60
11. Reuss ML, Kline J, Santos R, Levin B, Timor-Tritsch I: Age and the ovarian follicle pool assessed with transvaginal ultrasonography. *Am J Obstet Gynecol* 1997;174:624–627
12. Tomas C, Nuojua-Huttunen S, Martikainen H: Pretreatment transvaginal ultrasound examination predicts ovarian responsiveness to gonadotropins in in-vitro fertilization. *Hum Reprod* 1997;12:220–223
13. Chang M-Y, Chiang C-H, Hsieh T-T, Soong Y-K, Hsu K-H: Use of the antral follicle count to predict the outcome of assisted reproductive technologies. *Fertil Steril* 1998;69:505–510
14. Adashi EY, Rock JA, Rosenwaks Z: *Reproductive Endocrinology, Surgery, and Technology*, Chapter 10: The Ovarian Follicle. Lippincott-Raven Publishers, 1996
15. Jacobs SL, Metzger DA, Dodson WC, *et al.*: Effect of age on response to human menopausal gonadotropin stimulation. *J Clin Endocrinol Metab* 1990;71:1525–1530
16. Pellicer A, Ardiles G, Neuspiller F, Remohi J, Simon C, Bonilla-Musoles F: Evaluation of the ovarian reserve in young low responders with normal levels of follicle-stimulating hormone using three-dimensional ultrasonography. *Fertil Steril* 1998;70:671–675