J Assist Reprod Genet (2013) 30:657–665
DOI 10.1007/s10815-013-9975-3

ASSISTED REPRODUCTION TECHNOLOGIES

Antral follicle count determines poor ovarian response better than anti-müllerian hormone but age is the only predictor for live birth in in vitro fertilization cycles

Mehmet Firat Mutlu · Mehmet Erdem · Ahmet Erdem · Sule Yildiz · Ilknur Mutlu · Ozgur Arisoy · Mesut Oktem

Received: 29 November 2012 / Accepted: 6 March 2013 / Published online: 19 March 2013

Abstract
Purpose To determine the predictive value of serum anti-müllerian hormone (AMH) concentrations and antral follicle counts (AFC), on ovarian response and live birth rates after IVF and compare with age and basal FSH.
Methods Basal levels of AMH, FSH and antral follicle count were measured in 192 patients prior to IVF treatment. The predictive value of these parameters were evaluated in terms of retrieved oocyte number and live birth rates.
Results Poor responders in IVF were older, had lower AFC and AMH but higher basal FSH levels. In multivariate analysis AFC was the best and only independent parameter among other parameters and AMH was better than age and basal FSH to predict poor response to ovarian stimulation. Addition of AMH, basal FSH, age and total gonadotropin dose to AFC did not improve its prognostic reliability. Area under curve (AUC) for each parameter according to ROC analysis also revealed that AFC performed better in poor response prediction compared with AMH, basal FSH and age. The cut-off point for mean AMH and AFC in discriminating the best between poor and normal ovarian response cycles was 0.94 ng/mL (with a sensitivity of 70 % and a specificity of 86 %) and 5.5 (with a sensitivity of 91 % and a specificity of 91 %), respectively. However, age was the only independent predictor of live birth in IVF as compared to hormonal and ultrasound indices of ovarian reserve.
Conclusion AFC is better than AMH to predict poor ovarian response. Although AMH and AFC could be used to predict ovarian response they had limited value in live birth prediction. The only significant predictor of the probability of achieving a live birth was age.

Keywords Anti-müllerian hormone · Antral follicle counts · Ovarian response · Live birth

Capsule AFC is better than AMH to predict poor ovarian response. Both could be used to predict ovarian response. Only significant predictor of the probability of achieving live birth was age.

M. F. Mutlu · M. Erdem · A. Erdem · S. Yildiz · M. Oktem
Department of Obstetrics & Gynecology, Gazi University Faculty of Medicine, Ankara, Turkey

I. Mutlu
Department of Obstetrics & Gynecology, Elazig Training and Research Hospital, Elazig, Turkey

O. Arisoy
Department of Obstetrics & Gynecology, Soke State Hospital, Aydın, Turkey

M. F. Mutlu (✉)
Angora Caddesi, Beysu Green Park Konutları, Beysukent, Ankara, Turkey
e-mail: firatmutlu78@hotmail.com

Introduction

It has long been known that female fecundity decreases with increasing chronologic age. The age-related decline in female fertility is most likely due to the decline in both the quantity and quality of oocytes. As a result of diminished ovarian reserve, the ability of women to conceive naturally is restricted after the age of 40 [18]. It is important to assess ovarian reserve in IVF-ICSI cycles, where the success of treatment depends on multifollicular development. Poor response to ovarian stimulation in IVF cycles directly affects the prognosis in the form of low oocyte retrieval. Assessment of poor ovarian response and correct identification of poor responders before entering an IVF program may help to direct the management of the patient with regard to gonadotropin dosing and denial of treatment [30, 37, 61].

Numerous tests have been developed to predict IVF outcome in terms of the number of oocytes retrieved and occurrence of pregnancy. Basal FSH, estradiol, and inhibin B have been used in IVF cycles to predict ovarian response and pregnancy rates, but they have limited use since they have a low predictive value, show cycle-dependent variations, may be subject to disparities between laboratory assays, and lack clear cut-off values [3, 10, 38, 54–56]. The number of antral follicles in the early follicular phase directly correlates with ovarian reserve [21] and decreased antral follicle numbers on basal ultrasound is a sign of ovarian aging, which is a feature observed prior to an increase in FSH levels [53]. It was previously demonstrated that an antral follicle count (AFC) cut-off value of 3 to 7 indicates a significant decline in ovarian reserve and subsequently poor ovarian response in IVF cycles [1, 12, 14, 22, 27, 29, 31, 42, 43, 47, 58, 64]. Although AFC was previously considered to be a better prognostic indicator compared to other endocrine markers, cycle-to-cycle variations of measurements have rendered this count somewhat disadvantageous [1, 43].

AMH, a dimeric glycoprotein, has been identified in the ovary in the granulosa cells of growing follicles up to the antral stage or to a diameter of approximately 6 mm [13, 63]. AMH production diminishes as the follicles become FSH dependent [63]. Serum levels are not affected by the day of the menstrual cycle, are most probably not be manipulated by exogenous steroid administration, and are closely correlated with reproductive age. Hence, AMH has been used to predict poor as well as excessive response in IVF [46].

Ovarian reserve tests (ORT) have been compared to predict poor response and pregnancy outcome in many studies [7, 8, 35]. However, the data are insufficient regarding the methodology and number of patients involved in the studies, the prevalence of poor responders, and the lack of efficient statistical methodology. Live births, the main goal of IVF treatment, is the main outcome measure in a few studies. For this reason, we aimed to determine the value of AMH and AFC to predict poor response in a large group of patients undergoing IVF treatment and compare their effectiveness with other traditional markers of ovarian reserve like age and basal FSH. We also aimed to determine a value of AMH and AFC that would best predict live birth rates following IVF treatment.

Materials and methods

This prospective study included 192 infertile patients referred to our Infertility and IVF clinic for IVF therapy from January 2009 to April 2010. The inclusion criteria were as follows: age between 18 and 44 years, regular menstrual cycles (21–35 days), no endocrine disorders (polycystic ovarian syndrome or polycystic ovaries on ultrasonography, normal thyroid function and normal prolactin levels), no prior ovarian surgery, and intact ovaries visible on ultrasonography. Azospermic patients in whom testicular sperm extraction techniques yielded no sperm were excluded from the study. The study protocol was approved by our Faculty Ethics Committee and written informed consent was obtained from all patients.

Serum samples were taken on days 2–3 of spontaneous menstrual cycles prior to one month of IVF/ICSI treatment in agonist and microdose flare protocol cycles for serum FSH, estradiol, and AMH measurements. In antagonist cycles, serum samples were taken on day 3 of the treatment cycle. Serum samples were separated within 1 hour of collection and stored at −60 °C until assayed for AMH. AFC measurements were carried out on cycleday 3 using an Aloka SSD-1000 (Japan) with a 5 MHz transvaginal probe by one of the two authors for each case (M.E., A.E.). In 103 (54 %) cycles AFC measurement was performed by A.E. and in the remaining 89 (46 %) cycles it was performed by M.E. After the localization of both ovaries, round or oval sonolucent structures in the ovaries were regarded as follicles. Follicles measuring less than 10 mm in diameter were counted in both ovaries to determine the antral follicle cohort. The total number of follicles in both ovaries was used as the antral follicle count with an interobserver coefficients of variation <5 %.

A pituitary suppression protocol was individually determined for each cycle. A long down-regulation protocol was used routinely in patients expected to have a normal ovarian response with a starting dose of 150–225 IU. An antagonist or microdose flare up protocol with a starting dose of 300 IU gonadotropin was used in patients expected to have a poor response. In the traditional long down-regulation protocol, 1 mg of subcutaneous (s.c.) leuprolide acetate (Lucrin;Abbot,xxx) initiated on day 21 of the previous cycle was reduced to 0.5 mg after adequate ovarian suppression was achieved and continued until the day of hCG injection. Ovarian suppression was demonstrated by the lack of ovarian activity on ultrasound and serum E_2 levels <50 pgl/mL. After adequate suppression, an appropriate dose of gonadotropin was started on an individual basis. The stimulation protocol included 150 to a maximum of 300 units of exogenous gonadotropins in the form of either recombinant or urinary FSH or FSH + hMG combination. Gonadotropin doses were adjusted individually throughout the stimulation period. In cycles using a microdose flare up protocol, patients received 21 days of oral contraceptive pills (OCs) containing 0.03 mg estradiol and 0.150 mg levonorgesterol (Microgynon; Schering AG) beginning on the third day of the menstrual cycle preceding the treatment cycle. Three days after the discontinuation of OCs, patients began taking s.c. leuprolide acetate (Lucrin; Abbot,xxx) 40 μg every 12 hours. Gonadotopins were initiated the following day. In cycles with an antagonist protocol, gonadotropins were started at an appropriate dose on day 2 of the menstrual cycle. A multiple dose antagonist protocol (Cetrorelix (Cetrotide; Serono, xxx)) was used for pituitary suppression when a leading follicle reached 12–14 mm or E_2 levels

were >500 ng/mL and continued until the day of hCG. Patients then received recombinant hCG (Ovitrelle, 250 μg, Merck Serono, xxx) when there were two or more follicles ≥18 mm in diameter with accompanying follicles ≥14 mm and an adequate E_2 response. Cycles were cancelled when there were 2 or less dominant follicles (≥14 mm) in spite of adequate gonadotropin stimulation. Poor response was defined as cycles that were cancelled due to inadequate ovarian response or the number of retrieved oocytes being less than 4. Transvaginal ovum pick up was performed 35 hours after hCG administration under ultrasound guidance. ICSI was routinely performed in all cases. Embryo transfer (ET) was performed 2 to 3 days after ovum pick-up. Up to 3 embryos were transferred and all patients received luteal support with vaginal progesterone (Crinone® 8 % gel, Merck Serono, xxx) until a pregnancy test was performed 12 days after ET. Luteal support was continued up to 10 gestational weeks. Clinical pregnancy was confirmed by increasing β-hCG concentrations and sonographic evidence of an intrauterine gestational sac after ET. Live birth was defined as a birth of an infant as a result of a treatment cycle after 24 weeks of pregnancy. For this study, a multiple pregnancy was regarded as one pregnancy.

Hormonal assays

FSH levels were measured by chemiluminescent immunometric assay with intraassay and interassay coefficients of variation of 2.6 % and 3.3 %, respectively, and E_2 levels were measured by competitive immunoassay with intraassay and interassay coefficients of variation of 2.3 % and 2.4 %, respectively (Abbot Laboratories, Illinois, USA). AMH levels were measured in batches by using a commercial kit (enzyme-linked immunosorbent assay) provided by DSL (Webster, TX, USA), with values presented in concentration of nanograms per milliliter. All of the kits used in the study were AMH Gen 1. Inter- and intraassay coefficients of variation of the assay were 4.6 % and 5.2 %, respectively.

Statistical analysis

All data were analyzed by the use of SPSS (SPSS Inc. Chicago, IL). Hormone levels were normally distributed among the study groups, and comparisons of mean serum hormone levels and ultrasound measurements between normoresponder and poor responder patients were performed using an unpaired t test. A X^2 test was used for the comparison of live birth rates between different subgroups. Multiple logistic regression analysis was used to determine the effect of variables. Age, basal FSH, AFC, and AMH were used as determinants of ovarian response. Gonadotropin dose adjustments could affect the response of patients; for that reason, total gonadotropin dose used in the cycle was also used as a determinant of ovarian response. Source of the sperm could have an impact on live birth rates. For that reason, we compared the live birth rates between the groups. There was no significant difference between the groups with regard to the source of the sperm and so it was not included in the regression analysis. ROC analysis was used to analyze the predictive accuracy of variables. Age, basal FSH, AMH, and AFC were used as variables in the multiple logistic regression analysis as determinants to predict live birth. The results are expressed as mean ± SD. A p value of <0.05 was considered significant for all analyses.

Results

Our study included 192 patients in the IVF/ICSI program with different infertility etiologies (male 36.5 %, tubal 14.5 %, anovulation 3.5 %, and unexplained 45.5 %). Of these, 117 (60.9 %) patients received a long GnRH agonist down-regulation protocol, 39 (20.3 %) patients received a microdose flare up GnRH agonist protocol, and 36 (18.8 %) patients received a GnRH antagonist protocol. The mean age of the patients was 32.6±5.8 (range 19–43). The mean infertility period was 89±68 months (range 10–360) and the mean body mass index was 25.2±4.7 kg/m^2 (range 17–40). In 10 patients, testicular sperm extraction was performed. Two patients had live births (20 %).

Clinical pregnancy rates per cycle and per embryo transfer were 34 % and 36.1 %, respectively. Eleven of the 65 pregnancies (17 %) resulted in miscarriage. All births were viable and the live birth rate was 28 %. The mean AMH levels in cycles ending with a singleton live birth were higher than those in nonpregnant women (3.5±2.6 vs. 2.2±2.5; p=0.02). The mean number of transferred embryos was 2.2±0.5 for women with singleton pregnancies and 2.1±0.4 for nonpregnant women (p>0.05). There were 13 twins (24 %) and 3 triplets (5 %). For twins the mean AFC was 13.3±6.5 and mean AMH was 3.4±2.0, while for triplets the mean AFC was 12.0±2.0 and mean AMH was 3.8±1.5. The mean number of transferred embryos was 2.5±0.5 for twins and 3.0±0.0 for triplets.

In order to assess the relation between hormonal and ultrasound parameters of ovarian reserve and the patient's ART performance, we divided the patients into subgroups of poor and good responders according to response to ovarian stimulation and number of oocytes retrieved. Thirty-one of the 192 cycles (16.1 %) were cancelled; 22 (11.5 %) were due to poor ovarian response (the others were fertilization failure in 5 cases, embryonic developmental arrest in 3 cycles, and progesterone rise on the day of hCG in one cycle). The mean age of 22 patients with cancelled cycles due to poor ovarian response was 36.4±4.7; AMH level was 0.72± 0.84 ng/ml and AFC was 3.0±2.3. Among the 161 cycles that ended in embryo transfer, 27 cycles yielded less than 4

oocytes and this was regarded as a poor response. A total of 49 (25.5 %) patients were regarded as poor responders. Poor responders were older, and had lower AFC, AMH, peak E_2 levels, peak progesterone levels, retrieved oocytes, and metaphase II oocytes in comparison with normoresponders. In addition, basal FSH levels, the duration of gonadotropin stimulation, and the total gonadotropin dose used throughout the cycle were higher in poor responders. Clinical pregnancy and live birth rates were significantly higher in normal responders (Table 1).

Univariate and multivariate regression analysis was done to assess the effects of variables in the prediction of poor response to ovarian stimulation (Table 2). According to univariate regression analysis, AMH was better than age and basal FSH in predicting poor response to ovarian stimulation (OR=0.30; 95 % CI=0.20–0.47; sensitivity=63.3 %; specificity=90.9 %; and prognostic reliability=83.9 %). Univariate regression analysis also revealed that AFC was best for predicting poor response to ovarian stimulation (OR =0.52; 95 % CI =0.43–0.64; sensitivity=83.0; specificity =89.5; and prognostic reliability=87.9 %). The addition of age to AMH and AFC did not improve the predictive value of these parameters. When AMH, AFC, age, basal FSH, and total gonadotropin dose were analyzed together, AFC was the only independent and significant factor in predicting ovarian response. Combination of other parameters with AFC did not improve its prognostic reliability.

The area under the curve (AUC) for each parameter according to ROC analysis revealed that AFC was the most accurate of the four parameters assessed to discriminate between poor and normoresponders (Table 3). The cut-off point for mean AFC that best discriminated between poor and normal ovarian response cycles was 5.5, with a sensitivity of 89 % and a specificity of 87 %. The cut-off point for mean AMH that best discriminated between poor and normal ovarian response cycles was 0.94 ng/mL, with a sensitivity of 71 % and a specificity of 85 %. Patients with both AFC and AMH levels below the cut-off values were than categorized as a group and it was observed that live births were 5.4 times lower (5.6 % vs. 30.4 %) in patients with lower AMH and AFC below the cut-off. Yet there were three live births in patients with both AFC and AMH under the cut-off values.

Patients with live births were younger than nonpregnant patients (30.1±5.1 vs. 33.5±5.7, p<0.001). Patients with live births also had significantly higher AFC (11.8±6.1 vs. 8.3±5.7, $p<0.001$) and AMH levels (3.51±2.60 vs. 2.21±2.49, $p<0.01$). Mean basal FSH levels did not differ between the groups. Table 4 shows multiple logistic regression analysis and effects of variables on prediction of live birth after IVF. Age was the only independent variable to predict live birth after IVF.

Discussion

Prediction of response to ovarian stimulation with gonadotropins before IVF treatment could help clinicians to establish an optimal treatment strategy and prevent cycle cancellations for poor responders. Whether a priori identification of actual poor responders before an IVF cycle has any prognostic value for their chances of conception remains to be established, but a number of ovarian reserve tests have been widely used before treatment to predict adequate ovarian response to gonadotropin stimulation and optimization of the gonadotropin dose. AMH and AFC have clear correlation with oocyte yield in the extremes of the response to

Table 1 Patient and cycle characteristics of study groups according to ovarian response

	Normoresponders n=143	Poor responders n=49	p value
Age (Years)	31.2±5.4	36.5±4.7	<0.001
Body mass index (kg/m^2)	24.9±4.4	25.8±5.3	NS
Duration of infertility (months)	86.6±63.1	97.8±82.2	NS
Mean antral follicle counts (n)	11.3±5.5	3.3±2.4	<0.001
Mean AMH (ng/ml)	3.17±2.70	0.87±1.03	<0.001
Basal FSH levels (mIU/ml)	6.3±2.2	11.7±8.3	<0.001
Duration of stimulation (days)	10.5±2.0	11.6±3.3	<0.05
Total gonadotropin dose (IU)	2953±1978	3838±1692	<0.01
Peak E_2 (pg/ml)	1941.1±1389.1	659. 8±804.2	<0.001
Peak progesterone (ng/ml)	0.97±0.79	0.67±0.68	<0.05
Retrieved oocyte number (n)	13.2±6.8	1.9±1.0	<0.001
Metaphase II oocyte number (n)	10.5±6.0	1.5±1.0	<0.001
Clinical pregnancy rate per cycle (%)	42	10.2	<0.001
Live birth per cycle (%)	34.3	10.2	<0.01

NS Not significant

Table 2 Multiple logistic regression analysis for hormonal, sonographic and stimulation variables and age to predict ovarian response to stimulation in IVF cycles

	Odds ratio	95 % CI	Sensitivity	Specificity	Prognostic reliability (%)	*P* value
Univariable models						
AMH	0.30	0.20–0.47	63.3	90.9	83.9	<0.001
AFC	0.52	0.43–0.64	83.0	89.5	87.9	<0.001
FSH	1.34	1.20–1.50	36.7	96.4	80.6	<0.001
Age	1.21	1.12–1.30	30.6	96.5	79.7	<0.001
Multivariable models						
Age &AMH						
Age	0.36	0.22–0.56	53.1	88.8	79.7	<0.001
AMH	1.11	1.01–1.20				<0.001
Age & AFC						
Age	0.53	0.42–0.65	83.0	89.5	87.9	<0.001
AFC	1.02	0.92–1.12				NS
Age & FSH						
Age	1.28	1.14–1.44	49.0	94.9	82.8	<0.001
FSH	1.17	1.13–1.34				<0.001
AMH&AFC&Age &FSH&TGD						
AMH	0.89	0.52–1.50	84.8	92.7	90.7	NS
AFC	0.58	0.45–0.74				<0.001
Age	1.01	0.91–1.13				NS
FSH	1.09	0.97–1.22				NS
Total gonadotropin dose	1					NS

NS Not significant

TGD Total gonadotropin dose

ovarian stimulation in IVF [46, 48]. However, the performance of AMH and other ORTs in predicting live birth, the only goal of treatment, seems more indefinite. For that reason, we aimed to ascertain the clinical value of AMH as a predictor of ovarian response and live birth rate compared to other ORTs.

In our study, we found that the best and the only independent parameter for predicting poor response was AFC. The addition of age and AMH did not improve the accuracy of AFC. ROC analysis also revealed that AFC was the most accurate of all tests in predicting poor response to ovarian stimulation; AUC for AMH was lower than AFC but better than basal FSH and age. Our cut-off point for AFC that discriminated poor response from normal response was 5.5, with a sensitivity of 89 % and specificity of 87 %. These results are in line with various studies that revealed an AFC cut-off value of 3 to 7 [1, 12, 14, 22, 27, 29, 42, 43, 47, 58, 64]. There are conflicting data regarding the performance of AMH compared to that of AFC in predicting poor response to gonadotropins in IVF cycles. AMH was better than AFC in two studies [20, 41], whereas five studies reported a similar performance of these two markers [17, 28, 40, 42, 62] and two studies found AFC to be superior to AMH [16, 32].

Table 3 Area under curve for each test according to ROC analysis to discriminate between poor and normal responded treatment cycles

	AUC (ROC)	95 % CI
Basal FSH	0.75	0.66–0.85
Age	0.76	0.68–0.84
AMH	0.86	0.80–0.92
AFC	0.93	0.89–0.98

The discrepancies between studies could be related to the heterogeneity of the study populations as well as the definitions of poor response to ovarian stimulation. Therefore we used the ESHRE Bologna criteria, a recent consensus made to standardize the definition of poor response [19]. It is well known that the prevalence of poor responders in study populations affects the predictive value of the test used to discriminate between poor and normoresponders. In our study, 25.5 % of our patients were poor responders, a relatively high rate when compared to other studies. It is well known that the size of the population as well as the prevalence of the target population (poor response in our study) affects the performance of the test used to predict

Table 4 Multiple logistic analysis for hormonal and sonographic variables and age to predict live birth in IVF cycles

	Odds ratio	95 % CI	Sensitivity	Specificity	Prognostic reliability (%)	*P* value
Multivariable models						
AMH & AFC& Age & FSH						
AMH	1.03	0.88–1.20	15.7	94	72.3	NS
AFC	1.05	0.97–1.13				NS
Age	0.92	0.86–0.99				<0.05
FSH	0.96	0.86–1.07				NS

NS Not significant

outcome. Our study comparing AMH with AFC had the largest number of patients. The performance of AMH to predict poor response and live birth in IVF in comparison to basal FSH and age was analyzed in two larger study populations than ours, but AFC was not included as a parameter in those studies [24, 46].

AFC directly represents the follicle cohort in the ovaries, which is associated with the number of oocytes retrieved in IVF. For that reason, AFC was accepted as a direct marker of the recruitable follicular cohort. However, AMH is a dimeric glycoprotein synthesized from granulosa cells of preantral and small antral follicles and represents the indirect activity of the follicular pool. In concordance with our results, it was found that AFC was superior to AMH and other ORTs in predicting ovarian response [16, 32]. However, in a systematic review of the literature, it has been shown that the hormonal (basal FSH, inhibin-B, AMH), sonographic (ovarian volume and antral follicle count), and dynamic tests (Clomiphene Citrate Challenge Test, exogenous FSH ORT (EFFORT) and gonadotropin agonist stimulating test (GAST)) have modest to poor predictive properties and clinical use of these tests seems to be irrelevant in patient counseling [7]. In a meta-analysis, it has been shown that AMH and AFC have the same level of accuracy and clinical value for the prediction of poor response [8].

The major limitations of the present study are as follows: (1) the heterogeneous nature of the population; (2) the use of individualized gonadotropin doses specific to the patients; (3) ultrasonography performed by two authors, which may have led to variability in AFC measurements.

Heterogeneity is still a limitation of the present study as well as of most of the studies that assess ovarian reserve. In the current study, a wide range of patients with advanced age were added to the study population to better determine the exact impact of age on IVF outcome in other ovarian reserve tests. In fact, a wide variety of patients with different etiologies, ages, and infertility durations apply for IVF and it is not possible to form a homogeneous population in studies that assess the impact of ovarian reserve on IVF. In the current study, many potential confounders were analyzed other than age, like source of sperm and duration of infertility for the prediction of ovarian response as well as live birth, and it was found that they were not related to the outcome and they were not used in multivariate analyses.

One other possible limitation of the present study was the use of individualized gonadotropin doses, which might have had an impact on ovarian response. In fact, the total gonadotropin dose used throughout treatment could affect the actual response of the patients. However, since this had not been previously underlined in current studies, we decided to choose total gonadotropin dose as a covariate in the multivariate analysis to predict poor responders to ovarian stimulation. We found that total gonadotropin dose did not affect the performance of tests in predicting poor ovarian response in IVF.

Although AFC proved to be a useful predictor of stimulation outcome in IVF, there might be differences in AFC measurements between observers. However, it was demonstrated in several studies that AFC has good interobserver reliability [2, 22, 49, 52]. In our study the authors performing AFC measurement had good interobserver reproducibility due to the fact that there were only two sonographers in a single center study in contrast to having several sonographers assessing AFC in a multicenter study where interobserver reproducibility is likely to be more challenging. Therefore, interobserver variability does not seem to have had a major impact on the results of the current study.

Numerous studies have reported that serum AMH levels do not show fluctuations during the menstrual cycle [26, 36, 60]. Because of the stability of the levels throughout the menstrual cycle, AMH levels can be measured randomly on any day of the menstrual cycle, providing an advantage over other endocrine and sonographic markers that should be measured in the early follicular phase. There is consensus in the literature that AMH is better in prediction of poor response than age and basal FSH. In a systemic review, Broekmans et al. found that the FSH cut-off levels of >10 U/L had a specificity of 80–90 % and a lower sensitivity of 10–30 % for the prediction of poor ovarian response to gonadotropins in IVF [7]. The lack of a clear cut-off point with reasonable sensitivity and specificity in addition to intercycle variations of FSH measurements limits the reliability and the use of basal FSH in IVF practice.

In this study, the cut-off value for AMH that discriminated a poor response from a normal response was 0.94 ng/ml,

with sensitivity and specificity of 71 % and 85 %, respectively, making it a moderately performing test. In the current literature, the AMH cut-off value for predicting poor ovarian response is between 0.30 and 1.40 [4, 11, 15, 23, 25, 28, 32, 33, 41, 44–46, 50, 51, 57, 59, 62]. Although the sensitivity and specificity of AMH performance do not exhibit significant differences among studies, there is a wide variation of cut-off values in the literature for ovarian response prediction, probably resulting from the use of different assays and different poor responder prevalence between studies. While our study adds to the literature on the predictive power of AMH and AFC, clearly more prospective data are needed to validate the cut-off points presented here in other populations. Two different kits have been developed for AMH measurement [Immunotech–Beckman Coulter and Diagnostic System Laboratories (DSL)]. Results obtained with these two assays do not always seem to overlap, with those of the Beckmann-Coulter assay being approximately 3–4 times higher than the DSL measurements [5, 23]. Standardization of assays should be achieved in order to compare the results between studies more accurately.

The main goal of IVF/ICSI treatment cycles is to achieve a pregnancy resulting in a live birth. However, clinical or ongoing pregnancies up to 12 weeks were mostly used in the majority of studies so far. In the present study, we clearly found that the best and the only independent parameter for predicting live birth rate was age. These results are in line with a recent meta-analysis by Broer et al. that included 28 studies and argued that the best predictor for ongoing pregnancy was age. In addition, the same authors concluded that neither AMH nor AFC added any information to female age to predict ongoing pregnancy after IVF [9]. AMH and other ORTs have been used to predict live birth as an outcome in only 4 studies, so far [24, 34, 39, 46]. Nelson et al. analyzed the impact of AMH, FSH, and age by a stepwise regression model to predict live birth and extremes of response in IVF [46]. They found that AMH was the only independent predictor of live birth. However, oocyte yield was the only predictor of live birth when added to the analysis. In two recent studies, Gleicher et al. and La Marca et al. used only AMH and age in a regression analysis to predict live births after IVF [24, 34] and they both found that age and AMH were independent variables to predict live birth after IVF. Furthermore, a very recent prospective study including 336 patients undergoing their first IVF cycles clearly demonstrated that, among the ovarian reserve tests, AMH and age were the sole predictive factors of live birth for women >35 years, whereas the number of good-quality embryos was the only factor to predict live birth in women <35 [39]. In contrast to these studies, we used AFC and basal FSH in a regression model, in addition to AMH and age, which could explain the different results obtained from the two studies in comparison to ours. There are numerous factors such as age, embryo quality, transfer technique, and endometrial receptivity that determine the chance of achieving pregnancy other than the cohort size and number of retrieved oocytes after IVF [6]; this appears to be the reason why tests are not sensitive enough to predict pregnancy outcome after IVF.

Counseling and management of the cycle with knowledge gained only from the ORTs is a matter of debate. In fact many women whose tests results were under the cut-off value were able to achieve pregnancy after IVF. In our study, 17 % of live births were observed in patients who had AMH levels below the cut-off determined for poor ovarian response prediction. For that reason AMH measurement, similarly to that of other ovarian reserve markers, should not be used to exclude couples from IVF. The addition of patients with low AFC to the patients who had AMH level below the cut-off decreased the false positive rates in our study. When both of these parameters are under the cut-off value used in poor ovarian response prediction (0.94 ng/ml for AMH, 5.5 for AFC) the probability of achieving live birth is only 5.6 %, nearly 5.4 times lower than that of the group who have at least one parameter above the cut-off, regardless of age. Yet there were three live births when both AFC and AMH were under the cut-off values. Patients that had live births and had values under the cut-off were all below the age of 40. Pregnancy may occur even at extreme cut-offs for an abnormal test result for most of the tests commonly used to predict ovarian response like AMH, AFC, and FSH; therefore, IVF treatment should not be denied based on these tests, especially in patients who are seeking their first cycle of treatment. It should be emphasized that values below the cut-offs in patients with advanced age (>40 years of age) may indicate a lower rate of live birth.

In conclusion, this study clearly indicates that none of the current hormonal indices of ovarian reserve can predict ovarian response to gonadotropins better than sonographic evaluation of the antral follicle cohort. However, age is the only determinant of live birth prediction in IVF cycles. Patients could have pregnancies in the lower extremes of even AMH, the most widely used hormonal test nowadays; for that reason IVF should not be denied based on hormonal ORT.

Acknowledgment This research received funding from the Scientific Research Projects Coordination Unit of Gazi University (grant 01/2005-2).

References

1. 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(2):328–36.
2. Bancsi LF, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization. Fertil Steril. 2004;81(1):35–41.

3. Bancsi LF, Huijs AM, den Ouden CT, Broekmans FJ, Looman CW, 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(3):552–7.
4. Barad DH, Weghofer A, Gleicher N. Comparing anti-Mullerian hormone (AMH) and follicle-stimulating hormone (FSH) as predictors of ovarian function. Fertil Steril. 2009;91(4 Suppl):1553–5. doi:10.1016/j.fertnstert.2008.09.069.
5. Bersinger NA, Wunder D, Birkhauser MH, Guibourdenche J. Measurement of anti-mullerian hormone by Beckman Coulter ELISA and DSL ELISA in assisted reproduction: differences between serum and follicular fluid. Clin Chim Acta Int J Clin Chem. 2007;384(1–2):174–5. doi:10.1016/j.cca.2007.05.011.
6. Boomsma CM, Macklon NS. What can the clinician do to improve implantation? Reprod Biomed Online. 2007;14(Spec No 1):27–37. doi:10.1016/S1472-6483(10)61456-3.
7. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Updat. 2006;12(6):685–718. doi:10.1093/humupd/dml034.
8. Broer SL, Mol BW, Hendriks D, Broekmans FJ. The role of antimullerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. Fertil Steril. 2009;91(3):705–14. doi:10.1016/j.fertnstert.2007.12.013.
9. Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC, Bossuyt P, et al. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Updat. 2013;19(1):26–36. doi:10.1093/humupd/dms041.
10. Bukman A, Heineman MJ. Ovarian reserve testing and the use of prognostic models in patients with subfertility. Hum Reprod Updat. 2001;7(6):581–90.
11. Buyuk E, Seifer DB, Younger J, Grazi RV, Lieman H. Random anti-Mullerian hormone (AMH) is a predictor of ovarian response in women with elevated baseline early follicular follicle-stimulating hormone levels. Fertil Steril. 2011;95(7):2369–72. doi:10.1016/j.fertnstert.2011.03.071.
12. 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(3):505–10.
13. Durlinger AL, Kramer P, Karels B, de Jong FH, Uilenbroek JT, Grootegoed JA, et al. Control of primordial follicle recruitment by anti-Mullerian hormone in the mouse ovary. Endocrinology. 1999;140(12):5789–96.
14. Durmusoglu F, Elter K, Yoruk P, Erenus M. Combining cycle day 7 follicle count with the basal antral follicle count improves the prediction of ovarian response. Fertil Steril. 2004;81(4):1073–8. doi:10.1016/j.fertnstert.2003.08.044.
15. Ebner T, Sommergruber M, Moser M, Shebl O, Schreier-Lechner E, Tews G. Basal level of anti-Mullerian hormone is associated with oocyte quality in stimulated cycles. Hum Reprod. 2006;21(8):2022–6. doi:10.1093/humrep/del127.
16. Eldar-Geva T, Ben-Chetrit A, Spitz IM, Rabinowitz R, Markowitz E, Mimoni T, et al. Dynamic assays of inhibin B, anti-Mullerian hormone and estradiol following FSH stimulation and ovarian ultrasonography as predictors of IVF outcome. Hum Reprod. 2005;20(11):3178–83. doi:10.1093/humrep/dei203.
17. Elgindy EA, El-Haieg DO, El-Sebaey A. Anti-Mullerian hormone: correlation of early follicular, ovulatory and midluteal levels with ovarian response and cycle outcome in intracytoplasmic sperm injection patients. Fertil Steril. 2008;89(6):1670–6. doi:10.1016/j.fertnstert.2007.05.040.
18. Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: implications for forecasting menopause. Hum Reprod. 1992;7(10):1342–6.
19. Ferraretti AP, La Marca A, Fauser BC, Tarlatzis B, Nargund G, Gianaroli L. ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum Reprod. 2011;26(7):1616–24. doi:10.1093/humrep/der092.
20. Ficicioglu C, Kutlu T, Baglam E, Bakacak Z. Early follicular antimullerian hormone as an indicator of ovarian reserve. Fertil Steril. 2006;85(3):592–6. doi:10.1016/j.fertnstert.2005.09.019.
21. 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(3):512–7.
22. Frattarelli JL, Levi AJ, Miller BT, Segars JH. A prospective assessment of the predictive value of basal antral follicles in in vitro fertilization cycles. Fertil Steril. 2003;80(2):350–5.
23. Freour T, Mirallie S, Bach-Ngohou K, Denis M, Barriere P, Masson D. Measurement of serum anti-Mullerian hormone by Beckman Coulter ELISA and DSL ELISA: comparison and relevance in assisted reproduction technology (ART). Clin Chim Acta Int J Clin Chem. 2007;375(1–2):162–4. doi:10.1016/j.cca.2006.06.013.
24. Gleicher N, Weghofer A, Barad DH. Anti-Mullerian hormone (AMH) defines, independent of age, low versus good live-birth chances in women with severely diminished ovarian reserve. Fertil Steril. 2010;94(7):2824–7. doi:10.1016/j.fertnstert.2010.04.067.
25. Gnoth C, Schuring AN, Friol K, Tigges J, Mallmann P, Godehardt E. Relevance of anti-Mullerian hormone measurement in a routine IVF program. Hum Reprod. 2008;23(6):1359–65. doi:10.1093/humrep/den108.
26. Hehenkamp WJ, Looman CW, Themmen AP, de Jong FH, Te Velde ER, Broekmans FJ. Anti-Mullerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. J Clin Endocrinol Metab. 2006;91(10):4057–63. doi:10.1210/jc.2006-0331.
27. Hsieh YY, Chang CC, Tsai HD. Antral follicle counting in predicting the retrieved oocyte number after ovarian hyperstimulation. J Assist Reprod Genet. 2001;18(6):320–4.
28. Jayaprakasan K, Campbell B, Hopkisson J, Johnson I, Raine-Fenning N. A prospective, comparative analysis of anti-Mullerian hormone, inhibin-B, and three-dimensional ultrasound determinants of ovarian reserve in the prediction of poor response to controlled ovarian stimulation. Fertil Steril. 2010;93(3):855–64. doi:10.1016/j.fertnstert.2008.10.042.
29. Klinkert ER, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. The antral follicle count is a better marker than basal follicle-stimulating hormone for the selection of older patients with acceptable pregnancy prospects after in vitro fertilization. Fertil Steril. 2005;83(3):811–4. doi:10.1016/j.fertnstert.2004.11.005.
30. Klinkert ER, Broekmans FJ, Looman CW, Te Velde ER. A poor response in the first in vitro fertilization cycle is not necessarily related to a poor prognosis in subsequent cycles. Fertil Steril. 2004;81(5):1247–53. doi:10.1016/j.fertnstert.2003.10.030.
31. Kupesic S, Kurjak A, Bjelos D, Vujisic S. Three-dimensional ultrasonographic ovarian measurements and in vitro fertilization outcome are related to age. Fertil Steril. 2003;79(1):190–7.
32. Kwee J, Schats R, McDonnell J, Themmen A, de Jong F, Lambalk C. Evaluation of anti-Mullerian hormone as a test for the prediction of ovarian reserve. Fertil Steril. 2008;90(3):737–43. doi:10.1016/j.fertnstert.2007.07.1293.
33. La Marca A, Giulini S, Tirelli A, Bertucci E, Marsella T, Xella S, et al. Anti-Mullerian hormone measurement on any day of the menstrual cycle strongly predicts ovarian response in assisted reproductive technology. Hum Reprod. 2007;22(3):766–71. doi:10.1093/humrep/del421.
34. La Marca A, Nelson SM, Sighinolfi G, Manno M, Baraldi E, Roli L, et al. Anti-Mullerian hormone-based prediction model for a live birth in assisted reproduction. Reprod Biomed Online. 2011;22(4):341–9. doi:10.1016/j.rbmo.2010.11.005.

35. La Marca A, Sighinolfi G, Radi D, Argento C, Baraldi E, Artenisio AC, et al. Anti-Mullerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). Hum Reprod Updat. 2010;16(2):113–30. doi:10.1093/humupd/dmp036.
36. La Marca A, Stabile G, Artenisio AC, Volpe A. Serum anti-Mullerian hormone throughout the human menstrual cycle. Hum Reprod. 2006;21(12):3103–7. doi:10.1093/humrep/del291.
37. 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(4):964–9.
38. Lass A, Gerrard A, Abusheikha N, Akagbosu F, Brinsden P. IVF performance of women who have fluctuating early follicular FSH levels. J Assist Reprod Genet. 2000;17(10):566–73.
39. Lee TH, Liu CH, Huang CC, Hsieh KC, Lin PM, Lee MS. Impact of female age and male infertility on ovarian reserve markers to predict outcome of assisted reproduction technology cycles. Reprod Biol Endocrinol. 2009;7:100. doi:10.1186/1477-7827-7-100.
40. Lekamge DN, Barry M, Kolo M, Lane M, Gilchrist RB, Tremellen KP. Anti-Mullerian hormone as a predictor of IVF outcome. Reprod Biomed Online. 2007;14(5):602–10.
41. McIlveen M, Skull JD, Ledger WL. Evaluation of the utility of multiple endocrine and ultrasound measures of ovarian reserve in the prediction of cycle cancellation in a high-risk IVF population. Hum Reprod. 2007;22(3):778–85. doi:10.1093/humrep/del435.
42. Muttukrishna S, McGarrigle H, Wakim R, Khadum I, Ranieri DM, Serhal P. Antral follicle count, anti-mullerian hormone and inhibin B: predictors of ovarian response in assisted reproductive technology? BJOG Int J Obstet Gynaecol. 2005;112(10):1384–90. doi:10.1111/j.1471-0528.2005.00670.x.
43. Nahum R, Shifren JL, Chang Y, Leykin L, Isaacson K, Toth TL. Antral follicle assessment as a tool for predicting outcome in IVF–is it a better predictor than age and FSH? J Assist Reprod Genet. 2001;18(3):151–5.
44. Nakhuda GS, Sauer MV, Wang JG, Ferin M, Lobo RA. Mullerian inhibiting substance is an accurate marker of ovarian response in women of advanced reproductive age undergoing IVF. Reprod Biomed Online. 2007;14(4):450–4.
45. Nardo LG, Gelbaya TA, Wilkinson H, Roberts SA, Yates A, Pemberton P, et al. Circulating basal anti-Mullerian hormone levels as predictor of ovarian response in women undergoing ovarian stimulation for in vitro fertilization. Fertil Steril. 2009;92(5):1586–93. doi:10.1016/j.fertnstert.2008.08.127.
46. Nelson SM, Yates RW, Fleming R. Serum anti-Mullerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles–implications for individualization of therapy. Hum Reprod. 2007;22(9):2414–21. doi:10.1093/humrep/dem204.
47. Ng EH, Tang OS, Ho PC. The significance of the number of antral follicles prior to stimulation in predicting ovarian responses in an IVF programme. Hum Reprod. 2000;15(9):1937–42.
48. Ocal P, Sahmay S, Cetin M, Irez T, Guralp O, Cepni I. Serum anti-Mullerian hormone and antral follicle count as predictive markers of OHSS in ART cycles. J Assist Reprod Genet. 2011;28(12):1197–203. doi:10.1007/s10815-011-9627-4.
49. 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(4):638–42.
50. Penarrubia J, Fabregues F, Manau D, Creus M, Casals G, Casamitjana R, et al. Basal and stimulation day 5 anti-Mullerian hormone serum concentrations as predictors of ovarian response and pregnancy in assisted reproductive technology cycles stimulated with gonadotropin-releasing hormone agonist–gonadotropin treatmen. Hum Reprod. 2005;20(4):915–22. doi:10.1093/humrep/deh718.
51. Riggs RM, Duran EH, Baker MW, Kimble TD, Hobeika E, Yin L, et al. Assessment of ovarian reserve with anti-Mullerian hormone: a comparison of the predictive value of anti-Mullerian hormone, follicle-stimulating hormone, inhibin B, and age. Am J Obstet Gynecol. 2008;199(2):202. doi:10.1016/j.ajog.2008.05.004. e201-208.
52. Scheffer GJ, Broekmans FJ, Bancsi LF, Habbema JD, Looman CW, Te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: reproducibility of antral follicle counts. Ultrasound Obstet Gynecol. 2002;20(3):270–5. doi:10.1046/j.1469-0705.2002.00787.x.
53. 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(5):845–51.
54. Scott Jr RT, Hofmann GE, Oehninger S, Muasher SJ. Intercycle variability of day 3 follicle-stimulating hormone levels and its effect on stimulation quality in in vitro fertilization. Fertil Steril. 1990;54(2):297–302.
55. Seifer DB, Lambert-Messerlian G, Hogan JW, Gardiner AC, Blazar AS, Berk CA. Day 3 serum inhibin-B is predictive of assisted reproductive technologies outcome. Fertil Steril. 1997;67(1):110–4.
56. Seifer DB, Scott Jr RT, Bergh PA, Abrogast LK, Friedman CI, Mack CK, et al. Women with declining ovarian reserve may demonstrate a decrease in day 3 serum inhibin B before a rise in day 3 follicle-stimulating hormone. Fertil Steril. 1999;72(1):63–5.
57. Smeenk JM, Sweep FC, Zielhuis GA, Kremer JA, Thomas CM, Braat DD. Antimullerian hormone predicts ovarian responsiveness, but not embryo quality or pregnancy, after in vitro fertilization or intracyoplasmic sperm injection. Fertil Steril. 2007;87(1):223–6. doi:10.1016/j.fertnstert.2006.06.019.
58. Soldevila PN, Carreras O, Tur R, Coroleu B, Barri PN. Sonographic assessment of ovarian reserve. Its correlation with outcome of in vitro fertilization cycles. Gynecol Endocrinol Off J Int Soc Gynecol Endocrinol. 2007;23(4):206–12. doi:10.1080/09513590701253776.
59. Tremellen KP, Kolo M, Gilmore A, Lekamge DN. Anti-mullerian hormone as a marker of ovarian reserve. Aust N Z J Obstet Gynaecol. 2005;45(1):20–4. doi:10.1111/j.1479-828X.2005.00332.x.
60. Tsepelidis S, Devreker F, Demeestere I, Flahaut A, Gervy C, Englert Y. Stable serum levels of anti-Mullerian hormone during the menstrual cycle: a prospective study in normo-ovulatory women. Hum Reprod. 2007;22(7):1837–40. doi:10.1093/humrep/dem101.
61. van Disseldorp J, Eijkemans MJ, Klinkert ER, te Velde ER, Fauser BC, Broekmans FJ. Cumulative live birth rates following IVF in 41- to 43-year-old women presenting with favourable ovarian reserve characteristics. Reprod Biomed Online. 2007;14(4):455–63.
62. van Rooij IA, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, de Jong FH, et al. Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve. Hum Reprod. 2002;17(12):3065–71.
63. Weenen C, Laven JS, Von Bergh AR, Cranfield M, Groome NP, Visser JA, et al. Anti-Mullerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. Mol Hum Reprod. 2004;10(2):77–83.
64. Yong PY, Baird DT, Thong KJ, McNeilly AS, Anderson RA. Prospective analysis of the relationships between the ovarian follicle cohort and basal FSH concentration, the inhibin response to exogenous FSH and ovarian follicle number at different stages of the normal menstrual cycle and after pituitary down-regulation. Hum Reprod. 2003;18(1):35–44.