Human Reproduction Vol.20, No.4 pp. 915–922, 2005 doi:10.1093/humrep/deh718
Advance Access publication January 21, 2005

Basal and stimulation day 5 anti-Müllerian hormone serum concentrations as predictors of ovarian response and pregnancy in assisted reproductive technology cycles stimulated with gonadotropin-releasing hormone agonist–gonadotropin treatment

Joana Peñarrubia[1], Francisco Fábregues[1], Dolors Manau[1], Montserrat Creus[1], Gemma Casals[1], Roser Casamitjana[2], Franciso Carmona[1], Juan A.Vanrell[1] and Juan Balasch[1,3]

[1]Institut Clínic of Gynecology, Obstetrics and Neonatology and [2]Hormonal Laboratory, Faculty of Medicine–University of Barcelona, Hospital Clínic–Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[3]To whom correspondence should be addressed at: Institut Clínic of Gynecology and Obstetrics, Hospital Clínic, C/Casanova 143, 08036 Barcelona, Spain. E-mail: jbalasch@ub.edu

BACKGROUND: Anti-Müllerian hormone (AMH) has been recently proposed as a marker for ovarian ageing and poor ovarian response to controlled ovarian hyperstimulation in assisted reproduction cycles. The present study was undertaken to investigate the usefulness of baseline cycle day 3 AMH levels and AMH serum concentrations obtained on the fifth day of gonadotropin therapy in predicting ovarian response and pregnancy in women undergoing ovarian stimulation with FSH under pituitary desensitization for assisted reproduction. METHODS: A total of 80 women undergoing their first cycle of IVF/intracytoplasmic sperm injection (ICSI) treatment were studied. Twenty consecutive cycles which were cancelled because of a poor follicular response were initially selected. As a control group, 60 women were randomly selected from our assisted reproduction programme matching by race, age, body mass index, basal FSH and indication for IVF/ICSI to those in the cancelled group. For each cancelled patient, three IVF/ICSI women who met the matching criteria were included. RESULTS: Basal and day 5 AMH serum concentrations were significantly lower in the cancelled than in the control group. Receiver-operating characteristic (ROC) analysis showed that the capacity of day 5 AMH in predicting the likelihood of cancellation in an assisted reproduction treatment programme was significantly higher than that for basal AMH measurement. However, the predictive capacity of day 5 AMH was not better than that provided by day 5 estradiol. In addition, neither basal nor day 5 AMH or estradiol measurements were useful in the prediction of pregnancy after assisted reproductive treatment. CONCLUSIONS: AMH concentrations obtained early in the follicular phase during ovarian stimulation under pituitary suppression for assisted reproduction are better predictors of ovarian response than basal AMH measurements. However, AMH is not useful in the prediction of pregnancy. Definite clinical applicability of AMH determination as a marker of IVF outcome remains to be established.

Key words: AMH/IVF/low responders/ovarian reserve/ovarian response

Introduction

Recruitment and development of multiple ovarian follicles in response to gonadotropin stimulation are necessary for successful assisted reproductive treatment. The ability of the ovaries to respond to gonadotropins with adequate follicular development has been referred to as ovarian reserve. Although ovarian reserve declines with age, it is a biological and not just a chronological function. In fact, the most important aspect of diminished ovarian reserve is that the timing of its onset is highly variable (Scott and Hofmann, 1995). Thus, a useful biomarker of ovarian reponse to controlled ovarian hyperstimulation in assisted reproduction is needed.

Many tests have been developed to screen for diminished ovarian reserve. Traditional methods used to predict prospectively response to ovarian stimulation have included mainly the measurement of baseline cycle day 3 serum concentrations of hormones such as FSH, estradiol and inhibins, or ultrasonographic tests such as pretreatment ovarian volume and the number of early antral follicles (Bukman and Heineman, 2001). On their own, however, normal baseline values are not a guarantee that an endocrine organ is

functioning normally, and non-response to ovarian stimulation in normogonadotropic, normogonadal women has been reported (Wallach, 1995; Farhi *et al.*, 1997). As recently stressed (Tarlatzis *et al.*, 2003), despite the plethora of predictive tests for low ovarian response, the poor responder is revealed definitely only during ovarian stimulation. Therefore, the ability to predict ovarian response early in the course of controlled ovarian hyperstimulation would provide the opportunity of obtaining essential information to assist in deciding whether to proceed with an ongoing cycle. Recent studies by us (Peñarrubia *et al.*, 2000) and others (Phelps *et al.*, 1998; Eldar-Geva *et al.*, 2000; Fawzy *et al.*, 2002) have addressed this issue and concluded that estradiol and inhibin B levels obtained after 3–4 days of gonadotropin therapy rather than baseline values were highly predictive of ovarian response in assisted reproduction treatment cycles.

Recently, anti-Müllerian hormone (AMH), also referred to as Müllerian-inhibiting substance, has been proposed as a novel marker for predicting ovarian response to gonadotropin stimulation (Seifer *et al.*, 2002; van Rooij *et al.*, 2002; Gruijters *et al.*, 2003). AMH is a member of the transforming growth factor β superfamily of growth and differentiation factors. It was identified as a factor which, being synthesized by testicular Sertoli cells, induces regression of the Müllerian ducts during male fetal development. In females, AMH is only expressed by the ovary, and mRNA studies in rat and mouse species revealed specific expression of AMH in granulosa cells of early growing, preantral and small antral follicles but not in non-atretic large antral follicles and all atretic follicles (Baarends *et al.*, 1995; Gruijters *et al.*, 2003). It has been shown that AMH affects two important regulatory steps during folliculogenesis in female mice (Durlinger *et al.*, 2002; Gruijters *et al.*, 2003). At initial recruitment, AMH inhibits recruitment of primordial follicles into the growing pool, whereas at cyclic recruitment AMH lowers the FSH sensitivity of follicles. In these ways, AMH plays an essential role in regulating ovarian follicular growth in rodents. Notably, very recent data indicate that AMH expression in the human follows a similar pattern as compared to the mouse and rat, thus suggesting important roles for AMH in human folliculogenesis (Weenen *et al.*, 2004).

It has been shown that human female serum contains measurable levels of AMH during the reproductive life (Lee *et al.*, 1996). AMH serum levels decline with increasing female age in normo-ovulatory women (de Vet *et al.*, 2002) and are more strongly correlated with the number of early antral follicles than the usual hormone markers such as FSH, LH, estradiol and inhibin B on cycle day 3 (Fanchin *et al.*, 2003a). Thus, AMH has been proposed as a marker for ovarian ageing (de Vet *et al.*, 2002; van Roij et al., 2002; Fanchin *et al.*, 2003a). In fact, poor ovarian response to controlled ovarian hyperstimulation in assisted reproduction cycles, which is considered as indicative of ovarian ageing (Beckers *et al.*, 2002), has been demonstrated to be associated with reduced early follicular phase AMH serum levels (Seifer *et al.*, 2002; van Rooij *et al.*, 2002; Fanchin *et al.*, 2003b).

On the above evidence, the present study was undertaken to investigate the usefulness of baseline cycle day 3 AMH levels and AMH serum concentrations obtained on the fifth day of gonadotropin therapy in predicting ovarian response and pregnancy in women undergoing ovarian stimulation under pituitary desensitization for assisted reproduction. The fifth day of gonadotropin therapy was used because ovarian stimulation is routinely started on Thursday in our assisted reproduction programme and the first evaluation of the ovarian response is performed on Monday (i.e. after 4 days of gonadotropin treatment).

Materials and methods

Patients studied

The study involved 80 women undergoing their first cycle of IVF/intracytoplasmic sperm injection (ICSI) treatment, thus avoiding possible bias from experience with previous cycles regarding ovarian response to exogenous gonadotropin stimulation. Twenty consecutive cycles which were cancelled because of a poor follicular response were initially selected. As a control group, 60 women having a completed IVF/ICSI cycle were randomly selected from our assisted reproduction programme matching by race, age (± 1 years), body mass index (BMI) (± 1 kg/m^2), basal FSH (± 0.5 IU/l) and indication for IVF/ICSI to those in the cancelled group. For each cancelled cycle, three IVF/ICSI women who met the matching criteria were included. Patients included in the current investigation underwent assisted reproductive treatment over the period October 2003 to April 2004. This case–control study design has been previously used by us (Balasch *et al.*, 1996; Creus *et al.*, 2000; Peñarrubia *et al.*, 2000) and others (Hall *et al.*, 1999) in studies investigating the usefulness of inhibins as predictors of assisted reproduction treatment outcome. This allows the use of appropriate matched patients having undergoing IVF within a similar and reasonable short timeframe when necessary (Hall *et al.*, 1999; Peñarrubia *et al.*, 2000).

All patients had both ovaries with no previous ovarian surgery and normal ovulatory function according to midluteal plasma progesterone concentrations and regular menses. In our assisted reproduction programme, basal FSH, LH and estradiol serum levels are routinely measured in the early follicular phase within the 3 months preceding IVF/ICSI treatment, and estradiol serum concentrations on the fifth day of gonadotropin therapy are routinely used to evaluate ovarian response. For the specific purpose of this study all subjects had serum AMH determinations on day 3 of their cycle within 3 months of the IVF/ICSI attempt and on the fifth day of gonadotropin therapy during the IVF/ICSI index cycle, which was measured on completion of the study in frozen blood samples.

Stimulation regimen

All patients received standard ovarian stimulation with FSH under pituitary suppression with GnRH agonist, according to a protocol previously reported (Peñarrubia *et al.*, 2003). In all women, pituitary desensitization was achieved by s.c. administration of triptorelin acetate (Decapeptyl 0.1 mg; Ipsen Pharma, Barcelona, Spain; 0.1 mg daily, which was reduced to 0.05 mg after ovarian arrest was confirmed) started in the mid-luteal phase of the previous cycle. Gonadotropin stimulation of the ovaries was started when serum estradiol concentrations declined to < 50 pg/ml and a vaginal ultrasonographic scan showed an absence of follicles >10 mm diameter. On days 1 and 2 of ovarian stimulation, 450 IU and 300 IU/day of recombinant human FSH (Gonal-F; Serono, Madrid, Spain), respectively, were administered subcutaneously. On days 3 and 4 of

ovarian stimulation, 150 IU per day of FSH were administered to each patient. From day 5 onward, FSH was administered on an individual basis according to the ovarian response, as assessed by sequential transvaginal ultrasonography and serum estradiol measurements. The criteria for human chorionic gonadotropin administration (recombinant human HCG; 250 μg) (Ovitrelle; Serono) were the presence of two or more follicles >18 mm in diameter with ≥4 follicles measuring ≥14 mm in association with a consistent rise in serum estradiol concentration. Oocyte aspiration was performed with vaginal ultrasonography 35–36 h after HCG administration. The maturational status of the oocytes and the embryo grading were recorded according to published criteria (Veeck, 1999); embryos of Veeck grades 1 or 2 were considered high quality. Up to three embryos per patient were replaced and the luteal phase was supported with vaginal micronized progesterone. The cycle was cancelled when there were <3 follicles with diameter ≥14 mm after 8–9 days of gonadotropin therapy (early cancellation) or after 4–5 additional treatment days without attaining, or the imminent prospect of attaining, the criteria for HCG administration (late cancellation).

Pregnancy was diagnosed by increasing serum concentrations of β-HCG after embryo transfer, and the subsequent demonstration of an intrauterine gestational sac by ultrasonography.

Hormone analyses and ultrasonography

Blood samples were drawn between 08:00–10:00 h and processed within 2 h after withdrawal. For this study two serum aliquots were obtained. FSH, LH and estradiol were measured in one of the serum aliquots for clinical monitoring, and the second aliquot was stored at −20 °C for later measurement of AMH. Frozen serum samples from each patient for AMH measurement were examined in one run within 6 months of collection.

FSH, LH, estradiol and AMH in serum were measured using commercially available kits as reported previously (Balasch *et al.*, 2001; Peñarrubia *et al.*, 2003; Pigny *et al.*, 2003). Estradiol concentrations in serum were estimated by a competitive immunoenzymatic assay (Immuno 1, Technicon; Bayer, Tarrytown, NY). The sensitivity was 10 pg/ml and the interassay coefficient of variation (CV) was 5%. FSH and LH serum concentrations were measured by an immunoenzymatic assay with two monoclonal antibodies (Immuno 1, Technicon; Bayer) and data expressed in terms of IRP 78/549 and 68/40 respectively. The sensitivity of the assays was 0.1 IU/l for FSH and 0.3 IU/l for LH, and interassay CV were 2.7 and 3.1%, respectively. Total β-HCG was measured by a solid-phase, two-site chemiluminiscent enzyme immunometric assay standardized against the Third International Standard 75/537 (Immulite, Diagnostic Products Co., Los Angeles, CA) with a detection limit of 2 IU/l. The inter-assay CV was 5.8%. Serum AMH levels were determined in duplicate using an ultrasensitive enzyme-linked immunosorbent assay (ELISA) (Immunotech-Coulter, Marseilles, France) according to the supplier's instructions. Results are expressed in pmol/l using human recombinant AMH as a standard. The detection limit of this assay using the ultrasensitive protocol is 0.7 pmol/l. Intra- and inter-assay coefficients of variation were <5.5% and <9%, respectively.

Ultrasonic scans were performed using a Toshiba Eccocee SAA-340A/EF unit (Toshiba Co., Tokyo, Japan) equipped with a 5–7 MHz endovaginal probe (PVF-641VT).

Statistics and probability testing

For statistical analysis the Mann–Whitney *U* test, the Wilcoxon matched-pairs signed ranks test and χ^2-test were used as appropriate. Results are expressed as mean ± standard error of mean (SEM). $P < 0.05$ was considered significant. The discrimination attained between two study groups (cancelled vs non-cancelled cycles, and conception vs non-conception cycles) was evaluated with receiver-operating characteristic (ROC) analysis (Hanley and McNeil, 1982; Zweig and Campbell, 1993). ROC curves are plots of all the sensitivity and specificity pairs which are possible for all levels of a particular parameter. They are constructed by plotting the false positive rate or 100-specificity on the *x*-axis. The *y*-axis shows the true positive rate or sensitivity. The best cut-off value discriminating between two conditions is the value located at the greatest distance from the diagonal.

Calculation of the area under the ROC curve (AUC_{ROC}) provides the quantitative measure of accuracy, i.e. the ability of a particular parameter (e.g. AMH serum concentrations) to discriminate between two conditions (e.g. cancelled vs punctured cycles). Sensitivity, specificity and the AUC_{ROC} were obtained for each model. 95% confidence intervals (CI) were calculated for each of the estimates. The models' AUC_{ROC} values were compared using the method of Hanley and McNeil (1982). An ROC curve representing a parameter with no discrimination at all is a 45° diagonal line from the left lower corner (0% true positive rate and 0% false positive rate) to the upper right corner (100% true positive rate and 100% false positive rate) with an area under the curve of 0.5. Thus, an AUC_{ROC} whose CI includes 0.5 means no discrimination. A parameter with no overlap between the two conditions will discriminate perfectly and has a ROC curve passing along the *y*-axis to the upper left corner (100% true positive rate and 0% false positive rate) to end again in the upper right corner with an area under the curve of 1.0.

Data were analyzed by Statistics Package for Social Sciences (SPSS version 10.0, Chicago, IL).

Results

Table I shows patient characteristics, basal and day 5 AMH serum concentrations, dose of gonadotropin used, and ovarian response observed in cancelled and control groups of patients. As expected, mean age and BMI, as well as basal FSH serum levels, were very similar in both groups. Indications for assisted reproductive treatment were obviously identical for both groups. Both basal and day 5 serum concentrations of AMH were significantly higher in the control than in the cancelled group ($P < 0.05$ and $P < 0.001$, respectively). AMH serum levels on day 5 were significantly lower than baseline AMH concentrations both in cancelled ($P < 0.01$) and non-cancelled ($P < 0.05$) groups. The early cancellation group included 10 patients lacking any follicular growth after 8–9 days of gonadotropin therapy while in the remaining 10 cancelled cases there were only 1–2 growing follicles after 4–5 additional gonadotropin treatment days. Basal and stimulation day 5 AMH serum levels were similar in the early cancellation (25.6 ± 10.2 pmol/l and 18.8 ± 10.1 pmol/l, respectively) and late cancellation (19.7 ± 7.8 pmol/l and 12.6 ± 2.8 pmol/l, respectively) groups (P = N.S.).

There were no differences regarding basal LH and estradiol levels, duration and amount of gonadotropin treatment between cancelled cycles and controls but, as expected, estradiol peak concentrations and the number of follicles recruited were significantly higher in the control group. The number of oocytes retrieved and the clinical pregnancy rate per puncture in the control group were 9.1 ± 0.54 and 45% (27/60),

Table I. Patient characteristics, basal and day 5 AMH concentrations and ovarian response in the two groups studied

Variable	Cancelled group ($n = 20$)	Non cancelled group ($n = 60$)	*P*
Age (years)	35.1 ± 1.0	35.0 ± 0.4	NS
BMI (kg/m^2)	24.2 ± 0.8	23.9 ± 0.4	NS
Infertility factor			NS
Male factor, *n* (%)	11 (55)	33 (55)	
Tubal factor, *n* (%)	4 (20)	12 (20)	
Endometriosis, *n* (%)	3 (15)	9 (15)	
Unexplained, *n* (%)	2 (10)	6 (10)	
Basal FSH (IU/l)	9.2 ± 0.8	8.8 ± 0.2	NS
Basal LH (IU/l)	5.7 ± 0.7	5.8 ± 0.2	NS
Basal estradiol (pg/ml)	45.9 ± 6.9	37.4 ± 2.0	NS
Basal AMH (pmol/l) (range)	20.7 ± 4.4 (1.0–80)	29.9 ± 2.3 (3.1–87)	<0.05
Day 5 AMH (pmol/l) (range)	14.2 ± 4.3 (1.4–87)	25.1 ± 2.1 (3.1–69)	<0.001
Days of stimulation, *n*	12.2 ± 0.1	11.6 ± 0.2	NS
Total dose of gonadotropins, IU	3022 ± 132	2905.3 ± 143	NS
Peak estradiol level (pg/ml)	341.5 ± 72.2	2102.7 ± 108.3	<0.001
No. of follicles > 10 mm	1.79 ± 0.1[a]	12.5 ± 0.7[b]	<0.001

Values are mean ± SEM; BMI, body mass index; NS, not significant.
[a]Cancellation day.
[b]Day of HCG administration.

respectively. All control patients underwent embryo transfer and the mean number of embryos per replacement and the mean number of high quality embryos replaced were similar for pregnant (2.51 ± 0.12 and 1.78 ± 0.16, respectively) and non-pregnant women (2.42 ± 0.13 and 1.62 ± 0.17, respectively).

To analyze the diagnostic accuracy of both basal and day 5 AMH and estradiol determinations to discriminate between cancelled versus control cycles and pregnancy versus non-pregnancy cycles, the AUC_{ROC} values determined with ROC analysis for each hormone measurement are shown in Table II. The AUC_{ROC} for day 5 AMH in predicting the likelihood of cancellation in an assisted reproduction treatment programme was significantly higher than that for basal AMH measurement. The ROC curve analysis was used to determine the best threshold values for day 5 AMH serum concentrations in predicting ovarian response (Figure 1). The best criterion value discriminating between control and cancelled cycles was ≧4.9 pmol/l (sensitivity 53%, specificity 96%). Diagnostic accuracy of estradiol serum concentration measured on stimulation day 5 to discriminate between cancellation and no cancellation was similar to that afforded by day 5 AMH measurement (Table II). When the likelihood of pregnancy was analyzed, the AUC_{ROC} for basal AMH was similar to that for day 5 AMH, and none of them was useful in their prediction of the reproductive outcome (Table II; Figure 1). Similarly, basal and day 5 estradiol serum levels were not discriminatory for pregnancy and non-pregnancy cycles.

Discussion

Assisted reproduction is expensive, time-consuming and stressful for patients. Evaluations of IVF/ICSI performance rarely consider cancelled cycles, which usually result from an inadequate ovarian response to the stimulation treatment. The cycle cancellations further increase the cost and duration of therapy. Therefore, a major challenge to the IVF teams is to predict prospective patients who will be low responders and to appropriately counsel women who are potential candidates for assisted reproduction.

Traditional methodology used to assess ovarian reserve has consisted of baseline serum levels of hormones such as FSH, estradiol and inhibins, and chronological age (Scott and

Table II. Diagnostic accuracy of basal and day 5 AMH and estradiol measurements to discriminate between cancellation vs no cancellation and pregnancy vs non-pregnancy in assisted reproduction treatment cycles using ROC plots

Hormone	Sensitivity, % (95% CI)	Specificity, % (95% CI)	AUC_{ROC} (95% CI)
Cancellation			
Basal AMH	40.0 (19.2–63.9)	91.7 (81.6–97.2)	0.67 (0.56–0.77)[a]
Day 5 AMH	53.0 (27.2–72.8)	96.0 (87.2–99.1)	0.81 (0.69–0.87)[a]
Basal estradiol	50.1 (27.2–72.8)	73.3 (60.3–83.9)	0.45 (0.28–0.63)
Day 5 estradiol	80.0 (56.3–94.1)	76.7 (64.0–86.6)	0.82 (0.71–0.93)
Pregnancy			
Basal AMH	62.5 (40.6- 81.2)	55.2 (35.7–73.5)	0.55 (0.41–0.69)
Day 5 AMH	66.7 (44.7–84.3)	48.3 (29.5–67.5)	0.50 (0.36–0.64)
Basal estradiol	100.0 (85.6–100)	19.6 (10.2–32.4)	0.50 (0.35–0.66)
Day 5 estradiol	87.5 (67.6–97.2)	48.2 (34.7–62.0)	0.42 (0.26–0.58)

CI, confidence interval; ROC, receiver operating characteristic; AUC, area under the curve.
[a]Day 5 AMH is statistically better than basal AMH ($P < 0.05$).

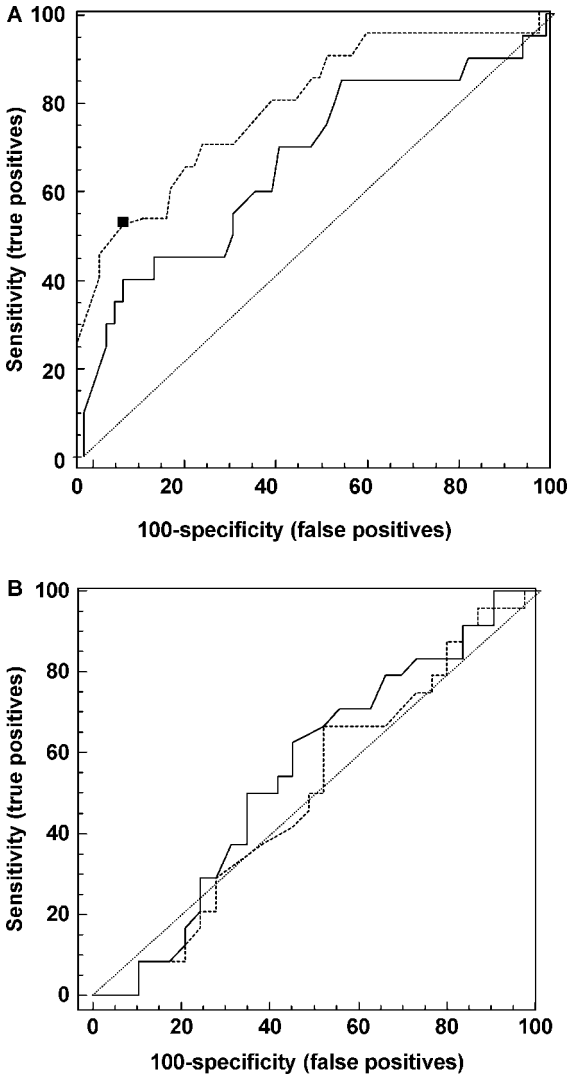

Figure 1. Receiver operating characteristic curves analysing the value of basal (solid line) and day 5 (dashed line) serum AMH concentrations for discriminating (A) cancelled vs non-cancelled cycles, and (B) conception versus non-conception cycles. The best cut-off value ■ discriminating between cancelled and control cycles for day 5 AMH was $\geqq 4.9$ pmol/l. The diagonal line is the line of no discrimination ($AUC_{ROC} = 0.5$).

Hofmann, 1995; Sharara and Scott, 1997; Karande and Gleicher, 1999; Bukman and Heineman, 2001). Also, a number of provocative tests have been devised to indirectly assess ovarian reserve and identify patients who might not be detected by basal hormone screening alone (Scott and Hofmann, 1995; Sharara and Scott, 1997; Bukman and Heineman, 2001). However, neither basal hormone measurements nor such dynamic tests provide direct information concerning the responsiveness of the ovaries to exogenous gonadotropins used in ovarian stimulation for assisted reproductive treatment. Thus, despite the validity of all these tests, there still remain patients who respond poorly to stimulation despite having normal tests of ovarian reserve. This supports the idea that ovarian reserve is not a simple static anatomic number of follicles but rather a dynamic process, the mechanism of which is not yet fully understood (Lass, 2001). In fact, it has been recently stressed that the ideal ovarian reserve test is the response of the ovaries to a 'normal' or 'standard' ovarian stimulation protocol (Karande and Gleicher, 1999; Tarlatzis *et al.*, 2003). Therefore, an early marker of ovarian responsiveness after the initiation of gonadotropin therapy would assist in deciding whether to proceed with an ongoing cycle. Ultimately, this will decrease the cost of continued monitoring and medication for patients in whom therapy will most likely fail.

The current study shows for the first time that AMH serum concentration obtained in the early follicular phase during ovarian stimulation with gonadotropins under pituitary suppression for assisted reproductive treatment is a better predictor of cancelled cycle than basal AMH. Thus, this report adds new data to recent work proposing cycle day 3 AMH measurement as a new marker for ovarian ageing and poor ovarian response to gonadotropin therapy in assisted reproduction cycles (de Vet *et al.*, 2002; Seifer *et al.*, 2002; van Roij et al., 2002; Fanchin *et al.*, 2003a,b) and supports the idea that dynamic tests seem better predictors of ovarian response than basal testing (Bukman and Heineman, 2001). A feature of the present investigation is that patients were matched for age, BMI, basal FSH and indication for assisted reproductive treatment. It has been reported that serum levels of AMH in normo-ovulatory women decrease over time and decrease with advancing age before changes occur in currently known aging-related variables (de Vet *et al.*, 2002). Both BMI and cause of infertility may influence ovarian response to gonadotropins (Crosignani *et al.*, 1994; Roseboom *et al.*, 1995; Tinkanen *et al.*, 1999; Loh *et al.*, 2002) and it has been reported that AMH serum levels tend to be lower in obese than in nonobese women (Pigny *et al.*, 2003). Basal serum FSH levels are an indication of biological age (Sharara and Scott, 1997) and FSH may be involved in the ontogenesis of AMH. However, data available in the literature in the latter respect are contradictory. Thus, it has been reported that FSH may down-regulate the AMH and AMH type II receptor expression in adult rat ovaries (Baarends *et al.*, 1995). Conversely, follicles from AMH knockout mice have been shown to be more sensitive to FSH than those from the wild type (Durlinger *et al.*, 2001). On the other hand, there are studies showing that AMH may foster FSH-induced follicular growth (McGee *et al.*, 2001) whereas it is well established that FSH is a positive regulator of testicular AMH gene expression in adults (Lukas-Croisier *et al.*, 2003). Therefore, our study design included matching for these variables, allowing the analysis of AMH as an independent marker of ovarian response. Although cohort studies can be more efficient than case–control studies, the latter can be accomplished in a shorter period of time, mainly when patients matching by pre-established criteria are used (Cramer, 1994; Schulz and Grimes, 2002). Performing IVF studies within a short time frame clearly contributes to precluding any bias due to possible changes in IVF/ICSI laboratory techniques. In this regard, a cohort study controlling for those well pre-established potential confounding variables would require too long a period of follow-up.

AMH serum levels obtained after 4 days of gonadotropin treatment were significantly lower than baseline AMH concentrations in both groups of patients studied. This is in agreement with recent work showing that serum AMH levels decline gradually during multiple follicular maturation, probably reflecting the dramatic reduction in the number of small antral follicles due to controlled ovarian hyperstimulation (Fanchin *et al.*, 2003b). These data suggest that AMH is preferentially secreted by small antral follicles and provide support to the hypothesis that differentiation of granulosa cells during follicular growth is likely to alter their ability of expressing AMH (Baarends *et al.*, 1995; Fanchin *et al.*, 2003b).

The question as to why AMH serum levels on stimulation day 5 would be a better predictor of ovarian response than basal AMH remains to be established for several reasons. First, the reduction in AMH serum levels observed during ovarian stimulation may be due to several factors such as a negative role of exogenous FSH administration or the supraphysiological increase in estradiol levels (La Marca *et al.*, 2004). Moreover, as stated above, the decrease in AMH may also be the result of a gonadotropin follicular growth stimulation. Second, it remains unclear what the relative contributions of the primordial pool and the preantral and early antral follicles may have in determining the serum concentrations of AMH, and the question whether AMH expression is lost in the follicles that are selected for dominance remains unanswered (Seifer *et al.*, 2002; Weenen *et al.*, 2004). Finally, uncertainties persist with respect to the control of granulosa cell AMH production and the possible effects of controlled ovarian hyperstimulation with exogenous gonadotropins on peripheral AMH levels (Fanchin *et al.*, 2003a). However, even though it is still questioned whether AMH is a marker of primordial follicles or later stages of follicle development or both, its serum level appears as a reliable marker of the ovarian follicle pool (Durlinger *et al.*, 2002; Pigny *et al.*, 2003). In addition, it is also possible that AMH function involves a role in selection of follicles that do not undergo atresia, and in preventing premature follicle maturation (Baarends *et al.*, 1995). Thus, it may be postulated that an ovarian dynamic test in the form of AMH detemination during gonadotropin treatment in the index cycle and reflecting the glandular response to stimulation may be a better marker of the overall follicular pool and activity of the ovary and AMH-related functions than basal AMH.

There are several potential limitations to the use of AMH as a marker of assisted reproduction treatment outcome. First, a wide range of serum AMH concentrations was found in the whole population studied both at baseline and on day 5 of gonadotropin therapy. This is in keeping with previous studies where the serum AMH concentration in normal prepubertal girls and normal adult women ranged from 0.7 to 73.9 pmol/l and 0.7 to 74.7 pmol/l, respectively (Long *et al.*, 2000). The wide range in values obtained for AMH would explain the low sensitivity (53%) of the best criterion value discriminating between control and cancelled cycles obtained in the present study. It could be argued that cancelled patients developing 1–2 follicles in response to prolonged gonadotropin treatment (late cancellation group) could still have a chance for pregnancy and thus results could be different if other cancellation criteria were used. However, from a practical point of view, it is considered that four oocytes are needed to reach an average of two embryos available for transfer (Bancsi *et al.*, 2002; Klinkert *et al.*, 2004). Thus, the collection of <4 oocytes at retrieval or cancellation of the cycle due to insufficient follicular growth (<3 developing follicles in response to exogenous gonadotropins) is the most widely used definition of poor response and it was used in many other studies because of the poor prognosis in such cases (Hanoch *et al.*, 1998; Hugues and Cedrin-Durnerin, 1998; Surrey *et al.*, 1998; De Placido *et al.*, 2000; García-Velasco et al., 2000; Bancsi *et al.*, 2002; El Toukhy *et al.*, 2002; Klinkert *et al.*, 2004). Furthermore, in the current study, basal and day 5 AMH serum levels were similar in the early cancellation and late cancellation groups.

Second, while cycle day 3 FSH measurement (which is the most routinely used ovarian reserve test) has clinical relevance, is inexpensive for the community and is not time-consuming for the medical team and so is easily repeatable, AMH assay is technically challenging and not readily available. Thus, there is as yet no international assay standard for AMH, which may explain discordance between different studies and makes comparison of results between laboratories difficult. In addition, the results of the current study indicate that stimulation day 5 estradiol serum level has similar predictive properties for ovarian performance in assisted reproductive treatment cycles as day 5 AMH determination. Third, as previously stressed, the pregnancy rate is the only important end-point of an ovarian reserve test (Bukman and Heineman, 2001). In this regard it is noteworthy that neither basal nor day 5 AMH or estradiol serum measurements were able to predict pregnancy in our study. This is in agreement with previous work by us and others stressing that hormone measurements may be helpful to evaluate ovarian response to stimulation but are less useful in their prediction of reproductive outcome (Commenges-Ducos *et al.*, 1998; Sharif *et al.*, 1998; Hall *et al.*, 1999; Creus *et al.*, 2000; Fábregues et al., 2000; Peñarrubia *et al.*, 2000; van Rooij *et al.*, 2003).

Finally, antral follicle counts were not considered in the present investigation but it has been reported that the number of antral follicles as counted early in the follicular phase provides better prognostic information on the occurrence of poor response during hormone stimulation for IVF than does the patient's chronological age and the currently used endocrine markers (Scheffer *et al.*, 1999; Bancsi *et al.*, 2002, Bancsi *et al.*, 2004). In addition, recent reports found a tight relationship between the AMH serum level and the antral follicular count assessed by ultrasonography in regularly menstruating infertile women studied at baseline day 3, before undergoing assisted reproduction (Van Rooij *et al.*, 2002; Fanchin *et al.*, 2003a). Therefore, the antral follicular count before ovarian stimulation may yield the same information and will better prevent unnecessary costs and effort.

In conclusion, AHM serum concentrations in the fifth day of gonadotropin therapy in women undergoing ovarian stimulation under pituitary desensitization for assisted reproduction

have a higher predictive value of ovarian response than basal AMH. This early assessment of ovarian stimulation reponse could help to decide early cancellation, thus avoiding further cost and therapy. However, the predictive capacity of day 5 AMH was not better than that provided by day 5 estradiol. Furthermore, neither basal nor day 5 AMH or estradiol measurements were useful in the prediction of pregnancy after assisted reproductive treatment. Therefore, definite clinical applicability of AMH determination as a marker of IVF outcome remains to be established.

Acknowledgements

The authors thank Mrs Paquita Antonell for technical assistance. This work was supported in part by grants from the Instituto de Salud Carlos III (RCMN C03/08) and the Comissionat per a Universitat i Recerca–Generalitat de Catalunya (2001SGR 00372) to J.B.

References

Baarends WM, Uilenbroek JT, Kramer P, Hoogerbrugge JW, van Leeuwen EC, Themmen AP and Grootegoed JA (1995) Anti-müllerian hormone and anti-müllerian hormone type II receptor messenger ribonucleic acid expression in rat ovaries during postnatal development, the estrous cycle and gonadotropin-induced follicle growth. Endocrinology 136,4951–4962.

Balasch J, Creus M, Fábregues F, Carmona F, Casamitjana R, Ascaso C and Vanrell JA (1996) Inhibin, follicle-stimulating hormone, and age as predictors of ovarian response in in vitro fertilization cycles stimulated with gonadotropin-releasing hormone agonist-gonadotropin treatment. Am J Obstet Gynecol 175,1226–1230.

Balasch J, Vidal E, Peñarrubia J, Casamitjana R, Carmona F, Creus M, Fábregues F and Vanrell JA (2001) Suppression of LH during ovarian stimulation: analysing threshold values and effects on ovarian response and the outcome of assisted reproduction in down-regulated women stimulated with recombinant FSH. Hum Reprod 16,1636–1643.

Bancsi LF, Broekmans FJ, Eijkemans MJ, de Jong FH, Habbema JD and te Velde ER (2002) Predictors of poor ovarian response in in vitro fertilization: a prospective study comparing basal markers of ovarian reserve. Fertil Steril 77,328–336.

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

Beckers NG, Macklon NS, Eijkemans MJ, Fauser BC (2002) Women with regular menstrual cycles and a poor response to ovarian hyperstimulation for in vitro fertilization exhibit follicular phase characteristics suggestive of ovarian aging. Fertil Steril 78,291–297.

Bukman A and Heineman MJ (2001) Ovarian reserve testing and the use of prognostic models in patients with subfertility. Hum Reprod Update 7, 581–590.

Commenges-Ducos M, Tricaud S, Papaxanthos-Roche A, Dallay D, Horovitz J and Commenges D (1998) Modelling of the probability of success of the stages of in-vitro fertilization and embryo transfer: stimulation, fertilization and implantation. Hum Reprod 13,78–83.

Cramer DW (1994) Case–control studies of infertility. Infertil Reprod Med Clin N Am 5,321–335.

Creus M, Penarrubia J, Fabregues F, Vidal E, Carmona F, Casamitjana R, Vanrell JA and Balasch J (2000) Day 3 serum inhibin B and FSH and age as predictors of assisted reproduction treatment outcome. Hum Reprod 15,2341–2346.

Crosignani PG, Ragni G, Parazzini F, Wyssling H, Lombroso G and Perotti L (1994) Anthropometric indicators and response to gonadotrophin for ovulation induction. Hum Reprod 9,420–423.

De Placido G, Alviggi C, Mollo A, Strina I, Varricchio MT and Molis M (2000) Recombinant follicle stimulating hormone is effective in poor responders to highly purified follicle stimulating hormone. Hum Reprod 15,17–20.

de Vet A, Laven JSE, de Jong FH, Themmen APN and Fauser BCJM (2002) Antimüllerian hormone serum levels: a putative maker for ovarian aging. Fertil Steril 77,357–362.

Durlinger AL, Gruijters MJ, Kramer P et al. (2001) Anti-Müllerian hormone attenuates the effects of FSH on follicle development in the mouse ovary. Endocrinology 142,4891–4899.

Durlinger ALL, Visser JA and Themmen APN (2002) Regulation of ovarian function: the role of anti-Müllerian hormone. Reproduction 124,601–609.

El Toukhy T, Khalaf Y, Hart R, Taylor A and Braude P (2002) Young age does not protect against the adverse effects of reduced ovarian reserve—an eight year study. Hum Reprod 17,1519–1524.

Eldar-Geva T, Robertson DM, Cahir N, Groome N, Gabbe MP, Maclachlan V and Healy DL (2000) Relationship between serum inhibin A and B and ovarian follicle development after a daily fixed dose administration of recombinant follicle-stimulating hormone. J Clin Endocrinol Metab 85, 607–613.

Fabregues F, Balasch J, Creus M, Carmona F, Puerto B, Quinto L, Casamitjana R and Vanrell JA (2000) Ovarian reserve test with human menopausal gonadotropin as a predictor of in vitro fertilization outcome. J Assist Reprod Genet 17,13–19.

Fanchin R, Schonaüer LM, Righini C, Guilbourdenche J, Frydman R and Taieb J (2003a) Serum anti-Müllerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. Hum Reprod 18,323–327.

Fanchin R, Schonaüer LM, Righini C, Frydman N, Frydman R and Taieb J (2003b) Serum anti-Müllerian hormone dynamics during controlled ovarian hyperstimulation. Hum Reprod 18,328–332.

Farhi J, Homburg R, Ferber A, Orvieto R and Ben-Rafael Z (1997) Non-response to ovarian stimulation in normogonadotrophic, normogonadal women: a clinical sign of impending onset of ovarian failure pre-empting the rise in basal follicle stimulating hormone levels. Hum Reprod 12, 241–243.

Fawzy M, Lambert A, Harrison RF, Knight PG, Groome N, Hennelly B and Robertson WR (2002) Day 5 inhibin B levels in a treatment cycle are predictive of IVF outcome. Hum Reprod 17,1535–1543.

Garcia-Velasco JA, Isaza V, Requena A, Martinez-Salazar FJ, Landazábal A, Remohí J, Pellicer A and Simón C (2000) High dose of gonadotrophins combined with stop versus non-stop protocol of GnRH analogue administration in low responder IVF patients: a prospective, randomized, controlled trial. Hum Reprod 15,2292–2296.

Gruijters MJG, Visser JA, Durlinger ALL and Themmen APN (2003) Anti-Müllerian hormone and its role in ovarian function. Mol Cell Endocrinol 211,85–90.

Hall JE, Welt CK and Cramer DW (1999) Inhibin A and inhibin B reflect ovarian function in assisted reproduction but are less useful at predicting outcome. Hum Reprod 14,409–415.

Hanley JA and McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143,29–36.

Hanoch J, Lavy Y, Holzer H, Hurwitz A, Simon A, Revel A and Laufer N (1998) Young low responders protected from untoward effects of reduced ovarian response. Fertil Steril 69,1001–1004.

Hugues JN and Cedrin Durnerin I (1998) Revisiting gonadotrophin-releasing hormone agonist protocols and management of poor ovarian responses to gonadotrophins. Hum Reprod Update 4,83–101.

Karande V and Gleicher N (1999) A rational approach to the management of low responders in in-vitro fertilization. Hum Reprod 14,1744–1748.

Klinkert ER, Broekmans FJM, Looman CWN and te Velde ER (2004) A poor response in the first in vitro fertilization cycles is not necessarily related to a poor prognosis in subsequent cycles. Fertil Steril 81, 1247–1253.

La Marca A, Malmusi S, Giulini S, Tamaro LF, Orvieto R, Levratti P, Volpe A (2004) Anti-Müllerian hormone plasma levels in spontaneous menstrual cycle and during treatment with FSH to induce ovulation. Hum Reprod 19,2738–2741.

Laas A (2001) Assessment of ovarian reserve–is there a role for ovarian biopsy? Hum Reprod 16,1055–1057.

Lee MM, Donahoe PK, Hasegawa T, Silverman B, Crist GB, Best S, Hasegawa Y, Noto RA, Schoenfeld D and MacLaughlin DT (1996) Müllerian inhibiting substance in humans: normal levels from infancy to adulthood. J Clin Endocrinol Metab 81,571–576.

Loh S, Wang JX and Matthews CD (2002) The influence of body mass index, basal FSH and age on the response to gonadotrophin stimulation in non-polycystic ovarian syndrome patients. Hum Reprod 17,1207–1211.

Long W, Ranchin V, Pautier P, Belville C, Denizto P, Cailla H, Lhommé C, Picard JY, Bidart JM and Rey R (2000) Detection of minimal levels of serum anti-Müllerian hormone during follow-up of patients with ovarian

granulosa cells tumor by means of a highly sensitive enzyme-linked immunosorbent assay. J Clin Endocrinol Metab 85,540–544.

Lukas-Croisier C, Lasala C, Nicaud J, Bedecarras P, Kumar TR, Dutertre M, Matzuk MM, Picard JY, Josso N and Rey R (2003) Follicle-stimulating hormone increases testicular Anti-Mullerian hormone (AMH) production through Sertoli cell proliferation and a nonclassical cyclic adenosine 5′-monophosphate-mediated activation of the AMH Gene. Mol Endocrinol 17,550–561.

McGee EA, Smith R, Spears N, Nachtigal MW, Ingraham H and Hsueh AJ (2001) Mullerian inhibitory substance induced growth of rat preantral ovarian follicles. Biol Reprod 64,293–298.

Peñarrubia J, Balasch J, Fábregues F, Carmona F, Casamitjana R, Moreno V, Calafell JM and Vanrell JA (2000) Day 5 inhibin B serum concentrations as predictors of assisted reproduction technology outcome in cycles stimulated with gonadotrophin-releasing hormone agonist-gonadotrophin treatment. Hum Reprod 15,1499–1504.

Peñarrubia J, Fábregues F, Creus M, Manau D, Casamitjana R, Guimerá M, Carmona F, Vanrell JA and Balasch J (2003) LH serum levels during ovarian stimulation as predictors of ovarian response and assisted reproduction outcome in down-regulated women stimulated with recombinant FSH. Hum Reprod 18,2689–2697.

Phelps JY, Levine AS, Hickman TN, Zacur HA, Wallach EE, Hinton EL (1998) Day 4 estradiol levels predict pregnancy success in women undergoing controlled ovarian hyperstimulation for IVF. Fertil Steril 69,1015–1019.

Pigny P, Merlen E, Robert Y, Cortet-Rudelli, Decanter C, Jonard S and Dewailly D (2003) Elevated serum level of anti-Mullerian hormone in patients with polycystic ovary syndrome: relationship to the ovarian follicle excess and to the follicular arrest. J Clin Endocrinol Metab 88, 5957–5962.

Roseboom TJ, Vermeiden JPW, Schoute E, Lens JW and Schats R (1995) The probability of pregnacy after embryo transfer is affected by the age of the patient, cause of infertility, number of embryos trasnsferred and the average morphology score, as revealed by multiple regression analysis. Hum Reprod 10,3035–3041.

Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW and te Velde ER (1999) Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril 72,845–851.

Schulz K and Grimes DA (2002) Case-control studies: research in reverse. Lancet 360,259–260.

Scott RT and Hofmann GE (1995) Prognostic assessment of ovarian reserve. Fertil Steril 63,1–11.

Seifer DB, MacLaughlin DT, Christian BP, Feng B and Shelden RM (2002) Early follicular serum müllerian-inhbiting substance levels are associated with ovarian response during assisted reproductive technology cycles. Fertil Steril 77,468–471.

Sharara FI and Scott RT (1997) Assessment of ovarian reserve and treatment of low responders. Infertil Reprod Med Clin N Am 8,501–522.

Sharif K, Elgendy M, Lashen H and Afnan M (1998) Age and basal follicle stimulating hormone as predictors of in vitro fertilisation outcome. Br J Obstet Gynaecol 105,107–112.

Surrey ES, Bower J, Hill DM, Ramsey J and Surrey MW (1998) Clinical and endocrine effects of a microdose GnRH agonist flare regimen administered to poor responders who are undergoing in vitro fertilization. Fertil Steril 69,419–424.

Tarlatzis BC, Zepiridis L, Grimbizis G and Bontis J (2003) Clinical management of low ovarian response to stimulation for IVF: a systematic review. Hum Reprod Update 9,61–76.

Tinkanen H, Bläuer M, Laippala P, Tuohimaa P and Kujansuu E (1999) Prognostic factors in controlled ovarian hyperstimulation. Fertil Steril 72, 932–936.

van Rooij IAJ, Broekmans FJM, te Velde ER, Fauser BCJM, Bancsi LFJMM, de Jong FH and Themmen APN (2002) Serum anti-Müllerian hormone levels: a novel measure of ovarian reserve. Hum Reprod 17, 3065–3071.

van Rooij IA, Bancsi LF, Broekmans FJ, Looman CW, Habbema JD and te Velde ER (2003) 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 79,482–488.

Veeck LL (1999) An Atlas of Human Gametes and Conceptuses: An Illustrated Reference for Assisted Reproductive Technology. CRC Press-Parthenon Publishers, New York, 1999.

Wallach EE (1995) Pitfalls in evaluating ovarian reserve. Fertil Steril 63, 13–14.

Weenen C, Laven JSE, von Bergh ARM, Cranfiels M, Groome NP, Visser JA, Kramer P, Fauser BCJM and Themmen APN (2004) Anti-Müllerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. Mol Hum Reprod 10,77–83.

Zweig MH and Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 39,561–577.

Submitted on October 5, 2004; accepted on December 7, 2004