Serum anti-Müllerian hormone levels: a novel measure of ovarian reserve


1Department of Reproductive Medicine, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, 2Department of Internal Medicine and 3Center of Reproductive Medicine, Erasmus MC, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands
4To whom correspondence should be addressed. E-mail: i.vanrooij@azu.nl

BACKGROUND: Anti-Müllerian hormone (AMH) is produced by the granulosa cells of preantral and small antral follicles and its levels can be assessed in serum. Since the number of ovarian follicles declines with increasing age, AMH levels might be used as a marker for ovarian ageing. Therefore, we studied the relationship between AMH levels and ovarian response during ovarian stimulation for IVF.

METHODS: A total of 130 patients undergoing their first IVF treatment cycle using a long protocol with GnRH agonist was prospectively included. Blood withdrawal was performed and the number of antral follicles was assessed by ultrasound on day 3 of a spontaneous cycle. Poor response and the number of oocytes were used as primary outcome measures. In a random subset of 23 patients a GnRH agonist stimulation test was performed to investigate whether a rise in FSH and LH would affect AMH levels. Blood withdrawal was performed and the number of antral follicles was assessed by ultrasound on day 3 of a spontaneous cycle. Poor response and the number of oocytes were used as primary outcome measures. In a random subset of 23 patients a GnRH agonist stimulation test was performed to investigate whether a rise in FSH and LH would affect AMH levels. RESULTS: The data of 119 patients were analysed. Serum AMH levels were highly correlated with the number of antral follicles (\(r = 0.77, P < 0.01\)) and the number of oocytes retrieved (\(r = 0.57, P < 0.01\)). A negative association was found between AMH levels and poor ovarian response (fewer than 4 oocytes or cycle cancellation; OR 0.82, 95% CI 0.75–0.90, \(P < 0.01\)). Inclusion of inhibin B and FSH concentrations to AMH in a multivariate model improved the prediction of ovarian response. The post GnRH agonist rise in FSH and LH levels did not influence AMH values. CONCLUSIONS: Poor response in IVF, indicative of a diminished ovarian reserve, is associated with reduced baseline serum AMH concentrations. In line with recent observations it appears that AMH can be used as a marker for ovarian ageing.

Key words: anti-Müllerian hormone/antral follicle count/IVF/ovarian reserve/poor response

Introduction

Anti-Müllerian hormone (AMH), a member of the transforming growth factor-β (TGF-β) family, was identified as a factor that causes regression of the Müllerian ducts during male fetal development (Jost, 1947; Behringer et al., 1994). In females, AMH, also known as Müllerian inhibiting substance, is produced in the granulosa cells of ovarian follicles (Vigier et al., 1984). Expression of AMH mRNA was detected in granulosa cells of primary follicles immediately after their formation in neonatal rats and mice, and subsequently in granulosa cells of all secondary preantral stage follicles, small antral follicles during the first prepubertal wave of development, and during estrus cycling thereafter. AMH expression starts to fade from the small antral follicle stage onwards (Baarends et al., 1995). This pattern of AMH mRNA and protein expression in follicles can also be observed in the human ovary (Rey et al., 2000). AMH may be functionally active in the female gonad (Lee and Donahoe, 1993; for review), affecting the transition from resting primordial follicles into growing follicles (Durlinger et al., 1999). Furthermore, AMH may be involved in the recruitment of FSH-sensitive follicles in the early antral stage (Durlinger et al., 2001). Notwithstanding the involvement of several transcription factors, such as SF1, Dax1 and GATA4 in the regulation of AMH expression, the hormonal regulation of AMH synthesis in the adult female gonad remains unclear (Nachtwal et al., 1998; Arango et al., 1999; Tremblay et al., 2001). No direct effect of FSH on AMH mRNA expression was seen in cultured human adult granulosa cells retrieved after gonadotropin stimulation (Voutilainen and Miller, 1987).

Human female serum contains measurable amounts of AMH during the reproductive life span (Lee et al., 1996). Since AMH is solely produced in the growing ovarian follicles, serum levels may be used as a marker for ovarian reserve, representing the quantity and quality of the ovarian follicle pool (te Velde and Pearson, 2002). Recent preliminary reports indeed indicate that AMH levels decline with increasing female age (de Vet et al., 2002) and that initial AMH is associated with ovarian response in IVF patients with normal FSH levels (Seifer et al., 2002).
In the present study we prospectively assessed the significance of AMH as a marker for ovarian response in a large unselected IVF population. In addition, the predictive performance of serum AMH levels towards poor response in relation to other ovarian reserve tests was investigated. Finally, as little is known concerning the regulation of AMH production in the human, we also investigated whether serum AMH levels are affected by a rise in endogenous FSH and LH induced by a single, high dose GnRH agonist administration.

Subjects and methods
A total of 130 patients who were going to have their first IVF treatment, was prospectively included according to the following criteria: (i) regular menstrual cycles (25–35 days), (ii) presence of both ovaries, (iii) no evidence of endocrine disorders (normal thyroid stimulating hormone, prolactin, testosterone and androstenedione), (iv) age ≤ 46 years, and (v) written informed consent. The Institutional Review Board approved this study. Of the 130 patients included, 112 were scheduled for conventional IVF, whereas the remaining 18 patients were scheduled for ICSI.

On day 3 of a spontaneous cycle within the 3 months preceding IVF treatment, patients underwent a transvaginal ultrasound examination to assess the number of antral follicles, measuring 2–5 mm, as described previously (Bancsi et al., 2002). On the same day a venous blood sample was obtained for the measurement of AMH, FSH, estradiol (E2) and inhibin B. Serum and plasma samples were centrifuged at 1700 g within 2 h and stored at −20°C until assayed.

In a subset of 23 patients a GnRH agonist stimulation test (GAST) was performed. These patients received 0.1 mg triptorelin (Decapeptyl®, Ferring, Hoofddorp, The Netherlands) s.c. directly after blood sampling on cycle day 3 and returned exactly 24 h later for a second blood sampling for measurement of AMH, FSH, E2 and inhibin B.

FSH and E2 were assessed in plasma with the AxSYM immunoanalyser (Abbott Laboratories, Abbott Park, IL, USA). The World Health Organization Second International Reference Preparation for human FSH (78/549) was used as a standard in the FSH assay. For FSH, interassay coefficients of variation were found to be 6.0, 6.6 and 8.0% at the levels of 5, 25 and 75 IU/l respectively (n = 46). The E2 assay is standardized to gas chromatography/mass spectrometry. Interassay variation of the E2 assay at 300, 1105 and 2626 pmol/l was 12.5, 7.5 and 4.9% (n = 29) respectively. Serum inhibin B levels were measured using an immuno-enzymometric assay (Serotec, Oxford, UK) as described by Groome et al. (Groome et al., 1996). Intra- and interassay coefficients of variation were <14.6 and <14.0% respectively. An ultra-sensitive immuno-enzymometric assay kit (Immunotech-Coulter, Marseille, France) was used for the estimation of AMH as described elsewhere (Long et al., 2000). The limit of detection (defined as blank + 3 SD of blank) was 0.05 μg/l. Intra- and interassay coefficients of variation were <5 and 8% respectively. If a GAST was performed both samples were analysed in the same run.

The IVF treatment has been described in detail in a previous publication (van Kooij et al., 1996). In brief, patients started with leuprolide acetate (Lucrin®, Abbott, Hoofddorp, The Netherlands) in the midluteal phase to achieve pituitary desensitisation. After menstruation, the ovarian hyperstimulation started with a fixed dose protocol of 150 IU follitropin alpha (rFSH, Gonal-F®, Serono Benelux BV, The Hague, The Netherlands). After 7 days follicular growth was assessed by ultrasound and E2 measurement. If necessary the dose of rFSH was adjusted. When at least three leading follicles developed, 10 000 IU hCG, (Profasi®, Serono Benelux) was administered and 36 h later oocyte retrieval was performed. A maximum of two embryos was transferred in women <38 years of age. Above this age a maximum of three embryos was transferred. To support the luteal phase, either hCG (Profasi®) or micronized progesterone (Progesteran®, Nourypharma BV, Oss, The Netherlands) was used.

The main outcome measures of the study were the number of oocytes retrieved and poor ovarian response. As described previously (Bancsi et al., 2002), poor response was defined as fewer than 4 oocytes at follicle puncture or as cancellation due to impaired (fewer than 3 follicles) or absent follicular growth in response to ovarian hyperstimulation. With a mean fertilization rate of 50–60% a minimum of four oocytes is necessary to transfer at least two embryos. In the analysis of poor ovarian response, the group of ‘normal’ responders also included patients with cancelled cycles due to an exaggerated response. Patients were considered high responders in case of collection of more than 20 oocytes at ovum retrieval or when the cycle was cancelled due to exaggerated response (more than 30 follicles in both ovaries and/or peak E2 >15 000 pmol/l). High response was considered a secondary outcome measure, and in the analysis of high response both the poor and normal responders are considered as one group.

Another secondary outcome measure was ongoing pregnancy, defined as a viable pregnancy assessed by ultrasound of at least 11 weeks gestation. Data from patients whose cycles were cancelled due to either risk of ovarian hyperstimulation syndrome (OHSS) or poor response (fewer than 3 follicles) to hormone stimulation were not included in the pregnancy analysis, because it cannot be excluded that such patients would have become pregnant if IVF were performed. However, patients with complete absence of follicle growth and E2 <200 pmol/l were considered to have a zero chance of pregnancy, and therefore data were on their cycles was included in the analysis of pregnancy.

Data were analysed with the Statistical Program for Social Sciences (SPSS Inc., Chicago, IL, USA). Values are presented as median and range. To compare normal with poor responders the Mann–Whitney test or χ²-test was performed whenever appropriate. The correlation between different parameters is expressed as Spearman’s correlation coefficient. Univariate and multivariate logistic regression with the main outcome measure poor response and secondary outcome measures of high response and ongoing pregnancy were performed. For each single variable used in the univariate analysis and for the models, the ability to discriminate between patients with a poor response and patients with a normal response was assessed by calculating the area under the receiver operating characteristics curves (ROC_AUC) (Harrell et al., 1996). The ROC_AUC may vary between 0.5 (no discriminative power) to 1.0 (perfect discrimination). For the comparison of E2, inhibin B and AMH before and after the GAST a Wilcoxon signed rank test was used. Statistical significance was considered to be reached at P-value <0.05.

Results
Of the 130 patients included, 10 dropped out before the IVF treatment. Six patients conceived spontaneously, two had a serious disease necessitating interruption of the study protocol and fertility treatment and two withdrew their consent. From one patient no AMH could be measured, because of lack of stored serum. Hence, 119 patients were included in the analysis, of which 101 underwent an IVF treatment and 18 an ICSI treatment.

Patient and ovarian reserve test characteristics of the complete group and of normal and poor responders separately are...
Serum AMH and ovarian reserve

Table I. Patient and ovarian reserve test characteristics in the total group of IVF patients, and in normal and poor responders separately

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total (n = 119)</th>
<th>Normal responders (n = 84)</th>
<th>Poor responders (n = 35)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>33.8 (22.3–44.0)</td>
<td>33.8 (24.4–44.0)</td>
<td>36.3 (22.3–43.3)</td>
<td>NS (0.09)a</td>
</tr>
<tr>
<td>Duration of infertility disorder (months)</td>
<td>30.0 (12.0–240.0)</td>
<td>30.0 (12.0–83.0)</td>
<td>30.0 (12.0–240.0)</td>
<td>NSa</td>
</tr>
<tr>
<td>Primary infertility (%)</td>
<td>65 (54.6)</td>
<td>48 (57.1)</td>
<td>17 (48.6)</td>
<td>NSb</td>
</tr>
<tr>
<td>Diagnosis of infertility</td>
<td>Tube pathology (%)</td>
<td>22 (18.5)</td>
<td>17 (20.2)</td>
<td>5 (14.3)</td>
</tr>
<tr>
<td></td>
<td>Male factor (%)</td>
<td>59 (49.6)</td>
<td>50 (59.5)</td>
<td>9 (25.7)</td>
</tr>
<tr>
<td></td>
<td>Unexplained</td>
<td>38 (31.9)</td>
<td>17 (20.2)</td>
<td>21 (60.0)</td>
</tr>
<tr>
<td></td>
<td>AMH (ug/l)</td>
<td>0.9 (0.0–6.2)</td>
<td>1.4 (0.0–6.2)</td>
<td>0.2 (0.0–1.7)</td>
</tr>
<tr>
<td></td>
<td>AFC (n)</td>
<td>8 (0–35)</td>
<td>11 (0–35)</td>
<td>4 (0–15)</td>
</tr>
<tr>
<td></td>
<td>FSH (IU/l)</td>
<td>6.6 (3.2–58.9)</td>
<td>6.0 (3.2–17.1)</td>
<td>10.5 (4.5–58.9)</td>
</tr>
<tr>
<td></td>
<td>E2 (pmol/l)</td>
<td>158 (41–1796)</td>
<td>160 (41–389)</td>
<td>158 (50–1796)</td>
</tr>
<tr>
<td></td>
<td>Inhibin B (ng/l)</td>
<td>103 (0–304)</td>
<td>115 (29–304)</td>
<td>73 (0–155)</td>
</tr>
<tr>
<td></td>
<td>Number of oocytes (n = 96)c</td>
<td>7 (1–28)</td>
<td>9 (4–28)</td>
<td>2 (1–3)</td>
</tr>
</tbody>
</table>

Values are presented as median (range) or as number (percentage). AFC = antral follicle count. N/A = not applicable.
aMann–Whitney test and bχ²-test are performed to compare normal and poor responders.
cNumber of patients is 77 in the normal responder group (seven cancellations due to high response) and 19 in the poor responder group (including 16 cancellations).

Presented in Table I. The 35 poor responders were somewhat older, and their AMH, FSH and inhibin B levels and antral follicle counts were statistically different from those in the 84 normal responders. As expected, patients with a poor response were more often treated for unexplained infertility. The number of oocytes retrieved in the normal and poor responders groups are also depicted in Table I. While seven patients in the normal response group had their ovum retrieval cancelled because of the risk of OHSS, no ovum retrieval took place in 16 poor responders because of insufficient follicle growth (0–2 follicles). Normal responders showed higher median peak E2 levels during ovarian stimulation compared with poor responders [6695 (2027–18 300) pmol/l versus 1354 (100–6570) pmol/l ; P < 0.001].

The correlation coefficients for the association between AMH levels on one hand and several ovarian reserve test variables and the total number of oocytes retrieved on the other hand are depicted in Figure 1. AMH was highly correlated with the number of antral follicles (AFC) and the number of oocytes retrieved after ovarian hyperstimulation. Besides AMH, only AFC was significantly correlated with chronological age (R = –0.29, P < 0.001). A good correlation was observed between AFC and the number of oocytes retrieved (R = 0.58, P < 0.001), comparable with that of AMH and number of oocytes.

In Table II the results of the logistic regression analysis for the prediction of poor response are given. AFC presented the highest ROC_AUC of 0.86, indicating a good discriminating potential for predicting poor ovarian response. The ROC_AUC for AMH was almost identical, followed by the AUC values for FSH and inhibin B. Age was not significantly related to poor response. Using a multivariate analysis AFC, inhibin B and FSH were selected. Due to the high correlation between AFC and AMH, AMH was not selected in the model. Leaving AFC out of the model led to the selection of AMH together with inhibin B and FSH. Both models showed a comparable discriminative potential towards poor response prediction.

When a similar analysis was performed with high response as the outcome measure, AFC and AMH were again the best performing variables. The AUC values were 0.89 and 0.88 for AFC and AMH respectively, in the univariate analysis. In the multivariate analysis AFC and inhibin B were selected, and as in the analysis of poor response, AFC could be exchanged for AMH, giving a similar performing model for the prediction of high response. In the analysis of ongoing pregnancy 106 patients could be included. None of the variables showed a statistically significant relationship with ongoing pregnancy.

To obtain further support for the notion that AMH is a measure of the ovarian follicle population and not of the gonadotrophic hormonal status of the patient, we performed a GAST in a subset of 23 patients. In these patients a significant rise of E2 and inhibin B values was observed as a response to endogenous FSH and LH elevation after s.c. injection of GnRH agonist, whereas no changes in AMH levels were seen (Table III).

Discussion

In this study we prospectively investigated whether serum AMH levels can predict ovarian response during first IVF treatment cycles. We observed a high correlation of AMH with ovarian response, as expressed by the number of oocytes retrieved. Ovarian response during exposure to high levels of gonadotrophins can be considered to be a measure of the selectable cohort of antral follicles. As this number of antral follicles appears to be related to the size of the primordial follicle pool (Gougeon, 1984), ovarian response can be regarded as a reflection of the ovarian reserve. Recent studies have shown that a low response to exogenous gonadotrophin stimulation is associated with an early menopause, supporting the idea that ovarian response indeed reflects the ovarian ageing process (de Boer et al., 2002; Nikolaou et al., 2002). The excellent correlation between initial AMH levels and subsequent ovarian response in IVF therefore implies that AMH is a promising marker for ovarian reserve.
Figure 1. Correlation of age (A), number of oocytes after ovum retrieval (B), cycle day 3 number of antral follicles (C), cycle day 3 E\textsubscript{2} (D), cycle day 3 inhibin B (E) and cycle day 3 FSH (F) with AMH in IVF patients. R is Spearman’s correlation coefficient followed by the P-value.

Ovarian reserve comprises two elements: the size of the stock of primordial follicles and the quality of the oocytes (te Velde and Pearson, 2002). From the primordial follicle pool, primary follicles will start a maturation process and develop through secondary (preantral) follicles into the pool of antral follicles from which the monthly follicle to be ovulated is selected (Fauser and Van Heusden, 1997). In rodents, AMH is produced immediately after the transition of primordial to primary follicles, and AMH expression disappears when follicles either are selected for ovulation or become atretic (Baarends et al., 1995). In the human ovary, AMH protein expression is also seen in granulosa cells of follicles from the primary stage up to the larger antral stage when follicles have gained FSH dependence (Rajpert-De Meyts et al., 1999). Since the size of the primordial follicle stock is difficult to measure directly, a marker that reflects all numbers...
of follicles that have made the transition from the primordial follicle pool to the growing pool may be a good indirect measurement. AMH might be such a marker, as it is involved in the regulation of primordial follicle recruitment (Durlinger et al., 1999), an important mechanism for the depletion of the primordial follicle pool. AFC is a direct ultrasound measure of this pool, whereas the initial inhibin B and E2 levels are considered to reflect, directly or indirectly, the size of the antral (2–5 mm) follicle pool. AFC is a direct ultrasound measure of this pool, whereas the initial inhibin B and E2 levels are considered to be greatly dependent on the number of antral follicles in the early menstrual days. FSH levels are regulated by a negative feedback action of these two granulosa cell products and hence are a more indirect reflection of antral follicle number.

The available tests for assessment of ovarian reserve all reflect, directly or indirectly, the size of the antral (2–5 mm) follicle pool. AFC is a direct ultrasound measure of this pool, whereas the initial inhibin B and E2 levels are considered to be greatly dependent on the number of antral follicles in the early menstrual days. FSH levels are regulated by a negative feedback action of these two granulosa cell products and hence are a more indirect reflection of antral follicle number.

We have recently shown that AFC gives the best prognostic information with regard to the occurrence of poor response in IVF (Bancsi et al., 2002). Moreover, the combination of the three ovarian reserve tests, AFC, inhibin B and FSH, in a multivariate logistic model appeared to improve the response prediction. In the present study AMH is found to have a predictive performance comparable with that of AFC. The
multivariate analysis on the prediction of poor ovarian response to hyperstimulation in the study revealed that AMH will contribute independently to this prediction but only if AFC is removed from the analysis. Obviously, the high interrelation between the follicle numbers in earlier stages and the antral follicle number prohibits any additional information to be obtained from this test. Both models need external validation, which is currently being performed. In the prediction of high response, comparable discriminative performance was found for the variables AFC and AMH, supporting the close relationship between these variables and ovarian response.

The present study confirms and extends results of a recent study where lower serum AMH levels were found in patients having six or fewer retrieved oocytes compared with patients having 11 or more oocytes (Seifer et al., 2002). We studied a group of unselected patients having their first IVF treatment with a fixed dose of gonadotrophins. Therefore we could compare AMH as a predictor of ovarian reserve with the other endocrinological markers and AFC.

There are advantages of the use of AMH over AFC in the multivariate model for the prediction of ovarian response, since all predictive information is obtained with blood sampling and no extra ultrasound is needed. Furthermore, since there is no change in AMH levels in response to gonadotrophins, AMH can be measured throughout the cycle in contrast to the other parameters, which can only be determined during the early follicle phase, an advantage for both patients and clinicians. Obviously, AMH intra-cycle and cycle-to-cycle variation should be further analysed, but the small fluctuation in serum AMH levels at three different time points during the menstrual cycle (Cook et al., 2000), supports the feasibility of AMH assessment throughout the cycle.

Application of AMH (this study), AFC (Bancsi et al., 2002), inhibin B (Seifer et al., 1997; Hall et al. 1999) and FSH (Sharif et al., 1998; Bancsi et al., 2000) as a predictor of ongoing pregnancy appears to be limited in view of the fact that they only represent the quantitative aspect of ovarian reserve, whereas pregnancy is also dependent on the oocyte quality. Furthermore, it is also possible for patients with normal ovarian reserve not to become pregnant, for instance as a result of fertilization failure. Nevertheless, the ability to predict poor response may be a valuable tool for patient counselling, since poor responders have a lower probability of pregnancy. A further potential application for the prediction of poor response is the augmentation of the starting dose of gonadotrophins in predicted poor responders. It is not certain that this may lead to higher pregnancy rates (Land et al., 1996), but randomized prospective data on this issue are still lacking. Also, patients normally denied IVF treatment because of advanced age might profit from response prediction. As patients >40 years of age have a better outcome in the case of a normal response (Roest et al., 1996), those patients with a normal response prediction can proceed to treatment.

In conclusion, we have found that AMH serum levels are associated with ovarian response in IVF patients and may serve as a novel marker for ovarian reserve. The predictive value of AMH for poor ovarian response is comparable with that of AFC and therefore AFC could be replaced by AMH in the prediction of ovarian response to controlled ovarian hyperstimulation in IVF.

References


Serum AMH and ovarian reserve


Submitted on June 7, 2002; accepted on August 1, 2002