Ovarian response prediction in GnRH antagonist treatment for IVF using anti-Müllerian hormone

O. Hamdine1,*, M.J.C. Eijkemans1,2, E.W.G. Lentjes3, H.L. Torrance1, N.S. Macklon1,4, B.C.J.M. Fauser1, and F.J. Broekmans1

1Department of Reproductive Medicine and Gynecology, University Medical Centre Utrecht, 3584 CX Utrecht, The Netherlands 2Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, 3848 CG Utrecht, The Netherlands 3Department of Clinical Chemistry and Hematology, University Medical Centre Utrecht, 3584 CX Utrecht, The Netherlands 4Academic Unit of Human Development and Health, Princess Anne Hospital, University of Southampton, Southampton SO16 5YA, UK

*Correspondence address. Department of Reproductive Medicine and Gynecology, University Medical Centre Utrecht, F05.126, Heidelberglaan 100, PO Box 85300, 3508 GA Utrecht, The Netherlands. E-mail: o.hamdine@umcutrecht.nl

Submitted on July 2, 2014; resubmitted on August 22, 2014; accepted on September 19, 2014

STUDY QUESTION: What is the clinical value of anti-Müllerian hormone (AMH) for the prediction of high or low ovarian response in controlled ovarian stimulation for IVF using GnRH antagonist treatment?

SUMMARY ANSWER: AMH as a single test has substantial accuracy for ovarian response prediction in GnRH antagonist treatment for IVF, with a higher accuracy for predicting a high response than for low response.

WHAT IS KNOWN ALREADY: The role of AMH and other patient characteristics in ovarian response prediction has been studied extensively in long GnRH agonist protocols; however, little information is available regarding the clinical value in GnRH antagonists.

STUDY DESIGN, SIZE, DURATION: This is an observational (retrospective) substudy as part of an ongoing cohort study. A total of 487 patients scheduled for IVF/ICSI between 2006 and 2011 were included in the study.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Patients with a regular cycle who underwent their first IVF/ICSI cycle with GnRH antagonist treatment while receiving a starting dose of 150 or 225 IU recombinant FSH were included in the study. Patients were divided into three subgroups according to the following ovarian response categories: high (> 15 oocytes or cycle cancellation), normal (4–15 oocytes) and low (< 4 oocytes or cycle cancellation). Serum samples collected prior to IVF treatment were used to determine serum AMH levels.

MAIN RESULTS AND THE ROLE OF CHANCE: According to the predefined ovarian response categories, 58 patients were classified as high, 326 as normal and 101 as low responders, and the ongoing pregnancy rates did not differ among groups (19.0, 22.1 and 16.8%, respectively, P = 0.9). For the prediction of high response, AMH had an area under the receiver-operating characteristic curve (AUC) of 0.87. Both female age and BMI had lower accuracy (AUC 0.66 and 0.58, respectively). For low response prediction, again AMH had a better accuracy (AUC 0.79) than female age and BMI (AUC 0.59 and 0.56, respectively). In a multivariate model, including the factors age, AMH, BMI, smoking, type and duration of subfertility, only BMI added some predictive value to AMH for both high and low response prediction. Clinical test characteristics demonstrated that using a specificity of ~90%, the detection rate of AMH for high and low response, corresponding with a test cut-off of 4.5 and 0.8 μg/l, was ~60 and ~45%, respectively.

LIMITATIONS, REASONS FOR CAUTION: The impact of the antral follicle count (AFC) on ovarian response prediction in GnRH antagonists was not assessed; however, previously studies demonstrated that for GnRH antagonists, AMH has a better accuracy than AFC.

WIDER IMPLICATIONS OF THE FINDINGS: The current study demonstrates that AMH is an adequate predictor for both high and low response in GnRH antagonist cycles, showing a similar accuracy to GnRH agonists, as reported previously. The optimization and individualization of GnRH antagonist protocols may be improved by using an AMH-tailored approach.

STUDY FUNDING/COMPETING INTEREST(S): This study was funded by the Academic Institutional Resources of the Department of Reproductive Medicine of the UMC Utrecht. O.H., M.I.C.E, E.W.G.L and H.L.T. have nothing to declare. N.S.M. has received fees and/or grant support from the following companies (in alphabetic order): Anecova, Ferring, Informa, Merck Serono and MSD. B.C.J.M.F. has received fees...
Introduction

The optimization and individualization of controlled ovarian stimulation (COS) for IVF have become increasingly important. Clinicians often use patient characteristics, such as female age, menstrual cycle length, BMI and results from previous IVF cycles to select a treatment protocol (La Marca and Sunkara, 2014). Treatment individualization has been hampered by disagreement as to which ovarian marker provides an accurate estimation of potential success for patients prior to IVF treatment. Several ovarian markers, including basal FSH, antral follicle count and anti-Müllerian hormone (AMH), have been suggested as predictors of ovarian response and clinical outcome (Broekmans et al., 2006). The role of these ovarian response tests (ORTs) has generally been studied in patients treated with a long GnRH agonist protocol. It has been demonstrated that AMH is an accurate predictor of both high (Broer et al., 2013a) and low ovarian response in GnRH agonist treatment (Broer et al., 2013b), suggesting it would be an ideal marker for the individualization of COS strategies. Indeed, the use of an AMH-tailored approach has previously been suggested by several investigators (Nelson et al., 2009; Yates et al., 2011; van Tilborg et al., 2012).

These days, many clinicians choose to use GnRH antagonists as a tool for luteinizing hormone (LH) surge suppression, as the availability of GnRH antagonists has enabled a reduction in complexity and costs, and has led to improved safety compared with GnRH agonists, without a clear difference in ongoing pregnancy rate (OPR) and live birth rate (Heijn et al., 2007; Al-Inany et al., 2011).

The accuracy of ORTs in ovarian response prediction for GnRH antagonist treatment regimens may differ from that in GnRH agonist protocols as there is a difference in the endocrine profile, early follicle recruitment and synchronization of follicular development, ultimately leading to a difference in number of oocytes retrieved (Huurre et al., 2007). Hence, predictive models cannot be extrapolated from GnRH agonist to GnRH antagonist protocols.

Only a limited number of studies have addressed the value of ORTs for ovarian response prediction in GnRH antagonists (Andersen et al., 2011; Polyzos et al., 2012, 2013; Arce et al., 2013). In one study, both AMH and basal FSH were found to be predictive factors for high response, whereas AMH was the only significant factor for low response (Andersen et al., 2011). Others also observed a high accuracy of AMH for the prediction of high and low response (Arce et al., 2013; Polyzos et al., 2013). However, among oocyte donors treated with a GnRH antagonist protocol, the predictive ability of AMH was only modest (Polyzos et al., 2012).

The difference in the accuracy of AMH found among these studies may be caused by the use of different definitions for ovarian response.

The question therefore remains whether AMH is able to correctly predict ovarian response to GnRH antagonist treatment with similar accuracy as GnRH agonist treatment. The aim of this study was to determine the accuracy and clinical value of AMH in the prediction of ovarian response in IVF using GnRH antagonist treatment.

Materials and Methods

Subject selection

All patients were selected from a prospective cohort study, registered at www.clinicaltrials.gov (Protocol ID 13-109) and approved by our Institutional Review Board. The purpose of this cohort study is to assemble data for research questions specifically related to IVF treatment and is currently still recruiting. All patients are treated at the IVF outpatient clinic of the Department of Reproductive Medicine and Gynecology of the University Medical Centre Utrecht, The Netherlands. Informed consent is obtained from all patients for the banking and use of both serum and DNA samples for research purposes regarding assisted reproduction. Blood samples are collected, irrespective of the cycle day (CD), during routine screening for hepatitis B and C and human immunodeficiency virus prior to IVF treatment. Blood samples were stored at −20°C until December 2012 and stored at −80°C as of January 2013. The baseline and IVF treatment data are prospectively recorded in our electronic infertility patient data file system.

For the current substudy, we selected women from the prospective cohort study which at that time comprised a total of 1031 patients. The selection was limited to patients with a regular cycle who underwent their first IVF/ICSI cycle with GnRH antagonist treatment while receiving a starting dose of 150 or 225 IU recombinant FSH (recFSH). Patients who had achieved a live birth after the previous IVF/ICSI treatment and requested a new treatment were also eligible candidates. Patients who fulfilled the inclusion criteria were treated between 2006 and 2011. Serum samples were retrieved from our biobank for the current study to determine serum AMH levels.

COS

COS was performed with recFSH (Gonal-f; Merck Serono, or Puregon; MSD, the Netherlands). A GnRH antagonist (Cetrotide; Merck Serono, or Orgalutran; MSD, the Netherlands) was used to prevent a premature LH surge. The patients were not pretreated with oral contraceptives. According to local protocol, recFSH (150—225 IU) was started on CD 2 or 3. The reasons to administer a starting dose of 225 IU recFSH/day instead of the standard dose of 150 IU recFSH in our clinic were female age >41 years, a previous treatment leading to live birth whilst using 225 IU recFSH daily or incipient ovarian failure (defined as a regular cycle, basal FSH >10 IU/l and age <40 years). GnRH antagonist treatment was commenced on stimulation Day 5. hCG (10 000 IU, Pregnyl; MSD. The Netherlands or 6500 IU Ovistrell; Merck Serono) was administered to induce final oocyte maturation when at least three follicles ≥17 mm diameter in diameter were visualized by ultrasound. Oocyte retrieval was performed 36 h after hCG administration. One or two embryos were transferred 3 or 4 days after oocyte retrieval. The luteal phase was supplemented with a daily dose of 600 mg vaginally administered micronized natural progesterone (Utrogesteran; Besins Healthcare, Brussels, Belgium).

AMH assay

After blood collection, serum for assay of AMH was separated and frozen in aliquots within 3—4 h. In December 2011, stored frozen samples were thawed overnight in the refrigerator in order to determine AMH levels. All
measurements were performed in a batch analysis using a DS2 ELISA robot with a single lot reagent (AMH Gen II ELISA, A79765, Beckman Coulter; Inc., USA). In the laboratory work up, a buffer was robotically pipetted first, and then the serum sample was added. The lower limit of AMH detection was 0.16 µg/l. Inter-assay variation was 10% at 0.27 µg/l and 4.7% at 3.9 µg/l (n = 18). The maximum time interval between serum sampling and the start of COS was 7 months (range 1 day–7 months).

Outcome measures
The primary outcome measure was ovarian response category. Since ovarian response definitions for GnRH antagonists are lacking, the most commonly used definitions for both high and low response in GnRH agonists were adapted for this study. A high response was arbitrarily defined as >15 oocytes retrieved or cancellation due to an anticipated risk of ovarian hyperstimulation syndrome (OHSS, Broer et al., 2013a). A low response was defined as <4 retrieved oocytes or cancellation due to low ovarian response (<3 dominant follicles >12 mm diameter), or a switch to intrauterine insemination (Broer et al., 2013b). A normal response was therefore defined as 4–15 oocytes retrieved. Secondary outcomes included the duration of stimulation; total cumulative dose of recFSH; number of oocytes retrieved; number of 2PN oocytes; number of suitable embryos for transfer and cryopreservation; ongoing implantation rate and OPR per started cycle.

Data analysis
All analyses for high and low response were performed for the whole group. Subgroup analyses were performed for the groups that received 150 and 225 IU recFSH. Data for continuous variables are presented as mean values and standard deviation. Between-group statistical comparisons of the mean values were performed with analysis of variance tests. $\chi^2$ tests were used for categorical data. Differences were considered to be statistically significant if $P<0.05$. Receiver operating characteristic (ROC) curves were constructed to demonstrate the predictive accuracy of AMH and other patient characteristics as single predictors and in combination using univariate and multivariate logistic regression, for both high and low ovarian response. The corresponding area under the curve (AUC) was calculated for both response groups in order to express the overall accuracy. In order to illustrate the clinical usefulness of AMH for the prediction of both high and low response, the sensitivity, specificity, likelihood ratios, pre- and post-test probability and the percentage of women with an abnormal test result, were calculated for several cut-off values of AMH, which were derived from the ROC curve. To explore the association between the number of oocytes and OPR per started cycle, for each oocyte number, the mean and 95% confidence interval (CI) was calculated for OPR per started cycle. Finally, different cut-off values for the number of oocytes used in the literature to define a high or low response were used to compare our results with previously published GnRH antagonist studies.

Results

Patient demographics and clinical outcome
Supplementary data, Figure S1 depicts the number of patients at each stage of the selection process. A total of 487 patients were included in the study; of which, 389 were scheduled to undergo an IVF treatment, with ICSI to be performed in 98 patients. Two patients withdrew from treatment prior to oocyte retrieval for personal reasons. The remaining patients were divided into three subgroups according to the ovarian response category; high (n = 58), normal (n = 326) and low (n = 101).

The baseline characteristics of the total group are shown in Table I. The subgroups differed significantly in age, bodyweight, BMI, AMH and type of subfertility. Table II demonstrates the stimulation characteristics and clinical outcome per started cycle. The majority of patients received a dosage of 150 IU recFSH daily (n = 439), while in a small subset a dose of 225 IU (n = 48) was administered. The recFSH dose was increased to 225 IU in 23 patients during stimulation. Cycle cancellation due to a low response or switch to intrauterine insemination occurred in 15 and 17 cases, respectively. Premature ovulation occurred once. Ten patients did not receive hCG due to a high response and thus a risk of OHSS. Twelve cases of mild OHSS and one case of severe OHSS, requiring hospitalization, were observed.

There was a significant between-group difference for the number of oocytes retrieved, number of 2PN oocytes, number of suitable and transferred or cryopreserved embryos and the percentage of single-embryo transfer. The OPR per started cycle did not differ among the three ovarian response groups. However, logistic regression demonstrated that OPRs rose with an increasing number of oocytes retrieved up to six oocytes [odds ratio (OR) 1.27, 95% CI 1.06–1.52, P = 0.009], whereas a non-significant trend towards lower pregnancy rates was observed beyond 15 oocytes (OR 0.84, 95% CI 0.62–1.14, P = 0.3). This is also illustrated in Fig. 1 which depicts the number of retrieved oocytes in relation to the OPR per started cycle.

Prediction of high and low ovarian response
To examine the predictive accuracy of AMH and other possible predictors of ovarian response, the parameters listed in Table I were analysed by univariable and multivariable logistic regression. ROC curves were plotted for single and combined predictors (Fig. 2). The levels of accuracy, as expressed by the AUCs, for ovarian response prediction are depicted in Table III, as described below.

High response
For the prediction of high response, AMH had the highest accuracy (AUC 0.87) when compared with age and BMI (Table III). Although the AUC of the combination ‘AMH and BMI’ remained similar (Table III), multivariate logistic regression demonstrated that the addition of BMI slightly improved the predictive accuracy of AMH (OR 0.89, 95% CI 0.81–0.98, P = 0.01). Bodyweight (AUC 0.59) also added some predictive value to AMH (OR 0.97, 95% CI 0.94–1.00, P = 0.03), similar to BMI. The addition of age, smoking and type or duration of subfertility did not add prognostic value to the AMH model (P = 0.4, 0.8, 0.2 and 0.4, respectively). A separate analysis (data not shown) of the subgroups that used 150 or 225 IU recFSH demonstrated similar AUCs for age, AMH, BMI and the multivariate model.

Table IV illustrates the clinical value of different AMH cut-offs for ovarian response prediction. When choosing a higher test cut-off level, the sensitivity and proportion of abnormal test results decreased, whereas the specificity, positive likelihood ratio and the post-test probability increased. At a specificity level of 90% and test cut-off of 4.5 µg/l, the test seemed to have the best performance level, indicating that in the case of an abnormal test result, the chance of having an excessive response is 50%, while at the same time, 60% of all true excessive responders will be identified.
Table I Baseline characteristics by subgroup for women undergoing IVF/ICSI in GnRH antagonist cycles.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>High response (n = 58)</th>
<th>Normal response (n = 326)</th>
<th>Low response (n = 101)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>32.4 ± 4.7</td>
<td>34.7 ± 4.3</td>
<td>35.7 ± 3.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Range</td>
<td>(22.8–41.6)</td>
<td>(21.2–44.0)</td>
<td>(24.2–42.7)</td>
<td></td>
</tr>
<tr>
<td>Bodyweight (kg)</td>
<td>66.0 ± 11.0</td>
<td>69.1 ± 12.7</td>
<td>72.2 ± 14.5</td>
<td>0.01</td>
</tr>
<tr>
<td>Range</td>
<td>(45.0–97.0)</td>
<td>(43.0–133.0)</td>
<td>(47.0–122.0)</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.9 ± 3.1</td>
<td>23.8 ± 4.1</td>
<td>24.9 ± 4.9</td>
<td>0.01</td>
</tr>
<tr>
<td>Range</td>
<td>(15.9–31.7)</td>
<td>(17.2–47.1)</td>
<td>(17.5–39.8)</td>
<td></td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td>11 (19.0)</td>
<td>46 (14.1)</td>
<td>19 (18.8)</td>
<td>0.5*</td>
</tr>
<tr>
<td>AMH prior to treatment (µg/l)</td>
<td>5.6 ± 3.4</td>
<td>2.5 ± 1.8</td>
<td>1.2 ± 1.2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Fertility characteristics

| Primary subfertility, n (%)            | 50 (86.2)              | 219 (67.2)                 | 73 (72.3)              | 0.01*   |
| Secondary subfertility, n (%)         | 8 (13.8)               | 107 (32.8)                 | 28 (27.7)              |         |
| Duration of subfertility (years)      | 3.2 ± 2.0              | 3.3 ± 2.0                  | 3.2 ± 2.4              | 1.0     |
| Cause of subfertility, n (%)          |                       |                            |                        |         |
| Male factor                            | 28 (48.3)              | 139 (42.6)                 | 41 (40.6)              | 0.4*    |
| Unexplained                            | 21 (36.2)              | 120 (36.8)                 | 34 (33.7)              |         |
| Tubal factor                           | 8 (13.8)               | 50 (15.3)                  | 15 (14.9)              |         |
| Endometriosis                          | 1 (1.7)                | 5 (1.5)                    | 2 (2.0)                |         |
| Incipient ovarian failure              | 0 (0.0)                | 6 (1.8)                    | 8 (7.9)                |         |
| Other                                   | 0 (0.0)                | 7 (2.1)                    | 1 (1.0)                |         |

High response: >15 oocytes or cancellation due to the risk of OHSS. Normal response: 4–15 oocytes. Low response: <4 oocytes or cancellation due to poor response. Data are presented as means ± SD. The ranges for age, bodyweight and BMI are depicted in italic font. P-values are for between-group analysis of variance (ANOVA) tests unless otherwise stated. P-values in bold are statistically significant. AMH, anti-Müllerian hormone. *P-value for between-group difference from χ² tests.

Table II Stimulation characteristics and clinical outcome.

<table>
<thead>
<tr>
<th>Stimulation characteristics</th>
<th>High response (n = 58)</th>
<th>Normal response (n = 326)</th>
<th>Low response (n = 101)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total dose of recombinant FSH (IU)</td>
<td>1347 ± 298</td>
<td>1432 ± 392</td>
<td>1476 ± 502</td>
<td>0.2</td>
</tr>
<tr>
<td>Total duration of stimulation (days)</td>
<td>8.7 ± 1.4</td>
<td>9.0 ± 1.8</td>
<td>9.1 ± 2.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Clinical outcome per started cycle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of oocytes</td>
<td>15.5 ± 7.8</td>
<td>8.4 ± 3.3</td>
<td>1.6 ± 1.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of 2PN oocytes</td>
<td>8.0 ± 5.9</td>
<td>4.5 ± 2.7</td>
<td>1.0 ± 1.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Suitable embryos for transfer or cryopreservation</td>
<td>6.7 ± 5.3</td>
<td>3.8 ± 2.5</td>
<td>0.9 ± 1.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of embryos transferred</td>
<td>1.0 ± 0.7</td>
<td>1.2 ± 0.6</td>
<td>0.7 ± 0.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Embryos cryopreserved</td>
<td>3.0 ± 4.0</td>
<td>1.2 ± 1.7</td>
<td>0.2 ± 0.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Single-embryo transfer (%)</td>
<td>33 (56.9)</td>
<td>197 (60.4)</td>
<td>41 (40.6)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Ongoing implantation rate per embryo (%)</td>
<td>20 ± 37.5</td>
<td>20.7 ± 38.5</td>
<td>25.0 ± 42.1</td>
<td>0.7*</td>
</tr>
<tr>
<td>Ongoing pregnancy per started fresh cycle, n (%)</td>
<td>11 (19.0)</td>
<td>72 (22.1)</td>
<td>17 (16.8)</td>
<td>0.9*</td>
</tr>
</tbody>
</table>

High response: >15 oocytes or cancellation due to the risk of OHSS. Normal response: 4–15 oocytes. Low response: <4 oocytes or cancellation due to poor response. Data are presented as means ± standard deviation (SD) and P-values are for between-group difference from ANOVA tests unless otherwise stated. P-values in bold are statistically significant. *P-value for between-group difference from χ² tests.
Low response

AMH had an AUC of 0.79 for low response prediction, whereas age and BMI had poor accuracy (AUC 0.59 and 0.56, respectively, Table III). In a multivariate logistic regression analysis, again BMI added some predictive value to AMH (OR 1.08, 95% CI 1.02–1.15, \(P = 0.01\)). Bodyweight (AUC 0.57) also had some added independent predictive value to AMH (OR 1.03, 95% CI 1.01–1.05, \(P = 0.01\)), similar to BMI. The factors age, smoking, type and duration of subfertility did not add prognostic value to AMH. Again, the subgroups that used 150 or 225 IU recFSH demonstrated similar AUCs for age, AMH, BMI and the multivariate models (data not shown). Table IV demonstrates that the performance of AMH as a test for the prediction of low response was limited, as reflected by the low sensitivities corresponding with lower AMH thresholds. The optimal cut-off point, if any, seems to lie at a level of 0.80 \(\mu\)g/l, thus identifying 50% of all low responders and with a probability of 50% having a low response in the case of an abnormal test.

Discussion

This prospective cohort study demonstrates that AMH as a single test has substantial accuracy in the prediction of ovarian response using GnRH antagonist down-regulation for IVF. Furthermore, the accuracy curves indicate that AMH is a better predictor for high than for low ovarian response.

The findings from the present study are in line with two systematic reviews on the predictive value of AMH for ovarian response using GnRH agonist treatment. These reviews summarize a large number of studies, from which solid information has become available demonstrating an AUC of 0.81 for the prediction of high response (Broer et al., 2013a) and 0.78 for the prediction of low response (Broer et al., 2013b).

The significance of AMH in a GnRH antagonist system has been addressed in only three other studies (Andersen et al., 2011; Arce et al., 2013; Polyzos et al., 2013) using a fixed start GnRH antagonist protocol with recFSH dosages of 150–225 IU. The reported accuracies of AMH for the prediction of high and low response differed among these studies. This may have been caused by differences in study population [maximum age 34 and 36 years, respectively, in Arce et al. (2013) and Polyzos et al. (2013)] or by the use of different definitions for ovarian response making it difficult to compare results. Hence, we reanalysed our data using the different oocyte number cut-offs which were applied in the aforementioned GnRH antagonist studies. To define a high response,
Arce et al. (2013), Andersen et al. (2011) and Polyzos et al. (2013) used cut-offs of >15, >18 and >20 oocytes, resulting in an accuracy of AMH of 0.81, 0.82 and 0.80, respectively. The use of these definitions substantially affected the accuracy of AMH in our data set (AUCs 0.87, 0.79 and 0.74, respectively). To define a low response, Arce et al. (2013) and Polyzos et al. (2013) used a similar cut-off (<4 oocytes), whereas Andersen et al. (2011) used a cut-off of <6 oocytes, demonstrating an accuracy of 0.90, 0.72 and 0.88, respectively. In contrast with high response prediction, using these definitions in our data set did not improve the accuracy of AMH (AUC 0.79). Thus, the results of the present study largely seem to confirm most of the findings in previous studies.

The definitions commonly applied for ovarian response categories are mostly based on GnRH agonists. It can be debated whether the use of a higher number of oocytes to define a high response should be applied in GnRH antagonist regimens, as these are associated with a lower number of oocytes retrieved compared with GnRH agonists (Al-Inany and Aboulghar, 2002; Huirne et al., 2007). The present study observed a non-significant trend towards lower pregnancy rates beyond 15 oocytes, which is in line with Sunkara et al. (2011) who demonstrated that the optimal number of oocytes associated with the chance of achieving live birth was \( \approx 15 \), whereas a decline was shown with >20 oocytes. The high estradiol levels associated with a high response, as well as the possible untimely changes in progesterone levels, may explain the lower OPR in this group of patients due to impaired endometrial receptivity and oocyte/embryo quality (Basir et al., 2001; Valbuena et al., 2001; Bosch et al., 2003, 2010; Macklon et al., 2006; Kolibianakis et al., 2012).

Additionally, the presence of \( \geq 18 \) follicles following GnRH antagonist treatment has previously been associated with an increased risk of developing OHSS (Papanikolaou et al., 2006). However, judging from the low incidence of \( \geq 18 \) oocytes in the present study \((n = 20, 4.1\%)\), defining a high response as \( \geq 15 \) oocytes may be adequate in GnRH antagonists.

A low response following conventional ovarian stimulation is regarded as a sign of advanced ovarian ageing (Beckers et al., 2002; Tarlatzis et al., 2003), as this type of stimulation will induce maximal stimulation of the ovaries.
ovaries (Huirne et al., 2007; Sterrenburg et al., 2011). Conversely, the retrieval of a low number of oocytes following GnRH antagonist treatment is associated with much more favourable pregnancy chances and may not reflect a ‘poor’ ovarian response (Hohmann et al., 2003; Verberg et al., 2009). The present study demonstrated that pregnancy rates rose up to six oocytes retrieved, which is similar to what has been observed in GnRH agonist treatment cycles (van der Gaast et al., 2006). Hence, defining low response at a lower cut off than the accepted level of <4 oocytes may yield a subgroup with sufficiently poor prospects for pregnancy that predicting such a group would be clinically relevant. Larger studies will be needed to support such a claim.

The current study demonstrated that the accuracy of AMH was slightly improved by the addition of BMI and bodyweight, whereas no improvement was observed after the addition of age, type or duration of subfertility and smoking. The biological availability of recFSH has been shown to be reduced in obese women. Moreover, BMI has been negatively associated with ovarian response (Wittener et al., 2000; Nichols et al., 2003; Steinkampf et al., 2003; Fedorcsak et al., 2004; Verberg et al., 2007; Pinborg et al., 2011). However, others have found no additional value of BMI for AMH in the prediction of high response (Broer et al., 2013a). In the present study, the accuracy of BMI as a single predictor of ovarian response was very low which is in line with previous GnRH agonist (Broer et al., 2013b) and GnRH antagonist (Andersen et al., 2011) studies. Nevertheless, BMI or bodyweight could be used in GnRH antagonist protocols to improve the accuracy of AMH.

Ideally, a test for ovarian response prediction would identify all women with a high or low response and exclude all women with a normal response. Table IV, however, demonstrates that the performance of AMH for predicting high and low response is not optimal. Judging by the abnormal test result corresponding with the optimal AMH cut-off for high response prediction, a considerable number of patients with a false-positive test would be treated with a lower FSH dose and may therefore become a low responder. Furthermore, the performance of AMH as a test for the prediction of low response was rather limited, as reflected by the low sensitivities corresponding to lower AMH thresholds. It remains to be established whether increased stimulation dosages in expected poor responders will result in better pregnancy prospects when using a GnRH antagonist protocol. Strengths of this study include the absence of selection bias as all women starting their first IVF/ICSI in our hospital were asked to participate in this study, and the absence of verification bias as AMH values were not available at the start of and during the IVF treatment. Furthermore, AMH was not measured during COS which has been shown to decrease AMH levels (Lee et al., 2010; Patrelli et al., 2012). A possible limitation is the time interval between serum sampling and initiation of COS in the present study (7 months). However, this is not likely to have influenced the results as a time interval up to 12 months between serum sampling and initiation of stimulation has been shown not to affect the predictive ability of AMH (Polyzos et al., 2013).

Accurate prediction of ovarian response prior to IVF is important to individualize stimulation regimens by, for example, adjusting the starting dose of gonadotrophins. Previous studies utilizing AMH to tailor IVF treatment have shown a reduction in the incidence of high and low response as well as improved pregnancy rates compared with non-individualized treatment cycles (Nelson et al., 2009; Yates et al., 2011). In the present study, no difference in pregnancy rates between the high, normal and low response groups was found. This may indicate that predicting a low response is clinically less relevant as opposed to predicting a high response, as here patient safety issues also play a role. One of the proposed preventive strategies for OHSS is the use of a GnRH agonist trigger to induce final oocyte maturation with or without cryopreservation of all embryos, which has resulted in a decreased incidence of OHSS in high-risk patients with no change in reproductive outcome (Humaidan et al., 2013). However, severe OHSS has recently been reported after GnRH agonist triggering, with and without luteal-phase hCG supplementation, despite cryopreservation of all embryos (Seyhan et al., 2013; Fatemi et al., 2014). Therefore, it remains crucial to individualize IVF treatment in order to decrease the incidence of OHSS, even though a direct benefit in terms of increased pregnancy rates remains to be established.

In conclusion, the current study demonstrates that AMH is an adequate predictor of both high and low response using GnRH antagonist treatment. The individualization of GnRH antagonist protocols may be further improved by using an AMH-tailored approach. However, since AMH has a higher accuracy for the prediction of high response than for low response, and the significance of a low response in GnRH antagonists regarding pregnancy prospects has become doubtful, studies into individualization may best focus on preventive strategies towards high response.

**Supplementary data**

Supplementary data are available at http://humrep.oxfordjournals.org/.

**Authors’ roles**

Study design: F.J.B., N.S.M. and B.C.J.M.F.. Data collection: O.H., H.L.T. and E.W.G.L. Analyses: O.H. and M.J.C.E. Interpretation of the results and writing of the article: all authors. Revision of the article: all authors.

**Funding**

This study was funded by the Academic Institutional Resources of the Department of Reproductive Medicine of the UMC in Utrecht, The Netherlands.

**Conflict of interest**

O.H., M.J.C.E, E.W.G.L. and H.L.T. have nothing to declare. N.S.M. has received fees and/or grant support from the following companies (in alphabetic order): Aeneova, Ferring, Informa, Merck Serono and MSD. B.C.J.M.F. has received fees and/or grant support from the following companies (in alphabetic order): Childhealth, CVON, Ferring, Ova-Science, PregLem, Roche and Watson laboratories. F.J.B. has received fees and/or grant support from the following companies (in alphabetic order): Merck Serono and MSD.

**References**


Al-Inany HG, Youssef MA, Aboulghar M, Broekmans F, Sterrenburg M, Smit J, Bou-Setta AM. Gonadotrophin-releasing hormone antagonists for


Valbuena D, Martin J, de Pablo JL, Remohi J, Pellicer A, Simon C. Increasing levels of estradiol are deleterious to embryonic implantation because they directly affect the embryo. *Fertil Steril* 2001;5:962–968.


