Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: from theory to practice

Antonio La Marca1,* and Sesh Kamal Sunkara2

1Mother-Infant Department, Institute of Obstetrics and Gynaecology, University of Modena and Reggio Emilia, 41100 Modena, Italy 2King’s College London, London, UK

Correspondence address. Mother-Infant Department, Institute of Obstetrics and Gynecology, University of Modena and Reggio Emilia, Policlinico di Modena, Largo del Pozzo, 41100 Modena, Italy. Tel: +39-338-4795303; Fax: +39-059-4224394; E-mail: antlamarca@libero.it

Submitted on February 20, 2013; resubmitted on June 15, 2013; accepted on June 21, 2013

TABLE OF CONTENTS

- Introduction
- The Complexity of Individualization of Therapy in IVF
- Methods
  - Results
  - Measuring AMH and AFC
  - Identification of Expected Poor Responders
  - Identification of Expected Hyper-responders
  - Selection of the GnRH Analogue Can Be Dictated by the Anticipated Ovarian Response
  - AMH, AFC and Ovarian Response to Exogenous Gonadotrophins
  - Individualization of Gonadotrophin Starting Dose by Using AMH and AFC: From Simple to Complex Models
- Conclusions

BACKGROUND: The main objective of individualization of treatment in IVF is to offer every single woman the best treatment tailored to her own unique characteristics, thus maximizing the chances of pregnancy and eliminating the iatrogenic and avoidable risks resulting from ovarian stimulation. Personalization of treatment in IVF should be based on the prediction of ovarian response for every individual. The starting point is to identify if a woman is likely to have a normal, poor or a hyper response and choose the ideal treatment protocol tailored to this prediction. The objective of this review is to summarize the predictive ability of ovarian reserve markers, such as antral follicle count (AFC) and anti-Mullerian hormone (AMH), and the therapeutic strategies that have been proposed in IVF after this prediction.

METHODS: A systematic review of the existing literature was performed by searching Medline, EMBASE, Cochrane library and Web of Science for publications in the English language related to AFC, AMH and their incorporation into controlled ovarian stimulation (COS) protocols in IVF. Literature available to May 2013 was included.

RESULTS: The search generated 305 citations of which 41 and 25 studies, respectively, reporting the ability of AMH and AFC to predict response to COS were included in this review. The literature review demonstrated that AFC and AMH, the most sensitive markers of ovarian reserve identified to date, are ideal in planning personalized COS protocols. These sensitive markers permit prediction of the whole spectrum of ovarian response with reliable accuracy and clinicians may use either of the two markers as they can be considered interchangeable. Following the categorization of expected ovarian response to stimulation clinicians can adopt tailored therapeutic strategies for each patient. Current scientific trend suggests the elective use of the GnRH antagonist based regimen for hyper-responders, and probably also poor responders, as likely to be beneficial. The selection of the appropriate and individualized gonadotrophin dose is also of paramount importance for effective COS and subsequent IVF outcomes.

CONCLUSION: Personalized IVF offers several benefits; it enables clinicians to give women more accurate information on their prognosis thus facilitating counselling especially in cases of extremes of ovarian response. The deployment of therapeutic strategies based on selective use of
GnRH analogues and the fine tuning of the gonadotrophin dose on the basis of potential ovarian response in every single woman can allow for a safer and more effective IVF practice.

**Key words:** IVF / individualization / ovarian reserve / antral follicle count / anti-Mullerian hormone

**Introduction**

According to the Italian national assisted reproduction technique (ART) register, 52,676 IVF cycles were performed in Italy in 2010. Of these cycles 9.9% (5215 cycles) were cancelled before oocyte retrieval; of which 6.7% of IVF cycles were cancelled for poor ovarian response, 1.5% for the risk of ovarian hyperstimulation syndrome (OHSS) and 1.7% for other causes. In other words in Italy alone a number close to 4500 cycles are cancelled every year due to an abnormal response to stimulation with gonadotrophins. The register also reports important data regarding rates of suspension or cancellation of IVF cycles according to the age of women. For women younger than 35 years, 8.1% of IVF cycles were cancelled, which is rather high. It could be assumed that within this younger population a significant number of cycles are likely to be cancelled because of the risk of OHSS rather than because of a poor ovarian response. On the other hand, for women over 40 years of age the cycle cancellation rate was higher, ranging from 11.5% (for women aged 40--42 years) to 17.4% (for women aged ≥43 years) and is more likely to be related to a poor ovarian response.

The main objective of individualization of treatment in IVF is to offer every single woman the best treatment tailored to her unique characteristics, thus maximizing success, eliminating iatrogenic risks, such as OHSS, and minimizing the risk of cycle cancellation (Fig. 1). In particular, the reduction in cycle cancellations would lead to reduced costs and possibly a lower number of couples dropping out of ART programmes. A study from the Netherlands showed that 40% of couples abandon IVF after a single cycle (Verberg et al., 2008). This study revealed that among the most common causes was the physical and psychological burden of the treatment and accounted for 35% of dropout. Another common cause for the dropout was an inadequate response to ovarian stimulation, which was unexpected in most cases, with 10% of couples quitting the IVF programme simply because of an inadequate response in the first cycle (Verberg et al., 2008). Hence, in some way a reduction in the rate of abnormal ovarian response to gonadotrophin stimulation could possibly reduce the dropout rate in an ART programme.

**The Complexity of Individualization of Therapy in IVF**

Although personalization of IVF treatment may lead to an improvement in patient compliance and better clinical practice, it is far from easy. The difficulty derives from the vast number of drugs and choices available for controlled ovarian stimulation (COS), such as the GnRH analogues, the gonadotrophin preparations and other adjuvant therapies, and from the lack of a clear evidence-based therapeutic approach for different subgroups of patients.

Clinicians usually choose therapies according to anamnestic and/or clinical criteria, the most important being the outcome of previous IVF cycles. The selection of a clinical protocol appears much easier in women who have undergone previous IVF attempts. If a previous cycle had a good performance, the clinician is likely to conform to the protocol. Conversely, if a previous cycle had an undesirable outcome, the protocol is likely to be modified. If no previous cycle has been performed, the choice is likely to be empirical, and based on either the clinician’s or a centre’s preference. The clinical criteria used by most clinicians to select a protocol usually include the woman’s age, BMI, menstrual cycle characteristics, features suggestive of polycystic ovary syndrome...
(PCOS), such as hyperandrogenism, and previous ovarian surgery (Homburg and Insler, 2002; Arslan et al., 2005).

A key factor determining the outcome of COS and subsequent IVF outcome is selection of the gonadotrophin starting dose. The need for individualizing gonadotrophin dosage derives from the assumption that variability in the functional ovarian reserve (the pool of recruitable follicles) is very wide (Gougeon and Lefèvre, 1983; Gougeon, 1998; Almog et al., 2011; La Marca et al., 2011a; Monget et al., 2012) and consequently a standard fixed dose of gonadotrophin may not be suitable for all women. Correct individualization of the gonadotrophin start dose is an extremely important clinical decision. For example, in a woman with either a normal or an elevated ovarian reserve, the choice of an unduly low gonadotrophin dose could lead to a mono or pauci follicular development, not always desired in IVF cycles. On the other hand, the choice of an excessive dose could provoke an excessive ovarian response with subsequent OHSS risk. In recent years, the prediction of the extremes of ovarian response and consequent dose adjustment has been the subject of interest amongst IVF experts (Broekmans et al., 2006; Nelson et al., 2009; La Marca et al., 2010; Broer et al., 2011; Devroey et al., 2011; Nelson, 2013). Obtaining detailed background information regarding an individual’s ovarian potential should be considered as vital before commencing stimulation. Recent studies have suggested that the prescription of standard medications is unacceptable both from an ethical and legal point of view, as it could have negative results for the woman (Nelson et al., 2009; Nardo et al., 2011).

As stated above, the correct individualization of treatment protocols in IVF should be based on the correct prediction of ovarian response especially the extremes, namely poor and hyper response. The aim is then to choose the ideal treatment protocol according to this prediction. The prediction of a poor or hyper response also allows clinicians to give women more accurate information regarding the likelihood of these scenarios occurring during their IVF cycle. Patients may receive more accurate information on possible protracted treatment, cycle cancellation, OHSS, treatment burden and reduced success. Finally, if personalization is based on the accurate prediction of ovarian response, then the prediction of ovarian response should be based on the most sensitive markers of ovarian reserve. In this article we discuss the use of the most recently identified markers of ovarian reserve, namely antral follicle count (AFC) and anti-Mullerian hormone (AMH), to categorize women based on their anticipated ovarian response. The marker-based strategy of assessing ovarian reserve in women in order to select the ideal therapeutic approach in IVF is reviewed.

**Methods**

A literature search was carried out for studies that addressed the ability of AMH and AFC to predict poor and/or excessive ovarian response in IVF cycles. A systematic search of Medline, EMBASE, Cochrane library and Web of Science databases was carried out using the keywords, anti-Mullerian hormone, AMH Mullerian inhibiting substance, antral follicles, AFC and several synonyms of IVF and ICSI. Criteria were identified in the title and/or abstract of the publications. Additional journal articles were identified from the bibliographies of included studies as well as textbooks. Literature available up to May 2013 was included. Searches were conducted by both the authors. Any article that could possibly be of value for the association between AMH or AFC and IVF outcome were preselected. Only studies reporting cut-off values for at least one of the two markers were included in the review for discussion. Moreover any article that could possibly be proposing individualization of the IVF therapy on the basis of such a prediction were included.

**Results**

The searches generated 305 citations. Of these, 160 articles were excluded on the basis of title and abstract. Another 90 studies were excluded on the basis of the fully read article. Finally, 41 studies that reported on the ability of AMH to predict response to COS and 25 studies that reported on the ability of AFC were included in this review (Fig. 2).

![Figure 2](https://academic.oup.com/humupd/article-abstract/20/1/124/887248/7268)
Measuring AMH and AFC

In-depth analysis of problems related to the measurement of AMH and AFC has been previously detailed (Broekmans et al., 2010; Nelson and La Marca 2011; La Marca et al., 2013). Over the last years there has been an evolution of AMH assays from the laboratory versions (Hudson et al., 1990; Long et al., 2000; Al-Qahtani et al., 2005), through to the commercially available Diagnostic Systems Lab (DSL) and Immunotech Beckman Coulter (IBC) assays, and published studies have used either the DSL assay or the IBC assay. As these assays utilize two different antibodies against AMH, the values reported by different authors have varied substantially with the IBC assay giving values for AMH that are higher than with the DSL assay. Issues regarding different AMH assays have now been resolved with Beckman Coulter purchasing the patents for all previous versions and establishing the AMH Generation II assay. This assay retains the cross-species specificity of the DSL assay and is calibrated to the IBC standard. Generally, values found with the DSL assay can be converted to the IBC assay by multiplying by 1.39 (Wallace et al., 2011) while the new Generation II assay should give values similar to old IBC assay (Kumar et al., 2010).

AFC is the number of antral follicles present in the ovaries and detectable by transvaginal ultrasound scan. AFC is strongly related to circulating AMH levels since the hormone is produced by antral follicles themselves (Weenen et al., 2004). Although ovarian follicles smaller than 2 mm diameter can be detected with modern ultrasound equipment, it is however still not clear what class of antral follicles may better correlate with AMH levels and the number of retrieved oocytes (Jeppesen et al., 2013), with different authors suggesting that probably the 2–5 or 4–6 mm diameter categories may better represent the age-dependent proportion of the visible antral follicle pool (Jeppesen et al., 2013). However, the number of smaller antral follicles (2–5 mm) is highly correlated with the number of larger antral follicles (2–10 mm) (Jayaprakasan et al., 2010a, b). It was therefore suggested that counting all identifiable antral follicles of 2–10 mm in diameter would provide the most practical method for assessment of AFC in clinical practice (Broekmans et al., 2010). A major technical improvement in ultrasound has been the development of three-dimensional (3D) automated follicular tracking, which can substantially decrease both intra- and inter-observer variability (Deb et al., 2011, 2013). Although a limited number of studies of the new 3D technique have been published, the increasing attention of clinicians on the new volumetric imaging suggests that the automated follicular tracking may become the gold standard for AFC measurement in the future.

Identification of Expected Poor Responders

The recent European Society of Human Reproduction and Embryology Consensus Conference established a standardized definition of poor ovarian response as the retrieval of <4 oocytes following a standard IVF protocol, i.e. following maximal stimulation (Ferraretti et al., 2011). The incidence of poor ovarian response in IVF ranges from 10 to 20% and the prevalence varies depending on the woman’s age, with a lower prevalence among women aged <34 years and increasing considerably with advancing age, reaching 50% in women aged 43–44 years (Ferraretti et al., 2011). Prediction of poor response is vital for the counselling and management of these women in clinical practice. As previously explained, the prediction criteria used by clinicians mainly include anamnestic characteristics, such as the outcome of previous IVF cycles, the woman’s advanced age, the presence of short menstrual cycles (a clinical manifestation of ovarian ageing) and previous ovarian surgery (Ferraretti et al., 2011). Markers of ovarian reserve are regularly used to predict poor ovarian response and a suboptimal response to gonadotrophin stimulation is suspected in the presence of high levels of FSH and/or estradiol, or more recently on the basis of a low AFC or reduced levels of AMH (Toner et al., 1991; Seifer et al., 2002; van Rooij et al., 2002; Hazout et al., 2004; La Marca et al., 2010, 2012a; Broer et al., 2011).

An important factor when using ovarian reserve markers as predictors of ovarian response is to establish the most sensitive markers and acceptable cut-off levels for these markers. By appropriate cut-off levels we mean those values that can distinguish with sufficient accuracy women who are likely to have a normal response from those likely to have abnormal responses to ovarian stimulation. Studies reporting cut-off values, sensitivity and specificity for AMH and AFC in the prediction of poor response in IVF are described in Table I. Although the number of studies is constantly increasing, the vast majority of prospective studies have been limited to a small number of patients. AMH and AFC values reported in literature are very variable thus creating difficulties for clinicians in selecting the best cut-off values based on evidence. The variability could be explained by factors such as studies involving small sample sizes, the different definitions of poor ovarian response adopted by various authors consequently resulting in variations in the predictive values of markers of ovarian reserve and to the varying methods used to estimate the individual ovarian reserve markers.

Cut-off levels of AMH values for poor ovarian response reported in the literature vary between 0.1 and 2.97 ng/ml, which is within the range of normal values for AMH in healthy women (Table I). As is always necessary, clinicians adopting a cut-off value from published studies should carefully and critically review the literature identifying studies that in some way may better reflect their clinical setting. The two largest prospective studies published to date have included 340 and 356 women, respectively (Nelson et al., 2007; Al-Azemi et al., 2011). In Nelson et al. (2007) the best cut-off value for AMH was 5 pmol/l (0.7 ng/ml) (DSL assay), which was associated with a sensitivity of 75% and specificity of 91%. Al-Azemi et al. (2011) found an AMH value of 1.36 ng/ml (9.7 pmol/l) (IBC assay) to be associated with 75.5% sensitivity and 74.8% specificity. According to published data a cut-off value of AMH ranging between 0.7–1.3 ng/ml may be considered acceptable for the prediction of poor response in IVF.

Like AMH, AFC can be used to reliably predict ovarian response in IVF but there is considerable variability in agreed AFC cut-off levels (Table II). Cut-off values used for predicting poor response vary between an AFC < 3 (Chang et al., 1998) and <12 (Melo et al., 2009). A possible reason for such variability is the absence of a standardized measurement of antral follicles with different studies measuring different follicle populations; 2–5, 2–9 or 5–9 mm. Importantly, the impact of the technological improvements in ultrasound equipment is difficult to analyse but with few doubts the most recent papers may be based on technologies similar to those that are available now. Hence, remaining focused on the most recent papers, the most frequently reported cut-off values of AFC for prediction of poor response ranged between <5 and <7 (Frattarelli et al., 2003; Jayaprakasan et al., 2010a, b).

Following the selection of appropriate cut-off values for AFC and AMH, the prediction of poor response is fairly easy and is certainly useful for counselling women especially of the possible negative IVF outcomes such as cancellation of cycle, prolonged treatment, increased treatment burden and reduced pregnancy rates (Fig. 3). This could
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>n</th>
<th>Assay used</th>
<th>Cut-off value</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Conversion to AMH gen II assay¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Rooij et al. (2002)</td>
<td>Prospective</td>
<td>119</td>
<td>IBC</td>
<td>0.3⁴ 2.1</td>
<td>60</td>
<td>89</td>
<td>0.3³ 2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muttukrishna et al. (2004)</td>
<td>Prospective</td>
<td>69</td>
<td>IBC</td>
<td>0.1³ 0.7</td>
<td>87.5</td>
<td>72.2</td>
<td>0.1³ 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muttukrishna et al. (2005)</td>
<td>Retrospective</td>
<td>108</td>
<td>IBC</td>
<td>0.2³ 1.4</td>
<td>87</td>
<td>64</td>
<td>0.2³ 1.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tremellen et al. (2005)</td>
<td>Prospective</td>
<td>75</td>
<td>IBC</td>
<td>1.1 8.1⁴ 80</td>
<td>85</td>
<td>1.1 8.1⁴ 80</td>
<td>0.68 4.9⁵</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peñarrubia et al. (2005)</td>
<td>Prospective</td>
<td>80</td>
<td>IBC</td>
<td>0.68 4.9⁵ 53</td>
<td>96</td>
<td>0.68 4.9⁵ 53</td>
<td>0.68 4.9⁵</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebner et al. (2006)</td>
<td>Prospective</td>
<td>141</td>
<td>IBC</td>
<td>1.66⁶ 11.9</td>
<td>69</td>
<td>86</td>
<td>1.66⁶ 11.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ficicioglu et al. (2006)</td>
<td>Prospective</td>
<td>50</td>
<td>DSL</td>
<td>2.5³ 17.9</td>
<td>90.9</td>
<td>90.9</td>
<td>3.47 24.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Marca et al. (2007a)</td>
<td>Prospective</td>
<td>48</td>
<td>IBC</td>
<td>0.75³ 5.4</td>
<td>80</td>
<td>93</td>
<td>0.7⁵ 9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fréour et al. (2007)</td>
<td>Prospective</td>
<td>69</td>
<td>IBC</td>
<td>1.3³ 9.3</td>
<td>44</td>
<td>100</td>
<td>1.3³ 9.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smeenk et al. (2007)</td>
<td>Prospective</td>
<td>80</td>
<td>IBC</td>
<td>1.4³ 10</td>
<td>62</td>
<td>73</td>
<td>1.4³ 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McIlveen et al. (2007)</td>
<td>Prospective</td>
<td>128</td>
<td>IBC</td>
<td>1.15 14</td>
<td>73</td>
<td>73</td>
<td>1.15 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kwee et al. (2008)</td>
<td>Prospective</td>
<td>77</td>
<td>DSL</td>
<td>0.35³ 2.5</td>
<td>90.1</td>
<td>81.8</td>
<td>0.48 3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakhuda et al. (2007)</td>
<td>Prospective</td>
<td>75</td>
<td>DSL</td>
<td>0.75³ 5.4</td>
<td>80</td>
<td>93</td>
<td>0.7³ 9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lekamge et al. (2007)</td>
<td>Prospective</td>
<td>126</td>
<td>IBC</td>
<td>1.96 14⁶ 73</td>
<td>73</td>
<td>73</td>
<td>1.9 14⁶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson et al. (2007)</td>
<td>Prospective</td>
<td>340</td>
<td>DSL</td>
<td>0.7 5³ 75</td>
<td>91</td>
<td>0.97 6.95</td>
<td>0.97 6.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gnoth et al. (2008)</td>
<td>Prospective</td>
<td>132</td>
<td>DSL</td>
<td>1.26³ 9</td>
<td>97</td>
<td>41</td>
<td>1.75 12.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jayaprakasan et al. (2008b)</td>
<td>Prospective</td>
<td>135</td>
<td>DSL</td>
<td>0.99³ 7.1</td>
<td>100</td>
<td>73</td>
<td>0.99 7.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riggs et al. (2008)</td>
<td>Retrospective</td>
<td>123</td>
<td>DSL</td>
<td>0.83³ 5.9</td>
<td>83</td>
<td>79</td>
<td>0.83 7.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elgindy et al. (2008)</td>
<td>Prospective</td>
<td>33</td>
<td>IBC</td>
<td>2.7³ 19.3</td>
<td>83.3</td>
<td>82.4</td>
<td>2.7³ 19.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nardo et al. (2009)</td>
<td>Prospective</td>
<td>165</td>
<td>IBC</td>
<td>1³ 7.1</td>
<td>87</td>
<td>67</td>
<td>1.39 9.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barad et al. (2009)</td>
<td>Retrospective</td>
<td>76</td>
<td>DSL</td>
<td>0.5³ 3.6</td>
<td>87</td>
<td>84</td>
<td>0.69 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riggs et al. (2011)</td>
<td>Retrospective</td>
<td>78</td>
<td>DSL</td>
<td>1.5³ 10.7</td>
<td>86</td>
<td>78</td>
<td>2.1 14.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al-Azemi et al. (2011)</td>
<td>Prospective</td>
<td>356</td>
<td>IBC</td>
<td>1.36³ 9.7</td>
<td>75.5</td>
<td>74.8</td>
<td>1.36³ 9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2011a, b)</td>
<td>Prospective</td>
<td>172</td>
<td>IBC</td>
<td>1.08³ 7.7</td>
<td>95</td>
<td>76</td>
<td>1.08³ 7.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buyuk et al. (2011)</td>
<td>Retrospective</td>
<td>73</td>
<td>DSL</td>
<td>0.6³ 4.3</td>
<td>70</td>
<td>70</td>
<td>0.6³ 4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurt et al. (2011)</td>
<td>Prospective</td>
<td>180</td>
<td>DSL</td>
<td>2.9³ 21.2</td>
<td>100</td>
<td>89.6</td>
<td>4.1 29.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2011a)</td>
<td>Retrospective</td>
<td>1538</td>
<td>DSL</td>
<td>0.68³ 4.8</td>
<td>64.7</td>
<td>85.1</td>
<td>0.94 6.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fridén et al. (2011)</td>
<td>Retrospective</td>
<td>127</td>
<td>DSL</td>
<td>0.7 5³ 75</td>
<td>75</td>
<td>0.97 6.95</td>
<td>0.97 6.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yoo et al. (2011)</td>
<td>Retrospective</td>
<td>91</td>
<td>IBC</td>
<td>0.95³ 6.8</td>
<td>73.3</td>
<td>82.1</td>
<td>0.95³ 6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tolikas et al. (2011)</td>
<td>Prospective</td>
<td>90</td>
<td>DSL</td>
<td>2.74³ 19.6</td>
<td>69</td>
<td>70.5</td>
<td>3.8 27.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonilla-Musoles et al. (2012)</td>
<td>Prospective</td>
<td>143</td>
<td>IBC</td>
<td>1.3 9.2⁶ 69</td>
<td>64</td>
<td>1.3 9.2⁶ 69</td>
<td>1.3 9.2⁶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satwik et al. (2012)</td>
<td>Prospective</td>
<td>198</td>
<td>DSL</td>
<td>2³ 14.3</td>
<td>20</td>
<td>98</td>
<td>2.78 19.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Lee et al. (2012)  Prospective  162  IBC  1.08<sup>b</sup>  7.7  85.8  78.6  1.08<sup>b</sup>  7.7
Honnma et al. (2013)  Retrospective  456  IBC  1.4  10<sup>b</sup>  72.2  75.7  1.4  10<sup>b</sup>
Mutlu et al. (2013)  Prospective  192  DSL  0.94<sup>b</sup>  6.7  71  85  1.3  9.3
Arce et al. (2013)  Retrospective  759  AMH gen II  1.68  12<sup>b</sup>  92  83  1.68  12<sup>b</sup>
Polyzos et al. (2013)  Retrospective  210  AMH gen II  1.37<sup>b</sup>  9.78  74.1  77.5  1.37<sup>b</sup>  9.78

Hyper response

van Rooij et al. (2002)  Prospective  114  IBC  3.5<sup>b</sup>  25  40  95  3.5<sup>b</sup>  25
Eldar-Geva et al. (2005)  Prospective  53  IBC  3.5<sup>b</sup>  25  72  89  3.5<sup>b</sup>  25
La Marca et al. (2007a)  Prospective  48  IBC  2.6<sup>b</sup>  18.6  86  56  2.6<sup>b</sup>  18.5
Kwee et al. (2008)  Prospective  110  DSL  5<sup>b</sup>  35.7  53  91  6.95  49.6
Nelson et al. (2007)  Prospective  340  DSL  3.5  25<sup>b</sup>  60  94.9  4.8  34.7
Riggs et al. (2008)  Retrospective  123  DSL  1.59<sup>b</sup>  11.3  84  67  2.21  15.7
Lee et al. (2008)  Prospective  262  DSL  3.36<sup>b</sup>  23.9  62  87  4.67  33.2
Nardo et al. (2009)  Prospective  165  DSL  3.5<sup>b</sup>  25  88  70  4.8  34.7
Aflatoonian et al. (2009)  Prospective  159  IBC  4.83<sup>b</sup>  34.5  93  78  4.83<sup>b</sup>  34.4
Riggs et al. (2011)  Retrospective  78  DSL  3  21.4  70  71  4.17  29.7
Ocal et al. (2011)  Retrospective  695  DSL  3.3<sup>b</sup>  23.6  90  71  61  94  4.6  32.6
Honnma et al. (2012)  Retrospective  456  IBC  2.46  17.6<sup>b</sup>  69  75  2.46  17.6<sup>b</sup>
Anckaert et al. (2012)  Retrospective  731  IBC  4.17  29.8<sup>b</sup>  82.5  70.4  4.17  29.8<sup>b</sup>
Lee et al. (2012)  Prospective  162  IBC  3.57  25.5  94.4  83.3  3.57  25.5
Arce et al. (2013)  Retrospective  759  AMH gen II  3.9  28<sup>b</sup>  78  67  3.9  28<sup>b</sup>
Polyzos et al. (2013)  Retrospective  210  AMH gen II  3.52<sup>b</sup>  25.1  89.5  83.8  3.52<sup>b</sup>  25.1

PPV, positive predictive value; NPV, negative predictive value.

<sup>a</sup>Values from the original study have been converted to the recent AMH gen II assay by using conversion factor reported in Wallace et al. (2011) and Kumar et al. (2010).

<sup>b</sup>Indicates the unit of measurement used in the study.
prepare the woman embarking on a psychologically demanding treat-
ment cycle and perhaps reduce the number of dropouts particularly
among women with an expected poor outcome. With regards to a
refusal of IVF treatment for women who have been predicted as poor
responders, it is commonly agreed not to exclude anyone from their
first IVF attempt only on the basis of the ovarian reserve test, as the ac-
curacy of these tests can be poor for the prediction of pregnancy
(Smeenk et al., 2007; Lie Fong et al., 2008; La Marca et al., 2010; Broer
et al., 2011). Furthermore, while AFC and AMH are the best predictive
markers of ovarian reserve available, neither is completely reliable, with a
false positive rate of 10–20%. Moreover, even though the possibility of
pregnancy is reduced, pregnancy rates in younger poor responders are
still considered acceptable (Ulug et al., 2003; Klinkert et al., 2005;
Oudendijk et al., 2012). Finally, it has been shown that the chance of con-
ception through IVF is not negligible even with prediction of a very low
ovarian reserve (Weghofer et al., 2011) although it is lower than in a
woman of the same age with a good ovarian reserve, as both age
and ovarian reserve are independent predictors of live birth after IVF
(Gleicher et al., 2010; Holte et al., 2011; La Marca et al., 2011b; Lee
et al., 2011a, b; Jayaprakasan et al., 2012; Khader et al., 2013).

### Identification of Expected Hyper-responders

The term ‘hyper response’ refers to the retrieval of > 15 (La Marca et al.,
2010; Broer et al., 2011) or > 20 (Nelson et al., 2007) oocytes following a
standard COS protocol. The prevalence rate in IVF cycles is estimated to
be around 7% and varies with the woman’s age: it is around 15% in
women aged ≤30 years, declining with advancing age. It is of great im-
portance to accurately predict women who are likely to have a high re-
sponse to COS as it is the main risk factor for OHSS (Humaidan et al.,
2010a, b). OHSS could be a life threatening condition and is character-
ized by cystic ovarian enlargement and by a dramatic and rapid shift of

### Table II Cut-off values of antral follicle count (AFC) for the prediction of poor- and hyper response in IVF cycles.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>n</th>
<th>AFC cut-off</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Poor response</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang et al. (1998)</td>
<td>Prospective</td>
<td>149</td>
<td>3</td>
<td>73</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharara and McClamrock (1999)</td>
<td>Prospective</td>
<td>127</td>
<td>4</td>
<td>53</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frattarelli et al. (2000)</td>
<td>Retrospective</td>
<td>278</td>
<td>10</td>
<td>87</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsieh et al. (2001)</td>
<td>Prospective</td>
<td>372</td>
<td>3</td>
<td>61</td>
<td>94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nahum et al. (2001)</td>
<td>Prospective</td>
<td>224</td>
<td>6</td>
<td>95</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frattarelli et al. (2003)</td>
<td>Prospective</td>
<td>267</td>
<td>4</td>
<td>30</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jarvela et al. (2003)</td>
<td>Prospective</td>
<td>45</td>
<td>4</td>
<td>86</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yong et al. (2003)</td>
<td>Prospective</td>
<td>46</td>
<td>4</td>
<td>9</td>
<td>97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buncsi et al. (2004)</td>
<td>Prospective</td>
<td>120</td>
<td>4</td>
<td>61</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durmusoglu et al. (2004)</td>
<td>Retrospective</td>
<td>91</td>
<td>6.5</td>
<td>85</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ng et al. (2005)</td>
<td>Prospective</td>
<td>131</td>
<td>4</td>
<td>33</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muttukrishna et al. (2005)</td>
<td>Retrospective</td>
<td>108</td>
<td>5</td>
<td>89</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ficicioglu et al. (2006)</td>
<td>Prospective</td>
<td>44</td>
<td>7</td>
<td>77</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soldevila et al. (2007)</td>
<td>Prospective</td>
<td>327</td>
<td>8</td>
<td>62</td>
<td>74.8</td>
<td>59.1</td>
<td>77</td>
</tr>
<tr>
<td>Jayaprakasan et al. (2007)</td>
<td>Prospective</td>
<td>100</td>
<td>7</td>
<td>100</td>
<td>92.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kwee et al. (2008)</td>
<td>Prospective</td>
<td>110</td>
<td>6</td>
<td>41</td>
<td>95</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Mela et al. (2009)</td>
<td>Prospective</td>
<td>1074</td>
<td>12</td>
<td>71.1</td>
<td>69.2</td>
<td>83.3</td>
<td>52.6</td>
</tr>
<tr>
<td>Jayaprakasan et al. (2010a, b)</td>
<td>Prospective</td>
<td>135</td>
<td>11</td>
<td>93</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tolias et al. (2011)</td>
<td>Prospective</td>
<td>90</td>
<td>4.5</td>
<td>72.4</td>
<td>80.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonilla-Musoles et al. (2012)</td>
<td>Retrospective</td>
<td>143</td>
<td>7</td>
<td>72</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutlu et al. (2013)</td>
<td>Retrospective</td>
<td>192</td>
<td>5.5</td>
<td>91</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyzos et al. (2013)</td>
<td>Retrospective</td>
<td>210</td>
<td>8</td>
<td>72.2</td>
<td>84.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hyper response</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ng et al. (2000)</td>
<td>Prospective</td>
<td>128</td>
<td>9</td>
<td>60</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Rooy et al. (2002)</td>
<td>Prospective</td>
<td>114</td>
<td>14</td>
<td>92</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eldar-Geva et al. (2005)</td>
<td>Prospective</td>
<td>56</td>
<td>14</td>
<td>94</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kwee et al. (2007)</td>
<td>Prospective</td>
<td>110</td>
<td>14</td>
<td>81</td>
<td>89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afatoonian et al. (2009)</td>
<td>Prospective</td>
<td>159</td>
<td>16</td>
<td>89</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ocal et al. (2011)</td>
<td>Retrospective</td>
<td>82</td>
<td>8*</td>
<td>78</td>
<td>65</td>
<td>52</td>
<td>86</td>
</tr>
<tr>
<td>Polyzos et al. (2013)</td>
<td>Retrospective</td>
<td>210</td>
<td>16</td>
<td>80</td>
<td>84.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Prediction of ovarian hyperstimulation syndrome; AFC, antral follicle count.
intravascular fluid into the third space; liver/renal derangement, ovarian haemorrhage and thromboembolism. As with poor response, prediction of high response is currently based on clinical criteria and anamnestic characteristics, such as young age, long menstrual cycles, evidence of symptoms of PCOS and hyper response in a previous cycle (Humaidan et al., 2010a, b; Papanikolaou et al., 2011).

Markers of ovarian reserve, in particular AMH and AFC, once again allow for considerable enhancement in identifying women who are likely to show a high response to COS. Few studies on the use of AMH for the prediction of a high response have been published to date (Table I). These studies have used either the DSL assay or the IBC assay and only two studies used the AMH Gen II assay for the prediction of hyper response (Arce et al., 2013; Polyzos et al., 2013). Studies based on the IBC assay have reported AMH cut-off levels between 2.6 and 4.83 ng/ml while for the ‘old’ DSL assay values ranging from 1.59 to 5 ng/ml have been reported. The two studies based on the AMH Gen II (Arce et al., 2013; Polyzos et al., 2013) found 3.9 and 3.52 ng/ml, respectively, as acceptable cut-off values for the prediction of hyper response (Table I). The diagnostic accuracy of basal pretreatment AMH level for the prediction of OHSS (cut-off 3.36 ng/ml, DSL assay) has been shown to be good and comparable to the estradiol levels or number of ovarian follicles on the day of hCG administration (Lee et al., 2008), thus allowing for its use reliably in routine clinical practice with the advantage of the pretreatment prediction, thereby allowing for preventive strategies. Interestingly there are even fewer studies on AFC specifically in the context of predicting a high response (Table II). The largest prospective study published to date was based on 159 women undergoing IVF. Aflatoonian et al. (2009) reported an AFC value of >16, with an apparent sensitivity of 89% and a specificity of 92%, for the prediction of high response. Other smaller prospective studies found values ranging between 9 and 14 as the most appropriate cut-off to identify hyper-responders (Ng et al., 2000; Van Rooij et al., 2002; Eldar-Geva et al., 2005; Kwee et al., 2007). Obviously more research, possibly based on well-designed prospective studies and on a larger number of patients, is urgently needed to confirm the best values for both AMH and AFC to be adopted in clinical practice to identify women likely to hyper respond to COS.

![Figure 3](https://academic.oup.com/humupd/article-abstract/20/1/124/887248)

**Figure 3** Ovarian reserve testing before the first IVF cycle would permit to categorize patients as expected poor-, normal- or hyper-responders. Since there is no evidence of superiority of one approach over another in the treatment of poor responders, the protocol associated with reduced discomfort and treatment burden should be preferred. In hyper-responder patients, one of the most important objectives of medical counselling is to prevent OHSS. Hence the first line protocol would be based on administration of low doses of FSH in a GnRH-antagonist-based scheme. AFC, antral follicle count; AMH, anti-Mullerian Hormone.

Selection of the GnRH Analogue Can Be Dictated by the Anticipated Ovarian Response

The GnRH agonist long regimen is still the most frequently used COS regimen in IVF treatment (Daya, 2000). GnRH antagonists which prevent a premature LH surge by their more direct action were introduced as an alternative to the GnRH agonists allowing a shorter duration of treatment. Moreover, as the GnRH antagonist regimen avoids the profound suppression of endogenous FSH and LH concentrations in the early follicular phase at the stage of follicular recruitment, this was considered to be an advantage for some subgroups of patients (i.e. poor responders) (Kenisberg et al., 1984). However several trials and meta-analyses showed that the GnRH agonist long and the GnRH antagonist regimen are comparable in their efficacy for the outcome of IVF for poor responders (Pu et al., 2011; Sunkara et al., 2013). However, given that the use of the GnRH agonist regimen was associated with higher...
gonadotropin consumption and longer duration of stimulation compared with the GnRH antagonist regimen, the antagonist regimen would perhaps be the suggested regimen for poor responders. Shorter duration of stimulation with the GnRH antagonist compared with the agonist (Pandian et al., 2010) is likely to improve patient compliance.

Although there are several studies comparing GnRH agonist and GnRH antagonist protocols in women with a previous poor response (De Placido et al., 2006; Tazegül et al., 2008; Kahraman et al., 2009; Devesa et al., 2010; Pu et al., 2011), there are very few studies in women with anticipated poor response undergoing their first IVF cycle. In the study by Nelson et al. (2009), treatment with a GnRH antagonist protocol reduced the treatment burden in anticipated poor responders (identified on the basis of low AMH levels <5 pmol/l, DSL assay) but did not vary in other outcomes when compared with a GnRH agonist protocol. The GnRH antagonist protocol was associated with a substantial drop in cycle cancellation [odd ratio (OR) 0.20 (95% confidence interval (CI) 0.06–0.65)] and required fewer days of gonadotrophin administration. This leads to a smaller exogenous gonadotrophin administration. This results in a smaller number of growing follicles when compared with the standard long GnRH antagonist protocol. To summarize, a modified therapeutic protocol should aim to gain patient compliance (Domar et al., 2009, 2010; Lainas et al., 2010; Pu et al., 2011). In conclusion, prediction of poor response can therefore have positive results in terms of patient compliance and reduction of costs. On the other hand however, current evidence suggests that this prediction may not result in a significant improvement of IVF outcome (Louradis et al., 2008; Pandian et al., 2010; Oudendijk et al., 2012).

Not only is prediction of a high response prior to an IVF cycle useful in counselling patients on the risk of OHSS but also gives a real possibility of modifying the stimulation protocol and reduce the incidence of a high response and OHSS. Recent studies have demonstrated that the use of antagonists is associated with a reduction of the occurrence of a high response and a significant reduction in the incidence of OHSS or of cycle cancellation because of the risk of OHSS (Al-Inany et al., 2007, 2011; Hosseini et al., 2010; Lainas et al., 2010; Tehraninejad et al., 2010).

With the GnRH antagonist protocol, initial follicular recruitment and selection is undertaken by endogenous endocrine factors prior to starting the exogenous gonadotrophin administration. This leads to a smaller number of growing follicles when compared with the standard long GnRH agonist protocol and this is undoubtedly an advantage in women with a high ovarian reserve and hence at risk of OHSS. In the study by Nelson et al. (2009), the safety of the GnRH antagonist regimen was superior to the GnRH agonist regimen for the treatment of high responders. The antagonist protocol required fewer days of stimulation than the GnRH agonist protocol (9 days versus 13 days) was associated with elimination of the need for cryopreservation of embryos due to excess response and reduced hospitalization for OHSS (13.9% in the agonist group versus 0.0% in the antagonist group) (Nelson et al., 2009). These results are consistent with those of other studies that have demonstrated a reduced incidence of OHSS with GnRH antagonist protocols compared with agonist protocols in women with PCOS, which can be considered the most significant risk factor for OHSS (Kolibianakis et al., 2006; Lainas et al., 2008). The antagonist protocol in high responders was also associated with significantly higher clinical pregnancy rates (61.7 versus 31.8%, P < 0.05) (Nelson et al., 2009). A recent retrospective study investigated the utility of an AMH-dictated strategy on 769 women undergoing their first IVF cycle (Yates et al., 2011). Women were treated with either a conventional stimulation protocol (n = 346) or a personalized protocol tailored to their AMH levels (n = 423). The study adopted the use of the GnRH antagonist protocol for presumed high responders with an AMH level > 28.6 pmol/l (DSL assay). The cycle cancellation rate due to OHSS was significantly lower with the AMH-tailored protocol (2.3 versus 6.9%, P < 0.05). Furthermore the overall cost for the clinical management of OHSS was reduced by 43% in the AMH group (Yates et al., 2011).

Finally, by using the GnRH antagonist protocol for COS, induction of a LH surge comparable to that occurring physiologically at mid-cycle to trigger ovulation could be obtained by administering a single bolus of GnRH agonist (Griesinger et al., 2006; Humaidan et al., 2011). This regimen may prove highly effective in terms of OHSS prevention (Humaidan et al., 2010a) and gives an additional reason for women who are anticipated to be hyper-responders to be preferentially treated with a GnRH antagonist protocol. To summarize, a modified therapeutic protocol with low gonadotrophin doses and GnRH antagonist seems to be ideal for women at a high risk of OHSS (Fig. 3). Consequently, identification of high responders on the basis of ovarian reserve markers must be considered as invaluable in women undergoing IVF.

**AMH, AFC and Ovarian Response to Exogenous Gonadotrophins**

Although exogenous gonadotrophin administration for ovarian stimulation has been used for decades and millions of cycles have been performed worldwide, criteria to select the ideal gonadotrophin starting dose have not yet been completely identified. In stimulation protocols, exogenous gonadotrophin administration leads to supraphysiological circulating levels of FSH which facilitate recruitment of multiple follicles by exceeding the ovarian FSH sensitivity threshold (Fauser and Van Heusden, 1997; Fleming et al., 2006). When exogenous FSH is administered, the number of mature follicles recruited largely depends upon the number of follicles attaining FSH sensitivity. Hence in women with a large antral follicle pool the administration of a high FSH dose may induce excessive ovarian response consequently leading to a high risk of OHSS. On the other hand, administration of an inappropriately low gonadotrophin dose may lead to the growth of a low number of follicles resulting in an ‘iatrogenic’ poor response with possible negative consequences for the outcome of the cycle, i.e. cycle cancellation for inadequate response, low number of retrieved oocytes, low number of good quality embryos available for transfer. On the contrary, a low functional ovarian reserve would never be compensated by an increase in the exogenous FSH over the maximal dose. Accordingly, different studies performed on women anticipated to be poor responders on the basis of low AMH or AFC showed that increasing the FSH dose was ineffective for preventing a negative ovarian response in those women (Klinkert et al., 2005; Lekamge et al., 2008; Berkkanoglu and Ozgur, 2010).

In a randomized trial (Klinkert et al., 2005), 52 patients with a basal AFC of <5 follicles were randomized to receive either 150 IU (n = 26) or 300 IU (n = 26) recombinant FSH (rFSH) as a starting dose:
The authors found a median number of three oocytes in both groups and most importantly the rate of poor response was similar (63 versus 62% respectively). In a more recent prospective randomized trial, 119 women with an AFC ≤ 12 were randomized to receive 300 IU (n = 38), 450 IU (n = 39) or 600 IU (n = 42) rFSH (Berkkanoglu and Ozgur, 2010). There was no significant difference in the mean number of oocytes retrieved (5.2, 6.3 and 6.6 respectively) nor the cycle cancellation rate (10.5, 15.3 and 14.2%, respectively). In a retrospective study including a total of 122 women aged < 36 years and having likely low ovarian reserve based on a serum AMH measurement below 14 pmol/l (IBC assay), 35 women were administered the standard gonadotrophin dose of 150 IU/day, while the remaining 87 received a higher starting dose of 200–300 IU/day (Lekamge et al., 2007). The mean number of retrieved oocytes was 6.8 and 7 with no significant difference (Lekamge et al., 2007). In conclusion the maximum number of oocytes that could be retrieved in women is strongly limited by the number of recruitable antral follicles in the ovaries and it is obvious that a gonadotrophin dose higher than the maximum will never compensate for the lack of ‘substrate’.

Reproductive medicine clinicians often prescribe the gonadotrophin starting dose based on the woman’s age, increasing proportionally with age. Although a woman’s ability to respond to ovarian stimulation declines with advancing age, age alone is not a reliable indicator of ovarian response (Fauser et al., 2008; La Marca et al., 2010; Broer et al., 2013). Besides, women of similar age may have a wide variability in the pool of recruitable antral follicles (Gougeon, 1998; Almog et al., 2011; La Marca et al., 2011a) thereby questioning the rationale of basing the gonadotrophin dose on age alone. According to data from our centre, the relationship between age and number of retrieved oocytes in IVF cycles has a low correlation (R² = 0.06) meaning that the variability in number of oocytes retrieved can be explained by the age of women in only 6% of cases. If we look at the correlation between number of retrieved oocytes and markers of ovarian reserve, such as FSH, AMH or AFC, this relationship seems much stronger, especially for AMH and AFC which perform much better than FSH. The variability of these two markers explains 22–23% of the variability in the number of retrieved oocytes. The similar performance of AFC and AMH in predicting oocyte yield reflects the strong and similar associations of the two markers with the size of the primordial follicle pool and follicular recruitment rates (Hansen et al., 2011; Kelsey et al., 2012). A large body of evidence clearly indicates that AMH and AFC may be considered interchangeable and globally perform better than all other known markers of ovarian response in IVF (Seifer et al., 2002; van Roeij et al., 2002; Fanchin et al., 2003; Hazout et al., 2004; Muttkrishna et al., 2004; Ficicioglu et al., 2006; Kwee et al., 2007; La Marca et al., 2007a; Mcliveen et al., 2007; Nelson et al., 2007; Elgindy et al., 2008; Jayaprakasan et al., 2008b; Lekamge et al., 2008). Therefore, if AMH and AFC allow a better prediction of the number of oocytes retrieved, we are justified in questioning why a woman’s age alone is the most commonly used sole criterion for choosing the COS protocol in an unknown but surely high number of centres. Probably this is because age is an easy and cheap ‘marker of ovarian reserve’. However, as shown in Table III, although the use of age as a marker of ovarian reserve has several advantages, such as the lack of variability between cycles and the fact that it is an ‘easy marker’, age appears to have the worst performance when it comes to predicting the ‘extremes’ of ovarian response in IVF (namely poor and hyper response).

### Table III Comparison of characteristics of the most widely used markers of ovarian reserve (modified with permission from La Marca et al. (2010)).

<table>
<thead>
<tr>
<th>Characteristics for a Good Marker</th>
<th>Age</th>
<th>AMH</th>
<th>FSH</th>
<th>AFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction of poor response</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Prediction of hyper response</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Low inter-cycle variability</td>
<td>+++</td>
<td>++</td>
<td>−</td>
<td>++</td>
</tr>
<tr>
<td>Low intra-cycle variability</td>
<td>+++</td>
<td>++</td>
<td>−</td>
<td>++</td>
</tr>
<tr>
<td>Applicable to all patients</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Economic</td>
<td>+++</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

−, not appropriate; +, not very appropriate; +++, very appropriate; AFC, antral follicle count; AMH, anti-Mullerian Hormone.
gonadotrophin dose was 150 IU for expected high responders and 300 IU for expected poor responders. This treatment strategy based on the AMH levels led to a reduction of both an excessive response and cancelled cycles (Nelson et al., 2009). A recent retrospective study comparing the study group undergoing a therapeutic protocol based on AMH levels versus a control group undergoing treatment based on pretreatment FSH levels confirmed that tailored treatment based on AMH reduced the incidence of OHSS (Yates et al., 2011). Moreover, the study showed a significant increase in both pregnancy (17.9 versus 27.7%) and live birth rates (15.9 versus 23.9%) in the study group compared with the control group (Yates et al., 2011). This seems to confirm that individualized therapy can improve IVF outcomes. Finally, and not least importantly, the study group also showed an important reduction of costs probably due to a reduced incidence of OHSS and drug consumption (Yates et al., 2011). While the two studies were conceptually similar, they had different discriminating AMH levels for the study groups. For Nelson et al. (2009), women with AMH levels higher than 15 pmol/l (DSL assay) were presumed to be high responders while Yates et al. (2011) considered women with serum AMH levels higher than 28.6 pmol/l (DSL assay) as high responders. Another possible limitation of the two models is that they were based on AMH measured with the ‘old’ DSL assay. As previously stated, the values generated by the current AMH generation II assays are 40% higher than the previous ‘old’ DSL version (Wallace et al., 2011). However, the absence of a linear conversion factor between the two assays makes it necessary to recalculate the AMH categories of ovarian response using the new commercial assay. A first attempt, not yet tested in a specific clinical study, has been carried out recently (Nelson, 2013). Finally, it remains to emphasize the need to validate any proposed models in independent and prospective studies.

**Multivariate models**

The concept of using multivariate models to identify the most appropriate gonadotrophin starting dose for individual women derives from the observations that ovarian response is a complex outcome and different variables may independently contribute to its prediction (Popovic-Todorovic et al., 2003a; Fauser et al., 2008; Al-Azemi et al., 2011). Combining multiple markers for prediction of outcomes is of course not new for our specialty, the classic example of which is prenatal screening for Down’s syndrome (Nicolaides, 2011). A recent individual patient data (IPD) meta-analysis demonstrated that an optimal response prediction was achieved by combining age, AMH and AFC with a statistically significant increase in the area under the receiver operating characteristic curve (from 0.61 to 0.80) when all three variables were used when compared with age alone (Broer et al., 2013). Some initial studies examined and tested complex models based on multiple phenotypic, ultrasound derived and biochemical indices to dictate starting doses of exogenous gonadotrophins in IVF cycles (Popovic-Todorovic et al., 2003a; Howles et al., 2006). An initial prospective study showed that the combination of age, AFC, ovarian volume, Doppler ovarian score and smoking status may allow clinicians to choose the appropriate starting doses of gonadotrophins (Broer et al., 2013).

**Figure 4**

Strategic modelling of controlled ovarian stimulation on the basis of ovarian reserve markers. The introduction of individualized AMH-tailored controlled ovarian stimulation utilizing agonist and antagonist protocols has been reported as associated with improved IVF cycle, i.e. increased pregnancy rate. Similarly a reduction in the incidence of adverse outcomes, such as OHSS, has been reported (modified with permission from Nelson et al. (2009) and Yates et al. (2011). (AMH was measured with the DSL assay). AMH; anti-Mullerian Hormone.
gonadotrophin dose in IVF cycles (Popovic-Todorovic et al., 2003a). In a subsequent study the proposed model was based on age, BMI, Day 3 serum FSH and AFC (Howles et al., 2006). Both models were validated in successive prospective trials and demonstrated that the application of an individualized versus standard gonadotrophin dose was associated with a reduced cancellation rate for abnormal ovarian response, a reduced need for adjusting the dose during treatment and increased occurrence of an adequate ovarian response (Popovic-Todorovic et al., 2003b, 2004; Olivennes et al., 2009). While appearing to be useful, both models were rather complex and have not had a wide clinical application. The algorithm by Popovic-Todorovic et al. (2003a) incorporated variables such as Doppler score of ovarian stromal blood flow that are not commonly measured in daily clinical practice. The model created by Howles et al. (2006) and later tested in the CONSORT study (Olivennes et al., 2009) predicted gonadotrophin starting doses that were relatively low compared with routine practice; for 48 of the 161 women recruited the calculated gonadotrophin dose was 75 IU/day and as would be expected a high proportion of these women (25%) did not reach oocyte retrieval as a consequence of an inadequate ovarian response (Olivennes et al., 2009). Most importantly the coefficients for computing the algorithm were not published and hence the formula cannot be used by clinicians in their daily clinical practice.

An easy to use algorithm to calculate the gonadotrophin dose based on AFC has recently been published (La Marca et al., 2013). This model, although interesting, requires validation in an independent cohort as it is based on a retrospective analysis. The multivariate regression analysis showed that independent predictors of ovarian response expressed in terms of retrieved oocytes were age, AFC and Day 3 serum FSH, with AFC being the most significant predictor. The nomogram calculated the gonadotrophin dose based on the age of the

![Figure 5](https://academic.oup.com/humupd/article-abstract/20/1/124/887248/7268) Nomogram for calculation of the FSH starting dose based on age, AFC and Day 3 serum FSH. In the example, for a 30-year-old woman with AFC = 16 and d3FSH = 4 IU/l, the FSH starting dose is 152 IU/day. Since the new FSH delivery system will have the dosage dial based on doses of FSH of 12.5 IU, on the right side of the FSH starting dose column, the FSH dose as selected for the delivery system is reported (150 IU/day, for example), (from La Marca et al. (2013) with permission). AFC; antral follicle count.
woman, Day 3 serum FSH level and AFC. For example in a woman aged 30 years, with a Day 3 FSH of 4 IU/l and an AFC of 16 the most appropriate gonadotrophin dose is 150 IU daily (Fig. 5). Overall this model predicted a daily dose of 12.5 IU gonadotrophin in 50.2% of women aged ≤35 years (12.8, 16.9 and 20.5% of women had a predicted daily dose of 150, 150–187.5 and 187.5–225 IU, respectively). The percentage of women aged >35 years with a predicted dose >225 IU was 18.1% (3.2, 7.8 and 7.1% of women had a predicted daily dose of <150, 150–187.5 and >187.5–225 IU, respectively).

A similar nomogram based on AMH had previously been developed by the same group (La Marca et al., 2012b). The choice of developing two different nomograms based on AMH or AFC followed the recognition that clinicians usually rely on measuring one marker of ovarian response, either AFC or AMH. As largely discussed throughout this review, clinicians may use one of the two markers since they can be considered interchangeable. The multivariate AMH based model was developed on 346 women undergoing ovarian stimulation with the same protocol (the long GnRH agonist standard protocol) and the same dose of gonadotrophin. The variables analysed as possible predictors of ovarian response to stimulation were Day 3 serum FSH, estradiol, AMH, BMI and smoking status. A multivariate regression analysis showed that independent predictors of ovarian response, expressed in terms of retrieved oocytes, were age, AMH and Day 3 serum FSH with AMH being the most significant predictor. According to the model, for women of similar age, the number of retrieved oocytes per unit of gonadotrophin was reduced with decreasing levels of basal AMH and increasing levels of Day 3 serum FSH. The multivariate model was the basis of a nomogram for the selection of the most appropriate gonadotrophin dose.

**Figure 6** The nomogram for the calculation of the FSH starting dose based on age, serum AMH and FSH. In the example, for a 30-year-old woman with serum AMH level of 4 ng/ml and FSH level of 4 IU/l, the FSH starting dose is 152 IU/day. Since the new upcoming FSH delivery system will have the dosage dial based on doses of FSH of 12.5 IU, on the right side of the FSH starting dose column, the FSH dose as selected for the delivery system is reported (150 IU/day for the example). (AMH was measured with the IBC assay. AMH conversion factor: 1 ng/ml = 7.143 pmol/l) (from La Marca et al. (2012b), with permission). AMH, anti-Mullerian Hormone.
appropriate gonadotrophin starting dose (Fig. 6). Based on the nomogram a woman aged 30 years with a Day 3 serum FSH of 4 IU/l and AMH of 4 ng/ml would require a gonadotrophin dose of 150 IU/daily (Fig. 6). As with the AFC-based nomogram, the model incorporating AMH needs to be validated in an external and independent population before adoption into routine clinical practice. However it is almost intuitive that in the future several multivariate prediction models will be available to guide decisions in an individualized approach. The two proposed nomograms may be considered as a first step on this path.

Conclusions
After decades of practice using IVF, it is now very clear that the ‘one size fits all’ approach may no longer exist. Individualization of treatment is not new to the field of medicine, although this concept is relatively fresh in reproductive medicine. The availability of new markers of ovarian reserve, the improvement in methodology for their measurement and the huge amount of clinical data have supported the view that individualization in IVF is the way forward. Ovarian response in IVF is a complex puzzle for which we now know the most important pieces. The correct measurement of markers of ovarian reserve allows a scientific estimate of the pool of follicles that potentially respond to ovarian stimulation. Published studies indicate an important role for both AFC and AMH in the prediction of the extremes of ovarian response and for enabling the subsequent individualization of a therapeutic strategy. This is the basis for the correct selection of women for use of the different GnRH analogues and, for the fine tuning of the gonadotrophin dose. The ultimate goal would be the selection of an effective protocol for ovarian stimulation which has to be well balanced between the risk of maximal and suboptimal ovarian response. The benefits of a personalized therapy may include reduced incidence of risks and dropout as well as a reduced treatment burden. Nevertheless, a clear definition for modality of a correct application of the individualized therapy is still required to optimize efficacy and daily clinical management.

Authors’ roles
All authors (A.L.M. and S.K.S.) participated in the design, bibliographic search and selection, data analyses, manuscript writing and final manuscript approval.

Funding
This work has been supported by a grant of the Italian Ministry of Health (grant no. GR-2009-1580036).

Conflict of interest
None declared.

References


Buyuk E, Sefer DB, Younger J, Grazi RV, Lieman H. Rapid anti-Mullerian hormone (AMH) is a predictor of ovarian response in women with elevated baseline early follicular follicle-stimulating hormone levels. Fertil Steril 2011;95:2369–2372.


Ng EH, Tang OS, Ho PC. The significance of the number of antral follicles prior to stimulation in predicting ovarian responses in an IVF programme. Hum Reprod 2000;15:1937–42.


Yong PY, Baird DT, Thong KJ, McNelly AS, Anderson RA. Prospective analysis of the relationships between the ovarian follicle cohort and basal FSH concentration, the inhibin response to exogenous FSH and ovarian follicle number at different stages of the normal menstrual cycle and after pituitary down-regulation. Hum Reprod 2003;18:35–44.