Predictors of ovarian response: progress towards individualized treatment in ovulation induction and ovarian stimulation


1Department of Reproductive Medicine and Gynecology, University Medical Center, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands; 2Department of Obstetrics and Gynecology, University Clinic of Schleswig-Holstein, Campus Luebeck, 23538 Luebeck, Germany; 3Center of Reproductive Medicine, Free University Brussels, Brussels, Belgium

4Correspondence address. E-mail: b.c.fauser@umcutrecht.nl

Ovarian stimulation is applied in the clinic to restore mono-ovulatory cycles in anovulatory women (ovulation induction) or to induce the development of multiple dominant follicles for assisted reproduction. Ovarian response is the endocrine and follicular reaction of the ovaries to stimulation. Achieving an appropriate ovarian response to anti-estrogens or exogenous gonadotrophins is central to ovulation induction and ovarian stimulation protocols. However, achieving an adequate response, without cycle cancellation or adverse events related to under- or over-stimulation, is complicated by high intra- and inter-individual variability. To predict each patient’s ovarian response to medication for ovarian stimulation and to individualize the starting dose of exogenous gonadotrophin or the need for exogenous luteinizing hormone, various clinical, endocrine, ovarian ultrasonographic and genetic characteristics have been explored. Some of these features have been incorporated into prediction models. In this review, the methodology behind predictive factors and prediction models and their potential clinical applicability across ovulation induction and ovarian stimulation are explored.

Keywords: predictive factors; predictive models; ovulation inductions; multifollicular stimulation

Introduction

Ovarian response can be defined as the endocrine and follicular reaction of the ovaries to a stimulus. The term ovarian response is used in clinical research and practice both qualitatively (e.g. achieving growth of a single-dominant follicle and ovulation in anovulatory women undergoing ovulation induction) and quantitatively [e.g. the extent of multifollicular development in ovulating women undergoing ovarian stimulation for in vitro fertilization (IVF)]. Achieving a distinct ovarian response usually represents the desired outcome of pharmacological interventions on the hypothalamic–pituitary–ovarian axis in ovulation induction and ovarian stimulation. The considerable individual variability in ovarian response to stimulation, however, necessitates close monitoring and dose adjustment for each patient.

In contrast, ovarian reserve refers to whatever remains of the ever-declining pool of primordial follicles in the ovaries at a given time point and the reproductive potential of each oocyte. Ovarian reserve thus reflects the reproductive age of an individual woman (Broekmans et al., 2007). Declining ovarian reserve has been suggested as a cause of the decrease in live birth rate that occurs after natural conception at ~31 years of age, and at ~35 years in IVF cycles (van Noord-Zaadstra et al., 1991; Templeton et al., 1996). Although ovarian reserve is likely to be linked to the ovarian response to exogenous stimulation, to date, it is unclear whether a linear relationship exists or whether ovarian response declines only once ovarian reserve falls below a distinct threshold level.

It is important to acknowledge that a strong inter-individual variability for ovarian reserve exists within the same chronological age group. In addition, results of ovarian reserve tests show not only inter-individual variability but also considerable intra-individual variability (Scott et al., 1990; Scheffer et al., 1999, 2002; Hansen et al., 2003; Kwee et al., 2004; Elter et al., 2005). Finally, the likelihood of pregnancy in a woman undergoing ovulation induction or ovarian stimulation is subject to a large number of factors other than ovarian reserve and ovarian response.

Nevertheless, it is of high clinical relevance to identify predictors of ovarian response that will enable clinicians to individualize...
ovulation induction and ovarian stimulation treatment, thereby minimizing complications and the risk of treatment failure while maximizing the chance of ongoing pregnancy. The conventional paradigm in many areas of reproductive medicine has been ‘one size fits all’ or a choice of therapy based on physicians’ experience from their own clinical practice, which may have low reproducibility (Wiegerinck et al., 1999). To improve consistency between clinics, various clinical, endocrine and ovarian ultrasonographic and genetic characteristics have been explored for use as predictors of ovarian response (van Santbrink et al., 2005). However, ‘the use of observed relationships to make predictions about individuals is an area with many pitfalls; just as it is dangerous to generalize from the particular, we must be very careful about particularizing from the general’ (Altman, 1980).

This review will evaluate the clinical applicability of predictive factors and predictive models across different clinical issues in ovarian stimulation, from anti-estrogens as first-line therapy in ovulation induction, to the use of gonadotrophins in mono- or multifollicular stimulation protocols. It will appraise whether these models can improve the safety and efficacy of treatment.

**Prediction Factors and Models**

A prediction model is, by definition, used to predict a particular outcome given the presence of a variety of independent variables. Prediction models are built over three stages: the first stage is to define the predictive factors, the second to form the model and the third to validate the model. To test which variables are predictive requires large prospective exploratory studies in which the patient is observed until the outcome occurs; only this design ensures absence of measurement bias, as the data are collected before the outcomes are known and will not influence the clinical management that leads to the outcomes (Enskog et al., 1999). Variables that may be identified by clinical consensus or univariate analysis are then built into the prediction model using regression analysis (linear regression for continuous outcome data, logistic regression for dichotomous data or proportional hazards analysis, also known as Cox regression, for time-to-event data). The model can be validated internally by split sample or boot strapping in the cohort from which it was developed. External validation is preferable, using data from a new but similar group of patients, usually from another treatment centre. The internal validity of the model (apparent validity) is not generalizable and hence is not as useful as external validity derived from testing in a separate cohort.

The validation procedures test for precision [how narrow are the 95% confidence intervals (CIs)], for reliability or calibration (how well does the prediction agree with the observed events) and for discrimination or accuracy (how well does the prediction model distinguish between patients who do or do not have events). The discriminatory ability of the model is assessed by the area under the receiver operator characteristics curve (ROC-AUC) or the $c$-statistic which has a value ranging from 0.5 (no discriminating ability) to 1.0 (perfect discrimination). The ROC-AUC and $c$-statistic are virtually synonymous and describe how much of the known and unknown variability in the event of interest is accounted for by the model. This corresponds to the $R^2$ from a linear regression, which gives the percentage of the variability in a dependent variable which is explained by an independent variable or set of variables (Harrell et al., 1996). In the following studies, the ROC-AUC is by far the most frequently used quality measure.

This process of identifying predictive factors, constructing a prediction model and validating the model are clinically relevant only to the extent that the model can be applied in clinical practice. One approach is to convert the regression coefficients into a simple and memorable score that can be used in a clinical calculator or a nomogram, so that the physician can input a given patient’s characteristics and estimate his/her individual prognosis. A second approach is to enter the variables into a computer programme, allowing the use of a more complex algorithm without requiring the physician to make any calculations. For use in every-day clinical practice, an approach i.e. easy to use would have an advantage over any complicated or unwieldy system.

**Predictors of Ovarian Response in Ovulation Induction**

**Predicting Response to Anti-estrogen Therapy**

Anovulation is a common cause of infertility and is present in at least a quarter of couples facing conception difficulties (Smith et al., 2003). In many patients, induction of ovulation with anti-estrogen therapy continues to be first-line therapy. Anti-estrogen is effective in inducing ovulation in 73% of women treated, giving a live birth rate of ~29% (pooled results from 5268 women) (Homburg, 2005). Using the best evidence to identify patients who will remain anovulatory despite anti-estrogen therapy can direct these patients towards alternative treatment approaches such as exogenous gonadotrophins, laparoscopic ovarian surgery, insulin sensitizing agents (Legro et al., 2007) or more complex assisted reproductive technology (ART) procedures, especially in women of advanced reproductive age. Furthermore, the process of identifying prognostic factors also provides an insight into ovarian abnormalities and the pathophysiology of anovulation. Three models that have been developed to predict the chances of success with anti-estrogen-induced ovulation in women with World Health Organization (WHO) group II infertility are described below (Table 1), together with a nomogram that combines predictive factors from two of the models (Fig. 1).

The predictive value of baseline characteristics was investigated in a prospective study of 201 women with WHO II anovulatory infertility, who underwent 432 cycles of clomiphene citrate (CC) ovulation induction, with all but 45 achieving ovulation (Imani et al., 1998). The most predictive characteristics were the free androgen index (FAI, calculated from the testosterone to sex hormone-binding globulin ratio) and body mass index (BMI), with AUCs of 0.76 and 0.70, respectively. Entering FAI, BMI, ovarian volume and cycle history (oligomenorrhea versus amenorrhea) into a regression model achieved a fairly accurate prediction, with an overall AUC of 0.82. By scoring each characteristic based on its value at screening (e.g. a patient with a BMI over 35 kg/m$^2$ would gain 15 points for that characteristic, whereas a patient whose BMI was <20 kg/m$^2$ would gain no points) a total score can be calculated. A higher score would predict a greater chance of that patient remaining anovulatory (Imani et al., 1998). Although this model had a moderately good predictive power the requirement to assess FAI has limited its use, as this variable is not commonly measured.
Table 1: Prediction models for treatment response in ovulation induction

<table>
<thead>
<tr>
<th>Treatment (study)</th>
<th>Outcome</th>
<th>Patients (n, achieving outcome/total in study)</th>
<th>Predictive factors</th>
<th>AUC/c-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clomiphene citrate (Imani et al., 1998)</td>
<td>Ovulation</td>
<td>156/201</td>
<td>Amenorrhoea, BMI, FAI</td>
<td>0.82</td>
</tr>
<tr>
<td>Clomiphene citrate (Imani et al., 1999)</td>
<td>Pregnancy</td>
<td>73/159</td>
<td>Age, oligomenorrhoea, IGF-I, testosterone, age</td>
<td>AUC not calculated</td>
</tr>
<tr>
<td>FSH (Mulders et al., 2003a)</td>
<td>Ongoing pregnancy</td>
<td>57/154</td>
<td>Oligomenorrhoea, FAI, duration of infertility</td>
<td>0.72</td>
</tr>
<tr>
<td>FSH (van Wely et al., 2005)</td>
<td>Ongoing pregnancy</td>
<td>57/85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clomiphene citrate/FSH (Eijkemans et al., 2003)</td>
<td>Live birth</td>
<td>134/240</td>
<td>Age, insulin:glucose, duration of infertility</td>
<td>0.61</td>
</tr>
</tbody>
</table>

The two models described above are designed to predict the chances of a woman failing to ovulate after anti-estrogen treatment. A model that predicts the chances of conception in women in whom ovulation is induced is the next step in predicting outcome for individual patients. In a proportional hazards analysis, the patient’s age and her cycle history were the only factors identified as predictors of time to conception (Imani et al., 1999). The disparity between the characteristics predictive of conception and the characteristics previously shown to be predictive of ovulation (body weight and hyperandrogenaemia) is most probably because ovarian response is only one of many variables associated with pregnancy likelihood. However, this observation also raised an interesting hypothesis. These results suggested that the regulation of endogenous follicle-stimulating hormone (FSH) to stimulate follicle growth and ovulation may differ from the regulation of endogenous FSH needed to ensure oocyte quality. It is the latter threshold that predicts the chances for conception in ovulatory cycles (Imani et al., 1999). Combining prediction models for success in ovulation induction and success in conception would allow prediction of the likelihood of conception before anti-estrogen therapy is initiated, allowing patients with a low percentage chance of a live birth to be directed towards another first-line treatment modality. This has been achieved through use of an integrated double nomogram that uses the predictive factors for anovulation (Imani et al., 1998) in one section and those for pregnancy (Imani et al., 1999) in another section. Although the nomogram was based on these earlier studies, it was tailored for use in clinical practice by including only characteristics that are routinely measured (Imani et al., 2000). The investigators found a negative predictive value of 80% (95% CI: 60–99%), indicating that the nomogram could

Figure 1: Nomogram to calculate the probability of ovulation and conception resulting in a live birth within 6 months of starting clomiphene citrate treatment. In the first step, ovulation is predicted from the patient’s FAI, BMI and cycle history. This result is then transposed to the second half of the nomogram, where the patient’s age and cycle history are both plotted. The resulting line transects a point on the scale showing the percentage chance of conception within 6 months of clomiphene citrate treatment leading to a live birth. For example, a 29-year-old woman with amenorrhoea, a FAI of 9.3 [testosterone x 100/sex hormone-binding globulin (SHBG)], and a BMI of 32 kg/m² has a 50% chance of ovulating and a 19% chance of pregnancy according to the nomogram. This figure was published in Imani et al. (2002b); copyright Elsevier 2002.
identify 80% of non-responders to clomiphene citrate; nevertheless, they considered it insufficiently accurate for clinical use (Ghobadi et al., 2007).

**Predicting Response to Gonadotrophins**

Gonadotrophins are commonly used as a second-line treatment to restore ovarian function in patients with WHO group II anovulation who have not responded to anti-estrogen therapy. Models have been developed to predict: the chances of pregnancy in women using clomiphene citrate therapy first line and gonadotrophin therapy second line (Eijkemans et al., 2003), ovulation in women in whom clomiphene citrate has failed (Mulders et al., 2003a), ovulation in women with PCOS (van Wely et al., 2005) and the gonadotrophin dose threshold (Imani et al., 2002a). These models are discussed below (Table 1).

To predict which patients with WHO group II infertility will not achieve pregnancy through first-line clomiphene citrate and second-line gonadotrophin treatment, 240 women were prospectively followed through clomiphene citrate and, if necessary, gonadotrophin treatment (Eijkemans et al., 2003). Predictor variables were entered into a Cox regression analysis to construct a multivariate prediction model. The final model included three variables that were negatively correlated with pregnancy at 12 months leading to a singleton live birth: the age of the woman, the insulin-to-glucose ratio and the duration of infertility. The c-statistic for the model was 0.61 (optimism-corrected) indicating only a moderate ability to discriminate between outcomes. To use this model in clinical practice, physicians would need to arbitrarily select the most appropriate cut-off for their clinical setting, offering patients an alternative first-line treatment for which the chances of success were only 10%, 20% or whatever level they considered acceptable. If a 30% chance of success is taken to represent a poor prognosis, the model predicted that 25 of 240 patients (10%) would be beneath this cut-off (Eijkemans et al., 2003).

The above model predicts success for patients from when they start clomiphene citrate therapy, which, for some, will lead to therapy with gonadotrophins. If the patient has already shown resistance to clomiphene citrate-induced ovulation, it is appropriate to assess her chances of success using a model specific for gonadotrophin induction of ovulation in patients with clomiphene citrate-resistant anovulation. Furthermore, to a far greater extent than with clomiphene citrate treatment, failure with gonadotrophins includes the failure to control ovulation leading to hyper-response, as well as the failure to induce ovulation. Predicting the chance of hyper-response is important in limiting cycle cancellations or, more rarely, ovarian hyperstimulation syndrome (OHSS).

To predict the individual outcome of ovulation induction with gonadotrophins in women for whom clomiphene citrate induction of ovulation was unsuccessful, a model has been developed based on characteristics at screening. Women (n = 154) who underwent a total of 544 gonadotrophin cycles in a prospective follow-up study formed the cohort for the model; the first cycle always followed a low-dose step-up protocol; the second cycle followed a step-down protocol (Mulders et al., 2003a). The factors identified as most strongly predictive of ongoing pregnancy were the woman’s age, testosterone concentration and insulin-like growth factor-I (IGF-I) levels. For this multivariate model, however, the AUC was only 0.67 (Mulders et al., 2003a). Factors most predictive of multifollicular growth were androstenedione concentration and the number of ovarian follicles (AUC 0.62). A separate study in patients with PCOS found that oligomenorrhoea, shorter duration of infertility and a lower FAI were associated with a higher chance of ongoing pregnancy (van Wely et al., 2005). The predictive model had a moderate discriminatory power (AUC 0.72). This allowed women with a ≤5% probability of attaining an ongoing pregnancy to be distinguished from those with a ≥25% chance.

The correct balance between under- and over-stimulation with gonadotrophins can be difficult to achieve because of the wide inter-individual variation in the dose of exogenous FSH required to induce ongoing follicle development (the FSH threshold). Two strategies are employed in achieving this balance: in the chronic low-dose step-up regimen, the dose is progressively increased from a low starting point. The limitation of this regimen is that in some women the threshold dose necessary to induce ongoing follicular growth may be reached only after prolonged treatment. In the step-down regimen, the patient starts treatment with a high dose, which progressively decreases over the following days. A declining dose is a better approximation of the normal physiological pattern of FSH exposure than an increasing dose; however, the high initial doses in the step-down approach can trigger an immediate hyper-response in some women (van Santbrink and Fauser, 2003). An alternative and potentially more successful approach than step- or step-down dosing would be a prediction of each patient’s individual FSH dose threshold using the carefully analysed experience of many women. The woman’s age is one of several factors that predict gonadotrophin success; other factors are summarized below.

Imani et al. (2002a) have developed a model to predict a woman’s FSH dose threshold from characteristics measured at screening and during cycle monitoring. In this prospective cohort study, normogonadotrophic, anovulatory women received daily exogenous FSH in a low-dose, step-up regimen (from 75 IU/day with weekly increments of 37.5 IU/day). The FSH dose threshold was defined as the FSH dose on the day that follicle growth exceeded 10 mm in diameter. Multiple regression analysis model of the association between clinical characteristics and FSH dose was: \[ \text{FSH} = [4 \times \text{BMI} (\text{kg/m}^2)] + [32 \times \text{clomiphene citrate resistance (yes = 1 or no = 0)] + [7 \times \text{initial free IGF-I (ng/ml)]} + [6 \times \text{initial serum FSH (IU/L)]} - 51 \]. The accuracy of the model was expressed by \( R^2 \), with a value of 0.54, and the average error in dose prediction was 31 IU (Imani et al., 2002a). To make the model easier to use, free IGF-I was substituted for insulin-to-glucose ratio, which is more often measured in clinical practice. The \( R^2 \) decreased from 0.54 to 0.49 indicating that the modified model explained ~49% of the variability in FSH dose.

This model was also validated externally. The cohort of women in the external validation (n = 85) had PCOS and none had ovulated with clomiphene citrate treatment (some women in the development cohort had ovulated but failed to conceive with clomiphene citrate treatment). The clinical characteristics of the two populations were similar, with the exception of more pronounced hyperandrogenism in the PCOS validation population (van Wely et al., 2006). The model overestimated the FSH threshold dose by 25 IU on average in the validation cohort, with higher discrepancies at higher predicted doses. Prescribing a dose higher than the stimulation threshold may lead to cycle
cancellations through over-stimulation. The $R^2$ of the model in the test cohort was 0.11, meaning that it could explain only 11% of the variation in FSH dose threshold between women (van Wely et al., 2006). This emphasizes the necessity to validate a model before routine clinical application and, furthermore, it implies that the external validity of a model will depend on how closely the external validation cohort resembles the original development cohort of patients.

In other studies not designed to develop prediction models, various characteristics have been identified that are associated with a good response to gonadotrophins. For example, women with small ovaries respond better to ovulation induction with gonadotrophins, but their likelihood of conceiving is similar to that seen in women with larger ovaries (Lass et al., 2002). To identify predictive factors that are common to all studies, a systematic review and meta-analysis assembled data from earlier studies of gonadotrophin ovulation induction in women with WHO group II anovulation. The combined results of 13 eligible studies suggested that obesity and insulin resistance are both associated with adverse outcomes, including increased total dose of FSH administered, cancelled cycles, and decreased ovulation and pregnancy rates (Mulders et al., 2003b). These predictive factors would need prospective validation before use in clinical practice.

**Predictors of Ovarian Response in Multifollicular Stimulation**

Multifollicular stimulation is used for both intrauterine insemination (IUI) and IVF/intracytoplasmic sperm injection (ICSI) protocols. The type of procedure determines the ideal number of mature follicles achieved through stimulation: from 2 to 3 follicles for IUI to ~10 follicles for IVF/ICSI. This, in turn, determines the dose of gonadotrophin used: just above the follicular-response threshold for IUI procedures but in excess of this threshold for IVF/ICSI procedures. These differences mean that results from prediction studies based on ovarian stimulation for IVF/ICSI may not be valid for IUI.

To our knowledge, only two studies have analysed predictive factors for ovarian response to gonadotrophin therapy in IUI protocols. The retrospective study of Ng et al. (2005) in women using menopausal gonadotrophin (HMG) for first-cycle IUI, found that BMI was the only significant parameter that predicted the number of follicles $>14$ mm in diameter, whereas antral follicle count (AFC) was the only significant predictor of the duration of stimulation. In a similar but prospective study of low-dose FSH stimulation for IUI, Freiesleben et al. (2006) found that among the nine parameters investigated, body weight and AFC were significant independent predictors of the number of mature follicles. The following sections address how ovarian response is predicted in women undergoing stimulation for IVF/ICSI.

**Predicting hyper-response**

Severe OHSS is the most serious iatrogenic complication of multifollicular ovarian stimulation. It is thought to follow from a series of events that are triggered by human chorionic gonadotrophin (hCG). Through the release of various mediators, vascular permeability is increased and fluid is lost into the third space (Rizk and Smitz, 1992). OHSS that presents after 9 days of hCG reflects endogenous stimulation from pregnancy and is likely to be more severe and of longer duration than early OHSS (Mathur et al., 2000). Depending on the timing of presentation, cycle cancellation (withholding hCG) may be necessary. Fortunately, severe OHSS has a low prevalence, affecting 0.5–5% of women (Delvigne and Rozenberg, 2002; Abouighar and Mansour, 2003).

Factors associated with a hyper-response and an increased risk of OHSS include patient history (Abouighar and Mansour, 2003), the presence of PCOS, younger age and lower BMI (Danninger et al., 1996; Enskog et al., 1999). The most important clinical predictor of severe OHSS is PCOS (Rizk and Smitz, 1992). In a systematic review that included 10 studies there was a ‘significant and consistent’ relationship between polycystic ovaries and OHSS (Tummon et al., 2005). The down-regulation protocol for ovarian stimulation also appears to influence the risk of OHSS. Switching from a gonadotrophin-releasing hormone (GnRH) agonist to an antagonist ovarian stimulation protocol may be beneficial in reducing the incidence of OHSS (Ragni et al., 2005; Al-Inany et al., 2006).

Identifying hyper-responders at an early stage of the stimulation phase would allow adaptation of the stimulation protocol to minimize potential complications. However, studies of endocrine, follicular and ovarian reserve tests have given disappointing results. Estradiol ($E_2$) is the best defined endocrine predictor for OHSS as the cascade of events that leads to the development of OHSS is almost always accompanied by elevated $E_2$ levels (Danninger et al., 1996; Abouighar, 2003; Miao and Huang, 2006). However, Hendriks et al. found that acceptable specificity with moderate sensitivity was achieved only at higher cut-off levels of $E_2$ for predicting both hyper-response and extreme response (Table 2). The authors concluded that the modest sensitivity and high false-positive rate limits the clinical value of $E_2$ (Hendriks et al., 2004). The results suggest that low $E_2$ levels in the late follicular phase may be a result of highly suppressed luteinizing hormone (LH) concentrations, without necessarily signifying a lower risk of OHSS. The study of Papanikolaou et al. (2006) also found high levels of $E_2$ to be unreliable in predicting risk of OHSS, but found follicle number to be significantly better ($P = 0.001$) (Table 2). A threshold of $\geq 13$ follicles (diameter $\geq 11$ mm) on the day of hCG would have predicted 100% of early OHSS and 87% of severe cases. However, there was only a low probability that OHSS was present when the test was positive. Among dynamic tests, neither the exogenous FSH ovarian reserve test (EFORT) nor the clomiphene citrate challenge test (CCCT) is adequate alone to predict hyper-response (Kwee et al., 2006). For hyper-response, the inhibin B increment in the EFORT was the best predictor, but had a low maximal accuracy of 0.78. Multiple logistic regression analysis did not produce a better prediction (Kwee et al., 2006).

Anti-Mullerian hormone (AMH) may also be a marker for patients at risk for OHSS. Baseline pretreatment serum levels of AMH in 16 patients who experienced OHSS were found to be 6-fold higher than in normal age- and weight-matched controls ($P = 0.0036$) (Nakhuda et al., 2006). AMH belongs to the transforming growth factor-β superfamily (Josso et al., 2001) and is expressed in the granulosa cells from follicles at the pre-antral and small antral stage (Durlinger et al., 2002; Weenen et al., 2004). Of great interest is the stability of this new marker, which appears not to fluctuate in concentration during the
menstrual cycle (La Marca et al., 2006). When serum levels of AMH were determined in 48 women, on any day of their menstrual cycles, all cycles that were cancelled due to absent response were in women whose AMH level was in the lowest quartile (≤0.4 ng/ml); in contrast, all cycles that were cancelled because of a risk of OHSS were in women whose AMH was in the highest quartile (>7 ng/ml) (Hehenkamp et al., 2006; La Marca et al., 2007). AMH could be the first serum marker of ovarian response that can be measured on any day of the menstrual cycle (La Marca et al., 2007; Seifer and Maclaughlin, 2007). Many other potential predictors of hyper-response have been investigated, such as total ovarian volume (Oyesanya et al., 1995), interleukin-10 (Enskog et al., 2001), vascular endothelial growth factor (Ludwig et al., 1999) and inhibins (Baird and Smith, 1993; Miao and Huang, 2006). However, to enable such associations to be clinically useful these characteristics need to be easily and reliably assessed in clinical practice and the associations need to clearly discriminate between normal and hyper-responders.

Predicting hypo-response

The ideal ovarian reserve test would reliably measure the quantity of the primordial follicle pool and reflect the overall quality of its oocytes. In reality, ovarian reserve tests provide an impression of the cohort of recruited antral follicles appearing in the FSH window at the start of each cycle (Fig. 2) (Fauser and Van Heusden, 1997; McGee and Hsueh, 2000). The relation between test results and true ovarian reserve is unknown but is probably moderate for the quantitative aspect and low for the qualitative aspect of ovarian reserve. Both quantity and quality of follicles are difficult to establish as the development from primordial follicles into antral follicles takes at least 6 months, during which time the morphology, endocrine responsiveness and steroidogenic activity develops (Gougeon, 1998; McGee and Hsueh, 2000). Ovarian reserve tests assess the number of recruited follicles, either directly through the AFC or indirectly through other assays, such as FSH.

Predictors for ovarian reserve in ART fall into the categories of clinical predictors (age, BMI and the cause of infertility) and predictive tests. Currently available and applied tests are either ultrasonographic (AFC, ovarian volume, ovarian blood flow), endocrine (early follicular phase serum FSH, E2, inhibin B, AMH) or dynamic (CCCT, EFORT, gonadotrophin agonist stimulation test). Broekmans et al. (2006) have systematically reviewed all of the currently available tests, calculating the ROC for each and expressing how likely a given test result is using likelihood ratios (LRs). The LR of an abnormal test result (LR+) is equivalent to the true-positive rate divided by the false-positive rate (sensitivity/(1 – specificity)). LR+ or ratios of true- to false-positive rates from 5 to 10 are considered moderately useful. The LR of a normal test result (LR−) is (1 – sensitivity/specificity) or (false-negative rate/true-negative rate), and values of 0.2–0.1 are considered moderately useful.

The authors showed that the accuracy of known ovarian reserve tests for predicting poor ovarian response to ovarian stimulation is modest, and that none of the tests are accurate predictors of pregnancy. Of all the tests, AFC and basal FSH had the best sensitivity and specificity for predicting ovarian response (Fig. 3) (Broekmans et al., 2006). If the prevalence of a poor response was 20%, an AFC LR+ of ~8 would imply a post-test probability of poor ovarian response around 67%, which would make the AFC test a clinically valuable test, but this LR+ is associated with such a low number of antral follicles that it would be found in only 12% of patients. For FSH, an LR+ of about 8 in a clinical setting where

<table>
<thead>
<tr>
<th>Study</th>
<th>Patients (n)</th>
<th>Predictive factors</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hendriks et al. (2004)</td>
<td>108 (first IVF treatment)</td>
<td>E2 concentrations on day 3&lt;br&gt;E2 concentrations on day 5</td>
<td>0.75&lt;br&gt;0.81 for extreme response&lt;br&gt;0.81</td>
</tr>
<tr>
<td>Papanikolaou et al. (2006)</td>
<td>1801 (2524 IVF cycles)</td>
<td>E2 on day of hCG (data given for optimal threshold of 2560 ng/l)&lt;br&gt;AFC on day of hCG (data given for optimal threshold of ≥13 follicles of ≥11 mm diameter)&lt;br&gt;E2 on day of hCG (5000 ng/L) AFC on day of hCG (≥18 follicles of ≥11 mm diameter)</td>
<td>0.680 (53% sensitivity, 77% specificity)&lt;br&gt;0.823 (86% sensitivity, 69% specificity)&lt;br&gt;83% sensitivity, 84% specificity</td>
</tr>
<tr>
<td>Kwee et al. (2006)</td>
<td>10 (first IVF treatment)</td>
<td>CCCT before stimulation started&lt;br&gt;EFORT (inhibin B increment) before stimulation started</td>
<td>0.82&lt;br&gt;0.92</td>
</tr>
</tbody>
</table>

Extreme response was defined as cancellation of the cycle or OHSS.
the prevalence of a poor response was 20% would imply a post-test likelihood of about 67%, but this LR+ implies a high basal FSH level that would occur in only 1% of patients. AFC and FSH may be replaced over the next few years by AMH as a factor predictive of a poor response (La Marca et al., 2007; Seifer and Maclaughlin, 2007). However, more evidence is required (Broekmans et al., 2006).

Overall, Broekmans et al. (2006) concluded that ovarian reserve tests had a modest clinical utility because of their limited predictive properties and advised that such tests should not be used routinely in all patients. The authors commented that ‘if a high threshold is used, to prevent couples from wrongly being refused IVF, a very small minority of indicated cases (~3%) were identified as having unfavourable prospects in an IVF treatment cycle (pregnancy rate for that cycle of 5%).’ Indeed, even when the LR+ of 8 is the cut-off for treatment, for every eight couples correctly denied treatment one couple would be unfairly refused IVF because of a false-positive result. With such a modest predictive ability, the use of these tests to screen patients may be questioned.

The authors did, however, hypothesize that ovarian response in the first IVF cycle could be used as a surrogate ovarian response test. If a woman had a poor response in the first IVF cycle despite maximal stimulation, and this was confirmed by a subsequent poor response, both results are likely to reflect a truly diminished ovarian reserve and further IVF cycles would be ill-advised. If, however, the poor response was not confirmed by a low post-hoc result, continuing IVF could still be worthwhile (Klinkert et al., 2004; Hendriks et al., 2005). This hypothesis has, as yet, no clinical applicability until it is confirmed in prospective studies, but is attractive in that it compensates for variability between test results. As with most biological data, results from ovarian reserve tests are subject to random fluctuations. In women with a normal ovarian reserve, a low test result in the first cycle is likely to fluctuate back to the mean (a phenomenon known as ‘regression to the mean’) in subsequent cycles (Scott et al., 1990; Scheffer et al., 1999; Hansen et al., 2003; Kwee et al., 2004; Elter et al., 2005). Variability between cycles is further confounded by intra-observer variability (Scheffer et al., 2002).

Predicting gonadotrophin dosing

Although gonadotrophin regimens have been used for ovarian stimulation for more than two decades, the lack of prospective, randomized trials in the early years has meant that optimal starting doses have not been established (van Hooff, 1995). Most centres have empirically chosen a ‘standard’ dose for a ‘standard’ patient who is defined as younger than 40 years of age, having two ovaries, a normal menstrual cycle (21–35 days) and a normal basal FSH level. The doses used for this population range from 100 to 250 IU/day, according to the criteria of success: from the few oocytes required in mild ovarian stimulation protocols to the large number of oocytes considered appropriate in more aggressive stimulation regimens. Empirical dosing does not, however, account for the large variation in ovarian response between patients. This variation stems from differences in the functional capacity of the ovaries and the pharmacodynamics of FSH and leads to wide variation in the yield of oocytes.

Several recent studies have compared starting doses of FSH, including 100 versus 200 IU/day (Out et al., 1999, 2001; Hoomans and Mulder, 2002), 150 versus 250 IU/day (Out et al., 1999; Latin-American Puregon IVF Study Group, 2001) and 150 versus 225 IU/day (Yong et al., 2003). These studies were conducted in well-defined populations of ‘standard’ patients using GnRH agonist down-regulation, although the inclusion criteria were not restricted to first treatment cycles. The common primary end-point was the number of retrieved oocytes. Across

Figure 3: Predictive factors for ovarian response in patients undergoing assisted reproductive technology

For each factor an estimated ROC curve and sensitivity–specificity points for all studies reporting on the performance in predicting a poor response are shown: (a) basal AFC; (b) basal FSH [Broekmans et al. (2006), by permission from Oxford University Press and the European Society of Human Reproduction and Embryology]
these studies, administration of a higher dose led to the retrieval of more oocytes and similar pregnancy rates, but increased dose did not compensate for the age-related decline in ovarian function. In all of the studies there was a large variability in ovarian response, irrespective of the dose used, with the number of retrieved oocytes ranging from 1 to ≥30.

Only two studies have assessed different starting doses in ‘standard’ patients using GnRH antagonist cycles (Wikland et al., 2001; Out et al., 2004). In the trial of Wikland et al., 60 patients received 150 IU/day and 60 received 225 IU/day. Although significantly more oocytes were retrieved in the higher FSH dose group there was no difference in ongoing pregnancy rates. In the study of Out et al. (2004), there was no difference in the ovarian response or pregnancy rates in women randomized to FSH doses of 150 (n = 131) or 200 (n = 126) IU/day.

The aim of choosing a dose of gonadotrophin with which to achieve an ‘appropriate’ response is to obtain a balance between efficacy (to retrieve an adequate number of oocytes) and risks (to avoid OHSS and cycle cancellation due to insufficient response). A clinically appropriate ovarian response may be defined as retrieval of 5–14 oocytes per patient (Popovic-Todorovic et al., 2003b). As the number of oocytes increases there is a steady increase in pregnancy rates upon fresh embryo transfer, but beyond a certain number of oocytes the increase in pregnancy rates levels off (De Vries et al., 1999; Sharma et al., 2002). In a population of 7422 women, van der Gaast et al. (2006) showed that the mean number of oocytes associated with the highest chance of conceiving per embryo transfer and per started cycle was 13.1. This pattern was not due to the embryo transfer rate, since transfer remained stable at 93–95% when four or more oocytes were obtained. Side effects and the risk of OHSS were, however, higher as the number of retrieved oocytes increased, limiting the increase in pregnancy rate. This concept is illustrated graphically in Fig. 4, which shows the relationship between oocyte numbers, benefits (pregnancies) and risks. Ideally, patients should be in the high benefit–low risk window. Whether the incidence of inappropriate responses can be lowered by individualizing the dose of gonadotrophin is, therefore, of great clinical concern.

Individualizing the dose of FSH, from the ‘standard’ dose for the ‘standard’ patient, is common at the beginning of stimulation, during the course of stimulation and in consecutive treatment cycles. The starting point for dose individualization during the first treatment cycle is the wide range in patient characteristics within the population of ‘standard’ patients. The most common clinical practice is to adjust starting FSH doses according to age (Tinkanen et al., 1999), basal FSH level or both (Harrison et al., 2001). These dose adjustments are, however, based solely on clinical judgement and experience. Scientific evidence is lacking as there have been no well-designed, prospective, randomized trials to assess the impact of dose adjustment during the course of ovarian stimulation. Two trials have investigated the effects of dose adjustment, but interpretation of both sets of results is hampered by limitations in the designs of the studies.

An early randomized controlled trial by van Hooff et al. (1993) found that doubling the FSH dose during the course of stimulation in patients with a low response at day 5 had no effect on overall ovarian response. Methodological shortcomings include the small sample size (n = 46), the inclusion of patients over 40 years of age and those with only a single ovary and differing stimulation protocols between patients. The impact of increasing the dose following 5 days of stimulation was also investigated in the retrospective study of Khalaf et al. (2002). On day 6, patients with an E2 level ≤100 pg/ml had the dose increased to 450 IU/day, whereas in patients with an E2 level ≥100 pg/ml no gonadotrophin dose alterations were implemented (patients started on 225 or 300 IU/day depending on whether they were aged ≥35 or <35 years). The authors concluded that increasing the gonadotrophin dose in the course of stimulation did not rectify an initial poor response. Unfortunately, as with the previous study, the limitations in the methodology preclude the conclusions of the study being applied in practice.

Dose adjustments in the second treatment cycle according to response in the first are supported by the results of studies that have shown a generally consistent ovarian response (Lashen et al., 1998; Hoveyda et al., 2002) and pregnancy rates (Croucher et al., 1998) across consecutive IVF or ICSI cycles, although there appears to be an age-independent deterioration in response (Kolibianakis et al., 2002). However, to date, all published studies are retrospective and as such exhibit sampling variability and clinical heterogeneity. Land et al. (1996) analysed the effects of doubling the starting dose of HMG in the second cycle in patients who had a low response (defined as ≥5 follicles on the day of hCG administration in the first treatment cycle). More oocytes were retrieved in the second treatment cycle, but the pregnancy rate was extremely low (3.2%). Lashen et al. (1998) found that more follicles and oocytes were retrieved in the second of two consecutive cycles when the dose was increased. However, the starting dose in the first treatment cycle was not the same for all patients. During the retrospective study of Popovic-Todorovic et al. (2004), ‘standard’ patients who had failed to achieve pregnancy in the first IVF/ICSI cycle either remained at the same FSH starting dose (150 IU, n = 170) or had their dose increased (>150 IU, n = 193) or decreased (<150 IU, n = 22) according to their response in the previous cycle. More than 50% of these ‘standard’ patients required gonadotrophin dose adjustment in the second treatment cycle. Women whose dose was increased had significantly more oocytes retrieved in their second cycle than in their.
first; those whose dose remained the same had no change; those whose dose decreased had fewer oocytes retrieved than previously. These results show that adjusting the dose of FSH in the second IVF/ICSI treatment cycle based on the response in the first cycle had a significant impact on the ovarian response in terms of the mean number of oocytes retrieved. The impact on the proportion of women achieving an appropriate ovarian response was less pronounced (Popovic-Todorovic et al., 2004).

Identifying independent predictors of ovarian response to FSH would allow individualization of the FSH dose from the first cycle, based on the patient’s characteristics at screening. Although there has been extensive research to define factors predictive of ovarian response to gonadotrophin stimulation, only recently has a gonadotrophin dosage nomogram based on predictive factors been designed and tested (Popovic-Todorovic et al., 2003a,b). To assess the predictive ability of a number of factors, a multiple regression analysis was undertaken using data from a prospective study of 145 ‘standard’ patients treated with 150 IU/day of FSH during their first IVF/ICSI cycle (Popovic-Todorovic et al., 2003b). A standard patient was defined as a woman aged ≤40 years with a regular menstrual cycle and a normal basal FSH level. Baseline factors (age, BMI, cycle length and smoking status) and factors measured on days 2–5 of stimulation (total ovarian volume, total number of antral follicles ≥10 mm diameter, total Doppler score of the ovarian stromal blood flow, serum FSH, LH, E2, inhibin B and testosterone) were examined as possible predictive factors. Using backward stepwise regression analysis (regression coefficient, P-value), the total number of retrieved oocytes was predicted from the total number of antral follicles (0.249; P < 0.001), total power Doppler score (1.295; P = 0.001), smoking status (1.840; P = 0.015) and serum testosterone level (1.457; P = 0.060). The final model explained 38% of the variability in the number of oocytes (adjusted R² = 0.379).

To allow these findings to be implemented in clinical practice they were incorporated into a recombinant human FSH (r-hFSH) dosage nomogram to ascertain the dose of r-hFSH that would yield an appropriate number of oocytes, arbitrarily defined as 5–14. The nomogram comprised the total number of antral follicles on days 2–5, total Doppler score on days 2–5, total ovarian volume on days 2–5, age and smoking status (Table 3). By using an individual r-hFSH dose regimen it was hypothesized that a more uniform oocyte distribution would be achieved than by giving a standard dose to all patients. To test the use of the FSH dosage nomogram in clinical practice, a randomized trial compared ovarian response in women assigned either to an individual dose of FSH based on her score or to a ‘standard’ dose of 150 IU/day (Popovic-Todorovic et al., 2003a). All 262 women were ‘standard’ patients undergoing IVF/ICSI treatment using down-regulation with a long GnRH agonist protocol. In the individual-dose group, a higher proportion of patients had an appropriate ovarian response, defined as retrieval of between 5 and 14 oocytes, than women in the standard-dose group (101 versus 86 patients; P = 0.04) and more women in the standard-dose group required dose adjustment than in the individualized-dose group from day 8 onwards (86 versus 59%; P = 0.001). Individual dosage regimens in a well-defined standard patient population increased the proportion of appropriate ovarian responses and decreased the need for dose adjustments during the course of ovarian stimulation. A higher ongoing pregnancy rate was observed in the individual-dose group (37%, 48/131 versus 24%, 32/131; P = 0.03). The data from this randomized trial, therefore, justifies a tailored approach to starting doses from the first treatment cycle in a well-defined group of ‘standard’ patients.

An alternative FSH dosing algorithm has been developed through meta-analysis of data from 1378 normo-ovulatory patients aged ≤35 years (Howles et al., 2006). The factors most predictive of ovarian response for ART, basal FSH, BMI, age and the number of follicles (diameter <11 mm) at baseline were weighted and modelled into a dosing algorithm to calculate the starting dose of recombinant FSH (rFSH). The use of the dosing algorithm has recently been tested in a prospective trial. It is clear that this area deserves further study to determine more accurately the starting dose at which to reduce both poor and hyper-response.

**Table 3: r-hFSH dosage nomogram**

<table>
<thead>
<tr>
<th>Measure</th>
<th>r-hFSH score (IU/day)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of follicles ≥10 mm Days 2–5</td>
<td>&lt;15</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>15–25</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>≥25</td>
<td>50</td>
</tr>
<tr>
<td>Total ovarian volume days 2–5 (ml)</td>
<td>&lt;9</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>9–13</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>&gt;13</td>
<td>50</td>
</tr>
<tr>
<td>Total Doppler score days 2–5</td>
<td>2–3</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Age (years)</td>
<td>&gt;35</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>&gt;30–L ≤35</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>≤30</td>
<td>0</td>
</tr>
<tr>
<td>Smoking habits (cigarettes/day)</td>
<td>&gt;10</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>≤10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Non-smoker</td>
<td>0</td>
</tr>
</tbody>
</table>

Popovic-Todorovic et al. (2003b) by permission of Oxford University Press and the European Society of Human Reproduction and Embryology.
**Predicting the need for LH supplementation**

LH is an important regulator of the normal menstrual cycle and is supplemented in women with hypogonadotropic hypogonadism undergoing ovulation induction. There is, however, no relation between endogenous LH levels and pregnancy rates when all women with normal ovulation or oligo-ovulation undergoing IVF are grouped together (Kolibianakis et al., 2006). For example, a prospective study that measured LH from stimulation day 5 to the administration of hCG (Penarrubia et al., 2003), a retrospective cohort analysis that measured LH on stimulation days 3 and 10 (Cabrera et al., 2005) and a review of patient records in which LH was measured on day 1 (Bjercke et al., 2005) all showed that LH levels were not predictive of the outcome of IVF or ICSI in women undergoing down-regulation with a GnRH agonist. The change in LH concentrations over the course of stimulation may, however, be important (Kol, 2005). Women in whom LH fell by 50% from the early- to mid-follicular phase of stimulation may, however, be important (Kol, 2005). Although not an empiric use of LH, studies have shown some women to benefit from LH supplementation; these have been reviewed previously (Caglar et al., 2005; Alviggi et al., 2006; Griesinger and Diedrich, 2006; Humaidan, 2006). The first subgroup of women who may gain from LH comprises older patients. In the randomized study of Humaidan et al. supplementation with LH from day 8 improved pregnancy rates in women older than 35 years (Humaidan et al., 2004). Similarly, when women were randomized to FSH with or without additional LH from day 6, implantation rates were higher in women aged over 35 years who were receiving LH than in those who were not (Marrs et al., 2004). The second subgroup comprises women with a reduced ovarian response to FSH. When such patients were randomized to FSH or FSH plus LH, pregnancy rates and live birth rates were higher in the women receiving LH than those receiving just FSH, despite FSH-dose elevation (Ferraretti et al., 2004; De Placido et al., 2005). A third group comprises normogonadotropic patients who have LH concentrations above 1.99 IU/l on stimulation day 8 after down-regulation with a GnRH agonist. Within this group of patients, the implantation rate was higher when women were randomized to LH supplementation, compared with FSH only (Humaidan et al., 2004).

**Genetic Predictors**

Since the human genome was mapped, much progress has been made in the search for genes related to ovarian function (Layman, 2006). Genetic polymorphisms such as single nucleotide polymorphisms (SNPs) may become the preferred predictive factors of ovarian response. The genetic test closest to reaching the clinic is that for polymorphisms of the FSH receptor (FSHR), which may help to predict the most appropriate dose of FSH for each woman. Mutations in the FSHR are associated with primary amenorrhoea (Doherty et al., 2002; Meduri et al., 2003), and a common SNP in the FSHR gene (rs6166, causing a change from an asparagine (A) to a serine (S) residue at codon position 680; p.S680N) is associated with a different sensitivity to both exogenous (Perez Mayorga et al., 2000) and endogenous (Greb et al., 2005) FSH. Moreover, anovulatory patients may have a different FSHR genotype compared to normo-ovulatory controls (Laven et al., 2003). As a group, women with the S/S genotype have a higher FSH threshold than those with the A/A genotype (Sudo et al., 2002; Greb et al., 2005; de Koning et al., 2006) and may benefit from a higher dose of FSH when undergoing multifollicular stimulation (Behre et al., 2005; Jun et al., 2006). The question of whether this polymorphism is associated with pregnancy rates remains controversial (Jun et al., 2006; Klinkert et al., 2006) and requires further study in larger populations. Furthermore, recent observations suggest that AMH and AMH receptor type II polymorphism is also associated with FSH sensitivity in the human ovary (Kevenaar et al., 2007).

Although progress has been slow, genetic factors may eventually help in predicting ovarian response and the likelihood of OHSS. Currently, most progress has come from initiatives to identify the contribution of genetic factors to ovarian dysfunction in patients with PCOS (Escobar-Morreale et al., 2005; Diamanti-Kandarakis and Piperi, 2005). Distinct SNPs in genes involved in steroid biosynthesis and in the hypothalamic–pituitary–gonadal axis have been identified in patients with WHO type II anovulation and PCOS. A common SNP in the aromatase gene (AR) may also be of interest. Other PCOS genes of interest include AMH and AMH receptors.

The challenge will be to study whether a certain SNP pattern related to ovarian dysfunction in PCOS is also associated with ovarian response to stimulation. In addition, certain SNP patterns may be identified related to FSH sensitivity in normo-ovulatory women. This, again, may impact on optimal dosing required for ovarian stimulation for IVF. It seems likely that—with many novel molecular research tools currently available—much attention in clinical research will focus on this crucial area in the near future. This may reveal entirely new possibilities for making individualized ovarian stimulation protocols a reality.

**Conclusions**

Predicting and managing the variability between patients is a significant clinical challenge in mono- or multifollicular ovarian stimulation protocols. Research into predictive factors and the construction of multivariate models are the first steps towards evidence-based individualized treatment. As yet, however, predictive models have a limited use in clinical practice because of their limited power and the need for validation.

Predictive power will improve when more factors are identified, particularly genetic factors. Validation will improve with further studies that apply the prediction model prospectively in a different patient population but with similar characteristics to that in which the model was developed. Only when these criteria have been met can the validation be trusted. So far, the results from validation studies that have met these criteria have been encouraging. Practical considerations also need attention: it is important for a
prediction model to be simple enough for physicians to remember and incorporate into daily work and to include only variables that are routinely measured.

Despite problems in using the current predictive tests in clinical practice, the wide variation in patients’ characteristics mean that individualized, patient-tailored approaches remain mandatory for safe and effective ovarian stimulation. The current practice of individualized treatment is based only on clinical experience and has poor reproducibility. The challenge is to design studies to identify better response prediction and further test the added value of individualized approaches.

Acknowledgements

The authors would like to thank Drs Veronika Alam, Aliza Eshkol and Krisztina Bagamery (Merck Serono International S.A., Geneva, Switzerland, an affiliate of Merck KGaA, Darmstadt, Germany) for their participation in the meeting and Dr Polly Field (Caudex Medical, Oxford, UK) for her assistance in drafting the manuscript. The second Evian Annual Reproductive (EVAR) Workshop was held in February 2007.

Funding

The workshop and the preparation of this manuscript were both sponsored by an unrestricted educational grant from Merck Serono International S.A., Geneva, Switzerland.

Appendix

The Evian Annual Reproduction (EVAR) Workshop Group 2007: Mohamed A. Aboughar (Cairo University, Cairo, Egypt), Anders N. Andersen (Copenhagen University Hospital, Copenhagen, Denmark), Philippe Bouchard (Hôpital Saint Antoine, Paris, France), John Collins (Nova Scotia, Canada), Benoît Destenaves (Merck Serono International S.A., Geneva, Switzerland), Georg Griesinger (University Clinic of Schleswig-Holstein, Luebeck, Germany), Joop S.E. Laven (Erasmus Medical Centre, Rotterdam, The Netherlands), François Olivennes (Clinique de La Muette, Paris, France), Biljana Popovic-Todorovic (University Hospital, Brussels, Belgium), Gamal Serour (Al-Azhar University, Cairo, Egypt), Manuela Simoni (University of Münster, Münster, Germany), B.C. Tarlatzis (Papageorgiou General Hospital, Aristotle University of Thessaloniki, Greece) and Egbert R. Te Velde (University Hospital, Utrecht, The Netherlands).

References

Aboughar M. Prediction of ovarian hyperstimulation syndrome (OHSS). Estradiol level has an important role in the prediction of OHSS. Hum Reprod 2003;18:1140–1141.


Eskog A, Nilsson L, Brannstrom M. Low peripheral blood levels of the immunosuppressive cytokine interleukin 10 (IL-10) at the start of gonadotrophin stimulation indicates increased risk for development of ovarian hyperstimulation syndrome (OHSS). J Reprod Immunol 2001;49:71–85.


Predictors of ovarian response


Fauser et al.


Submitted on July 6, 2007; resubmitted on August 13, 2007; accepted on October 10, 2007