Prediction of over-response to ovarian stimulation in an intrauterine insemination programme

H.Lashen1,2, M.Afman1, L.McDougall1 and P.Clark2

1Assisted Conception Unit, Birmingham Women’s Hospital, Edgbaston, Birmingham, B15 2TG and 2Regional Endocrine Laboratory, University Hospital, Birmingham B29 6JD, UK

Prediction of poor-response is of equal importance to prediction of over-response in intrauterine insemination programmes. The gonadotrophin-releasing hormone agonist (GnRHa) stimulation test (GAST) was assessed as a predictor of over-response to ovarian stimulation in 81 patients. Blood samples were taken on cycle day 2 (before and 24 h after starting the GnRHa). Day 2 and 3 samples were assayed for oestradiol, follicle stimulating hormone (FSH) and luteinizing hormone (LH). Linear and logistic regression analyses were used to assess age, day 2 FSH, day 2 FSH/LH, oestradiol ratio (oestradiol on day 3/ oestradiol on day 2) and FSH ratio (FSH on day 3/FSH on day 2) as predictors of the number of follicles (total and ≥14 mm), oestradiol on HCG day, and clinical pregnancy rate as appropriate. Several parameters were also compared between the patients who produced ≤3 (≥14 mm) follicles (group A) and those who produced >3 (≥14 mm) follicles (group B). The mean ± SEM age of the patients in the study was 32 ± 0.4 years. The mean total dose of recombinant FSH was 800 ± 20 IU and the mean duration of stimulation was 7.6 ± 0.2 days. Nine (11%) and 12 (15%) patients were cancelled for poor and over-response respectively. The oestradiol ratio was significantly positively correlated with oestradiol on HCG day (P < 0.001), and with the number of mature follicles (≥14 mm) (P = 0.01). Age, day 2 FSH and FSH ratio were not significantly correlated with oestradiol on HCG day, total follicles and follicles ≥14 mm. None of the above-mentioned variables was correlated with clinical pregnancy rate. Group A had significantly lower oestradiol ratio (P = 0.007), longer duration of stimulation (P = 0.002), higher total FSH dose (P = 0.001), and lower oestradiol on HCG day (P = 0.001). GAST is therefore useful in predicting the high responders to gonadotrophin stimulation.

Key words: GnRH stimulation test/intrauterine insemination/ovarian response

Introduction

Ovulation induction and intrauterine insemination (IUI) has been used to treat couples with unexplained and mild male factor infertility (Dodson et al., 1987; Cruz et al., 1986; Serhal et al., 1988). The extent of ovarian stimulation in this modality of treatment is crucial, mainly to avoid the recruitment of an excessive number of follicles with the attendant risk of multiple pregnancy. The multiple pregnancy rate has increased dramatically with the introduction of assisted reproduction techniques. Cohen (1998) reviewed the data on the incidence of multiple pregnancy associated with in-vitro fertilization (IVF) treatment and recommended several approaches to reduce such risk with its associated perinatal mortality and morbidity. However, reducing the incidence of multiple pregnancy in IVF is relatively easy compared with that in ovarian stimulation/IUI due to the difficulties encountered in controlling the number of follicles produced given the wide variability in women’s response to ovarian stimulation. High responders in this treatment category are usually cancelled unless IVF treatment is available. Therefore, prediction of ovarian response in this group of patients will help avoiding unnecessary cost and anguish to the patients. Age, basal follicle stimulating hormone (FSH) concentration, clomiphene citrate challenge test (CCCT), and gonadotrophin-releasing hormone (GnRH) agonist stimulation test (GAST) have been reported as good predictors of ovarian response and outcome in IVF patients (Fédération CECOS 1982; Navot et al., 1987; Padilla et al., 1990; Toner et al., 1991). As these tests were developed and tested mainly in IVF patients, the main aim was to predict poor response and the pregnancy outcome. However, predicting the patients’ over-response to stimulation is equally important in the IUI programme. We prospectively assessed GAST as a predictor of ovarian over-response to stimulation.

Materials and methods

Subjects

During the period from October 1, 1996 to September 30, 1997, patients with unexplained infertility undergoing IUI were invited to take part in the study. Those who agreed were asked to give a written consent and were then consecutively enrolled in the study. Unexplained infertility was diagnosed by confirming tubal patency with hysterosalpingography or laparoscopy and dye test, normal semen fluid analysis (semen volume ≥2 ml, count ≥20 x 10⁶/ml, motility ≥50% and normal forms ≥20%), and confirming ovulation by a serum progesterone concentration >30 nmol/l in the mid-luteal phase.

Intrauterine insemination protocol

The ovarian stimulation protocol involved ultrasound scanning on cycle day 2 to rule out any ovarian pathology before starting intranasal nafarelin acetate 400 µg twice daily (Synarel, Searle, High Wycombe, Bucks, UK) from cycle day 2 until the administration of
human chorionic gonadotrophin (HCG). On cycle days 3–5 the
patients were given 150 IU of recombinant FSH (rFSH, Gonal-F;
Serono UK Ltd, Garden City, Herts, UK) daily which was then
reduced to 75 IU daily from day 6 onwards. The patients' response
was monitored with ultrasound scanning starting on cycle day 9, and
daily thereafter. When the leading follicle measured \( \geq 16 \, \text{mm} \), 10,000
IU HCG was administered.

A maximum of five follicles measuring \( \geq 14 \, \text{mm} \) was the threshold
above which the cycle was either cancelled or converted to IVF. A
total of nine patients who failed to produce at least one follicle \( \geq 14 \, \text{mm} \)
after 10 days of gonadotrophin stimulation were cancelled for poor response.

**Design**

Blood samples were taken between 0800 and 1100 h on cycle day 2
(before starting the GnRHa) and on cycle day 3. A third blood sample
was taken on HCG day. Day 2 sample was assayed for oestradiol,
FSH and luteinizing hormone (LH). Day 3 sample was assayed for
oestradiol and FSH, and the HCG day sample was assayed for
oestradiol.

**Hormonal assays**

The blood samples were centrifuged at 3000 g, the serum was
separated and frozen at \(-20^\circ\text{C}\) and assayed in one batch at the end
of the study. Oestradiol was measured by direct radioimmunoassay
(Sorin Biomedica Diagnostica, Wokingham, Berks, UK) with interas-
say coefficient of variation (CV) of 9.1% at 212 pmol/l and 8.5% at
763 pmol/l. FSH was measured by an automated chemilumimetric assay
(Chiron Diagnostics, Halstead, Essex, UK) calibrated against the
WHO second International Reference Preparation (IRP) 78/549. CV were <10% over the range 8.6–51 IU/ l. LH was measured by chemilumimetric assay (Chiron Diagnostics)
calibrated against WHO second IRP 80/552. CV were <9% over the
range 4.6–53 IU/l.

**Data analysis**

Linear and logistic regression analyses were used where appropriate
to establish the relationship between age, day 2 FSH, FSH/LH,
oestradiol ratio (oestradiol day 3/oestradiol day 2), and FSH ratio
(FSH day 3/FSH day 2) as predictors, and the number of follicles
(total and \( \geq 14 \, \text{mm} \)), oestradiol on HCG day, and clinical pregnancy
rate as the response. The \( t \)-value as well as the regression coefficient
were provided since we believe that the \( t \)-value is more informative
than the regression coefficients alone.

We also compared the age, oestradiol ratio, FSH ratio, day 2 FSH,
and FSH/LH between the poor responders, over-responders and
adequate responders (within the specified threshold of five follicles
\( \geq 14 \, \text{mm} \)). The significance of GAST was assessed at an arbitrary
threshold of three mature follicles \( \geq 14 \, \text{mm} \), which is the recom-
ended threshold of the British Fertility Society (Balen, 1997). The
patients were divided into two groups: A (1–3 follicles) and B (\( \geq 3 \)
follicles) and the two groups were compared with regard to age,
oestradiol ratio, FSH ratio, day 2 FSH, day 2 oestradiol, and
FSH/LH, duration of stimulation, total gonadotrophin dose, clinical
pregnancy rate per insemination and multiple pregnancy rate.

ANOVA (analysis of variance) test was used to carry out the
comparative analyses. \( P < 0.05 \) was considered significant. Statistical
analysis was carried out using Minitab for Windows (Minitab Inc.,
State College, PA, USA).

Ethical committee approval and patients’ written consent were
obtained before starting the study.

**Results**

A total of 81 patients with unexplained infertility was recruited
for the study. The mean age \( \pm \text{SEM} \) was 32 \( \pm \) 0.4 years
(range 22–42). The mean total rFSH dose was 800 \( \pm \) 20 IU
and the mean duration of stimulation by rFSH was 7.6 \( \pm \) 0.2
days. Nine (11%) and 12 (15%) patients were cancelled for poor
and over-response respectively giving a 26% cycle
cancellation rate. The response of the remaining 60 patients
was within the accepted protocol. For comparison purposes,
the results of the 12 over-responders were included where
appropriate. The clinical pregnancy rates (CPR) per started
cycle and per insemination were 13/81 (16%) and 13/60 (22%)
respectively. The multiple pregnancy rate (MPR) was 5/13
(38%) of whom three were twins and two were quadruplets.

**Correlation analysis results**

Oestradiol ratio was correlated significantly with oestradiol on
HCG day (regression coefficient = 1017.8, \( t = 4, \, P < 0.001 \)),
and the number of follicles \( \geq 14 \, \text{mm} \) (regression coefficient =
0.5, \( t = 2.6, \, P = 0.01 \)). Day 2 FSH/LH significantly inversely
related with the total number of follicles (regression coefficient =
\(-0.7, \, t = -2 \) and \( P = 0.02 \)).

Age, day 2 FSH, day 2 oestradiol, and FSH ratio were not
significantly correlated with oestradiol on HCG day, the total
number of follicles, or the number of follicles \( \geq 14 \, \text{mm} \).
Neither age, oestradiol ratio, FSH ratio, day 2 FSH, day 2
oestradiol nor FSH/LH were significantly correlated with the
clinical pregnancy rate.

**Poor, over- and adequate responders**

The oestradiol ratio was significantly different between the
over-responders and the adequate responders \( P = 0.04 \)
while no significant difference was found between the poor
responders and the other two groups (Table I). Considering
the oestradiol ratio, the 95% CI for the poor responders \((n =
9)\), adequate responders \((n = 60)\) and over-responders \((n = 12)\)
respectively were 2.1–3.8, 2.9–3.7 and 3.5–5.4. No significant
difference was found between the three groups with regard to
age, day 2 oestradiol, day 2 FSH, and FSH/LH. The mean \( \pm \text{SD} \)
of the compared parameters is summarized in Table I.

**Comparison of patients according to number of mature
follicles \( \geq 14 \, \text{mm} \) produced**

Forty-one patients produced 1–3 mature follicles (group A),
and 31 patients produced \( \geq 3 \) follicles (group B; including 12
over-responders whose cycles were cancelled). The two groups
were similar with regard to age, FSH ratio, day 2 FSH, day 2
oestradiol, and FSH/LH ratio (Table II). However, the oestradiol
concentration on HCG day and the oestradiol ratio were
significantly different between the two groups. In groups A
and B, the 95% CI for the oestradiol ratio were 2.6–3.4 and
3.6–4.3 respectively.

**Sensitivity and predictive value of the test at different
thresholds**

The sensitivity and positive and negative predictive values of
the GAST in predicting over-response \( \geq 3 \) mature follicles)
Over-response to ovarian stimulation

Table I. Adequate, poor, and over-responders

<table>
<thead>
<tr>
<th></th>
<th>PR</th>
<th>AR</th>
<th>OP</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 9)</td>
<td>(n = 60)</td>
<td>(n = 12)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>34.1 ± 5.2</td>
<td>32 ± 3.9</td>
<td>31.2 ± 3.3</td>
<td>NS</td>
</tr>
<tr>
<td>Day 2 FSH (IU)</td>
<td>10.1 ± 5.2</td>
<td>8.7 ± 3.6</td>
<td>8.6 ± 3.3</td>
<td>NS</td>
</tr>
<tr>
<td>Day 2 oestradiol (pmol/l)</td>
<td>99.3 ± 32</td>
<td>109 ± 35</td>
<td>108.2 ± 22.3</td>
<td>NS</td>
</tr>
<tr>
<td>FSH/LH on day 2</td>
<td>2.1 ± 1.5</td>
<td>2 ± 1</td>
<td>1.7 ± 0.7</td>
<td>NS</td>
</tr>
<tr>
<td>Oestradiol ratio</td>
<td>2.9 ± 1.1</td>
<td>3.35 ± 1.5</td>
<td>4.4 ± 1.5</td>
<td>0.04*</td>
</tr>
<tr>
<td>FSH ratio</td>
<td>2.1 ± 1</td>
<td>2.4 ± 0.6</td>
<td>2.3 ± 0.8</td>
<td>NS</td>
</tr>
</tbody>
</table>

Data presented as mean ± SD.
PR = poor responders (no follicles >14 mm after 10 days stimulation); AR = adequate responders (1-5 follicles >14 mm); OP = over-responders (>5 follicles >14 mm).
*Significant difference between over- and adequate responders only; NS = not significant.
FSH = follicle stimulating hormone; LH = luteinizing hormone.

Table II. Clinical characteristics of patients in group A (1-3 follicles) and in group B (>3 follicles, including 12 over-responders whose cycles were cancelled)

<table>
<thead>
<tr>
<th></th>
<th>&gt;/3 follicles</th>
<th>&gt;/3 follicles</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 41)</td>
<td>(n = 31)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>32 ± 3.6</td>
<td>31.6 ± 4</td>
<td>NS</td>
</tr>
<tr>
<td>Oestradiol ratio</td>
<td>3.1 ± 1.5</td>
<td>4 ± 1.4</td>
<td>0.005*</td>
</tr>
<tr>
<td>Day 2 oestradiol (pmol/l)</td>
<td>116.8 ± 36.3</td>
<td>100.7 ± 26.3</td>
<td>NS</td>
</tr>
<tr>
<td>FSH ratio</td>
<td>2.3 ± 0.1</td>
<td>2.4 ± 0.6</td>
<td>NS</td>
</tr>
<tr>
<td>Day 2 FSH (IU)</td>
<td>9.4 ± 4.1</td>
<td>7.6 ± 3.2</td>
<td>NS</td>
</tr>
<tr>
<td>FSH/LH on day 2</td>
<td>1.9 ± 0.9</td>
<td>2 ± 1</td>
<td>NS</td>
</tr>
<tr>
<td>Oestradiol on HCG day (pmol/l)</td>
<td>2213 ± 1423</td>
<td>5142 ± 3363</td>
<td>0.0004*</td>
</tr>
<tr>
<td>Total gonadotrophin dose (IU)</td>
<td>830 ± 171.3</td>
<td>766.9 ± 767</td>
<td>NS</td>
</tr>
<tr>
<td>Duration of gonadotrophin stimulation (days)</td>
<td>7.9 ± 1.9</td>
<td>7.2 ± 2.7</td>
<td>NS</td>
</tr>
<tr>
<td>CPR/started cycle (%)</td>
<td>10/41 (25)</td>
<td>3/31 (9.7)%</td>
<td>NS</td>
</tr>
<tr>
<td>MPR (%)</td>
<td>2/10c (20)</td>
<td>3/3 (100)</td>
<td>–</td>
</tr>
</tbody>
</table>

Data presented as mean ± SD.
*Significant; NS = not significant.
Duration of gonadotrophin stimulation in days.
Two pairs of twins.
One pair of twins and two sets of quadruplets.
CPR = clinical pregnancy rate; MPR = multiple pregnancy rate.
FSH = follicle stimulating hormone; LH = luteinizing hormone.

Table III. Sensitivity and predictive values of gonadotrophin-releasing hormone agonist stimulation test at different threshold values of oestradiol ratio

<table>
<thead>
<tr>
<th>Oestradiol ratio</th>
<th>Sensitivity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7</td>
<td>0.52</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>4.0</td>
<td>0.52</td>
<td>0.59</td>
<td>0.67</td>
</tr>
<tr>
<td>4.3</td>
<td>0.42</td>
<td>0.65</td>
<td>0.65</td>
</tr>
</tbody>
</table>

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

were calculated at the following thresholds: 3.7, 4.0 and 4.3, which represent the mean ± SEM of oestradiol ratio in group B. The findings are summarized in Table III.

Discussion

Considering the personal commitment and expenditure required by patients in assisted conception, sufficient prognostic parameters are not readily available. The stage of ovarian stimulation is crucial to the success of assisted reproduction techniques. Prediction of the individual patient’s poor response to stimulation may be attempted using the patient’s age, day 3 FSH, with or without oestradiol concentrations, FSH/LH ratio, CCCT, and GAST in IVF patients. However, there is a different challenge when ovarian stimulation is used for IUI where prediction of over-response and poor response are of equal importance; it is also important to stimulate only a certain number of follicles so that the risk of multiple pregnancy may be reduced. There has not been any test available to predict over-response to ovarian stimulation. Such a test is important in centres where conversion to IVF treatment is not available. In this study we assessed the possibility of predicting ovarian over-response using GAST in a group of patients undergoing ovarian stimulation for IUI. We modified the GAST which was reported to be a sensitive predictor of ovarian response in flare-up IVF cycles (Winslow et al., 1991). The reason for modifying the test in this study was 3-fold: (i) measuring the oestradiol concentration beyond the second day post GnRHa stimulation, in the study reported by Winslow et al. (1991), was not of added significance; (ii) a value (oestradiol ratio) rather than a pattern of oestradiol rise facilitates the statistical analysis and the identification of a useful threshold for future use; (iii) since the absolute value of oestradiol is assay dependent, a ratio of two values measured with the same assay could be more reproducible.

In this study, oestradiol ratio was the only factor that was significantly correlated with the number of mature follicles (>14 mm), and oestradiol on HCG day. Furthermore oestradiol ratio was distinctly different between the adequate (1-5 follicles >14 mm in diameter) and the over-responders (>5 follicles >14 mm) despite the similarity between the two groups with regard to age, day 2 FSH, FSH ratio, day 2 oestradiol, and FSH/LH ratio. The ability of the GAST to detect a difference in oestradiol response to GnRHa despite the small number of patients was repeated when the threshold was arbitrarily changed to three mature follicles. The sensitivity and predictive values of the test at a threshold of oestradiol ratio of 4.0 (mean oestradiol ratio in group B) was the best compromise for an acceptable sensitivity and positive and negative predictive values as shown in Table III. However, it should be noted that the GAST is suggested only as a screening test and no major decisions should be made based on the result of the test alone. Though the predictive values are fairly low, they can be considered acceptable given the population size in the study and the incidence of the event screened for.

The use of GnRHa in conjunction with gonadotrophin prior to IUI has been a controversial issue (Dodson et al., 1991, Galgiardi et al., 1991, Sengoku et al., 1994, Manzi et al., 1995). Nevertheless, we believe it offers a degree of flexibility with regard to monitoring, cycle cancellation, and when conversion to IVF seems appropriate; hence it was used for ovarian stimulation in the IUI programme. Nevertheless, the use of the GAST does not require the continuation of the GnRHa provided the values of oestradiol ratio obtained in this study are validated accordingly.

The similarity of the oestradiol ratio, day 2 FSH, day 2
oestradiol, and day 2 FSH/LH ratio between the poor responders and the other groups could be due to the small number of the poor responders in the study and/or the heterogeneity of this group of patients. It should also be noted that the character of ovarian response is relative to the stimulation protocol and the gonadotrophin dose used. Therefore, in any prospective assessment of this study, validation of the test according to the stimulation protocol and the relevant population is essential. Although the number of pregnancies is too small to draw a conclusion, the data suggest that similar pregnancy and lower multiple pregnancy rates would be obtained when ≤3 follicles were recruited, as when a threshold of five mature follicles is used (as in this study).

We can conclude that prediction of ovarian over-response to ovarian stimulation in an IUI programme could be achieved using the modified GAST as suggested in this study and the patients could be counselled regarding the risk of cycle cancellation or conversion to IVF and multiple pregnancy. Furthermore, gonadotrophin dose adjustment could be decided prior to starting the treatment. The use of GnRH antagonist with recombinant FSH in the late follicular phase to induce a modest increase in the follicle numbers over the natural cycle has been suggested (Edwards et al., 1996). Whichever approach is used, it will have to be tested prospectively in a comparative study.

References