Key words: infertility/inhibin B/IVF/ovarian response

Introduction

Several approaches have been used to predict ovarian reserve and IVF outcome in response to gonadotrophin stimulation. For example, the measurement of day 3 FSH in the preceding normal cycle (Scott et al., 1989; Cahill et al., 1994; Scott and Hofman, 1995; Fawzy et al., 1997; Sharara et al., 1998); follicular blood flow (Nargund et al., 1996a; Van Blerkom et al., 1997; Chui et al., 1997; Engmann et al., 1999), ovarian volume (Syrop et al., 1995, 1999; Lass et al., 1997) or pre-treatment transvaginal ultrasound of the antral follicles (Tomâs et al., 1997; Chang et al., 1998) have all been investigated. Of these, the day 3 basal FSH levels has proved the most useful, in that a poorer IVF outcome was associated with higher levels of the hormone even within the normal range (Fawzy et al., 1997), and these results are in broad agreement with those reported by previous investigators (Cahill et al., 1994; Seifer et al., 1997; Peñarrubia et al., 2000).

With respect to inhibin, early reports indicated that this family of hormones may have potential as a marker of follicular development and numbers due to its production being FSH dependent (McLachlan et al., 1986; Hughes et al., 1990; Matson et al., 1991; Buckler et al., 1992; Burger, 1993; Mitchell et al., 1996). However, most of these studies were complicated by specificity problems associated with the radio-immunoassay used at that time. With the help of the new specific enzyme-linked immunoassays it became possible to measure the different species of inhibin (Groome et al., 1994, 1995, 1996) and it is now clear that inhibin B rises in the early follicular phase of both unstimulated (Groome et al., 1996) and FSH-stimulated cycles (Lockwood et al., 1996; Fawzy et al., 2001) and can be used as a direct measure of ovarian follicular reserve (Seifer et al., 1997). Furthermore, changes in the serum inhibin B levels may precede those found in basal FSH levels as ovarian reserve declines (Siefer et al., 1999). With regard to the value of basal day 3 inhibin levels, the early study of Siefer et al. indicated that day 3 inhibin B levels were predictive of IVF outcome (Siefer et al., 1997). However, several recent studies failed to find clinical value in measuring basal inhibin B with regard to IVF outcome (Corson et al., 1999; Hall et al., 1999; Ravhon et al., 2000).

Day 5 inhibin B levels in a treatment cycle are predictive of IVF outcome

M.Fawzy1, A.Lambert4,6, R.F.Harrison1, P.G.Knight2, N.Groome3, B.Hennelly1 and W.R.Robertson4,5

1Human Assisted Reproduction Unit, Rotunda Hospital, Dublin 1, Ireland, 2Animal and Microbial Sciences, University of Reading, Reading, 3Biological and Molecular Sciences, Oxford Brookes University, Oxford and 4Nuffield Department of Obstetrics and Gynaecology, John Radcliffe Hospital, Oxford OX3 9DU, UK
5Present address: Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland
6To whom correspondence should be addressed. E-mail: ann.lambert@obs-gyn.ox.ac.uk

BACKGROUND: Day 5 serum inhibin B during IVF treatment has been investigated as a predictor of outcome. METHODS: A total of 54 women (≤39 years, normal menses and endocrine profiles) were treated with urinary gonadotrophins or recombinant FSH following pituitary down-regulation. Serum day 3 FSH in a preceding cycle was <8.5 IU/l. Plasma inhibin B, inhibin A and estradiol were determined after 4 days of gonadotrophin administration (day 5). RESULTS: Day 5 inhibin B was the most highly correlated with the number of mature follicles (>14 mm), oocytes retrieved and fertilized. Receiver operating characteristic analysis gave high accuracy for day 5 inhibin B in predicting ovarian response and indicated that a threshold of 400 pg/ml may be helpful in the decision as to whether to continue treatment. Women with <400 pg/ml (n = 16) had lower numbers of follicles, mature follicles, oocytes retrieved, fertilized and cleaved compared with those >400 pg/ml (n = 36) and this threshold gave a positive likelihood ratio of 30, 92.9% sensitivity, 95.0% specificity and 86.7% positive predictive value to detect poor ovarian response. Day 5 inhibin B was the best predictor of pregnancy (no live births and four cases cancelled, low inhibin group; nine live births and no cancelled cycles, high inhibin group). CONCLUSIONS: Normogonadotrophic, normogonadal women with day 5 inhibin B <400 pg/ml in down-regulated cycles have a poor response to ovarian stimulation and are less likely to conceive compared with women with higher day 5 inhibin B.
Furthermore, on their own, normal baseline values are not a guarantee that an endocrine organ is functioning normally (Wallach, 1995) in that patients with normal baseline measurements may also have compromised ovarian reserve (Farhi et al., 1997; Kim et al., 1998). A clearer picture of the likely ability of the ovary to respond, and what hormone markers are of value in predicting this reserve, may only become evident during gonadotrophin treatment itself, and a number of studies have been designed to address this. For example, a recent retrospective study (Peñarrubia et al. 2000) found that day 5 inhibin B levels in treatment cycles correlated well with the lack of an ovarian response but not to pregnancy. Further, a multicentre study on the efficacy of a daily fixed dose stimulation regimen of 100 or 200 IU recombinant FSH (Eldar-Geva et al., 2000) showed that the rise in serum inhibin B levels after 4–10 days of FSH treatment was correlated with the number of recruited follicles and oocytes retrieved.

To date, to our knowledge, there has not been a prospective study primarily designed to investigate the predictive value of serum inhibin B at the early follicular phase of an IVF cycle in comparison with other hormones. The need for such studies has been stressed (Galtier-Dereure et al., 1996; Fabregues et al., 2000). To that end, we have performed a prospective assessor-blind controlled trial on 54 women to evaluate the value of measuring the levels of early follicular phase (day 5) inhibin B, inhibin A and estradiol in predicting IVF outcome in terms of ovarian response and pregnancy.

Materials and methods

Subjects and study design

Women presenting with unexplained or tubal infertility at the Human Assisted Reproduction Unit in Ireland (HARI), Rotunda Hospital, Dublin were recruited to take part in this study. Ethical Committee approval and specific consent to participate in the study were obtained.

The study was prospective, controlled and assessor-blind. The inclusion criteria were female age at the time of screening (20–39 years), body mass index (BMI, 20–28 kg/m²), >2 years duration of infertility with a full fertility profile (Harrison et al., 1981), a normal menstrual cycle (range 24–35 days), no previous attempts at any kind of assisted conception techniques, the presence of two ovaries, no previous ovarian surgery and a normal day 3 FSH (≤8.5 IU/l) (Fawzy et al., 1997) measured in a preceding cycle. The exclusion criteria were endocrine abnormalities including hyperprolactinemia and polycystic ovarian syndrome (PCOS), and previous exposure to gonadotrophin treatment. Those women with male partners having a total sperm concentration >1536 x 10⁶/ml, with sperm motility and normal morphology (>85%) were excluded from the study.

A total of 54 women were randomized to receive 225 IU/day of either urinary gonadotrophins in the form of Humegon (n = 13; Organon Laboratories Ltd, Cambridge, UK), Normegon (n = 11; Organon), Orgafol (n = 14; Organon) or recombinant FSH (n = 16; Puregon, Organon; and Gonafol; Serono Laboratories Ltd, Welwyn, Herts, UK) in their first IVF cycle. The randomization of gonadotrophins used was done by the pharmacist using a computerized code and was blind to the clinician. The urinary gonadotrophin preparations were from a single batch for each drug and all the gonadotrophic injections were administered i.m. by the clinic nurses. This was not a double-blind study because the gonadotrophin preparations differed in appearance. However, all the gonadotrophin drugs were administered i.m. by the nursing staff independent of the clinician’s activities in evaluating the scans and making the clinical decisions. Concealment of treatment allocation avoided any bias towards cycle cancellation or changing the dose of specific drug during the treatment. This also ensured patient compliance with regard to drug administration. Despite the fact that diagnostic test studies require a cohort design, the assessor-blind randomization between the four gonadotrophins avoided potential bias. In IVF clinics, usually a mixture of drugs is available to the clinician (Fawzy and Harrison, 1995) and a predictive test for IVF outcomes should not be influenced by the kind of drug administered. In addition we found no difference in the number of follicles stimulated by the four different preparations used (Table I). Furthermore, no differences were found in the concentration of inhibin B, inhibin A and estradiol per follicle (≥14 mm) on day 5 or pregnancy rate for the four drugs administered (Table II).

Following screening by vaginal ultrasound on day 1 of the menstrual cycle, patients were placed on a long protocol down-regulation regime (Kondaveeti-Gordon et al., 1996). Buserelin (Superfact; Hoechst, Uxbridge, UK) administered as a nasal spray, 150 µg 6 hourly, was used for pituitary down-regulation. An initial check for down-regulation was made at 14 days and then if necessary rechecked until serum estradiol was ≤100 pmol/l with quiescent ovaries on ultrasound scan (no cysts ≥10 mm). IVF treatment was performed as described previously (Harrison et al., 1992, 2001; Gordon et al., 2001).

Ovarian follicles were measured by 5 MHz vaginal transducer attached to a sector scanner (RT3600; General Electric, Milwaukee, WI, USA) and the mean of the two largest perpendicular diameters was taken. Cycles were cancelled if the ovarian response was poor with <3 follicles at day 8 of the cycle. No cycle was cancelled in this study because of excessive response. Maturity of oocytes was based on published criteria (Veeck et al., 1983) and mature oocytes (metaphase II) had a generous expansion of cumulus-corona and a distinct polar body. Pregnancy was diagnosed by increasing concentrations of β-HCG 14 days post embryo transfer and the subsequent demonstration of an intrauterine gestational sac by transvaginal scan 2 weeks later. The pregnancy rate is defined as the clinical pregnancy per initiated cycle and birth rate was defined as the percentage of liveborn infants per initiated cycle.

No differences were observed in the ovarian response or pregnancy rate to urinary or recombinant gonadotrophins. Therefore for the purpose of this study the data were combined into a single group (n = 54 cases).

Hormone assays

Blood samples were taken to examine the secretory profiles of immunoactive inhibin B, inhibin A, estradiol (E₂) and FSH. Days chosen for the study, according to the clinic’s policy, were prior to down-regulation (pre-treatment), the day of down-regulation (day 1), the day of the first clinic visit to monitor the ovarian stimulation (day 5) after 4 days of gonadotrophin administration, and then on day 8 of the treatment cycle. The samples were taken 24 h post gonadotrophin injection and prior to the next injection. The blood samples were allowed to clot, and the serum was separated and stored at −40°C until hormone assay (Fawzy et al., 2001).

Inhibin A was measured by enzyme-linked immunosorbent assay (ELISA) using a commercial kit (Serotech, Kidlington, Oxford, UK). The assay has minimal cross-reaction with inhibin B, pro-αC or activins. The limit of detection is 4 pg/ml and intra- and inter-assay coefficient of variation (CV) was <6 and <16% over the working range 4–500 pg/ml.

Inhibin B was measured by ELISA (Groome et al., 1996) using a commercial kit (Serotech). The assay has minimal cross-reaction with...
interval (CI) where appropriate. Unpaired two-tailed Student’s t-test was used to compare normally distributed continuous variables and Fisher’s exact test was used to compare categorical outcomes; analysis of variance was used to compare multiple variables. Correlation coefficients given are Spearman’s or Pearson’s, dependent on the distribution. P < 0.05 was considered to be significant.

Ovarian outcomes are presented as discrete counts and the clinical pregnancy as dichotomous (yes versus no) data. In order to make the analyses consistent for the main outcome, the number of mature oocytes was transformed into dichotomized data (yes or no). The threshold to delineate poor ovarian responders was chosen arbitrarily as 8 mature oocytes, the 25th percentile of the observed distribution of this variable, i.e. less than 8 mature oocytes (= yes) or >8 (= no).

A multiple regression and multiple logistic regression model was then used for all possible prognostic variables (hormones) to evaluate the relationship between different hormone concentrations and ovarian response. Age and duration of infertility were included in the regression models.

The diagnostic accuracy of different hormones (tests), described as their ability to discriminate between patients with a poor or normal ovarian response and non-pregnant and pregnant patients, was assessed by the area under the receiver operating characteristic (ROC) curve (Beck and Schultz, 1986; Zweig and Campbell, 1993). ROC plots sensitivity (true positives) for a given concentration of the hormone (y-axis) against 1 – specificity (false positives, x-axis). The area under the ROC (AUCROC) is determined statistically using the Wilcoxon and Mann–Whitney U-statistics (Hanley and McNeil, 1982). AUCROC is a quantitative measure of accuracy of the test, where the closer the plot is to the upper left corner the higher the overall accuracy of the test variable (Chard and Lilford, 1991) were calculated at the discriminating cut-off point for the prevalence of the study population. Sensitivity is the percentage of the abnormal cases (poor ovarian response or non-pregnancy) correctly identified as such by the diagnostic test (i.e. true positives). Specificity is the percentage of the normal cases (normal ovarian response or non-pregnancy) correctly identified as such by the diagnostic test (i.e. false positives). The positive predictive value is calculated from the ROC analysis by plotting the percentage of true positives against the percentage of true positives at the best discriminating cut-off. The positive likelihood ratio is the ratio between the probability of a positive result given the presence of a poor ovarian response and the probability of a positive test result given the absence of a poor ovarian response.

Statistical analyses were performed on GraphPad InStat version 3.0a.

### Table I. Gonadotrophin dose, duration of stimulation and numbers of follicles stimulated by four gonadotrophins

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Humegon (n = 13)</th>
<th>Normegon (n = 11)</th>
<th>Orgafol (n = 14)</th>
<th>rFSH (n = 16)</th>
<th>ANOVA</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSH (IU)</td>
<td>2950 ± 1217</td>
<td>2230 ± 648</td>
<td>2313 ± 482</td>
<td>2760 ± 817</td>
<td></td>
<td>2</td>
<td>0.1 (NS)</td>
</tr>
<tr>
<td>Duration of stimulation (days)</td>
<td>9.7 ± 1.2</td>
<td>9.1 ± 1.0</td>
<td>9.2 ± 1.0</td>
<td>9.4 ± 1.3</td>
<td></td>
<td>0.5</td>
<td>0.8 (NS)</td>
</tr>
<tr>
<td>No. of follicles ≥14 (mm)</td>
<td>14 ± 7.6</td>
<td>17 ± 5.5</td>
<td>16 ± 6.2</td>
<td>14 ± 6.5</td>
<td></td>
<td>1.5</td>
<td>0.2 (NS)</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD (95% confidence interval). F ratio: the ratio of the variable mean squared to the error mean square.

rFSH = recombinant FSH; ANOVA = analysis of variance; NS = not significant.

### Table II. Day 5 hormone per mature follicle (≥14 mm) at day 5 of IVF stimulation cycle in the four gonadotrophin groups

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Humegon (n = 13)</th>
<th>Normegon (n = 11)</th>
<th>Orgafol (n = 14)</th>
<th>rFSH (n = 16)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 5 inhibin B/ follicle (pg/ml)</td>
<td>44.9 ± 14</td>
<td>57.3 ± 12</td>
<td>49.9 ± 26</td>
<td>49 ± 26</td>
<td>0.6 (NS)</td>
</tr>
<tr>
<td>Day 5 inhibin A/ follicle (pg/ml)</td>
<td>7.9 ± 4.9</td>
<td>12.5 ± 7.9</td>
<td>8.3 ± 5.0</td>
<td>8.1 ± 9.4</td>
<td>0.8 (NS)</td>
</tr>
<tr>
<td>Day 5 estradiol/ follicle (pmol/l)</td>
<td>48.6 ± 23</td>
<td>38.3 ± 18</td>
<td>37.9 ± 18</td>
<td>44 ± 24</td>
<td>0.5 (NS)</td>
</tr>
<tr>
<td>Pregnancies</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0.7 (NS)</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD (95% confidence interval).

rFSH = recombinant FSH; NS = not significant.

inhibin pro-αC or activins, and ~1% cross-reaction with inhibin A. The limit of detection is 16 pg/ml and intra- and inter-assay CV was <6% and <13% over the working range of 16–1000 pg/ml.

Serum estradiol was measured by Delfia kit (Wallace, Milton Keynes, Bucks, UK) according to the manufacturer’s protocol. The assay has a limit of detection of 50 pmol/l and intra- and inter-assay CV of <10% and <13% over the range 100–1200 pmol/l.

Serum samples were assayed for FSH using the Delfia time-resolved fluorescence immunoassay kit following the manufacturer’s instructions. The limit of detection of the assay was 1 IU/l and the intra- and inter-assay CV was <6% and <13% over the range 1–250 IU/l.
Results

Table III shows the clinical characteristics of the patients \( (n = 54) \). Four cycles (7.4%) were cancelled because of poor ovarian response (less than 3 follicles \( <10 \) mm in diameter at day 5 or 8), two cases had no fertilization of their retrieved oocytes and two cases had no available embryos (no cleavage). In the rest of the cases, 1–3 embryos were transferred back to the uterus. There were in total 10 pregnancies (18.5% per cycle). Two ended in miscarriages (6 and 13 weeks gestation); there were seven singletons and one twin delivery. In this study, no women were admitted to the hospital because of severe ovarian hyperstimulation syndrome (OHSS) (Golan et al., 1989); however, six patients had mild/moderate OHSS (more than 20 oocytes, range 22–76).

Hormone levels during treatment

The mean level of inhibin B rose \( \sim20\) -fold from day 1 (36 pg/ml, 95% CI: 28–45) to day 5 (747 pg/ml, 95% CI: 598–895). Similarly, the mean inhibin A concentration increased 14-fold from day 1 (10 pg/ml, 95% CI: 7–13) to day 5 (141 pg/ml, 95% CI: 101–181) whilst the E\(_2\) concentration at day 5 (597 pmol/l, 95% CI: 474–720) was 8.6-fold that of its concentration at day 1 (69 pmol/l, 95% CI: 62–76).

Mean levels of day 5 inhibin B were 747 pg/ml (all patients); cancelled cycles only had 67 pg/ml whereas patients with OHSS had 1572 pg/ml (Table IV).

Table III. Clinical characteristics of the patients

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
</table>
| Age (years)    | 33 (24–39)
| Body mass index (kg/m\(^2\)) | 24 (21–28)
| Type of infertility |          |
| Unexplained    | 37 (68.5%)
| Tubal          | 17 (31.5%)
| Cause of infertility |      |
| Primary        | 44 (81.5%)
| Secondary      | 10 (18.5%)
| Duration of infertility (years) | 4.6 (2–10)
| Day 3 FSH (IU/l) | 6.3 (4.4–8.2)\(^a\) (6–6.6)\(^b\)
| Days of stimulation | 9 (5–12)\(^c\)
| Gonadotrophin units used | 2550 (1350–5250)\(^a\)
| Cycles         |          |
| Total          | 54        |
| No. cancelled  | 4 (7.4%)  |

Values are \(^a\)mean (range), \(^b\)95% confidence interval.

Table IV. Hormone levels in all cycles, cancelled and ovarian hyperstimulation (OHSS) cycles

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All cycles ((n = 54))</th>
<th>Cancelled cycles ((n = 4))</th>
<th>OHSS cycles ((&gt;20) follicles, (n = 6))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibin B day 5 (pg/ml)</td>
<td>747 (598–895)</td>
<td>67 (26–108)</td>
<td>1572 (783–2366)</td>
</tr>
<tr>
<td>Inhibin A day 5 (pg/ml)</td>
<td>141 (101–181)</td>
<td>15 (11–42)</td>
<td>294 (82–507)</td>
</tr>
<tr>
<td>Estradiol day 5 (pmol/l)</td>
<td>597 (474–720)</td>
<td>137 (54–221)</td>
<td>1143 (593–1744)</td>
</tr>
</tbody>
</table>

Data are presented as mean (95% confidence interval).

Figure 1. The distribution of day 5 inhibin (\(\square\)) levels (pg/ml). The arrow denotes the 25th percentile.
Prediction of IVF outcome

between patients likely to respond adequately to treatment from poor or non-responders. Using the value of 373 for inhibin B (rounded up to 400) pg/ml as the cut-off point, day 5 inhibin B gave the highest prediction of 30.1 positive likelihood ratio, 92.9% sensitivity, 95.0% specificity, 86.7% positive predictive value and 97.4% negative predictive value for ovarian response for the study population. Using a threshold of less than 7 oocytes (n = 14 patients), day 5 inhibin B predicted ovarian response in terms of retrieved mature oocytes (odds ratio 1.015, 95% CI 1.002–1.0304, P = 0.04) with AUCROC of 0.964 and a specificity of 93%. Using a threshold of less than 6 oocytes (n = 11 patients), day 5 inhibin B predicted ovarian response in terms of retrieved mature oocytes (odds ratio 1.028, 95% CI 1.006–1.056, P = 0.04) with ROC AUC 0.987 and a specificity of 91%. Too few patients (n = 6) produced less than 5 oocytes for meaningful statistical analysis.

The diagnostic accuracy of the different hormone variables to discriminate between ongoing pregnancy and birth cases and non-ongoing pregnant IVF cases was also analysed using the ROC plot and the best predictor was again the day 5 inhibin B (0.699 + 0.07; 95% CI: 0.558–0.841, P = 0.002).

Following the above analysis, the patients were divided into two groups: low day 5 inhibin (<400 pg/ml, n = 16; group 1) and high day 5 inhibin (>400 pg/ml, n = 38; group 2). There were no differences (data not shown) in the age, BMI, type of infertility, cause of infertility, duration of infertility, duration of stimulation and dose of gonadotrophins used between group 1 and group 2. Further, pre-treatment hormone levels (FSH, E2 and inhibin B) were similar (data not shown) between the groups whereas day 5 inhibin B per follicle (≥14 mm) was (53 pg/ml, 95% CI: 44–63) in group 1, ~2.5-fold less than in group 2 (129 pg/ml, 95% CI: 114–144).

Women in group 1 had lower (P < 0.0001) mean day 5 inhibin A and B serum levels. Day 5 serum estradiol levels were directly proportional (P < 0.0001) to day 5 inhibin B levels in the two groups. There was a strong association between serum day 5 inhibin B concentration with the dose of gonadotrophin administered (P = 0.003), numbers of follicles at day 5, mature follicles >14 mm, mature oocytes (P = 0.0001), fertilized oocytes (P = 0.03), developed embryos (P = 0.03) and transferred embryos (P = 0.03). Additionally, it was noted that all the cancelled cycles (45/54 cases) were in the low day 5 inhibin B group (Table V).

All the ongoing pregnancies and births (n = 9) were in the day 5 inhibin B >400 pg/ml group (live-birth rate/cycle 21%), while the low inhibin group had only one pregnancy, which resulted in early miscarriage (live-birth rate 0%). The difference in pregnancies between the two groups did not reach significance (P = 0.09), probably due to the small numbers of patients. When the total cases (n = 54) were divided into groups of pregnant (n = 10) and non-pregnant (n = 40) cases, there were no differences in terms of patient profiles and outcome parameters. However, the mean day 5 inhibin B (946; 95% CI: 699–1194) in the ongoing pregnant group was higher than the day 5 inhibin B (662; 95% CI: 526–799) in the non-pregnant group (P = 0.07, not significant).

Discussion

Poor responses to ovarian stimulation in normogonadotropic, normogonal women have been reported (Farhi et al., 1997). Although several studies, which have been the subject of recent comprehensive reviews (Broekmans et al., 1998; Sharara et al., 1998), have attempted to develop tests for predicting ovarian reserve; these tests have limited value (Galitier-Dereure et al., 1996) and doubts remain about their accuracy and interpretation (Barnhart and Osheroff, 1999). Thus the clinician is often left puzzled by the discordant responses seen in this group of patients.

Basal FSH levels in the current study, measured in the previous menstrual cycle, were normal at <8.5 IU/l and were inversely proportional (P < 0.002) to day 5 inhibin B levels in the treatment cycle. These FSH data confirm the findings of earlier studies (Cahill et al., 1994; Fawzy et al., 1997).

It has been suggested that inhibin B, a granulosa cell product, may be more sensitive to the stimulus of FSH than inhibin A and estradiol in the early phase of the menstrual cycle (Burger et al., 1998) or stimulated IVF cycle (Lockwood et al., 1996; Anderson et al., 1998). It has also been demonstrated that a decrease in inhibin B secretion is an early marker of declining reproductive potential which is linked to both quantitative and qualitative changes in oocytes (Seki et al., 1997; Seifer et al., 1997, 1999; Welt et al., 1999a). Studies in the mouse and in the human have also shown that inhibin B is an indicator of the best quality follicles, which were higher in number in subjects who became pregnant (Smitz and Cortvrindt, 1998; Hall et al., 1999).

Although previous studies have demonstrated that inhibin B can serve as the earliest index of FSH-dependent growth of antral follicles and as a direct measure of ovarian reserve.

Women in group 1 had lower (P < 0.0001) mean day 5 inhibin A and B serum levels. Day 5 serum estradiol levels were directly proportional (P < 0.0001) to day 5 inhibin B levels in the two groups. There was a strong association between serum day 5 inhibin B concentration with the dose of gonadotrophin administered (P = 0.003), numbers of follicles at day 5, mature follicles >14 mm, mature oocytes (P = 0.0001), fertilized oocytes (P = 0.03), developed embryos (P = 0.03) and transferred embryos (P = 0.03). Additionally, it was noted that all the cancelled cycles (45/54 cases) were in the low day 5 inhibin B group (Table V).

All the ongoing pregnancies and births (n = 9) were in the day 5 inhibin B >400 pg/ml group (live-birth rate/cycle 21%), while the low inhibin group had only one pregnancy, which resulted in early miscarriage (live-birth rate 0%). The difference in pregnancies between the two groups did not reach significance (P = 0.09), probably due to the small numbers of patients. When the total cases (n = 54) were divided into groups of pregnant (n = 10) and non-pregnant (n = 40) cases, there were no differences in terms of patient profiles and outcome parameters. However, the mean day 5 inhibin B (946; 95% CI: 699–1194) in the ongoing pregnant group was higher than the day 5 inhibin B (662; 95% CI: 526–799) in the non-pregnant group (P = 0.07, not significant).

Discussion

Poor responses to ovarian stimulation in normogonadotropic, normogonal women have been reported (Farhi et al., 1997). Although several studies, which have been the subject of recent comprehensive reviews (Broekmans et al., 1998; Sharara et al., 1998), have attempted to develop tests for predicting ovarian reserve; these tests have limited value (Galitier-Dereure et al., 1996) and doubts remain about their accuracy and interpretation (Barnhart and Osheroff, 1999). Thus the clinician is often left puzzled by the discordant responses seen in this group of patients.

Basal FSH levels in the current study, measured in the previous menstrual cycle, were normal at <8.5 IU/l and were inversely proportional (P < 0.002) to day 5 inhibin B levels in the treatment cycle. These FSH data confirm the findings of earlier studies (Cahill et al., 1994; Fawzy et al., 1997).

It has been suggested that inhibin B, a granulosa cell product, may be more sensitive to the stimulus of FSH than inhibin A and estradiol in the early phase of the menstrual cycle (Burger et al., 1998) or stimulated IVF cycle (Lockwood et al., 1996; Anderson et al., 1998). It has also been demonstrated that a decrease in inhibin B secretion is an early marker of declining reproductive potential which is linked to both quantitative and qualitative changes in oocytes (Seki et al., 1997; Seifer et al., 1997, 1999; Welt et al., 1999a). Studies in the mouse and in the human have also shown that inhibin B is an indicator of the best quality follicles, which were higher in number in subjects who became pregnant (Smitz and Cortvrindt, 1998; Hall et al., 1999).

Although previous studies have demonstrated that inhibin B can serve as the earliest index of FSH-dependent growth of antral follicles and as a direct measure of ovarian reserve.

Figure 2. Receiver operating characteristic (ROC) curve for prediction of ovarian response by day 5 inhibin (AUCROC = 0.955). Each point represents the actual day 5 inhibin B level (°). The main inflexion point in the graph is indicated by the arrow. The diagonal line is the line of no discrimination (AUCROC = 0.05).
In our study, there was only one pregnancy, which ended in miscarriage, in the low day 5 inhibited B (<400 pg/ml) group compared with nine pregnancies resulting in one miscarriage and nine live births (eight singleton and one twin; live-birth rate/cycle 21%) in the high day 5 inhibited group. Peñarrubia et al. (2000) stated that day 5 inhibited B, inhibited A and estradiol had similar, but poor, predictive properties on pregnancy outcome while Eldar-Geva et al. (2000) failed to report on the pregnancy outcome, in their study. In the present study, day 5 inhibited B alone was also the best predictor of the birth rate out of all the hormones studied alone or combined with age and duration of infertility. However, this result should be interpreted with caution because of the small number of cases.

Poor ovarian response, in terms of low numbers of mature follicles (>14 mm diameter) and mature oocytes retrieved, is usually associated with poor pregnancy rates and poor oocyte quality (Keay et al., 1997). Peñarrubia et al. (2000) reported that a cut-off of day 5 inhibited B of 141 pg/ml gave a predictive efficiency (product of sensitivity and specificity) of 91.03% for cancellation. However, their cancellation rate (25%) was high compared with the internationally published rate of ~8.6% (Human Fertilisation and Embryology Authority, 2000; Society for Assisted Reproductive Technology/American Society for Reproductive Medicine, 2000). In the current study, the mean day 5 inhibited B was 67 pg/ml in the four (7.5%) cycles cancelled because of poor ovarian response.

In contrast, the mean day 5 inhibited B level was 1572 pg/ml in the six patients who suffered mild/moderate OHSS (more than 20 oocytes). In agreement with this finding, one study (Enskog et al., 2000) reported that the elevation of serum inhibited B 3 days prior to oocyte retrieval was greater in patients who later developed OHSS. Concerns have been expressed about the value of plasma estradiol measurements in predicting OHSS, a condition which is apparently exacerbated by the use of recombinant FSH preparations which lead to a reduction of the amount of estradiol secreted per follicle (Devroey et al., 1994; Lockwood et al., 1996; Jacob et al., 1998; Fawzy et al., 2001). Consequently, day 5 inhibited B measurements may be a useful tool in predicting and monitoring OHSS. However, further studies with a larger number of patients are needed to support these findings.

### Table V. Ovarian response and IVF outcome

<table>
<thead>
<tr>
<th></th>
<th>Group 1 (n = 16)</th>
<th>Group 2 (n = 36)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days of stimulation (days)</td>
<td>8.9 ± 1.1</td>
<td>9 ± 0.2</td>
<td>NS</td>
</tr>
<tr>
<td>Gonadotrophin (IU)</td>
<td>394 ± 101</td>
<td>284 ± 101</td>
<td>0.003</td>
</tr>
<tr>
<td>Day 5 follicles</td>
<td>4.4 ± 2</td>
<td>7.4 ± 2.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>Follicles ≥14 mm</td>
<td>7.4 ± 4.6</td>
<td>17.3 ± 4.6</td>
<td>0.0001</td>
</tr>
<tr>
<td>Mature oocytes</td>
<td>5.1 ± 4.8</td>
<td>14.5 ± 4.6</td>
<td>0.0001</td>
</tr>
<tr>
<td>Fertilized oocytes</td>
<td>2.3 ± 1.8</td>
<td>6.4 ± 3.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cleaved oocytes</td>
<td>1.8 ± 1.4</td>
<td>3.0 ± 1.9</td>
<td>0.03</td>
</tr>
<tr>
<td>Embryos transferred</td>
<td>1.6 ± 1.3</td>
<td>2.3 ± 1.0</td>
<td>0.03</td>
</tr>
<tr>
<td>Cancelled cycles</td>
<td>4</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>Ongoing pregnancies</td>
<td>0</td>
<td>9</td>
<td>0.09 (NS)</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD.

Group 1 includes patients with day 5 inhibited B <400 pg/ml and group 2 includes patients with day 5 inhibited B ≥400 pg/ml.

NS = not significant.
The results of the present study suggest that a serum day 5 inhibin B level <100 pg/ml may be an indication for cancellation of that cycle. Conversely, patients with a high day 5 inhibin B level (>1000 pg/ml) may benefit from a reduction in the gonadotrophin dose and closer follow-up during stimulation. However, cut-off values should act only as guidelines and should not exclude patients from treatment, but should allow for proper counselling and better planning of the most appropriate treatment. For example, patients with diminished ovarian reserve with low serum day 5 inhibin B level (<400 pg/ml) may benefit from stimulation with a higher dose of gonadotrophin or possibly a modification in GnRH agonist administration. No test of ovarian reserve is likely to be precise enough to merit total reliance on the results, and all relevant findings such as patient history, scan results and patient’s age should be taken into account.

With respect to estradiol, the levels of this hormone in the early follicular phase have been reported to be highly predictive of successful ovulation stimulation and pregnancy outcome (Phelps et al., 1998; Khalaf et al., 2000) but these authors did not report on basal day 3 hormones or inhibin(s). We found the predictive value of day 5 estradiol to be inferior to day 5 inhibin B with respect to both ovarian stimulation and pregnancy outcome.

In conclusion, serum day 5 inhibin B measured during gonadotrophin stimulation was a better marker for oocyte maturity and pregnancy rate than the other day 5 hormones such as estradiol and inhibin A. It is also a useful predictor of poor-response and for the monitoring of over-response. This information may be helpful to counsel patients who would otherwise expect a different outcome. Our study and previous studies provide further evidence that stimulated serum inhibin B measurement in the early follicular cycle, either alone or in combination with other ovarian predictive measures, may be an essential part of the future management of patients undergoing ovarian stimulation. A large prospective multicentre study is needed to confirm our findings and to set a cut-off point of inhibin B for either discontinuation or monitoring of the stimulation cycle for poor or over-response.

Acknowledgements
The authors wish to thank the Royal College of Surgeons, Ireland for financial support. The authors are also grateful to the staff at the Human Assisted Reproduction Unit, Ireland who contributed expertly to the care of the patients studied and to the staff of the Clinical Biochemistry Departments at Rotunda Hospital, Dublin and Hope Hospital, Salford for their help with the hormone analyses.

References


Prediction of IVF outcome


Submitted on August 15, 2001; resubmitted on December 13, 2001; accepted on February 2, 2002