Outcome from consecutive in-vitro fertilization/intracytoplasmic sperm injection attempts in the final group treated with urinary gonadotrophins and the first group treated with recombinant follicle stimulating hormone

S. Jacob¹, L. Drudy¹, R. Conroy² and R.F. Harrison¹,³

¹RCSI Department of Obstetrics and Gynaecology and Human Assisted Reproduction Ireland, Rotunda Hospital, Dublin 1 and ²Department of Statistics, Royal College of Surgeons in Ireland, St Stephen’s Green, Dublin 1, Ireland

³To whom correspondence should be addressed

In the absence of specific dose equivalency data, the aim of this study was to compare the clinical results during the cross-over from menopausal urinary products (human menopausal gonadotrophin; HMG) to recombinant follicle stimulating hormone (FSH) follitrophin beta (FSHr) in order to determine whether the manufacturer’s recommendation for equivalence of ampoule to ampoule (50 IU FSHr:75 IU HMG) would prove clinically correct. A total of 353 consecutive in-vitro fertilization (IVF)/intracytoplasmic sperm injection (ICSI) treatment cycles was studied between 1st September 1996 and mid-February 1997. This included cycles in the last 191 women receiving HMG and the first 162 taking FSHr. All were down-regulated using a gonadotrophin releasing hormone (GnRH) agonist long protocol method from day 1 of the cycle. Greater efficacy was seen in the HMG group in terms of days of stimulation required, need to increase dosage, cycle discontinuation, number of follicles punctured, the numbers of oocytes retrieved and their quality. The hormonal response to stimulation assessed by oestradiol concentrations on days 5, 8 and day of human chorionic gonadotrophin (HCG) was significantly lower in the FSHr group. The ratio of oestradiol per follicle and per oocyte was significantly lower in the FSHr group. There was a highly significant increase in cost with FSHr therapy. Clinical pregnancy rates were 14% per cycle with FSHr and 20% per cycle with HMG. Key words: efficacy/FSH/gonadotrophin/IVF/recombinant/urinary

Introduction

Human menopausal gonadotrophins (HMG) of various types have been used in in-vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) almost since their initial introduction (Fawzy and Harrison, 1994). In recent times, however, availability has been a problem. Batch-to-batch comparability, as with any biological product that uses a bioassay technique to assess potency, can be difficult to maintain. The response of patients can also be unpredictable. The introduction of recombinant human follicle stimulating hormone (FSH) Follitrophin Beta (Gonal F; Serono, Geneva, Switzerland) and Follitrophin Beta (FSHr, Puregon; Organon, Oss, The Netherlands) would therefore appear to be a milestone for infertility treatment. Availability, potency and therefore perhaps a predictable reaction to therapy should be guaranteed. Extensive testing has demonstrated safety and efficacy for ovarian stimulation in down-regulated IVF clinical trial situations for both the recombinant products (Out et al., 1995; Recombinant Human FSH Study Group, 1995). Indeed, prospective randomized multicentre trials found FSHr (Out et al., 1995) and Follitropin Alpha (Berg et al., 1997) to be more effective than urinary FSH (Urofollitrophin, Metrodin; Serono) in inducing follicular development and achieving pregnancy. Higher bioactivity of FSHr (Mannaerts et al., 1996), and a greater purity (Hard et al., 1990) and batch-to-batch consistency has also been demonstrated (Out et al., 1996).

As a participant in the above multicentre study our unit had had positive experience with FSHr. It was nevertheless disconcerting to learn that availability was to be in ampoules of 50 IU rather than the 75 IU used for the study. However, we decided to acquiesce with the manufacturer’s initial suggestion that 50 IU FSHr should be considered equivalent to 75 IU HMG, based on data from the study by Out et al. (1995), although it had not been designed with this question in mind. Eager anticipation soon turned to scepticism as our unit reported an apparent decrease in quality of all outcome parameters. Insufficient FSHr was thought to be the cause, but anecdote alone should not lead to change. Hence it was decided to study our experiences to date with FSHr compared with the last cohort of patients who had HMG (Humegon or Orgafol; Organon BV).

Materials and methods

All IVF and ICSI treatment cycles undergone between 1st September 1996 and 15th February 1997 were scrutinized retrospectively. Of the 353 consecutive cycles undergone, 107 were ICSI cycles; 158 patients received FSHr (50 ICSI) and 191 had HMG (57 ICSI); 95 of the latter received Humegon and 96 Orgafol (Organon). Up to 20th November, HMG only was used (total 181 cycles, Humegon = 91, Orgafol = 90); from then until 29th December all three products were used (total 49 cycles, Humegon = 4, Orgafol = 6, FSHr = 39) and from 1st January, FSHr only was used (total 119 cycles). The cycles were divided into two groups according to whether HMG or FSHr was used for ovarian stimulation; they did not differ significantly in general terms (Table I). Growth hormone was added to the treatment in two patients using FSHr and four on HMG. The additional four patients treated during the time with Follitrophin Alpha (Gonal F; Serono) were not included in the study. Of 204 patients treated for the first time, 91 received FSHr; of the remaining 145 patients having a second or subsequent treatment cycle, 67 received FSHr.
All couples had already been previously investigated, the main reason for their therapy defined and informed consent given by both partners to treatment. The programme was carried out in the Human Assisted Reproduction Unit, at the Rotunda Hospital Dublin. The modus operandi of this unit has been detailed fully elsewhere (Harrison et al., 1992), as has the down-regulation regime (Kondaveeti-Gordon et al., 1996).

Ovarian stimulation commenced when down-regulation was confirmed by ultrasound assessment of the ovaries and oestradiol assay (oestradiol <100 pmol/l). For patients on their first cycle where their basal day 3 FSH was <8.5 U/l, three ampoules (150 IU FSHr or 225 IU HMG) were used as the starter dosage, whereas if the day 3 FSH was >8.5 U/l, six ampoules (300 IU FSHr or 450 IU HMG) were employed. For patients who were in a subsequent treatment cycle the maximum drug dose used in the previous cycle was the starter dose. The drug dose was adjusted on day 5 depending on the ovarian response as assessed by vaginal ultrasound and serum oestradiol concentrations. These regimes were based on previous internal audit of our results modifying suggestions from Scott et al. (1989).

When deemed as an adequate response, the same dosage was continued through to the day of human chorionic gonadotrophin (HCG), but if the response was poor for those starting at three ampoules (150 IU/225 IU of FSHr/HMG, respectively) the dose was doubled to six ampoules (300 IU/450 IU) on day 5 of stimulation. Where women had already started at the higher dosage the number of ampoules was increased by two (400 IU/600 IU of FSHr/HMG, respectively). Poor ovarian response was defined as less than three follicles identified on ultrasound or serum oestradiol concentration of <300 pmol/l (conversion factor, 3.671) on day 5 of stimulation. All were re-assessed on day 8 after commencement of gonadotrophin and either allowed to continue if progress was satisfactory, or the cycle was stopped if the response was inadequate. For those who continued, the dosage was maintained until at least three or more follicles measuring >17 mm in diameter were obtained when HCG (Pregnyl; Organon) 10 000 IU i.m. was administered at a time to allow morning oocyte collection 33–36 h later.

Following appropriate laboratory procedure, embryos were replaced at the 2- to 4-cell stage 2 days later. Luteal phase support was given depending on the number of oocytes retrieved. If >10 oocytes were retrieved, progesterone pessaries (Cyclogest; Hoechst-Roussel, Uxbridge, UK) 200 mg were given for 14 days, while two doses of HCG 5000 IU i.m. were given on the day of embryo transfer and 2 days later if the number of oocytes retrieved was <10.

The clinical end-points examined were those as used in the majority of units to assess efficacy and efficiency of treatment. These were the treatment length (days), need to increase drug dosage on day 5 of ovarian stimulation, number of cancelled cycles, oestradiol concentrations on days 5, 8 and the day of HCG administration, number of follicles punctured at oocyte retrieval, oestradiol per follicle ratio, number of oocytes retrieved, fertilization rate, pregnancy rate and drug costs per treatment cycle. As it is only truly possible at present to assess oocyte maturity with ICSI and follow this directly into the quality of resultant embryos, the ICSI group of patients was also used specifically to compare the two drugs in terms of metaphase II oocytes and the quality of embryos transferred.

### Data analysis

The data were analysed using Data Desk 6.0b (Data Description Inc., Ithaca, New York, USA) and Stata 5.0 (Stata Corporation College Station, Texas, USA). \( \chi^2 \) tests were used to compare categorical outcomes, and t-tests used for normally distributed continuous variables. Medians rather than means are reported as the data were frequently skewed. Some variables were log transformed before being analysed, resulting in satisfactory correlations with the expected normal distribution (all \( r >0.900 \)). The odds ratios for the quality of embryos in each group was calculated using adjustment for clustered data implemented in Stata 5.0 procedure svylogit (Eltinge and Sribney, 1996).

### Results

#### General outcome

There was no significant difference between the 95 cycles treated with HMG, with a 1:1 ratio of FSH/luteinizing hormone (LH) and the 96 cycles treated with urinary FSH. Hence, these cycles were considered as one HMG group of 191 for comparison with the FSHr group. Of the 349 cycles analysed, 107 were ICSI attempts and were distributed equally between the two groups [FSHr 50/158 (31.6%); HMG 57/191 (29.8%)].

### Gonadotrophin stimulation requirement

The mean days required for adequate ovarian stimulation was lower (\( P = 0.004 \)) in the HMG group (9 days) than in the FSHr group (9.5 days), as shown in Table II.

A higher proportion of patients treated with FSHr had their dose increased (38.6% compared with 28.3% in the HMG

---

### Table I. General characteristics of the two groups

<table>
<thead>
<tr>
<th>Parameter</th>
<th>FSHr group (n = 158)</th>
<th>HMG group (n = 191)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (years)</td>
<td>35</td>
<td>34.48</td>
<td>NS</td>
</tr>
<tr>
<td>Mean duration of infertility (years)</td>
<td>5.56</td>
<td>5.70</td>
<td>NS</td>
</tr>
<tr>
<td>Mean day 3 cycle FSH (IU/l)</td>
<td>6.71</td>
<td>6.73</td>
<td>NS</td>
</tr>
<tr>
<td>Percentage of IVF or ICSI first attempts</td>
<td>59%</td>
<td>57%</td>
<td>NS</td>
</tr>
</tbody>
</table>

Main cause of infertility

<table>
<thead>
<tr>
<th>Male factor</th>
<th>57</th>
<th>58</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tubal</td>
<td>23</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Unexplained</td>
<td>41</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Endometriosis</td>
<td>26</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Anovulation</td>
<td>11</td>
<td>20</td>
<td>0.705</td>
</tr>
</tbody>
</table>

---

### Table II. Comparison of main outcome parameters

<table>
<thead>
<tr>
<th>Outcome</th>
<th>FSHr</th>
<th>HMG</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cycles treated</td>
<td>158</td>
<td>191</td>
<td></td>
</tr>
<tr>
<td>Ovarian stimulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days of stimulation (mean)</td>
<td>9.5</td>
<td>9</td>
<td>0.004</td>
</tr>
<tr>
<td>No. of cycles dose increased</td>
<td>61</td>
<td>24</td>
<td>0.041</td>
</tr>
<tr>
<td>No. of cycles stopped on day 8</td>
<td>19</td>
<td>18</td>
<td>NS</td>
</tr>
<tr>
<td>Oestradiol/follicle (pmol/l)</td>
<td>351.7</td>
<td>402.3</td>
<td>0.0026</td>
</tr>
<tr>
<td>Oestradiol/oocyte (pmol/l)</td>
<td>410.8</td>
<td>546</td>
<td>0.0007</td>
</tr>
<tr>
<td>Oocyte retrieval and fertilization rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follicles aspirated</td>
<td>11</td>
<td>13</td>
<td>0.008</td>
</tr>
<tr>
<td>No. of oocytes retrieved</td>
<td>9</td>
<td>10</td>
<td>0.044</td>
</tr>
<tr>
<td>Fertilization rate (% of oocytes)</td>
<td>51.7</td>
<td>53.4</td>
<td>NS</td>
</tr>
</tbody>
</table>

---

*Based on t-test on raw data.

**Based on t-test on log transformed data.

\( \chi^2 \) test.

NS, not significant.
group). This 10% difference was statistically significant (Pearson’s \( \chi^2 \) test, \( P = 0.041 \)).

There was a significant difference in the total drug dosage between FSHr and HMG (Table III). This was expected as we were equating 75 IU of HMG with 50 IU of FSHr and hence a patient would be started with either 225 IU of HMG (three ampoules) or with 150 IU of FSHr (three ampoules).

### Ovarian stimulation outcome

Overall (Table II), 37 patients (10.6%) had to be stopped on day 8 due to a poor ovarian response. This was not related to the drug used for stimulation (FSHr, \( n = 19 \); HMG, \( n = 18 \); not statistically significant), cause of infertility, number of attempts or the duration of infertility. In 18 cycles stopped (49%), the age of the female partner was \( > 40 \) years (FSHr, \( n = 10 \); HMG, \( n = 8 \)); however, all had basal day 3 FSH within parameters considered treatable (<15 U/l), as reported elsewhere (Scott et al., 1989; Mukherjee et al., 1996).

The median number of follicles aspirated on the day of oocyte retrieval was 11 in the FSHr group, significantly lower than the 13 aspirated in the HMG group (\( P = 0.008 \)). The median number of oocytes obtained was also significantly lower (\( P = 0.044 \)); nine in the FSHr group and 10 in the HMG group (total number = 3209 oocytes; FSHr = 1344, HMG = 1865). Distribution of the metaphase II oocytes retrieved as a percentage of the total oocytes in the 107 ICSI cycles where treatment progressed as far as oocyte collection was almost identical between the two groups: 318 metaphase II oocytes of a total 484 (65.7%) were obtained in the FSHr group while 441 metaphase II oocytes of a total 680 (64.8%) were obtained in the HMG group.

The median basal oestradiol concentration on the day of down-regulation in the FSHr group was 75.5 pmol/l, compared with 78.9 pmol/l in the HMG group (\( t = 1.56, P = 0.12 \)). However, median oestradiol concentrations on day 5, day 8 and the day of HCG were significantly lower in the FSHr group (\( P = 0.0002, <0.0001 \) and \(<0.0001 \), respectively) (Table IV). Similarly, patients treated with FSHr had significantly lower median concentrations of oestradiol per follicle (\( P = 0.0026 \)) and per oocyte (\( P = 0.0007 \)) than patients treated with HMG (Table II).

### Fertilization initiated and clinical pregnancies achieved

Out of the 312 cycles which reached the stage of oocyte retrieval, total failure of fertilization occurred in 30 cases (9.6%) (FSHr, \( n = 16 \); HMG, \( n = 14 \)). In those cases where fertilization of the oocytes was occurring (within 14–21 h post-insemination), no significant difference was found between the two groups in terms of fertilization rate (FSHr = 51.7%, HMG = 53.4%; Table II). Of the 1344 oocytes retrieved in the FSHr group, 700 were fertilized; by comparison, 1005 of a total 1865 oocytes in the HMG group were fertilized. As Table V shows, in patients treated with ICSI, where it is possible to select the best quality oocytes for insemination, there was no significant difference in the number or quality (number of cells and grade) of >2-cell-stage embryos replaced. It must be pointed out, however, that the confidence interval is wide, indicating that the odds of obtaining a higher quality zygote in the FSHr group could be anything from half as great to twice as great as in the HMG group.

In the 349 treatment cycles, 60 pregnancies were achieved of which 10 resulted in twins, two in triplets and one in quadruplets. (The quadruplets were delivered by Caesarean section at 30 weeks, and are alive and well.) Twenty-two pregnancies were achieved in the FSHr group (pregnancy rate 18.3% per transfer, 15.8% per oocyte retrieval and 13.9% per cycle started), while 38 were achieved in the HMG group (pregnancy rate 24.4% per transfer, 22% per oocyte retrieval and 20% per cycle started). The pregnancy rate was not significantly different between the two groups. The pregnancy rate achieved using HMG was similar to the certified overall pregnancy rate per cycle started for the year of 1996, which was 23% (Rotunda Hospital Dublin, Annual Clinical Report, 1996).

### Costs per cycle started (IRE)

The average drug list cost of the FSHr alone part of the treatment was IRE1824 while that of HMG was IRE720. The drugs bill for FSHr was significantly higher than that for HMG (\( t = 11.8, P < 0.0001 \)). It should also be noted that 25% of patients treated with FSHr were charged more than IRE2432, while the charge to the 25% most expensive cases treated with urinary gonadotrophins was >IRE11000. The cost per pregnancy achieved, as calculated from the overall drug bills of the groups, was IRE13 338 in the FSHr group, which was much higher than in the HMG group (IRE3923).
Discussion

There has long been a perceived need by those endeavouring to optimize ovarian stimulation with gonadotrophins before IVF/ICSI to have improved reliability, availability and predictability. The advent of recombinant FSH was therefore most welcome, particularly when clinical studies exploring these qualities were positive (Out et al., 1995; Recombinant Human FSH Study Group, 1995). The power of the study (Out et al., 1995) comparing Follitrophin Beta (FSHr) with urinary FSH seems to have tempted the manufacturers to extrapolate from their original study intentions to suggest that their brand of FSHr as used in that study was more effective than urinary FSH, and that lower doses and shorter treatment periods were therefore needed to achieve pre-ovulatory conditions (Edwards, 1997). It should be noted, however, that the study was based on increments of 75 IU FSH per ampoule and not the 50 IU subsequently made available and used in the study described here. While it may be possible to suggest that there is an increased potency of FSHr over HMG and that a lower dosage in terms of international units could be sufficient to gain the same response while at the same time lowering the risk of hyperstimulation, such claims need the support of published, verified dosage equivalency studies between HMG and FSHr.

Our investigation certainly has scientific weaknesses: it was designed ad hoc to study a clinical situation in which we found ourselves; it was retrospective; and there was the possibility of observer bias. All these problems had also bedevilled the randomized multicentre Follitrophin Beta study (Out et al., 1995). However, our statistician (R.C.) analysed the data blind. FSHr was not new to us. The period of study (Out et al., 1995) was bedevilled the randomized multicentre Follitrophin Beta study (Out et al., 1995). The power of the study (Out et al., 1995) comparing Follitrophin Beta (FSHr) with urinary FSH seems to have tempted the manufacturers to extrapolate from their original study intentions to suggest that their brand of FSHr as used in that study was more effective than urinary FSH, and that lower doses and shorter treatment periods were therefore needed to achieve pre-ovulatory conditions (Edwards, 1997). It should be noted, however, that the study was based on increments of 75 IU FSH per ampoule and not the 50 IU subsequently made available and used in the study described here. While it may be possible to suggest that there is an increased potency of FSHr over HMG and that a lower dosage in terms of international units could be sufficient to gain the same response while at the same time lowering the risk of hyperstimulation, such claims need the support of published, verified dosage equivalency studies between HMG and FSHr.

Our investigation certainly has scientific weaknesses: it was designed ad hoc to study a clinical situation in which we found ourselves; it was retrospective; and there was the possibility of observer bias. All these problems had also bedevilled the randomized multicentre Follitrophin Beta study (Out et al., 1995). However, our statistician (R.C.) analysed the data blind. FSHr was not new to us. The period of study (Out et al., 1995) was bedevilled the randomized multicentre Follitrophin Beta study (Out et al., 1995). The power of the study (Out et al., 1995) comparing Follitrophin Beta (FSHr) with urinary FSH seems to have tempted the manufacturers to extrapolate from their original study intentions to suggest that their brand of FSHr as used in that study was more effective than urinary FSH, and that lower doses and shorter treatment periods were therefore needed to achieve pre-ovulatory conditions (Edwards, 1997). It should be noted, however, that the study was based on increments of 75 IU FSH per ampoule and not the 50 IU subsequently made available and used in the study described here. While it may be possible to suggest that there is an increased potency of FSHr over HMG and that a lower dosage in terms of international units could be sufficient to gain the same response while at the same time lowering the risk of hyperstimulation, such claims need the support of published, verified dosage equivalency studies between HMG and FSHr.

As stated, the study was prompted by fears of inadequate dosage of FSHr as perceived by those working within the department. Therefore it was not a surprise to see that there was a statistically significant difference between the drugs in efficacy parameters: length of ovarian stimulation, the need to increase drug dosage, number of follicles and oocytes, and their quality. There was a 10% greater need to increase the drug dosage with FSHr on day 5 of stimulation which reached significance \( P = 0.041 \). This, coupled with the significantly lower oestradiol concentrations wherever measured in the treatment cycle, however, does suggest that while FSHr may eventually prove in appropriate studies to be more potent than comparable HMG preparations, advice that 50 IU FSHr is equivalent in effect to 75 IU HMG may be an underestimation. Indeed, it has been reported by Jansen and Van Os (1996) that the total drug dosage needed for stimulation may be more with FSHr than HMG, even though the difference they found was not significant.

Our findings in terms of lower oestradiol production in the FSHr cycles are in contrast to those of other studies (Out et al., 1996), but not all recombinant FSH products give identical results with IVF or are superior to urinary FSH in oestradiol response (Recombiant Human FSH Study Group, 1995). An explanation for our findings could be that the absence of LH in the FSHr preparation might affect follicle function. In down-regulated cycles using HMG with differing FSH/LH ratios, as well as FSHr, data from this unit (Gordon et al., 1997) show that the most favourable overall response in IVF-stimulated cycles is in patients receiving urinary gonadotrophins containing 75 IU FSH and 75 IU LH (1:1 ratio) per ampoule.

Minimum amounts of LH are required for increased FSH receptors on the granulosa cells (Shoham et al., 1993). Since FSHr has no LH as an impurity (but LH present in HMG), the increase in FSH receptors in granulosa cells may be affected, resulting in lower oestradiol response. Studies have shown, however, that the amount of remaining endogenous LH after profound pituitary down-regulation is still sufficient to support FSHr-induced oestrogen biosynthesis (Devroey et al., 1994). In our study, oestradiol response to stimulation was positively associated with oestradiol per follicle which was significantly lower in the FSHr group. Nevertheless, this did not influence pregnancy rates as reported elsewhere (Loumaye et al., 1997). Those cycles that did achieve pregnancy had a wide range of oestradiol per follicle.

There was an apparent 5% increase in pregnancy in the HMG group over the FSHr group which was not statistically significant. A study of this limited size, while making it possible to detect significant differences in many parameters, is unlikely to detect more than trends in pregnancy rates. Assuming that FSHr treatment gives rise to a 20% pregnancy rate and HMG to a pregnancy rate of 15%, then two groups each of 1250 subjects would be needed in order to have a 90% power to detect this difference. This study has only a 23% power of detecting a difference of 5% in pregnancy rate.

Although pregnancy rates were not significantly different, the significant difference in parameters such as cycle length, increased dosage requirement due to poor response, hormonal
response to stimulation and number of oocytes retrieved makes it difficult to justify the use of FSHr stimulation regimes in assisted reproduction programmes where patients are already under great emotional stress (Slade et al., 1997). In addition, the significant difference in cost–benefit between HMG and FSHr is worrying. In the Republic of Ireland the retail price per IU of FSHr is €0.76 compared with €0.24 per IU of HMG. This has led to a significant increase in the cost of treatment. If one considers pregnancy to be the ultimate outcome of treatment, the reality is that for the same pregnancy rate to be achieved, the cost of treatment with the FSHr product is actually three-fold higher.

This study was prompted by staff worries in the unit that the starter doses of FSHr used as recommended gave inferior results to those anticipated from previous experience. There was also unease at the apparent absence of true scientific data on which to base the recommended equivalence of 50 IU FSHr and 75 IU HMG. This study shows that, while fears of massive differences in efficacy can be discounted between recombinant product and urinary extract human gonadotrophin, as used in this study, there is doubt as to equivalency of potency, particularly when linking outcome parameters, oestradiol concentrations and cost–benefits. The urinary product of our choice, however, which gave us better results is no longer available, leaving not only ourselves but also other units with a dosage conundrum that urgently needs to be solved in a truly scientific manner. Dosage equivalency and prescribed regimes based on the previous menopausal urine extract product experience are clearly inadequate.

References


