Meta-analysis of recombinant versus urinary-derived FSH: an update

Hesham Al-Inany\textsuperscript{1,2,3}, Mohamed Aboulghar\textsuperscript{1,2}, Ragaa Mansour\textsuperscript{1} and Gamal Serour\textsuperscript{1}

\textsuperscript{1}The Egyptian IVF-ET Center, Maadi, Cairo and \textsuperscript{2}Cairo University, Department of Obstetrics and Gynecology, Cairo, Egypt

\textsuperscript{3}To whom correspondence should be addressed at: The Egyptian IVF-ET Center, 3 Street 161, Hadayek El-Maadi, Maadi, Cairo 11431, Egypt. E-mail: ivf@link.net

BACKGROUND: The study aim was to analyse the results of randomized controlled trials (RCTs) comparing recombinant FSH and urinary-derived FSH gonadotrophins [hMG, urinary purified FSH (FSH-P) and highly purified FSH (FSH-HP)] in an IVF/ICSI programme. METHODS: All published truly RCTs using a long protocol of GnRH agonists for down-regulation, were reviewed. Data of pregnancy rate per started cycle were extracted, and odds ratios (OR) calculated using a fixed effect model. Subgroup analysis was carried out to compare recombinant FSH (rFSH) with each product (hMG alone, FSH-P alone and FSH-HP alone). RESULTS: There was no statistically significant difference in the pregnancy rate per started cycle between rFSH and urinary-derived FSH gonadotrophins (OR 1.07; 95% CI 0.94–1.22). Subgroup analysis showed no statistically significant difference in the pregnancy rate per started cycle between rFSH versus hMG (OR 0.81; 95% CI 0.63–1.05), rFSH versus FSH-P (OR 1.24; 95% CI 0.98–1.58) and rFSH versus FSH-HP (OR 1.14; 95% CI 0.94–1.40). There was no significant heterogeneity of treatment effect across the trials. CONCLUSIONS: There is no evidence of clinical superiority in clinical pregnancy rate for rFSH over different urinary-derived FSH gonadotrophins. Additional factors should be considered when choosing a gonadotrophin regimen, including the cost, patient acceptability, safety and drug availability.

Key words: meta-analysis/ovarian stimulation/RCTs/recombinant FSH/urinary FSH

Introduction

Pharmaceutical preparations of human gonadotrophins play an important role in the treatment of human infertility, and have been used widely to stimulate follicular development in infertile women. During the 1970s, urinary hMG was the only gonadotrophin used in infertility treatment, but since the 1980s a variety of subproducts of urinary hMG have been produced with the intention of eliminating most or all of the LH content (Zafeiriou et al., 2000). During the mid-1990s, recombinant FSH (rFSH) was produced \textit{in vitro} from hamster ovarian cell cultures, and this step was considered a landmark in the production of gonadotrophins (Out et al., 1997).

The manufacture of human FSH using recombinant DNA technology (rFSH) makes its production independent of urine collection, and also guarantees a high availability of a biochemically pure FSH preparation (specific activity $>10\,000$ IU FSH/mg) that is free from urinary protein contaminants. The production process yields FSH with minimal batch-to-batch discrepancy (Bergh, 1999). The high purity and low immunogenicity allows subcutaneous administration. Many reports have demonstrated the efficacy of rFSH in ovarian stimulation (Recombinant Human FSH Study Group, 1995; Aboulghar et al., 1996; Out et al., 1996).

A meta-analysis has demonstrated that the use of urinary FSH was associated with a significantly higher clinical pregnancy rate than hMG (Daya et al., 1995), while a further meta-analysis showed rFSH to be superior to both purified FSH (FSH-P) and highly purified FSH (FSH-HP) in achieving clinical pregnancy rate (Daya and Gunby, 1999). Although it may be assumed that rFSH is more effective than hMG, this was not the case with recent randomized controlled trials (RCTs) that showed equivalent efficacy (Gordon et al., 2001; Ng et al., 2001; Strehler et al., 2001; Westergaard et al., 2001; Diedrich, 2002).

The aim of the present study was to update the evidence comparing rFSH and urinary-derived FSH gonadotrophins. The concept was that urinary FSH-P and FSH-HP are subproducts of hMG, and hence should be grouped together when compared with rFSH, after which each is compared separately. In support of this concept, in clinical practice these products are given for the same purpose, for the same patients with similar effects, and in similar doses.

Materials and methods

On conducting a MEDLINE search and searching the Cochrane Menstrual Disorders and Subfertility Review Group specialized register of randomized controlled trials, as well as the abstracts of the European Society for Human Reproduction and Embryology (ESHRE) and the American Society for Reproductive Medicine
### Table I. Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Methods</th>
<th>Participants</th>
<th>Interventions</th>
<th>Concealment of allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alvino et al.</td>
<td>RCT, unconcealed allocation</td>
<td>Couples with different causes of infertility, female age (mean 31 years)</td>
<td>Long luteal GnRH agonist protocol, then rFSH alpha s.c. versus uFSH 150 IU daily for 5 days, then adjusted</td>
<td>C</td>
</tr>
<tr>
<td>Berger et al.</td>
<td>Multicentre RCT, allocation not described</td>
<td>Couples with all causes of infertility including male factor, female age &lt;40 years</td>
<td>Long protocol, rFSH beta 150 IU daily fixed dose versus uFSH-HP225 150 IU daily fixed dose</td>
<td>B</td>
</tr>
<tr>
<td>Bergh et al.</td>
<td>Multicentre RCT, allocation using a computer-generated list</td>
<td>Couples with different causes of infertility female age (mean 32 years)</td>
<td>Long protocol then rFSH alpha versus uFSH-HP 150 IU s.c. daily for 6 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Franco et al.</td>
<td>RCT, allocation using a randomization table</td>
<td>Male factor infertility undergoing ICSI, female age (mean 31 years)</td>
<td>Long protocol, then rFSH alpha versus uFSH-HP 150 IU s.c. daily for 6 days, then adjusted</td>
<td>B</td>
</tr>
<tr>
<td>Frydman et al.</td>
<td>Multicentre RCT, allocation, using coded medications</td>
<td>Couples with different causes of infertility, female age (mean 31 years)</td>
<td>Long protocol, then rFSH alpha versus uFSH-HP 150 IU s.c. daily for 6 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Ghosh et al.</td>
<td>RCT with allocation by sealed envelopes</td>
<td>Couples with different causes of infertility</td>
<td>Long protocol, then rFSH alpha 150 IU daily versus uFSH-HP 225 IU daily for 5 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Gordon et al.</td>
<td>4-arm RCT, assessor-blind, allocation concealed using sealed envelopes</td>
<td>Couples with different causes of infertility, female age (median 32 years)</td>
<td>Long protocol, then hMG (Humegon) versus rFSH (Pregure, Follitropin beta) 225 IU daily for 5 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Hedon et al.</td>
<td>Multicentre RCT, allocation concealed using coded medications</td>
<td>Couples with different causes of infertility, female age (mean 32 years)</td>
<td>Long protocol, then rFSH beta versus uFSH 150–225 IU i.m. daily for 4 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Hoomans et al.</td>
<td>Multicentre RCT, allocation concealed, using coded medications. Couples with different causes of infertility. PCOS cases were excluded; female age (mean 33) years</td>
<td>Long protocol, then GnRH agonist in the long luteal or long follicular protocol with buserelin, then rFSH 150 IU s.c. daily fixed dose versus uFSH-HP 225 IU s.c. daily fixed dose</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Lenton et al.</td>
<td>Multicentre RCT, allocation by using sealed envelopes</td>
<td>Couples with different causes of infertility, female age (mean 32 years). PCOS cases were excluded; female age (mean 33) years</td>
<td>Long protocol, then rFSH alpha versus uFSH-HP 150 IU s.c. daily for 6 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Machado et al.</td>
<td>RCT, allocation by drawing straws</td>
<td>Couples with different causes of infertility, female age (mean 35 years)</td>
<td>Long protocol, then rFSH beta versus uFSH-HP</td>
<td>B</td>
</tr>
<tr>
<td>O’Dea et al.</td>
<td>Multicentre RCT, method of allocation not described</td>
<td>Couples with different causes of infertility female age 18–38 years</td>
<td>Long protocol, then GnRH agonist protocol with leuprolide acetate 0.5 mg s.c. daily, then rFSH alpha versus uFSH 225 IU daily for 5 days, then adjusted</td>
<td>B</td>
</tr>
<tr>
<td>Out et al.</td>
<td>Multicentre RCT, concealed allocation, using coded medications</td>
<td>Couples with different causes of infertility (male factor excluded); female age (mean 32 years)</td>
<td>Long protocol, then rFSH beta versus uFSH 150 or 225 IU i.m. daily for 4 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>RHFSG</td>
<td>Multicentre RCT, allocation by sealed envelope</td>
<td>Couples with different causes of infertility, female age (mean 32 years)</td>
<td>Long protocol, then rFSH alpha s.c. versus uFSH 150 or 225 IU i.m. daily, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Schats et al.</td>
<td>Multicentre RCT, assessor-blind, allocation by sealed envelopes</td>
<td>Couples with different causes of infertility, female age (mean 31 years)</td>
<td>Long protocol, then rFSH alpha versus uFSH-HP at a fixed dose of 150 IU s.c. daily (dose reduction permitted if response excessive</td>
<td>A</td>
</tr>
<tr>
<td>Diedrich</td>
<td>Multicentre RCT, allocation by randomization list, randomization blocks</td>
<td>Couples with different causes of infertility except PCO, female age (mean 31 years)</td>
<td>Long protocol, then hMG (Menopur) versus rFSH (Follitropin Beta, Puregon), 225 IU s.c. daily for 5 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Ng et al.</td>
<td>RCT. Allocation by computerized randomization</td>
<td>Severe male factor; female age (median 33 years)</td>
<td>Long protocol, then hMG (Pergonal) versus rFSH (Follitropin alpha, Gonal-F) given as 300 IU in the first 2 days, followed by 150 IU daily. Long protocol, then hMG (Menogon) versus rFSH (Follitropin alpha, Gonal-F) given at 225 IU for 7 days.</td>
<td>A</td>
</tr>
<tr>
<td>Westergaard et al.</td>
<td>RCT. Allocation by computerized randomization</td>
<td>Couples with different causes of infertility; PCOS excluded. Female age (mean 31 years)</td>
<td>Long protocol, then follitropin alpha versus FSH-HP 225 IU/d for 3 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Germond et al.</td>
<td>RCT, allocation by random list Multicentre RCT. Computer-generated randomization blocks</td>
<td>Couples with different causes of infertility, female age (mean 37.7 years)</td>
<td>Long protocol, then follitropin beta versus FSH-HP 225 IU/d for 5 days, then adjusted</td>
<td>A</td>
</tr>
<tr>
<td>Dickey et al.</td>
<td>RCT, allocation by random list Multicentre RCT. Computer-generated randomization blocks</td>
<td>Couples with different causes of infertility, female age (mean 32 years)</td>
<td>Long protocol, then follitropin alpha versus FSH-HP 225 IU/d for 5 days, then adjusted</td>
<td>A</td>
</tr>
</tbody>
</table>
(ASRM) meetings from 1999 to 2001, all RCTs comparing rFSH with urinary-derived FSH gonadotrophins were identified.

The methodology used herein included only the true RCTs comparing rFSH with urinary-derived FSH gonadotrophins for ovarian stimulation in subfertile women undergoing IVF/ICSI. Quasi-randomized trials were excluded because they are known to give inflated treatment effects. Only those trials in which pituitary down-regulation was achieved using the long protocol were included, as amalgamation of results of the different protocols would be of uncertain value (Agrawal et al., 2000). The long protocol was selected as it has been the most widely used protocol for pituitary down-regulation during the past two decades (Al-Inany and Aboulghar, 2002).

Studies were identified by a literature search using a combination of the following key words: FSH, recombinant, urinary, gonadotrophins, hMG, uFSH-Purified, uFSH-Highly Purified, pregnancy, and randomized controlled trial. Review articles and abstracts of major scientific meetings and conference proceedings [ESHRE, ASRM, International Federation of Fertility Societies (FFS)] from 1999 until 2002 were reviewed. The main outcome measure was limited to clinical pregnancy rate per cycle started. Data of clinical pregnancy rate per cycle started were extracted (Al-Inany and Aboulghar, 2002).

The dichotomous data results for each study were expressed as an odds ratio (OR) with 95% confidence intervals (CI). These results were combined for meta-analysis with RevMan software (using the Mantel–Haenszel method) (Mantel and Haenszel, 1959). In the graphical display of meta-analyses, a benefit from rFSH would be displayed graphically to the right of the centre-line, while a benefit from urinary-derived FSH gonadotrophins would be displayed graphically to the left of the centre-line. Differences between the studies were tested using the Breslow–Day test for homogeneity performed across all trials (Breslow and Day, 1980).

In the present meta-analysis, the results were pooled using a fixed-effects model only after confirming that statistical heterogeneity was not present (i.e. the observed treatment effects in individual trials were not statistically significantly different from the overall pooled estimate of the treatment effect). A funnel plot analysis was performed in order to detect any publication bias.

Subgroup analysis was carried out to check the stability of the results reached by pooling data of all studies in general because urinary-derived FSH gonadotrophins are not identical in their chemical structure, despite belonging to one family.

Results

The present meta-analysis included 20 studies (Table 1), 15 of which were reported in the updated meta-analysis comparing rFSH versus urinary FSH (Daya and Gunby, 2002). A total of 12 trials was identified after the updated meta-analysis (Daya and Gunby, 2002) had been published. Included among these 12 trials were three that compared rFSH and hMG (Ng et al., 2001; Westergaard et al., 2001; The European and Israeli Study Group on highly purified menotropin versus recombinant follicle-stimulating hormone, 2002), and two that compared rFSH with FSH-HP (Germond et al., 2000; Dickey et al., 2002). The other trials were excluded due to no down-regulation (Soong et al., 1999), the use of a GnRH agonist short protocol (Strehler et al., 2001), the use of rFSH versus combined rFSH and hMG (Mahmoud et al., 2001), and the non-RCT nature of the study (Gomez-Parga et al., 1999; Sharma et al., 2001; Meo et al., 2002).

Two other studies were also excluded (Manassiev et al., 1997, which was cited in Daya and Gunby, 2002; and Serhal et al., 2000, which was identified during the search). Both studies used a quasi-randomization method: Manassiev et al. randomized subjects according to their residence area, while Serhal et al. randomized subjects by alternating weeks. One other trial (Ferraretti et al., 1999, cited in Daya and Gunby, 2002) was also excluded as the authors did not use down-regulation in their study. Another trial (Kornilov et al., 1999) was also excluded as the authors reported pregnancy rate per embryo transfer rather than per started cycle. In addition, the groups were non-matching (40 subjects received hMG and 28 received rFSH), and there was a significant age difference between the two groups despite claimed randomization. The method of randomization was not clear, and the authors were contacted for additional information; no response was obtained, however.

Although many of the included studies were in fact small, pooling the data from all 20 (giving a total of 4610 IVF/ICSI cycles) resulted in no statistically significant differences in the clinical pregnancy rate per cycle started between rFSH and urinary-derived FSH gonadotrophins (Figure 1) (OR 1.07; 95% CI 0.94–1.22) or between rFSH and various types of urinary-derived FSH gonadotrophins (hMG, FSH-P and FSH-HP) (Figures 2, 3 and 4).

Although the Kornilov trial (Kornilov et al., 1999) was excluded, adding these data to the meta-analysis did not change the overall significance (OR 1.09; 95% CI 0.95–1.24). Likewise, the addition of data from both the Manassiev trial (Manassiev et al., 1997) and the Serhal trial (Serhal et al., 2000) did not affect the overall results (OR 1.05; 95% CI 0.93–1.20).

It was planned to undertake sensitivity analyses if there were more than 10 trials included in the meta-analysis to examine the stability of the results in relation to the influence of pharmaceutical companies (Figures 5 and 6). There was still no significant difference seen between rFSH and urinary-derived FSH gonadotrophins in the studies, whether they were sponsored by pharmaceutical companies, or not. A funnel plot analysis confirmed that selective publication was unlikely to have been a source of bias in the present meta-analysis (Figure 7).

Discussion

hMG contains FSH and LH in a 1:1 ratio with urinary proteins. Purified hMG can be processed so that LH is separated from bulk material by using highly specific monoclonal antibodies. Thus, FSH together with minimal amounts of LH and urinary protein are collected and lyophilized for use as FSH-P. More recently however, a more direct process was used in which highly specific monoclonal antibodies could be selectively bound to FSH molecules in the hMG bulk material. The unbound urinary protein could then be removed along with the LH, thus creating FSH-HP. Accordingly, the FSH content and type is the same in all types of the urinary-derived FSH gonadotrophins, the only difference lying in the content of LH and urinary proteins. The aim of the present study was to
compare rFSH with all types of urinary-derived FSH gonadotrophins (hMG, FSH-P and FSH-HP) together. Furthermore, a subgroup analysis was carried out to compare, separately, rFSH with each of the three types of urinary-derived FSH gonadotrophins.

It might be argued that hMG and urinary FSH are not equal, as hMG contains equal amounts of FSH and LH (75 IU of each per ampoule); by contrast, the FSH-P preparation contains only a small amount (<5%) of LH, while FSH-HP contains <1% LH. Therefore, it may not be justified to include the hMG/rFSH...
trials in the meta-analysis on urinary FSH versus rFSH. However, this argument is not believed valid, as FSH-P and FSH-HP are subproducts from hMG, and have the same type and content of FSH. These drugs may not be similar, but all of them contain the same dose of the same family of FSH—the only differences lie in their LH and protein contents. Accordingly, FSH-P and FSH-HP should be grouped together when compared with rFSH, after which subgroup analysis can be carried out between each type of gonadotrophin to rFSH. In support of this concept, a recent report (Sykes et al., 2001) has grouped the three forms of urinary-derived FSH gonadotrophins together (hMG, FSH-P and FSH-HP) in comparing their cost-effectiveness with that of rFSH.

In the present meta-analysis, a subgroup analysis was carried out to confirm the stability of results among all groups. There was no superiority for recombinant FSH over either hMG, FSH-P or FSH-HP (Figures 2, 3 and 4).

A subgroup analysis according to IVF or ICSI (Daya and Gunby, 1999) was not carried out because it is believed that as long as the trials were truly randomized, then any differences observed in pregnancy rate could be attributed to the effect of gonadotrophins rather than to either IVF or ICSI. The purpose

---

**Figure 3.** Comparison between recombinant FSH and purified urinary FSH (FSH-P).

**Figure 4.** Comparison between recombinant FSH and highly purified urinary FSH (FSH-HP).
of randomization was to generate both control and experimental groups that were likely to be similar with respect to known and unknown co-variates. Accordingly, any differences observed in pregnancy rate could be attributed to the effect of gonadotrophins, whether recombinant or urinary in origin.

Neither was any subgroup analysis according to the type of rFSH (Puregon® or Gonal-F®) performed, as was carried out by others (Daya and Gunby, 1999). This subgroup analysis does not allow direct comparison between both drugs, and this markedly limits any conclusion that can be drawn from such analysis. Bearing in mind that several prospective controlled trials have now been published in the medical literature comparing Puregon and Gonal-F (Tulppala et al., 1999; Brinsden et al., 2000; Harlin et al., 2000), it was found inappropriate to carry out such subgroup analysis. These trials each showed a non-significant difference between the two...
recombinant drugs. Interestingly, no direct RCT has been carried out to compare FSH-HP with FSH-P, most likely because rFSH was developed soon after FSH-HP and there was no benefit in comparing the two. This demonstrates the lack of available evidence to support the efficacy of FSH-HP.

Validity score assessment (Daya and Gunby, 1999) was not carried out as the policy of the Cochrane Menstrual Disorders Subfertility Group does not recommend the use of a validity scoring system. Because there is no ‘gold standard’ for the ‘true’ validity of a trial, the possibility of validating any proposed scoring system is limited. While it is possible to apply basic principles of measurement to the development of a scale to assess the validity of randomized trials, the relationship between such a score and the degree to which a study is free from bias is not clear. None of the currently available scales for measuring the validity or ‘quality’ of trials can be recommended without reservation (Clarke and Oxman, 2002).

Thus, the present meta-analysis showed that there is no clinical superiority for rFSH over other urinary gonadotrophins. Moreover, there are certain concerns regarding the use of rFSH. First, it has been suggested that GnRH agonist down-regulation in some normogonadotropic women may result in profound suppression of LH concentration, impairing adequate estradiol synthesis (Fleming et al., 2000). Therefore, in such cases when rFSH is used for ovarian stimulation after GnRH agonist down-regulation, very low serum LH concentrations may adversely affect IVF outcome (Levy et al., 2000).

Second, in spite of the proven efficacy of rFSH, its widespread use has been hampered by its relatively high cost as compared with urinary-derived FSH gonadotrophins (Sykes et al., 2001). In many countries (including Egypt), patients pay for assisted reproductive treatment, and this has subsequent financial implications for both the infertile couple and the healthcare system. The decision to adopt a more expensive treatment could result in fewer couples receiving IVF treatment. An economic analysis is therefore required in order to guide both couples and aid decision-makers, based on the new data presented in the present meta-analysis.

Recently, the National Institute of Clinical Excellence (NICE) announced that it will be analysing the cost-effective-
the cost-effectiveness of one intervention over another based on the most up-to-date evidence available.

Acknowledgements
The authors had no potential conflicts of interest during these studies.

References


