Dear Sir,

We would like to react on the ongoing discussion of the meta-analysis on all performed recombinant versus urinary FSH (rFSH versus uFSH) clinical trials in assisted reproduction (Daya and Gunby, 1999) by making the following points.

We endorse the initiative by Daya and Gunby in performing this meta-analysis. It was reassuring to see that their meta-analysis basically confirmed our own analysis using only follitropin beta data, i.e. that rFSH results in statistically and clinically significantly more pregnancies when compared with urinary gonadotrophins (Out et al., 1997a).

However, we very much regret that the paper has been misinterpreted by many people as showing that follitropin alpha is more superior to uFSH than follitropin beta. Although Daya and Gunby state that that was never the intention of their analysis and that they feel that ‘any such conclusion is invalid and inappropriate’ (Daya and Gunby, 2000), they themselves gave food to this suggestion by speculating about it in their paper; they suggest that slight differences in the isohormone profile between follitropin alpha and beta might account for the differences in embryo quality and, hence, in the pregnancy rates.

As Daya and Gunby point out, we have indicated in the
Dear Sir,

Thank you for the opportunity to respond to the comments made by Out et al.

We would like to begin by pointing out that their meta-analysis involved only three trials, and included one trial that compared follitropin β with human menopausal gonadotrophin (HMG) (Out et al., 1997a). Our meta-analysis (Daya and Gunby, 1999) was more comprehensive, because it included many more trials, and more homogeneous, because it was restricted to comparing recombinant FSH (rFSH) only with urinary FSH (uFSH). We have previously demonstrated that uFSH is more efficacious than HMG (Daya et al., 1995). Consequently, it should be of no surprise that the inclusion of the trial involving the clinically less effective HMG would bolster the finding that follitropin β is more efficacious than urinary gonadotrophins. However, when the comparison is restricted to follitropin β and uFSH, a difference in favour of rFSH still is observed but (based on the currently available total sample size in this subgroup), it is not statistically significant, a conclusion that remains in the updated meta-analysis we have performed in which a total of 18 trials comparing 3421 subjects was evaluated (Daya and Gunby, 2000a).

The rest of their comments focus on the subgroup analysis we undertook to compare follitropin α and β against uFSH. Most of the points they make have already been made in a previous letter-to-the-editor, to which we have responded (Daya and Gunby, 2000b). Nevertheless, it is important to reiterate our position. The fact that the two follitropins should be evaluated separately has already been stressed by Dr Out himself on the grounds that the manufacturing and purification procedures are different and lead to dissimilar pharmaceutical formulations (Out et al., 1997b). The small difference in magnitude of the absolute treatment effect regarding the pooled clinical pregnancy rates in IVF when trials of follitropin α versus uFSH (5.2%, 95% confidence limits = 0.2, 10.2) were compared with trials of follitropin β versus uFSH (3.7%, 95% confidence limits = −1.5, 8.8) suggests that variability from sampling error may be one explanation for this observation. However, the minor difference in isohormone profiles also implies that there may be subtle differences in clinical outcomes between these two follitropins. We readily admit that our meta-analysis was not capable of directly evaluating the two follitropins against each other, and eagerly await the results of adequately powered randomized trials that compare these preparations with clinically relevant end-points. Out et al. refer to data from one randomized study (n = 344) that showed similar clinical pregnancy rates (33.5% per cycle with follitropin α versus 32.9% with follitropin β) (Tulppala et al., 1999) and one non-randomized cohort study (Harlin et al., 2000) (which, by design, is fraught with bias), also demonstrating similar pregnancy rates (29.1 and 28.1% respectively). A smaller, randomized trial (n = 44) showing a much greater difference in the estimates of clinical pregnancy per cycle (31.8% for follitropin α versus 18.2% for follitropin β) was not mentioned (Brinsden et al., 2000).

Harlin, J., Czemiczky, G., Wramsby, H. and Fried, G. (2000) Recombinant follicle-stimulating hormone (follitropin beta) against human menopausal gonadotrophin (HMG) (Out et al., 1997a). The subgroup analysis of follitropin alpha and beta (as carried out by Daya and Gunby) therefore, makes sense. Meanwhile, two published direct comparisons between both follitropins (Tulppala et al., 1999; Harlin et al., 2000) justify the conclusion that the slight differences between the two molecules do not have any clinical significance. Therefore, in view of these new data, the subgroup analysis loses its relevance.

Finally, it is worthwhile noting that the uFSH control groups differ significantly between the follitropin alpha and beta groups (see Table I). The fact that in the IVF studies, the odds favour follitropin alpha and not beta appears to be mainly due to the fact that, compared to the follitropin beta studies, the control patients perform significantly worse in the follitropin alpha studies (difference: −5.5%; 95% confidence interval: −10.7 to −0.3%; two-tailed P = 0.037). This should have been recognized by the authors.

References


Henk J. Out1

Organon Laboratories, Science Park, Cambridge CB4 0FL, UK
Eric H.M.Hoomans and W.N.G.M.de Laat
N.V. Organon, Oss, The Netherlands

1To whom correspondence should be addressed. E-mail: h.out@organon.cge.akzonobel.nl

Table I. Pregnancy rates in recombinant (r) FSH and urinary (u) FSH-treated women

<table>
<thead>
<tr>
<th>Studies</th>
<th>Product</th>
<th>Pregnant/cycle</th>
<th>Pregnancy rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follitropin alpha studies</td>
<td>Follitropin alpha</td>
<td>151/542</td>
<td>27.9</td>
</tr>
<tr>
<td></td>
<td>uFSH</td>
<td>115/530</td>
<td>21.7*</td>
</tr>
<tr>
<td>Follitropin beta studies</td>
<td>Follitropin beta</td>
<td>224/725</td>
<td>30.9</td>
</tr>
<tr>
<td></td>
<td>uFSH</td>
<td>139/511</td>
<td>27.2*</td>
</tr>
</tbody>
</table>

*p = 0.037.
It is readily apparent that the sample sizes of the two published randomized trials are woefully inadequate; any claims that the two follitropins are equivalent will require supporting evidence from much larger studies with sample size of several thousand subjects. Consequently, it is important to emphasize that, with smaller studies, the absence of statistical significance when the pregnancy rates between the two follitropins are compared does not imply that the two preparations are equivalent. Thus, based on the evidence from the two studies cited by Out et al., their statement that ‘the subgroup analysis [we performed] loses its relevance’ is premature, invalid, and biased.

Much fuss has been made about the different clinical pregnancy rates observed in the control groups. Out et al. try to readdress this issue, on which we have already commented, (Daya and Gunby, 2000b) by ‘demonstrating’ in their table that there is a significant difference in the pregnancy rates obtained with uFSH in the follitropin α studies compared with the follitropin β studies. This comparison is statistically incorrect and is as meaningless as a comparison of outcomes with follitropin α and β across studies. Differences in pregnancy rates among centres are expected, which is the reason why inferences must be made only from direct comparisons within studies and not across studies.

In conclusion, we contend that the arguments made by Out et al. are invalid because they are based on the use of methodologically poor studies, underpowered controlled trials, statistically incorrect hypothesis testing procedures and selective inclusion of trials to support their position. It is important that the results of systematic reviews be presented accurately and completely, as we have done, so that inappropriate and biased inferences can be avoided. We believe we have achieved these objectives.

References


