One additional consequence of using an inappropriately high starting dose of rFSH is also worth addressing. The authors report that the increased number of oocytes retrieved in the recrFSH group was not accompanied by a higher proportion of top-quality embryos (TQE). The authors then speculate that the higher proportion of TQEs in the HP hMG group may be due to the LH-activity present in HP hMG. This could be misleading, as proportions are dependent on both the numerator and the denominator. As noted, more oocytes were retrieved from women receiving rFSH. If only 2–3 TQEs are generated per cohort/retrieval, the proportion will be lower when more oocytes are retrieved. It seems reasonable to speculate, then, that above a certain threshold no additional TQEs will be obtained. This alternate speculation is consistent with the findings of this study and again reflects the choice of a higher-than-necessary dose of rFSH rather than the consequence of the added LH.

I think the following quote from Mark Twain seems particularly apt: ‘There is something fascinating about science - one gets such wholesale returns of conjecture out of such a trifling investment of fact.’

I believe that it is more interesting to explore how protocols can be optimized and individualized rather than indiscriminately adding LH (i.e. hMG to all IVF patients) without evidence of significant benefit. It has been demonstrated that such an individualized approach to gonadotropin stimulation, based on predictive factors, increased the proportion of appropriate ovarian responses, decreased the need for dose adjustments, and was associated with improved outcomes (Popovic-Todorovic et al., 2003).

Finally, I believe that an optimized treatment protocol should also consider other important factors such as tolerability, convenience, compliance and patient preference. Unfortunately, tolerability is not discussed in this article despite a notation adding LH (i.e. hMG to all IVF patients) without evidence of significant benefit. It has been demonstrated that such an individualized approach to gonadotropin stimulation, based on predictive factors, increased the proportion of appropriate ovarian responses, decreased the need for dose adjustments, and was associated with improved outcomes (Popovic-Todorovic et al., 2003).

We would like to thank Dr Trew for his interest in our study and his acknowledgement of the careful design of the MERIT (Menotrophin versus Recombinant FSH in vitro fertilisation trial). We appreciate this opportunity to explain the considerations supporting specific methodology choices made and to clarify the potential impact of these decisions on the outcome of the trial.

Before we address the remarks to the protocol design, we will first clarify the issue of superiority as this is important to understand the underlying assumptions in the study and the associated conclusions. MERIT was powered to detect an odds ratio of one treatment versus the other treatment of 1.67, corresponding to an absolute difference between treatments in ongoing pregnancy rates of 10%: 32% with HP-hMG (Menopur) and 22% with rFSH (Gonal-F). The actual ongoing pregnancy rates in the study were 27% with HP-hMG and 22% with rFSH. Although the magnitude and direction of this finding is in line with the 4–5% difference found in the most recent meta-analysis of menotrophins versus rFSH (van Wely et al., 2003), superiority was not established in this trial because the assumption of a 10% difference was not achieved. Rather than contradicting the existing evidence, the findings of MERIT reinforce the current literature suggesting a higher efficacy with menotrophins over rFSH in a long agonist protocol. As mentioned in our article (Nyboe Andersen et al., 2006), the issue of statistical significance should be addressed by an updated meta-analysis.

If concomitant fertility products are used in this type of efficacy trial, these should be approved products (i.e. with documented efficacy and safety), readily available to the clinicians and used within the recommended labelling. MERIT followed that principle and used approved products.

Letters to the Editor

Reply: Comparing highly purified hMG and rFSH in patients undergoing IVF

Sir,

We would like to thank Dr Trew for his interest in our study and his acknowledgement of the careful design of the MERIT (Menotrophin versus Recombinant FSH in vitro fertilisation trial). We appreciate this opportunity to explain the considerations supporting specific methodology choices made and to clarify the potential impact of these decisions on the outcome of the trial.

Before we address the remarks to the protocol design, we will first clarify the issue of superiority as this is important to understand the underlying assumptions in the study and the associated conclusions. MERIT was powered to detect an odds ratio of one treatment versus the other treatment of 1.67, corresponding to an absolute difference between treatments in ongoing pregnancy rates of 10%: 32% with HP-hMG (Menopur) and 22% with rFSH (Gonal-F). The actual ongoing pregnancy rates in the study were 27% with HP-hMG and 22% with rFSH. Although the magnitude and direction of this finding is in line with the 4–5% difference found in the most recent meta-analysis of menotrophins versus rFSH (van Wely et al., 2003), superiority was not established in this trial because the assumption of a 10% difference was not achieved. Rather than contradicting the existing evidence, the findings of MERIT reinforce the current literature suggesting a higher efficacy with menotrophins over rFSH in a long agonist protocol. As mentioned in our article (Nyboe Andersen et al., 2006), the issue of statistical significance should be addressed by an updated meta-analysis.

If concomitant fertility products are used in this type of efficacy trial, these should be approved products (i.e. with documented efficacy and safety), readily available to the clinicians and used within the recommended labelling. MERIT followed that principle and used approved products.

References


and doses of GnRH agonist for down regulation, hCG for triggering final maturation and progesterone for luteal support. A protocol design issue raised by Dr Trew is the choice of gering final maturation and progesterone for luteal support, and doses of GnRH agonist for down regulation (100 μg triptorelin/day). It should be stressed that no study has been properly designed to address if there are differences in pregnancy rate between triptorelin 100 μg and other GnRH agonists. The largest available randomized study (n = 246) in women undergoing IVF comparing triptorelin 100 μg once daily to buserelin 0.3 mg twice daily and leuprolrein 0.2 mg once daily (Parinaud et al., 1992) indicated similar results among the three GnRH agonists in terms of cycle cancellation rates and clinical pregnancy rates. The data on comparative efficacy of triptorelin 100 μg versus lower unapproved doses in terms of pregnancy rates are inadequate and the limited existing data are not suggestive of higher pregnancy rates with lower dose compared to those obtained with 100 μg (Janssens et al., 2000). Furthermore, the more interesting point to note is that the data from two large trials comparing menotrophins and rFSH (Westergaard et al., 2001; Diedrich et al., 2003) showed that the type/form/route of administration of GnRH agonist did not affect the magnitude of the difference in pregnancy rate between HP-hMG and rFSH (Arce et al., 2005). Thus, there are no data supporting that the type of GnRH agonist chosen introduced a bias in favour of the HP-hMG group. The hypothesis of a substantially different effect of using another GnRH agonist or lower doses of triptorelin on clinically relevant outcomes is thus speculative. A reduction in the triptorelin dose approved for down regulation would clearly need to be evidence-based and cannot be introduced as clinical strategy based on hypotheses.

Dr Trew refers to 225 IU FSH as an inappropriately high starting dose. It must be made clear that the starting dose of 225 IU for down regulated patients is within the approved European labelling and is the approved starting dose in the United States for both gonadotrophin preparations. Results from a comparative study of 150 IU versus 225 IU of rFSH (Gonal-F) indicated a cancellation rate for inadequate response of 15.0% for 150 IU compared with 3.3% for 225 IU and similar pregnancy rates between both dose groups (Yong et al., 2003). In MERIT, less than 5% of the cycles in both groups were cancelled because of inadequate response, the moderate/severe early OHSS rate was 2%, the cancellation rate because of excessive response was 2% and 90% of the patients had the dose maintained (60–66%) or even increased (25–33%) at day 6 of stimulation. Thus, the starting dose of 225 IU was not associated with safety concerns and did not compromise efficacy evaluation.

The elevated serum progesterone at the end of stimulation in the rFSH group compared with the HP-hMG group is a more complex issue than the effect of the FSH dose. Actually, in MERIT, the total FSH dose administered and the serum FSH concentration on the day of hCG administration were significantly higher in the HP-hMG group compared with the rFSH group, and yet the progesterone concentration was significantly higher in rFSH treated patients than in HP-hMG treated patients at the end of stimulation. The development of more follicles with rFSH could not entirely explain the increased progesterone tonus observed with rFSH, as serum progesterone remained significantly higher with rFSH compared with HP-hMG when adjusting for the number of developed follicles (Smitz et al., 2007). The link between FSH exposure and elevations of serum progesterone levels during stimulation has also been observed with starting doses of 150 IU FSH (Filicori et al., 2002). We have recently published the endocrine data for serum and follicular fluid from MERIT, including a discussion of the paracrine regulation factors that could explain differences in progesterone production in the follicle (Smitz et al., 2007).

Live birth is the outcome of interest to couples seeking infertility treatment and is now considered the most relevant outcome for couples undergoing ART rather than number of oocytes retrieved. Ongoing pregnancy represents an appropriate primary end point for efficacy trials in ART as it is a practical surrogate for live birth (Arce et al., 2005). We do not consider the higher number of oocytes retrieved in the rFSH group as an expression of enhanced efficiency of rFSH or a protocol-driven finding. The presence of LH-activity in the HP-hMG preparation results in a more selective follicle recruitment process than an FSH-only gonadotrophin (Smitz et al., 2007). The lower number of oocytes retrieved in the HP-hMG group is caused by the modulating effect of LH-activity. Therefore, HP-hMG is not to be viewed as less efficient; especially as a difference between 10.0 and 11.8 oocytes (average number of oocytes retrieved with HP-hMG and rFSH, respectively) is irrelevant in relation to pregnancy rates (van der Gaast et al., 2006). The proportion of top-quality embryos is a relevant measure, as the number of oocytes retrieved (denominator) is significantly different between the two treatment groups being compared. Thus, the absolute number of top-quality embryos does not provide the full picture of the oocyte/embryo quality and needs to be complemented with the relative frequency. We believe that the aspect of quantity versus quality is very interesting and the findings from MERIT are well in line with the current change of success criterion from obtaining many oocytes to obtaining an adequate cohort of top-quality embryos. We cannot speculate on whether there is a ‘threshold’ number of top-quality embryos generated per cohort but, as described in our paper, the average absolute number of top-quality embryos was 1.1 per patient irrespective of treatment group, and was thus not lower with HP-hMG compared with rFSH. In this context, it is also important to keep in mind that only 50 and 47%, respectively, of the patients treated with HP-hMG and rFSH had top-quality embryos.

Regarding safety end points, there were no differences in MERIT between HP-hMG and rFSH in adverse event profile, moderate/severe OHSS rates, ectopic pregnancies, pregnancy losses or neonatal outcome. Comparative data related to specific evaluations of local tolerability are available from a large (n = 184) randomized controlled trial in anovulatory patients, which showed comparable injection site reactions in terms of redness, pain, itching, swelling and bruising self-assessed 1 h and 24 h after administration of HP-hMG and rFSH preparations (Platteau et al., 2006). Based on this evidence from randomized controlled trials, there is no reason to
speculate on differences in safety issues, including local tolerability.

As one of us is senior author to the cited publication by Popovic-Todorovic (2003), we certainly agree with Dr Trew that an individualized approach to gonadotrophin stimulation dose regimens, based on predictive factors, may increase the proportion of appropriate ovarian responses, decrease the need for dose adjustments, and may be associated with improved outcomes. However, individualizing dose regimens in an efficacy trial comparing two drugs impose a source of post-randomization variability that could violate the strict design of the study.

Speculations on how the outcome of MERIT would have been if the design had been different are of course interesting, but in our times of evidence-based medicine, such hypotheses must be tested and confirmed before they can be given weight in comparison to data obtained from well-conducted randomized controlled trials. As Dr Trew did, we would also like to cite Mark Twain: ‘How empty is theory in the presence of fact!’.

We hope that we have clarified the main concerns raised by Dr Trew. We strongly believe that the exchange of opinions and data, such as this correspondence, contributes to a better understanding of the complexity of clinical trials, and we welcome future discussion of methodological and protocol design aspects.

References


Anders Nyboe Andersen1,4, Paul Devroey2 and Joan-Carles Arce3

1Rigshospitalet, Fertility Clinic, Copenhagen, Denmark, 2Center for Reproductive Medicine of the Vrije Universiteit Brussel, Brussels, Belgium and 3Ferring Pharmaceuticals A/S, Obstetrics and Gynaecology, Clinical Research and Development, Copenhagen, Denmark

4Correspondence address. Rigshospitalet, Fertility Clinic, Blegdamsvej 9, 2100 Copenhagen, Denmark. E-mail: anders.nyboe.andersen@rh.hosp.dk
doi:10.1093/humrep/dem050
Advance Access publication April 5, 2007