Title:

Highly purified HMG versus recombinant FSH for ovarian stimulation in IVF cycles

Authors:

Peter Platteau, Anders Nyboe Andersen, Anne Loft, Johan Smitz, Pascal Danglas, Paul Devroey

Journal:

RBM Online 2008

Article

Highly purified HMG versus recombinant FSH for ovarian stimulation in IVF cycles

Dr Platteau is a Consultant in Reproductive Medicine at the University Hospital, Vrije Universiteit Brussel, Belgium. He graduated from Medical School at the University of Antwerp, Belgium, and is a member of the Royal College of Obstetricians and Gynaecologists, having qualified as a specialist in Obstetrics and Gynaecology at the University of Newcastle-upon-Tyne, United Kingdom. He also holds a Specialty Diploma in Tropical Medicine. His research work has focused in particular on the role of preimplantation genetic diagnosis for aneuploidy screening in infertility treatment. He has published in several medical books and journals and presented at international congresses.

Dr Peter Platteau

Peter Platteau[1,4], Anders Nyboe Andersen[2], Anne Loft[2], Johan Smitz[1], Pascal Danglas[3], Paul Devroey[1]
[1]Centre for Reproductive Medicine of the Academisch Ziekenhuis Vrije Universiteit Brussel, Brussels, Belgium; [2]Rigshospitalet, Fertility Clinic, Copenhagen, Denmark; [3]Ferring Pharmaceuticals, St Prex, Switzerland
[4]Correspondence: Tel.: +32 496 122327; Fax: +32 2 3050357; e-mail: peterplatteau@telenet.be

Abstract

The objective of this study was to compare the live birth rates resulting from ovarian stimulation with highly purified human menopausal gonadotrophin (HP-HMG), which combines FSH and human chorionic gonadotrophin-driven LH activities, or recombinant FSH (rFSH) alone in women undergoing IVF cycles. An integrated analysis was performed of the raw data from two randomized controlled trials that were highly comparable in terms of eligibility criteria and post-randomization treatment regimens with either HP-HMG or rFSH for ovarian stimulation in IVF, following a long down-regulation protocol. All randomized subjects who received at least one dose of gonadotrophin in an IVF cycle (HP-HMG, n = 491; rFSH, n = 495) were included in the analysis. Subjects who underwent intracytoplasmic sperm injection cycles were excluded. The superiority of one gonadotrophin preparation over the other was tested using the likelihood ratio test in a logistic regression analysis. The live birth rate per cycle initiated was 26.5% (130/491) with HP-HMG and 20.8% (103/495) with rFSH (P = 0.041). The odds ratio in favour of HP-HMG was 1.36 (95% confidence interval: 1.01–1.83). Thus, the findings of this integrated analysis demonstrate that ovarian stimulation with HP-HMG, following a long down-regulation protocol, in IVF cycles results in significantly more live births than stimulation with rFSH alone.

Keywords: *highly purified HMG, IVF, LH activity, pregnancy outcome, randomized controlled trial, recombinant FSH*

Introduction

Ovarian stimulation in IVF and intracytoplasmic sperm injection (ICSI) cycles is successfully performed using menotrophins that exhibit both FSH and LH activities as well as recombinant FSH (rFSH) alone, but it remains to be elucidated if there is any difference between the two treatments in terms of ongoing pregnancy rate or live birth rate, which are the appropriate outcome measures for assisted reproductive technology trials (Dickey *et al*., 2004; Min *et al*., 2004; Wennerhom and Bergh, 2004; Arce *et al*., 2005). A meta-analysis of truly randomized controlled studies (van Wely *et al*., 2003) revealed that in women who had down-regulated cycles, using a gonadotrophin-releasing hormone (GnRH) agonist in a long protocol, ovarian stimulation with menotrophins resulted in an insignificant increase in ongoing pregnancy or live birth rate as compared with stimulation with rFSH alone. The authors' conclusion was that data from a larger population are needed to precisely estimate any difference between menotrophins and rFSH. In that investigation (van Wely *et al*., 2003), the studies included in the analysis were conducted using various preparations of menotrophins, which may not be necessarily equivalent in terms of effectiveness because they are different types of exogenous compounds with an LH activity either due to authentic LH or to human chorionic gonadotrophin (HCG) (Wolfenson *et al*., 2005). Moreover, a separate analysis for IVF cycles and ICSI cycles was not performed. The choice of different methods of fertilization (IVF or ICSI cycles) may reflect different reasons for infertility and certainly implies different oocyte/embryo handling. Although there is no clear evidence that the effectiveness of certain menotrophins and rFSH varies depending on the fertilization method, the results of a study by Platteau *et*

al. (2004) indicate that a possible difference in pregnancy rate between a specific menotrophin, the highly purified HMG (HP-HMG) Menopur (Ferring Pharmaceuticals A/S, Copenhagen, Denmark), and rFSH may be observed in IVF cycles but not in ICSI cycles. Menopur differs from other menotrophins because most of its exogenous LH activity derives from HCG content rather than authentic LH (Wolfenson *et al.*, 2005). Thus, this HP-HMG may be superior to rFSH alone for ovarian stimulation in IVF because exposure to its HCG-driven LH activity from the start of ovarian stimulation, following a long GnRH agonist-induced down-regulation protocol, may produce hormonal changes that have beneficial effects on the embryo quality and endometrial receptivity (Smitz *et al.*, 2007).

The objective of the present study was to compare, by retrospective analysis of individual data, the live birth rates resulting from ovarian stimulation in IVF with the HP-HMG Menopur and the rFSH follitropin alpha (Gonal-F, Merck Serono, Geneva, Switzerland) in two large randomized controlled trials that were conducted on similar populations, following a similar protocol for ovarian stimulation, and were evaluated for ongoing pregnancy, a practical surrogate for live birth (Arce *et al.*, 2005), as the primary outcome of a fresh IVF cycle.

Materials and methods

Trials included in the integrated analysis

The two large randomized controlled trials included in this integrated evaluation were the European and Israeli Study Group on highly purified menotrophin versus recombinant follicle-stimulating hormone (EISG) trial (2002) and the menotrophin versus recombinant FSH *in vitro* fertilization trial (MERIT) (Andersen *et al.*, 2006).

The EISG trial was a randomized, open-label, multicentre, multinational study conducted between 1999 and 2000 in 22 fertility clinics in five European countries and in Israel. The trial was designed as a non-inferiority study with respect to the primary outcome measure, ongoing pregnancy rate. Randomization was stratified by planned fertilization procedure (either IVF or ICSI). A total of 781 subjects were randomized in the study and 727 underwent ovarian stimulation: 373 subjects were treated with HP-HMG and 354 were treated with rFSH. Subjects were premenopausal women aged 18–38 years, with regular menstrual cycles and normal uterus, ovaries and adnexa. They had been infertile for at least 1 year before randomization, with the exception of proven bilateral tubal infertility. Serum FSH concentrations were within normal limits, and the body mass index (BMI) was in the range 18–29 kg/m^2. Prior to the study cycle, subjects had undergone at least one menstrual cycle without use of hormonal drugs. Subjects with poor response or severe ovarian hyperstimulation syndrome (OHSS) in a previous cycle were excluded. Subjects with clinically relevant systemic diseases or endocrine abnormalities, contraindications to the use of GnRH or gonadotrophins, smoking habit (more than 10 cigarettes per day) or an abuse of alcohol or drugs were not allowed to participate. Randomization to HP-HMG or rFSH was done before the start of down-regulation. Down-regulation was conducted using the long protocol with a GnRHagonist (triptorelin), which was started in the mid-luteal phase. The starting gonadotrophin dose was 225 IU subcutaneously for the first 5 days. Thereafter, the gonadotrophin dose could be adjusted individually, according to the subject's follicular response. When the established HCG criterion of ≥3 follicles, each with a diameter of ≥16 mm and/or oestradiol concentrations ≥1000 pmol/l per follicle with a diameter of ≥16 mm, was met, a single injection of HCG (5000–10,000 IU either subcutaneously or intramuscularly) was given to induce final follicular maturation. Women returned to the clinic within 32–42 h after HCG administration for oocyte retrieval, and fertilization was performed using standard IVF or ICSI procedures. One to three normally developed embryos were transferred according to the standard practice at the participating clinics. Suitable progestogens were given as luteal support in conformity with the trial centre's standard practice. A serum or urine biochemical pregnancy test was performed 2–3 weeks after embryo transfer (embryo transfer) and, if the test result was positive, subjects were investigated to verify clinical pregnancy (presence of fetal cardiac activity at 4 weeks or later after oocyte retrieval) and ongoing pregnancy (presence of fetal cardiac activity at 10 weeks or later after oocyte retrieval) by transvaginal ultrasound testing. Data on delivery outcome, including neonatal health, were collected during the post-study follow-up period (December 2003 to January 2004).

The MERIT trial was a randomized, assessor-blind, open-label, multicentre, multinational study, conducted in 2004 in 37 fertility clinics in nine European countries and in Israel. The study was designed as a superiority study (convertible to non-inferiority) with respect to the primary outcome measure, ongoing pregnancy rate. Randomization was stratified by age (<35 or 35–37 years). A total of 731 subjects was randomized to treatment with HP-HMG (n = 363) or rFSH (n = 368). Differing from the EISG trial, where 781 subjects were randomized in the study before down-regulation and 727 underwent ovarian stimulation and were treated with either HP-HMG or rFSH, in the MERIT study all randomized subjects underwent ovarian stimulation and were exposed to the gonadotrophins. Subjects were premenopausal women aged 21–37 years, with regular menstrual cycles and normal uterus, ovaries and adnexa. The inclusion and exclusion criteria were similar to those indicated for the EISG trial. Subjects underwent the long protocol with a GnRH agonist (triptorelin) for down-regulation, starting in the mid-luteal phase. Randomization took place on the day of the administration of the first dose of gonadotrophin. The starting gonadotrophin dose was 225 IU subcutaneously for the first 5 days. Thereafter, the gonadotrophin dose could be adjusted individually, according to the subject's follicular response. When the established HCG criterion of ≥3 follicles, each with a diameter of ≥17 mm, was met, a single injection of recombinant HCG (250 µg subcutaneously) was given to induce final follicular maturation. Oocytes were retrieved 36 h (± 2 h) after HCG administration and inseminated using the standard IVF procedures. One or two embryos that fulfilled pre-specified minimum criteria (transferable embryos) were transferred on day 3. Vaginal progesterone was given as luteal support from the day of embryo transfer until confirmation of clinical pregnancy. Subjects were investigated for positive serum βHCG test at 13–15 days after embryo transfer. Clinical pregnancy and ongoing pregnancy were confirmed by transvaginal ultrasound testing at 5–6 weeks (presence of intrauterine gestational sac with fetal heart beat) and at 10–11 weeks (presence of a viable fetus) after embryo transfer respectively. All pregnancies were followed up to delivery. Data on delivery outcome, including neonatal health, were collected as post-study follow-up.

As reported in the original published articles (European and Israeli Study Group on Highly Purified Menotrophin versus Recombinant Follicle-Stimulating Hormone, 2002; Andersen *et al.*, 2006), the protocols of the EISG and MERIT studies were approved by the ethics committees or institutional review boards at the participating centres before patients were enrolled in the trials.

Trial(s) excluded from the integrated analysis

Analysis of the database available at Ferring Pharmaceuticals did not reveal the existence of published or unpublished reports of other completed well-controlled randomized studies comparing similar treatment regimens with the same preparations of HP-HMG and rFSH in a similar long GnRH agonist protocol for ovarian stimulation in IVF, covering the outcome of fresh cycles and having ongoing pregnancy rate as primary outcome measure. In particular, the Dutch First IVF–ICSI Cycle Recombinant FSH versus Menotropin (FIRM) open-label, prospective, randomized study (Hompes *et al.*, 2007), which compared the effectiveness of HP-HMG and rFSH for ovarian stimulation in IVF/ICSI cycles, was not included for the following reasons: (i) the treatment regimen with each gonadotrophin was very different from that adopted in the EISG and MERIT trials because patients received a fixed dose of HP-HMG or rFSH (150 IU/daily); (ii) two different preparations of rFSH were used in the study (either Gonal-F or Puregon); (iii) two different GnRH-agonists were employed for down-regulation in the FIRM trial (either triptorelin or leuprorelin-acetate); (iv) like the EISG trial, this study enrolled women who underwent ovarian stimulation in either IVF or ICSI cycles but, differing from the EISG trial, randomization was not stratified by the fertilization procedure; and (v) in the FIRM trial, most cumulative outcome data were determined during 1 year after embryo transfer in the collecting cycle and live birth rate per started cycle, with the fresh transfer, was not included as an outcome measure.

Consistency of data across trials

Both trials included in this analysis were randomized, controlled multicentre, multinational studies conducted at fertility clinics in Europe and Israel (a total of 59 fertility clinics in 13 countries) and were highly comparable in terms of inclusion/exclusion criteria. The study design and conduct were in line with the treatment practices in the participating countries. Both trials were designed with ongoing pregnancy as the primary end-point. It should be noted that the design of the MERIT trial was based on that of the EISG study and that the possibility of an integrated analysis was mentioned in the protocol for the MERIT study. Subjects in the ICSI stratum of the EISG study were excluded. Both trials were conducted on large populations, and stimulation with HP-HMG or rFSH was performed according to an identical protocol. The protocol for concomitant fertility treatment was similar for all the women enrolled in the two controlled randomized trials, including the long protocol used for down-regulation with the same GnRH agonist, the use of HCG for inducing final follicular maturation and the administration of progesterone as luteal support. However, as mentioned above, randomization to treatment with either HP-HMG or rFSH in the EISG trial was performed before down-regulation. To ensure proper integration of the data with those from the MERIT study, where all randomized subjects were exposed to the gonadotrophins, only data from subjects who received at least one dose of gonadotrophin for IVF in both studies were included in the integrated analysis. Parameters that were only investigated in one of the trials, such as embryo quality, which was assessed only in the MERIT study, were not included in the integrated analysis. Data that were collected differently in the two trials, such as follicular development, were also excluded from the integrated analysis.

Individual data abstraction and integrated analysis

The information source for the retrospective analysis was the database available for the two randomized controlled trials at Ferring Pharmaceuticals, the manufacturer of the HP-HMG preparation used in these trials. Data on individual subjects from each study were integrated for evaluation. Therefore, the integrated analysis was based on the raw data from the trials and was not a meta-analysis. The period of analysis was one IVF cycle and covered the outcome of fresh cycles.

Although the main objective of this analysis was to assess possible differences in live birth rates between treatments, both live birth and ongoing pregnancy were considered as primary end-points because the EISG and the MERIT trials were designed with ongoing pregnancy rate as the primary outcome measure. A live birth cycle was defined as a cycle that resulted in at least one live born neonate, regardless of the number of other neonates and whether they were live born or still born. Ongoing pregnancy was defined either as the presence of fetal cardiac activity at 10 weeks or later after oocyte retrieval (EISG study) or as the presence of one viable fetus 10–11 weeks after embryo transfer upon transvaginal ultrasound testing (MERIT study). Secondary outcome measures were clinical pregnancy rate, implantation rate and ongoing implantation rate. Clinical pregnancy was defined either as the presence of fetal cardiac activity at 4 weeks or later following oocyte retrieval (EISG study) or as the presence of an intrauterine gestational sac with fetal heart beat 5–6 weeks after embryo transfer (MERIT study) upon transvaginal ultrasound testing. Implantation rate was defined as the number of fetuses with heart beat at the gestational age of 6–8 weeks per number of embryos transferred (in percentage). Ongoing implantation rate was defined as the number of viable fetuses or fetuses with heart beat at the gestational age of 12–13 weeks per number of embryos transferred (in percentage). Adverse events were collected from signed informed consent to the end-of-study visit (including ongoing pregnancy visit, if applicable). Adverse events with onset after the first dose of gonadotrophins were considered treatment-emergent adverse events and were included in the analysis. The numbers of moderate and severe OHSS cases occurring during or after gonadotrophin treatment were combined to cover the clinically relevant OHSS cases and to minimize differences between studies. A miscarriage was defined as a pregnancy loss in the first trimester, with a positive pregnancy test but no ongoing pregnancy. Ectopic pregnancy was defined as implantation of an embryo in extrauterine tissue. Delivery outcomes included live birth, still birth and pregnancy loss from ongoing pregnancy onwards. Neonatal health parameters included gender distribution, gestational age and

birth weight, as well as the presence of genetic or congenital abnormalities.

The efficacy and safety populations of this integrated analysis consisted of all subjects who were randomized and received at least one dose of gonadotrophin for ovarian stimulation in IVF cycles (intention-to-treat analysis). Subjects were grouped and analysed according to actual treatment allocation and planned fertilization method, if the assisted reproductive technology method used in the trials was not exclusively IVF. For the primary end-points, the superiority of one gonadotrophin preparation over the other was tested using the likelihood ratio test in a logistic regression analysis. Treatment, study, and age (<35 years and ≥35 years) were included as factors in the model. For each end-point, the difference between treatments was expressed as odds ratio (OR) with 95% confidence interval (CI) (likelihood ratio based) and *P*-value. For the secondary end-points, all dichotomous data were analysed in the same way as the primary end-points. Continuous data were analysed by analysis of variance, with treatment, study, and age (<35 years and ≥35 years) included as factors in the linear model. Subgroup analyses of the primary and secondary outcome variables were conducted in women with embryo transfer and the statistical methods were identical to those applied for the intention-to-treat analyses. Heterogeneity between studies was investigated by a model where interaction between study and treatment was included, in addition to the main effect of study and treatment. There was no statistically significant heterogeneity between the two trials for any of the parameters tested.

Results

Trial population and subject disposition

All subjects in the MERIT study (n = 731) and the subjects in the IVF stratum of the EISG trial (n = 255) were included in the analysis (HP-HMG, n = 491; rFSH, n = 495) (**Table 1**). Age, BMI, duration of infertility and number of previous IVF cycles were comparable between treatment groups (**Table 1**). Tubal and unexplained infertility were the primary reasons for infertility in both treatment groups (**Table 1**).

The percentage of subjects who discontinued prematurely from the studies was comparable in both treatment groups: 12.2% (n = 60/491) in the HP-HMG group and 14.5% (n = 72/495) in the rFSH group. The most frequent reason for discontinuation was non-compliance with the protocol, which was observed in 4.7% of subjects (n = 23/491) on HP-HMG and 6.5% of subjects (n = 32/495) on rFSH. The main deviation was not attending the study visit as planned in the protocol. Among the efficacy-related reasons for discontinuation, excessive response leading to cancellation of the cycle was reported in 0.6% (n = 3/491) of subjects in the HP-HMG group and in 1.6% (n = 8/495) of subjects in the rFSH group, while inadequate response was observed in 2.6% of the subjects in both treatment groups (n = 13/491 for the HP-HMG group and n = 13/495 for the rFSH group).

IVF cycle variables and treatment efficiency

Among the subjects who started an IVF cycle, 95.7% (n = 470/491) in the HP-HMG group and 94.1% (n = 466/495) in the rFSH group had oocyte retrieval following ovarian stimulation. The mean (± SD) number of oocytes retrieved was significantly higher (P < 0.001) in women stimulated with rFSH (12.3 ± 6.7) than in those stimulated with HP-HMG (10.5 ± 5.9). The difference [HP-HMG – rFSH] in mean number of oocytes retrieved was –1.6 (95% confidence interval [CI]:–2.4 to –0.8). The mean (± SD) fertilization rate was 53.6% (± 29.8%) in the HP-HMG group (n = 469) and 54.1% (± 28.8%) in the rFSH group (n = 466). The percentage of subjects with embryo transfer in an IVF cycle was 83.7% (n = 411/491) in women stimulated with HP-HMG and 83.6% (n = 414/495) in women stimulated with rFSH. The mean (± SD) number of embryos transferred in an IVF cycle was 1.8 ± 0.6 in both treatment groups.

The data on treatment efficiency are presented in **Table 2**. Subjects who underwent an IVF cycle were on average

Table 1. Demographics and baseline characteristics.

	HP-HMG (n = *491*)	*rFSH* (n = *495*)
Age (mean years ± SD)	31.0 ± 3.4	31.1 ± 3.5
BMI range (mean kg/m^2 ± SD)	22.7 ± 2.8	22.3 ± 2.7
Duration of infertility (mean years ± SD)	3.9 ± 2.3	4.0 ± 2.4
First treatment cycle [*n* (%) of subjects]	344 (70.1)	327 (66.1)
Primary cause of infertility [*n* (%)]		
Tubal factor	182 (37.1)	173 (34.9)
Unexplained infertility	187 (38.1)	210 (42.4)
Endometriosis	36 (7.3)	37 (7.5)
Mild male factor	72 (14.7)	63 (12.7)
Other	14 (2.8)	12 (2.4)

BMI = body mass index; HP-HMG = highly purified human menopausal gonadotrophin; rFSH = recombinant FSH.

stimulated for 10.7 days (SD = 1.9 days) with HP-HMG and for 10.5 days (SD = 1.9 days) with rFSH. The duration of gonadotrophin treatment in an IVF cycle was 0.2 days shorter in the rFSH group than in the HP-HMG group, which was statistically significant (P = 0.033). The mean total dose of gonadotrophin used in an IVF cycle was 2593.5 IU (SD = 767.8 IU) for HP-HMG and 2476.4 IU (SD = 688.6 IU) for rFSH, and the difference between treatment groups was statistically significant (P = 0.006).

Live birth rate and ongoing pregnancy rate

The live birth rate per cycle initiated was significantly higher (P = 0.041) in the HP-HMG group (26.5%) than in the rFSH group (20.8%) (**Table 3**). The odds of a live birth were 36% higher with HP-HMG than with rFSH (OR: 1.36; 95% CI: 1.01 to 1.83) (**Table 3**). The differences between treatment groups in terms of live birth rate per embryo transfer were also statistically significant and in favour of HP-HMG (P = 0.034; **Table 3**). The superiority of HP-HMG over rFSH for ovarian stimulation in IVF was confirmed by analysis of the ongoing pregnancy rate per cycle initiated and the ongoing pregnancy rate per embryo transfer, which were both significantly higher (P < 0.05) in the HP-HMG group than in the rFSH group (**Table 3**).

Secondary efficacy outcome measures

For the secondary efficacy outcome measures clinical pregnancy rate and implantation rate, the treatment direction was again in favour of HP-HMG (**Table 3**). Both the clinical pregnancy rate per cycle initiated and the clinical pregnancy rate per embryo transfer were significantly higher (P < 0.025) in the HP-HMG group than in the rFSH group. The OR were 1.42 (95% CI: 1.06–1.89) and 1.46 (95% CI: 1.08–1.96) respectively. The implantation rate was 24.2% in the HP-HMG group and 18.8% in the rFSH group (P = 0.030). The odds of implantation were 37% higher with HP-HMG than with rFSH (**Table 3**). For the other secondary efficacy outcome measures ongoing implantation rate, the treatment direction was also in favour of HP-HMG, but the difference between treatment groups did not reach statistical significance (**Table 3**).

Safety evaluation

There were no clinically relevant differences between the HP-HMG and the rFSH groups in terms of treatment-emergent adverse events. Overall, adverse events in an IVF cycle were reported by 52.7% (n = 259/491) of the subjects treated with HP-HMG and by 51.1% (n = 253/495) of the subjects treated with rFSH. It should be noted that the recording of these events covered the period from first gonadotrophin dose to the end-of-study visit, including ongoing pregnancy visit (if applicable), thus exceeding the duration of the actual gonadotrophin treatment. The most commonly reported adverse event in both treatment groups was vaginal haemorrhage (18.1% of subjects treated with HP-HMG, n = 89/491; and 19.0% of subjects treated with rFSH, n = 94/495). In the vast majority of cases, the event consisted of menstrual bleeding after oocyte retrieval in women who did not become pregnant. Besides vaginal haemorrhage, the most frequently reported adverse events were the following: headache (10.2% of subjects in the HP-HMG group, n = 50/491; and 10.1% of subjects in the rFSH group, n = 50/495), OHSS (4.7% of subjects in the HP-HMG group, n = 23/491; and 3.4% of subjects in the rFSH group, n = 17/495), spontaneous abortion (4.5% of subjects in the HP-HMG group, n = 22/491; and 5.3% of subjects in the rFSH group, n = 26/495), pelvic pain (4.3% of women in the HP-HMG group, n = 21/491; and 4.4% of women in the rFSH group, n = 22/495) and nausea (3.5% of subjects in the HP-HMG group, n = 17/491; and 4.0% of subjects in the rFSH group, n = 20/495). There were no statistically significant differences between treatment groups in terms of miscarriages, multiple pregnancies, ectopic pregnancies or moderate/severe OHSS (**Table 4**)

Considering the delivery outcome, the singleton live birth rate was significantly higher (P = 0.047) in women stimulated with HP-HMG (19.3%, n = 95/491) than in women stimulated with rFSH (14.5%, n = 72/495). The OR was 1.41 (95% CI: 1.01–1.98). Delivery of a singleton occurred in 73.1% (n = 95/130) of the live birth cycles with HP-HMG and in 69.9% (n = 72/103) of the live birth cycles with rFSH. Twin deliveries accounted for 26.2% (n = 34/130) and 30.1% (n = 31/103) of the deliveries in the HP-HMG group and in the rFSH group respectively. There was one triplet delivery in the HP-HMG group and none in the rFSH group. The frequency of ongoing pregnancies that did not result in the delivery of a live born child was 4.6% (n = 8/175) in the HP-HMG group and 3.6% (n = 5/140) in the rFSH group. There were two cases of stillbirth in both groups. Late pregnancy loss/abortion occurred for two fetuses in the HP-HMG group and three fetuses in the rFSH group. Four fetuses in the HP-HMG group were removed by selective fetal reduction: two women with three viable fetuses had selective termination of one or two fetuses; a woman with two viable fetuses had selective termination of one fetus with trisomy 21. A fetus with hydrocephalus was removed by selective termination in the rFSH group. One subject in each treatment group was lost to follow-up from ongoing pregnancy to delivery. Both were singleton pregnancies.

There were no appreciable differences between treatment groups with respect to neonatal health among the live born children. The frequency of boys among the live born children was 54.2% (n = 90/166) in the HP-HMG group and 56.0% (n = 75/134) in the rFSH group. The mean gestational age at delivery was 263 days for both treatment groups. The mean gestational age among singletons was 276 days in the HP-HMG group and 274 days in the rFSH group. The mean gestational age among twins was 247 days in the HP-HMG group and 251 days in the rFSH group. The overall incidence of pre-term birth (gestational age below week 37) was 35.5% (n = 59/166) in the HP-HMG group and 35.1% (n = 47/134) in the rFSH group. Birth weight of at least 2500 g was recorded in 68.1% (n = 113/166) and 64.4% (n = 85/132) of the newborns in the HP-HMG and rFSH groups, respectively. The mean birth weight among singletons was 3323 g in the HP-HMG group and 3202 g in the rFSH group. The average birth weight for twins was 2232 g in the HP-HMG group and 2283 g in the rFSH group. Pre-term birth and low/very low birth weight was primarily observed among twins and triplets. The following genetic or congenital abnormalities were observed at birth: one singleton in the HP-HMG group had genetic glucose-6-phosphate dehydrogenase deficiency; one live born child in the same group showed hypertelorism, macroglossia, omphalocoele and an upturned nose while two siblings (triplet delivery) did not

Table 2. Duration of treatment and gonadotrophin consumption in an IVF cycle.

	HP-HMG (n = 491)	*rFSH (n = 495)*	P-*value*
Treatment duration (days)			
Mean	10.7	10.5	0.033
SD	1.9	1.9	
Median	10.0	10.0	
Total dose administered (IU)			
Mean	2593.5	2476.4	0.006
SD	767.8	688.6	
Median	2475.0	2250.0	
Number of dose adjustments [n (% of subjects)]			
0	186 (37.9)	217 (43.8)	
1	236 (48.1)	214 (43.2)	
2	57 (11.6)	50 (10.1)	
3	7 (1.4)	13 (2.6)	
≥4	5 (1.0)	1 (0.2)	

HP-HMG = human menopausal gonadotrophin; rFSH = recombinant FSH.

Table 3. Efficacy outcome measures for IVF cycles using either highly purified human menopausal gonadotrophin (HP-HMG) or recombinant FSH (rFSH) for ovarian stimulation.

Outcome measures	*HP-HMG % (n)*	*rFSH % (n)*	*OR (95% CI)*	P-*value*
Clinical pregnancy rate per cycle initiated	28.9 (142/491)	22.2 (110/495)	1.42 (1.06 to 1.89)	0.017
Clinical pregnancy rate per embryo transfer	34.5 (142/411)	26.6 (110/414)	1.46 (1.08–1.96)	0.013
Implantation rate	24.2 (183/757)	18.8 (144/764)	1.37 (1.03–1.82)	0.030
Ongoing pregnancy rate per cycle initiated	27.3 (134/491)	21.2 (105/495)	1.38 (1.03–1.86)	0.030
Ongoing pregnancy rate per embryo transfer	32.6 (134/411)	25.4 (105/414)	1.42 (1.05–1.92)	0.024
Ongoing implantation rate	22.9 (173/757)	18.2 (139/764)	1.33 (1.00–1.77)	NS
Live birth rate per cycle initiated	26.5 (130/491)	20.8 (103/495)	1.36 (1.01–1.83)	0.041
Live birth rate per embryo transfer	31.6 (130/411)	24.9 (103/414)	1.39 (1.03–1.89)	0.034

CI = confidence interval; embryo transfer = embryo transfer; NS = not statistically significant; OR = odds ratio.

Table 4. Clinically relevant safety parameters for IVF cycles using either highly purified human menopausal gonadotrophin (HP-HMG) or recombinant FSH (rFSH) for ovarian stimulation.

Parameters	*HP-HMG % (n)*	*rFSH % (n)*	*OR (95% CI)*
Miscarriage rate[a]	24.3 (43/177)	31.4 (48/153)	0.73 (0.45–1.20)
Multiple pregnancy rate per ongoing pregnancy	29.1 (39/134)	32.4 (34/105)	0.77 (0.43–1.37)
Ectopic pregnancy rate	0.6 (3/491)	1.0 (5/495)	0.60 (0.12–2.48)
Moderate/severe OHSS	1.8 (9/491)	1.8 (9/495)	0.99 (0.38–2.56)

[a]The miscarriage rate is based on the proportion of women with positive pregnancy test (n = 177 in the HP-HMG group and n = 153 in the rFSH group) who had a pregnancy loss in the first trimester (before ongoing pregnancy). There were no statistically significant differences between the two groups. CI = confidence interval; OHSS = ovarian hyperstimulation syndrome; OR = odds ratio.

show any evidence of congenital abnormalities; a fetus in the rFSH group was diagnosed with acrania and died after birth, while no abnormalities were observed in the remaining live born twin.

Discussion

Live birth is the outcome of interest to couples seeking infertility treatment and it is the most relevant standard of success in assisted reproductive technology (Dickey *et al.*, 2004). However, ongoing pregnancy late in the first trimester represents an appropriate primary outcome for efficacy trials in assisted reproductive technology because it is a practical surrogate for live birth and it is defined within a time frame for which potential co-interventions can be limited and controlled in large multicentre studies (Arce *et al.*, 2005). Therefore, both live birth rate and ongoing pregnancy rate are considered clinically relevant outcome measures for establishing the efficacy of infertility treatment and are mandatory outcome measures for all infertility trials, according to the Food and Drug Administration in the USA.

The present investigation was addressed to compare HP-HMG and rFSH alone for ovarian stimulation in IVF cycles using the clinically relevant efficacy outcome measures mentioned above and focusing in particular on the live birth rate. For this purpose, an integrated, patient-by-patient evaluation of two large randomized controlled trials comparing similar treatment regimens with identical preparations of HP-HMG and rFSH was performed. The most important characteristic of the present investigation, when compared with the previous meta-analysis by van Wely *et al.* (2003), is the inclusion of true randomized controlled trials that were highly comparable in terms of inclusion and exclusion criteria, down-regulation protocol, method used for fertilization, gonadotrophin preparations, and treatment protocol. Moreover, both trials were designed with ongoing pregnancy, the practical surrogate for live birth, as the primary outcome.

The exclusion of data from women who underwent ovarian stimulation in ICSI in the EISG trial (European and Israeli Study Group on highly purified menotrophin versus recombinant follicle-stimulating hormone, 2002) was clinically justified because the clinical decision on the fertilization procedure to be used in the subjects enrolled in that trial was made before initiating the treatment cycle. It should be noted that subjects are generally assigned to a specific fertilization procedure in relation to the reason for infertility and that laboratory manipulation of oocytes differs markedly between the IVF and ICSI procedures. The effectiveness of certain menotrophins and rFSH may vary depending on the fertilization method, because it has been reported that HP-HMG results in higher ongoing pregnancy rate than rFSH alone when used for ovarian stimulation in IVF cycles but not in ICSI cycles (Platteau *et al.*, 2004). From a biological perspective, the interaction between oocytes and cumulus cells after oocyte retrieval occurs only in IVF cycles. While the cumulus cells surround the oocyte for about 1 day after retrieval in IVF cycles, the oocyte is stripped from the cumulus cells immediately after retrieval in ICSI cycles. It has been proposed that the beneficial effect of the exogenous LH activity, associated with the use of menotrophins from the start of ovarian stimulation, may materialize during the first hours of insemination in the IVF procedure and may be related to an enhanced effect of the cumulus cells on the oocytes resulting from increased LH activity in the stimulation period when menotrophins are administered (Platteau *et al.*, 2004). This hypothesis is supported by recent data on the cumulus–oocyte complex gene expression profile (Assou *et al.*, 2006) and by the evidence that the cumulus granulosa cell gene expression profile can be used to predict fertilization and embryo selection in women undergoing ovarian stimulation in IVF cycles (McKenzie *et al.*, 2004). The possibility that subjects undergoing the IVF or the ICSI procedure following ovarian stimulation represent distinct populations, in terms of response to different gonadotrophins, is in keeping with the absence of a statistically significant heterogeneity between the EISG and the MERIT trials for any of the end-points analysed after exclusion of the subjects who were assigned to the ICSI procedure in the EISG study.

This integrated analysis of data from subjects who underwent the IVF procedure following ovarian stimulation with HP-HMG or rFSH demonstrated that ovarian stimulation with HP-HMG resulted in a significantly higher live birth rate than stimulation with rFSH ($P < 0.05$). The odds of having at least one live born neonate in an IVF cycle were 36% higher for women treated with HP-HMG than for women treated with rFSH. Similar results were obtained for the second primary end-point, ongoing pregnancy. The odds of an ongoing pregnancy in an IVF cycle were 38% higher with HP-HMG than with rFSH. The superiority of HP-HMG over rFSH for ovarian stimulation in IVF cycles was confirmed when considering secondary outcome measures that are related to live birth and ongoing pregnancy, such as the clinical pregnancy rate and the implantation rate. In order to translate these results into terms that can be applied directly to individual women in clinical practice, it is possible to use the data presented in **Table 3** for calculation of the number needed to treat (NNT) (Cook and Sackett, 1995; Chatellier *et al.*, 1996; McQuay and Moore, 1997). Despite numerous limitations (Cook and Sackett, 1995; Chatellier *et al.*, 1996; McQuay and Moore, 1997; Smeeth *et al.*, 1999; Cates, 2002), the NNT defines the treatment-specific effect of an intervention that has been found to be superior to a given comparator, in terms of statistical significance, and it is commonly used for making therapeutic decisions about individual patients (McQuay and Moore, 1997). For the main primary end-point of this study, an NNT of 17.5 can be calculated as the inverse of the absolute difference between the live birth rate per cycle initiated in the HP-HMG arm (130/491 = 0.265) and the live birth rate per cycle initiated in the rFSH arm (103/495 = 0.208). This means that for every 18 women who are given HP-HMG for ovarian stimulation in an IVF cycle, one will have at least a single live born neonate who otherwise would not have if she had received rFSH. The clinical relevance of a given NNT largely depends on the trial outcome (McQuay and Moore, 1997). An NNT of 18 is within the limits that are considered clinically relevant in studies where the end-point is a major event comparable to live birth, such as death prevention, and is comparably affected by numerous variables (McQuay and Moore, 1997). In terms of clinical decision making, these findings complement the results of a recent economic analysis that compared the use of HP-HMG and rFSH for ovarian stimulation in IVF, on the basis of data from the MERIT trial, and demonstrated the cost-saving potential of HP-HMG (Wechowski *et al.*, 2007).

Considering the ovarian response to stimulation with the two gonadotrophins in the present study, the mean number of oocytes retrieved in an IVF cycle was significantly higher in the rFSH group than in the HP-HMG group ($P < 0.001$). However, this measure of ovarian response is considered of less relevance than the other outcome measures because the number of oocytes retrieved is not a good predictor for pregnancy (Arce *et al*., 2005). ovarian stimulation with different gonadotrophins may affect the quality of oocytes and/or the endocrine profile in a different way and, through these effects, may indirectly influence the embryo quality and/or the endometrial receptivity (Lisi *et al*., 2005; Andersen *et al*., 2006). As discussed above, the exogenous LH activity associated with the use of menotrophins in ovarian stimulation may have beneficial effects on these variables, which cannot be obtained with FSH-only preparations. These beneficial effects may primarily manifest in IVF cycles where a long GnRH agonist protocol is used for pituitary suppression because the use of the GnRH analogue in the long protocol can deprive the growing follicles of sufficient endogenous LH activity (Fleming *et al*., 1998; Filicori, 1999; van Wely *et al*., 2003; Mochtar *et al*., 2007).

It is important to note that Menopur combines FSH and LH activities in a 1:1 ratio and that most of its exogenous LH activity derives from HCG rather than LH content (Wolfenson *et al*., 2005). A recent study demonstrated a significant positive correlation between serum concentrations of HCG on day 6 after the start of treatment with Menopur and ongoing pregnancy rate in women undergoing IVF after ovarian stimulation with this menotrophin, following a long GnRH-agonist protocol (Smitz *et al*., 2007). In these subjects, serum concentrations of HCG on day 6 after Menopur treatment start was also significantly positively related to the number of top-quality embryos (defined as 4–5 cells on day 2, ≥7 cells on day 3, equally sized blastomeres and ≤20% fragmentation on day 3, and no multinucleation) and the number of top-quality embryos on day 3 after oocyte retrieval. Therefore, exposure to the LH activity provided by HCG from the start of ovarian stimulation, after GnRH agonist-induced down-regulation, may explain the superior effect of Menopur on the pregnancy outcome in IVF cycles as compared with rFSH alone, irrespective of the magnitude of the ovarian response. Indeed, the results of the analysis confirm and extend in an ample population the findings of a previous study comparing Menopur and rFSH in IVF cycles (Platteau *et al*., 2004), which demonstrated higher implantation rate and ongoing pregnancy rate in the HP-HMG group despite the retrieval of fewer oocytes. While the analysis was conducted on data concerning the transfer of 'fresh' embryos, the study mentioned above (Platteau *et al*., 2004) recorded the total number of embryos, including cryopreserved embryos, which were the same in the two treatment groups (Ziebe *et al*., 2007). Therefore, it is unlikely that the increased number of oocytes resulting from stimulation with rFSH leads to more pregnancies after cryopreservation. The results of this integrated analysis at the embryo level showed that the odds of implantation in an IVF cycle was 37% higher in women stimulated with HP-HMG than in women stimulated with rFSH, although the average number of transferred embryos was similar in the two treatment groups. Thus, the findings support the hypothesis that the use of Menopur for ovarian stimulation in IVF may indirectly increase the embryo quality and/or the endometrial receptivity by combining an HCG-driven LH activity with the FSH activity (Andersen *et al*., 2006; Smitz *et al*., 2007; Ziebe *et al*., 2007).

The present results are also in keeping with the data from the FIRM trial (Hompes *et al*., 2007), which compared the effectiveness of fixed doses of HP-HMG and rFSH in women who underwent ovarian stimulation in either IVF or ICSI cycles. Despite the lower number of oocytes retrieved, treatment with HP-HMG resulted in a similar ongoing pregnancy rate per started collecting cycle and in a slightly higher ongoing pregnancy rate per embryo transfer as compared with treatment with rFSH. Although randomization was not stratified by the fertilization procedure, a post-hoc exploratory analysis revealed that the trend towards higher ongoing pregnancy rate per embryo transfer in the HP-HMG group than in the rFSH group was particularly evident in the subpopulation of women who underwent ovarian stimulation in IVF. In the FIRM trial, cumulative outcome data were determined during 1 year after embryo transfer in the collecting cycle and the higher oocyte yield after stimulation with rFSH did not result in more pregnancies even when the results of cryo cycles were included.

The integrated analysis did not reveal any significant difference between HP-HMG and rFSH in terms of treatment-emergent adverse events and discontinuations due to adverse events. Importantly, there was no significant difference between HP-HMG and rFSH in terms of maternal risks of miscarriages, multiple pregnancy, ectopic pregnancy or moderate/severe OHSS, which are considered the most clinically relevant safety end-points in infertility trials (Arce *et al*., 2005). These results are in apparent contrast with the findings of the FIRM study (Hompes *et al*., 2007), which reported an increased incidence of ovarian hyperresponse and/or OHSS in women stimulated with rFSH as compared with women stimulated with HP-HMG, despite the relatively low daily dose of 150 IU used in that trial. However, the lack of an unequivocal definition of OHSS and inclusion of mild cases in the FIRM trial may contribute to explain these inconsistencies. Considering other clinically relevant safety end-points (Arce *et al*., 2005), it is worth mentioning that the integrated database from the two trials included in this investigation contains a description of delivery outcome and neonatal outcome as well, with only 0.8% of all ongoing pregnancies and 0.6% of all fetuses lost to follow-up. The data suggest that pregnancy loss after the first trimester was not affected by the type of gonadotrophin used for ovarian stimulation. The rate of pre-term delivery, low birth weight and very low birth weight was higher in multiple pregnancies than in singleton pregnancies, irrespective of the type of gonadotrophin used for ovarian stimulation. Although the investigation did not include studies designed specifically for single embryo transfer, it is noteworthy that the singleton live birth rate was significantly higher in women stimulated with HP-HMG than in women stimulated with rFSH. The likelihood of having one single live born neonate in an IVF cycle was 1.41 times higher in women who received HP-HMG for ovarian stimulation than in women who received rFSH ($P = 0.047$). The neonatal health of the newborns was related to the multiplicity of pregnancies rather than to the type gonadotrophin used for ovarian stimulation. Overall, there were no major differences between treatment groups in terms of number of genetic abnormalities or malformations observed at the fetal level or after birth.

In conclusion, the present study provides an integrated comparative analysis of the efficacy and safety of HP-HMG and rFSH for ovarian stimulation in IVF cycles, using the most clinically relevant outcomes (live birth and maternal risks). The cumulative evidence indicates that ovarian stimulation with a HP-HMG preparation that provides FSH and HCG-driven LH activity from treatment start results in significantly more live births than treatment with rFSH alone in IVF cycles where a long down-regulation protocol is used ($P < 0.05$). Because the maternal risks and neonatal outcomes are similar in women treated with HP-HMG or with rFSH and in their offspring, respectively, the benefit–risk ratio seems to favour the use of HP-HMG over rFSH for ovarian stimulation in IVF cycles. Although there is at present no evidence that HP-HMG is more effective than rFSH alone for ovarian stimulation in ICSI cycles, HP-HMG may potentially exert beneficial effects in ICSI by enhancing endometrial receptivity. Therefore, further studies on large populations are needed to verify the relative effectiveness of Menopur in different fertilization procedures.

Acknowledgements

The study was supported by Ferring Pharmaceuticals A/S, Copenhagen, Denmark. Pascal Danglas is an employee of Ferring Pharmaceuticals. The other authors have no relevant commercial or financial interests to disclose.

References

Andersen AN, Devroey P, Arce JC 2006 Clinical outcome following stimulation with highly purified hMG or recombinant FSH in patients undergoing IVF: a randomized assessor-blind controlled trial. *Human Reproduction* **21**, 3217–3227.

Arce JC, Nyboe Andersen A, Collins J 2005 Resolving methodological and clinical issues in the design of efficacy trials in assisted reproductive technologies: a mini-review. *Human Reproduction* **20**, 1757–1771.

Assou S, Anahory T, Pantesco V *et al.* 2006 The human cumulus–oocyte complex gene-expression profile. *Human Reproduction* **21**, 1705–1719.

Cates CJ 2002 Simpson's paradox and calculation of number needed to treat from meta-analysis. *BioMed Central Medical Research Methodology* **2**, 1.

Chatellier G, Zapletal E, Lemaitre D *et al.* 1996 The number needed to treat: a clinically useful nomogram in its proper context. *British Medical Journal* **312**, 426–429.

Cook RJ, Sackett DL 1995 The number needed to treat: a clinically useful measure of treatment effect. *British Medical Journal* **310**, 452–454.

Dickey RP, Sartor BM, Pyrzak R 2004 What is the most relevant standard of success in assisted reproduction? No single outcome measure is satisfactory when evaluating success in assisted reproduction: both twin births and singleton births should be counted as successes. *Human Reproduction* **19**, 783–787.

European and Israeli Study Group on Highly Purified Menotrophin versus Recombinant Follicle-Stimulating Hormone 2002 Efficacy and safety of highly purified menotrophin versus recombinant follicle-stimulating hormone in in vitro fertilization/intracytoplasmic sperm injection cycles: a randomized, comparative trial. *Fertility and Sterility* **78**, 520–528.

Filicori M 1999 The role of luteinizing hormone in folliculogenesis and ovulation induction. *Fertility and Sterility* **71**, 405–414.

Fleming R, Lloyd E, Herberr M *et al.* 1998 Effects of profound suppression of luteinizing hormone induces different hormone profiles compared with menotrophins, dependent upon the route of administration and endogenous luteinizing hormone activity. *Human Reproduction* **13**, 1788–1792.

Hompes PGA, Broekmans FJ, Hoozemans DA, Schats R 2007 Effectiveness of highly purified human menopausalgonadotropin vs. recombinant follicle-stimulating hormone in first-cycle in vitro fertilization-intracytoplasmic sperm injection patients. *Fertility and Sterility* [published online ahead of print] Doi: 10.1016/j.fertnstert.2007.05.039.

Lisi F, Rinaldi L, Fishel S *et al.* 2005 Evaluation of two doses of recombinant luteinizing hormone supplementation in an unselected group of women undergoing follicular stimulation for in vitro fertilization. *Fertility and Sterility* **83**, 309–315.

McKenzie LJ, Pangas SA, Carson SA, *et al.* 2004 Human cumulus granulosa cell gene expression: a predictor of fertilization and embryo selection in women undergoing IVF. *Human Reproduction* **19**, 2869–2874.

McQuay HJ, Moore RA 1997 Using numerical results from systematic reviews in clinical practice. *Annals of Internal Medicine* **126**, 712–720.

Min JK, Breheny SA, MacLachlan V, Healy DL 2004 What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. *Human Reproduction* **19**, 3–7.

Mochtar MH, Van der Veen, Ziech M, van Wely M 2007 Recombinant luteinizing hormone (rLH) for controlled ovarian hyperstimulation in assisted reproductive cycles. *Cochrane Database of Systematic Reviews* **2**, CD005070.

Platteau P, Smitz J, Albano C *et al.* 2004 Exogenous luteinizing hormone activity may influence the treatment outcome in in vitro fertilization but not in intracytoplasmic sperm injection cycles. *Fertility and Sterility* **81**, 1401–1404.

Smeeth L, Haines A, Ebrahim S 1999 Numbers needed to treat derived from meta-analysis—sometimes informative, usually misleading. *British Medical Journal* **318**, 1548–1551.

Smitz J, Andersen AN, Devroey P, Arce JC 2007 Endocrine profile in serum and follicular fluid differs after ovarian stimulation with HP-hMG or recombinant FSH in IVF patients. *Human Reproduction* **22**, 676–687.

van Wely M, Westergaard LG, Bossuyt PMM, Van der Veen F 2003 Human menopausal gonadotrophin versus recombinant follicle stimulation hormone for ovarian stimulation in assisted reproductive cycles. *Cochrane Database of Systematic Reviews* 1: CD003973.

Wechowski J, Connolly M, McEwan P, Kennedy R 2007 An economic evaluation of highly purified HMG and recombinant FSH based on a large randomized trial. *Reproductive BioMedicine Online* **15**, 500–506.

Wennerholm UB, Bergh C 2004 What is the most relevant standard of success in assisted reproduction? Singleton live births should also include preterm births. *Human Reproduction* **19**, 1943–1945.

Wolfenson C, Groisman J, Couto AS *et al.* 2005 Batch-to-batch consistency of human-derived gonadotrophin preparations compared with recombinant preparations. *Reproductive BioMedicine Online* **10**, 442–454.

Ziebe S, Lundin K, Janssens R *et al.* 2007 Influence of ovarian stimulation with HP-hMG or recombinant FSH on embryo quality parameters in patients undergoing IVF. *Human Reproduction* **22**, 2404–2413.

The results of this study were presented in abstract form at the 22nd Annual Meeting of the European Society for Human Reproduction and Embryology, Prague, 18–21 June 2006.

Received 5 December 2007; refereed 19 December 2007; accepted 19 March 2008.