Low-dose human menopausal gonadotrophin versus clomiphene citrate in subfertile couples treated with intrauterine insemination: a randomized controlled trial

Karen Peeraer¹, Sophie Debrock¹, Peter De Loecker¹, C. Tomassetti¹, A. Laenen², M. Welkenhuysen¹, L. Meeuwis¹, S. Pelckmans¹, B.W. Mol³, C. Spiessens¹, D. De Neubourg¹, and T.M. D’Hooghe¹,*

¹Leuven University Fertility Center, UZ Leuven Campus Gasthuisberg, 3000 Leuven, Belgium ²Leuven Biostatistics and Statistical Bioinformatics Centre, Leuven, Belgium ³The Robinson Institute, School of Paediatrics and Reproductive Health University of Adelaide, 5000 Adelaide SA, Australia

*Correspondence address. Tel: +32-16340819; E-mail: karen.peeraer@uzleuven.be

Submitted on November 6, 2014; resubmitted on February 17, 2015; accepted on February 20, 2015

STUDY QUESTION: Can controlled ovarian stimulation with low-dose human menopausal gonadotrophin (hMG) improve the clinical pregnancy rate when compared with ovarian stimulation with clomiphene citrate (CC) in an intrauterine insemination (IUI) programme for subfertile couples?

SUMMARY ANSWER: Ovarian stimulation with low-dose hMG is superior to CC in IUI cycles with respect to clinical pregnancy rate.

WHAT IS KNOWN ALREADY: IUI after ovarian stimulation is an effective treatment for mild male subfertility, unexplained subfertility and minimal-mild endometriosis, but it is unclear which medication for ovarian stimulation is more effective.

STUDY DESIGN, SIZE, DURATION: A total of 330 women scheduled for IUI during 657 cycles (September 2004–December 2011) were enrolled in an open-label randomized clinical trial to ovarian stimulation with low-dose hMG subcutaneous (n = 334, 37.5–75 IU per day) or CC per oral (n = 323, 50 mg/day from Day 3–7). Assuming a difference of 10% in ‘clinical pregnancy with positive fetal heart beat’, we needed 219 cycles per group (alpha-error 0.05, power 0.80).

PARTICIPANTS/MATERIALS, SETTING, METHODS: We studied subfertile couples with mild male subfertility, unexplained subfertility and minimal-mild endometriosis. Further inclusion criteria were failure to conceive for ≥12 months, female age ≤42 years, at least one patent Fallopian tube and a total motility count (TMC) >5.0 million spermatozoa after capacitation. The primary end-point was clinical pregnancy. Analysis was by intention to treat and controlled for the presence of multiple measures, as one couple could have more randomizations in multiple cycles. Linear mixed models were used for continuous measures. For binary outcomes we estimated the relative risk using a Poisson model with log link and using generalized estimating equations.

MAIN RESULTS AND THE ROLE OF CHANCE: When compared with ovarian stimulation with CC, hMG stimulation was characterized by a higher clinical pregnancy rate (hMG 48/334 (14.4%) versus CC 29/323 (9.0%), relative risk (RR) 1.6 (95% confidence interval (CI) 1.1 – 2.4)), higher live birth rate (hMG 46/334 (13.8%) versus CC 28/323 (8.7%), RR 1.6 (95% CI 1.0 –2.4)), low and comparable multiple live birth rate (hMG 3/46 (6.5%) versus CC 1/28 (3.6%), P > 0.99), lower number of preovulatory follicles (hMG 1.2 versus CC 1.5, P < 0.001), increased endometrial thickness (hMG 8.5 mm versus CC 7.5 mm, P < 0.001), and a lower cancellation rate per started cycle (hMG 15/322 (4.7%) versus CC 46/298 (15.4%), P < 0.001).

LIMITATIONS, REASONS FOR CAUTION: We randomized patients at a cycle level, and not at a strategy over multiple cycles.
Introduction

Subfertility, defined as the failure to conceive after 12 months of regular intercourse, affects ~1 in 10 couples wishing to have a child. Intrauterine insemination (IUI) indicated for unexplained subfertility, minimal-mild endometriosis, male subfertility and physical disability or psychosexual problems, is widely used as first line treatment for subfertile couples (Van Weert et al., 2004; Kennedy et al., 2005; Practice Committee of the American Society for Reproductive Medicine 2006, 2012a,b; Verhulst et al., 2006; Chamley and Clarke, 2007; Steures et al., 2007; Francavilla et al., 2009; Collège National des Gynécologues et Obstétriciens Français (CNGOF), 2010; NICE guidelines, 2013). In Europe, close to 200,000 IUI cycles were registered in 2009 (Ferraretti et al., 2013). Unfortunately no data exist for the USA since public reporting of IUI cycles is not registered in the USA. For unexplained subfertility and endometriosis there is sufficient evidence that IUI after controlled ovarian stimulation (COS) significantly improves the pregnancy rate over IUI in natural cycle and over expected management (Cohen, 2005; Kennedy et al., 2005; Practice Committee of the American Society for Reproductive Medicine 2006, 2012a,b; Verhulst et al., 2006; Crosignani, 2009). Although there is still insufficient evidence to conclude that IUI is effective for male factor subfertility (Bensdorp et al., 2007), it is considered as an acceptable first line treatment when at least 1 million motile spermatozoa are available for IUI (Van Weert et al., 2004; Merviel et al., 2010; Ombelet and Cohen, 2013).

Despite its widespread use, the role and type of COS combined with IUI is controversial. Clomiphene citrate, an anti-estrogen, is mostly used as first choice for COS in the context of IUI since clomiphene can be administered orally and is cheaper than gonadotrophin injections (Ecochard et al., 2000; Dankert et al., 2007; Berker et al., 2011). Clomiphene is a selective estrogen receptor modulator (SERM), a non-steroidal estrogen that binds to the estrogen receptors at multiple sites throughout the reproductive tract and can act as an estrogen agonist or as an antagonist. Clomiphene binds to estrogen receptors in the hypothalamus, inhibiting negative feedback of estrogen on gonadotrophin release. Subsequent up-regulation of the hypothalamic–pituitary–gonadal axis leads to growth of the ovarian follicle(s). Gonadotrophins are glycoprotein hormones that can be extracted from urine of menopausal women or can be manufactured in recombinant variants. They stimulate follicular growth by acting directly on ovarian FSH receptors, and have no anti-estrogenic effect on cervical mucus or endometrium like clomiphene.

Contradictory results have been reported in multiple randomized controlled trials (RCTs) comparing reproductive outcome after IUI and COS with clomiphene and gonadotrophins for various clinical indications or contexts (unexplained infertility, endometriosis, mild male factor, donor insemination) (Table I). The latest systematic review (Cantineau et al., 2007) included seven RCTs comparing COS with clomiphene or gonadotrophins in women with various IUI treatment indications (n = 556) and showed a significantly higher pregnancy rate after COS with gonadotrophins than after COS with clomiphene (odds ratio (OR) 1.8, 95% CI 1.2 to 2.7). However, since the largest RCT showed no effect, more multi-centre RCTs are needed to confirm this observation and to reassure that better effectiveness does not increase multiple pregnancy rates (Karlstrom et al., 1993; Karande et al., 1995; Matorras et al., 2002; Berker et al., 2011). Therefore, we performed a multi-centre RCT to test the hypothesis that the clinical pregnancy rate after IUI is higher after COS with low-dose human menopausal gonadotrophin (hMG) (20%) when compared with COS with clomiphene citrate (CC) (10%).

Materials and Methods

Ethical approval

The study was approved by the Institutional Review Board (ML2438) of Leuven University Hospitals. Since trial registration was not mandatory at the start of the study in 2003–2004, registration was done at Clinical-Trials.gov (NCT01569945) after the last patient was enrolled.

Study population

Between September 2004 and December 2011 we performed an open-label randomized clinical trial in the Divisions of Reproductive Medicine, Departments of Obstetrics and Gynecology at University Hospital Gasthuisberg (Leuven, Belgium) and at St. Augustinus Hospital (Antwerp, Belgium).

Subfertile couples underwent a complete infertility evaluation including a medical history, physical examination, serum hormone assays between Day 2 and 5 of the menstrual cycle, mid luteal serum progesterone in women with regular menstrual cycles, pelvic ultrasound, assessment of tubal patency either by hysterosalpingography or laparoscopy and semen analysis. If an implantation abnormality was seen on ultrasound, a hysteroscopy and/or laparoscopy was also performed. Semen samples with a concentration of >20 million/ml, type A (fast progressive) + type B (medium progressive) motility >50% and normal morphology >15% were considered normal (World Health Organization, 1999). Couples were eligible when they had tried to conceive for at least 12 months, when female age was ≤42 years, with at least one patent Fallopian tube and with a total motility count (TMC) ≥5.0 million spermatozoa after capacitation and a written informed consent.
<table>
<thead>
<tr>
<th>First Author</th>
<th>Type of study</th>
<th>Sperm origin</th>
<th>No. of patients</th>
<th>No. of cycles</th>
<th>Indications</th>
<th>Type of Gns</th>
<th>Ovulation triggering with hCG</th>
<th>Starting dose</th>
<th>Mean no. of follicles</th>
<th>Mean age (years)</th>
<th>Mean duration of subfertility (months)</th>
<th>Pregnancy rate/cycle (%)</th>
<th>Multiple pregnancies (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karlstrom, 1993</td>
<td>RCT</td>
<td>Partner</td>
<td>32</td>
<td>32</td>
<td>Unexplained, endometriosis</td>
<td>hMG</td>
<td>10,000 IU&lt;sup&gt;a&lt;/sup&gt;</td>
<td>150 IU/d from CD 2–3</td>
<td>32</td>
<td>59</td>
<td>3/15 (20%)</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Balasch, 1994</td>
<td>RCT</td>
<td>Partner</td>
<td>100</td>
<td>192</td>
<td>Male factor, unexplained</td>
<td>uFSH</td>
<td>10,000 IU</td>
<td>75 IU/d IM from CD 7</td>
<td>78</td>
<td>73</td>
<td>12/94 (13%)</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Karande, 1995</td>
<td>RCT</td>
<td>Partner</td>
<td>NS</td>
<td>120</td>
<td>Unexplained, female factor, male factor, endometriosis</td>
<td>hMG</td>
<td>yes, dose NS</td>
<td>150 mg/d</td>
<td>10/76 (13%)</td>
<td>NS</td>
<td>3/10 (30%)</td>
<td>0/3 (0%)</td>
<td></td>
</tr>
<tr>
<td>Kamel, 1995</td>
<td>RCT</td>
<td>Partner</td>
<td>54</td>
<td>54</td>
<td>Unexplained, male factor</td>
<td>hMG</td>
<td>10,000 IU</td>
<td>75 IU/d IM from CD 3</td>
<td>12/28 (14%)</td>
<td>2/26 (7%)</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Manganiello, 1997</td>
<td>Observational</td>
<td>Partner</td>
<td>83</td>
<td>204</td>
<td>Female factor, male factor, unexplained, mixed, combined</td>
<td>hMG</td>
<td>5000 IU</td>
<td>150 IU/d from CD 3</td>
<td>3/44 (7%)</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hannoun, 1998</td>
<td>Observational</td>
<td>Partner</td>
<td>147</td>
<td>544</td>
<td>Unexplained, male factor</td>
<td>hMG</td>
<td>5000 IU&lt;sup&gt;a&lt;/sup&gt;</td>
<td>150 IU/d from CD 4, 6, 8 and 9</td>
<td>30</td>
<td>43</td>
<td>6/84 (7%)</td>
<td>13/90 (14%)</td>
<td>NS</td>
</tr>
<tr>
<td>Karlstrom, 1998</td>
<td>RCT</td>
<td>Partner</td>
<td>74</td>
<td>74</td>
<td>Unexplained, endometriosis, male factor, cervical factor</td>
<td>hMG</td>
<td>10,000 IU&lt;sup&gt;a&lt;/sup&gt;</td>
<td>150 IU/d from CD 2–3</td>
<td>32</td>
<td>48</td>
<td>15/278 (5%)</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Ecochard, 2000</td>
<td>RCT, crossover</td>
<td>Partner</td>
<td>58</td>
<td>174</td>
<td>Female factor, male factor, unexplained</td>
<td>hMG</td>
<td>5000 IU</td>
<td>150 IU/d from CD 4, 6, 8 and 9</td>
<td>30</td>
<td>48</td>
<td>16/264 (6%)</td>
<td>6/30 (20%)</td>
<td>15/238 (13%)</td>
</tr>
<tr>
<td>Matorras, 2002</td>
<td>RCT</td>
<td>Donor</td>
<td>100</td>
<td>502</td>
<td>Male factor</td>
<td>uFSH</td>
<td>5000 IU</td>
<td>150 IU/d from CD 2–3</td>
<td>32</td>
<td>64</td>
<td>16/264 (6%)</td>
<td>6/30 (20%)</td>
<td>15/238 (13%)</td>
</tr>
<tr>
<td>Dankart, 2007</td>
<td>RCT</td>
<td>Partner</td>
<td>138</td>
<td>406</td>
<td>Unexplained, male factor</td>
<td>rFSH</td>
<td>5000 IU</td>
<td>75 IU/d from CD 3</td>
<td>32</td>
<td>32</td>
<td>23/207 (11%)</td>
<td>27/199 (14%)</td>
<td>12/24 (4%)</td>
</tr>
<tr>
<td>Berker, 2011</td>
<td>RCT</td>
<td>Partner</td>
<td>189</td>
<td>189</td>
<td>Unexplained, male factor</td>
<td>rFSH</td>
<td>10,000 IU</td>
<td>75 IU/d from CD 2–4 depending on BMI</td>
<td>44</td>
<td>44</td>
<td>9/93 (10%)</td>
<td>2/15 (13%)</td>
<td>0/9 (0%)</td>
</tr>
<tr>
<td>Our study, 2014</td>
<td>RCT</td>
<td>Partner, donor</td>
<td>306</td>
<td>620</td>
<td>Female factor, male factor, unexplained, mixed</td>
<td>hMG</td>
<td>5000 IU</td>
<td>75–150 IU/d from CD 2–3</td>
<td>30</td>
<td>30</td>
<td>40/322 (12%)</td>
<td>20/298 (7%)</td>
<td>1/20 (5%)</td>
</tr>
</tbody>
</table>

NS, not stated; Gn, gonadotrophin; CD, cycle day.
<sup>a</sup>Ovulation triggering only in gonadotrophin group.
<sup>b</sup>Ongoing pregnancy rate: pregnancy proceeding beyond 20 weeks of gestation.
Study design
We randomized eligible couples at the level of the treatment cycle. The participants were randomized for COS with either hMG (Menopur®, Feringa, Aalst, Belgium) (hMG COS group) or clomiphene citrate (Clomid®, Sanofi, Diegem, Belgium; Pergotim®, Merck, Overijse, Belgium) (CC COS group) by an independent investigator. Treatment allocation for each participant was performed by opening opaque sealed envelopes only after written informed consent of the female partner was obtained. We used blocked randomization per 10 envelopes for each recruiting centre, containing 5 in the CC group and 5 in the hMG group. Division of a batch of 30 (3 × 10) envelopes was as follows: 20 envelopes in the University Hospital Gasthuisberg (10 in control of an independent investigator and 10 envelopes in control of an independent trial midwife) and 10 envelopes under responsibility of an independent investigator in the St Augustinus Hospital. A participant was allowed to participate more than once in either one or both of the hMG COS and CC COS groups, each time after new randomization.

Patients allocated to receive hMG were stimulated from Day 2 of the menstrual cycle with a starting dose of 37 or 75 IU in a low-dose step-up protocol as recommended (Cohlen 2005; Cantineau et al., 2007) and based on patient’s age, BMI, basal serum FSH levels and previous medical history. Patients allocated to CC group were treated with a starting dose of 50 mg/day from Day 3 until Day 7 of the menstrual cycle. They also received oral ethinyl estradiol (EE2) 50 µg from Day 8 until Day 12 to support endometrial growth, in order to prevent the anti-estrogenic effect of CC reducing endometrial thickness (Martinez et al., 1990; Yagel et al., 1992; Dickey et al., 1993; Gelety and Buyalos, 1993; Nakamura et al., 1997; Gerli et al., 2000; Unfer et al., 2001; Dehbashi et al., 2003; Haritha and Rajagopalan, 2003; Satriapod et al., 2014). In both groups, ovarian follicular growth was monitored by ultrasound and serum hormonal analysis (17 beta-estradiol, progesterone, luteinizing hormone, follicle-stimulating hormone) starting on Day 6 or Day 7 of the menstrual cycle in the hMG COS group and on Day 10–12 of the cycle in the CC COS group. The follicular diameter was measured by ultrasound and calculated as the average of 2 perpendicular dimensions between the inner walls of each follicle. In all cycles, hCG 5000 IU was administered when the leading follicle had reached an average diameter of at least 17–18 mm, according to the protocol for ovulation triggering in our centre for all cycles stimulated for ovulation induction or for IUI. A single IUI was performed later depending on morning serum LH levels. When a serum LH peak was detected on the day of hCG administration, the IUI was performed 1 day (27–30 hours) later. When serum LH was within normal limits on the day of hCG administration, the IUI was performed 2 days (51–54 h) later. When three or more follicles of 14 mm or larger were present at the time of hCG administration, selective ultrasound-guided follicular aspiration was offered before IUI or the cycle was cancelled, as described previously (Spiessens et al., 2003; Ghersquere et al., 2007). A Frydman Classic Catheter (Laboratoire CCD, Paris, France) equipped with a sterile tuberculin syringe was used for each intrauterine insemination. Semen samples used for insemination were processed within 1 h after ejaculation. The sperm was washed in a 3-layer discontinuous gradient centrifugation by using iSolate® (Irvine Scientific, Santa Ana, CA, USA), and semen analysis was performed according to the World Health Organization criteria (World Health Organization, 1999). After IUI, women had bed rest for 15 min (Custers et al., 2009). Pregnancy was determined as positive serum hCG levels (≥ 25 IU/l) 2 weeks after IUI. Clinical pregnancy was defined as the presence of a fetus intrauterine or extraterine with a positive heartbeat, on ultrasound at 6–8 weeks amenorrhoea (Zegers-Hochschild et al., 2009). Live birth (LB) was defined as the live birth of a child beyond 24 weeks of gestation. Multiple live birth (MLB) was defined as the birth of two or more infants. Follow-up of pregnancies and deliveries was performed in our hospital and in other hospitals. There was no specific protocol for follow-up of pregnancies in our study design. All obstetrical data were reported according to the compulsory registration of IVF cycles to the Belgian Register for Assisted Procreation (BELRAP) and therefore available in our databank (http://www.belrap.be; De Neubourg et al., 2013).

Outcome
The primary outcome was clinical pregnancy. Secondary outcomes included live birth, multiple live birth, endometrial thickness and number of preovulatory follicles on the day of hCG administration and discontinuation rate during treatment.

Statistical considerations
Both groups were compared with respect to baseline characteristics (Table II), as well as endometrial thickness, number of dominant ovarian follicles on the day of hCG administration and discontinuation rate during treatment.

Sample size calculation
The aim of this study was to test the hypothesis that hMG stimulation would increase the clinical pregnancy rate when compared with CC stimulation. We assumed the clinical pregnancy rate after COS with hMG to be 20% as based on the average pregnancy rate in our centre at the time of initiation of this study (Spiessens et al., 2003). The pregnancy rate per IUI cycle after COS with CC was estimated to be 10% based on literature data (7% pregnancy rate per cycle for IUI with CC without EE2 supplementation with an estimated increase to 10% due to a potentially added value of EE2 co-treatment during ovarian stimulation with CC) (Gerli et al., 2000). Accepting an α error of 0.05 and 80% power, we needed 219 cycles in each group to demonstrate a significant difference in the primary outcome which is clinical pregnancy.

Table II Baseline clinical characteristics at cycle level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CC (N = 323)</th>
<th>hMG (N = 334)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean female age (± SD)</td>
<td>31.6 ± 3.7</td>
<td>31.9 ± 4.1</td>
</tr>
<tr>
<td>Mean BMI (± SD)</td>
<td>22.6 ± 3.5</td>
<td>23.1 ± 4.1</td>
</tr>
<tr>
<td>Mean FSH Basal (± SD)</td>
<td>6.8 ± 2.5</td>
<td>6.8 ± 2.4</td>
</tr>
<tr>
<td>Mean cycle duration (± SD)</td>
<td>29.6 ± 4.1</td>
<td>30.2 ± 5.6</td>
</tr>
<tr>
<td>Type of infertility, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>223 (69%)</td>
<td>249 (75%)</td>
</tr>
<tr>
<td>Secondary</td>
<td>100 (31%)</td>
<td>85 (25%)</td>
</tr>
<tr>
<td>Duration of infertility in months (± SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>25.3 ± 19.0</td>
<td>28.0 ± 18.5</td>
</tr>
<tr>
<td>Secondary</td>
<td>22.5 ± 16.0</td>
<td>21.1 ± 13.0</td>
</tr>
<tr>
<td>Tubal patency, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unilateral</td>
<td>13 (4%)</td>
<td>9 (3%)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>310 (96%)</td>
<td>325 (97%)</td>
</tr>
<tr>
<td>Indication for treatment, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anovulation</td>
<td>25 (8%)</td>
<td>26 (8%)</td>
</tr>
<tr>
<td>Endometriosis</td>
<td>58 (18%)</td>
<td>73 (22%)</td>
</tr>
<tr>
<td>Implantation</td>
<td>10 (3%)</td>
<td>21 (6%)</td>
</tr>
<tr>
<td>Male factor</td>
<td>75 (23%)</td>
<td>50 (15%)</td>
</tr>
<tr>
<td>Mixed</td>
<td>56 (17%)</td>
<td>58 (17%)</td>
</tr>
<tr>
<td>Tubal factor</td>
<td>13 (4%)</td>
<td>9 (3%)</td>
</tr>
<tr>
<td>Unexplained</td>
<td>86 (27%)</td>
<td>97 (29%)</td>
</tr>
</tbody>
</table>
follicles, incidence of selective follicular aspiration before IUI and sperm quality data at the time of IUI (Table III). Following randomization, an intention-to-treat analysis was done on the cycle level for all cycles from all participating couples. Since couples could participate more than once, and the data show clustering of cycles within couples, we applied statistical methods that took into account this clustering. Linear mixed models were used for the analysis of continuous measures, where a random intercept accounts for clustering (Verbeke and Molenberghs, 2000). For binary variables we used logistic regression analysis and the relative risk was estimated using a Poisson model with log link. Both models used generalized estimating equations (GEE) to account for clustering (Liang and Zeger, 1986). In secondary analyses, we analysed data per started COS cycle (all cycles where COS was started, with or without IUI subsequently performed) and per inseminated cycle. All analyses were performed using SAS (SAS software, version 9.2 of the SAS System for Windows).

### Results

Between September 2004 and December 2011, 330 couples were randomized who started 657 cycles; 334 cycles were allocated to hMG COS and 323 cycles were allocated to CC COS (Fig. 1). Only 13 treatment cycles were randomized in 2004–2006 due to logistic problems; the majority of patients were recruited during the latter 5 years of the study (2007–2011). Both groups were comparable with respect to baseline clinical variables (Table II). One hundred and twenty-seven couples participated in one cycle, 99 couples in two cycles, 86 couples in three cycles, 16 couples in four cycles and 2 couples in five cycles. The mean interval between two consecutive IUI cycles was about 2 months (62.1 ± 77.5 days). A total of 18 women (18/657 or 3% of all randomized cycles) conceived spontaneously before therapy was started (Fig. 1). There were 252 inseminations in the CC group and 307 in the hMG group. When compared with CC cycles, hMG cycles were marked by a lower number of dominant follicles (diameter of 14 mm or more) at the time of hCG administration (hMG 1.2 versus CC 1.5, P < 0.001), and were marked by increased endometrial thickness (hMG 8.5 mm versus CC 7.5 mm, P < 0.001) (Table III).

The primary outcome clinical pregnancy occurred significantly more frequently in the hMG group (48/334; 14.4%) than in the CC group (29/323; 9.0%) (RR 1.6 (95% CI 1.0–2.4)). Live birth rate (LBR) was also significantly higher in the hMG group (46/334; 13.8%) than in the CC group (28/323; 8.7%) (RR 1.6 (95% CI 1.0–2.4)) (Table IV). The multiple live birth rate (MLBR) was low and did not differ between both groups (hMG 3/46 (6.5%) versus CC (1/28, (3.6%), P > 0.99 Fisher). IUI was more often cancelled in the CC group (71/323; 22.0% (95% CI 17.6–27.1%)) when compared with the hMG group (27/334; 8.1% (95% CI 5.6–11.5%)). Reasons for cancellation were comparable between groups (Fig. 1) (Supplementary Table SI).

Overall, only one major adverse event (general allergic reaction) was observed in a patient allocated to the CC group, whose treatment was cancelled. Hospitalizations were not reported in any participant during the trial. Perinatal mortality was not observed among babies delivered by the participants of this trial. In newborns, congenital malformations were absent, except for one case of unique left kidney with double excretory system diagnosed in one of the babies of a twin pregnancy in the CC group. Admission to the neonatal unit was not needed for any of the 29 babies born in the CC group, but required for 2 out of 49 babies born in the hMG group. One baby was admitted during 24 days (preterm birth) and the other one was admitted for 1 day only (difficult neonatal adaptation).

When we limited our analysis to the cycles in which COS was started, irrespective of subsequent IUI, the clinical pregnancy rate was also significantly higher in the hMG group (40/322; 12.4%) than in the CC group (20/298; 6.7%) (RR 1.9 (95% CI 1.1–3.1)) (Table IV). Selective ultrasound-guided follicular aspiration before IUI was significantly lower in the hMG group (9/322; 2.8% (95% CI 1.5–5.3)) than in the CC group (28/298; 9.4% (95% CI 6.5–13.4)) (P = 0.001) (Table III). When we only analysed cycles in which IUI was performed, the clinical pregnancy rate was also significantly higher in the hMG group (40/307; 13.0%) than in the CC group (18/252; 7.1%) (RR 1.8 (95% CI 1.1–3.1)).

### Discussion

In this randomized controlled trial, ovarian stimulation with low-dose gonadotrophins was superior to clomiphene with respect to clinical pregnancy and live birth rates, without increased incidence of multiple live birth rate.

Our study is marked by several strengths. Its results are applicable in daily clinical practice since patients with diverse causes of subfertility were eligible, and because recruitment at cycle level reflects real life clinical practice where often the type of controlled ovarian stimulation used in IUI cycles is selected individually on the cycle level, not on the patient level and is based on dynamic and shared decision making between doctors and patients. Another significant strength of the trial is the low
starting doses of gonadotrophin and the rigid cancellation requirements, including options to allow aspiration of extra follicles enabling continue participation in a cycle. Our data suggest that this practice is useful and could also be considered in other countries. A potential bias is caused by the fact that each patient could participate with more than one cycle in the study in both treatment groups. Ideally, one would have compared two strategies involving multiple cycles. However, recruitment for such studies is difficult, as couples are asked to stick to a particular

Table IV  Reproductive outcome per randomized cycle (intention to treat analysis), per started COS cycle and per IUI cycle.

<table>
<thead>
<tr>
<th></th>
<th>CC</th>
<th>hMG</th>
<th>Relative risk</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per randomized cycle</td>
<td>n = 323 cycles</td>
<td>n = 334 cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHB + pregnancy rate: percentage (95% CI)</td>
<td>9.0 (6.5;13.1)</td>
<td>14.4 (11.5;19.1)</td>
<td>1.6 (95% CI 1.0–2.4)</td>
<td>0.03</td>
</tr>
<tr>
<td>LBR per cycle: percentage (95% CI)</td>
<td>8.7 (6.3;12.8)</td>
<td>13.8 (11.0;18.6)</td>
<td>1.6 (95% CI 1.0–2.4)</td>
<td>0.03</td>
</tr>
<tr>
<td>Per started COS cycle</td>
<td>n = 298 cycles</td>
<td>n = 322 cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHB + pregnancy rate: percentage (95% CI)</td>
<td>6.7 (4.5;10.5)</td>
<td>12.4 (9.5;16.8)</td>
<td>1.9 (95% CI 1.1–3.1)</td>
<td>0.02</td>
</tr>
<tr>
<td>LBR per cycle: percentage (95% CI)</td>
<td>6.4 (4.2;10.1)</td>
<td>11.8 (9.0;16.3)</td>
<td>1.9 (95% CI 1.1–3.1)</td>
<td>0.02</td>
</tr>
<tr>
<td>Per IUI cycle</td>
<td>n = 252 cycles</td>
<td>n = 307 cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHB + pregnancy rate: percentage (95% CI)</td>
<td>7.1 (4.7;11.4)</td>
<td>13.0 (10.0;17.6)</td>
<td>1.8 (95% CI 1.1–3.1)</td>
<td>0.02</td>
</tr>
<tr>
<td>LBR per cycle: percentage (95% CI)</td>
<td>6.8 (4.4;11.0)</td>
<td>12.4 (9.5;17.0)</td>
<td>1.8 (95% CI 1.1–3.1)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

FHB, fetal heart beat; LBR, live birth rate.

*The P-value is calculated using statistical tests taking into account repeated measurements per individual, to correct for the fact that individuals were allowed to participate more than once in this study (see Methodology section).

*Absolute difference between proportions of treatments.
treatment in case of lack of conception. In our design, repeated randomization after each cycle allowed couples to have both treatments, thus increasing the willingness to participate. Indeed, our study is the largest in this area thus far. The repeated randomization makes our study principally different from a crossover study, in which randomization determines the allocation in the first cycle, but then ‘crosses over’ to the other intervention, thus generating bias (Khan et al., 1996). Repeated randomization prevented this type of bias. Furthermore, our statistical analysis accounts for patient participation to multiple cycles, no matter whether cycles of the same patient appeared in the same or different treatment groups. The clustered study design (multiple observations per couple) does not lead to a bias in the sense of an over- or underestimated treatment effect. A possible problem with clustered data might be related to the precision of the estimated effect. The precision could be over- or underestimated, depending on whether couples were randomized more likely within the same or different treatment groups. This could, respectively, lead to too narrow or too wide confidence intervals for the treatment effect, and hence too liberal or too conservative P-values. However, statistical techniques to correct for such clustering effect are nowadays commonly used and applied in this study (Aerts et al., 2002). Therefore, we would like to argue that the results of our study are both unbiased and with correct precision estimates.’

Since it is relevant to analyse statistically independent couples instead of cycles, we performed an additional post hoc analysis restricted to the first IUI cycle only per unique couple, and restricted to IUI cycles rank 2—7 separately. The results are shown in Supplementary Table SII. It appears that the overall result of our study (primary outcome higher clinical pregnancy rate in hMG group than in CC group) can be explained by the results obtained in cycles 2—7, as the clinical pregnancy rate was not significantly different between both groups in the first cycle per unique couple (P = 0.577, hMG 15/117 (12.82%) versus CC 19/124 (15.32%). However, we stress that the study was not large enough to analyse the data by first cycle only per unique couple since 438 cycles in total were needed according to the sample size calculation. When we calculated only the first cycle per unique couple there were only 227 cycles in total.

It can be questioned why we continued the study after sample calculation was reached. Indeed, even though the 438 cycles needed according to sample calculation were attained in early 2010, we then decided to continue our study until December 2011 for the following reasons. Firstly, a considerable amount of trial medication was left; secondly, we agreed that the quality and power of our study would increase if the final number of randomized patients was higher than the minimally required number. Belgian patients were not affected in a negative way by continuing the study because new Belgian legislation in 2007 regulated that gonadotrophins for IUI cycles are only reimbursed after four failed IUI cycles stimulated with CC. Furthermore, we stress that no interim analysis was done after the initial sample size was reached and that data analysis was only done once at the end of the study. A 10% difference between both treatment groups was assumed for the initial sample size calculation. Such a 10% difference could not be demonstrated in our study, where the difference in proportions was estimated as 5%. The initial sample size calculation would have led to 47% of power for demonstrating such difference. As a result of the incremented sample size, the 5% difference could be demonstrated as statistically significant.

The pregnancy rates in our study were comparable to previously published studies with a similar research design (4–14% for CC COS and 7–20% for gonadotrophin COS) and confirm the higher pregnancy rates observed after IUI combined with gonadotrophin COS than after IUI combined with CC COS (Table I) (Karlstrom et al., 1993; Balasch et al., 1994; Kamel et al., 1995; Karande et al., 1995; Manganelli et al., 1997; Hanoun et al., 1998; Matorras et al., 2002; Berker et al., 2011). However, two RCTs, that failed to confirm our results, used different protocols for COS with clomiphene (starting dose of 50 mg CC in our study, but 100 mg CC (Dankert et al., 2007) and 50 or 100 mg CC (Ecochard et al., 2000) in the other studies) and for COS with gonadotrophins (hMG in our study versus recombinant FSH in one other study (Dankert et al., 2007); low dose step up COS with starting dose of 37.5 or 75 IU in our study versus much higher starting dose of 150 IU IM given on Days 4, 6, and 9 of the cycle in one other study (Ecochard et al., 2000)).

Natural conception can occur during expectant management in women with unexplained infertility (Pandey et al., 2014). However, in our study only 18 women (18/657 or 3% of all randomized cycles) conceived spontaneously before the IUI cycle was started and they were equally divided between both groups. Indeed, these spontaneous pregnancies have been accounted for in the intention-to-treat analysis (Table IV). Our observation is in line with previous publications (Manganelli et al., 1997; Dankert et al., 2007; Steevers et al., 2007) and can possibly be explained by the relatively long randomization-to-intervention interval as well as the fact that at least some couples had a reasonable prognosis themselves. Indeed, randomization was done immediately after the patient had given informed consent, and not at the beginning of the cycle just before controlled ovarian stimulation was started.

After COS therapy was started (secondary analysis), we observed a higher cancellation rate per started cycle in the CC group than in the hMG group, mostly due to insufficient ovarian response and to premature LH and/or progesterone rising (Fig. 1)( Supplementary Table SI). A similar trend was found in the study of Dankert et al. (2007) and Ecochard et al. (2000) but not in the study of Matorras et al. (2002). Other studies fail to report the endometrial thickness and cancelled cycles. We hypothesize that the lower cancellation rate per started cycle and the increased endometrial thickness at the time of hCG injection both contributed to the higher ongoing pregnancy and live birth rates in the hMG group when compared with the CC group.

In the present study, it was surprising that the mean number of pre-ovulatory follicles was lower in the hMG group, when compared with the CC group (Table IIII), since other RCTs have demonstrated that COS with gonadotrophins is associated with more multifollicular development and more multiple pregnancies than COS with clomiphene (Table I). However, in most other studies the mean number of preovulatory follicles was not reported (Table I). Nevertheless, one other group (Balasch et al., 1994) also observed significantly more cycles with ≥2 mature follicles after COS with CC than after COS with FSH 75 IU started on cycle day 7. We hypothesize that this difference can be explained by the fact that multifollicular development is dependent on the starting dose of gonadotrophins, and the starting dose of gonadotrophins was higher in most other studies (Kamel et al., 1995; Matorras et al., 2002; Berker et al., 2011) than the gonadotrophin starting dose of 37–75 IU used in our study. The low mean number of pre-ovulatory follicles in this trial also explains why the overall multiple pregnancy rate after COS with IUI in our study was low (7%) and similar in the CC (5%) and hMG (8%) groups, much lower than the multiple pregnancy rate reported in other studies after COS with gonadotrophins
This may be again related to the low starting dose of gonadotrophins used in our study (Table I). Multifollicular development and increased multiple pregnancy rate after COS with gonadotrophins can be avoided by the use of a low-dose stimulation protocol, the possibility of selective follicular aspiration before IUI and/or strict cancellation criteria (allowing only maximum two mature follicles), while maintaining an acceptable live birth rate, as has been recommended by several authors (Cohlen 2005; Cantineau et al., 2007; Ghesquière et al., 2007; van Rumste et al., 2008; Dickey, 2009; Practice Committee of the American Society for Reproductive Medicine, 2011a,b).

Endometrial thickness at the time of hCG injection was significantly lower in the CC group than in the hMG group, in contrast with our assumption that co-treatment of CC with ethinyl estradiol would neutralize the anti-estrogenic endometrial effect of CC documented in many other studies (Gerli et al., 2000; Unfer et al., 2001, 2004; Dhabhadi et al., 2003; Haritha and Rajagopalan, 2003; Saturapod et al., 2014). This concept was based on one RCT (Gerli et al., 2000) comparing the endometrial thickness and pregnancy rate between treatment with CC with or without ethinyl estradiol (EE2) and demonstrated that ethinyl estradiol can reverse the deleterious effects of CC on endometrial thickness contributing to a significantly lower miscarriage rate and significantly higher ongoing pregnancy rate in the CC + EE2 group compared with CC alone. The value of co-treatment of CC with estrogens requires more study before it can be recommended in clinical practice according to a Cochrane analysis (Cantineau et al., 2007).

In conclusion, our study demonstrated that in an IUI programme, COS with low-dose hMG is associated with better reproductive outcome (higher clinical pregnancy and live birth rates, similar and low multiple pregnancy and multiple live birth rates) per randomized cycle, per started cycle and per IUI cycle, when compared with COS with CC. COS with hMG resulted in a lower cancellation rate per started cycle and a lower follicular aspiration rate before IUI. When we express the results (pregnancy and live birth rate) as number needed to treat (NNT), the NNT equals 18 for pregnancy rate and 19 for live birth rate. This means that 18 or 19 women needed to be treated to get, respectively, an extra pregnancy or live birth, after hMG treatment when compared with CC treatment before IUI. Based on these results, which are in line with those from previous less powered studies, we recommend that IUI combined with low-dose gonadotrophins is the treatment of choice for patients with an indication for IUI treatment, as long as they are willing to accept the concept of daily subcutaneous injections and to pay the extra cost associated with gonadotrophin treatment. A health economic analysis of our data, including the cost of failed cycles and multiple pregnancies and births, is planned to test the hypothesis that COS with low-dose hMG in IUI is associated with increased cost-effectiveness when compared with COS with CC in IUI.

**Supplementary data**

Supplementary data are available at http://humrep.oxfordjournals.org.

**Acknowledgements**

The authors wish to thank the medical, paramedical and technical staff of the Leuven University Fertility Center.

**Authors’ roles**

T.M.D. and K.P. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: T.M.D. Acquisition of data: All authors. Analysis and interpretation of data: K.P., S.D., D.D.N., T.M.D. Drafting of the manuscript: K.P. Critical revision of the manuscript for important intellectual content: D.D.N., B.W.M., T.M.D. Statistical analysis: A.L. Administrative, technical or material support: M.W.

**Funding**

T.M.D. and K.P. were supported by the Clinical Research Foundation of UZ Leuven, Belgium. This study was also supported by the Ferring company (Copenhagen, Denmark) which provided free medication (Menopur) required for the group of patients who were randomized in the hMG COS group.

**Conflict of interest**

The authors have no conflicts of interest to declare. The Ferring company was not involved in the study design, data analysis, writing and submission of the paper.

**References**


Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H, van Dop PA. A randomized clinical trial of clomiphene citrate versus low dose recombinant FSH for ovarian hyperstimulation in...


Matorras R, Diaz T, Corcostegui B, Ramón O, Pijoan JI, Rodríguez-Escudero FJ. Ovarian stimulation in intrauterine insemination with donor sperm; a randomized study comparing clomiphene citrate in fixed protocol versus highly purified urinary FSH. *Hum Reprod* 2002;17:2107–2111.


