RBM ONLINE

ELSEVIER

Article

A randomized controlled trial investigating the use of a predictive nomogram for the selection of the FSH starting dose in IVF/ICSI cycles

Adolfo Allegra [a,*], ***Angelo Marino*** [a], ***Aldo Volpes*** [a], ***Francesco Coffaro*** [a], ***Piero Scaglione*** [a], ***Salvatore Gullo*** [b], ***Antonio La Marca*** [c]

[a] *Reproductive Medicine Unit, ANDROS Day Surgery Clinic, Palermo, Italy*
[b] *Medical Statistics Unit, ANDROS Day Surgery Clinic, Palermo, Italy*
[c] *University of Modena and Reggio Emilia, Modena, Italy*

Adolfo Allegra is a gynaecologist and an endocrinologist. He is aggregate Professor of Surgery in Reproductive Medicine at the University of Palermo and is the Director of the Reproductive Medicine Unit of ANDROS Day Surgery Clinic, Palermo. His main clinical and research interests are in the field of reproductive endocrinology and reproductive endoscopy. He has published more than 100 scientific papers and has given a large number of international presentations.

KEY MESSAGE
This trial found that the FSH starting dose in IVF/ICSI cycles may be selected according to patient's age and serum AMH and FSH concentrations. This strategy increased the proportion of patients with an optimal response. The possible effects of this approach on pregnancy and live-birth rates needs further investigation.

A B S T R A C T

The number of oocytes retrieved is a relevant intermediate outcome in women undergoing IVF/intracytoplasmic sperm injection (ICSI). This trial compared the efficiency of the selection of the FSH starting dose according to a nomogram based on multiple biomarkers (age, day 3 FSH, anti-Müllerian hormone) versus an age-based strategy. The primary outcome measure was the proportion of women with an optimal number of retrieved oocytes defined as 8–14. At their first IVF/ICSI cycle, 191 patients underwent a long gonadotrophin-releasing hormone agonist protocol and were randomized to receive a starting dose of recombinant (human) FSH, based on their age (150 IU if ≤35 years, 225 IU if >35 years) or based on the nomogram. Optimal response was observed in 58/92 patients (63%) in the nomogram group and in 42/99 (42%) in the control group (+21%, 95% CI = 0.07 to 0.35, $P = 0.0037$). No significant differences were found in the clinical pregnancy rate or the number of embryos cryopreserved per patient. The study showed that the FSH starting dose selected according to ovarian reserve is associated with an increase in the proportion of patients with an optimal response: large trials are recommended to investigate any possible effect on the live-birth rate.

* Corresponding author.
E-mail address: adolfo.allegra@clinicaandros.it (A. Allegra).
http://dx.doi.org/10.1016/j.rbmo.2017.01.012

Introduction

The number of oocytes retrieved is considered a relevant prognostic marker in women undergoing IVF/ intracytoplasmic sperm injection (ICSI) cycles, and consistent evidence shows that an optimal – rather than a maximal – oocyte yield is the preferred achievement after controlled ovarian stimulation (COS) when fresh embryo transfer is scheduled. In fact, live-birth rates steadily increase when an optimal number of oocytes is collected, whereas low response and hyperresponse are associated with lower implantation rates, increased obstetrical risks and, at least when considering hyper response, increased risk of ovarian hyperstimulation syndrome (OHSS) in the fresh cycle (Baker et al., 2015; Sunkara et al., 2011, 2015). While the use of different drugs to control the spontaneous LH surge may affect the ovarian response to stimulation, with protocols based on gonadotrophin-releasing hormone (GnRH) antagonist usually associated with a reduced duration of ovarian stimulation and the total FSH dose needed, it is universally recognized that choosing different doses of gonadotrophins for different patients is the most important clinical decision in the planning of IVF cycles for infertile couples (Fauser *et al.*, 2008; La Marca and Sunkara, 2014; Moolenaar et al., 2011). However, although exogenous FSH has been used for decades and millions of cycles have been performed worldwide, criteria for selecting the proper starting dose of FSH in daily clinical practice have not yet been clearly identified (Fauser *et al.*, 2008). Clinicians commonly choose the FSH starting dose in accordance with clinical history and criteria, the most important being the ovarian response to stimulation in previous IVF cycles. If no previous cycles have been performed, the choice will be based on such criteria as women's age and markers of ovarian reserve (Fleming et al., 2013; Howles *et al.*, 2006). Currently used markers of ovarian reserve include FSH, anti-Mullerian hormone (AMH) and antral follicle count (AFC), with the last two markers having the best performance in predicting ovarian response to exogenous FSH (Broer *et al.*, 2011, 2013a, 2014; Fleming et al., 2015; Iliodromiti and Nelson, 2015; La Marca et al., 2010; Lan et al., 2013; Nelson et al., 2007). In particular, AMH and AFC are nowadays considered two markers with similar diagnostic performance (Broer et al, 2013b), even if recent evidence seems to suggest some superiority of AMH over AFC, at least when considering their performance at the multicentric level; this is due to the lower variability of AMH when compared with AFC (Anderson et al., 2015). On the use and efficacy of the single value of AMH in tailored treatment, two studies have been published (Nelson et al., 2009; Yates et al., 2011). In both studies, which were not randomized controlled trials (RCT), three different FSH doses were proposed for three different AMH strata levels, i.e. the higher the serum AMH the lower the starting dose of FSH. Both studies indicated that the 'AMH-stratified care' may lead to a reduction of cancelled cycles and increased pregnancy rate.

Recently, an easy to use nomogram has been proposed in order to calculate the most appropriate FSH starting dose in IVF cycles when the long GnRH agonist protocol is used (La Marca et al., 2012a). The nomogram is based on a patient's age, serum day 3 FSH and AMH, and may be the basis for the individualization of the FSH dose for patients entering the IVF programme. In this model, AMH is the leading predictor, explaining the large part of the model variability followed by serum FSH and female age that can improve just by a little, even if significantly, the accuracy of the model (**Figure 1**). Such a nomogram, however, needs to be externally validated. In fact, before clinicians can adopt any treatment model in routine clinical practice, the accuracy of the model should be independently evaluated in a population different from the one on which the model was elaborated. External validation of the model is therefore crucial to assess the generalizability to other populations.

The objective of this RCT was to investigate the performance of the nomogram in selecting the most appropriate FSH starting dose in IVF/ICSI cycles. In particular, women undergoing IVF/ICSI were randomized to receive a starting dose of recombinant (human) FSH (rFSH) selected merely on the basis of their age (150 IU if ≤35 years, 225 IU if >35 years) or on the basis of their ovarian reserve, by using the nomogram including age, day 3 serum FSH and AMH.

Materials and methods

Participants

This two-arm, single-centre, prospective, randomized, interventional trial involved 194 couples attending their first IVF/ICSI cycle at the Andros Day Surgery Clinic, Palermo, Italy. All the women had been trying to conceive for at least 12 months and had undergone a fertility workup.

The couples were only included if all the following inclusion criteria were satisfied: first IVF/ICSI cycle, female age between 18 and 40 years, body mass index (BMI) between 18 and 25 kg/m^2, serum AMH concentrations between 1.0 and 4.0 ng/ml, basal serum day 3 FSH ≤15 IU/l, normal regular menstrual cycles, ranging from 25 to 33 days in length, normal thyroid-stimulating hormone (TSH) and prolactin concentrations, normal uterine cavity as assessed by hysteroscopy or sonohysterography or three-dimensional ultrasound and presence of both ovaries. Clinical exclusion criteria were: irregular menstrual cycles, polycystic ovary syndrome, severe endometriosis (stage III-IV of the American Society for Reproductive Medicine revised classification, American Society for Reproductive Medicine, 1997), previous ovarian surgery, presence of ovarian cysts, use of hormonal contraception in the previous three months, any known metabolic or endocrinological disease.

Interventions

COS was performed after pituitary down-regulation with a GnRH agonist (buserelin acetate, Sanofi, Italy; 0.1 ml subcutaneously twice per day), beginning from day 21 of the previous cycle until the day of recombinant human chorionic gonadotrophin administration. Multifollicular development was achieved by daily injections of rFSH (Gonal F; Merck Serono, Italy), commencing after at least 12 days of pituitary down-regulation. Ovarian suppression was demonstrated by thin endometrium and low oestradiol concentrations.

Three couples were excluded because they dropped out (one for a spontaneous pregnancy, this being between the recruitment and the starting of the ovarian stimulation, two for personal reasons), thus 191 couples constituted the population included in the statistical analysis.

In the control group (n = 99), the starting dose of rFSH was fixed and established on the basis of the female age: 150 IU of rFSH if female age was ≤35 years or 225 IU of rFSH if female age was >35 years. In the study group (n = 92), an individualized starting dose of rFSH was decided on the basis of the nomogram based on female age, AMH and basal FSH.

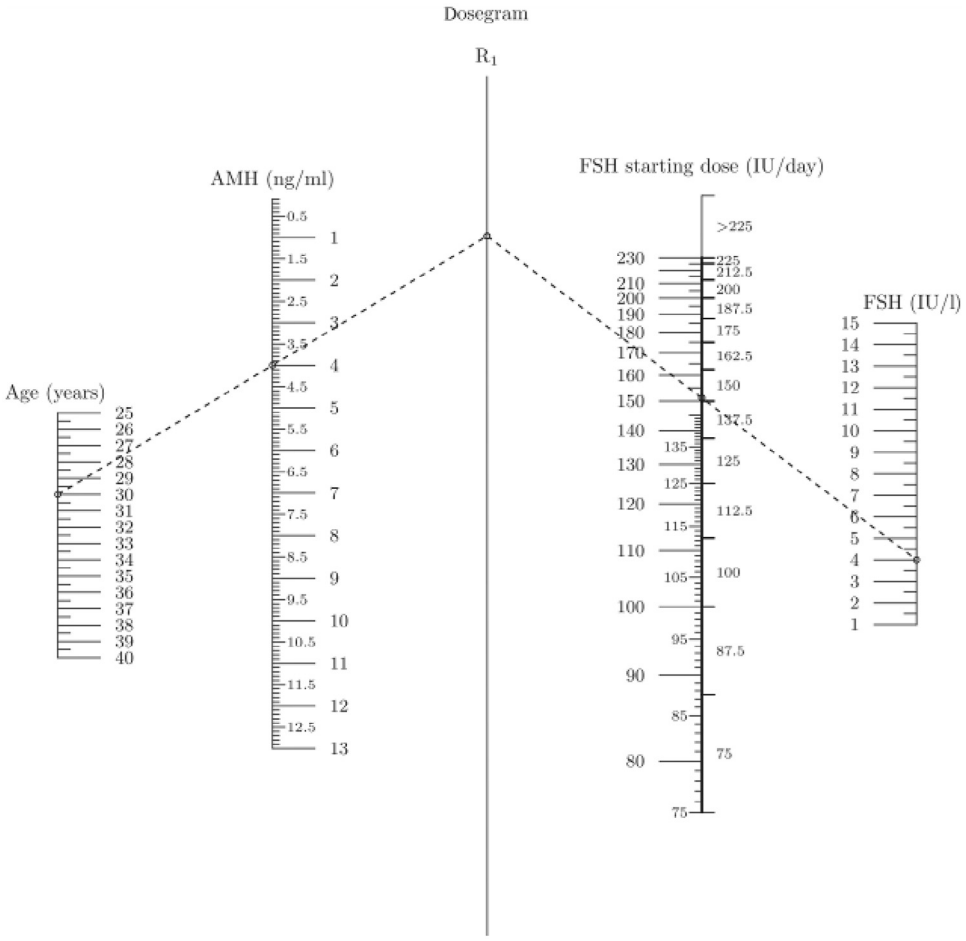

Figure 1 – The nomogram for the calculation of the FSH starting dose on the basis of age, serum day 3 FSH and anti-Müllerian hormone. In the example, for a 30-year-old patient with serum AMH and day 3 FSH concentrations of 4 ng/ml and 4 IU/l, respectively, the FSH starting dose should be 160 IU per day. This has been corrected to 150 IU/daily, since the dosage dial of the FSH delivery system is based on multiples of 12.5 IU (with permission from La Marca et al., 2012a).

In both groups, the rFSH dose was then maintained up to the first ultrasound follicular control (day 5 or 6 of stimulation). Ovarian stimulation was monitored on alternate days from day 5 or 6 of ovarian stimulation and the dosage of rFSH was maintained, increased or decreased depending on the patient's follicular response (Anckaert et al., 2012; Devroey *et al.*, 2012). Oocyte maturation was triggered with an injection of 250 µg subcutaneously of recombinant human chorionic gonadotrophin (HCG) (Ovitrelle; Merck Serono, Italy), when three or more follicles ≥17 mm in diameter were observed. Criteria for cycle cancellation before the day of HCG administration were either an inability to reach the HCG criterion (less than three growing follicles) or more than 20 follicles with a diameter of ≥10 mm.

The oocyte retrieval procedures were performed 36 h after the triggering of oocyte maturation. The IVF/ICSI procedures were performed in accordance with established protocols described in detail elsewhere (Volpes et al., 2004). All embryo transfers were carried out using a short Frydman set (Laboratoire CCD, Paris, France) between 48 and 120 h after oocyte retrieval. Generally, embryo transfer was performed on day 3 if less than four good quality embryos were available, while it was on day 5 if four or more embryos were available. Micronized progesterone, 600 mg vaginally (Prometrium, Rottapharm S.p.A., Italy) was used for luteal support starting from the day of oocyte retrieval. Pregnancy was confirmed by determining serum β-HCG concentration 14 days after oocyte retrieval in all patients. When the pregnancy test resulted positive, a second test was performed two days later. Ultrasound evaluations were performed 28 – 32 days after oocyte retrieval, and only gestational sacs with a clear fetal heartbeat were diagnosed as clinical pregnancies.

Assays

Blood samples for hormonal basal evaluation were taken in the early follicular phase (day 3), before any IVF-related drug administration. The IVF/ICSI procedures were performed in the next month or two after the blood sampling. The AMH measurement was performed by the General Laboratory of Andros Day Surgery Clinic, Palermo, Italy, using the Gen II ELISA assay (Beckman Coulter, Italy). The modified AMH Gen II assay was used for all AMH analyses in this study. The AMH Gen II assay is a two-step, sandwich-type enzymatic, microplated assay. Problems with the robustness of the Gen II assay were solved with a modified version of the AMH Gen II assay kit (reference A79765), including an additional assay step before calibrators were added (premixing). This additional step eliminates the complement and thereby overcomes the non-optimal assay reproducibility of the original AMH

Gen II assay. The standards cover a range of 0 – 22 ng/ml. The sensitivity is reported to be 0.1 – 0.21 ng/ml. Reported intra- and inter-assay coefficient of variations were <2% and <12%, respectively. Serum FSH was measured by a chemiluminescent assay (ADVIA Centaur, Siemens Healthcare Diagnostics, Milan,Italy). The sensitivity of the assay was 0.3 IU/l; intra-assay and inter-assay coefficients of variation were 2.7 and 3.1%, respectively.

Statistical analysis

Primary and secondary outcomes

The main outcome measure of the study was the proportion of women with an appropriate number of retrieved oocytes. The optimal number of oocytes was defined as ranging between 8 – 14 (La Marca and Sunkara, 2014; Polyzos and Sunkara, 2015). The same range of oocytes to define the 'appropriate ovarian response' has been adopted in other large multicentric trials (Arce et al., 2014; Bosch et al., 2015).

The secondary outcome measures were: total rFSH dose employed, treatment duration, serum oestradiol concentrations on rHCG day, number of growing follicles (≥11 mm) on rHCG day; number of large ovarian follicles ≥17 mm on rHCG day, embryos obtained, embryos transferred, fertilization rate, implantation rate, clinical pregnancy rate, OHSS rate, the number of cryopreserved embryos and the proportion of patients with cryopreserved embryos.

Sample size

In the present study it was expected that the control group might have an optimal response (8 – 14 oocytes) in 40% of patients, in accordance with a previously published randomized, open-label, assessor-blind, parallel-group, multicentre, multinational study (Andersen et al., 2006), in which women younger than 38 were treated with a long standard protocol + 225 IU of gonadotrophins and 40% of patients had an appropriate response. The optimal response rate in the nomogram group was assumed to be 40% under the null hypothesis H0 and 55% under the alternative hypothesis H1. A 15% increase in the percentage of women with an optimal response was considered clinically sufficient to promote the employment of an ovarian reserve marker-based strategy in IVF. Under these assumptions, using a one-sided Z-test and set at an alpha level of 0.05, it was calculated that sample sizes of 177 in the nomogram group and 177 in the control group achieve 80% power to detect a difference in the optimal response rate of 15%. Therefore, it was established that the study was conducted on a total sample of 360 patients, with 180 for each of the two arms.

A formal interim analysis was pre-planned at 50% of recruitment, by using the O'Brien-Fleming procedure (O'Brien and Fleming, 1979). In this approach, early on, the stopping criteria are conservative and they successively are reduced as the results become more reliable and stable. If the interim analysis reveals a superiority of the nomogram group with a difference significant at an α less than 0.005 (Schulz and Grimes, 2005), the study has to be prematurely stopped at a power of 50% and superiority is thus proven.

After informed consent was obtained, subjects who complied with all the selection criteria were randomly assigned to one of two treatment groups by giving them a code number from a randomization sequence (in order of enrolment). The randomization sequence was generated by a computer program software (PASW-17) using a simple randomization method. To guarantee the concealment of allocation, a staff member, who was not involved in the study, was in possession of the randomization sequence; in this way, after receiving information from the physician recruiting the couples, the staff member followed the randomization sequence allocating each couple to one of the two groups, without knowledge of which patient would receive which treatment.

All summaries and analyses are based on the intention-to-treat population defined as all randomized and exposed-to-treatment patients. Between the groups, differences in continuous variables were assessed with parametric or non-parametric statistics, as appropriate. Z-tests were conducted comparing the proportion of individuals who reported an optimal outcome response. Secondary endpoints were analyzed by using Fisher's exact test and chi-squared test to compare dichotomous characteristics of the two groups, and an independent *t*-test was applied to compare means for continuous variables. Effect sizes were calculated using Cohen's d to indicate the difference in magnitude. Cohen's effect sizes are understood as negligible (≥–0.15 and <0.15), small (≥0.15 and <0.40), medium (≥0.40 and <0.75), and large (≥0.75 and 1.10) (Cohen, 1988).

Statistical analyses were conducted by PASS 14 (NCSS, Kaysville, Utah, USA) and PASW 17 (SPSS Westlands Centre, Westlands Road, Quarry Bay, Hong Kong) by a professional statistician.

This study was approved by the local Ethical Committee (01/MR/13) on 18 January 2013. This RCT was registered to ClinicalTrials.gov (ID registration code: NCT01816789; date of trial registration: 21 March 2013).

Results

A total of 194 patients were randomized into two arms: 99 were assigned to the control group and 95 were assigned to the nomogram group. From this group, three couples dropped out before starting any treatment: two for personal reasons; one for a spontaneous pregnancy. Hence the intention-to-treat population consisted of 191 couples who were exposed to the treatment, 99 in the control group and 92 in the nomogram group, respectively.

The characteristics of the 191 participants were similar and not statistically different in the two treatment groups (**Table 1**). There were no significant differences between the two groups with regard to age, BMI, duration of infertility, AFC, serum AMH and basal FSH concentrations.

The most prevalent cause of infertility was the male factor in both groups (55/99; 56% and 52/92; 57% for the control and nomogram groups, respectively), followed by unexplained infertility, tubal factor

Table 1 – Demographic characteristics of patients included in the study.

	Control group (*n* = 99)	Nomogram group (*n* = 92)
Age (years)	34.4 ± 3.9	33.5 ± 4.3
BMI (kg/m^2)	22.7 ± 2.2	22.4 ± 2.3
Infertility duration (years)	3.32 ± 2.12	3.55 ± 2.59
Basal FSH (IU/l)	7.9 ± 5.3	7.1 ± 2.5
AFC (*n*)	10.8 ± 4.9	11.7 ± 5.5
AMH (ng/ml)	2.5 ± 0.9	2.4 ± 1.01

Values are mean ± SD. There were no statistically significant differences between the two groups.
AFC, antral follicle count; AMH, anti-Müllerian hormone; BMI, body mass index.

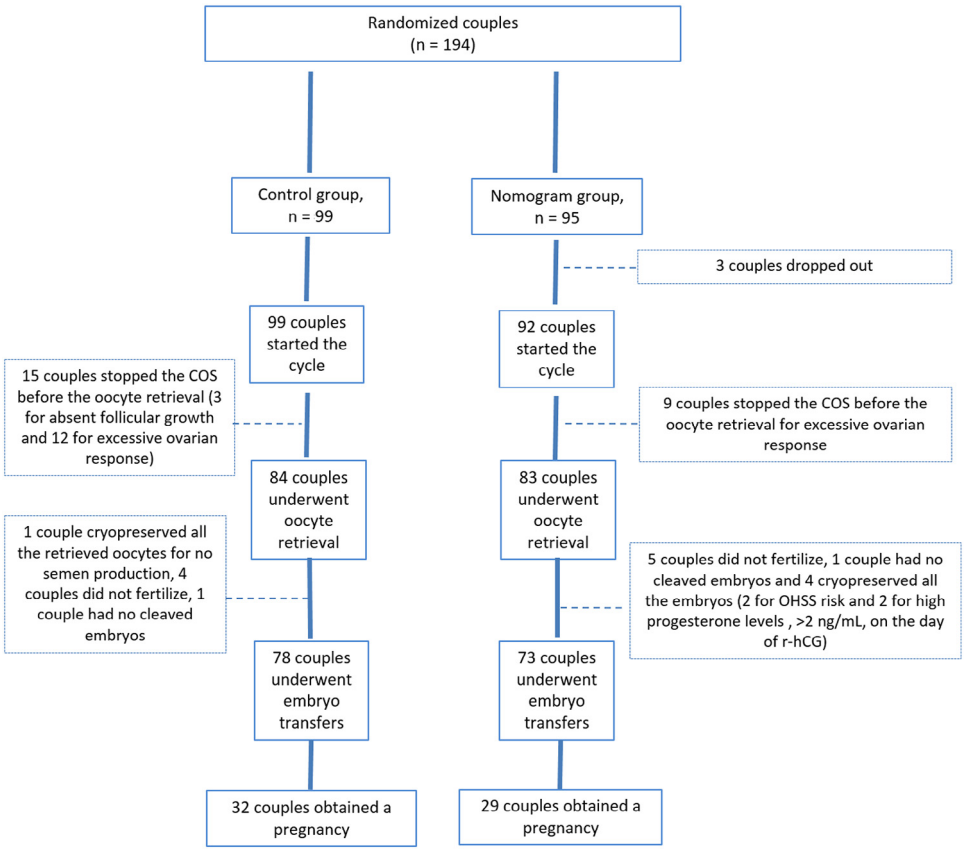

Figure 2 – Flow diagram of trial participants.

and mixed. Causes of infertility were similarly distributed in the two groups (Z = –0.38).

Twenty-four of the 191 patients (13%) stopped the treatment for hyper- (*n* = 21) or hypo- (*n* = 3) response; the suspended cycles were not differently distributed between the two groups (15/99; 15% and 9/92; 10% in control and nomogram group, respectively; Z = 1.04). Therefore, 167/191 patients (87%) underwent the oocyte retrieval: 84/99 (85%) in the control group and 83/92 (90%) in the nomogram group. A total of 151 patients underwent embryo transfer procedures, which means that 79% of patients who started the cycles (90% of those who underwent oocyte retrieval) underwent an embryo transfer procedure. The participant flow diagram is shown in **Figure 2**.

To test the hypothesis of superiority of the new AMH-based individualized treatment, the difference between the nomogram and control group with regard to the primary and secondary outcomes was analyzed. The primary outcome was the percentage of patients with the optimal number (8 – 14) of retrieved oocytes. The rate of patients with an optimal outcome was 63% (58 out of 92 patients) in the nomogram group and 42% (42 out of 99 patients) in the control group. The difference between the groups (d = 21%, 95% confidence interval [CI] = 0.07 to 0.35) was statistically significant with a *P*-value of 0.0037 (**Figure 3**). According to Schulz and Grimes (2005), it was decided to terminate the trial since the difference reached a *P* -value lower than the stopping level of 0.005 determined by the O'Brien-Fleming procedure for the interim analysis (**Table 2**). The rate of patients with suboptimal ovarian response (<8 oocytes) was significantly reduced in patients treated in accordance with the nomogram (24/92; 26% versus 40/99; 40%; d = –14% 95% CI = 0.01 to 0.28, *P* = 0.040). The proportion of women with more than 14 oocytes in the nomogram group was lower than in the control group, but the statistical significance was not reached (10/92; 11% versus 17/99; 17%; d = –6% 95% CI = –0.05 to 0.15, *P* = 0.328). The frequency distribution of retrieved oocytes in the two groups is shown in **Figure 4**.

Secondary endpoints are summarized in **Tables 3 and 4**. The mean starting dose of rFSH was higher in the nomogram group than in the control group (*P* = 0.001, **Table 3**), whereas the mean total IU of rFSH administered per cycle was not significantly different between the two groups. The frequency distribution of the starting dose of rFSH is reported in **Figure 5**. As shown, in the control group, 56/99 (57%) and 43/99 (43%) of patients received a starting dose of 150 and 225 IU respectively. In the nomogram group, 44/92 (48%) of patients received the dose of 225 IU, while the remaining patients received a dose ranging between 125 IU and 212.5 IU per day. In the fixed-dose and personalized group 72 and 56 women (73% and 61%), respectively had the dose of rFSH adjusted during ovarian stimulation, and the difference was statistically significant (*P* = 0.01).

There were no significant differences between groups in serum oestradiol and progesterone concentrations, number of growing follicles (≥11 mm), number of large (≥17 mm) ovarian follicles on the day of HCG, total FSH dose employed and treatment duration. The mean number of retrieved and mature oocytes, and the percentage of mature oocytes were not significantly different in the nomogram and the control group. Patients in the nomogram and control group showed comparable rates of fertilization, cleavage and implantation (**Table 4**). Furthermore, when the clinical pregnancy rate was compared between the nomogram and the control group, results did not

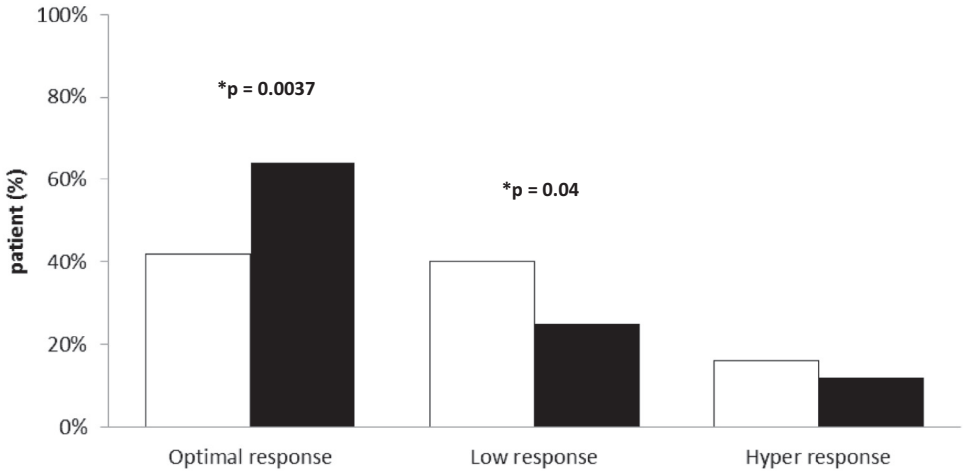

Figure 3 – The proportion of women with appropriate, low or hyper ovarian response in the control group (white) and in the nomogram group (black).

show significant differences, either for subjects who started the cycle or for those reaching the embryo transfer (**Table 4**). The percentage of subjects with at least one cryopreserved embryo was not significantly different between the two groups (24/83; 28.9% versus 17/84; 20.2% for the nomogram and control group respectively); also, the difference between the mean number of frozen embryos per patient did not reach statistical significance. No cases of moderate or severe OHSS were observed in this study.

Table 2 – Optimal response and interim analysis results.

	Control group (*n* = 99) *n* (%)	Nomogram group (*n* = 92) *n* (%)	Z	*P*-value	O'Brien-Fleming spending function	
					Z-value Bound (95 % CI)	*P*-value Bound (95% CI)
Women with optimal (8–14) retrieved oocytes	42 (42.4)	58 (63.0)	2.85	0.0037	–	–
Interim–1 look (50%)	–	–	–	–	*2.52 (2.43–2.59)*	*0.005 (0.005–0.007)*
Interim–2 look (100%)	–	–	–	–	*1.67 (1.63–1.73)*	*0.050 (0.042–0.051)*
Women with <8 retrieved oocytes	40 (40.4)	24 (26.1)	2.11	0.040	–	–
Women with >14 retrieved oocytes	17 (17.2)	10 (10.9)	1.04	NS	–	–

CI, confidence interval; NS, not statistically significant.

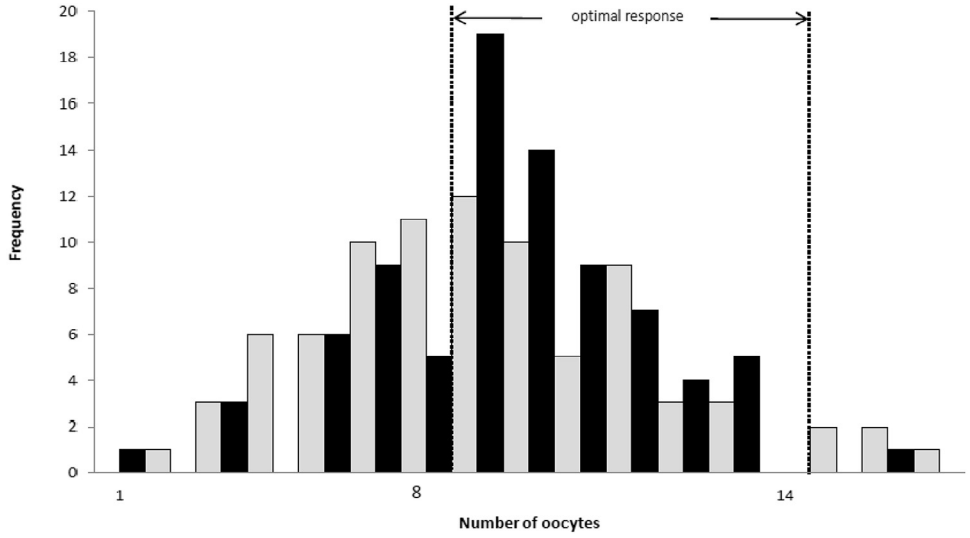

Figure 4 – The frequency of retrieved oocytes in the two groups (black: nomogram group; grey: control group).

Table 3 – Outcome of ovarian stimulation in women treated with a fixed (control) versus personalized (nomogram) rFSH dose.

Parameter	Control group	Nomogram group	*P*-value
	n = 99	*n* = 92	
Mean starting dose of rFSH (IU)[a]	182.6 ± 37.4	201.1 ± 28.4	0.001
Total rFSH used (IU)[a]	2048 ± 681	2037 ± 733	NS
Patients with dose adjustment; *n* (%)	72 (73)	56 (61)	0.01
Serum oestradiol on HCG day (pg/ml)[a]	1688.6 ± 484	1782.3 ± 536	NS
Serum progesterone on HCG day (ng/ml)[a]	0.9 ± 0.4	1.0 ± 0.6	NS
Growing follicles (≥11 mm)[a]	11.4 ± 3.4	12.1 ± 3.3	NS
Large follicles (≥17 mm)[a]	7.4 ± 2.4	7.8 ± 2.3	NS
Treatment duration (days)[a]	11.2 ± 1.5	10.8 ± 1.6	NS
Retrieved oocytes (*n*)[a]	8.2 ± 3.2	8.5 ± 2.6	NS
Mature oocytes (*n*)[a]	6.6 ± 3.0	7.1 ± 2.5	NS
Mature oocytes (%)	81	83	NS

[a] Values are mean ± SD.
HCG, human chorionic gonadotrophin; NS, not statistically significant; rFSH, recombinant (human) FSH.

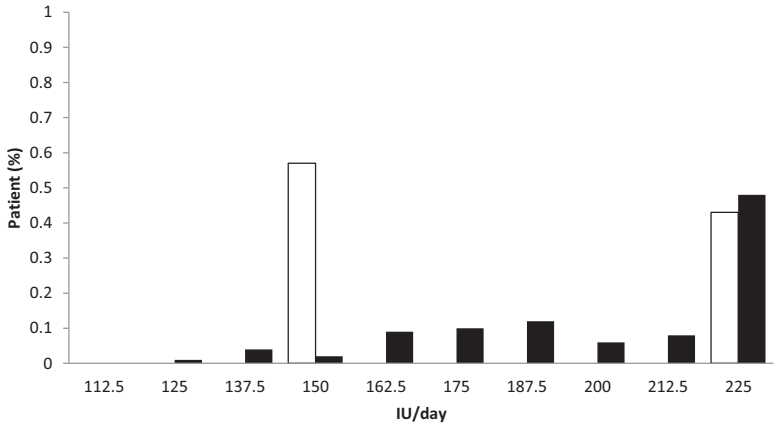

Figure 5 – The frequency distribution of the starting doses in the two groups (black: nomogram group; white: control group).

Discussion

The results of this RCT demonstrate that the personalization of the FSH starting dose in women with a normal ovarian reserve at their first IVF cycle, is associated with a significant increase in the rate of women with optimal ovarian response. Whereas in repeated IVF cycles, the choice of the FSH starting dose is mainly based on the response observed in previous attempts, in the first IVF cycle, the choice of the dose is mainly based on empirical methods and usually responds to a personal criterion by which the FSH dose rises with an increase in female age. Of course, this method, albeit useful, may be improved by adding other variables to the decision-making algorithm.

Results of previous studies of recombinant FSH have clearly shown that ovarian response to FSH depends mainly on the status of ovarian reserve (Arce et al., 2014; Fauser *et al.*, 2008; La Marca et al., 2010, 2012b, 2013; Nelson et al., 2009). Hence, an accurate measurement of functional ovarian reserve will give the clinicians a very useful tool for individualizing treatment. On the other hand, whilst tailored therapy based on markers of ovarian reserve appears to be an agreed-upon approach by the majority, studies suggesting ways of determining individualized therapy (compared with 'one size fits all') are scarce.

While the use of one single marker of ovarian reserve may be sufficient to correctly inform IVF (Lan et al., 2013; Nelson et al., 2009; Yates et al., 2011), recent studies clearly showed that algorithms based

Table 4 – Embryological data relative to the IVF cycle in women treated with a fixed (control) or personalized (nomogram) rFSH dose.

	Control group	Nomogram group
	n = 99	*n* = 92
Fertilization rate[a] (%)	356/551 (64.6)	386/586 (65.8)
Cleavage rate (%)	278 (78.1)	283 (73.3)
No. of obtained embryos[b]	4.3 ± 2.7	4.7 ± 3.0
No. of cleaved embryos[b]	3.4 ± 2.2	3.5 ± 2.3
No. of transferred embryos[b]	1.8 ± 0.4	1.8 ± 0.5
Implantation rate[c] (%)	35/142 (24.6)	34/129 (26.4)
Day 2 embryo transfer, *n* (%)	20 (25.6)	18 (24.6)
Day 3 embryo transfer, *n* (%)	38 (48.8)	38 (52.1)
Day 5 embryo transfer, *n* (%)	20 (25.6)	17 (23.3)
Clinical pregnancy rates[d] (%)	32 (32.3)	29 (31.5)
Clinical pregnancy rates[c] (%)	32 (41.0)	29 (39.7)
No. women with at least one cryopreserved embryo, *n* (%)	17/84 (20.2)	24/83 (28.9)
Mean number of cryopreserved embryos per patient	0.31 ± 0.68	0.58 ± 1.17

[a] Calculated as number of fertilized oocytes/number of mature inseminated oocytes.
[b] Values are mean ± SD. There were no statistically significant differences between the two groups.
[c] Per embryo transfer.
[d] Per started cycle.

on multiple biomarkers may improve the accuracy of the prediction of ovarian response hence creating the rationale for using complex models for the individualization of the FSH dose for patients at individual level. In the past, several algorithms based on multiple markers and including serum FSH and AFC have been proposed and externally validated (Howles *et al.*, 2006; La Marca and Sunkara, 2014; Olivennes *et al.*, 2009, 2015; Popovic-Todorovic et al., 2003a, 2003b).

Some years ago, the use of an algorithm including female age, serum FSH and AMH was proposed for identifying the most correct FSH starting dose on an individual patient basis (La Marca et al., 2012a). In the original study, the model predicted a dose lower than 150 IU in 7% of patients, a dose between 150 and 187.5 IU in 9% of patients, or between 187.5 and 225 IU in 23% and 225 IU in 61% of patients, and suggested that the FSH starting dose should be finely tuned in line with the extent of functional ovarian reserve (La Marca et al., 2012a).

The same concept, albeit in an indirect way, was also confirmed in a recent study in which a clear, direct, dose-response relationship was reported between the FSH starting dose and the ovarian response, both in terms of retrieved oocytes (Arce et al., 2014) or steroid production (Bosch et al., 2015). When the women were divided into groups in accordance with low or high serum AMH, it was clear that the slopes of the FSH dose–response curves differed significantly between the two AMH strata, demonstrating that a 10% increase in dose resulted in 0.5 and 1 more oocyte in the low and high AMH stratum, respectively (Arce et al., 2014).

In clinical practice, this translates as the idea that the higher the serum AMH, the larger the functional ovarian reserve, and the lower the FSH starting dose if a limited oocyte yield is the therapeutic objective.

The present study represents the first external validation of the nomogram developed by La Marca et al. (2012a) for calculation of the optimal starting dose for COS, based on easily available ovarian reserve markers such as AMH, FSH and age.

In the present study, the nomogram indicated a FSH dose of 225 IU in 48% of women while in the remaining 52% the dose was between 125 and 212.5 IU in accordance with the steps of 12.5 IU, hence clearly demonstrating an objective ability of personalization of the treatment when this nomogram is used.

In particular, in this RCT, it was clearly demonstrated that the use of an objective tool to select the FSH starting dose is followed by a significant increase in the proportion of women with an optimal ovarian response (63% versus 42%; $P = 0.0037$). The percentage of women with high ovarian response (>14 oocytes) was decreased, even if not significantly (11% versus 17%). At least in part the absence of statistical significance for this category of ovarian response may be due to the prematurely halted study at the first ad-interim analysis. However, on the basis of this study's finding, it may be calculated that 690 patients should be included in a trial aiming to demonstrate a significant reduction in women with a hyper-response to gonadotrophins. Interestingly, it was found a significant decrease in the proportion of women with suboptimal ovarian response (<8 oocytes) from 40% to 26% ($P < 0.05$) in control and nomogram group, respectively. All these figures clearly indicate increased efficacy of the IVF programme when the treatment for ovarian stimulation is personalized at the individual level. They also indicate that the proposed nomogram seems to be a possible tool for personalizing treatment on a patient's ovarian reserve, reducing the inter-operator variability derived from personal clinical experience.

In the present trial, the outcome of IVF, including clinical pregnancy rate and the number of cryopreserved embryos, was very similar in the two groups of patients, but this is not surprising since the study was not designed to detect any other difference that was not the primary outcome (the rate of women with an appropriate response).

The present trial has several limitations including the single-centre design and the primary outcome limited to ovarian response (oocytes retrieved). While the clinical pregnancy or live-birth rates are considered to be the 'gold' primary outcome of trials, when specifically investigating ovarian stimulation, the ovarian response (i.e. retrieved oocytes) remains an interesting outcome. Other trials have indeed adopted the number of retrieved oocytes as the main outcome (Arce et al., 2014; Bosch et al., 2015; Humaidan et al., 2015). In this study the optimal number of retrieved oocytes has been considered as ranging between 8 to 14; of course, we are aware that this topic has been a subject of debate for many years but a meta-analysis demonstrated that the ideal number of retrieved oocytes after conventional stimulation should be around 10 (Verberg et al., 2009). More interestingly the same range of oocytes (8–14) has been used in other large upcoming trials to define an adequate ovarian response in IVF (Arce et al., 2014; Bosch et al., 2015).

In the present trial, patients at the two 'extremes' of ovarian response (poor and high ovarian reserve) have been virtually excluded by the adopted inclusion and exclusion criteria. While recognizing that mainly women at the two extremes of ovarian reserve would benefit from personalization of the treatment, it seemed appropriate in this first study exploring the ovarian reserve-based strategy, to focus on women with normal ovarian reserve. Hence similarly to many other trials, women were excluded if affected by polycystic ovarian syndrome (indicative of high ovarian reserve) or with high serum FSH (indicative of low ovarian reserve). Future specific study should definitely investigate this therapeutic strategy in women with high and low ovarian reserve.

Another limitation is that in the trial, the long-standard GnRH agonist protocol has been used and the results, as a consequence, cannot also be considered valid for the GnRH antagonist cycles. The decision of using the GnRH agonist was because the nomogram was developed on IVF cycles performed with such a protocol (La Marca et al., 2012a). Moreover, different authors proposed the GnRH agonist long protocol as the most useful in expected normal responders (Nelson, 2013; Nelson et al., 2009; Yates et al., 2011). In the clinical practice, although the growing use of GnRH antagonist, the long standard protocol is still applied in a very high percentage of cycles, hence giving high clinical relevance to the present study.

In the present study, the FSH dose for the control group was just based on female age (150 IU if ≤35 and 225 IU if >35). While it is clear that in the 'real IVF', clinicians base the choice of the FSH dose on many criteria other than age alone, in a clinical trial setting, fixed rules should be set. Accordingly, in a very recent large multicentric trial, the FSH dose for patients was selected exactly as this study did (150 IU and 225 IU in young and old patients, respectively) (Toftager et al., 2016), hence reinforcing the design of the present study.

Clearly, the results of this study may be used to design proper multicentric trials, presupposing live birth to be the main outcome. At the same time, the present study has sufficient statistical potency to conclude that the personalization of the FSH dose increases the proportion of women with an appropriate ovarian response.

In conclusion, results from this study suggest that the nomogram could be a useful and easily available tool for the choice of a

tailored starting-dose and that it could be useful in optimizing treatment in a relevant percentage of patients. Further model validation, based on prospective and possibly randomized studies, is needed to demonstrate clearly whether the use of the AMH based nomogram can deliver significant advantages in terms of rate of optimal oocyte yield and, lastly, with regard to the outcome of the IVF cycle.

Acknowledgements

The authors thank Professor Bart Fauser for his constructive comments on a draft of the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ARTICLE INFO

Article history:
Received 30 July 2016
Received in revised form 11 January 2017
Accepted 17 January 2017
Declaration: The authors report no financial or commercial conflicts of interest.

Keywords:
Anti-Müllerian hormone
IVF/ICSI
ovarian reserve
ovarian response
predictive nomogram

REFERENCES

American Society for Reproductive Medicine, 1997. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. Fertil. Steril. 67, 817–821.

Anckaert, E., Smitz, J., Schiettecatte, J., Klein, B.M., Arce, J.C., 2012. The value of anti-Mullerian hormone measurement in the long GnRH agonist protocol: association with ovarian response and gonadotrophin-dose adjustments. Hum. Reprod. 27, 1829–1839.

Andersen, A.N., Devroey, P., Arce, J.C., 2006. Clinical outcome following stimulation with highly purified hMG or recombinant FSH in patients undergoing IVF: a randomized assessor-blind controlled trial. Hum. Reprod. 21, 3217–3227.

Anderson, R.A., Anckaert, E., Bosch, E., Dewailly, D., Dunlop, C.E., Fehr, D., Nardo, L., Smitz, J., Tremellen, K., Denk, B., Geistanger, A., Hund, M., 2015. Prospective study into the value of the automated Elecsys antimüllerian hormone assay for the assessment of the ovarian growing follicle pool. Fertil. Steril. 103, 1074–1080.

Arce, J.C., Andersen, A.N., Fernández-Sánchez, M., Visnova, H., Bosch, E., García-Velasco, J.A., Barri, P., de Sutter, P., Klein, B.M., Fauser, B.C., 2014. Ovarian response to recombinant human follicle-stimulating hormone: a randomized, antimüllerian hormone-stratified, dose-response trial in women undergoing in vitro fertilization/intracytoplasmic sperm injection. Fertil. Steril. 102, 1633–1640 e5.

Baker, V.L., Brown, M.B., Luke, B., Conrad, K.P., 2015. Association of number of retrieved oocytes with live birth rate and birth weight: an analysis of 231,815 cycles of in vitro fertilization. Fertil. Steril. 103, 931–938.

Bosch, E., Nyboe Andersen, A., Barri, P., García-Velasco, J.A., de Sutter, P., Fernández-Sánchez, M., Visnova, H., Klein, B.M., Mannaerts, B., Arce, J.C., 2015. Follicular and endocrine dose responses according to anti-Müllerian hormone levels in IVF patients treated with a novel human recombinant FSH (FE 999049). Clin. Endocrinol. 83, 902–912.

Broer, S.L., Dólleman, M., Opmeer, B.C., Fauser, B.C., Mol, B.W., Broekmans, F.J., 2011. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. Hum. Reprod. Update 17, 46–54.

Broer, S.L., Dólleman, M., van Disseldorp, J., Broeze, K.A., Opmeer, B.C., Bossuyt, P.M., Eijkemans, M.J., Mol, B.W., Broekmans, F.J., IPD-EXPORT Study Group, 2013a. Prediction of an excessive response in in vitro fertilization from patient characteristics and ovarian reserve tests and comparison in subgroups: an individual patient data meta-analysis. Fertil. Steril. 100, 420–429 e7.

Broer, S.L., van Disseldorp, J., Broeze, K.A., Dolleman, M., Opmeer, B.C., Bossuyt, P., Eijkemans, M.J., Mol, B.W., Broekmans, F.J., IMPORT Study Group, 2013b. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum. Reprod. Update 19, 26–36.

Broer, S.L., Broekmans, F.J., Laven, J.S., Fauser, B.C., 2014. Anti-Müllerian hormone: ovarian reserve testing and its potential clinical implications. Hum. Reprod. Update 20, 688–701.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Erlbaum, Hillsdale, New Jersey, Hove and London.

Devroey, P., Pellicer, A., Nyboe Andersen, A., Arce, J.C., Menopur in GnRH Antagonist Cycles with Single Embryo Transfer Trial Group, 2012. A randomized assessor-blind trial comparing highly purified hMG and recombinant FSH in a GnRH antagonist cycle with compulsory single-blastocyst transfer. Fertil. Steril. 97, 561–571.

Fauser, B.C., Diedrich, K., Devroey, P., Evian Annual Reproduction Workshop Group 2007, 2008. Predictors of ovarian response: progress towards individualized treatment in ovulation induction and ovarian stimulation. Hum. Reprod. Update 14, 1–14.

Fleming, R., Broekmans, F., Calhaz-Jorge, C., Dracea, L., Alexander, H., Nyboe Andersen, A., Blockeel, C., Jenkins, J., Lunenfeld, B., Platteau, P., Smitz, J., de Ziegler, D., 2013. Can anti-Müllerian hormone concentrations be used to determine gonadotrophin dose and treatment protocol for ovarian stimulation? Reprod. Biomed. Online 26, 431–439.

Fleming, R., Seifer, D.B., Frattarelli, J.L., Ruman, J., 2015. Assessing ovarian response: antral follicle count versus anti-Müllerian hormone. Reprod. Biomed. Online 31, 486–496.

Howles, C.M., Saunders, H., Alam, V., Engrand, P., FSH Treatment Guidelines Clinical Panel, 2006. Predictive factors and a corresponding treatment algorithm for controlled ovarian stimulation in patients treated with recombinant human follicle stimulating hormone (follitropin alfa) during assisted reproduction technology (ART) procedures. An analysis of 1378 patients. Curr. Med. Res. Opin. 22, 907–918.

Humaidan, P., Schertz, J., Fischer, R., 2015. Efficacy and Safety of Pergoveris in Assisted Reproductive Technology–ESPART: rationale and design of a randomised controlled trial in poor ovarian responders undergoing IVF/ICSI treatment. BMJ Open 5, e008297.

Iliodromiti, S., Nelson, S.M., 2015. Ovarian response biomarkers: physiology and performance. Curr. Opin. Obstet. Gynecol. 27, 182–186.

La Marca, A., Sunkara, S.K., 2014. Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: from theory to practice. Hum. Reprod. Update 20, 124–140.

La Marca, A., Sighinolfi, G., Radi, D., Argento, C., Baraldi, E., Artenisio, A.C., Stabile, G., Volpe, A., 2010. Anti-Mullerian hormone (AMH) as a

predictive marker in assisted reproductive technology (ART). Hum. Reprod. Update 16, 113–130.
La Marca, A., Papaleo, E., Grisendi, V., Argento, C., Giulini, S., Volpe, A., 2012a. Development of a nomogram based on markers of ovarian reserve for the individualisation of the follicle-stimulating hormone starting dose in in vitro fertilisation cycles. BJOG. 119, 1171–1179.
La Marca, A., Argento, C., Sighinolfi, G., Grisendi, V., Carbone, M., D'Ippolito, G., Artenisio, A.C., Stabile, G., Volpe, A., 2012b. Possibilities and limits of ovarian reserve testing in ART. Curr. Pharm. Biotechnol. 13, 398–408.
La Marca, A., Grisendi, V., Giulini, S., Argento, C., Tirelli, A., Dondi, G., Papaleo, E., Volpe, A., 2013. Individualization of the FSH starting dose in IVF/ICSI cycles using the antral follicle count. J. Ovarian. Res. 6, 11.
Lan, V.T., Linh, N.K., Tuong, H.M., Wong, P.C., Howles, C.M., 2013. Anti-Müllerian hormone versus antral follicle count for defining the starting dose of FSH. Reprod. Biomed. Online 27, 390–399.
Moolenaar, L.M., Broekmans, F.J., van Disseldorp, J., Fauser, B.C., Eijkemans, M.J., Hompes, P.G., van der Veen, F., Mol, B.W., 2011. Cost effectiveness of ovarian reserve testing in in vitro fertilization: a Markov decision analytic model. Fertil. Steril. 96, 889–894.
Nelson, S.M., 2013. Biomarkers of ovarian response: current and future applications. Fertil. Steril. 99, 963–969.
Nelson, S.M., Yates, R.W., Fleming, R., 2007. Serum anti-Müllerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles–implications for individualization of therapy. Hum. Reprod. 22, 2414–2421.
Nelson, S.M., Yates, R.W., Lyall, H., Jamieson, M., Traynor, I., Gaudoin, M., Mitchell, P., Ambrose, P., Fleming, R., 2009. Anti-Müllerian hormone-based approach to controlled ovarian stimulation for assisted conception. Hum. Reprod. 24, 867–875.
O'Brien, P.C., Fleming, T.R., 1979. A multiple testing procedure for clinical trials. Biometrics 35, 549–556.
Olivennes, F., Howles, C.M., Borini, A., Germond, M., Trew, G., Wikland, M., Zegers-Hochschild, F., Saunders, H., Alam, V., CONSORT study group, 2009. Individualizing FSH dose for assisted reproduction using a novel algorithm: the CONSORT study. Reprod. Biomed. Online 18, 195–204.
Olivennes, F., Trew, G., Borini, A., Broekmans, F., Arriagada, P., Warne, D.W., Howles, C.M., 2015. Randomized, controlled, open-label, non-inferiority study of the CONSORT algorithm for individualized dosing of follitropin alfa. Reprod. Biomed. Online 30, 248–257.
Polyzos, N.P., Sunkara, S.K., 2015. Sub-optimal responders following controlled ovarian stimulation: an overlooked group? Hum. Reprod. 30, 2005–2008.
Popovic-Todorovic, B., Loft, A., Lindhard, A., Bangsbøll, S., Andersson, A.M., Andersen, A.N., 2003a. A prospective study of predictive factors of ovarian response in 'standard' IVF/ICSI patients treated with recombinant FSH. A suggestion for a recombinant FSH dosage normogram. Hum. Reprod. 18, 781–787.
Popovic-Todorovic, B., Loft, A., Bredkjaeer, H.E., Bangsbøll, S., Nielsen, I.K., Andersen, A.N., 2003b. A prospective randomized clinical trial comparing an individual dose of recombinant FSH based on predictive factors versus a 'standard' dose of 150 IU/day in 'standard' patients undergoing IVF/ICSI treatment. Hum. Reprod. 18, 2275–2282.
Schulz, K.G., Grimes, D.A., 2005. Multiplicity in randomised trials II: subgroup and interim analyses. Lancet 365, 1657–1661.
Sunkara, S.K., Rittenberg, V., Raine-Fenning, N., Bhattacharya, S., Zamora, J., Coomarasamy, A., 2011. Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. Hum. Reprod. 26, 1768–1774.
Sunkara, S.K., La Marca, A., Seed, P.T., Khalaf, Y., 2015. Increased risk of preterm birth and low birth weight with very high number of oocytes following IVF: an analysis of 65 868 singleton live birth outcomes. Hum. Reprod. 30, 1473–1480.
Toftager, M., Bogstad, J., Bryndorf, T., Løssl, K., Roskær, J., Holland, T., Prætorius, L., Zedeler, A., Nilas, L., Pinborg, A., 2016. Risk of severe ovarian hyperstimulation syndrome in GnRH antagonist versus GnRH agonist protocol: RCT including 1050 first IVF/ICSI cycles. Hum. Reprod. 31, 1253–1264.
Verberg, M.F., Eijkemans, M.J., Macklon, N.S., Heijnen, E.M., Baart, E.B., Hohmann, F.P., Fauser, B.C., Broekmans, F.J., 2009. The clinical significance of the retrieval of a low number of oocytes following mild ovarian stimulation for IVF: a meta-analysis. Hum. Reprod. Update 15, 5–12.
Volpes, A., Sammartano, F., Coffaro, F., Mistretta, V., Scaglione, P., Allegra, A., 2004. Number of good quality embryos on day 3 is predictive for both pregnancy and implantation rates in in vitro fertilization/intracytoplasmic sperm injection cycles. Fertil. Steril. 82, 1330–1336.
Yates, A.P., Rustamov, O., Roberts, S.A., Lim, H.Y., Pemberton, P.W., Smith, A., Nardo, L.G., 2011. Anti-Mullerian hormone-tailored stimulation protocols improve outcomes whilst reducing adverse effects and costs of IVF. Hum. Reprod. 26, 2353–2362.