Human Reproduction, Vol.32, No.4 pp. 811–819, 2017
Advanced Access publication on February 8, 2017 doi:10.1093/humrep/dex012

human reproduction

ORIGINAL ARTICLE *Infertility*

The addition of anti-Müllerian hormone in an algorithm for individualized hormone dosage did not improve the prediction of ovarian response—a randomized, controlled trial

Å. Magnusson[1,*], L. Nilsson[1], G. Oleröd[2], A. Thurin-Kjellberg[1], and C. Bergh[1]

[1]Department of Obstetrics and Gynaecology, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, Reproductive Medicine, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden [2]Department of Clinical Chemistry, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden

*Correspondence address. Department of Obstetrics and Gynaecology, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, Reproductive Medicine, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden. E-mail: asa.magnusson@vgregion.se

Submitted on September 28, 2016; resubmitted on December 23, 2016; accepted on January 11, 2017

STUDY QUESTION: Does the addition of anti-Müllerian hormone (AMH) to a conventional dosage regimen, including age, antral follicle count (AFC) and BMI, improve the rate of targeted ovarian response, defined as 5–12 oocytes after IVF?

SUMMARY ANSWER: The addition of AMH did not alter the rate of targeted ovarian response, 5–12 oocytes, or decreased the rate of ovarian hyperstimulation syndrome (OHSS) or cancelled cycles due to poor ovarian response.

WHAT IS KNOWN ALREADY: Controlled ovarian hyperstimulation (COH) in connection with IVF is sometimes associated with poor ovarian response resulting in low pregnancy and live birth rates or leading to cycle cancellations, but also associated with excessive ovarian response, causing an increased risk of OHSS. Even though it is well-established that both AMH and AFC are strong predictors of ovarian response in IVF, few randomized trials have investigated their impact on achieving an optimal number of oocytes.

STUDY DESIGN, SIZE AND DURATION: Between January 2013 and May 2016, 308 patients starting their first IVF treatment were randomly assigned, using a computerized randomization program with concealed allocation of patients and in the proportions of 1:1, to one of two dosage algorithms for decisions on hormone starting dose, an algorithm, including AMH, AFC, age and BMI (intervention group), or an algorithm, including only AFC, age and BMI (control group). The study was blinded to patients and treating physicians.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Women aged >18 and <40 years, with a BMI above 18.0 and below 35.0 kg/m^2 starting their first IVF cycle where standard IVF was planned, were eligible. All patients were treated with a GnRH agonist protocol and recombinant FSH was used for stimulation. The study was performed as a single-centre study at a large IVF unit at a university hospital.

MAIN RESULT AND THE ROLE OF CHANCE: The rate of patients having the targeted number of oocytes retrieved was 81/152 (53.3%) in the intervention group versus 96/155 (61.9%) in the control group ($P = 0.16$, difference: −8.6, 95% CI: −20.3; 3.0). Cycles with poor response (<5 oocytes) were more frequent in the AMH group, 39/152 (25.7%) versus the non-AMH group, 17/155 (11.0%) ($P < 0.01$), while the number of cancelled cycles due to poor ovarian response did not differ 7/152 (4.6%) and 4/155 (2.6%) ($P = 0.52$). An excessive response (>12 oocytes) was seen in 32/152 (21.1%) and 42/155 (27.1%) patients, respectively ($P = 0.27$). Moderate or severe OHSS was observed among 5/152 (3.3%) and 6/155 (3.9%) patients, respectively ($P = 1.0$).

Live birth rates were 48/152 (31.6%) and 42/155 (27.1%) per started cycle.

LIMITATIONS, REASONS FOR CAUTION: The categorization of AMH values in predicted low, normal and high responders was originally established using the Diagnostic Systems Laboratories assay and was translated to more recently released assays, lacking international standards and well-established reference intervals. The interpretation of AMH values between different assays should therefore be made with some caution.

WIDER IMPLICATIONS OF THE FINDINGS: An individualised dosage regimen including AMH compared with a non-AMH dosage regimen in an unselected patient population did not alter the number of women achieving the targeted number of oocytes, or the cancellation rate due to poor response or the occurrence of moderate/severe OHSS.

However, this study cannot answer the question if using an algorithm for dose decision of FSH is superior to a standard dose and neither which ovarian reserve test is the most effective.

STUDY FUNDING/COMPETING INTEREST: Financial support was received through Sahlgrenska University Hospital (ALFGBG-70 940) and unrestricted grants from Ferring Pharmaceuticals and the Hjalmar Svensson Research Foundation. None of the authors declares any conflict of interest.

TRIAL REGISTRATION: The study was registered at www.clinicaltrials.gov NCT02013973.

TRIAL REGISTRATION DATE: 6 December 2013.

DATE OF FIRST PATIENT RANDOMIZED: 14 January 2013.

Key words: anti-Müllerian hormone / individual / dose of FSH / algorithm / ovarian / response / prediction

Introduction

It is well known that controlled ovarian hyperstimulation (COH) in connection with IVF is sometimes associated with poor ovarian response, resulting in low pregnancy and live birth rates or leading to cycle cancellations but also associated with excessive ovarian response, causing an increased risk of ovarian hyperstimulation syndrome (OHSS) (van der Gaast *et al.*, 2006; Sunkara *et al.*, 2011; Ji *et al.*, 2013; Steward *et al.*, 2014).

In COH, it is thus important to optimize the number of oocytes. In addition to OHSS, excessive response has been associated with a negative impact on embryo quality (Baart *et al.*, 2007) and endometrial receptivity (Valbuena *et al.*, 1999).

The optimal number of oocytes for live birth has been found to be between 5– and 15 oocytes (van der Gaast *et al.*, 2006; McAvey *et al.*, 2011; Ji *et al.*, 2013). We chose 5–12 oocytes as a balance between efficacy and safety when deciding about number of optimal oocytes retrieved (Chen *et al.*, 2015).

Initially, factors for the prediction of ovarian response after COH, and identified in observational studies, included antral follicle count (AFC), total ovarian volume, Doppler score, age and smoking (Popovic-Todorovic *et al.*, 2003a) and, in another study, FSH, BMI, age and AFC (Howles *et al.*, 2006). These predictors have been used to create algorithms and have been tested in randomized controlled trials (RCT) (Popovic-Todorovic *et al.*, 2003b; Olivennes *et al.*, 2015). Popovic-Todorovic *et al.* found that an individualised dosage algorithm, compared with a standard dose, resulted in a more optimal ovarian response. In the trial by Olivennes *et al.*, the dosage algorithm compared with a standard dose, resulted in significantly fewer oocytes while the rates of poor and excessive ovarian response, did not differ significantly.

Subsequent studies have focused on anti-Müllerian hormone (AMH) and AFC as predictors of ovarian response during COH and a strong correlation between the two biomarkers has been described (Brodin *et al.*, 2015). It is now well-established from numerous observational studies (Nelson *et al.*, 2007, 2009, 2015; Brodin *et al.*, 2015) and systematic reviews (La Marca *et al.*, 2010; Broer *et al.*, 2013; La Marca and Sunkara, 2014; Iliodromiti *et al.*, 2015) that both AMH and AFC are good predictors of ovarian response.

Despite this very strong support for AMH as a predictor of ovarian response, no RCT evaluating AMH for individualised hormonal stimulation in unselected patients, in order to obtain an optimal ovarian response, have been published.

The aim of this RCT was to investigate whether the addition of AMH to a conventional dosage regimen, including AFC, age and BMI, would increase the number of patients receiving the targeted number of oocytes retrieved during IVF and thereby reduce poor and excessive response. In addition, we wanted to compare live birth rates.

Materials and Methods

Patients

This was a prospective, randomized, controlled, single-centre study conducted at the Department of Reproductive Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden. Patients were recruited between January 2013 and May 2016. The study was approved by the regional ethics committee at Gothenburg University and all patients signed informed consent.

The inclusion criteria were women aged >18 and <40 years, with a BMI above 18.0 and below 35.0 kg/m^2, having their first standard IVF planned and AMH not previously measured, thus representing the general patient population at our university hospital clinic. The exclusion criteria were male factor infertility where ICSI was planned, cycles planned for oocyte donation or PGD. Only one cycle per patient was included.

Procedures

From infertile women, agreeing to participate and signing informed consent, a blood sample for AMH was taken and immediately stored at −70°C.

Patients were downregulated with a GnRH agonist (Suprecur, Sanofi AB, France) in a long protocol (estradiol (E2) <200 pmol/L).

On Day 2–4 of menstrual bleeding, and after downregulation was achieved, a vaginal sonography was performed to estimate AFC using a standardized method for assessing AFC, suggested by Broekmans (Broekmans *et al.*, 2010), using a GE Voluson Logic P six sonograph with an E eight CS vaginal probe. AFC was defined as the total number of follicles 2–10 mm in both ovaries. In order to minimize inter-observer variation, only two physicians participated in the AFC procedure (Å.M. and A.T.K.). On the same visit, patients were randomized to one of two treatment groups by the study nurse; either a dose-decision algorithm group, including age, BMI, AFC and serum AMH (AMH group), or a conventional dose-decision algorithm group, including only age, BMI and AFC (non-AMH group). Randomization was performed with a computerized randomization program with concealed allocation of patients and in the proportions of 1:1. Optimal allocation was applied according to Pocock's minimization technique for sequential randomization (Pocock, 1983), taking account of the age of the woman, duration of infertility, parity, polycystic ovary syndrome, smoking, BMI and AFC. The study was blinded to patients, physicians managing patients during IVF treatment and the statistician. For practical reasons, the two physicians performing the AFC estimations were unblinded to the starting dose of recombinant FSH (rFSH).

If the patient was randomized to the AMH group, her blood sample was thawed and analysed for AMH using the Beckman Coulter AMH Gen II ELISA (Chaska, MN, USA). In late June 2013, Beckman Coulter announced that the assay might not be reliable due to complement interference. By then, blood samples of 26 patients had been analysed. Since complement interference appears to be a problem in serum samples stored at room temperature but not in frozen samples (Welsh *et al.*, 2014), we regarded the AMH values of these first 26 samples as reliable. Later in 2013, a new modified Beckman Coulter Gen II ELISA, where the complement interference was eliminated, was released and considered stable. This modified assay was used for all the other patients in the study. The total coefficient of variance (total CV) for serum AMH was 10.6% at the level of 2.6 ng/mL, 4.9% at the level of 4.9 ng/mL and 10.2% at the level of 8 ng/mL. The lower limit of quantification was 0.2 ng/mL.

A large cohort study (Nelson *et al.*, 2009) suggested that the level of AMH in serum can be used to classify infertile women having COH and IVF into three different categories, high responders, normo-responders and low responders. We used these categories, after translation from Diagnostic Systems Laboratories (Nelson and La Marca, 2011; Wallace *et al.*, 2011), as part of an algorithm for an individualised dosage regimen of rFSH. The reference intervals for the categories were revised for the modified Beckman Coulter Gen II assay (personal communication Professor Richard Fleming).

Patients were stimulated with an rFSH ((Gonal-F, Merck, Germany) or (Puregon, MSD, USA)) starting dose of 75, 100, 150, 225 or 300 IU, according to the algorithm and used for the non-AMH group and the AMH group assessed as having normal AMH values (1.55–2.95 ng/mL). For the group having high AMH (>2.95 ng/mL), the rFSH was decreased by one step and, for the group having low (<1.55 ng/mL), the dose was increased one step (Supplementary Table SI). The lowest starting dose was 75 IE and the highest 300 IE.

Hormonal stimulation was monitored by E2 and vaginal sonography. Dose adjustment was allowed at the earliest on Day 7 and only in predefined steps and if E2 on stimulation day six was either <350 or >1.500 pmol/L. These cut-off limits were arbitrarily chosen from clinical experience.

Vaginal sonography was performed on stimulation Day 9–11 and the number of follicles of ≥11 mm was recorded. Further dose adjustment was made if >15 follicles of >11 mm were observed. Coasting was allowed if the physician considered there to be a high risk of OHSS. Ovulation was triggered by 6500 IE recombinant human chorion gonadotrophin (Ovitrelle, Merck), when ≥2 follicles of ≥17 mm were detected at ultrasound. Oocyte retrieval, fertilization and culture followed standard techniques. In case of a low sperm count ICSI was allowed.

Single embryo transfer (SET) was performed on Day 2 or 3. In the event of no good-quality embryos, double-embryo transfer was allowed. Luteal phase support with vaginally administered progesterone (Lutinus, Ferring, Switzerland) was given from the day of oocyte retrieval and for 15 days until a pregnancy test (urinary hCG).

In the event of impending OHSS, where a decision was made to cryopreserve all embryos, the outcome of the first transfer of a cryopreserved embryo was included in the study.

For patients achieving a pregnancy, a vaginal sonography was performed in pregnancy Week 7–8. A live birth was defined as a delivery with at least one live-born child regardless of gestational age.

The patients were asked to contact the IVF unit at any time after oocyte retrieval if they experienced subjective symptoms of OHSS. Classification of OHSS was done according to Navot *et al.* (1992) and Golan *et al.* (1989), using strict criteria for moderate and severe OHSS.

The primary outcome of the trial was the number of patients achieving the targeted number of oocytes, defined as 5–12 oocytes. Secondary outcomes included the number of patients with a poor and excessive response (<5 oocytes and >12 oocytes, respectively), cancellation rate due to poor or excessive ovarian response, number of patients with moderate and severe OHSS, number of patients with OHSS-prevention strategies (coasting, cryopreservation of all embryos), number of good-quality embryos, number of cryopreserved embryos, total dose of gonadotrophins, pregnancies and live births.

Statistical analysis and calculation of sample size

An analysis at Sahlgrenska University Hospital of all IVF cycles in 2010–2011 revealed that, in 63% of the cycles, 5–12 oocytes were obtained, in 16% <5 oocytes and in 21% >12 oocytes were obtained at oocyte retrieval. We calculated that 138 patients per group would be required to detect a 15% increase (from 65 to 80%) in patients receiving the targeted number of oocytes with two-sided Fisher's exact test with a power of 80% and significance level 5%. We aimed to recruit 150 patients per group.

The main analyses were performed on the Full Analyses Set (FAS) and complementary analyses were performed on the Per Protocol (PP) set. The FAS consisted of all randomized patients who had at least one follow-up variable (Fig. 1). The patient was evaluated in the group to which she was randomized, regardless of the treatment she received, or whether or not she completed the intervention. A PP analysis, taking the actual treatment given and the compliance with intervention into consideration, was also performed. The PP population consisted of all randomized subjects who completed the study according to pre-defined rules.

The main analyses between the two groups were performed unadjusted with Fisher's exact test for dichotomous variables, the Mantel–Haenszel chi-square test for ordered categorical variables, Fisher's non-parametric permutation test for continuous variables and Pearson's chi-square test for non-ordered categorical variables. For dichotomous outcome variables, 95% CIs were calculated for the difference in proportions. Complementary analyses for primary and selected secondary efficacy variables were performed on the FAS adjusted for the minimization variables (data not shown).

Continuous variables are described with the mean, SD, median, minimum and maximum and categorical variables with number and percentages. All significance tests were two-sided and conducted at the 5% significance level. For correlation analysis between AMH and AFC Spearman's rank correlation coefficient (r_s) was used. For analysis of inter-observer agreement regarding AFC the distribution of the difference between the raters, limits of agreement and intraclass correlation coefficient

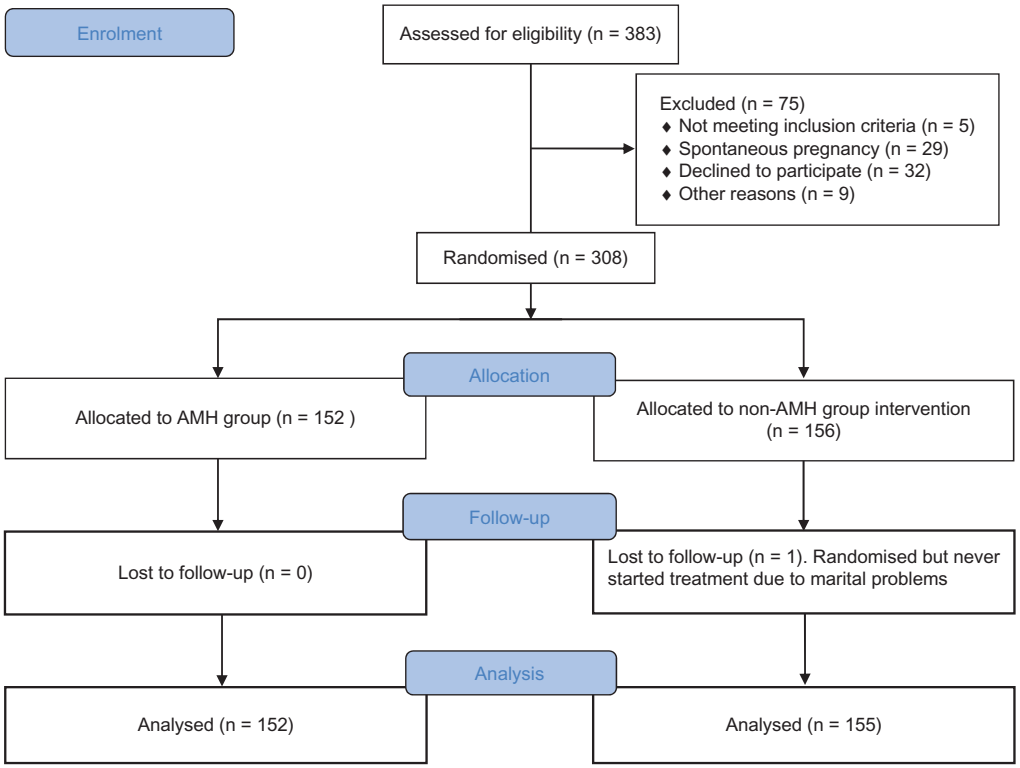

Figure 1 Flow chart of a randomized trial comparing two dose algorithms for starting dose of FSH in IVF. AMH, anti-Müllerian hormone.

(ICC) were given. Wilcoxon Signed Rank test was used to analyse systematic differences between the two raters.

Results

A total of 383 patients were assessed for eligibility. The reasons for exclusion are listed in Fig. 1.

Overall, 308 patients were randomized and 307 patients had at least one follow-up measurement. One patient was randomized but never started treatment because of marital problems and she was excluded from the FAS population.

All the other 307 patients were included in the FAS population. We also performed a PP analysis, excluding 30 patients where dose adjustment did not follow the study protocol (25 patients) or where patients received the wrong starting dose (five patients). The PP analysis thus comprised 277 patients.

No significant differences were noted in baseline patient characteristics between groups (Table I).

The treatment outcomes are summarized in Table II. There was no significant difference between the groups regarding the primary efficacy variable, rate of patients with 5–12 oocytes 81/152 (53.3%) versus 96/155 (61.9%) ($P = 0.16$, difference: −8.6, 95% CI: −20.3; 3.0). Similar results were seen in the PP analysis.

The number of cancelled cycles due to poor ovarian response was 7/152 (4.6%) and 4/155 (2.6%), respectively ($P = 0.52$), while no cycles were cancelled due to an excessive response. Significantly more cycles developed a poor response (<5 oocytes) in the AMH group, 39/152 (25.7%) versus the non-AMH group, 17/155 (11.0%) ($P < 0.01$).

An excessive response (>12 oocytes) was achieved in 32/152 (21.1%) and 42/155 (27.1%), respectively ($P = 0.27$). Moderate or severe OHSS was observed among 5/152 (3.3%) and 6/155 (3.9%) patients, respectively ($P = 1.0$) and OHSS-prevention strategies were implemented in 6/152 (3.9%) and 12/155 (7.7%) in the AMH and non-AMH groups ($P = 0.24$).

Live birth rate was 48/152 (31.6%) and 42/155 (27.1%) per started cycle and 35.3 and 30.4% per embryo transfer.

More than 95% of the patients in this study had a SET.

When dividing the patients according to the starting dose, low dose (75–100 IU), normal dose (150 IU) or high dose (225–300 IU), no significant differences between AMH and non-AMH groups in the distribution of oocytes were noted (Fig. 2).

A correlation analysis between AMH and AFC showed a significant correlation, $r_s = 0.70$ ($P < 0.0001$).

Analysis of the inter-observer agreement regarding AFC showed a high ICC = 0.94. The mean difference between the raters was −2.30 (SD = 5.50, range: −17; 3), limits of agreement (−13.1 to 8.48). No significant difference ($P = 0.17$) in AFC was found between the two raters.

In a post hoc analysis of all patients in the AMH group we investigated how many patients that would have had a different starting dose if AMH had not been included in the algorithm. In total 69/152 (45%) patients would have had a higher dose, 52/152 (34%) the same dose and 26/152 (17%) a lower dose. Thus a large proportion of patients

Table I Demographics and baseline characteristics by randomized group (FAS population).

Variable	AMH (n = 152)	No AMH (n = 155)
Age at first dose (years)	32.3 (4.0)	32.3 (3.8)
	32.4 (21.4; 39.3)	32.5 (20.2; 39.3)
	n = 152	n = 155
Duration of infertility (months)	32.4 (14.7)	32.0 (14.9)
	30 (1; 120)	30 (0; 108)
	n = 152	n = 155
Reason for infertility		
Other	2 (1.3%)	4 (2.6%)
Tubal factor	27 (17.8%)	23 (14.8%)
Endometriosis	11 (7.2%)	14 (9.0%)
Hormonal	10 (6.6%)	17 (11.0%)
PCOS	10 (6.6%)	9 (5.8%)
Male factor	2 (1.3%)	1 (0.6%)
Unexplained infertility	90 (59.2%)	87 (56.1%)
Parity		
0	40 (87.0%)	43 (86.0%)
1	2 (4.3%)	5 (10.0%)
2	3 (6.5%)	2 (4.0%)
3	1 (2.2%)	0 (0.0%)
PCO		
No	127 (83.6%)	126 (81.3%)
Yes	25 (16.4%)	29 (18.7%)
Smoker		
No	143 (94.1%)	147 (94.8%)
Yes	9 (5.9%)	8 (5.2%)
BMI (kg/m^2)	23.6 (3.7)	23.5 (3.6)
	22.8 (18.1; 35.1)	22.9 (18.0; 35.0)
	n = 152	n = 155
Ethnicity		
Other	1 (0.7%)	5 (3.2%)
Caucasian	144 (94.7%)	140 (90.3%)
African	2 (1.3%)	1 (0.6%)
Asian	5 (3.3%)	9 (5.8%)
AFC	21.6 (12.0)	21.3 (11.3)
	19 (3; 73)	18 (6; 70)
	n = 152	n = 155
AMH	4.03 (3.53)	
	2.95 (0.20; 18.20)	
	n = 148	
AMH		
Low < 1.55 ng/ml	36 (24.3%)	
Normal 1.55–2.95 ng/ml	38 (25.7%)	
High > 2.95 ng/ml	74 (50.0%)	

For categorical variables, n (%) is presented.
For continuous variables, the mean (SD)/median (min; max)/n= is presented.
For comparisons between groups, Fisher's exact test (lowest one-sided P-value multiplied by 2) was used for dichotomous variables, the Mantel–Haenszel chi-square test was used for ordered categorical variables, the chi-square test was used for non-ordered categorical variables and Fisher's non-parametric permutation test was used for continuous variables.
No significant differences between groups in any variable.
PCOS, polycystic ovary syndrome; AFC, antral follicle count; AMH, Anti-Müllerian hormone.

Table II Efficacy analyses by randomized group (FAS population).

Variable	AMH (n = 152)	No AMH (n = 155)	P-value	Difference between groups mean (95% CI)
5–12 oocytes retrieved	81 (53.3%)	96 (61.9%)	0.16	−8.6 (−20.3; 3.0)
<5 oocytes retrieved	39 (25.7%)	17 (11.0%)	0.0013	14.7 (5.5; 23.9)
>12 oocytes retrieved	32 (21.1%)	42 (27.1%)	0.27	−6.0 (−16.2; 4.1)
Cancelled cycles due to poor ovarian response	7 (4.6%)	4 (2.6%)	0.52	2.0 (−2.8; 6.8)
Cancelled cycles due to excessive response	0 (0.0%)	0 (0.0%)	1.00	
Cancelled cycles other reasons	0 (0.0%)	0 (0.0%)	1.00	
Severe or moderate OHSS	5 (3.3%)	6 (3.9%)	1.00	−0.6 (−5.4; 4.2)
OHSS-prevention strategies (coasting and/or freezing of all embryos)	6 (3.9%)	12 (7.7%)	0.24	−3.8 (−9.7; 2.1)
Coasting	3 (2.1%)	7 (4.6%)	0.37	−2.6 (−7.3; 2.2)
Freezing of all embryos due to excessive response	3 (2.0%)	6 (3.9%)	0.52	−1.9 (−6.3; 2.5)
Severe or moderate OHSS/OHSS and/or prevention strategies	9 (5.9%)	16 (10.3%)	0.23	−4.4 (−11.1; 2.3)
Dose adjustment at any time during stimulation	83 (54.6%)	81 (52.3%)	0.77	2.3 (−9.5; 14.2)
Total dose of gonadotropins	1685 (997)	1604 (701)	0.41	80.9 (−112.1; 273.3)
	1383 (600; 6000)	1450 (675; 4650)		
	n = 152	n = 155		
Follicles >12 mm at 0–2 days before hCG	9.4 (5.04)	10.8 (6.0)	0.022	−1.46 (−2.71; −0.22)
	9 (0; 30)	10 (1; 31)		
	n = 152	n = 155		
Follicles punctured at oocyte retrieval	12.1 (6.6)	13.8 (6.4)	0.034	−1.62 (−3.10; −0.13)
	11 (1; 40)	12 (2; 35)		
	n = 145	n = 151		
Number of oocytes retrieved (0 included)	8.9 (6.00)	10.0 (5.17)	0.085	−1.11 (−2.37; 0.15)
	8 (0; 32)	10 (0; 30)		
	n = 152	n = 155		
Fertilization rate	0.58 (0.232)	0.60 (0.237)	0.31	−0.028 (−0.083; 0.026)
	0.60 (0.0; 1.0)	0.63 (0.0; 1.0)		
	n = 140	n = 144		
Number of good quality embryos	2.7 (2.69)	2.9 (2.27)	0.67	−0.13 (−0.72; 0.46)
	2 (0; 14)	3 (0; 11)		
	n = 139	n = 141		
Number of transferred embryos				
1	126 (95.5%)	128 (98.5%)		−3.0 (−7.9; 1.9)
2	6 (4.5%)	2 (1.5%)	0.29	3.0 (−1.9; 7.9)
Day of transfer				
2	132 (100.0%)	129 (99.2%)		0.8 (−1.5; 3.0)
3	0 (0.0%)	1 (0.8%)	0.99	−0.8 (−3.0; 1.5)
Number of frozen embryos day 2 and/or day 5	1.9 (2.51)	2.2 (2.10)	0.35	−0.27 (−0.81; 0.28)
	1 (0; 12)	2 (0; 10)		
	n = 138	n = 140		
Biochemical pregnancy	64 (46.7%)	62 (44.6%)	0.82	2.1 (−10.4; 14.6)
Live birth rate (per patient)*	48 (31.6%)	42 (27.1%)	0.46	4.5 (−6.3; 15.3)
Live birth rate (per embryo transfer)*	48 (35.3%)	42 (30.4%)	0.47	4.9 (−7.0; 16.7)
Miscarriage rate	16 (25.0%)	20 (32.3%)	0.48	−7.3 (−24.6; 10.1)

For categorical variables, n (%) is presented.
For continuous variables, the mean (SD)/median (min; max)/n = is presented.
For comparisons between groups, Fisher's exact test (lowest one-sided P-value multiplied by 2) was used for dichotomous variables and Fisher's non-parametric permutation test was used for continuous variables.
*All women have given birth except one woman (due 19 Feb 2017).
OHSS, ovarian hyperstimulation syndrome.

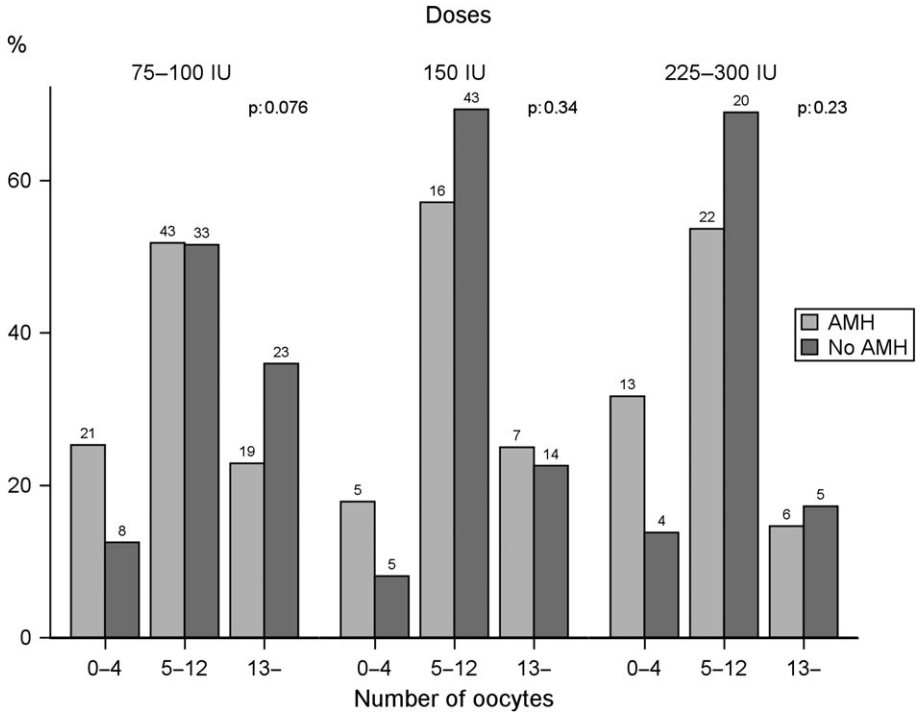

Figure 2 Number of oocytes by start dose and AMH/No AMH. Percentage of patients is shown on the *y*-axis and number of patients is shown above each bar. The *P*-values refer to significance tests in distribution of oocytes between AMH and non-AMH groups.

would have had a different dose, mainly a higher dose, without AMH in dose decision.

Beside the 11 cases of OHSS, six severe adverse events were registered (Supplementary Table SII). All patients were completely recovered.

Discussion

This randomized trial, comparing an individual dosage regimen for rFSH including AMH versus a non-AMH dosage regimen in an unselected patient population, showed that AMH did not alter the number of patients achieving the targeted ovarian response, defined as 5–12 oocytes. Significantly more patients achieved a low ovarian response in the AMH group, while the number of patients with cancelled cycles due to a poor response did not differ. There were no significant differences in the rates of OHSS or OHSS-prevention strategies between groups.

When dividing the patients according to starting dose, no significant difference in the distribution of oocytes was found between the randomized groups, indicating that the failure of AMH to improve the distribution of oocytes was a general phenomenon and not only observed for a certain category of patients.

These results might be unexpected in view of the strong support during the last decade for AMH in predicting ovarian response during COH (La Marca *et al.*, 2010; Broer *et al.*, 2013; La Marca and Sunkara, 2014; Iliodromiti *et al.*, 2015). Among published randomized studies, where an algorithm for dosage has been used compared with a fixed standard dose (Popovic-Todorovic *et al.*, 2003b; Olivennes *et al.*, 2015) or where two different doses have been compared in a population defined by ovarian reserve markers (Lefebvre *et al.*, 2015; Jayaprakasan *et al.*, 2010, Klinkert *et al.*, 2005), only the study by Popovic-Todorovic showed an advantage from individualised dosage.

In a very recent study (Nyboe-Andersen and Arce, 2016), published as an abstract at ESHRE 2016, a large group of women were randomized to either a group receiving treatment based on an individualised dosage algorithm, including AMH and body weight and being given a newly developed human rFSH (follitropin delta), or a group using a standard dose and follitropin alfa. Similar ongoing pregnancy rates were noted, while the individualised dosage regimen yielded more patients with an appropriate ovarian response and a reduced OHSS risk.

In another large trial (van Tilborg *et al.*, 2012) and also published at ESHRE 2016 (van Tilborg *et al.*, 2016), including patients predicted as poor responders according to AFC, individualised gonadotropin dosage did not improve live birth rates compared with standard treatment, even though a lower cancellation rate depending on poor ovarian response was noted. In the same trial, studying patients predicted as high responders (Oudshoorn *et al.*, 2016), individualised dosage did not improve live birth rates. A higher cancellation rate was noted in the individualised group, plus lower rates of mild and moderate OHSS, while the occurrence of severe OHSS was unaffected.

The choice of control group is extremely important in all randomized trials. A demonstration of superiority for an intervention is usually interpreted as an effect of that particular intervention, but it might also be an effect of inferior performance in the control group. In the present study, we chose a control group using an individualised dosage algorithm including AFC. We think this more accurately reflects usual

care than a control group with standard dosage for ovarian stimulation. AFC and age are commonly used in clinical practice today for dosage of rFSH and this is well supported in the literature (Hsu *et al.*, 2011; Broekmans *et al.*, 2006). The use of AFC has naturally also been facilitated by the rapid development of ultrasound equipment. The choice of this kind of control group was of particular interest in the present study, performed in an unselected population, including poor and high responders, as well as PCO patients. When evaluating predictors of ovarian reserve, the strength of AMH includes cycle- and investigator independence and stability, while limitations include the presence of different assays, translational issues between assays and lack of international standards on cut off levels. The strength of AFC is that it is usually included in gynaecological and sonographic evaluation before IVF is performed and thus without additional costs. The main weakness is the inter-observer variation.

In addition to the randomized design, the strength of the present study is the unselected population, which better reflects usual care, giving high generalizability to the study.

The main limitation is that the study was not powered for live birth.

In conclusion, this randomized trial, comparing an individual dosage regimen for rFSH including AMH versus an individual non-AMH dosage regimen in an unselected patient population, did not show any benefit for the intervention group in oocyte distribution, or in the cancellation rate due to poor response or the occurrence of moderate/severe OHSS. A high live birth rate was achieved in both groups after almost exclusive use of SET.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

The authors wish to thank Nils Gunnar Pehrsson and Mattias Molin, Statistical Consulting Group, Gothenburg, for valuable statistical support Christina Larsson, study nurse, for administrative work Niklas Svensson, Sharp Consulting, Gothenburg, for preparing the electronic case report form and the randomization program Professor Richard Fleming, Glasgow Centre for Reproductive Medicine, for valuable advice on defining reference values for the new modified Beckman Coulter Gen II ELISA Ferring Pharmaceuticals for support with an unrestricted grant.

Authors' roles

All the authors contributed to the conception and design of the study and to the analysis and interpretation of data. G.O. performed the AMH assays and contributed the technical data on the AMH assays. Å.M. drafted the article. Å.M. and C.B. finalized it and all the authors revised the article and approved the final version.

Funding

Financial support was received through Sahlgrenska University Hospital (ALFGBG-70 940), grants from Ferring Pharmaceuticals and the Hjalmar Svensson Research Foundation.

Conflict of interest

None declared.

References

Baart EB, Martini E, Eijkemans MJ, Van Opstal D, Beckers NGM, Verhoeff A, Macklon NS, Fauser BC. Milder ovarian stimulation for in-vitro fertilization reduces aneuploidy in the human preimplantation embryo: a randomized controlled trial. *Hum Reprod* 2007;**22**:980–988.

Brodin T, Hadziosmanovic N, Berglund L, Olofsson M, Holte J. Comparing four ovarian reserve markers-associations with ovarian response and live births after assisted reproduction. *Acta Obstet Gynecol Scand* 2015; **94**:1056–1063.

Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. *Hum Reprod Update* 2006;**12**:685–718.

Broekmans FJM, de Ziegler D, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: practical recommendations for better standardization. *Fertil Steril* 2010;**94**:1044–1051.

Broer SL, Dólleman M, van Disseldorp J, Broeze KA, Opmeer BC, Bossuyt PMM, Eijkemans MJC, Mol BW, Broekmans FJM, on behalf of the IPD-EXPORT Study group. Prediction of an excessive response in in vitro fertilization from patient characteristics and ovarian reserve test and comparison in subgroups: an individual patient data meta-analysis. *Fertil Steril* 2013;**100**:420–429.

Chen YH, Xu XH, Wang Q, Zhang SD, Jiang LL, Zhang CL, Ge ZJ. Optimum oocyte retrieved and transfer strategy in young women with normal ovarian reserve undergoing a long treatment protocol: a retrospective cohort study. *J Ass Reprod Genet* 2015;**32**:1459–1467.

Golan A, Ron-El R, Herman A, Soffer Y, Weinraum Z, Caspi E. Ovarian hyperstimulation syndrome: an update review. *Obstet Gynecol Surv* 1989; **44**:430–440.

Howles CM, Saunders H, Alam V, Engrand P, FSH Treatment Guidelines Clinical Panel. 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* 2006;**22**:907–918.

Hsu A, Arny M, Knee AB, Bell C, Cook E, Novak AL, Grow DR. Antral follicle count in clinical practice: analyzing clinical relevance. *Fertil Steril* 2011;**95**:474–479.

Iliodromiti S, Anderson RA, Nelson SM. Technical and performance characteristics of anti-Mullerian hormone and antral follicle count as biomarkers of ovarian response. *Hum Reprod Update* 2015;**21**: 698–710.

Jayaprakasan K, Hopkisson J, Campbell B, Johnson I, Thornton J, Raine-Fenning N. A randomised controlled trial of 300 versus 225 IU recombinant FSH for ovarian stimulation in predicted normal responders by antral follicle count. *BJOG* 2010;**117**:853–862.

Ji J, Liu Y, Tong XH, Luo L, Ma J, Chen Z. The optimum number of oocytes in IVF treatment: an analysis of 2455 cycles in China. *Hum Reprod* 2013; **28**:2728–2734.

Klinkert ER, Broekmans FJ, Looman CW, Habbema JD, Te Velde ER. Expected poor responders on the basis of an antral follicle count do not benefit from a higher starting dose of gonadotrophins in IVF treatment: a randomized controlled trial. *Hum Reprod* 2005;**20**:611–615.

La Marca A, Sighinolfi G, Radi D, Argento C, Baraldi E, Carducci Artenisio A, Stabile G, Volpe A. Anti-Müllerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). *Hum Reprod Update* 2010;**16**:113–130.

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

Lefebvre J, Antaki R, Kadoch IJ, Dean NL, Sylvestre C, Bissonnette F, Benoit J, Ménard N, Lapensée L. 450 IU versus 600 IU gonadotropin for controlled ovarian stimulation in poor responders: a randomized controlled trial. *Fertil Steril* 2015;**104**:1419–1425.

McAvey B, Athena Zapantis A, Jindal SK, Liema HJ, Polotsky HJ. How many eggs are needed to produce an assisted reproductive technology baby: is more always better? *Fertil Steril* 2011;**96**:332–332.

Navot D, Bergh PA, Laufer N. Ovarian hyperstimulation syndrome in novel reproductive technologies: prevention and treatment. *Fertil Steril* 1992;**58**:249–261.

Nelson SM, Yates RW, Fleming R. Serum anti-Müllerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles—implications for individualization of therapy. *Hum Reprod* 2007; **22**:2414–2421.

Nelson SM, Yates RW, Lyall H, Jamieson M, Traynor I, Gaudoin M, Mitchell P, Ambrose P, Fleming R. Anti-Müllerian hormone-based approach to controlled ovarian stimulation for assisted conception. *Hum Reprod* 2009;**24**:867–875.

Nelson SM, La Marca A. The journey from the old to the new AMH assay: how to avoid getting lost in the values. *RBM Online* 2011;**23**:411–442.

Nelson SM, Klein BM, Arce JC. Comparison of anti-Müllerian hormone levels and antral follicle count as predictor of ovarian response to controlled ovarian stimulation in good-prognosis patients at individual fertility clinics in two multicenter trials. *Hum Reprod* 2015;**103**: 923–931.

Nyboe-Andersen A and Arce JCESTHER-1. Efficacy and safety of follitropin delta in an individualised dosing regimen: a randomised, assessor-blind controlled phase 3 trial in IVF/ICSI patients (ESTHER-1). *Abstract P-434, ESHRE*. Helsinki, 2016.

Olivennes F, Trew G, Borini A, Broekmans F, Arriagada P, Warne DW, Howles CM. Randomized, controlled, open-label, non-inferiority study of the CONSORT algorithm for individualized dosing of follitropin alfa. *RBM Online* 2015;**30**:248–257.

Oudshoorn SC, van Tilborg TC, Eijkemans MJC, Laven JSE, Koks CAM, de Bruin JP, Scheffer, GJ, van Golde RJT, Fleisher K, Hoek A *et al*. Optimization of outcome through dosing in predicted hyper responders undergoing IVF/ICSI; The OPTIMIST randomized controlled trial, NTR2657. *Abstract O-036 ESHRE*. Helsinki, 2016.

Pocock SJ. *Clinical Trials: A Practical Approach*. Chichester, England: John Wiley, 1983.

Popovic-Todorovic B, Loft A, Lindhard A, Bangsbøll S, Andersson AM, Nyboe Andersen A. 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* 2003b;**18**:781–787.

Popovic-Todorovic B, Loft A, Ejdrup Bredkjæer H, Bangsbøll S, Nielsen IK, Nyboe Andersen A. 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* 2003a;**18**:2275–2282.

Steward RG, Lan L, Shah AA, Yeh JS, Price TM, Goldfarb JM, Muasher SJ. Oocyte number as predictor for ovarian hyperstimulation syndrome and live birth: an analysis of 256,381 in vitro fertilization cycles. *Fertil Steril* 2014;**101**:967–973.

Sunkara SH, Rittenberg V, Raine-Fenning N, Bhattacharya S, Zamora J, Coomarasamy A. Association between the number of eggs and live birth in IVF treatment: an analysis of 400135 treatment cycles. *Hum Reprod* 2011;**26**:1768–1774.

Valbuena D, Jasper M, Remohí J, Pellicer A, Simón C. Ovarian stimulation and endometrial receptivity. *Hum Reprod* 1999;**14**:107–111.

van der Gaast MH, Beier-Hellwig K, Fauser BC, Beier HM, Macklon NS. Optimum number of oocytes for a successful first IVF treatment cycle. *Reprod Biomed Online* 2006;**13**:476–480.

van Tilborg TC, Eijkemans MJC, Laven JSE, Koks CAM, de Bruin JP, Scheffer GJ, van Golde RJT, Fleischer K, Hoek A, Nap AP *et al*. The OPTIMIST study: optimisation of cost effectiveness through individualised FSH stimulation dosages for IVF treatment. A randomised controlled trial. *BMC Women's Health* 2012;**12**:29.

van Tilborg TC, Oudshoorn SC, Eijkemans MJC, Manger PAP, Brinkhuis E, van Heusden AM, Kuchenbecker WKH, Smeenk JMJ, Kwee J, Verhoeve HR *et al*. Optimization of outcome through dosing in predicted poor responders undergoing IVF/ICSI; The OPTIMIST randomized controlled trial, NTR2657. *Abstract O-035 ESHRE*. Helsinki, 2016.

Wallace AM, Faye SA, Fleming R, Nelson SM. A multicenter evaluation of the new Beckman Coulter antimullerian hormone immunoassay (AMH-Gen II). *Ann Clin Biochem* 2011;**48**:370–373.

Welsh P, Smith K, Nelson SM. A single-centre evaluation of two new anti-Müllerian hormone assays and comparison with the current clinical standard assay. *Hum Reprod* 2014;**29**:1035–1041.