

Individualized FSH dosing based on ovarian reserve testing in women starting IVF/ICSI: a multicentre trial and cost-effectiveness analysis

Theodora C. van Tilborg^{1,*}, Simone C. Oudshoorn¹,
Marinus J.C. Eijkemans², Monique H. Mochtar³, Ron J.T. van Golde^{4,5},
Annemieke Hoek⁶, Walter K.H. Kuchenbecker⁷, Kathrin Fleischer⁸,
Jan Peter de Bruin⁹, Henk Groen¹⁰, Madelon van Wely³,
Cornelis B. Lambalk¹¹, Joop S.E. Laven¹², Ben Willem J. Mol^{13,14},
Frank J.M. Broekmans¹, and Helen L. Torrance¹, on behalf of the
OPTIMIST study group[†]

¹Department of Reproductive Medicine and Gynaecology, University Medical Centre Utrecht, Utrecht University, PO Box 85500, 3508 GA Utrecht, the Netherlands ²Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, PO Box 85500, 3508 GA Utrecht, the Netherlands ³Centre for Reproductive Medicine, Academic Medical Centre Amsterdam, University of Amsterdam, PO Box 22660, 1100 DD Amsterdam, The Netherlands ⁴Department of Reproductive Medicine, Maastricht University Medical Centre+, PO Box 5800, 6202 AZ Maastricht, The Netherlands ⁵GROW—School for Oncology and Developmental Biology, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands ⁶Centre for Reproductive Medicine, University of Groningen, University Medical Centre Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands ⁷Department of Obstetrics and Gynaecology, Isala Clinics, PO box 10400, 8000GK Zwolle, The Netherlands ⁸Department of Obstetrics and Gynaecology, Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands ⁹Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, PO box 90153, 5200 ME 's-Hertogenbosch, The Netherlands ¹⁰Department of Epidemiology, University of Groningen, University Medical Centre Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands ¹¹Centre for Reproductive Medicine, VU University Medical Centre, PO Box 7057, 1007 MB Amsterdam, the Netherlands ¹²Department of Obstetrics and Gynaecology, Erasmus Medical Centre, Postbus 2040, 3000 CA Rotterdam, The Netherlands ¹³The Robinson Research Institute, School of Paediatrics and Reproductive Health, University of Adelaide, SA 5006 Adelaide, Australia ¹⁴The South Australian Health and Medical Research Unit, PO Box 11060, SA 5001 Adelaide, Australia

*Correspondence address. E-mail: ctilborg@umcutrecht.nl

Submitted on June 10, 2017; resubmitted on August 22, 2017; accepted on October 13, 2017

STUDY QUESTION: Is there a difference in live birth rate and/or cost-effectiveness between antral follicle count (AFC)-based individualized FSH dosing or standard FSH dosing in women starting IVF or ICSI treatment?

SUMMARY ANSWER: In women initiating IVF/ICSI, AFC-based individualized FSH dosing does not improve live birth rates or reduce costs as compared to a standard FSH dose.

WHAT IS KNOWN ALREADY: In IVF or ICSI, ovarian reserve testing is often used to adjust the FSH dose in order to normalize ovarian response and optimize live birth rates. However, no robust evidence for the (cost-)effectiveness of this practice exists.

STUDY DESIGN, SIZE, DURATION: Between May 2011 and May 2014 we performed a multicentre prospective cohort study with two embedded RCTs in women scheduled for IVF/ICSI. Based on the AFC, women entered into one of the two RCTs (RCT1: AFC < 11; RCT2: AFC > 15) or the cohort (AFC 11–15). The primary outcome was ongoing pregnancy achieved within 18 months after randomization

[†]OPTIMIST study group: Theodora C. van Tilborg, Simone C. Oudshoorn, Marinus J.C. Eijkemans, Monique H. Mochtar, Carolien A.M. Koks, Ron J.T. van Golde, Harold R. Verhoeve, Annemiek W. Nap, Gabrielle J. Scheffer, A. Petra Manger, Annemieke Hoek, Bendictus C. Schoot, G. Jur E. Oosterhuis, Walter K.H. Kuchenbecker, Kathrin Fleischer, Jan Peter de Bruin, Alexander V. Sluijmer, Jaap Friederich, Arie Verhoeff, Marcel H.A. van Hooff, Evert J.P. van Santbrink, Egbert A. Brinkhuis, Jesper M.J. Smeenk, Janet Kwee, Corry H. de Koning, Henk Groen, Madelon van Wely, Cornelis B. Lambalk, Joop S.E. Laven, Ben Willem J. Mol, Frank J.M. Broekmans, Helen L. Torrance.

resulting in a live birth (delivery of at least one live foetus after 24 weeks of gestation). Data from the cohort with weight 0.5 were combined with both RCTs in order to conduct a strategy analysis. Potential half-integer numbers were rounded up. Differences in costs and effects between the two treatment strategies were compared by bootstrapping.

PARTICIPANTS/MATERIALS, SETTING, METHODS: In both RCTs women were randomized to an individualized (RCT1:450/225 IU, RCT2:100 IU) or standard FSH dose (150 IU). Women in the cohort all received the standard dose (150 IU). Anti-Müllerian hormone (AMH) was measured to assess AMH post-hoc as a biomarker to individualize treatment. For RCT1 dose adjustment was allowed in subsequent cycles based on pre-specified criteria in the standard group only. For RCT2 dose adjustment was allowed in subsequent cycles in both groups. Both effectiveness and cost-effectiveness of the strategies were evaluated from an intention-to-treat perspective.

MAIN RESULTS AND THE ROLE OF CHANCE: We included 1515 women, of whom 483 (31.9%) entered the cohort, 511 (33.7%) RCT1 and 521 (34.4%) RCT2. Live births occurred in 420/747 (56.3%) women in the individualized strategy and 447/769 (58.2%) women in the standard strategy (risk difference -0.019 (95% CI, -0.06 to 0.02), $P = 0.39$; a total of 1516 women due to rounding up the half integer numbers). The individualized strategy was more expensive (delta costs/woman = €275 (95% CI, 40 to 499)). Individualized dosing reduced the occurrence of mild and moderate ovarian hyperstimulation syndrome (OHSS) and subsequently the costs for management of these OHSS categories (costs saved/woman were €35). The analysis based on AMH as a tool for dose individualization suggested comparable results.

LIMITATIONS, REASONS FOR CAUTION: Despite a training programme, the AFC might have suffered from inter-observer variation. In addition, although strict cancel criteria were provided, selective cancelling in the individualized dose group (for poor response in particular) cannot be excluded as observers were not blinded for the FSH dose and small dose adjustments were allowed in subsequent cycles. However, as both first cycle live birth rates and cumulative live birth rates show no difference between strategies, the open design probably did not mask a potential benefit for the individualized group. Despite increasing consensus on using GnRH antagonist co-treatment in women predicted for a hyper response in particular, GnRH agonists were used in almost 80% of the women in this study. Hence, in those women, the AFC and bloodsampling for the post-hoc AMH analysis were performed during pituitary suppression. As the correlation between AFC and ovarian response is not compromised during GnRH agonist use, this will probably not have influenced classification of response.

WIDER IMPLICATIONS OF THE FINDINGS: Individualized FSH dosing for the IVF/ICSI population as a whole should not be pursued as it does not improve live birth rates and it increases costs. Women scheduled for IVF/ICSI with a regular menstrual cycle are therefore recommended a standard FSH starting dose of 150 IU per day. Still, safety management by individualized dosing in predicted hyper responders is open for further research.

STUDY FUNDING/COMPETING INTEREST(S): This study was funded by The Netherlands Organisation for Health Research and Development (ZonMW number 171102020). AMH measurements were performed free of charge by Roche Diagnostics. TCT, HLT and SCO received an unrestricted personal grant from Merck BV. AH declares that the department of Obstetrics and Gynecology, University Medical Centre Groningen receives an unrestricted research grant from Ferring pharmaceuticals BV, The Netherlands. CBL receives grants from Merck, Ferring and Guerbet. BWJM is supported by a NHMRC Practitioner Fellowship (GNT1082548) and reports consultancy for OvsEva, Merck and Guerbet. FJMB receives monetary compensation as a member of the external advisory board for Ferring pharmaceuticals BV (the Netherlands) and Merck Serono (the Netherlands) for consultancy work for Gedeon Richter (Belgium) and Roche Diagnostics on automated AMH assay development (Switzerland) and for a research cooperation with Ansh Labs (USA). All other authors have nothing to declare.

TRIAL REGISTRATION NUMBER: Registered at the ICMJE-recognized Dutch Trial Registry (www.trialregister.nl). Registration number: NTR2657.

Key words: antral follicle count / anti-Müllerian hormone / ovarian reserve / ovarian reserve test / IVF / ICSI / individualized / FSH / live birth / cost-effectiveness

Introduction

The use of IVF has increased over the last few decades. In countries where IVF is covered by a public health care system, 2–3% of the babies that are born annually are now conceived by this treatment (Kamphuis and Bhattacharya, 2014). Globally, the annual number of estimated IVF/ICSI cycles per year is 1.5 million, resulting in the birth of approximately 350 000 babies (<https://www.sciencedaily.com/releases/2012/07/120702134746.htm> (accessed 27 Oct 2016)) at a cost of almost 5 billion Euros. Gonadotrophins are used for controlled ovarian stimulation (COS) and contribute significantly to these total costs.

The objective of COS is to collect an appropriate number of oocytes to facilitate IVF. The daily gonadotrophin doses applied range from 100 IU to 600 IU (Nargund et al., 2007). Despite this wide range of doses, a considerable proportion of women still shows either a poor or hyper response to COS (Broer et al., 2013a,b), both of which are thought to negatively affect live birth rates (van der Gaast et al., 2006; Sunkara et al., 2011; Drakopoulos et al., 2015), with hyper response also jeopardizing safety (Delvinge and Rozenberg, 2002). Response prediction through ovarian reserve testing is superior to the use of chronological age alone (Broekmans et al., 2006; Broer et al., 2009; La Marca et al., 2009). Currently, the antral follicle count

(AFC) and serum anti-Müllerian hormone (AMH) are the most practical and reliable ovarian reserve tests (ORTs) available (Broer *et al.*, 2011, 2014; La Marca and Sunkara, 2014) and are most often used for individualized dose adjustment in daily clinical practice. However, there are no clear epidemiological data to support such practice (van Tilborg *et al.*, 2016).

Of the eight RCTs that evaluated ORT-based individualized gonadotrophin dosing (Harrison *et al.*, 2001; Popovic-Todorovic *et al.*, 2003; Klinkert *et al.*, 2005; Berkkanoglu and Ozgur, 2010; Jayaprakasan *et al.*, 2010; Lefebvre *et al.*, 2015; Olivennes *et al.*, 2015; Nyboe Andersen *et al.*, 2017), just one study reported higher ongoing pregnancy rates in women treated with an individualized dose compared to a standard daily dose of 150 IU (Popovic-Todorovic *et al.*, 2003). None of the RCTs primarily evaluated live birth rates or costs, and the largest studies thus far were limited to women with a good prognosis (Popovic-Todorovic *et al.*, 2003; Olivennes *et al.*, 2015; Nyboe Andersen *et al.*, 2017). Thus, there is no sound evidence for the widespread application of ORT-based dose adjustment (van Tilborg *et al.*, 2016). We therefore performed a nationwide multicentre prospective cohort study with two embedded RCTs testing the hypothesis whether individualized gonadotrophin dosing based on an AFC classification improves either cumulative live birth rates and/or cost-effectiveness as compared to a standard dose regimen in women starting IVF/ICSI treatment.

Materials and Methods

This study was designed as a pragmatic multicentre prospective cohort study with two embedded RCTs and was carried out within the Dutch Consortium for Research in Women's Health (www.studies-obsgyn.nl). Details about the study design, sample size calculation, study procedures and outcomes are described in the previously published study protocol (van Tilborg *et al.*, 2012).

Ethical approval

Ethical approval for the study was obtained from the Institutional Review Board of the University Medical Centre Utrecht (MEC 10-273) and the board of directors of all participating centres approved the study. The study was registered in the Dutch Trial Registry (NTR2657) and monitored by a Data Safety Monitoring Board. Data quality, protocol adherence and safety were monitored by qualified and trained staff.

Study design and participants

Infertile women younger than 44 years starting their first IVF or ICSI treatment, or the first IVF or ICSI treatment after a previous live birth, were eligible for inclusion. Participants had to have a regular menstrual cycle (on average 25–35 days) and no significant uterine or ovarian abnormalities on transvaginal ultrasound. Women with polycystic ovary syndrome (PCOS) (PCOS Consensus Workshop Group TRE-SPCWG, 2004), endocrine or metabolic abnormalities or women who opted for oocyte donation were not included in the trial. Eligible women were recruited and counselled by their physicians or by dedicated research nurses prior to the start of their treatment. All participants provided written informed consent before participation. Just prior to the start of their COS, women underwent ORT by measuring the AFC as well as serum sampling/banking for a pre-specified, post-hoc analysis with AMH. Based on the AFC result, women entered one of the two RCTs (RCT1 included predicted poor responders: AFC < 11, RCT2 included predicted hyper responders:

AFC > 15) or were followed in the cohort (predicted normal responders: AFC 11–15). The detailed methodology, including the results of the two RCTs, has been published in 'Human Reproduction' (Oudshoorn *et al.*, 2017 and van Tilborg *et al.*, 2017).

All women had their antral follicles (2–10 mm) counted following previously published practical recommendations (Broekmans *et al.*, 2010). The AFC was performed by properly trained physicians, as at the start of the trial several AFC methodology workshops were given and video instructions were provided to all participating centres. On the same day as the AFC was measured, a blood sample was taken, processed and stored at -80°C for later analysis of serum AMH. AMH levels were determined centrally by using the fully automatic Elecsys assay, on a cobas e601 system (Roche Diagnostics, Germany). The detection limit of the assay was 0.01 ng/mL (0.07 pmol/L) and the intermediate precision coefficient of variation of the system was 2.7–3.5% (Elecsys AMH package insert 2016-10, V 4.0).

Randomization and masking

Within each of the two RCTs, women were randomly assigned to the use of a fixed individualized (450/225 IU/day for expected poor responders and 100 IU/day for expected hyper responders) or standard FSH dose (150 IU/day) in a 1:1 ratio, using a web-based randomization program with variable block-sizes and stratification per centre. Women in the cohort were not randomized and all received a fixed standard FSH dose (150 IU/day). Clinicians, participants and investigators were not masked for the assigned treatment.

Procedures

Details regarding the IVF/ICSI procedures have been published previously (van Tilborg *et al.*, 2012) and further details can be found in the two separate RCT papers (Oudshoorn *et al.*, 2017; van Tilborg *et al.*, 2017). If women did not become pregnant in the first treatment cycle, they continued to use the assigned dose in subsequent cycles. However, in all standard dose groups and in the 100 IU dose group, small dose adjustments in the subsequent cycles were permitted with a maximum step of 25–50 IU/day under strict conditions (van Tilborg *et al.*, 2012).

Outcomes

The primary outcome was ongoing pregnancy achieved within 18 months after randomization leading to live birth. Live birth was defined as the delivery of at least one live foetus after 24 weeks of gestation. All treatment cycles, including cryo/thaw cycles, performed and all ongoing pregnancies achieved within 18 months after randomization, irrespective of the mode of conception, were recorded. According to the intention-to-treat principle, pregnancies after escape IUI, natural conceptions occurring between treatment cycles and pregnancies after a switch to gamete donation were all taken into account. For the perspective of the cost-effectiveness analyses, data on direct medical costs were collected. All outcomes were pre-specified in the study protocol and are reported in detail in the two RCT papers (Oudshoorn *et al.*, 2017; van Tilborg *et al.*, 2017).

Sample size calculation

For women with a predicted poor response (AFC < 11) we assumed an increase of 15% in cumulative ongoing pregnancy leading to live birth rate (from 25% to 40%) when they were treated by the individualized FSH strategy as compared to a standard FSH dose (Popovic-Todorovic *et al.*, 2003). In order to be able to detect this difference, 300 participants were needed in RCT1 (150 women per group; 80% power, alpha-error 5%). As we expected that 20% of the women would be classified as predicted poor

responder (Broer et al., 2013a), we aimed to include 1500 women in the overall cohort study.

Statistical analysis

All analyses were performed according to the intention-to-treat principle. As women from the non-randomized cohort were not allocated to an individualized or standard strategy, we added all data from the cohort to both strategy groups derived from the RCTs, with weight 0.5, to be able to perform a total strategy analysis (Fig. 1). For continuous data we calculated weighted means (SDs) and for categorical data (numbers (%)) weighted totals were determined. This could potentially result in half-integer numbers for the strategies if there was an uneven number of events in the cohort. We rounded up the half-integer number of events to a whole number for presentation in the tables. Bootstrapping by taking 5000 random samples with replacement was used to calculate 95% CI and *P*-values. Results were tested two-sided and if the *P*-value was below 0.05 the result was considered to indicate statistical significance. Data were analysed in SPSS for Windows (version 21.0, IBM Corp Armonk, NY, USA) and R (version 3.1.3, R Foundation for Statistical Computing, Vienna, Austria).

Effectiveness analysis

Time to ongoing pregnancy leading to live birth was graphically depicted in cumulative incidence curves. Women who did not reach the primary outcome or did not complete the follow-up period because of loss of contact (despite repeated attempts) were considered as 'not having an ongoing pregnancy leading to live birth' at 18 months of follow-up. A pre-specified per-protocol analysis of the primary outcome was performed by excluding women with protocol violations related to their first cycle FSH starting dose and by excluding treatment cycles from the moment a woman had their first protocol violation regarding within and between cycle dose adjustments. Furthermore, live births achieved after natural conception or a conception mode other than a regular IVF or ICSI treatment, including escape IUI, were not included in this per-protocol analysis.

Health economic evaluation

For both strategies we calculated the mean costs and effectiveness from the complete cases with the standard strategy as the reference.

As mentioned previously, data from the cohort were added to those from both strategy groups, with weight 0.5. Unit costs were derived for the following categories: ORT, stimulation phase, laboratory phase, luteal support, ovarian hyperstimulation syndrome (OHSS) and cryo/thaw cycles (Supplementary Table S1). Costs for medication were derived from the Dutch Formulary for medication (<https://www.farmacotherapeutischkompas.nl> (accessed 21 Sept 2016)). For all other costs, when available, we used unit costs derived from previous fertility studies (Merkus, 2006). If no previously calculated unit cost was available an expert panel on cost-effectiveness from the Dutch Consortium for Research in Women's Health provided unit costs based on an average of academic and general hospital prices (Supplementary Table S1). All costs were expressed in Euros. Costs calculations were performed using Dutch price levels in 2015, with adjustments to correct for monetary inflation according to the consumer price index (CBS, Statistics Netherlands, 2016) if necessary. In order to test the robustness of our cost-effectiveness result, bootstrap resampling with replacement was performed.

Pre-specified post-hoc AMH-based analysis

Women were eligible for this analysis if a serum sample had been stored. Published AMH cut-offs were not available for the Elecsys assay. Therefore, these were defined by using the cumulative percentages belonging to the AFC cut-offs, such that the proportion of predicted poor, normal and hyper responders remained the same. Women classified into a different response category based on the AMH cut-off (discordant test result) could have received a different FSH starting dose than based on the AFC classification (Supplementary Table SII). Therefore, the outcome of treatment may have been different. To correct for this, the statistical method of standardization was applied by replacing the observed outcome with a predicted outcome for their AMH-based dose group. For the development of this prediction model, we evaluated various potential predictors of live birth after IVF/ICSI (McLernon et al., 2016). The best fitting prediction model included female age, duration of infertility, AMH and the applied FSH dose group, with an accuracy of 0.7. The effectiveness analysis was then repeated for the following standardized outcomes: number of cancellations, number of oocytes yielded, number of poor responses, number of hyper responses, OHSS events, time to ongoing pregnancy leading to live birth and ongoing pregnancy leading to live birth.

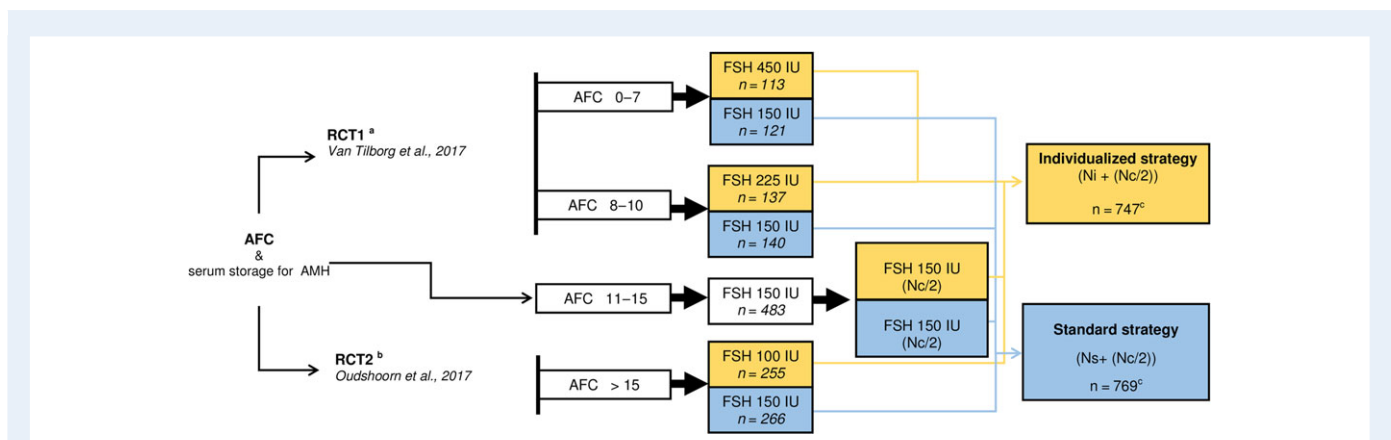


Figure 1 Study design and total strategy analysis. AFC = antral follicle count, AMH = anti-Müllerian hormone, Nc = number of events in the cohort (AFC 11–15) in case of a categorical variable, Ni = number of events in women randomized for the individualized dose in case of a categorical variable, Ns = number of events in the women randomized for the standard dose in case of a categorical variable. ^aResults from RCT1 are published in van Tilborg et al. (2017). ^bResults from RCT2 are published in Oudshoorn et al. (2017). ^cWe rounded up the half-integer number of events to a whole number for presentation in the figure.

Results

Between 12 May 2011 and 5 May 2014, we included 1515 women of whom 511 (33.7%) had an AFC < 11 and were included in RCT1, 521 (34.4%) had an AFC > 15 and were included in RCT2 and 483 (31.9%) had an AFC 11–15 and were included in the cohort (Fig. 2). Within the RCTs, 505 (33.3%) women were randomly allocated to the individualized FSH strategy and 527 (34.8%) to the standard FSH strategy. The results of the two RCTs have been published separately (Oudshoorn *et al.*, 2017; van Tilborg *et al.*, 2017).

Baseline characteristics of the two strategy groups were comparable (Table I and Supplementary Table SIII for baseline characteristics of the AFC subgroups). Owing to rounding up the half-integer numbers, a total of 1516 is displayed in the results. Almost all women (98.2%) received the assigned dose during the first treatment cycle (Fig. 2). The number of women lost to follow-up before 18 months was 81/747 (10.8%) in the individualized group versus 86/769 (11.1%) in the standard group, with an equal mean duration of follow-up (275.7 (142.6) days vs. 273.4 (146.8) days, respectively; MD 2.2 (95%CI, –34.4 to 39.6), $P = 0.91$).

Effectiveness analysis

Ongoing pregnancy leading to live birth occurred in 420/747 women (56.3%) assigned to the individualized FSH strategy as compared to 447/769 women (58.2%) in the standard strategy (risk difference (RD) –0.019 (95%CI, –0.06 to 0.02), $P = 0.38$; Table II and Fig. 3). Time to ongoing pregnancy leading to live birth did not differ between groups ($P = 0.89$). The mean number of IVF/ICSI treatment cycles per woman conducted in both strategies was similar (1.9 (0.9) vs. 1.9 (1.0),

MD –0.02 (95%CI, –0.10 to 0.06), $P = 0.56$; Table III and Supplementary Table IV). Individualized FSH dosing resulted in a higher incidence of normal responses and a lower incidence of hyper responses and any OHSS event (mild, moderate and severe), but the number of cycles with severe OHSS was 8 for both groups.

The results of the first cycle are summarized in Table IV and Supplementary Table IV. Despite differences in the number of retrieved oocytes and the proportion of poor and hyper responders, live birth rates were comparable (28.9% vs. 30.7%, RD –0.02 (95% CI, –0.06 to 0.02), $P = 0.32$). The OHSS rates were 4.9% versus 8.5% for women in the individualized and standard FSH strategy group (RD –0.04 (95% CI, –0.06 to –0.02), $P = 0.001$), but severe OHSS rates did not differ (0.8% vs. 0.8%, RD 0.0003 (95% CI, –0.007 to 0.008), $P = 0.86$).

The per-protocol analysis confirmed the similarity in cumulative live birth rates (individualized FSH strategy 323/712 (45.3%) vs. standard FSH strategy 363/744 (48.7%), RD –0.03, (95%CI, –0.08 to 0.01), $P = 0.11$).

Health economic evaluation

Since only a small proportion of the variables needed for the cost-effectiveness analyses were missing (0.7% and 1.1% in women allocated to the individualized or standard FSH strategy), the cost-effectiveness analysis was a complete case analysis. The total number of compared cycles was 1379 versus 1432 cycles in the individualized versus standard FSH strategy group, respectively. The mean costs per woman after 18 months of randomization were €5215 for the individualized and €4940 for the standard FSH strategy (MD €275, (95%CI, 40 to 499); Supplementary Table I). The total number of live births per strategy differed a little from the effectiveness analysis, as cycles with incomplete cost unit information were excluded, resulting in 409/744 (54.9%) versus 432/764 (56.5%) live birth events in the individualized versus standard strategy, respectively. The mean costs per woman for management of mild/moderate OHSS were €53 in the individualized and €88 in the standard FSH strategy group (costs saved/woman were €35).

Since individualized dosing was more expensive and led to slightly less live births, standard dosing was the dominant strategy and the average cost increase per percentage point of live birth rate increase (incremental cost-effectiveness ratio) could not be calculated. In order to test the robustness of the results of our cost-effectiveness analysis bootstrap resampling was performed. Figure 4 shows the bootstrap result of 5000 resamples (depicted as points in the plane) taken to represent the uncertainty of the calculated additional costs and observed effects on live birth rate of individualized compared to standard FSH dosing. As the majority of the points are in the upper left quadrant, this indicates that individualized FSH dosing tends to be less effective and more expensive than standard dosing.

A cost-effectiveness acceptability curve was then derived based on the bootstrapped results, showing the probability that individualized FSH dosing based on the AFC is cost-effective compared with a standard FSH regimen, given the observed data, for a range of values of the willingness to pay for an additional live birth (Fig. 5). The maximum probability of cost-effectiveness is approximately 20%.

Pre-specified post-hoc AMH-based analysis

Serum AMH levels were available in 1345 out of 1515 women (88.8%). According to the procedure described in the methods

Table I Baseline characteristics of the study population.

Baseline characteristics	Individualized (n = 747)	Standard (n = 769)
Age (years)	33.5 (4.2)	33.5 (4.4)
Duration of infertility (years)	2.7 (1.8)	2.6 (1.7)
Smoking	122 (17.0%)	133 (17.9%)
BMI (kg/m ²)	24.0 (3.6)	24.1 (3.7)
Primary infertility	457 (61.4%)	463 (60.2%)
Cause of infertility ^a		
Unexplained	245 (32.8%)	280 (36.4%)
Male factor	420 (56.2%)	420 (54.6%)
Endometriosis	38 (5.1%)	23 (3.0%)
Tubal factor	83 (11.1%)	75 (9.8%)
AFC	14.1 (6.8)	14.2 (7.1)
AMH (ng/mL)	2.23 (1.68)	2.18 (1.67)
GnRH analogue used		
GnRH agonist	590 (79.2%)	605 (78.8%)
GnRH antagonist	155 (20.8%)	163 (21.2%)

Data are n (%) or mean (SD). AFC = antral follicle count, AMH = anti-Müllerian hormone.

^aIncluding patients with >1 cause of infertility.

section, the following AMH cut-offs (ng/mL) were calculated: AMH \leq 0.769 (predicted poor responders, $n = 207$), AMH > 0.769 and ≤ 1.330 (predicted normal-poor responders, $n = 246$), AMH > 1.330 and ≤ 2.520 (predicted normal responders, $n = 428$) and AMH > 2.520 (predicted hyper responders, $n = 464$). In 323/1345 (24.0%) women, the observed outcomes were replaced by prediction model-

based outcomes (Supplementary Table II). Baseline characteristics of women included in this AMH analysis were comparable between strategies (Supplementary Table SV). The primary outcome of ongoing pregnancy leading to live birth occurred in 55.9% women allocated to the individualized strategy, compared to 58.3% women allocated to the standard strategy (RD -0.024 , (95%CI, -0.062 to 0.015), $P = 0.13$).

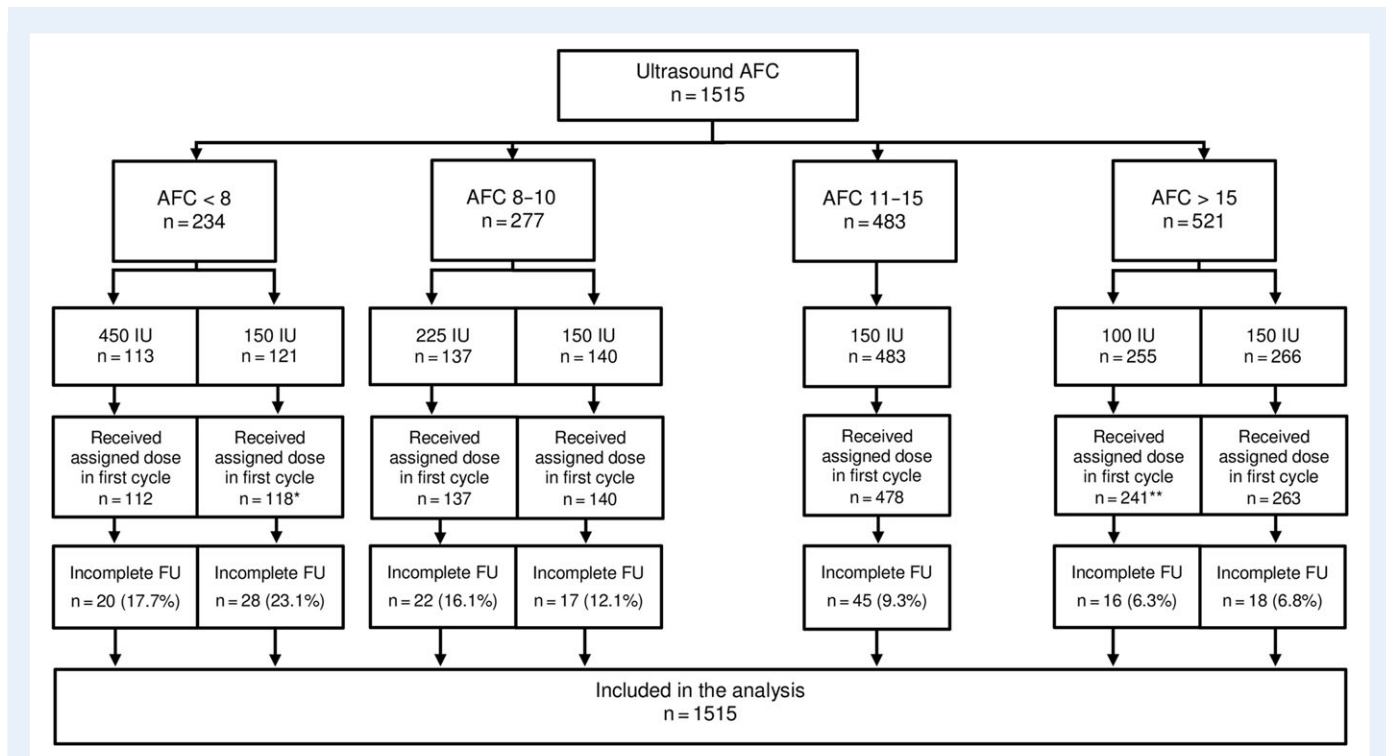


Figure 2 Flowchart of the study of individualized FSH dosing based on AFC in women starting IVF/ICSI. * $n = 1$ never started ovarian stimulation and $n = 1$ never started treatment; ** $n = 2$ no cycle information. FU = follow-up.

Table II Cumulative pregnancy results per woman.

Pregnancy data per woman ^a	Individualized ($n = 747$)	Standard ($n = 769$)	MD/RD (95% CI)	P-value
Ongoing pregnancy resulting in live birth within 18 months of FU	420 (56.3%)	447 (58.2%)	-0.019 (-0.06 to 0.02)	0.38
Multiple pregnancy ^b	16 (2.1%)	26 (3.3%)	-0.012 (-0.03 to 0.001)	0.07
Conception mode (% of live birth)				0.78
Spontaneous	35 (8.2%)	30 (6.6%)	0.016 (-0.01 to 0.05)	
IVF/ICSI fresh ^c	298 (71.0%)	314 (70.2%)	0.007 (-0.04 to 0.06)	
IVF/ICSI frozen	84 (19.9%)	95 (21.2%)	-0.013 (-0.06 to 0.03)	
Other ^d	4 (1.0%)	9 (2.0%)	-0.011 (-0.03 to 0.001)	
Time to ongoing pregnancy leading to live birth, days (mean)	188.0 (128.0)	187.1 (128.6)	0.95 (-13.6 to 15.4)	0.89

Data are n (%) or mean (SD). FU = follow-up, MD = mean difference, RD = risk difference. Bootstrapping by taking 5000 random samples with replacement was used to calculate 95% CI and P-values.

^aWith a maximum of one event per woman.

^bBased on the number of ongoing pregnancies.

^cIncluding escape IUI.

^dIncluding the following classifications: individualized strategy 4x unknown; standard strategy 6x unknown, 2x IUI in another clinic, unknown whether regular IUI or escape IUI, 1x IUI with donor semen.

Women in the AMH-based individualized strategy had less cycle cancellations, ovarian hyper responses and mild and moderate OHSS events (Supplementary Table SVI). The occurrence of severe OHSS did not differ between strategies.

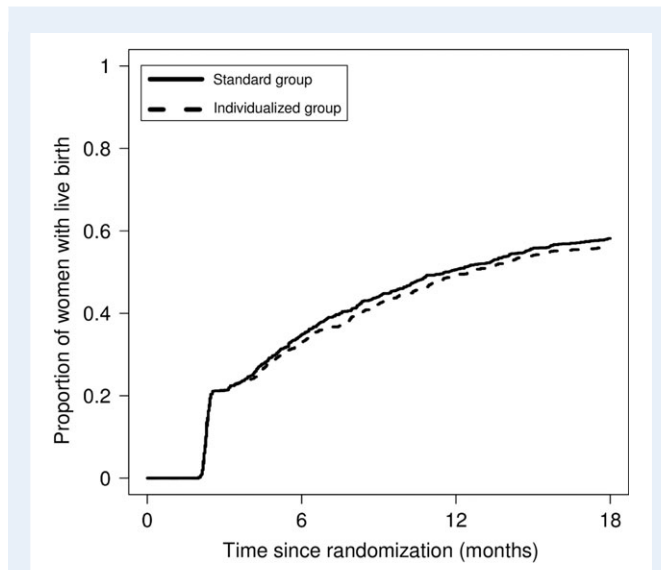


Figure 3 Cumulative incidence curve presenting time to ongoing pregnancy leading to live birth. Women with incomplete FU were deemed not to be pregnant at 18 months of FU. Standard: the group of women receiving the standard FSH dose of 150 IU per day irrespective of the result of the AFC. Individualized: the group of women receiving and individualized FSH dose of 100, 225 or 450 IU per day based on the AFC predicted response category.

Discussion

In this prospective cohort study with two embedded RCTs involving infertile women starting their first IVF/ICSI treatment, an individualized FSH strategy based on the AFC did not result in better live birth rates as compared to a standard FSH strategy. Also, individualized FSH dosing was slightly more expensive than standard dosing. The only beneficial effect of individualized FSH dosing in, for the most part GnRH agonist co-treated cycles, was a reduction of the rate of mild and moderate OHSS, but effects on the occurrence of severe OHSS per se could not be demonstrated. Finally, a post-hoc scenario analysis with AMH as ORT, using the statistical method of standardization to correct for non concordant test results, suggested comparable results.

Our effectiveness results are in contrast to the findings of Popovic-Todorovic *et al.* (2003) who reported improved ongoing pregnancy rates in women treated with an individualized FSH dose based on a multifactor algorithm compared to a standard daily dose of 150 IU. The robustness of this finding can be questioned since that study was not powered for ongoing pregnancy rates, included single cycle results only and may have limited generalizability as mainly women with a good prognosis were included. Indeed, a recent review failed to find compelling evidence that individualized dosing positively affects pregnancy rates as compared to standard dosing (van Tilborg *et al.*, 2016). As a previous cost-effectiveness analysis did suggest that individualization of the FSH dose according to an ORT could well be cost-effective, the urge for large studies was felt indeed (Moolenaar *et al.*, 2011).

One of the two large studies recently executed focussed on patients with a relatively good prognosis and compared individualized dosing based on AMH and body weight with conventional treatment in GnRH antagonist co-treated cycles (Nyboe Andersen *et al.*, 2017). In line with our study, Nyboe Andersen *et al.* did not find any difference in live birth rates. Their findings suggest improvement in safety for excessive

Table III Cumulative treatment results.

Cumulative cycle results	Individualized (n = 745)	Standard (n = 768)	MD/RD (95% CI)	P-value
Number of fresh cycles started/woman	1.9 (0.9)	1.9 (1.0)	-0.02 (-0.10 to 0.06)	0.56
Cycle cancellation ^a	166 (11.7%)	213 (14.5%)	-0.03 (-0.05 to -0.006)	0.01
Number of oocytes ^b	8.7 (5.0)	9.2 (5.6)	-0.53 (-0.88 to -0.18)	0.003
Poor response ^c	393 (28.0%)	425 (29.0%)	-0.01 (-0.04 to 0.02)	0.47
Hyper response ^c	149 (10.6%)	220 (15.0%)	-0.04 (-0.06 to -0.02)	<0.001
Number of fresh embryo transfers ^a	1120 (79.4%)	1133 (77.2%)	0.02 (-0.004 to 0.05)	0.08
Number of embryos per transfer	1.2 (0.44)	1.2 (0.43)	-0.003 (-0.03 to 0.03)	0.87
OHSS classification ^d				0.003
No OHSS	1343 (95.9%)	1371 (93.6%)	0.023 (0.01 to 0.04)	<0.001 ^e
Mild OHSS	44 (3.1%)	68 (4.6%)	-0.02 (-0.03 to -0.004)	0.01 ^e
Moderate OHSS	6 (0.4%)	18 (1.2%)	-0.008 (-0.01 to -0.004)	0.004 ^e
Severe OHSS	8 (0.6%)	8 (0.5%)	0.0002 (-0.005 to 0.005)	0.89 ^e
Number of cryo cycles with transfer/woman	0.8 (1.3)	0.8 (1.3)	-0.02 (-0.12 to 0.09)	0.72

Data are n (%) or mean (SD). OHSS = ovarian hyperstimulation syndrome, cryo = cryopreservation.

^aData available from 1411 cycles in the individualized and 1469 cycles in the standard strategy.

^bBased on the number of cycles with oocyte retrieval.

^cData available from 1403 cycles in the individualized and 1464 cycles in the standard strategy.

^dData available from 1400 cycles in the individualized and 1464 cycles in the standard strategy.

^ePost-hoc testing. P-value is considered to be significant when <0.013 (Bonferroni correction).

Table IV First treatment results.

First cycle results	Individualized (n = 745)	Standard (n = 768)	MD/RD (95% CI)	P-value
Total FSH used IU	2201 (1420)	1701 (451)	500 (395 to 611)	<0.001
Total duration stimulation days	11.5 (2.9)	11.3 (2.8)	0.18 (−0.07 to 0.43)	0.14
Cancellations ^a	103 (13.8%)	117 (15.2%)	−0.01 (−0.05 to 0.02)	0.34
Number of oocytes ^b	8.8 (5.2)	10.0 (6.0)	−1.2 (−1.69 to −0.67)	<0.001
Poor response ^c	220 (29.6%)	197 (25.7%)	0.04 (−0.001 to 0.08)	0.04
Hyper response ^c	93 (12.4%)	148 (19.2%)	−0.07 (−0.10 to −0.04)	<0.001
Number of fresh embryo transfers	570 (76.6%)	587 (76.5%)	0.001 (−0.04 to 0.04)	0.94
Number of fresh embryos per transfer	1.1 (0.3)	1.1 (0.3)	0.01 (−0.02 to 0.04)	0.51
OHSS classification ^c				0.005
No OHSS	707 (95.1%)	702 (91.5%)	0.04 (0.02 to 0.06)	0.001 ^d
Mild OHSS	28 (3.7%)	49 (6.3%)	−0.03 (−0.04 to −0.009)	0.004 ^d
Moderate OHSS	3 (0.4%)	11 (1.4%)	−0.01 (−0.02 to −0.005)	0.010 ^d
Severe OHSS	6 (0.8%)	6 (0.8%)	0.0003 (−0.007 to 0.008)	0.86 ^d
Number of cryo cycles with transfer/woman	0.5 (1.0)	0.5 (1.0)	−0.06 (−0.14 to 0.03)	0.21
Number of cryo embryos per transfer	1.1 (0.3)	1.1 (0.2)	0.01 (−0.02 to 0.04)	0.44
Live birth (fresh only) ^e	163 (21.9%)	169 (22.0%)	−0.001 (−0.03 to 0.03)	0.94
Live birth (fresh and cryo) ^e	215 (28.9%)	236 (30.7%)	−0.02 (−0.06 to 0.02)	0.32

Data are n (%) or mean (SD).
^aData available from 745 women in the individualized and 768 women in the standard strategy.
^bBased on the number of women who underwent oocyte retrieval.
^cData available from 744 women in the individualized and 768 women in the standard strategy.
^dPost-hoc testing, P-value is considered to be significant when <0.013 (Bonferroni correction).
^ePercentage calculated based on the number of women per group.

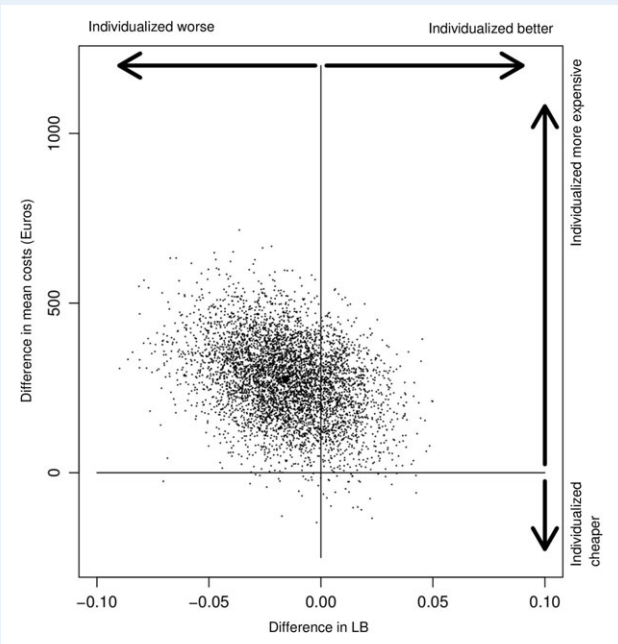


Figure 4 Cost-effectiveness plane of individualized FSH strategy compared with a standard FSH strategy. Each point represents the uncertainty of the additional costs and effects (measured as live birth) of individualized FSH dosing as compared with standard dosing after bootstrap resampling. LB = live birth.

responders, based on a reduced number of preventive interventions for OHSS and/or occurrences of moderate/severe OHSS. More studies on safety management in women who are predicted hyper responders are needed before FSH dose adjustment can be implemented.

The cost-effectiveness analysis was performed from a health care perspective and therefore focussed on direct medical costs during treatment. A previous study demonstrated that indirect costs of fertility treatment and indirect costs from a societal perspective will only have a small impact on the total costs (Reindollar et al., 2010). The inclusion of these indirect costs is therefore unlikely to change our overall cost-effectiveness result.

The present study used AFC as a tool to choose the FSH dose. As AMH results could not be obtained rapidly when this study was initiated, serum was stored simultaneously for later AMH analysis. AMH might be a better predictor of the ovarian response category, and thereby a better test for FSH dose individualization (Iliodromiti et al., 2015; Nelson et al., 2015). Still, from compilations of existing literature the accuracy of AFC and AMH in predicting poor and excessive response has been demonstrated to be equal (La Marca et al., 2009; Broer et al., 2013a,b). In our pre-specified analysis using serum AMH cut-offs for predicted response classifications, no beneficial effect on live birth rates was observed, which is in line with the AFC-based analyses.

The present study has several strengths. First, it was powered to assess cumulative live birth rates, an outcome measure that reflects day-to-day practice better than single cycle results and is the most important outcome from the patient's perspective (Scholten et al., 2016). Second, this

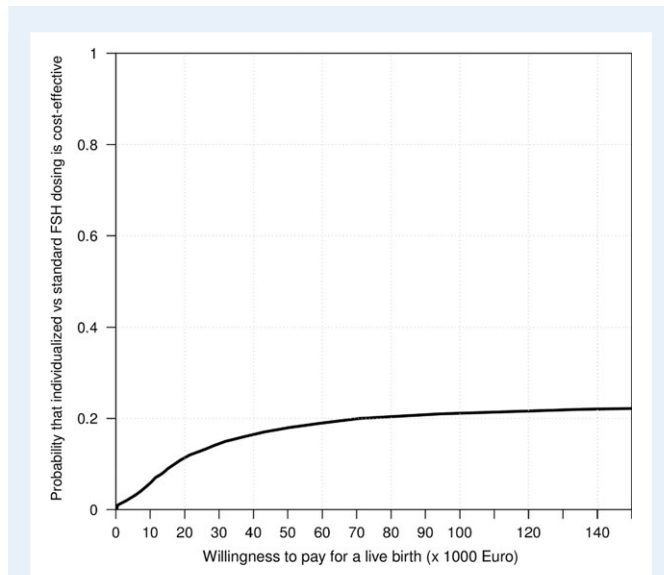


Figure 5 Cost-effectiveness acceptability curve. The cost-effectiveness acceptability curve shows the probability that individualized FSH dosing based on the AFC is cost-effective compared with a standard FSH regimen, given the observed data, for a range of values of the willingness to pay for an additional live birth. This figure shows that the maximum probability of cost-effectiveness is approximately 20%.

is the largest cohort that directly assessed both effectiveness and cost-effectiveness of a frequently used strategy in daily practice, the so-called 'individualized FSH dosing'. The performance of a cost-effectiveness analysis is essential, particularly in countries with a public health care system, before a strategy can be implemented as standard care. Third, due to the inclusion of a heterogeneous study population and the performance of an intention-to-treat analysis, this study has a significant potential for satisfactory external validation and therefore well reflects daily practice. Besides, our live birth results seem to be robust as the results from single cycles only, cumulative treatment cycles and the per-protocol analyses all pointed in the same direction.

Some limitations need to be addressed. First, the fact that the AFC was measured by observers in multiple centres is a potential weakness owing to known inter-observer variability (Scheffer *et al.*, 2002; Broekmans *et al.*, 2010). To improve the inter-observer reproducibility, AFC workshops were held and instructions for the AFC procedure were given prior to the start of the study (Broekmans *et al.*, 2010). The remaining variability could also be seen as a strength of the study as it reflects the daily practice variation. Second, in women with GnRH agonist co-treatment the assessment of the ORT was during pituitary down-regulation. This might have influenced the ORT results, but the final effect is unknown as contradictory results regarding the influence of down-regulation on AFC and AMH have been published (Sharara *et al.*, 1999; Yong *et al.*, 2003; Frattarelli, 2006; Wang *et al.*, 2007; Jayaprakasan *et al.*, 2008a,b; Su *et al.*, 2013; Tran *et al.*, 2016). Despite the possible presence of influence from GnRH agonist down-regulation, there seems to be no influence on the ovarian response predictive value as correlations between the ORT and number of oocytes were not affected (Yong *et al.*, 2003; Frattarelli, 2006; Wang *et al.*, 2007; Jayaprakasan *et al.*, 2008b; Tran *et al.*, 2016).

Another potential limitation of the study is that, between treatment cycles, dose adjustments were permitted under strict pre-specified conditions, which could have reduced the magnitude of the differences between the strategies. The dose adjustment rules were introduced both to ensure a high protocol adherence and to mimic current daily practice as much as possible, where adapting the dose based on previous 'inappropriate' ovarian response is very common. As first cycle live birth rates did not differ between the strategies, it is unlikely that the between cycle dose adjustments masked a potential benefit from individualized dosing. Furthermore, in recent years an increasing trend has been set in which women with a high risk for OHSS are treated with a GnRH antagonist, so a GnRH agonist trigger can be used if needed. At the time of the design of the OPTIMIST study, this was not established practice and so only 20% of women were treated with a GnRH antagonist. Moreover, only HCG triggering was allowed according to protocol. Knowing that the occurrence of mild/moderate OHSS was 1.5 times higher in the standard strategy group and assuming such preventive measures would have been able to reduce OHSS, a subsequent reduction in costs for management of OHSS may be expected. This, in turn, might favour the standard dosing strategy even more in the cost-effectiveness analysis.

Finally, as AMH was assessed post-hoc, statistical methods had to be applied to handle outcomes from women with discordant AFC and AMH results. Therefore these observations need to be interpreted with caution.

In conclusion, the present large study demonstrates that individualized FSH dosing based on the AFC does not improve live birth rates or reduce costs. ORT-based treatment for the IVF/ICSI population as a whole should therefore not be pursued. Women scheduled for IVF/ICSI with a regular menstrual cycle are therefore recommended to a standard FSH starting dose of 150 IU per day.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

We would like to thank the women who participated in this study, the staff of the participating hospitals and the office members of the Dutch Consortium (www.studies-obsgyn.nl) for their contributions.

Authors' roles

M.J.C.E., H.L.T., F.J.M.B. and B.W.J.M. designed the trial. T.C.T., H.L.T., F.J.M.B. and B.W.J.M. coordinated the trial. T.C.T., S.C.O., M.H.M., R.J.T.G., A.H., W.K.H.K., K.F., J.P.B., C.B.L., J.S.E.L., F.J.M.B. and all other members from the OPTIMIST study group collected data. M.J.C.E., H.G. and M.W. provided input on the statistical analyses. T.C.T. and M.J.C.E. performed the statistical analyses. T.C.T. drafted the manuscript. T.C.T., S.O., M.J.C.E., H.L.T., F.J.M.B. and B.W.J.M. interpreted the data. All authors revised the report, and approved the final submitted version.

Funding

The study was funded by The Netherlands Organisation for Health Research and Development (ZonMW number 171102020). The funder of the study had no role in any of the study proceedings (i.e., study design, data collection, data analysis, data interpretation or writing of the manuscript).

Conflict of interest

T.C.T., H.L.T. and S.C.O. received an unrestricted personal grant from Merck BV. A.H. declares that the department of Obstetrics and Gynecology, University Medical Centre Groningen receives an unrestricted research grant from Ferring pharmaceuticals BV, The Netherlands. C.B.L. receives grants from Merck, Ferring and Guerbet. B.W.J.M. is supported by a NHMRC Practitioner Fellowship (GNT1082548) and reports consultancy for OvsEva, Merck and Guerbet. F.J.M.B. receives monetary compensation as a member of the external advisory board for Ferring pharmaceuticals BV (the Netherlands) and Merck Serono (the Netherlands) for consultancy work for Gedeon Richter (Belgium) and Roche Diagnostics on automated AMH assay development (Switzerland) and for a research cooperation with Ansh Labs (USA). All other authors have nothing to declare.

References

- Berkkanoglu M, Ozgur K. What is the optimum maximal gonadotropin dosage used in microdose flare-up cycles in poor responders? *Fertil Steril* 2010;**94**:662–665.
- 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.
- 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.
- Broer SL, Broekmans FJM, Laven JSE, Fauser BCJM. Anti-Müllerian hormone: ovarian reserve testing and its potential clinical implications. *Hum Reprod Update* 2014;**20**:688–701.
- Broer SL, Dölleman M, Opmeer BC, Fauser BC, Mol BW, Broekmans FJM. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. *Hum Reprod Update* 2011;**17**:46–54.
- Broer SL, Dölleman M, van Disseldorp J, Broeze KA, Opmeer BC, Bossuyt P, Eijkemans MJ, Mol BW, Broekmans FJ, IPD-EXPORT Study Group. 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* 2013b;**100**:420–429.
- Broer SL, Mol BWJ, Hendriks D, Broekmans FJM. The role of antimüllerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. *Fertil Steril* 2009;**91**:705–714.
- Broer SL, van Disseldorp J, Broeze KA, Dölleman M, Opmeer BC, Bossuyt P, Eijkemans MJ, Mol BW, Broekmans FJ, IMPORT study group. 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* 2013a;**19**:26–36.
- Delvinge A, Rozenberg S. Epidemiology and prevention of ovarian hyperstimulation syndrome (OHSS): a review. *Hum Reprod Update* 2002;**8**:559–577.
- Drakopoulos P, Blockeel C, Stoop D, Camus M, de Vos M, Tournaye H, Polyzos NP. Conventional ovarian stimulation and single embryo transfer for IVF/ICSI. How many oocytes do we need to maximize cumulative live birth rates after utilization of all fresh and frozen embryos? *Hum Reprod* 2015;**31**:370–376.
- Frattarelli JL. A prospective analysis of the changes in ovarian morphology during hormonal pituitary suppression before in vitro fertilization. *Fertil Steril* 2006;**86**:577–582.
- Harrison RF, Jacob S, Spillane H, Mallon E, Hennelly B. A prospective randomized clinical trial of differing starter doses of recombinant follicle-stimulating hormone (folitropin- β) for first time in vitro fertilization and intracytoplasmic sperm injection treatment cycles. *Fertil Steril* 2001;**75**:23–31.
- <https://www.farmacotherapeutischkompas.nl> (September 21, 2016, date last accessed).
- <https://www.sciencedaily.com/releases/2012/07/120702134746.htm> (October 27, 2016, date last accessed).
- Iliodromiti S, Anderson RA, Nelson SM. Technical and performance characteristics of anti-Müllerian hormone and antral follicle count as biomarkers of ovarian response. *Hum Reprod Update* 2015;**21**:698–710.
- Jayaprakasan K, Campbell BK, Hopkisson JF, Clewes JS, Johnson IR, Raine-Fenning NJ. Effect of pituitary desensitization on the early growing follicular cohort estimated using anti-Müllerian hormone. *Hum Reprod* 2008a;**23**:2577–2583.
- Jayaprakasan K, Hopkisson JF, Campbell BK, Clewes J, Johnson IR, Raine-Fenning NJ. Quantification of the effect of pituitary down-regulation on 3D ultrasound predictors of ovarian response. *Hum Reprod* 2008b;**23**:1538–1544.
- 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: fertility and assisted reproduction. *BJOG* 2010;**117**:853–862.
- Kamphuis E, Bhattacharya S. Are we overusing IVF? *BMJ* 2014;**252**:1–5.
- Klinkert ER, Broekmans FJM, Looman CWN, Habbema JDF, 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, Artensio AC, Stabile G, Volpe A. Anti-Müllerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). *Hum Reprod Update* 2009;**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, Bissonette F, Benoit J, Ménard S, 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.
- McLemon DJ, Maheshwari A, Lee AJ, Bhattacharya S. Cumulative live birth rates after one or more complete cycles of IVF: a population-based study of linked cycle data from 178 898 women. *Hum Reprod* 2016;**31**:572–581.
- Merkus JMW. Fertility treatments: possibilities for fewer multiple births and lower costs—the ‘Umbrella’ study. *Ned Tijdschr Geneesk* 2006;**150**:1162–1164.
- Moolenaar LM, Broekmans FJM, Van Disseldorp J, Fauser BC, Eijkemans MJ, Hompes PG, van der Veen F, Mol BW. Cost effectiveness of ovarian reserve testing in in vitro fertilization: a Markov decision-analytic model. *Fertil Steril* 2011;**96**:889–894.
- Nargund G, Fauser BCJM, Macklon NS, Ombelet W, Nygren K, Frydman R. The ISMAAR proposal on terminology for ovarian stimulation for IVF. *Hum Reprod* 2007;**22**:2801–2804.
- Nelson SM, Klein BM, Arce J-C. Comparison of antimüllerian 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. *Fertil Steril* 2015;**103**:923–30.e1.
- Nyboe Andersen A, Nelson SM, Fauser BCJM, Garcia-Velasco JA, Klein BM, Arce J. Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial. *Fertil Steril* 2017;**107**:387–396.
- 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. *Reprod Biomed Online* 2015;**30**:248–257.
- Oudshoorn SC, van Tilborg TC, Eijkemans MJC, Oosterhuis GJE, Friederich J, van Hooff MHA, van Santbrink EJP, Brinkhuis EA, Smeenk JMJ, Kwee J et al. Individualized versus standard FSH dosing in predicted hyper responders starting IVF/ICSI: an RCT. Part 2: the predicted hyper responder. *Hum Reprod* 2017;**32**:2506–2514.
- PCOS Consensus Workshop Group TRE-SPCWG. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004;**81**:19–25.
- Popovic-Todorovic B, Loft A, Ejdrup Bredkjær H, Bangsbo 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* 2003;**18**:2275–2282.
- Reindollar RH, Regan MM, Neumann PJ, Levine BS, Thornton KL, Alper MM, Goldman MB. A randomized clinical trial to evaluate optimal treatment for unexplained infertility: the fast track and standard treatment (FASTT) trial. *Fertil Steril* 2010;**94**:888–899.
- Scheffer GJ, Broekmans FJM, Bancsi LF, Habbema JDF, Looman CWN, Te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: reproducibility of antral follicle counts. *Ultrasound Obstet Gynecol* 2002;**20**:270–275.
- Scholten I, Braakhekke M, Limpens J, Hompes PG, van der Veen F, Mol BW, Gianotten J. Reporting multiple cycles in trials on medically assisted reproduction. *Reprod Biomed Online* 2016;**33**:646–651.
- Sharara FI, Lim J, McClamrock HD. The effect of pituitary desensitization on ovarian volume measurements prior to in-vitro fertilization. *Hum Reprod* 1999;**14**:183–185.
- Su HI, Maas K, Sluss PM, Chang RJ, Hall JE, Joffe H. The impact of depot GnRH agonist on AMH levels in healthy reproductive-aged women. *J Clin Endocrinol Metab* 2013;**98**:1961–1966.
- Sunkara SK, 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 400 135 treatment cycles. *Hum Reprod* 2011;**26**:1768–1774.
- Tran ND, Aghajanova L, Kao CN, Cedars MI, Rosen MP. Impact of pituitary suppression on antral follicle count and oocyte recovery after ovarian stimulation. *Fertil Steril* 2016;**105**:690–696.
- van der Gaast MH, Eijkemans MJC, van der Net JB, de Boer EJ, Burger CW, van Leeuwen FE, Fauser BC, Macklon NS. Optimum number of oocytes for a successful first IVF treatment cycle. *Reprod Biomed Online* 2006;**13**:476–480.
- van Tilborg TC, Broekmans FJM, Dölleman M, Eijkemans MJ, Mol BW, Laven JS, Torrance HL. Individualized follicle-stimulating hormone dosing and in vitro fertilization outcome in agonist downregulated cycles: a systematic review. *Acta Obstet Gynecol Scand* 2016;**95**:1333–1344.
- van Tilborg TC, Eijkemans MJC, Laven JSE, Koks CA, de Bruin JP, Scheffer GJ, van Golde RJ, Fleischer K, Hoek A, Nap AW et al. The OPTIMIST study: optimisation of cost effectiveness through individualised FSH stimulation dosages for IVF treatment. A randomised controlled trial. *BMC Womens Health* 2012;**12**:29.
- van Tilborg TC, Torrance HL, Oudshoorn SC, Eijkemans MJC, Koks CAM, Verhoeve HR, Nap AW, Scheffer GJ, Manger AP, Schoot BC et al. Individualized versus standard FSH dosing in predicted poor responders starting IVF/ICSI: an RCT. Part I: The predicted poor responder. *Hum Reprod* 2017;**32**:2496–2505.
- Wang JG, Nakhuda GS, Guarnaccia MM, Sauer MV, Lobo RA. Müllerian inhibiting substance and disrupted folliculogenesis in polycystic ovary syndrome. *Am J Obstet Gynecol* 2007;**196**:77.e1–77.e5.
- Yong PYK, Baird DT, Thong KJ, McNeilly AS, Anderson RA. Prospective analysis of the relationships between the ovarian follicle cohort and basal FSH concentration, the inhibin response to exogenous FSH and ovarian follicle number at different stages of the normal menstrual cycle and after pituitary down-regulation. *Hum Reprod* 2003;**18**:35–44.