Ana Rita Batista[1], Ertugrul Akbas[2], Gafar Victoir Coulidiaty[3], Juan-Enrique Schwarze[4], Florian Kohlhepp[4,5], Shiv Shanker Gupta[6], Erica Velthuis[7], Akram Kharbat[8], Monica Lispi[4,11], Thomas D'Hooghe*[4,9,10]

[1]Merck Serono Ltd, London, UK, an affiliate of Merck KGaA, Darmstadt, Germany;
[2]Merck, Turkey, an affiliate of Merck KGaA, Darmstadt, Germany
[2]Merck, Turkey, an affiliate of Merck KGaA, Darmstadt, Germany
[3]Afrika Kommt! And Medical Affairs Fertility EMEA, Merck KGaA, Darmstadt, Germany;
[4]Global Medical Affairs Fertility, R&D Biopharma, Merck KGaA, Darmstadt, Germany;
[5]Medical Affairs Fertility LATAM and MEA, Merck KGaA, Darmstadt, Germany;
[6]Medical Affairs Fertility Europe, Merck KGaA, Darmstadt, Germany;
[7]Global Patient Safety, Merck B.V., Schiphol-Rijk, the Netherlands, an affiliate of Merck KGaA, Darmstadt, Germany;
[8]Global Medical Communications Fertility, Merck KGaA, Darmstadt, Germany;
[9]Department of Development and Regeneration, KU Leuven (University of Leuven), Leuven, Belgium;
[10]Department of Obstetrics and Gynecology, Yale University, New Haven, CT, USA;
[11]International PhD School in Clinical and Experimental Medicine (CEM), University of Modena and Reggio Emilia, Modena, Italy.

*Correspondence: Thomas D'Hooghe
Email: thomas.dhooghe@merckgroup.com

Online comment by Batista et al. regarding article "Randomized, assessor-blinded trial comparing highly purified human menotropin and recombinant follicle-stimulating hormone in high responders undergoing intracytoplasmic sperm injection"

We have read with great interest the article by Witz et al. describing the results of the Menopur in GnRH Antagonist Cycles with Single Embryo Transfer – High Responder (MEGASET-HR) trial which compares highly purified human menotropin (HP-hMG) and

recombinant human follicle-stimulating hormone (r-hFSH) in high responders undergoing intracytoplasmic sperm injection (Witz et al. 2020). With a standard starting dose of 150 IU for either r-hFSH or HP-hMG, Witz et al. report that HP-hMG was non inferior to r-hFSH in regards to ongoing pregnancy (primary endpoint) after fresh transfer (35.5% [HP-hMG] vs 30.7% [r-hFSH]; difference 4.7%, 95% CI, –2.7%, 12.1%). With respect to secondary outcomes, HP-hMG was associated with lower risk of ovarian hyperstimulation syndrome (OHSS; 21.4% vs. 9.7%; difference –11.7%, 95% CI –17.3%, –6.1%) and lower cumulative early pregnancy loss rate (25.5% vs. 14.5%; difference –11.0%, 95% CI –18.8%, –3.14%), compared with r-hFSH whereas the cumulative live birth rates (CLBR) were reported to be similar: 50.6% (HP-hMG) and 51.5% (r-hFSH) (Witz et al. 2020). Live birth rates (LBR) per embryo transfer in fresh cycles and in frozen cycles were higher for HP-hMG when compared with r-hFSH (fresh cycles: 52.2% vs 48.7%; difference 3.6%, 95% CI –6.4%, 13.4%; frozen cycles: 63.4% vs 50.8%; difference 12.7%, 95% CI –0.9%, 26.2%). In summary, the authors (Witz et al. 2020) claim that HP-hMG was consistently associated with a moderate stimulation profile, lower incidence of complications and pregnancy loss, and corresponding higher probability of ongoing pregnancy and live birth per transfer in fresh or frozen cycles.

We believe there are major methodological flaws in study design, statistical analysis, data collection and reporting, and interpretation of the results which raise serious questions regarding the validity of the authors' conclusions. In this response, we have outlined our main criticisms, subclassifying them under four separate sections referring to the study design, cycle cancellation and protocol violation, statistical analysis and power of the study, and study results and discussion.

I. Lack of robustness in study design

The MEGASET-HR study protocol pre-defined certain elements for the management of the study population (high responder patients) in a way which, as discussed below, is not sufficiently aligned with current evidence and available guidelines, exposing a considerable proportion of the study population in MEGASET-HR to an unnecessary high risk of OHSS.

1. Lack of individualization resulting in a too high r-hFSH starting dose, against good practice guidelines

Overall, it is well known that, during ovarian stimulation in the context of treatment with Assisted Reproductive Technology (ART), the ovarian response depends not only on the woman's age and her ovarian reserve but also on the stimulation protocol applied, namely on the starting dose of gonadotropins, which is associated with a clinically relevant variability in follicular recruitment and oocyte yield (La Marca et al. 2014). Consequently, it has been well established that a significant proportion of patients require different starting doses to reach optimal stimulation, emphasizing the need for an individualized approach (Popovic-Todorovic et al. 2003; La Marca et al. 2013; Papaleo et al. 2016; Allegra et al. 2017). Additionally, for any type of patient population, to achieve a maximal probability for a fresh embryo transfer but also to minimize the risk of OHSS and avoid cycle cancellation while optimizing LBR, the number of retrieved oocytes should be limited to around 15 (Sunkara et al. 2011; Steward et al. 2014), and this can only be achieved by using a personalized FSH starting dose (La Marca, 2013), rather than adopting a "one-size-fits-all" strategy .

It is clear that a significant subgroup of patients in the MEGASET-HR study had a very high risk for OHSS based on inclusion criteria (Antral Follicle Count [AFC] levels >10 and Anti-Mullerian Hormone [AMH] serum levels ≥5 ng/mL) and high mean baseline values for both AFC (30.7) and AMH (7.7 ng/mL). Indeed, the OHSS risk is well known to be increased when: serum AMH >3.6 ng/mL or AFC >24 (Lee et al. 2008; Jayaprakasan et al. 2012); when AMH >3.4 ng/mL or AFC >24 in PCOS patients (Practice Committee of the American Society for Reproductive Medicine 2016); when AMH >3.5 ng/mL or AFC ≥20 (Lunenfeld et al. 2019); or when more than 18 follicles ≥11 mm in size are present on day of oocyte maturation trigger and/or ≥18 oocytes are collected (European Society of Human Reproduction and Embryology 2019),

Knowing this, for this high responder population included in MEGASET-HR (women younger than 35 years old [mean age 30.2 years] presenting a mean AMH level of 7.7 ng/mL, mean AFC of 30.7 and an average Body Mass Index [BMI] of 24.3 kg/m^2 [amongst other baseline parameters reported in Table 1 of Witz et al. 2020]), we can state that a starting dose lower than 150 IU r-hFSH should have been used to optimize patients' clinical and safety outcomes, based on existing knowledge about specific patients' characteristics and related biomarkers, when individualizing the gonadotropin starting dose (La Marca 2013; Lunenfeld et al. 2019; Popovic-Todorovic et al. 2003; Griesinger et al. 2016; Yovich et al. 2016; European Society of Human Reproduction and Embryology

2019). These baseline characteristics and related biomarkers have been the basis of several published nomograms and algorithms that can be used to calculate the starting dose of r-hFSH, based on:

- Age, basal FSH and two markers of ovarian reserve (AMH or AFC) (La Marca et al. 2012 and La Marca 2013) These algorithms had been published and were available to the MEGASET-HR investigators before the MEGASET HR study was registered on clinical trials.gov in 2015.
- Age, BMI, AMH and AFC and BMI (Magnusson et al. 2017)
- Age, AFC, ovarian volume and ovarian stromal blood flow and smoking status (Popovic-Todorovic et al. 2003).

According to our calculations, the individualized r-hFSH starting dose range for the MEGASET-HR population, should have been between 75-120 IU according to three available nomograms (La Marca et al, 2012; La Marca et al, 2013; Magnusson et al, 2017), taking into account the MEGASET HR inclusion criteria (age 21-35 years old, BMI 18-30 kg/m^2, AFC >10 and AMH serum ≥5 ng/mL) and applying available MEGASET HR study data (age and BMI ranges; mean values of AFC and AMH [were used since no upper limits were available for these variables]) to these nomograms.
More specifically, the individualized r-hFSH starting dose range in the MEGASET-HR study, should have been:

- 103–120 IU, using the La Marca et al. 2012 AMH based nomogram
  Calculations for this dose range were done using the mean AMH level of 7.7 ng/mL. Since the MEGASET HR population included women aged between 21 and 35 years old, to find the lower dose limit, the mean AMH level was correlated with the nomogram lower age limit (25 years).The same was done for the higher dose limit, using in this case the MEGASET upper age limit (35 years) (La Marca et al. 2012);
- 90–106 IU, using the La Marca et al. 2013 AFC based nomogram
  Calculations for this dose range were done using the mean AFC level of 30.7. Since the MEGASET HR population included women aged between 21 and 35 years old, to find the lower dose limit, the mean AFC level was correlated with the nomogram lower age limit (22 years).The same was done for the higher dose limit, using in this case the MEGASET upper age limit (35 years) (La Marca et al. 2013);
- 75–100 IU, using the Magnusson et al. 2017 algorithm

Calculations for this starting dose range were done by correlating the MEGASET HR reported values for age, BMI, AFC and AMH with the applicable ranges used in Magnusson et al. 2017 (Supplementary Table SI) regarding BMI (<19; 19-25; >25), age (<30; 30-35), AFC (10-24; >24) and AMH (>2.95 ng/ml) (Magnusson et al. 2017).

Furthermore, there is good evidence that, in expected high responders, a fixed FSH starting dose of 150 IU has important disadvantages when compared to a fixed FSH starting dose of 100 IU (Oudshoorn et al, 2017): increased rate of hyper-response (38% vs 12%), increased cycle cancellation rate for excessive response (8% vs 2%), and increased OHSS rate (15% vs 5%), while both fresh live birth rates (25% vs 26%) and cumulative live birth rates (39% vs 36%) were comparable, respectively. The only "advantage" observed in the 150 IU FSH starting dose group was a lower cycle cancellation rate due to insufficient response, when compared to the 100 IU FSH starting dose group (3% vs 21%) (Oudshoorn et al, 2017). However, this "advantage" can be explained by the fact that the 100 IU starting dose was fixed for all patients regardless of BMI, not really individualized based on nomograms (La Marca et al, 2012; La Marca et al, 2013; Magnusson et al, 2017), allowing a range of starting doses below 150 IU (i.e. 75-120 IU as calculated above), and gonadotropin dose increase during ovarian stimulation was not allowed in patients with a low response at day 5 or 6 of ovarian stimulation. We hypothesize that this "advantage" would have disappeared if dose increase had been allowed in selected patients with low response during the first 5-6 days of ovarian stimulation.

In the MEGASET-HR study (Witz et al. 2020), the fixed r-hFSH starting dose of 150 IU was considerably higher (25% higher versus 120 IU, 100% higher versus 75 IU) than the 75-120 IU range of possible starting doses recommended in high responder populations calculated according to the three nomograms described above (La Marca et al, 2012; La Marca et al, 2013; Magnusson et al, 2017), higher than the <150 IU starting dose recommended by the ESHRE guidelines for Ovarian Stimulation during ART treatment (European Society of Human Reproduction and Embryology 2019), and higher than starting dose recommendations for r-hFSH in high responders published in recent papers (Lunenfeld et al. 2019; Velthuis et al. 2020).

In our opinion, the MEGASET HR study investigators were aware that a 150 IU gonadotropin starting dose of r-hFSH was likely to be too high for this high responder population, based on insights derived from previous Ferring studies: MERIT (Andersen et al. 2006) and MEGASET (Devroey et al, 2012), performed in normal responders. According to these studies, the optimal r-hFSH starting dose is 150 IU in normal responders (MEGASET, Devroey et al, 2012), avoiding the risks of increased serum progesterone rise and of OHSS associated with a 225 IU r-hFSH starting dose (MERIT, Andersen et al, 2006). It is unclear why the MEGASET-HR investigators did select a starting dose of 150 IU r-hFSH, appropriate for a normal responder population with baseline mean AMH concentration of 3.78 ng/ml and AFC value of 15.7 (MEGASET, Devroey et al, 2012), but not for a high responder patient population where baseline mean AMH concentration (7.7 ng/mL) and mean AFC value (30.7) were twice as high (MEGASET-HR, Witz et al, 2020) and where the increased OHSS risk was predictable, against the recommendations made in published nomograms (La Marca et al, 2012; La Marca et al, 2013; Magnusson et al, 2017), and against international guidelines (European Society of Human Reproduction and Embryology 2019).

The 150 IU gonadotropin starting dose used in high responders in the MEGASET-HR study has different implications for r-hFSH when compared to HP-hMG. In normal responders, multiple studies have shown that, using the same starting dose in both treatment arms, ovarian stimulation with r-hFSH is more effective than HP-hMG with respect to follicular development, number of retrieved oocytes (up to 1-2 more in r-hFSH group as can be seen in Table 1 below) and number of embryos (Bosch et al, 2008; Hompes et al. 2008; Andersen et al, 2006; Balasch et al, 2003; Devroey et al. 2012). This can be explained by the fact that, based on *in vivo* bio- and immuno-assays and other biological tests, r-hFSH shows higher specific bioactivity (biopotency/µg of proteins), higher purity, and lower batch-to-batch variability than HP-hMG (Lispi et al. 2006; Bassett et al. 2009; Leão Rde et al. 2014).
Considering this, in the high responder population included in the MEGASET study (Witz et al, 2020), the too high r-hFSH starting dose of 150 IU, combined with the higher biopotency of r-hFSH, not only led to a significantly higher number of oocytes and embryos in the r-hFSH group (as has been reported in all RCTs comparing r-hFSH with HP-HMG, Table 1), but also exposed more patients in the r-hFSH group to an increased risk of OHSS, which could have been avoided by a lower r-hFSH starting dose.

Table 1: Mean (± standard deviation) number of oocytes and embryos obtained in RCTs comparing ovarian stimulation with HP-HMG vs r-hFSH

	HP-hMG	**r-hFSH**
Witz et al. 2020, MEGASET HR, high responder patient population (mean AMH 7.7 ng/mL and AFC 30.7); HP-hMG 150 IU; r-hFSH 150 IU		
Oocytes	15.1 ± 10.1	22.2 ± 11.5
Fertilized oocytes	8.2 ± 5.9	12.9 ± 7.4
Embryos (Day 5)	5.6 ± 4.3	8.5 ± 5.7
Devroey et al. 2012, MEGASET, normal responder population (mean AMH 3.78 ng/mL and AFC 15.7); HP-hMG 150 IU; r-hFSH 150 IU		
Oocytes	9.1 ± 5.2	10.7 ± 5.8
Embryos (Day 3)	4.0 ± 3.0	4.8 ± 3.7
Embryos (Day 5)	2.7 ± 2.5	3.1 ± 3.0
Nyboe Andersen, et al. 2006, MERIT normal responder population (mean AFC 10.9); HP-hMG 225 IU; r-hFSH 225 IU		
Oocytes	10.0 ± 5.4	11.8 ± 5.7
Embryos (Day 3)	6.3 ± 4.7	7.4 ± 5.0
Hompes et al, 2008; Normal responder population HP-hMG 150 IU; r-hFSH 150 IU		
Oocytes (SD not available)	7.8	10.6
Fertilized oocytes (SD not available; N embryos D3 not available)	4.2	5.6
Bosch et al, 2008, Normo-ovulatory population;		

HP-hMG 225 IU; r-hFSH 225 IU		
Oocytes	11.3 ± 6.0	14.4 ± 8.1
Embryos	Not available	Not available
Balasch et al, 2003, Normo-ovulatory population HP-hMG 150 IU; r-hFSH 150 IU		
Oocytes	9.1 ± 0.9	11.8 ± 0.9
Embryos (day 2)	4.9 ± 0.5	7.9 ± 0.7

Data are mean ± standard deviation

In summary, taking into account the increased risk for OHSS in high responders; the insights gained in the MERIT and MEGASET studies; the existing nomograms that allow a calculation of r-hFSH starting dose range based on age, AMH and AFC; the current international guidelines recommending a <150 IU r-FSH starting dose in high responders; and the higher biopotency of r-hFSH compared with HP-hMG, we firmly believe that the 150 IU r-hFSH starting dose used in the MEGASET-HR study was too high, exposing patients to a predictable increased risk of exaggerated ovarian response and related OHSS. Therefore, in such a patient population of high responders, study participants should have received a r-hFSH starting dose <150 IU to reduce the incidence of OHSS, while optimizing follicular stimulation with no detrimental effect on clinical outcomes (European Society of Human Reproduction and Embryology 2019; Lunenfeld et al. 2019; Velthuis et al. 2020).

2. Too restrictive criteria to allow ovulation triggering with GnRH agonists followed by 'freeze all' only in patients with extreme ovarian response, against current good practice guidelines

Taking into account current global consensus, a high ovarian response to ovarian stimulation is present if more than 18-25 follicles >11-12 mm in size are present on the day of ovulation trigger, and/or if more than 18-25 oocytes are obtained at egg retrieval, and this high ovarian response is associated with an increased OHSS risk, which can be prevented by using GnRH agonist triggering followed by egg retrieval and freezing of all embryos (Practice Committee of the American Society for Reproductive Medicine 2016; European Society of Human Reproduction and Embryology 2019).

According to the ICMART, high ovarian response is defined as *"an exaggerated response to ovarian stimulation characterized by the presence of more follicles than intended. Generally, more than 20 follicles >12 mm in size and/or more than 20 oocytes collected following ovarian stimulation are considered excessive, but these numbers are adaptable according to ethnic and other variables"* (Zegers-Hochschild et al. 2017).
The ESHRE guidelines for Ovarian Stimulation during ART treatment define high ovarian response to be *"characterized by generally having more than 18 follicles ≥11 mm in size on day of oocyte maturation trigger and/or 18 oocytes collected (Griesinger et al. 2016) and defined by a risk increase in OHSS*" (European Society of Human Reproduction and Embryology 2019). For management of high responders, they recommend that "*applying a GnRH agonist trigger is certainly a way to improve safety. Finally, prevention of pregnancy derived hCG by freezing all embryos will be another logical step"* (European Society of Human Reproduction and Embryology 2019).
According to ASRM guidelines, there is fair evidence (level II-2) that PCOS, elevated AMH and AFC, increased peak estradiol (E_2) levels, multi-follicular development, and a high number of oocytes retrieved, increase the risk of OHSS in patients with predicted high response (AMH values >3.4 ng/mL, AFC >24, estradiol values >3500 pg/mL, development of ≥25 follicles, or ≥24 oocytes retrieved). For this population, ASRM recommends the *"use of a GnRH agonist to trigger oocyte maturation prior to oocyte retrieval in order to reduce the risk of OHSS if peak estradiol levels are high or multi follicular development occurs during stimulation"*, and highlights that there is *"fair evidence that cryopreservation prevents OHSS"* (Practice Committee of the American Society for Reproductive Medicine 2016).
Before these guidelines were developed, the use of GnRH antagonist protocols with a GnRH agonist to trigger ovulation, followed by a freeze all approach was already recommended by other investigators (Humaidan et al, 2014; Boothroyd et al, 2015; Smith, Osanlis and Vollenhoven, 2015) in case of AMH levels >3.36 ng/mL, and/or AFC ≥24 or the presence of >14 follicles with a diameter of 11 mm (Smith, Osanlis and Vollenhoven, 2015; Lee et al. 2008; Jayaprakasan et al. 2012).

In the MEGASET-HR study, GnRH agonist triggering followed by freeze all was only allowed for patients with an extreme ovarian response, classified as having >30 follicles with ≥12 mm diameter and/or serum E_2 levels ≥ 5000 pg/mL, and ultimately only applied in

54 patients receiving r-hFSH and 37 patients receiving HP-hMG. Despite the fact that all MEGASET HR study participants were high responders with high mean baseline levels for both AFC (30.7) and serum AMH (7.7 ng/mL), in patients with a high but not extreme ovarian response (those with 18-30 follicles with ≥12 mm and/or serum E2 levels between 3500-5000 pg/mL), physicians had no other option than the use of a standard dose of human chorionic gonadotropin (hCG) for triggering, followed by a fresh transfer according to MEGASET HR protocol. This study design choice is in violation with ASRM guidelines (Practice Committee of the American Society for Reproductive Medicine 2016), that recommend agonist triggering and freeze all in patients with predicted high ovarian response (AMH values >3.4 ng/mL, AFC >24, estradiol values >3500 pg/mL, development of ≥25 follicles), similarly to other international guidelines (European Society of Human Reproduction and Embryology 2019).
Considering this, we do not understand why the MEGASET HR investigators (Witz et al, 2020) exposed a number of study participants to an unnecessary high risk of OHSS, by allowing agonist triggering and freeze all only in extreme responders, denying this option to other eligible high responders according to criteria listed by ASRM (Practice Committee of the American Society for Reproductive Medicine 2016), also in line with other evidence available before patient recruitment started for the MEGASET HR study (Humaidan et al, 2014; Boothroyd et al, 2015; Smith, Osanlis and Vollenhoven, 2015).

It is now important for the medical-scientific community to have access to essential data that were not published in the MEGASET-HR study (Witz et al. 2020). More concretely, we would like to request the authors of the MEGASET HR study (Witz et al. 2020) to report, as provided in the MEGASET-HR protocol, the proportion of patients with >18 follicles ≥11 mm in size (as per ESHRE guidelines for Ovarian Stimulation during ART treatment), and ≥25 follicles ≥11–12 mm in size (as per ASRM guidelines), and how many of these patients received hCG triggering. We also request further details of how OHSS risk was further assessed in these patient subgroups. This information is essential to understand the number of patients who were at high risk of OHSS (even though they did not meet the MEGASET HR criteria for extreme response) and how many of them received hCG triggering followed by a fresh transfer, without freeze all.

Another demonstrated benefit of a freeze all strategy in high responders, is that it can help overcome the detrimental impact of ovarian hyper-response on endometrial receptivity.

Gonadotropin stimulation may lead to endometrial receptivity changes that, in turn, may affect reproductive outcomes in fresh cycles. For example, a high ovarian response with increased serum E_2 and P levels can result in an unfavorable endometrium for embryo implantation in fresh cycles, which can be overcome with a freeze all approach (Munch et al. 2017).
Consequently, we can infer that the higher ovarian response in MEGASET-HR in the r-hFSH versus HP-hMG group (reflected by the higher mean serum E_2 peak [3201 ± 2003 pg/mL vs 2809 ± 1783 pg/mL] and P levels [1.0 ± 0.9 ng/mL vs. 0.7 ± 0.9 ng/mL] on the day of hCG triggering), could have contributed to sub-optimal endometrial receptivity in r-hFSH-treated patients. Indeed, we speculate that the higher early pregnancy loss rates observed in the r-hFSH arm can, in part, be explained by the likely impairment of endometrium receptivity due to higher E_2 and P levels, in addition to the inappropriately high starting dose, the overall higher ovarian response, and the too restrictive criteria for GnRH agonist triggering.

In summary, the design of the MEGASET HR study exposed its participants to an increased OHSS risk and also to a possibly elevated risk of spontaneous miscarriage, by restricting the use of GnRH agonist triggering and freeze all to extreme responders only, which is not in line with ASRM practice recommendations, ESHRE guidelines, and other evidence available at the time of the study

II. Lack of details regarding cycle cancellation and protocol violations

There is lack of clarity in the MEGASET-HR publication regarding patient flow and reasons for cycle cancellation, which may have biased some of the study outcomes reported (Witz et al. 2020).

The supplementary material (Supplemental Figure 2 of Witz et al. 2020) contains details showing that the treatment discontinuation rate between Day 6 of ovarian stimulation and embryo transfer, was about twice as high in the HP-hMG treatment arm when compared with r-hFSH treatment (32/308 [10.3%] vs 18/309 [5.8%] patients, respectively). Apparently, the majority of these patients in the HP-hMG group discontinued treatment due to ‘protocol violation/cycle cancellation’ (Supplemental Figure 2 of Witz et al. 2020)

and the discontinuation rate due to this reason was three times higher in the HP-hMG group than in the r-hFSH patient group (21/308 [6.8%] vs 7/309 [2.3%] patients, respectively). However, the authors provide no explanation regarding the nature/content of these “protocol violations” nor the rationale for “cycle cancellations”, but we speculate that they may be related to either a suboptimal or an exaggerated ovarian response. We call upon the authors to provide further details about the exact nature of protocol violations, and the exact reasons for cycle cancellations.

When one examines the patient numbers in more detail it becomes apparent that, aside from the 32 HP-hMG patients and 18 r-hFSH (50 patients in total) that are described in Supplemental Figure 2 of Witz et al. 2020 as discontinuing participation, there is a considerably higher number of patients who are lost to follow-up. In **Table 2** below, based on the data published in the MEGASET HR study (Witz et al, 2020), we have summarized the number of patients at each stage of ART treatment, enabling us to calculate the number of patients lost to follow-up at each stage. According to our calculations, there were a total of 138/619 “missing “patients (73/310 [23%] in the HP-hMG group and 65/309 [21%] in the r-hFSH group).

Table 2: Number of patients at each stage as detailed in MEGASET-HR

	Described in manuscript by MEGASET HR authors [Supplemental Figure 2, Witz et al, 2020]		**Number Calculated as “loss to follow-up” by authors of this Fertil Steril Dialog paper**	
	HP-hMG	**r-hFSH**	**HP-hMG**	**r-hFSH**
Screened	1258			
Randomized	620			
Randomized and treated	619		1	
mITT	310	309	0	0
Completed Day 6 of stimulation	308	309	2	0
Received trigger **GnRH agonist**	293 37	307 54	15	2

	Described in manuscript by MEGASET HR authors [Supplemental Figure 2, Witz et al, 2020]		Number Calculated as "loss to follow-up" by authors of this Fertil Steril Dialog paper	
	HP-hMG	**r-hFSH**	**HP-hMG**	**r-hFSH**
hCG	256	253		
Oocyte retrieval	292	306	1	1
Fresh embryo transfer	201	191	55	62
Total calculated loss to follow-up			73	65

Data are number of patients except for oocyte retrieval

Examining these data stage-by-stage, it is evident that despite similar starting dose and opportunities for dose adjustment in both groups, the proportion of patients with cancelled treatment upon ovulation triggering was about 10 times higher in the HP-hMG group than in r-hFSH group (17/310 [5.5%] vs 2/309 [0.6%] patients, respectively [**Table 2 above**]). However, the reasons for cycle cancellation were not reported in the manuscript, nor was a description of appropriate criteria for cycle cancellation pre-defined in the protocol, to the best of our knowledge. More specifically, among patients who completed Day 6 of stimulation, the proportion of those not receiving any ovulation triggering was seven times higher in the HP-hMG group (15/308 patients [4.9%]) than in the r-hFSH group (2/309 patients [0.7%]).

This brings us to our main concern, which relates to the difference in the number of patients who were triggered with hCG versus those who subsequently received a fresh embryo transfer. The study methods state that "*patients receiving an hCG trigger underwent fresh transfer of a single blastocyst of best quality by morphology on day 5, following ICSI*", suggesting that all patients receiving an hCG injection followed by oocyte retrieval also had a fresh embryo transfer. However, as can be seen from our summary table (**Table 2** above), there is a considerable loss to follow-up in both treatment groups. In the HP-hMG group, 256 patients received the hCG trigger but only 201 of these patients (201/256 [79%]) underwent fresh embryo transfer. In the r-hFSH group, 253 patients

received the hCG trigger but only 191 of these patients (191/253 [75%]) underwent fresh embryo transfer. No explanations were provided as to why these 55/256 [21%] patients from the HP-hMG group, and 62/253 [25%] patients from the r-hFSH group did not receive an embryo transfer after hCG triggering. While we speculate that the reason to deny an embryo transfer to a quarter of all patients after HCG triggering can be related to the lack of availability of a blastocyst and/or to perceived increased OHSS risk, we call on the authors to provide the exact reasons as to why embryo transfer did not take place in these patients, as this may constitute a relevant protocol violation. This information is essential to assess possible bias in the selection of patients receiving a fresh embryo transfer, which may have affected reproductive outcomes such as pregnancy rate and (C)LBR.

In summary, we are concerned about the lack of clarity regarding causes for treatment discontinuation, rationale for cycle cancellation and nature/content of protocol violations, particularly between day 6 of ovarian stimulation, hCG triggering, oocyte retrieval, and embryo transfer, which may represent a potential bias in the study that could have impacted reproductive outcomes and overall study conclusions.

III. Issues with statistical analysis and power of the study

MEGASET-HR was designed as a non-inferiority trial, comparing ovarian stimulation of HP-hMG with r-hFSH and selected the clinical pregnancy rate per fresh embryo transfer as primary outcome. Although the sample size calculation required 275 patients receiving an embryo transfer per group, the study was underpowered as only 191/309 (62%) patients in the r-hFSH group and 201/310 (65%) patients in the HP-hMG group received a fresh embryo transfer. Furthermore, as the study was neither designed nor powered to calculate differences across treatment arms for pre-specified secondary outcomes, the authors need to acknowledge that any conclusions drawn from these secondary outcomes data can only be considered as hypothesis-generating. Moreover, no adjustment for multiple comparisons was reported in the manuscript. Therefore, any significant test could be explained by type 1 error.

Finally, the lack of clarity regarding denominators used to calculate the different study outcomes reduces the reliability of the data and limits the possibility to comprehensively discuss and compare outcomes. As an example, in Figure 1 of Witz et al. 2020, several study outcomes were displayed such as: ongoing pregnancy rate per cycle start, live birth rate per fresh embryo transfer and FET and CLBR without having a clear reference to

what denominator was used. Thus it is not clearly stated if the outcomes are calculated per initiated cycle, oocyte pick-up or embryo transfer. As final note on this topic, regarding the primary endpoint, there are additional concerning inconsistencies, as the sample size calculation was done with "ongoing pregnancy rate for fresh cycles" but however, in the protocol, under efficacy and safety outcome is stated "ongoing pregnancy rate/cycle start after the fresh IVF cycle".

IV. Issues with study results and discussion

Certain aspects of the results reported and discussed in MEGASET-HR could have been impacted by design or execution bias and therefore require a more thorough analysis, which we provide below. This analysis is also important to explain why the clinical relevance and applicability of some of the data and statements presented in the MEGASET-HR study (Witz et al. 2020) may be questionable, as they are not aligned with available clinical evidence nor existing guidelines for this group of patients (high responders). To put the results in context, we suggest that these points should ideally have been highlighted in the discussion and limitations section of the original manuscript.

1. Analysis and reporting issues with data from frozen embryo transfers (FET)

It is not possible to properly evaluate outcomes, nor infer conclusions, from the FET cycles with the data provided by the authors, since there are significant issues with how the data were analysed and reported.

The treatment selection and patient allocation in the FET cycles were not part of the randomized study design, implying that the reproductive outcomes resulting from FET can only be considered as post-trial and post-hoc outcomes. Furthermore, the results of FET cycles should not be considered independently; the data should be integrated into an analysis of cumulative ongoing pregnancy rate and CLBR per patient in the modified intent-to-treat population following life table analysis. It is not acceptable to present cumulative ongoing pregnancy rates and CLBR just by counting together the outcomes of fresh embryo transfers and FETs, as presented in the MEGASET-HR study results (Witz et al, 2020).

Finally, essential demographic, baseline and ART treatment information regarding cryopreservation results is missing for the following three subgroups: 1) patients who received hCG trigger, oocyte retrieval and fresh embryo transfer; 2) patients who received

hCG trigger and oocyte retrieval but no fresh embryo transfer; 3) patients who received GnRH agonist trigger and oocyte retrieval but no frozen embryo transfer. We call upon the authors to provide this missing information to enable a comprehensive and transparent analysis of FET cycles. These additional data are needed to better understand potential bias impacting the outcome of FET cycles, related to patient selection, treatment variables during FET cycles, and/or other unreported factors, and will hopefully provide the he basis for more robust conclusions than those currently presented in the paper (Witz et al, 2020).

2. **Follow-up not long enough to allow CLBR calculation**

According to the MEGASET-HR study design, data on CLBR were collected for up to three FET cycles within 6 months of randomization. We believe that this follow-up period is too short to use the full cryopotential of frozen blastocysts and therefore to adequately assess cumulative pregnancy/live birth outcomes. Additionally, the mean number of transfer cycles (number of total transfers/number of patients with oocyte retrieval) in the r-hFSH and HP-hMG groups was 1.09 and 0.99, respectively, with an overall mean number of transfer cycles of ~1 per patient. This signals that there was not a real "cumulation" of transfers, but rather a summation of first transfer cycles (fresh or frozen) that took place in most patients, followed by an uncontrolled and unspecified number of patients with subsequent FET cycles. Therefore, the results presented as “CLBR” in the MEGASET-HR paper (Witz et al 2020) do not reflect the true CLBR and do not document the true cryopotential of all frozen blastocysts.

Although a clear protocol for analysis was not provided, one would have expected additional details about the availability of euploid embryos and blastocysts, and a description about the FET policy, if present. From our own analysis, considering the stronger ovarian response observed in the r-hFSH when compared with the HP-hMG group (**Table 3** below) and knowing that the total euploid blastocyst number was ~1906 in r-hFSH and ~1308 in HP-hMG groups (calculated based on the provided mean number of blastocysts and the reported aneuploidy rate), we conclude that there would be ~600 more euploid blastocysts available in the r-hFSH group than in the HP-hMG group.

Table 3: MEGASET HR study: stronger ovarian response to lower total gonadotropin dose in r-hFSH group than in HP-hMG group (all values expressed as mean values published in Witz et al, 2020 [Table2])

	r-hFSH	HP-hMG
Lower total gonadotropin dose (Δ=615.6 IU)	1498.9 IU	2114.5 IU
Shorter duration of ovarian stimulation (Δ=1.6 days)	9.3 days	10.9 days
More oocytes (47% relative increase)	22.2	15.1
More metaphase II oocytes (57% relative increase)	15.9	10.1
More 2PN fertilized oocytes/patient (59% relative increase)	12.9	8.1
Higher fertilization rate (7% relative increase)	59%	55%
More blastocysts/patient (52% relative increase)	8.5	5.6
More excellent blastocysts/patient (30% relative increase)	3.9	3.0

Only a longer follow-up, with transparent reporting on how all frozen and transferable blastocysts were handled, allows the evaluation of the real cryopotential and the true calculation of the CLBR, as well as the number of remaining frozen blastocysts after a live birth had been achieved. It is unclear why the follow-up for CLBR was restricted to 6 months, as there was enough time between when the study was conducted (August 2015 – February 2018) and the publication of the results (May 2020) to facilitate a considerably longer CLBR follow-up period. The authors recognize this in the discussion: "*An open question is whether collection of live birth data from frozen transfers for longer periods of time could have revealed differences.*" Therefore, we ask the authors to now share this information (as raw data), to enable the calculation of the real potential for both treatment options and to facilitate a better comparison between treatments.

3. **Lack of detail regarding early pregnancy loss data, and questionable validity of the preimplantation genetic testing for aneuploidy (PGT-A) assessment**

As no details about patient profiles in embryo transfer cycles were reported in the MEGASET HR study, it is difficult to properly comment on potential explanations regarding early pregnancy loss rates. The lower proportion of patients receiving a FET cycle in the HP-hMG group (82/310 [26.5%]) compared to the r-hFSH group (130/309 [42.1%]), coupled with the absence of any reported baseline patient characteristics for these patients, make it inappropriate to directly compare early pregnancy loss outcomes across treatment groups using only the data provided. Following these remarks, we call upon the authors to provide additional details (including both baseline and ART treatment

characteristics) of all patients who underwent fresh and FET cycles and report additional parameters (corresponding number of retrieved oocytes and E2 and P values) to help clarify the reasons for higher proportions of patients experiencing early pregnancy loss in the r-hFSH versus HP-hMG groups and to enable a better interpretation of the results.

We were also surprised to observe (in both treatment arms) that the early pregnancy loss rate was similar after fresh transfers and after FET cycles in each group (18/126 patients [14.0%] and 9/62 [14.5%] respectively for HP-hMG group; 29/122 [23.8%] and 26/91 [28.6%] respectively for r-hFSH group), despite the fact that genetical untested blastocysts were transferred in fresh cycles and only euploid blastocysts were transferred during FET cycles. This unexpectedly high early pregnancy loss observed after transfer of euploid blastocysts in FET cycles is not concordant with an early pregnancy loss lower than 10% reported in the literature. Indeed, the early pregnancy loss rate was only ~5% in an analysis of >1800 FET cycles (where FET was performed using preimplantation genetic testing for aneuploidy [PGT-A] to guide embryo transfer [as occurred in MEGASET-HR]) (Simon et al. 2018). Similarly, the miscarriage rate for patients <35 years old was only 9.2% in a retrospective cohort study assessing the predictive value of patient characteristics, controlled ovarian stimulation and embryological parameters on live birth outcome after of 707 single euploid blastocyst transfers during FET cycles after PGT-A, (Boynukalin et al. 2020).

A very important additional bias affecting early pregnancy loss risk after FET cycles reported in the MEGASET-HR study (Witz et al, 2020) is introduced by the higher than anticipated error rate in the real time polymerase chain reaction (rt-PCR) assay used for PGT-A, due to reagent issues. This higher error rate may have reduced the overall reliability of PGT-A results, allowing embryo transfer of genetically abnormal embryos leading to increased early pregnancy loss. As the distribution of this error is unknown, not necessarily balanced between both groups, and the number of patients receiving a frozen transfer during a FET cycle was higher in the r-hFSH group (130/309, 42%) than in the HP-hMG group (82/310, 26.5%), it cannot be excluded that the higher pregnancy loss rate in the r-hFSH group is caused by a higher proportion of aneuploid blastocysts transferred, as PGT-A results were not sufficiently reliable.

In summary, the lack of relevant baseline and treatment data provided, that could impact the risk of early pregnancy loss, in addition to significant quality issues observed with the rt-PCR assay of the PGT-A assessment which may have biased outcomes from FET cycles, mean that no robust conclusions regarding early pregnancy loss across treatments should be made.

4. **OHSS results not discussed in the right context**

One of the most serious iatrogenic complications of controlled ovarian stimulation is OHSS. Its early identification and proper management have paramount importance.
As high responders, all patients started this study (Witz et al. 2020) at an elevated risk for OHSS compared with a more general population. Upon examination of mean baseline clinical characteristics, however, it is reasonable to suggest that some patients may have been at an even higher risk than others. For example, it has been shown that OHSS risk increases more meaningfully with AFC >14 (Kwee et al. 2007); in the MEGASET-HR study, the mean AFC was 30.7. Similarly, an AMH cut-off value of 3.36 ng/mL has previously been identified as a good predictor of OHSS (sensitivity of 90.5%, specificity of 81.3%) (Lee et al. 2008); the mean AMH was 7.7 ng/mL in MEGASET-HR. Furthermore, AMH >3.9 ng/mL has been shown to be increase the probability of a patient having PCOS (Sahmay et al. 2013); in MEGASET-HR, as PCOS was not specified as an exclusion criterion of the study, we then assume that patients with PCOS were included, and speculate that they represent the majority of the 106/619 (17%) patients with oligo-ovulation as cause of infertility (Table 1 reported in Witz et al, 2020) included in the study population and PCOS patients are well known to have an increased risk for OHSS after ovarian stimulation.

As previously described, current guidelines recommend reduction of the r-hFSH starting dose for ovarian stimulation (European Society of Human Reproduction and Embryology 2017 and 2019) and the use of an agonist triggering protocol followed by freeze all (European Society of Human Reproduction and Embryology 2017 and 2019; Practice Committee of the American Society for Reproductive Medicine 2016), to better manage high risk patients and reduce the risk of OHSS.

In MEGASET-HR, the overall OHSS rate was lower in the HP-hMG group (30/310 patients [9.7%]) than in the r-hFSH group (66/309 patients [21.4%]). This observation can be

explained by numerous different reasons we have already addressed at length in the previous sections. They can be summarized as follows: a too high r-hFSH starting dose, a higher biopotency of r-hFSH compared to HP-HMG resulting in a stronger ovarian response, a too restrictive ovulation trigger policy limiting the use of GnRH agonist trigger and freeze all to extreme responders only (defined as >30 follicles of ≥12 mm each and/or E_2 levels ≥5,000 pg/mL), and the high proportion of patients (~one-quater of population: 55/256 HP-hMG and 62/253 r-hFSH patients) who were triggered with hCG, had oocyte retrieval, but did not receive a fresh embryo transfer with no explanation provided.

There were also risk factors for OHSS that were present to a higher degree in the r-hFSH group than in the HP-hMG group. For example, the ESHRE guidelines for Ovarian Stimulation during ART treatment and other published evidence state that E_2 levels >3000 pg/mL significantly increase the risk of OHSS (European Society of Human Reproduction and Embryology 2017 and 2019; Abdallah et al. 2010; Levinsohn-Tavor et al. 2003). In MEGASET-HR, the mean serum E_2 level on the day of ovulation triggering was higher (3201 pg/mL) in the r-hFSH arm than in the HP-hMG arm (2809 pg/mL). Similarly, the ESHRE guidelines for Ovarian Stimulation during ART treatment and other evidence show that risk of OHSS significantly increases with the number of oocytes, particularly if >18 oocytes are developed and retrieved (Griesinger et al. 2016; Papanikolaou et al. 2006; European Society of Human Reproduction and Embryology 2019). In MEGASET-HR, the mean number of oocytes retrieved was much higher in the r-hFSH arm (n=22) than in the HP-hMG arm (n=15). In addition, it appeared that in the r-hFSH group a slightly higher number of oligo-ovulatory patients was present (56/309 [18%]) when compared with the HP-hMG group (50/310 [16%]), which is a sub-population of PCOS-like high responders, known to have an increased risk of OHSS (European Society of Human Reproduction and Embryology 2017).

While the overall OHSS rate should always be reported, it is the severe OHSS rate which is perhaps the more relevant safety outcome, as it is associated with patient hospitalization and will therefore likely have the most impact on patient health. Severe OHSS was reported in similarly low proportions of patients in both the r-hFSH (9/309 patients [2.9%]) and the HP-hMG (8/310 patients [2.6%]) groups, demonstrating an equally good safety profile with either treatment.

The low rate of both OHSS and severe OHSS with r-hFSH treatment has been demonstrated in a recent systematic review of 45 clinical trials (including high responder patients and those with PCOS), reinforcing its proven safety profile from clinical practice and previous studies. Overall 272 cases of OHSS were reported (5190 per 100,000 treatment cycles [5.19%]), only 10 of which were classified as severe (191 per 100,000 treatment cycles [0.19%]) (Velthuis et al. 2020).

Conclusion

The goal of assisted reproductive technology is to enable the parents' wish to have a healthy baby while ensuring safety. Every attempt should be made to put safety first and utilize the best personalized strategy to achieve the birth of single healthy baby. With this in mind, appropriately designed, executed and analyzed randomized controlled trials (RCTs) are the gold standard for comparing treatment strategies, answering clinically meaningful questions and helping define standards of care. Thus, great attention and scrutiny should be used when evaluating RCTs in terms of study design, validity, statistical analysis, and applicability of results to routine clinical practice and in accordance with clinical guidelines. In our opinion, many significant issues addressed by us here regarding the quality of the MEGASET-HR trial do question the conclusions made by the MEGASET-HR study authors (Witz et al, 2020), and do not support their claim that HP-hMG has a more optimized risk: benefit profile than r-hFSH.

We have detailed several methodological and statistical inaccuracies in the MEGASET HR study and have highlighted where the study design and related data are not fully in line with current guidelines and clinical practice for this patient population of high responders. Our main concerns are:

- too high starting dose of r-hFSH, considering the high OHSS risk of the patient population of high responders, and the higher biopotency of r-hFSH compared to HP-hMG
- too restrictive criteria for GnRH agonist triggering and 'freeze all' in patients at risk for OHSS, only allowing the inclusion of patients with extreme ovarian response, in violation of existing practice recommendations
- lack of explanation regarding treatment discontinuation, i.e. rationale for cycle cancellation, content/nature of protocol violations
- problems related to statistical analysis and power calculation

- lack of complete data and lack of transparency in data reporting and data analysis related to FET cycles
- inadequate reporting of CBLR and follow-up too short for proper assessment of CLBR
- lack of enough essential data with possible impact on risk of early pregnancy loss,
- questionable validity of the PGT-A assessment used in FET cycles
- lack of context provided to explain data related to OHSS

All these issues could have been addressed in the study design and/or protocol, should have been reported in a transparent way, and at a minimum would have deserved a description as limitations of the MEGASET HR study. The discussion section of the manuscript was, in our opinion, overly focused on secondary outcomes such as early pregnancy loss and OHSS with associated statistical analysis. As the trial was neither designed nor powered to calculate differences across arms for secondary outcomes, any conclusions should only be considered as hypothesis-generating and not a demonstration of non-inferiority/superiority. As such, the clinical relevance and applicability of some of the data analyzed and related statements presented in the MEGASET HR study are, at present, questionable. We, therefore, in the spirit of post publication peer review, call upon the authors to address our concerns and provide the missing information and data. We have summarized our requests in **Table 3**.

Table 3: Request for additional information in relation to the MEGASET-HR study

Topic	Request
Definition of increased OHSS risk (in relation to triggering and freeze all protocol)	• The proportion of patients with >18 follicles ≥11 mm in size (as per ESHRE guidelines for Ovarian Stimulation during ART treatment), and ≥25 follicles ≥11–12 mm in size (as per ASRM guidelines) • Clarification on how OHSS risk was further assessed in these patient subgroups
Cycle cancellations and protocol violations	• Additional details regarding reasons for discontinuation, specifically for: o The 32 HP-hMG patients and 18 r-hFSH patients who completed Day 6 of stimulation but did not have a fresh embryo transfer ▪ We are particularly interested in the 21 HP-hMG patients and 7 r-hFSH patients who had a

	‘protocol violation/cycle cancellation’ according to Supplementary Figure 2 o The 17 HP-hMG patients and 2 r-hFSH patients who had their treatment cancelled before ovulation triggering o The 15 HP-hMG patients and 2 r-hFSH patients who completed Day 6 of stimulation but did not receive ovulation triggering o The 55 HP-hMG patients and 62 r-hFSH patients who received hCG trigger but did not undergo embryo transfer
FET cycles	• Information regarding cryopreservation results for: o Patients who received hCG trigger, oocyte retrieval and fresh embryo transfer o Patients who received hCG trigger and oocyte retrieval but no embryo transfer o Patients who received agonist trigger and oocyte retrieval but no embryo transfer • An update on the live birth rate from frozen transfers
Early pregnancy loss	• Provide additional details, including both baseline and assisted reproductive technologies (ART) treatment characteristics of all patients who underwent FET cycles • Separate reporting of patient outcomes (e.g. number of oocytes retrieved in both groups and E_2 and P values) for a better analysis of whether or not there is a link between high ovarian response and poor outcomes and to enable better interpretation of the reasons for pregnancy loss in both trial arms.

While we await the missing information, due to the study’s significant shortcomings in design and execution which do not reflect daily clinical practice for at least some of high responder population, we advise readers to interpret the results of MEGASET-HR with caution.

References

Abdallah R, Kligman I, Davis O, and Rosenwaks Z. Withholding gonadotropins until human chorionic gonadotropin administration. *Semin Reprod Med* 2010: **28**; 486-492.

Allegra A, Marino A, Volpes A, Coffaro F, Scaglione P, Gullo S, and La Marca A. A randomized controlled trial investigating the use of a predictive nomogram for the selection of the FSH starting dose in IVF/ICSI cycles. *Reprod Biomed Online* 2017: **34**; 429-438.

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

Bassett R, Lispi M, Ceccarelli D, Grimaldi L, Mancinelli M, Martelli F, and Van Dorsselaer A. Analytical identification of additional impurities in urinary-derived gonadotrophins. *Reprod Biomed Online* 2009: **19**; 300-313.

Bergh C, Howles CM, Borg K, Hamberger L, Josefsson B, Nilsson L, and Wikland M. Recombinant human follicle stimulating hormone (r-hFSH; Gonal-F) versus highly purified urinary FSH (Metrodin HP): results of a randomized comparative study in women undergoing assisted reproductive techniques. *Hum Reprod* 1997: **12**; 2133-2139.

Boothroyd C, Karia S, Andreadis N, Rombauts L, Johnson N, Chapman M, and Australasian CREI Consensus Expert Panel on Trial evidence (ACCEPT) group. Consensus statement on prevention and detection of ovarian hyperstimulation syndrome. *Aust N Z J Obstet Gynaecol* 2015: **55**; 523-534.

Boynukalin FK, Gultomruk M, Cavkaytar S, Turgut E, Findikli N, Serdarogullari M, Coban O, Yarkiner Z, Rubio C, and Bahceci M. Parameters impacting the live birth rate per transfer after frozen single euploid blastocyst transfer. *PLoS One* 2020: **15**; e0227619.

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

European Society of Human Reproduction and Embryology. Guideline on Ovarian stimulation for IVF/ICSI 2017.

European Society of Human Reproduction and Embryology. Guideline on Ovarian Stimulation for IVF/ICSI 2019.

Griesinger G, Verweij PJ, Gates D, Devroey P, Gordon K, Stegmann BJ, and Tarlatzis BC. Prediction of ovarian hyperstimulation syndrome in patients treated with corifollitropin alfa or rFSH in a GnRH antagonist protocol. *PLoS One* 2016: **11**; e0149615.

Hompes PG, Broekmans FJ, Hoozemans DA, and Schats R. Effectiveness of highly purified human menopausal gonadotropin vs. recombinant follicle-stimulating hormone in first-cycle in vitro fertilization-intracytoplasmic sperm injection patients. *Fertil Steril* 2008: **89**; 1685-1693.

Humaidan P and Alsbjerg B. GnRHa trigger for final oocyte maturation: is HCG trigger history? *Reprod Biomed Online* 2014: **29**; 274-280.

Jayaprakasan K, Chan Y and Islam. Prediction of in vitro fertilization outcome at different antral follicle count thresholds in a prospective cohort of 1,012 women. *Fertil Steril* 2012: **98**(3); 657–663.

Kwee J, Elting ME, Schats R, McDonnell J, and Lambalk CB. Ovarian volume and antral follicle count for the prediction of low and hyper responders with in vitro fertilization. *Reprod Biol Endocrinol* 2007: **5**; 9.

La Marca A, Papaleo E, Grisendi V, Argento C, Giulini S, and Volpe A. Development of a nomogram based on markers of ovarian reserve for the individualisation of the follicle-stimulating hormone starting dose in *in vitro* fertilisation cycles. *BJOG* 2012b: **119**; 1171-1179.

La Marca A, Grisendi V, Giulini S, Argento C, Tirelli A, Dondi G, Papaleo E, and Volpe A. Individualization of the FSH starting dose in IVF/ICSI cycles using the antral follicle count. *Journal of ovarian research* 2013: **6**; 11-11.

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

La Marca A, Grisendi V, Spada E, Argento C, Milani S, Plebani M, Seracchioli R and Volpe A. Reference values in ovarian response to controlled ovarian stimulation throughout the reproductive period, *Gynecological Endocrinology* 2014: 30; 66-69.

Law YJ, Zhang N, Venetis CA, Chambers GM, and Harris K. The number of oocytes associated with maximum cumulative live birth rates per aspiration depends on female age: a population study of 221 221 treatment cycles. *Hum Reprod* 2019: **34**; 1778-1787.

Leão Rde B and Esteves SC. Gonadotropin therapy in assisted reproduction: an evolutionary perspective from biologics to biotech. *Clinics (Sao Paulo)* 2014: **69**; 279-293.

Lee TH, Liu CH, Huang CC, Wu YL, Shih YT, Ho HN, Yang YS, and Lee MS. Serum anti-Müllerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted reproduction technology cycles. *Hum Reprod* 2008: **23**; 160-167.

Levinsohn-Tavor O, Friedler S, Schachter M, Raziel A, Strassburger D, and Ron-El R. Coasting-what is the best formula? *Hum Reprod* 2003: **18**; 937-940.

Lee T, Liu C and Huang C. Serum anti-Müllerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted reproduction technology cycles. *Human Reprod* 2008: **23**(1); 160–167.

Lispi M, Bassett R, Crisci C, Mancinelli M, Martelli F, Ceccarelli D, De Bellis C, and Mendola D. Comparative assessment of the consistency and quality of a highly purified FSH extracted from human urine (urofollitropin) and a recombinant human FSH (follitropin alpha). *Reprod Biomed Online* 2006: **13**; 179-193.

Lunenfeld B, Bilger W, Longobardi S, Kirsten J, D'Hooghe T, Sunkara SK. Decision points for individualized hormonal stimulation with recombinant gonadotropins for treatment of women with infertility. *Gynecol Endocrinol*. 2019: **35**(12); 1027-1036.

Magnusson A, Nilsson L, Olerod G, Thurin-Kjellberg A, and Bergh C. The addition of anti-Mullerian hormone in an algorithm for individualized hormone dosage did not improve the prediction of ovarian response-a randomized, controlled trial. *Hum Reprod* 2017: **32**; 811-819.

Munch EM, Sparks AE, Zimmerman MB, Van Voorhis BJ, and Duran EH. High FSH dosing is associated with reduced live birth rate in fresh but not subsequent frozen embryo transfers. *Hum Reprod* 2017: **32**; 1402-1409.

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 women starting IVF/ICSI: an RCT. Part 2: The predicted hyper responder. *Hum Reprod* 2017: **32**; 2506-2514.

Papaleo E, Zaffagnini S, Munaretto M, Vanni VS, Rebonato G, Grisendi V, Di Paola R, and La Marca A. Clinical application of a nomogram based on age, serum FSH and AMH to select the FSH starting dose in IVF/ICSI cycles: a retrospective two-centres study. *Eur J Obstet Gynecol Reprod Biol* 2016: **207**; 94-99.

Papanikolaou EG, Pozzobon C, Kolibianakis EM, Camus M, Tournaye H, Fatemi HM, Van Steirteghem A, and Devroey P. Incidence and prediction of ovarian hyperstimulation syndrome in women undergoing gonadotropin-releasing hormone antagonist in vitro fertilization cycles. *Fertil Steril* 2006: **85**; 112-120.

Popovic-Todorovic B, Loft A, Bredkjaeer HE, Bangsboll S, Nielsen IK, and Andersen AN. 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.

Practice Committee of the American Society for Reproductive Medicine. Prevention and treatment of moderate and severe ovarian hyperstimulation syndrome: a guideline. *Fertil Steril* 2016: **106**; 1634-1647.

Sahmay S, Atakul N, Aydogan B, Aydin Y, Imamoglu M, and Seyisoglu H. Elevated serum levels of anti-Müllerian hormone can be introduced as a new diagnostic marker for polycystic ovary syndrome. *Acta Obstet Gynecol Scand* 2013: **92**; 1369-1374.

Schats R, Sutter PD, Bassil S, Kremer JA, Tournaye H, and Donnez J. Ovarian stimulation during assisted reproduction treatment: a comparison of recombinant and highly purified urinary human FSH. On behalf of The Feronia and Apis study group. *Hum Reprod* 2000: **15**; 1691-1697.

Shavit T, Shalom-Paz E, Samara N, Aslih N, Michaeli M, and Ellenbogen A. Comparison between stimulation with highly purified hMG or recombinant FSH in patients undergoing IVF with GnRH antagonist protocol. *Gynecol Endocrinol* 2016: **32**; 629-633.

Simon AL, Kiehl M, Fischer E, Proctor JG, Bush MR, Givens C, Rabinowitz M, and Demko ZP. Pregnancy outcomes from more than 1,800 *in vitro* fertilization cycles with the use of 24-chromosome single-nucleotide polymorphism-based preimplantation genetic testing for aneuploidy. *Fertil Steril* 2018: **110**; 113-121.

Smith V, Osianlis T, and Vollenhoven B. Prevention of ovarian hyperstimulation syndrome: a review. *Obstet Gynecol Int* 2015: **2015**; 514159.

Steward R, Lan L, Shah A, Yeh J, Price T, Goldfarb J and Muasher S. Oocyte number as a predictor for ovarian hyperstimulation syndrome and live birth: an analysis of 256,381 in vitro fertilization cycles. *Fertil Steril 2014*; 101(4): 967-73.Sunkara SK, Rittenberg V, Raine-Fenning N, Bhattacharya S, Zamora J, and 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.

Trew GH, Brown AP, Gillard S, Blackmore S, Clewlow C, O'Donohoe P, and Wasiak R. *In vitro* fertilisation with recombinant follicle stimulating hormone requires less IU usage compared with highly purified human menopausal gonadotrophin: results from a European retrospective observational chart review. *Reprod Biol Endocrinol* 2010: **8**; 137.

Velthuis E, Hubbard J, Longobardi S, and D'Hooghe T. The frequency of ovarian hyperstimulation syndrome and thromboembolism with originator recombinant human follitropin alfa (GONAL-f) for medically assisted reproduction: a systematic review. *Adv Ther* 2020: **37**; 4831-4847.

Witz CA, Daftary GS, Doody KJ, Park JK, Seifu Y, Yankov VI, Heiser PW, and Menopur in GnRH Antagonist Cycles with Single Embryo Transfer High Responder Trial Group. Randomized, assessor-blinded trial comparing highly purified human menotropin and recombinant follicle-stimulating hormone in high responders undergoing intracytoplasmic sperm injection. *Fertil Steril* 2020: **114**; 321-330.

Yovich JL, Alsbjerg B, Conceicao JL, Hinchliffe PM, and Keane KN. PIVET rFSH dosing algorithms for individualized controlled ovarian stimulation enables optimized pregnancy productivity rates and avoidance of ovarian hyperstimulation syndrome. *Drug Des Devel Ther* 2016: **10**; 2561-2573.

Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, Rienzi L, Sunde A, Schmidt L, Cooke ID*, et al.* The International Glossary on Infertility and Fertility Care, 2017. *Fertil Steril* 2017: **108**; 393-406.