Algorithm-based individualization methodology of the starting gonadotropin dose in IVF/ICSI and the freeze-all strategy prevent OHSS equally in normal responders: a systematic review and network meta-analysis of the evidence

Angelo Marino[1] · Salvatore Gullo[2] · Francesca Sammartano[1] · Aldo Volpes[1] · Adolfo Allegra[1]

Abstract

Purpose Ovarian hyperstimulation syndrome (OHSS) represents a rare but dangerous condition associated with controlled ovarian stimulation (COS) in IVF/ICSI. Over the last decades, many strategies have been introduced into clinical practice with the objective of preventing this potentially life-threatening condition. Among these, the freeze-all policy has gained great popularity, thanks to improvements in vitrification. Nevertheless, not all clinics have adequate skills in vitrification procedures and patients may be dissatisfied with a longer time to pregnancy.

Methods This study is a systematic review and network meta-analysis of randomized controlled trials comparing different strategies of ovarian stimulation in IVF/ICSI cycles (freeze-all policy, algorithm-based individualization of the starting dose, experience-based individualization of the starting dose, standard dose) in terms of reduction of OHSS, in normal responders.

Results The results indicate that only the algorithm-based individualization of the starting gonadotropin dose reduces OHSS similarly to the freeze-all strategy.

Conclusion Albeit in the era of the freeze-all policy, the personalization of the starting gonadotropin dose obtained by the use of algorithms should be pursued as a valid and safe option for IVF.

Keywords Freeze-all policy · Algorithm-based individualization of starting gonadotropin dose · Normal responders

Introduction

Ovarian hyperstimulation syndrome (OHSS) represents a rare but dangerous condition associated with controlled ovarian stimulation (COS) during IVF/ICSI cycles, and its development is mainly mediated by hCG.

The incidence of this condition is variable between 0.2 and 3%, depending on the cases series, but its real frequency is probably underestimated. The findings of the syndrome include ovarian enlargement, ascites, pleural effusion, hemoconcentration, hypercoagulability and electrolytic imbalance. OHSS is classified on the basis of the signs and symptoms as mild, moderate and severe. Once again, there is a great variability in the classification among different physicians, and this could explain the great inconsistency in the reported incidence rates. The last European Society of Human Reproduction and Embryology (ESHRE) report on ART procedures in Europe, referring to cycles performed by 39 participating countries in 2017, indicates an incidence rate of OHSS of 0.2% of all reported cycles [1].

Over the last decades, many strategies have been introduced into clinical practice with the objective of preventing this potentially life-threatening condition: these include metformin pretreatment in women with polycystic ovary syndrome (PCOS), use of a gonadotropin-releasing hormone (GnRH) antagonist protocol for pituitary suppression, use of clomiphene citrate for controlled ovarian stimulation, use of dopamine agonists and use of progesterone for luteal phase support. These strategies are effective in reducing the occurrence of OHSS, while not influencing or even improving

✉ Angelo Marino
   angelo.marino@clinicaandros.it

1   Reproductive Medicine Unit, ANDROS Day Surgery Clinic, Via Ausonia 43/45, 90144 Palermo, Italy

2   Department of Psychology, Educational Science and Human Movement-Statistics Unit, University of Palermo, Palermo, Italy

pregnancy outcomes, as described by a recent overview of Cochrane reviews [2].

In the last years, the freeze-all policy, which is the freezing of all good-quality embryos and their progressive transfer in subsequent cycles, has gained great popularity. Data from the US Centers for Disease Control and Prevention show that the proportion of frozen embryo transfers is increasing—from around 20% in 2005 to almost 50% in 2014—while the proportion of fresh embryo transfer procedures following IVF and ICSI decreased correspondingly [3].

This increase was possible thanks to improvement in vitrification procedures as an alternative to the more traditional method of slow freezing. Vitrification uses higher concentrations of cryoprotectants and ultra-rapid cooling, lowering the risk of ice nucleation and crystallization. It is now established that vitrification is much more efficient than slow freezing, regardless of the stage of embryo development [4].

Three recent meta-analyses indicated a significant OHSS reduction with elective frozen embryo transfer (eFET) when compared to fresh ET [5–7]. The reduction of OHSS and the demonstration that, in frozen cycles, the cumulative rate of live births (CLBR) constantly increases with the number of oocytes retrieved [8] resulted in a significant increase in freeze-all cycles. Consequently, a standardized ovarian stimulation, in scheduled freeze-all cycles, seems to be replacing the starting dose personalization, one of the historical cornerstones of the IVF cycles.

Nevertheless, not all IVF clinics are well-skilled in managing vitrification, and the patients may be dissatisfied with a longer time to pregnancy correlated with the freeze-all policy. Additionally, some other negative aspects correlated with this strategy should be considered: a recent systematic review and meta-analysis, in fact, showed that the risk of pre-eclampsia increased with eFET in comparison with fresh embryo transfer (RR = 1.79; 95% CI: 1.03–3.09) [5].

Furthermore, the freeze-all policy does not completely eliminate OHSS, even if the oocyte maturation trigger is obtained with GnRH agonists [9].

Another practical approach for minimizing OHSS is the tailoring of the starting gonadotropin dose by using specific mathematical tools, such as algorithms and/or nomograms.

As already shown, when a fresh embryo transfer is scheduled, an optimal—rather than a maximal—oocyte yield is the preferred accomplishment after controlled ovarian stimulation. In fact, live birth rates steadily increase when an optimal number of oocytes is collected, whereas low response and hyper response are associated with lower implantation rates, increased obstetrical risks and, at least when considering hyper response, increased risk of OHSS in the fresh cycle [8, 10–12].

Indeed, choosing different doses of gonadotropins for different patients has represented for decades one of the most important clinical decisions in the planning of IVF cycles for infertile couples.

Commonly, the choice of the FSH starting dose is made in accordance with clinical history and the ovarian response to stimulation in previous IVF cycles. If no previous cycles have been performed, the choice will be based on such criteria as women's age and markers of ovarian reserve. Currently used markers of ovarian reserve include FSH, anti-Müllerian hormone (AMH) and antral follicle count (AFC), with the last two biomarkers having the best performance in predicting ovarian response to exogenous FSH. The use of biomarkers of the functional ovarian reserve should define more patient-tailored dosing regimens that fulfil both the clinical efficacy and safety objectives for COS. Often, the evaluation of these functional markers is carried out by the physicians on the basis of their clinical experience. This way of personalizing the treatment is important, but we believe less precise and standardizable than a COS tailoring obtained with specific mathematical algorithms. A recent Cochrane review concluded that algorithms for dose individualization reduce the incidence of OHSS compared to standard dosing of 150 IU, although their use does not improve live birth rate [13].

In the freeze-all era, is the scenario presented above "a mere realm of the past", as stated by Broekmans in 2019 [14], or is there still room for a therapy based more on the specific characteristics of the patients?

The aim of this study is to conduct, for the first time, as far as we know, a meta-analysis of all available randomized controlled trials in order to make a comparison between four different strategies of COS (the freeze-all policy, the algorithm-based individualization of the starting gonadotropin dose, the clinical experience-based individualization of the starting gonadotropin dose and the standard gonadotropin dose of 150 IU) in terms of reduction of OHSS, in a normal responder population.

Since there are no comparative studies between freeze-all and algorithm-based strategy to date, we have made an indirect comparison thanks to the use of the statistical approach of the network meta-analysis.

Materials and methods

Search strategy and selection criteria

A systematic search, based on PubMed, Cochrane CENTRAL and EMBASE databases, was conducted to identify potential study areas. We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [15] and PRISMA extension for Network Meta-analysis [16].

The criteria for including the studies in the meta-analysis [16, 17] were the following: normal responder population; randomized design of the trials (RCTs); and combination of two or more of the subsequent interventions, which were freeze-all strategy, individualization of the starting dose based on algorithms, individualization of the starting dose based on clinical experience, or administration of a standard dose.

OHSS was considered as the primary outcome, while clinical pregnancy and live birth rates as secondary ones.

The individualization of the starting gonadotropin dose based on clinical experience is referred to a choice based on the experience of the physician generally depending on the woman's age, AMH levels, AFC and response to gonadotropins in any prior IVF cycle. The standard dose regimen is referred, instead, to a starting dose of gonadotropins of 150 IU, independently of the basal characteristics of the patients.

For a complete analysis, the primary outcome included not only the actual cases of OHSS but also those in which a suspension of the treatment had been determined due to a serious risk of the syndrome itself.

Network meta-analysis relies on the transitivity assumption that requires that all interventions compared are jointly randomizable, i.e. all interventions compared in a network meta-analysis should be clinically reasonable in a theoretical multi-arm RCT. Considering the types of treatment reported in the selected studies, we considered the transitivity assumption valid.

The present study was exempted from the Institutional Review Board approval, being a meta-analysis.

Data extraction and assessment of risk of bias

Electronic and manual search, from 1990 to August 2021, was first conducted to identify potential areas of study. After duplicates were removed, two researches (A. M. and S. G.) reviewed the studies independently, assessing the eligibility of all remaining citations. In the first screening, title and abstract were reviewed; in the second screening, the full texts of the potential studies were examined.

Any disagreement was resolved by discussion with a third reviewer (A.A.). Study quality was assessed by the two independent reviewers (A.M. and S.G.) using the Cochrane risk of bias assessment tool [18]. Again, in cases of disagreement, a third reviewer (A.A.) was asked in order to reach consensus. Assessment of quality included the following domains: randomization and sequence generation, allocation concealment, blinding of participants and outcome assessment, completeness of outcome data and selective outcome reporting. The Cochrane risk of bias assessment was performed in duplicate by two independent reviewers (Supplementary data, S1).

Statistical analysis

Network meta-analysis was conducted to simultaneously compare four different strategies of gonadotropin dosage. The network meta-analysis is a method that combines evidence from RCTs comparing different treatments for a given clinical population through direct and indirect estimates of the relative effect of each treatment [19]. The indirect comparison between the relative effects of two interventions (e.g. A vs B) is performed by comparing them with a third common intervention (C) through a series of direct trials (e.g. trial of A vs C and trial of B vs C).

All network meta-analyses were conducted within a random-effects multiple regression model and implemented in R (version 3.6.1). Where direct data were available, pairwise meta-analyses were performed trough the statistical package "metafor" (version 2.4-0, http://www.metafor-project.org https://github.com/wviechtb/me).

The network meta-analysis was carried out by using the frequentist model with the package "netmeta" (version 0.9-0, https://cran.rproject.org/web/packages/netmeta/index.html).

The results were expressed in terms of prevalence, OR and 95% confidence intervals (CI).

Through the same package, we produced comparison adjusted funnel plots to explore publication bias or other small study effects for all available comparisons. Symmetry around the effect estimate line indicates the absence of publication bias or small study effects. In pairwise meta-analyses heterogeneity was measured by the I^2 statistic (range: 0–100%); I^2 with p value < 0.05 indicated the presence of significant variability of results across studies [20].

Descriptive analyses involving four covariates (total dose of gonadotropins, levels of oestradiol peak, number of retrieved oocytes, BMI) were carried out and shown in Supplementary data, S2.

We also assessed global statistical heterogeneity across all comparisons using the τ^2 measure from the "netmeta" statistical package. The estimate of τ^2 is based on the generalized DerSimonian-Laird method. Estimates of τ^2 of approximately 0.04, 0.16 and 0.36 are considered to represent a low, moderate and high degree of heterogeneity, respectively. Inconsistency in the network analysis was firstly measured by generalized Cochran Q statistics as described by Krahn and colleagues [21]. Under the assumption that the indirect evidence must be consistent with the direct evidence, we also assessed inconsistency between direct and indirect estimates by using the "netsplit" function in the R package netmeta. Finally, the network heat plots have been calculated. These plots have grey squares, which represent the size of the contribution of the direct estimate in columns, compared with the network estimate in rows. The coloured squares around these represent the degree of inconsistency, with red

squares indicating "hotspots" of inconsistency that should be removed from the review (Supplementary data, S3).

We estimated the probability that each treatment was the most effective by calculating *P* score (possible rage: 0–1), which measures the mean extent of certainty that one management strategy is better than another, averaged over all competing strategies. Higher scores indicate a greater probability of the strategy being ranked as the best.

When we failed to find clear and conclusive evidence of low inconsistency, network meta-regression (NMR) analysis was used to investigate potential sources of inconsistency (Supplementary data, S4). NMR is an extension of network meta-analysis, which examines whether several treatment effects differ according to a covariate. NMR results consist of, for each comparison, a treatment effect estimated at the covariate value zero (unadjusted model) and a regression coefficient for the treatment by covariate interaction (adjusted model). The models are commonly compared

trough the deviance information criterion for which lower value indicates better model [22].

Results

Electronic and manual search yielded 655 citations. After the 62 duplicates were removed, two researches (A. M. and S. G.) reviewed the studies independently, excluding 544 citations after the first screening. Thirty-six out of the remaining 49 potential studies were excluded after the second screening (Fig. 1). This process left 13 randomized controlled trials that comprised 10,818 participants, 2,516 of which for freeze-all strategy, 2,152 for algorithm-based treatment, 3,726 for experience-based treatment and 2,424 treated with standard dose. All of these reported the incidence of OHSS. The search strategy details are shown in Supplementary data, S5.

Fig. 1 Flowchart for the studies identification and selection process according to preferred reporting items for systematic reviews and meta-analyses guidelines
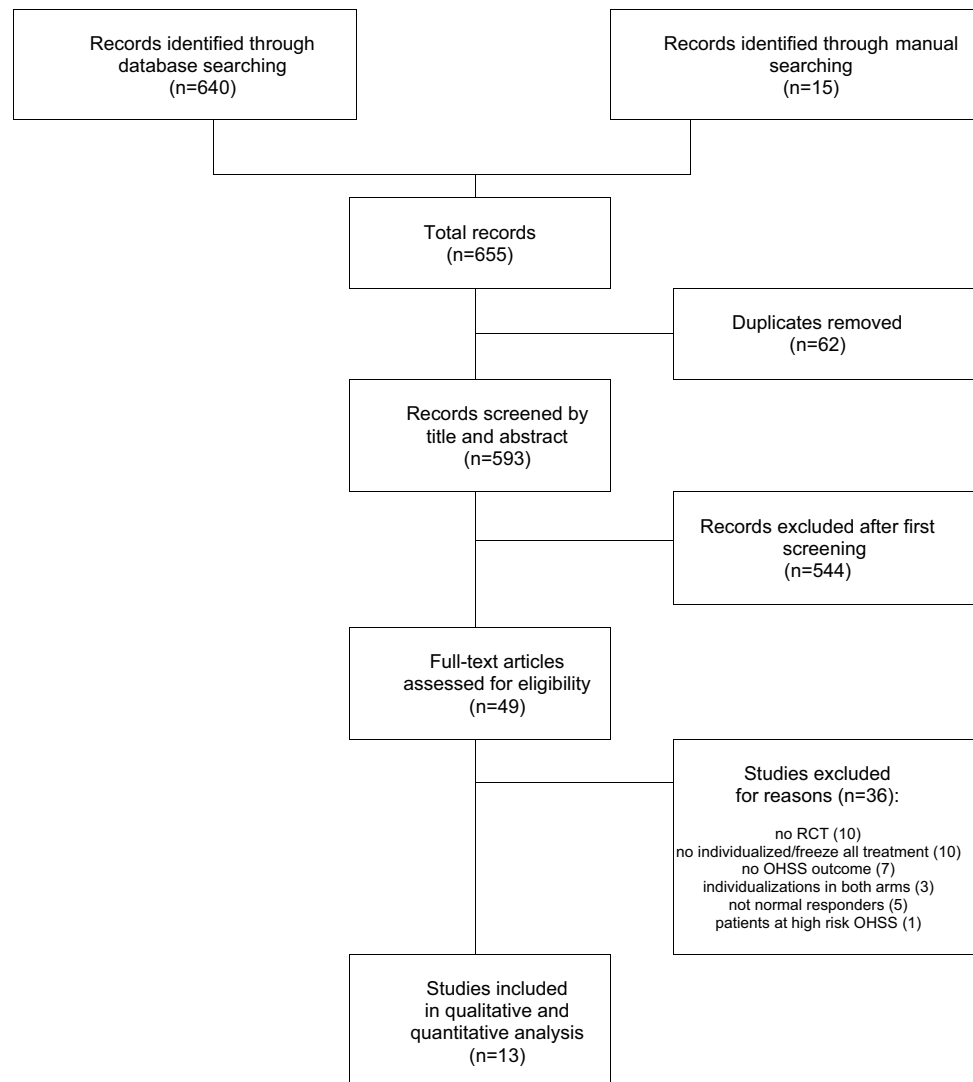
The study reported four different treatment strategies: freeze-all, individualization of the starting dose based on algorithms, individualization of the starting dose based on clinical experience and use of standard or non-individualized dose.

The details of the intervention arms for each study are described below.

Four out of the thirteen studies were referred to freeze-all strategy [23–26], eight were referred to algorithm-based strategy [27–34], and one was referred to individualization of the starting dose based on AFC [35]. This last strategy was classified as an individualization policy based on clinical experience.

With regard to the comparison groups, in the four studies referred to freeze-all strategy, the comparator was the individualization strategy based on clinical experience. Among the eight studies referred to algorithm-based strategy, in six [27–29, 32–34], the comparator was the standard gonadotrophin dose of 150 IU, whereas in two [30, 31], the comparator was the individualization strategy based on clinical experience.

In the last study [35], the comparison group was the standard gonadotrophin dose of 150 IU.

No studies directly compared freeze-all with algorithm-based individualization strategies and freeze-all with standard dose strategies, while direct comparisons were found for all the remaining treatment pairs. According to the PRISMA extension for network meta-analyses [16], the geometry of the treatment network was illustrated by producing a network plot (Fig. 2) with node and connection size corresponding to the number of study participants and number of studies, respectively.

In the four RCTs considering the freeze-all strategy as intervention group [23–26], a protocol with GnRH antagonists was used. The induction of oocyte maturation was obtained with recombinant chorionic gonadotropin [24], with urinary chorionic gonadotropin [23, 25], or with a gonadotropin-releasing hormone agonist [26].

Considering the eight trials focusing on the individualization algorithm-based of the starting dose, a long protocol with GnRH agonists was used in three [27, 28, 30], while in the remaining ones, a protocol with GnRH antagonists was used [29, 31–34]. In all these papers, a specific algorithm based on the ovarian reserve test was used for deciding the starting dose.

In more detail, the algorithms were based on:

- AFC, ovarian volume, ovarian stromal blood flow detected by Power Doppler (total Doppler score), age and smoking habits [27]
- Body weight and AMH [29, 31, 33, 34]
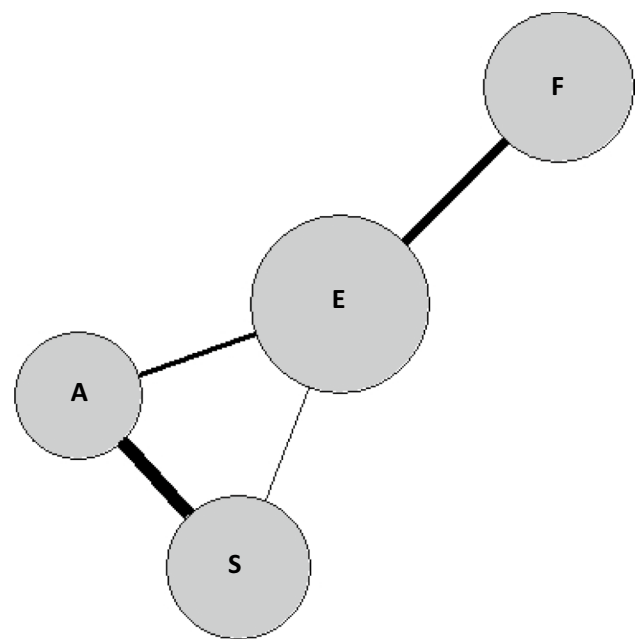- Age, BMI, basal FSH level and AFC [28]

Fig. 2 Network plot of treatment strategies. Options: F = Freeze-all strategy; E = individualized treatment based on clinical Experience; A = individualized treatment based on Algorithms; S = treatment with Standard dose. The size of the nodes represents the number of women randomized to each treatment option, and the thickness of the lines represents the number of randomized trials with head to head comparison between each treatment option

- Age, basal FSH level and AMH [30]
- AMH [32]

In the multicentre trial of van Tilborg et al. (2017) [35], protocols with GnRH agonists or antagonists were used both depending on the experience of each centre involved.

In six studies [29, 31–35], the individualized starting dose was maintained for the entire period of the stimulation. The induction of the oocyte maturation was obtained with rhCG or uhCG. In some studies, in selected cases, the GnRH agonist was used for triggering the final oocyte maturation [29, 31, 33, 34]. Table 1 reports the main characteristics for each study.

Descriptive analysis, obtained by pooling data across studies of the distinct categories, showed that the total dose of gonadotropins was on average 1,759.31 (± 559.76) IU for participants in the freeze-all group; 1,111.89 (± 377.52) IU for participants who received algorithm-based individualized strategy; 1,697.68 (± 459.33) IU for participants who received individualized treatment based on clinical experience; and 1,948.84 (± 354.04) IU for those treated with a standard dose. Levels of the oestradiol peak were not reported in five out of the thirteen RCTs [26, 29, 31, 32, 35]; for the remaining studies, the peak of oestradiol, on average, showed a value of 2,064.82 (± 1297.53) pg/ml in freeze-all;

Table 1 Study characteristics of the selected RCTs

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Shi et al., 2018 [23]	Multicentre, randomized, controlled trial	Women with regular menses who were undergoing their first cycle of IVF or ICSI because of tubal factors, male factors, or both. Eligible women were 20 to 35 years of age, had a normal menstrual cycle (defined as a spontaneous cycle length of ≥ 21 days and ≤ 35 days), and a duration of infertility of more than 1 year	Patients with normal ovaries at ultrasound	Freeze-all ($n = 1077$)	Individualization strategy based on clinical experience with fresh ET ($n = 1080$)
Vuong et al., 2018 [24]	Single-centre, randomized, controlled	Women without the polycystic ovary syndrome (on the basis of the Rotterdam criteria)	Patients with normal ovaries at ultrasound	Freeze-all ($n = 391$)	Individualization strategy based on clinical experience with fresh ET ($n = 391$)
Wei et al., 2019 [25]	Non-blinded, multicentre, randomized controlled trial	Women with regular menses who were undergoing the first cycle of IVF with or without ICSI with an indication of tubal, male, or unexplained infertility. Eligible women were aged 20–35 years, and had a menstrual cycle length of 21–35 days indicative of regular ovulation	Patients with normal ovaries at ultrasound	Freeze-all ($n = 825$)	Individualization strategy based on clinical experience with fresh ET ($n = 825$)
Stormlund et al., 2020 [26]	Multicentre randomized controlled trial	Women aged 18-39 years with a regular menstrual cycle (≥ 24 and ≤ 35 days) who were starting their first, second, or third treatment cycle of in vitro fertilization or intracytoplasmic sperm injection because of male, tubal, uterine, or unexplained infertility	Serum anti-müllerian hormone levels greater than 6.28 pmol/L	Freeze-all ($n = 223$)	Individualization strategy based on clinical experience with fresh ET ($n = 230$)

Table 1 (continued)

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Popovic-Todorovic et al., 2003 [27]	Prospective randomized dual-centre study	Women at the first IVF/ICSI treatment cycle; normal basal serum FSH level (with current assays, up to 12.5 IU/l); presence of both ovaries; regular spontaneous menstrual cycle (21 ± 35 days); maximum age 39 years at the onset of treatment and no evidence of endocrine disorders	Expected normal responders	Individualized dose based on algorithm (n=131)	Standard dose (150 IU) (n=131)
Olivennes et al., 2015 [28]	Prospective, randomized, controlled, multiregional, open-label, phase IV study	Women with spontaneous menstrual cycle of 21–35 days; BMI < 30 kg/m²; basal serum FSH level 12 IU/L or less; no previous poor response to ovarian stimulation in two or more assisted reproductive technique cycles; no previous hyper-response to ovarian stimulation (defined as 25 or more oocytes retrieved) in two or more assisted reproduction technique cycles; no previous severe OHSS; no PCOS or three or more spontaneous abortions	Patients with no poor or hyper response	Individualized dose based on algorithm (n = 96)	Standard dose (150 IU) (n = 104)
Nyboe-Andersen et al., 2017 [29]	Randomized, multicentre, assessor-blinded, non-inferiority trial (ESTHER-1).	Women undergoing their first IVF/ ICSI cycle; BMI 17.5–32.0 kg/m², regular menstrual cycles of 24–35 days, presence of both ovaries, and basal FSH serum concentration 1–15 IU/L, no endometriosis stage III–IV, no history of recurrent miscarriage and no use of hormonal preparations	Expected normal responders	Individualized dose based on algorithm (n = 665)	Standard dose (150 IU) (n = 661)

Table 1 (continued)

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Allegra et al., 2017 [30]	Two-arm, single-centre, prospective, randomized, interventional trial	Women at the first IVF/ICSI cycle, female age between 18 and 40 years, BMI between 18 and 25 kg/m², serum AMH concentrations between 1.0 and 4.0 ng/ml, basal serum day 3 FSH ≤ 15 IU/l, normal regular menstrual cycles, ranging from 25 to 33 days in length, normal uterine cavity, presence of both ovaries, no PCOS, no severe endometriosis, no previous ovarian surgery, absence of ovarian cysts, no use of hormonal contraception in the previous three months, absence of metabolic or endocrinological disease	Patients with normal values of AMH and basal FSH	Individualized dose based on algorithm ($n = 95$)	Individualization strategy based on clinical experience (150 IU ≤35 ys; 225 IU > 35 ys) ($n = 99$)

Table 1 (continued)

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Bosch et al., 2019 [31]	Randomized, controlled, assessor-blind trial	Women who had participated in cycle 1 (women undergoing their first IVF/ICSI cycle, 18–40 years of age, with regular menstrual cycles, diagnosed with tubal infertility, unexplained infertility, endometriosis stage I/II or with partners diagnosed with male factor infertility) and failed to achieve an ongoing pregnancy were eligible for cycle 2 and women who failed to achieve an ongoing pregnancy in cycle 2 were eligible for cycle 3. Patients with severe OHSS in a previous cycle, or patients with any clinically relevant change to any of the eligibility criteria or any clinically relevant medical history since the previous cycle were not eligible for enrolment	Normal responders	Individualized dose based on algorithm ($n = 347$)	Individualization strategy based on clinical experience ($n = 354$)
Friis Petersen et al., 2019 [32]	Dual-centre, randomized controlled trial	Women with normal menstrual cycles of 24-35 days interval presumed to be ovulatory; bodyweight < 75 kg; FSH ≤ 12IU/L; two ovaries accessible to oocyte retrieval; uterus with presumed normal function based on sonography; no PCOS; no endometriosis grade III/IV; no hydrosalpinx on ultrasound, no recurrent miscarriages (≥3), absence of major medical disorders	Expected normal responders	Individualized dose based on algorithm ($n = 149$)	Standard dose (150 IU) ($n = 72$)

Table 1 (continued)

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Ishihara et al., 2021 [33]	Randomized, controlled, assessor-blind, multicentre, non-inferiority trial	Japanese women between 20 and 40 years, BMI between 17.5 and 32 kg/m², first IVF/ICSI cycle. Eligibility criteria included women diagnosed with tubal infertility, unexplained infertility, endometriosis stage I/II or with partners diagnosed with male factor infertility. Additionally, women had to have regular menstrual cycles of 24–35 days, presence of both ovaries and early follicular phase FSH serum levels of 1–15 IU/l.	Expected normal responders	Individualized dose based on algorithm ($n = 170$)	Standard dose (150 IU) ($n = 177$)
Qiao et al., 2021 [34]	Randomized, controlled, multi-centre, assessor-blind, non-inferiority trial	Asian reproductive-aged women between 20 and 40 years, BMI between 17.5 and 32 kg/m², first IVF/ICSI cycle. Eligibility criteria included women diagnosed with tubal infertility, unexplained infertility, endometriosis stage I/II or with partners diagnosed with male factor infertility. Additionally, women had to have regular menstrual cycles of 24–35 days, presence of both ovaries and early follicular phase FSH serum levels of 1–15 IU/l. Women with endometriosis stage III/IV, history of recurrent miscarriage and with one or more follicles 10 mm observed prior to randomization were excluded	Expected normal responders	Individualized dose based on algorithm ($n = 499$)	Standard dose (150 IU) ($n = 510$)

Table 1 (continued)

Study reference	Type of study	Definition of population	Normal responders definition	Intervention group (sample size)	Comparison group (sample size)
Van Tilborg et al., 2017 [35]	Multicentre prospective cohort study with two embedded RCT	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), endocrine or metabolic abnormalities or women who opted for oocyte donation were not included in the trial	Expected normal responders	Individualization of the starting dose based on AFC (individualization policy based on clinical experience). (n = 747)	Standard dose (150 IU) (n = 769)

a value of 1,845.15 (± 867.23) pg/ml in algorithm-based individualized strategy; a value of 2,046.54 (± 1398.44) pg/ml in experience-based strategy; and a value of 2,217 (± 1331.32) pg/ml in standard dose. The number of oocytes retrieved was 13.20 (± 5.32) for freeze-all, 9.55 (± 5.27) for algorithm-based individualization, 12.70 (± 4.38) for experience-based strategy and 11.33 (± 6.11) for standard dose. The BMI (Kg/m^2) was 21.94 (± 2.61), 22.00 (± 2.93), 21.09 (± 2.64) and 21.97 (± 2.85) in average, respectively, for freeze-all, algorithm-based, experience-based and standard dose treatment (see Supplementary data, S2).

Lastly, the duration of stimulation was 9.42 days in freeze-all, 9.7 days in the algorithm-based strategy, 9.43 days in the experience-based strategy and 9.26 days in standard dose.

Risk of bias

The quality of included studies was overall moderate to good. Due to the nature of the treatments, which makes blinding of participants non-feasible, risk on this domain was assessed as unclear (9 studies; 70%) when not otherwise specified in 4 studies [27, 28, 30, 35]. In the other domains, the percentage of studies with unclear risk of bias was less than 10% for incomplete outcome and 15% for blinding outcome; high risk was found in 15% of studies regarding selective reporting and in 30% of cases regarding blinding outcome. Specifically, all the 13 studies included had a low risk of bias for randomization and for allocation concealment. Detection assessment (i.e. selective reporting) was judged to have a low risk of bias in 11 studies and was considered high in two studies [27, 29]; all the studies had a low risk of bias in outcome reporting (i.e. incomplete outcome) with the exception of one [31], where the risk of bias was unclear. The risk of blinding outcome was considered high in four RCTs [27, 28, 30, 35] and unclear in four [23, 25, 26, 32]; in the remaining studies, the risk was low. No potential conflict of interest was explicitly declared in two studies [30, 32]. A summary of risk of bias assessment in included trials is provided in Supplementary data, S1.

Primary outcome

OHSS rate

Regarding the primary outcome (OHSS), the thirteen studies showed little and non-significant level of heterogeneity ($\tau^2 = 0.014$; $Q = 9.36$, $df = 9$, $P = .405$) and the level of inconsistency among designs ($I^2 = 0.1\%$ [0.0–60.2%]) was found to be not significant after applying the Q statistics ($Q = 0.02$, $df = 1$, $P = .898$). Assessment of inconsistency, carried out also through comparison of direct and indirect evidence, revealed no significant difference between the two estimates

(z = .059, P = 0.953), and the network heat plot had no red "hotspots" of inconsistency (Supplementary data, S3).

The ranking P score for OHSS risk was: freeze-all strategy (0.950), individualization algorithm-based (0.708), individualization experience-based (0.183) and standard dose (0.159).

Direct comparisons

Results showed that individualized algorithm-based strategy showed significantly lower risk of hyperstimulation syndrome than standard dose treatment (6 studies [27–29, 32–34]; OR = 0.54, 95% CI 0.42–0.69) and non-significant lower risk than women in the experience-based group (2 studies, [30, 31]; OR = 0.57, 95% CI 0.29–1.10).

In the other direct comparisons, women who received treatment based on clinical experience had significantly higher risk of OHSS in comparison with the freeze-all group (4 studies, [23–26]; OR = 2.83, 95% CI 1.54–5.20), while the risk in individualized experience-based treatment was found to be not significantly different from standard dose treatment (1 study [35]; OR = 1.03, 95% CI 0.39–2.73) (Fig. 3).
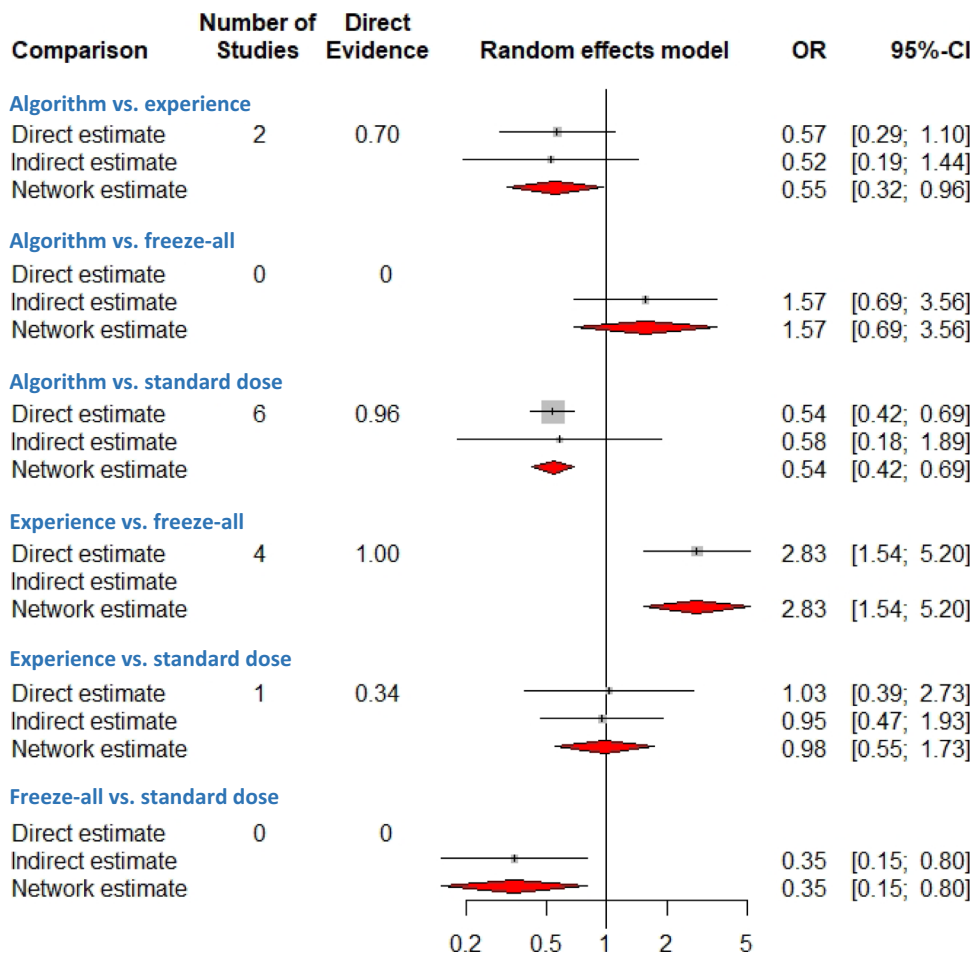
Indirect comparisons

Estimates from indirect comparison between freeze-all and individualized algorithm-based strategy showed there were no significant differences between these two strategies (OR = 1.57, 95% CI 0.69–3.56). Also, the indirect comparisons between algorithm-based and clinical experience-based strategies (OR = 0.52, 95% CI 0.19–1.44) and between algorithm-based strategy and standard dose (OR = 0.58, 95% CI 0.18–1.89) were found not significant.

Indirect comparison between clinical experience-based strategy and standard dose (OR = 0.95, 95% CI 0.47–1.93) resulted not significant, while the comparison between freeze-all and standard dose showed a lower significant risk of OHSS in favour of "freeze-all" (OR = 0.35, 95% CI 0.15–0.80) (Fig. 3).

Moreover, network estimate combines the contribution of direct and indirect evidence and allowed us to control for inconsistency in the estimates of individual comparisons. It is noteworthy to point out that network estimate showed a significant lower risk of OHSS for algorithm-based strategy in comparison with experience-based treatment (OR = 0.55, 95% CI 0.32–0.96) and with standard

Fig. 3 Direct and indirect comparison of OHSS between the different strategies

Comparison	Number of Studies	Direct Evidence	Random effects model	OR	95%-CI
Algorithm vs. experience					
Direct estimate	2	0.70		0.57	[0.29; 1.10]
Indirect estimate				0.52	[0.19; 1.44]
Network estimate				0.55	[0.32; 0.96]
Algorithm vs. freeze-all					
Direct estimate	0	0			
Indirect estimate				1.57	[0.69; 3.56]
Network estimate				1.57	[0.69; 3.56]
Algorithm vs. standard dose					
Direct estimate	6	0.96		0.54	[0.42; 0.69]
Indirect estimate				0.58	[0.18; 1.89]
Network estimate				0.54	[0.42; 0.69]
Experience vs. freeze-all					
Direct estimate	4	1.00		2.83	[1.54; 5.20]
Indirect estimate					
Network estimate				2.83	[1.54; 5.20]
Experience vs. standard dose					
Direct estimate	1	0.34		1.03	[0.39; 2.73]
Indirect estimate				0.95	[0.47; 1.93]
Network estimate				0.98	[0.55; 1.73]
Freeze-all vs. standard dose					
Direct estimate	0	0			
Indirect estimate				0.35	[0.15; 0.80]
Network estimate				0.35	[0.15; 0.80]

0.2   0.5   1   2   5

dose (OR = 0.54, 95% CI 0.42–0.69). For the other comparisons, network estimates confirmed the results of direct and indirect estimates.

Overall, the results suggest that freeze-all and algorithm-based treatments are more effective in comparison with the other two approaches in reducing risk of OHSS and not significantly different from each other.

Secondary outcomes

Live birth rate (LBR)

Eleven out of the thirteen studies reported rates of live birth [23–26, 28, 29, 31–35]. Overall, these studies comprised 2,516 participants for the freeze-all policy, 2,152 for the individualized algorithm-based strategy, 3,726 for the individualized clinical experience-based strategy and 2,424 for the standard dose. When data were pooled, there was a low, but statistically significant, level of statistical heterogeneity within studies ($\tau^2 = 0.046$; $Q = 25.36$; $P = 0.001$) and a level of inconsistency between studies of $I^2 = 66.10\%$ (95% CI 31.2–83.3%) was found to be not significant after applying the Q statistics ($Q = 0.30$, $df = 1$, $P = .583$). Assessment of inconsistency carried out also through comparison of direct and indirect evidence revealed no significant difference between the two estimates ($z = -0.299$, $P = 0.765$), and the network heat plot showed no red "hotspots" of inconsistency (Supplementary data, S3).

Individualized strategy based on algorithm was ranked first (P score = 0.693), followed by freeze-all strategy (P score = 0.664), standard dose strategy (P score = 0.366) and clinical experience ($P = 0.278$). Direct comparisons showed that LBR did not reveal significant differences in any of the four comparison pairs: algorithm-based vs clinical experience-based individualization (1 study, OR = 1.07, 95% CI 0.63–1.83), clinical experience-based individualization vs freeze-all (4 studies, OR = 0.88, 95% CI 0.69–1.13), algorithm-based individualization vs standard dose (5 studies, OR = 1.13, 95% CI 0.87–1.47) and clinical experience-based vs standard dose (1 study, OR = 0.93, 95% CI 0.58–1.46). Similarly, indirect estimates showed no significant difference in all the five comparison pairs: algorithm-based vs clinical experience-based individualization (OR = 1.22, 95% CI 0.72–2.07), algorithm-based individualization vs freeze-all (OR = 1.01, 95% CI 0.65–1.58), algorithm-based individualization vs standard dose (OR = 0.99, 95% CI 0.49–2.01), experience-based individualization vs standard dose (OR = 1.05, 95% CI 0.58–1.90) and freeze-all vs standard dose (OR = 1.10, 95% CI 0.71–1.71) (Fig. 4). Network estimates confirmed the results of direct and indirect estimates indicating that there were no significant differences among all the four strategies.
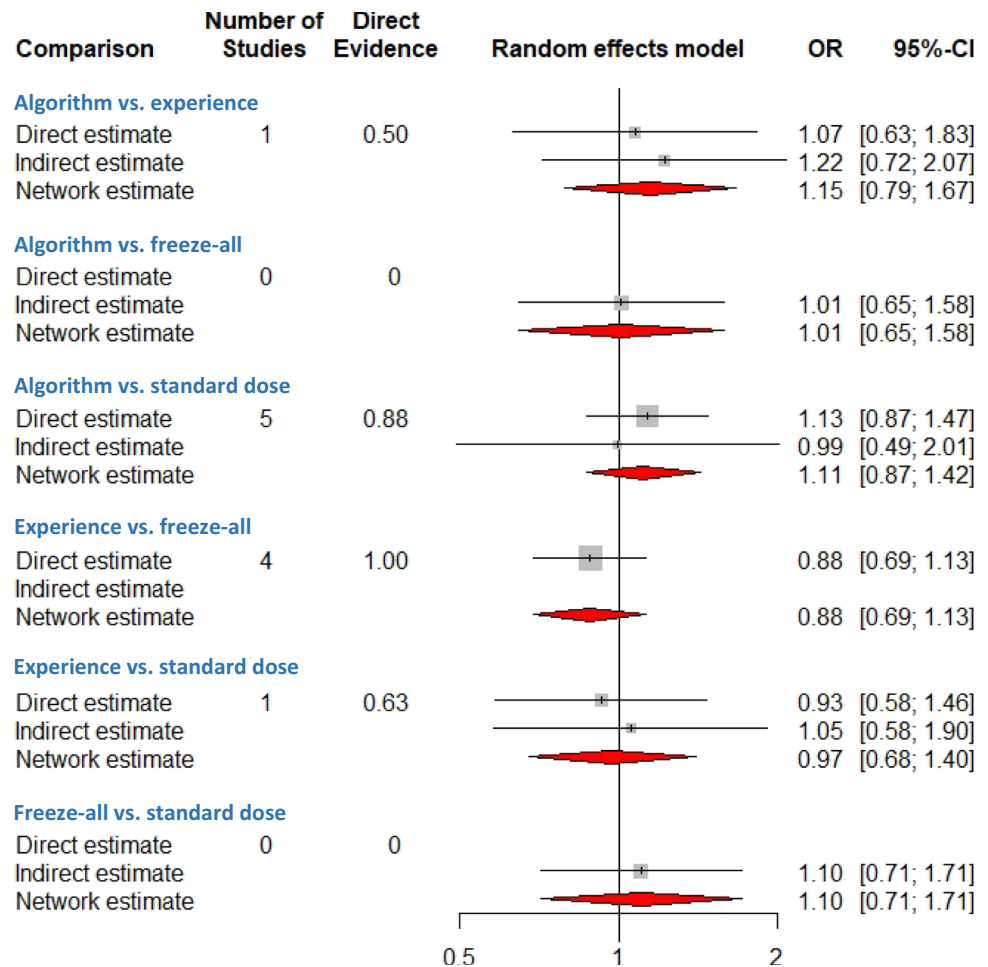
Clinical pregnancy rate (CPR)

All the thirteen RCTs reported CPR. When data were pooled, there was a low but significant level of statistical heterogeneity within studies ($\tau^2 = 0.051$; $Q = 30.30$; $p < 0.01$); with a level of inconsistency $I^2 = 67.30\%$ (38.4–82.6%), it was found to be not significant after applying the Q statistics ($Q = 0.25$, $df = 1$, $p = .615$). Assessment of inconsistency carried out also through comparison of direct and indirect evidence revealed no significant difference between the two estimates ($z = -0.251$, $p = 0.802$), and the network heat plot showed no red "hotspots" of inconsistency (Supplementary data, S3).

Freeze-all strategy was ranked first (P score = 0.708) followed by algorithm-based individualization (P score = 0.624), standard dose treatment (P score = 0.361) and clinical experience-based (P score = 0.306) individualization. Direct comparisons showed that CPR did not reveal significant differences in any of the four comparison pairs: algorithm-based vs experience-based individualization (2 studies, OR = 1.04, 95% CI 0.67–1.62), algorithm-based individualization vs standard dose (6 studies, OR = 1.10, 95% CI 0.86–1.41); clinical experience-based individualization vs freeze-all (4 studies, OR = 0.87, 95% CI 0.68–1.13) and clinical experience-based individualization vs standard dose (1 study, OR = 0.93, 95% CI 0.57–1.51). Similarly, indirect estimates showed no significant difference in the following comparison pairs: algorithm-based individualization vs freeze-all (OR = 0.96, 95% CI 0.63–1.47), algorithm-based vs clinical experience-based individualization (OR = 1.19, 95% CI 0.69–2.06), algorithm-based individualization vs standard dose (OR = 0.96, 95% CI 0.50–1.86); freeze-all vs standard dose (OR = 1.13, 95% CI 0.73–1.74) and clinical experience-based individualization vs standard dose (OR = 1.06, 95% CI 0.64–1.75) (Fig. 5). Network estimates confirmed the results of direct and indirect estimates indicating that there were no significant differences among all the four strategies.

Network meta-regression (NMR)

Although analysis of inconsistency provided evidence of low inconsistency between studies for the primary outcome (OHSS), it did not produce similar conclusive evidence regarding the other two outcomes (LBR and CPR). In order to explore potential influence of patient's characteristics on these outcomes, NMR models included main effects and interaction terms for the three available potential effect modifiers: number of retrieved oocytes, BMI and total dose of gonadotropins. Level of oestradiol peak was not included due to missing data. Results showed that the introduction of such moderators did not significantly produce better fit in comparison with the unadjusted model neither

Fig. 4 Direct and indirect comparison of live birth rate between the different strategies

Comparison	Number of Studies	Direct Evidence	Random effects model	OR	95%-CI
Algorithm vs. experience					
Direct estimate	1	0.50		1.07	[0.63; 1.83]
Indirect estimate				1.22	[0.72; 2.07]
Network estimate				1.15	[0.79; 1.67]
Algorithm vs. freeze-all					
Direct estimate	0	0			
Indirect estimate				1.01	[0.65; 1.58]
Network estimate				1.01	[0.65; 1.58]
Algorithm vs. standard dose					
Direct estimate	5	0.88		1.13	[0.87; 1.47]
Indirect estimate				0.99	[0.49; 2.01]
Network estimate				1.11	[0.87; 1.42]
Experience vs. freeze-all					
Direct estimate	4	1.00		0.88	[0.69; 1.13]
Indirect estimate					
Network estimate				0.88	[0.69; 1.13]
Experience vs. standard dose					
Direct estimate	1	0.63		0.93	[0.58; 1.46]
Indirect estimate				1.05	[0.58; 1.90]
Network estimate				0.97	[0.68; 1.40]
Freeze-all vs. standard dose					
Direct estimate	0	0			
Indirect estimate				1.10	[0.71; 1.71]
Network estimate				1.10	[0.71; 1.71]

for LBR ($DIC_{unadj} = 39.7$, $DIC_{ret\_ooc} = 39.9$, $DIC_{bmi} = 40.3$, $DIC_{t\_dos} = 42.7$) nor for CPR ($DIC_{unadj} = 46.2$, $DIC_{ret\_ooc} = 49.0$, $DIC_{bmi} = 47.7$, $DIC_{t\_dos} = 49.7$). All the 95% credible interval for the regression coefficient included zero showing that adjusting for retrieved oocytes, BMI and total dose did not explain substantial amounts of heterogeneity in the data. Outputs of NMR were shown in Supplementary data, S4.

Discussion

The present systematic review and network meta-analysis indicates that, in normal responders, both the algorithm-based individualization of the starting gonadotropin dose and the freeze-all strategy show a similar significant effect in reducing OHSS when compared with conventional treatments.

This is one of the main results of the present network meta-analysis. Another important point to be underlined is that only the individualized algorithm-based strategy for the starting dose reduces the OHSS risk in a similar way to freeze-all strategy. In fact, the individualization of the
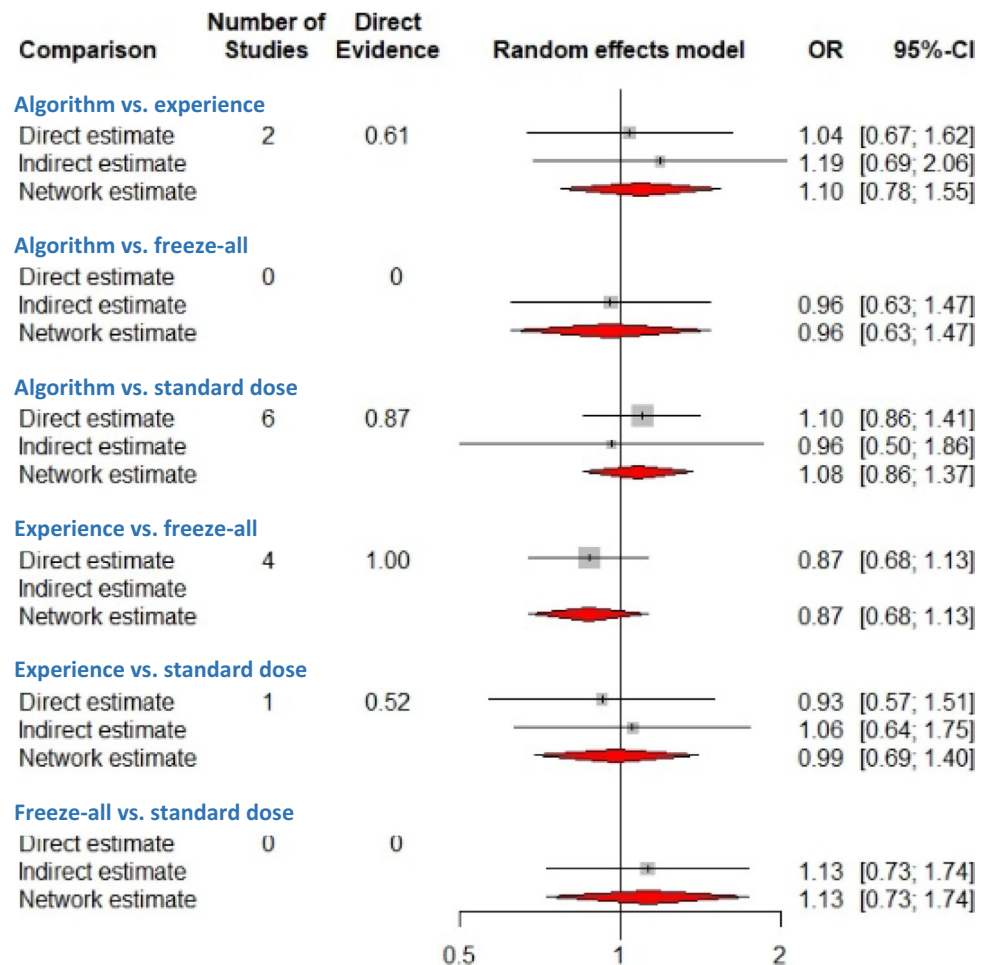
starting dose obtained through the experience of the physician, also using the ovarian reserve markers without a mathematic tool, and a standard dosage of 150 IU are associated with a higher significant OHSS risk in comparison with the freeze-all strategy. Furthermore, the algorithm-based strategy significantly reduces the OHSS risk if compared with the standard dose (direct estimates) and with the clinical experience strategy (network estimates).

Indeed, albeit in the era of the freeze-all policy, the personalization of the starting gonadotropin dose obtained by using specific mathematic tools should be considered a safe and valid option during an IVF program.

We are aware that the freeze-all policy has represented a great improvement in the management of an IVF cycle, apart from the reduction of OHSS, also for the improvement of the cumulative live birth rate with a single operation [8, 36].

Nevertheless, the possibility of carrying out the fresh embryo transfer may be considered more physiologic by some patients, and, usually, women are more likely to become pregnant and have a live birth in the shortest time. Couples who undergo an IVF cycle have the expectation that the cycle will end with the transfer of fresh embryos and,

Fig. 5 Direct and indirect comparison of clinical pregnancy rate between the different strategies

Comparison	Number of Studies	Direct Evidence	Random effects model	OR	95%-CI
Algorithm vs. experience					
Direct estimate	2	0.61		1.04	[0.67; 1.62]
Indirect estimate				1.19	[0.69; 2.06]
Network estimate				1.10	[0.78; 1.55]
Algorithm vs. freeze-all					
Direct estimate	0	0			
Indirect estimate				0.96	[0.63; 1.47]
Network estimate				0.96	[0.63; 1.47]
Algorithm vs. standard dose					
Direct estimate	6	0.87		1.10	[0.86; 1.41]
Indirect estimate				0.96	[0.50; 1.86]
Network estimate				1.08	[0.86; 1.37]
Experience vs. freeze-all					
Direct estimate	4	1.00		0.87	[0.68; 1.13]
Indirect estimate					
Network estimate				0.87	[0.68; 1.13]
Experience vs. standard dose					
Direct estimate	1	0.52		0.93	[0.57; 1.51]
Indirect estimate				1.06	[0.64; 1.75]
Network estimate				0.99	[0.69; 1.40]
Freeze-all vs. standard dose					
Direct estimate	0	0			
Indirect estimate				1.13	[0.73; 1.74]
Network estimate				1.13	[0.73; 1.74]

possibly, with a pregnancy. The freeze-all strategy is often seen as a delay in reaching the coveted goal. In fact, it is possible that the pregnancy may occur, on average, a longer time after the egg collection, in comparison with the fresh ET. Furthermore, when a fresh embryo transfer is scheduled, it is also possible that supernumerary embryos may be frozen for subsequent transfers, and, as described in the meta-analysis by Zaat et al. [7], freeze-all strategy is not superior to conventional strategy (ET fresh + frozen) in terms of cumulative live birth rate and ongoing pregnancy rate.

Additionally, in a recent study, it was observed that a prolonged storage time after vitrification negatively impacts on clinical pregnancy and live birth rates [37].

For all these reasons, we believe that the algorithm-based individualization of the starting dose followed by a fresh embryo transfer could not be considered a simple "realm of the past", as stated by Broekmans [14]. As a further demonstration of how important it is still considered to identify the correct starting dose of gonadotropins through a mathematical model, a very recent prospective observational study has been published including BMI, AMH and AFC as independent factors predictors of starting dose [38]. Moreover,

another recent retrospective, observational cohort study validated a nomogram for predicting the number of oocytes retrieved in COS cycles [39].

To date, as far as we know, no direct comparison between the algorithm-based individualization of the starting gonadotropin dose and the freeze-all policy has been performed; therefore, we conducted a network meta-analysis of all available randomized controlled trials which compared these two different strategies with conventional treatments. Network meta-analysis represents a credible way to test the effectiveness of different management strategies, even in the absence of trials making direct comparisons, and may arrange the treatment strategies to inform clinical decision making. The additional value of our meta-analysis is that we distinguished between an individualization of the starting dose based on algorithms by individualization based on the experience of the physician, demonstrating superiority in terms of safety of the first treatment in comparison with the second one.

This systematic review was performed according to the PRISMA statement, thereby securing a high methodological quality. The total number of patients treated with one of the two strategies compared was above 4,000, thereby

minimizing the effect of statistical heterogeneity correlated with small samples.

All the studies included were randomized, controlled trials, considering a population of normal responders. We decided to exclude patients at the two "extremes" of ovarian response (poor and high ovarian reserve). Patients affected by polycystic ovarian syndrome were also excluded. The reason for not considering patients defined as hyper-responders was dictated by the fact that in high responder patients, the freeze-all strategy, with the use of the GnRH agonist for the induction of oocyte maturation, is recommended [40]; thus, the fresh embryo transfer in this specific category of patients should be avoided, in order to minimize OHSS and considering also the improvement in live birth rate with the frozen embryo transfer [6]. On the contrary, the poor responders obviously do not experience OHSS, and therefore, they are outside the aim of the present study.

Our results are in line with the published literature. We showed that the freeze-all strategy reduces the incidence of OHSS in comparison with the fresh ET, as already indicated by two meta-analyses [5, 7]. In the same way, the algorithm-based individualization of the starting gonadotropin dose reduces the occurrence of OHSS in comparison with a standard dose as described in a recent meta-analysis [13].

The main novelty of our meta-analysis, differently from those described above, is that for the first time as far as we know, we tried to systemize the different approaches of ovarian stimulation which are generally not distinct in terms of strategies.

Important new information is added in the field through the fact that women who received a starting dose based on clinical experience had a significant higher risk of OHSS in comparison with the freeze-all group, with no difference being found between freeze-all and the individualized dose based on algorithms, demonstrating indirectly that the combination of different predictive factors in a mathematic model allows to establish the starting dose more objectively. Furthermore, the identification of a statistical significance in the network estimates between algorithm-based and clinical experience-based strategies for OHSS risk opens the possibility to better analyze this aspect in RCT focused on this intriguing feature.

In the studies focusing on the freeze-all policy, the trigger of oocyte maturation was induced by hCG, except in the study by Stormlund et al. (2020) [26]. Consequently, the described cases of OHSS, in the freeze-all group, are expressions of early-onset OHSS.

The induction of oocyte maturation with hCG may, thus, represent a reason for caution in the expression of the results of the present study, considering that the freeze-all policy is, in very recent times, often carried out with the induction of the final oocyte maturation by GnRH agonists, probably with a lower incidence of OHSS.

Nevertheless, we must not forget that, worldwide, in routine clinical activity, a long GnRH agonist protocol is frequently used in a normal responder population, and this obliges one to use hCG for triggering. Furthermore, the use of GnRH agonists for an ovulation trigger may not always be applicable, since some patients show abnormally low LH levels during late follicular phase, and this has a detrimental effect on oocyte maturation, ending up with the empty follicle syndrome [41]. Moreover, the latest guidelines provided by ESHRE [40] indicate in the predicted normal responders the use of hCG for the induction of the final oocyte maturation, in all kinds of stimulation protocol.

Additionally, when also using GnRH agonists for triggering, cases of severe OHSS have been reported [9, 42], probably correlated to the GnRH receptor, FSH receptor, or LH receptor gene mutations.

Considering the analyses for the secondary outcomes, no difference was observed for live birth rate and clinical pregnancy rate. For the freeze-all strategy, this is in line with the meta-analyses by Bosdou et al. (2019) [6] and Zaat et al. (2021) [7] and contradicts the meta-analysis by Roque et al. (2019) [5] where a positive effect of eFET in live birth rate was indicated. Concerning this latter study, the authors themselves noted that, after exclusion of a PGT-A study, low quality evidence indicated that there were no differences in LBR through the use of eFET in preference to fresh ET in the overall (non-PGT-A) population undergoing IVF/ICSI.

In the paper by Bosdou et al. (2019) [6], a positive effect of frozen ET was only observed in hyper-responders. The reason for this effect in hyper-responders may be due to the impairment of endometrial receptivity which a sustained ovarian stimulation might exert in a fresh cycle. Indeed, the development of the endometrium, in a subsequent cycle, under less intensive preparation regimens may provide a more favourable uterine environment for embryo implantation.

We might hypothesize that, in normal responders, especially so if the gonadotropins dose is individualized through an algorithm, the effect of supra-physiological oestradiol levels is much less pronounced.

We must highlight, however, that in the studies considered for the present meta-analysis, only the first frozen ET was considered, thus not including the others possible frozen ET which could, as widely demonstrated, improve the cumulative live birth rate.

For the algorithm-based individualization strategy, the absence of an improvement in the success outcomes of IVF led many physicians to not consider it in favour of a standardized dose.

We should emphasize that the primary objective of the individualization of the gonadotropin dose is to improve the

safety of the ovarian stimulation, rather than increase the live birth rate.

Another aspect which should be emphasized is that the freeze-all policy may be linked with certain concerns. Although most obstetrical and perinatal outcomes seem to be better when following a FET, some evidence supports the fact that FET may be associated with an increased incidence of large-for-gestational-age in singletons. Furthermore, evidence suggest that the risk of pre-eclampsia and pregnancy-complications is higher with FET than with fresh ET [5, 7]. A possible explanation for the increased risk of pre-eclampsia, proposed by Roque et al. (2019) [5], may be linked to endometrial priming with estrogens performed during artificial FET cycles. No difference was observed for miscarriage and ectopic pregnancy rates.

Furthermore, the "sine die" cryopreservation of the embryos implies ethical and legal concerns (in some countries, such as Italy) which should be taken in consideration.

Some limitations and reasons for caution are presented in this paper. The main limitation is due to the fact that the comparison of the studies was indirect because no published study directly compared the freeze-all strategy with the algorithm-based individualization of the starting dose. For this reason, randomized controlled trials comparing the freeze-all policy and the algorithm-based individualization strategy are necessary. Moreover, in the studies selected, the definition of normal responders, albeit similar, was not exactly the same.

In conclusion, the demonstration that an algorithm-based individualization strategy is equivalent to freeze-all policy in preventing OHSS is good news. In our opinion, there has been a great misunderstanding with the concept of personalization of the starting dose of gonadotropins, whose goal is not to increase the rate of live births (in this sense, probably, the embryology laboratory has a much greater role) but to improve the safety of the treatments. This goal may be achieved when a mathematical tool is used for tailoring the starting dose. The experience of the physicians alone, even based on ovarian reserve markers and the age of the woman, does not seem to be useful in reducing OHSS risk.

Furthermore, the role of modulating the starting dose of gonadotropins is much more significant when a GnRH agonist protocol is used, because the induction of oocyte maturation, in this case, may be obtained only by hCG, with the possibility of early-onset OHSS, even in the case of a freeze-all program.

Finally, we do not want to lean towards one option rather than the other, but we must emphasize that in cases where the freeze-all strategy is not desired or cannot be carried out, the personalization of the initial dose of gonadotropins, obtained by the use of specific algorithms based on the markers of ovarian reserve, should be pursued.

Even more than for the other branches of medicine, in reproductive medicine "primum non nocere" must represent the polar star for physicians. The concept of "more oocytes is better", which determines the use of more gonadotropins, as observed in freeze-all cycles, may be associated with important health concerns, apart from the OHSS, such as ovarian torsion and thromboembolic events, which although rare occur mainly when 15 or more oocytes are retrieved [43].

Therefore, the results of this systematic review and network meta-analysis must be kept in mind when the physician is scheduling an ovarian stimulation.

Declarations

References

1. European IVF-Monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Wyns C, De Geyter C, Calhaz-Jorge C, Kupka MS, Motrenko T, Smeenk J, Bergh C, Tandler-Schneider A, Rugescu IA, Vidakovic S, Goossens V. ART in Europe, 2017: results generated from European registries by ESHRE. Hum Reprod Open. 2021;3:hoab026.
2. Mourad S, Brown J, Farquhar C. Interventions for the prevention of OHSS in ART cycles: an overview of Cochrane reviews (review). Cochrane Database Syst Rev. 2017;1:CD012103.
3. Groenewoud ER, Cohlen BJ, Macklon NS. Programming the endometrium for deferred transfer of cryopreserved embryos: hormone replacement versus modified natural cycles. Fertil Steril. 2018;109:768–74.
4. Loutradi KE, Kolibianakis EM, Venetis CA, Papanikolaou EG, Pados G, Bontis I, Tarlatzis B. Cryopreservation of human embryos by vitrification or slow freezing: a systematic review and meta-analysis. Fertil Steril. 2008;90:186–93.
5. Roque M, Haahr T, Geber S, Esteves SC, Humaidan P. Fresh versus elective frozen embryo transfer in IVF/ICSI cycles: a systematic review and meta-analysis of reproductive outcomes. Hum Reprod Update. 2019;25:2–14.
6. Bosdou JK, Venetis CA, Tarlatzis BC, Grimbizis GF, Kolibianakis EM. Higher probability of live-birth in high, but not normal, responders after first frozen-embryo transfer in a freeze-only cycle strategy compared to fresh-embryo transfer: a meta-analysis. Hum Reprod. 2019;34:491–505.
7. Zaat T, Zagers M, Mol F, Goddijn M, van Wely M, Mastenbroek S. Fresh versus frozen embryo transfers in assisted reproduction. Cochrane Database Syst Rev. 2021;2:CD011184.
8. Polyzos NP, Drakopoulos P, Parra J, Pellicer A, Santos-Ribeiro S, Tournaye H, Bosch E, Garcia-Velasco J. Cumulative live birth rates according to the number of oocytes retrieved after the first ovarian stimulation for in vitro fertilization/intracytoplasmic sperm injection: a multicenter multinational analysis including approximately 15,000 women. Fertil Steril. 2018;110:661–70.

9. Fatemi HM, Popovic-Todorovic B, Humaidan P, Kol S, Banker M, Devroey P, García-Velasco JA. Severe ovarian hyperstimulation syndrome after gonadotropin-releasing hormone (GnRH) agonist trigger and "freeze-all" approach in GnRH antagonist protocol. Fertil Steril. 2014;101:1008–11.

10. Baker VL, Brown MB, Luke B, Conrad KP. Association of number of retrieved oocytes with live birth rate and birth weight: an analysis of 231.815 cycles of in vitro fertilization. Fertil Steril. 2015;103:931–8.

11. 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;26:1768–74.

12. Sunkara SK, La Marca A, Seed PT, Khalaf Y. Increased risk of preterm birth and low birth weight with very high number of oocytes following IVF: an analysis of 65.868 singleton live birth outcomes. Hum Reprod. 2015;30:1473–80.

13. Lensen SF, Wilkinson J, Leijdekkers JA, La Marca A, Mol BWJ, Marjoribanks J, Torrance H, Broekmans FJ. Individualised gonadotropin dose selection using markers of ovarian reserve for women undergoing in vitro fertilisation plus intracytoplasmic sperm injection (IVF/ICSI) (Review). Cochrane Database Syst Rev. 2018;2:CD012693.

14. Broekmans F. Individualization of FSH doses in assisted reproduction: facts and fiction. Front Endocrinol. 2019;10:181.

15. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L. PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Br Med J. 2015;350:g7647.

16. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP, Mulrow C, Catalá-López F, Gøtzsche PC, Dickersin K, Boutron I, Altman DG, Moher D. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med. 2015;162:777–84.

17. Haynes RB. Forming research questions. J Clin Epidemiol. 2006;59:881–6.

18. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JAC. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

19. Efthimiou O, Debray TPA, van Valkenhoef G, Trelle S, Panayidou K, Moons KGM, Reitsma J, Shang A, Salanti G. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods. 2016;7:236–63.

20. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res. Synth Methods. 2012;3:98–110.

21. Krahn U, Binder H, König J. A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol. 2013;13:35.

22. Dias S, Sutton AJ, Welton NJ, Ades AE. Heterogeneity: subgroups, meta-regression, bias and bias-adjustment [internet]. London: National Institute for Health and Care Excellence (NICE). 2012; NICE DSU Technical Support Document No. 3.

23. Shi Y, Sun Y, Hao C, Zhang H, Wei D, Zhang Y, Zhu Y, Deng X, Qi X, Li H, Ma X, Ren H, Wang Y, Zhang D, Wang B, Liu F, Wu Q, Wang Z, Bai H, et al. Transfer of fresh versus frozen embryos in ovulatory women. N Engl J Med. 2018;378:126–36.

24. Vuong LN, Dang VQ, Ho TM, Huynh BG, Ha DT, Pham TD, Nguyen LK, Norman RJ, Mol BW. IVF transfer of fresh or frozen embryos in women without polycystic ovaries. N Engl J Med. 2018;378:137–47.

25. Wei D, Liu JY, Sun Y, Shi Y, Zhang B, Liu JQ, Tan J, Liang X, Cao Y, Wang Z, Qin Y, Zhao H, Zhou Y, Ren H, Hao G, Ling X, Zhao J, Zhang Y, Qi X, et al. Frozen versus fresh single blastocyst transfer in ovulatory women: a multicentre, randomised controlled trial. Lancet. 2019;393:1310–8.

26. Stormlund S, Sopa N, Zedeler A, Bogstad J, Prætorius L, Nielsen HS, Kitlinski ML, Skouby SO, Mikkelsen AL, Spangmose AL, Jeppesen JV, Khatibi A, la Cour FN, Ziebe S, Polyzos NP, Bergh C, Humaidan P, Andersen AN, Løssl K, Pinborg A. Freeze-all versus fresh blastocyst transfer strategy during in vitro fertilisation in women with regular menstrual cycles: multicentre randomised controlled trial. BMJ. 2020;370:m2519.

27. Popovic-Todorovic B, Loft A, Bredkjaeer HE, Bangsbøll S, Nielsen IK, 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;18:2275–82.

28. 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;30:248–57.

29. Nyboe Andersen A, Nelson SM, Fauser BC, García-Velasco JA, Klein BM, Arce JC. ESTHER-1 study group. Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial. Fertil Steril. 2017;107:387–96.

30. Allegra A, Marino A, Volpes A, Coffaro F, Scaglione P, Gullo S, 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;34:429–38.

31. Bosch E, Havelock J, Martin FS, Rasmussen BB, Klein BM, Mannaerts B, Arce JC. ESTHER-2 Study Group. Follitropin delta in repeated ovarian stimulation for IVF: a controlled, assessor-blind phase 3 safety trial. Reprod Biomed Online. 2019;38:195–205.

32. Friis Petersen J, Løkkegaard E, Andersen LF, Torp K, Egeberg A, Hedegaard L, Nysom D, Nyboe AA. A randomized controlled trial of AMH-based individualized FSH dosing in a GnRH antagonist protocol for IVF. Hum Reprod Open. 2019;1:hoz003.

33. Ishihara O, Arce JC, for the Japanese Follitropin Delta Phase 3 Trial (STORK) Group. Individualized follitropin delta dosing reduces OHSS risk in Japanese IVF/ICSI patients: a randomized controlled trial. Reprod Biomed Online. 2021;42:909–18.

34. Qiao J, Zhang Y, Liang X, Ho T, Huang H-Y, Kim S-K, Goethberg M, Mannaerts B, Arce JC, on behalf of the Asian follitropin delta phase 3 trial (GRAPE) group. A randomised controlled trial to clinically validate follitropin delta in its individualised dosing regimen for ovarian stimulation in Asian IVF/ICSI patients. Hum Reprod. 2021;36:2452–62.

35. van Tilborg TC, Oudshoorn SC, Eijkemans MJC, Mochtar MH, van Golde RJT, Hoek A, Kuchenbecker WKH, Fleischer K, de Bruin JP, Groen H, van Wely M, Lambalk CB, Laven JSE, Mol BWJ, Broekmans FJM, Torrance HL, OPTIMIST study group. Individualized FSH dosing based on ovarian reserve testing in women starting IVF/ICSI: a multicentre trial and cost-effectiveness analysis. Hum Reprod. 2017;32:2485–95.

36. Zhao Z, Shi H, Li J, Zhang Y, Chen C, Guo Y. Cumulative live birth rates according to the number of oocytes retrieved following the "freeze-all" strategy. Reprod Biol Endocrinol. 2020;18:14.

37. Li J, Yin M, Wang B, Lin J, Chen Q, Wang N, Lyu Q, Wang Y, Kuang Y, Zhu Q. The effect of storage time after vitrification on pregnancy and neonatal outcomes among 24 698 patients following the first embryo transfer cycles. Hum Reprod. 2020;35:1675–84.

38. Li Y, Duan Y, Yuan X, Cai B, Xu Y, Yuan Y. A novel nomogram for individualized gonadotropin starting dose in GnRH antagonist protocol. Front Endocrinol (Lausanne). 2021;12:688654.

39. Ebid AHIM, Motaleb SMA, Mostafa MI, Soliman MMA. Novel nomogram-based integrated gonadotropin therapy individualization in in vitro fertilization/intracytoplasmic sperm injection: a modeling approach. Clin Exp Reprod Med. 2021;48:163–73.

40. The ESHRE Guideline Group on Ovarian Stimulation, Bosch E, Broer S, Griesinger G, Grynberg M, Humaidan P, Kolibianakis E, Kunicki M, La Marca A, Lainas G, Le Clef N, Massin N, Mastenbroek S, Polyzos N, Sunkara SK, Timeva T, Töyli M, Urbancsek J, Vermeulen N, Broekmans F. ESHRE guideline: ovarian stimulation for IVF/ICSI. Hum Reprod Open. 2020;2:hoaa009.

41. Dosouto C, Haahr T, Humaidan P. Advances in ovulation trigger strategies. Panminerva Med. 2019;61:42–51.

42. Gurbuz AS, Gode F, Ozcimen N, Isik AZ. Gonadotrophin-releasing hormone agonist trigger and freeze-all strategy does not prevent severe ovarian hyperstimulation syndrome: a report of three cases. Reprod Biomed Online. 2014;29:541–4.

43. Magnusson Å, Källen K, Thurin-Kjellberg A, Bergh C. The number of oocytes retrieved during IVF: a balance between efficacy and safety. Hum Reprod. 2018;33:58–64.