

Mild versus conventional ovarian stimulation for IVF in poor, normal and hyper-responders: a systematic review and meta-analysis

Adrija Kumar Datta  ^{1,*}, Abha Maheshwari  ², Nirmal Felix ¹,
Stuart Campbell ^{3,4}, and Geeta Nargund ^{4,5}

¹Create Fertility, Birmingham, UK ²Aberdeen Fertility Centre and NHS Grampian, Aberdeen, UK ³St George's University of London, London, UK ⁴Create Fertility, London, UK ⁵St Georges University Hospitals NHS Trust London, London, UK

*Correspondence address. Create Fertility, 6270 Bishops Court, Birmingham Business Park, Solihull Parkway, Birmingham B37 7YB, UK.
Tel: 03332407300; E-mail: adrija@createfertility.co.uk <https://orcid.org/0000-0002-3764-9298>

Submitted on March 23, 2020; resubmitted on July 6, 2020; editorial decision on July 21, 2020

TABLE OF CONTENTS

- Introduction
 - Criteria for including studies in this review
 - Search method
 - Search terms
 - Data collection and analysis
 - Data extraction and management
 - Assessment of risk of bias
 - Treatment effect
 - Missing data
 - Assessment of heterogeneity
 - Reporting bias
 - Subgroup analysis
 - Multi-arm studies
 - Sensitivity analysis
- Results
 - Characteristics of included studies
 - Risk of bias of the included studies
 - Primary outcomes
 - Secondary Outcomes
- Discussion
 - Findings of the review
 - Strength of this review
 - Limitations
 - Comparison with other reviews
 - Implications for clinical practice
 - Implications for research

BACKGROUND: Mild ovarian stimulation has emerged as an alternative to conventional IVF with the advantages of being more patient-friendly and less expensive. Inadequate data on pregnancy outcomes and concerns about the cycle cancellation rate (CCR) have prevented mild, or low-dose, IVF from gaining wide acceptance.

OBJECTIVE AND RATIONALE: To evaluate parallel-group randomised controlled trials (RCTs) on IVF where comparisons were made between a mild (≤ 150 IU daily dose) and conventional stimulation in terms of clinical outcomes and cost-effectiveness in patients described as poor, normal and non-polycystic ovary syndrome (PCOS) hyper-responders to IVF.

SEARCH METHODS: Searches with no language restrictions were performed using Medline, Embase, Cochrane central, Pre-Medicine from January 1990 until April 2020, using pre-specified search terms. References of included studies were hand-searched as well as advance access articles to key journals. Only parallel-group RCTs that used ≤ 150 IU daily dose of gonadotrophin as mild-dose IVF (MD-IVF) and compared with a higher conventional dose (CD-IVF) were included. Studies were grouped under poor, normal or hyper-responders as described by the authors in their inclusion criteria. Women with PCOS were excluded in the hyper-responder group. The risk of bias was assessed as per Cochrane Handbook for the included studies. The quality of evidence (QoE) was assessed according to the GRADE system. PRISMA guidance was followed for review methodology.

OUTCOMES: A total of 31 RCTs were included in the analysis: 15 in the poor, 14 in the normal and 2 in the hyper-responder group. Live birth rates (LBRs) per randomisation were similar following use of MD-IVF in poor (relative risk (RR) 0.91 (CI 0.68, 1.22)), normal (RR 0.88 (CI 0.69, 1.12)) and hyper-responders (RR 0.98 (CI 0.79, 1.22)) when compared to CD-IVF. QoE was moderate. Cumulative LBRs (5 RCTs, $n = 2037$) also were similar in all three patient types (RR 0.96 (CI 0.86, 1.07)) (moderate QoE). Risk of ovarian hyperstimulation syndrome was significantly less with MD-IVF than CD-IVF in both normal (RR 0.22 (CI 0.10, 0.50)) and hyper-responders (RR 0.47 (CI 0.31, 0.72)), with moderate QoE. The CCRs were comparable in poor (RR 1.33 (CI 0.96, 1.85)) and hyper-responders (RR 1.31 (CI 0.98, 1.77)) but increased with MD-IVF among normal responders (RR 2.08 (CI 1.38, 3.14)); all low to very low QoE. Although fewer oocytes were retrieved and fewer embryos created with MD-IVF, the proportion of high-grade embryos was similar in all three population types (low QoE). Compared to CD-IVF, MD-IVF was associated with less gonadotrophin use and lower cost.

WIDER IMPLICATIONS: This updated review provides reassurance on using MD-IVF not only for the LBR per cycle but also for the cumulative LBR, with moderate QoE. With risks identified with 'freeze-all' strategies, it may be time to recommend mild-dose ovarian stimulation for IVF for all categories of women i.e. hyper, poor and normal responders to IVF.

Key words: mild ovarian stimulation / low-dose stimulation / conventional IVF / poor responders / low ovarian reserve / hyper-responders / normal responders / systematic review / meta-analysis

Introduction

Mild stimulation IVF is defined as 'a protocol in which the ovaries are stimulated with gonadotrophins, and/or other pharmacological compounds, with an intention of limiting the number of oocytes following stimulation for IVF' according to the International Glossary on Infertility and Fertility Care by the International Committee for Monitoring Assisted Reproductive Technologies (ICMART) (Zegers-Hochschild et al., 2009). In contrast, conventional-dose IVF (CD-IVF) relies on the concept that, by increasing the stimulation dose and obtaining more oocytes the pregnancy rates can be improved and hence the more oocytes the better (Drakopoulos et al., 2016).

Despite reported benefits of less treatment-related stress, better tolerance (Hogaard et al., 2001; de Klerk et al., 2007) and lower treatment cost (Heijnen et al., 2005), low or mild-dose ovarian stimulation for IVF has not achieved wide acceptance in the Reproductive Medicine community, primarily due to the concerns about its clinical effectiveness. Several trials and systematic reviews on this topic have been published; however, the controversy about clinical and cost-effectiveness as well as patient acceptability of mild-dose IVF (MD-IVF) protocols continues (Nargund et al., 2017).

Although MD-IVF gained momentum a decade ago due to reduction in the risk of ovarian hyperstimulation syndrome (OHSS), this enthusiasm was curtailed with the widespread adoption of conventional stimulation with an aim to maximise oocyte number and, if needed,

'freeze-all embryos'. This policy is based on the concept of eliminating OHSS and completing a 'family' from a single oocyte collection cycle. However, current evidence shows that the risk of OHSS is not completely eliminated and there is no prospective study to prove that a single cycle allows a 'family' to be completed. Therefore, 'more is better' is not a reality (Nargund and Fauer, 2020). In addition, there is increasing uncertainty about the benefits of 'freeze all' strategy (Roque et al., 2019). As a result, there is a resurgence of interest around MD-IVF as a first-line treatment.

Several systematic reviews have compared MD-IVF with CD-IVF but the definition of 'mild' stimulation for IVF has varied; the majority compared studies with anti-oestrogens, clomiphene citrate (CC) or aromatase inhibitors (AIs) combined with low-dose gonadotropin versus conventional protocols without oral compounds (Gibreel et al., 2012; Bechtejew et al., 2017; Fan et al., 2017; Kamath et al., 2017), while others compared low- versus high-dose gonadotrophin only regimens (Sterrenburg et al., 2011), or analysed gonadotrophin only protocols as a separate subgroup (Youssef et al., 2018). The upper limit of the gonadotrophin dose to qualify as a 'mild'- or 'low'-dose IVF protocol was often not specified. Some reviews were limited to studies on poor responders (Song et al., 2016; Youssef et al., 2018) while others included unselected populations (Matsaseng et al., 2013) or presented data on the poor responders in a separate subgroup analysis (Fan et al., 2017; Kamath et al., 2017). The sole review to date on the hyper-responders was a narrative, without a meta-analysis (Gat et al., 2015).

The American Society for Reproductive Medicine (ASRM) Practice Committee proposed a daily dose of ≤ 150 IU gonadotrophin (with or without oral compounds) to be considered as 'mild ovarian stimulation' (Practice Committee of the American Society for Reproductive Medicine. Electronic Address: ASRM@asrm.org, 2018). Except for the review by ASRM, this definition has not been used in any meta-analysis previously. In addition, existing reviews were under-powered for sample size to compare live birth outcome and included mostly small studies with high risk of bias (RoB); thus, an updated systematic review including only randomised controlled trials (RCTs) that used no more than 150 IU daily dose of gonadotrophin with and without oral compounds (CC or AIs) as MD-IVF in all clinical settings (for poor, normal or hyper-responders of IVF) became necessary.

The objective of this systematic review was to evaluate MD-IVF (≤ 150 IU daily dose of gonadotrophin alone, or in combination with oral compounds) in randomised studies by comparing its clinical effectiveness, risks and cost with those of conventional (higher-dose stimulation) IVF protocols (CD-IVF) in patients identified as poor, normal and hyper-responders to IVF.

Methods

The Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green, 2011) was followed to conduct this review and meta-analysis and the findings were presented according to the PRISMA guideline. Registration number: PROSPERO 2018 CRD42018104879 (for poor responders of IVF), PROSPERO CRD42019150069 (for normal/high responders of IVF).

Criteria for including studies in this review

There was no restriction on language. We included studies from January 1990 (since the introduction of the concept of poor or high ovarian response in IVF) to April 2020. Abstracts or conference proceedings were also reviewed and included, avoiding duplication, only if all required information was available. Studies were excluded if complete information was not obtained despite personal request.

Type of study: RCT with parallel-group comparison

Participants. Couples underwent IVF/ICSI due to any cause, where the female partners were known or anticipated to have normal, high or poor response to ovarian stimulation. We went by the inclusion criteria as described by the authors to define the population as poor, normal (unselected) or hyper-responders and grouped the trials accordingly.

Poor responders: where women were predicted to have low ovarian reserve based on elevated basal follicle stimulation hormone (FSH) and/or low anti-Mullerian hormone (AMH) and/or low antral follicle count (AFC) and/or low ovarian response in the previous cycle and those who met the Bologna criteria (Ferraretti et al., 2011).

Normal responders: where the age of the women or ovarian reserve or previous ovarian response, as described by the authors, predicted to result in a not too low, or too high ovarian response. The definition of normal responders is based on predicted response only; some women might have had an unexpected exaggerated response

while some others an unpredicted poor response. This limitation has been accepted, in absence of any better marker to denote 'normal responders'.

Hyper-responders: where women were predicted to yield high ovarian response based on high AMH and/or high AFC and/or exaggerated follicular response in the previous cycle, except where a diagnosis of typical polycystic ovary syndrome (PCOS) was made.

If there is no mention of age or ovarian reserve in the primary study, we have classified them as 'unselected population' and included the data under the normal responders for meta-analysis.

Intervention. MD-IVF: Treatment protocol using 'mild' or low-dose (≤ 150 IU daily) gonadotrophin (FSH or hMG) alone, or in combination with oral compounds (e.g. CC/AIs) or oral compounds alone irrespective of agonist or antagonist protocol.

Comparison. CD-IVF: Protocols with gonadotrophin exposure higher than that of mild or low-dose arm in terms of daily dose and/or duration.

The search conformed to the standard descriptions of 'mild' and 'conventional' stimulation IVF protocols (Nargund et al., 2007; Zegers-Hochschild et al., 2009); but because of the varying description of these terms in the literature, we were obliged to define them on the basis of gonadotrophin dosage. This permitted the comparison of the outcomes of mild and conventional stimulation dosages of gonadotrophins (FSH and hMG) on the same population, whether daily or *de facto* 'cumulative'.

Exclusion criteria. Studies comparing oocyte or embryo yield only with no data on any of the primary outcomes measured in this review were excluded. Studies comparing a 'standard' 150 IU daily dose in one arm with a wide range of 'individualised' stimulation dosage in the other arm based on ovarian reserve were excluded.

Primary outcomes. Live birth rate (LBR) per woman randomised; OHSS and cycle cancellation rates (CCRs) per cycle started.

Secondary outcomes. Cumulative LBR, ongoing pregnancy rate (OPR), clinical pregnancy rates (CPRs) (with separate note on biochemical pregnancies) as defined in the ICMART glossary (Zegers-Hochschild et al., 2009), total dose of gonadotrophin used, number of oocytes, number of embryos, number of high-grade embryos per started cycle and cost comparison. The number of embryos transferred may not be a true reflection of total number of embryos created, therefore was not considered.

All outcomes were derived from the first or only treatment cycle with fresh embryo transfer conducted in the individual trials, except while reporting the cumulative outcomes. Cumulative live birth, whether adding data from all subsequent frozen embryo transfer cycle(s) or subsequent fresh cycles as well as frozen cycles in a given study period, were expressed as per-patient randomised. Cumulative secondary outcomes, e.g. incidence of OHSS, cycle cancellations or mean number of oocytes or embryos, were therefore reported on a per started cycle basis, counting the outcomes from all fresh cycles together.

Search method

An electronic search was conducted in Medline, Embase, PreMedline and Cochrane Central from January 1990 (inception of the concept of low or high responder) to April 2020. Databases were searched using

relevant medical subject headings, free-text terms and study type filters where appropriate, without language restrictions. Advance access articles of key journals were checked for related papers. The reference list of all reviews or individual RCTs was also hand-searched to find any additional RCT. Duplications arising from a conference abstract and subsequent full-text paper were excluded.

Search terms

((IVF, ICSI, ovarian stimulation) AND ((mild IVF stimulation, oral agents, aromatase inhibitors, clomiphene, letrozole, anastrozole) OR ((gonadotropin, FSH, follitropin, hMG, menotrophin) AND (dose, low dose))) AND randomised controlled trials]. Because of the diversity in protocols, the terms related to CD-IVF were not included in the electronic search; however, individual abstracts were reviewed to confirm eligibility of the CD-IVF protocols and to identify trials on poor, normal or high responders in IVF. The electronic search was performed by National Guideline Alliance (NGA) of Royal College of Obstetricians and Gynaecologists.

Data collection and analysis

First an electronic search was made using the search terms and databases described above. Full text of all shortlisted studies (RCTs) was reviewed by two reviewers (A.K.D. and N.F.) independently; conflict if any was resolved by any of the other reviewers (S.C. or G.N.). References of all included and excluded full-text papers and other related systematic reviews were hand-searched to look for additional RCTs. *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins and Green, 2011) was consulted to prepare the data-extraction form, obtain the features of included studies, assess RoB and outcome data. Review Manager 5 (version 5.3) software was used to construct the RoB graph, Funnel plots and Forest plots in this review (Review Manager (RevMan) (Computer program) Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014).

Data extraction and management

The following information and data were extracted.

Trial characteristics

Year and location of the trial (single or multi-centre), type of trials (2-arm/3-arm etc.), study population with sample size calculation, method of randomisation, method of allocation concealment, exclusion of participants after randomisation, proportion of and reasons for losses at follow up, reports of ethical approval and consent.

Participants

Age, ovarian reserve of the women, e.g. FSH, AMH, AFC, ovarian response in the previous IVF/ICSI cycles (if mentioned) to categorise women in poor, normal and high ovarian response groups. In addition, whether in accordance with Bologna criteria (Ferraretti et al., 2011) for poor responders, exclusion criteria of individual trials were also noted.

Intervention

Treatment protocols in the intervention and comparator group(s) with regards to the type of medications (oral and injectable), dose, time of

commencement, method of suppression of premature ovulation, dose adjustment or pre-treatment or co-intervention, if any, ovulation trigger type and dose, cancellation criteria and luteal phase regimen were noted.

Outcomes

What outcomes were reported, how the outcomes were defined and the timing of outcome measurement (e.g. per woman randomised/started cycle or per embryo transfer) were recorded. Cumulative live birth data were extracted as an aggregation of both the first fresh and all subsequent frozen transfer cycle(s) or further fresh cycle(s); data from each subsequent fresh or frozen cycle(s) were not analysed separately. In the cost analysis, whether total cost per cycle or per woman or cumulative cost of fresh and frozen cycles were noted.

Assessment of risk of bias

RoB was assessed under the headings of Sequence generation, Allocation concealment, Blinding of participants and assessors, Selective outcome reporting and Other sources of bias as outlined in *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins and Green, 2011). Blinding of patients and clinicians was neither possible nor applicable for this particular type of intervention and outcomes (e.g. pregnancy rates). We considered studies with absence of blinding as low RoB, as it was unlikely to influence outcomes. The RoB was considered 'unclear' if the information was insufficient in any type of bias.

Treatment effect

For dichotomous data, relative risk (RR) and for continuous data, mean differences (MD) between treatment groups were calculated with 95% CI. In case of anticipated heterogeneity, a random effect model was used. In presence of heterogeneous data, the standardised mean difference (SMD) was used.

Missing data

Authors were contacted for missing data by email at least twice.

Assessment of heterogeneity

The clinical and methodological characteristics of all included studies were examined (Table I); sub-group meta-analysis was performed as described below. Statistical heterogeneity was assessed by the χ^2 test. The I^2 statistic assessed the impact of the heterogeneity on the meta-analysis; an I^2 of $>50\%$ indicated significant heterogeneity, in which case a 'random effect model' was applied, otherwise, a 'fixed-effect model' was used as a default.

Reporting bias

A funnel plot was generated with all included studies on CCR outcome. We did not limit our search by any language or time.

Subgroup analysis

For each outcome, meta-analyses were performed separately for poor responders, normal and high responders. Subgroup analysis was performed with different types of mild stimulation protocols: low-dose

Table I Characteristics of included randomised controlled trials.

Author	Date	Place, Population	Trial type and Method	Participants	Interventions: Mild (MD) versus Conventional-dose (CD) protocols	Cancellatn. criteria	Outcomes (per woman randomised/ cycle)
Normal/ hyper-responders							
Baart <i>et al.</i> (2007)	2007	Single centre, 2-arm RCT; Power: for oocytes #; Consent, ethical approval: yes	Age <38 yrs; BMI 19-29*; Prev. poor response: No	MD: FSH 150 IU/d from D5, GnRH-ant (n=67) CD: Long GnRH-a DR; FSH 225 IU/d (n=44); Dose adj: No. Trigger: hCG 10000 IU. LPS: ? ET: 1-2 euploid embryo(s)	?	OPR: 19.0% vs 17.1% (NS); OHSS: 0% vs 2% (NS); CCR: 11% vs 2% (NS); # Oocytes: 8.3 vs 12.1 (p<0.01). Good quality embryos: 51% vs 35% (p=0.04)	
Blockeel <i>et al.</i> (2011)	2011	Single centre, 2-arm RCT; Power: for oocytes #; Consent, ethical approval: yes	Age 18-36 yrs; BMI 18-29*; IVF/ICSI- 1 st attempt; Regular cycle; FSH<12 IU/L; No endocrine disease/ >2° endometriosis	MD: rFSH 150 IU/d from D5 (n=40) CD: rFSH 150 IU/d from D2 (n=36); GnRH-ant in both; Dose adj: No. Trigger: hCG 10000 IU. LPS: P; pessary; ET: D3/5	<2 follicles	OPR: 25.0% vs 28 (NS); PR: 25% vs 36% (NS); OHSS: 0% vs 2.8% (NS); CCR: 17.5% vs 13.9% (NS); # Oocytes: mean 0.3 vs 8.9 (NS); Gn requirement: ↓ with MD	
Casano <i>et al.</i> (2012)	2012	Single centre, 2-arm RCT; Power: for oocyte #; Consent, ethical approval: yes	Age <38 yrs; 1 st IVF; AMH >2ng/ml; AFC >16	MD: FSH 150 IU/d from D4; GnRH-ant from D8 (n=205) CD: Long GnRH-a DR; FSH 150 IU/d (n=207); Dose adj: Yes. Trigger: hCG 10000 IU. LPS: P; pessary; ET: I-2 on D2	?	LBR: 24.8% vs 24.6% & Cum LBR: 42.7% vs 41.7% (NS); OHSS: 1.6% vs 2.0% (NS); CCR: 1% vs 0% (NS); # oocytes: mean 9.9 vs 10.3 (NS); Gn requirement: ↓ with MD	
Dhont <i>et al.</i> (1995)	1995	Single centre, 2-arm RCT; Power: for PR; Consent, ethical approval: ?	Unselected population; 1 st cycle IVF- a minority also had GIFT and ZIFT techniques	MD: CC 100 mg/d D2-6, then hMG 150 IU/d; No GnRH-ant (n= 151) CD: Long GnRH-a DR; hMG 300 IU/d (n= 152); Dose adj: Yes. COCP pre-Tx. Trigger: hCG 10000 IU. LPS: hCG or P; ET: I-3 if good, 5 if low quality embryo	<3 follicles or E2 <100 pg/ml or pre-ovulatory LH rise	LBRs: 18.5% vs 25.7% (NS); PR: 24.5% vs 36.8% (p= 0.02); OHSS: 0% vs 4.1% (p= 0.007); CCR: 25.5% vs 2.6% (p< 0.001); # Oocytes: mean 6.1 vs 14.2 (p< 0.0001); Gn requirement: ↓ with MD	
Elashar <i>et al.</i> (2016)	2016	Single centre, 2-arm RCT; Power: Not done; Consent, ethical approval: ?	Age 20-35 yrs; BMI 18-29*; AFC >5 in one ovary; AMH > 1ng/dl; Unexplained infertility	MD: Let 10 mg/d D3-7, FSH 75 IU/d from D5; D4; No GnRH-ant (n=40) CD: Long GnRH-a DR; FSH 150-225 IU/d (n= 40); Dose adj: ? Trigger: hCG. LPS: ?	?	PR: 12.5% vs 42.5% (p=0.01); # Oocyte (matre) 80% vs 79.4% (NS); # Good-quality embryo: 37.7% vs 38.8% (NS); Gn requirement: ↓ with MD	
Harrison <i>et al.</i> (1994)	1994	Single centre, 2-arm RCT; Power: not done; Consent: yes, ethical approval: ?	Unselected (not specified); 1 st IVF attempt	MD: CC 100 mg/d on D2-6, hMG 150 IU/d from D4; No GnRH-ant (n=50) CD: GnRH-a long acting DR (buserelin in 1 arm, triptorelin in another arm), hMG 225 IU/d (n=50); Dose adj: ? Trigger: hCG 10000. LPS: hCG/ P	?	LBRs: 24% vs 22% & 28% (NS); CCR: 32% vs 28% (NS); # Oocytes: mean 8.8 vs 12.03 & 11.5 (NS)	
Heijnen <i>et al.</i> (2007)	2007	Multi-centre, 2-arm RCT; Power: for LB; Consent, ethical approval: yes	Age <38 yrs; BMI 18-28*; Regular 25-35 days' cycle; 1 st IVF/ICSI attempt/ no LB from IVF.	MD: FSH dose 150 IU/d off from D5; GnRH-ant. (n=205) CD: Long GnRH-a DR; FSH 150 IU/d (n=199) Dose adj: No. Trigger: hCG. LPS: SET for MD & DET for CD	<3 follicles	LBRs: 15.8% vs 24.0% (p= 0.003); Cum LBRs: 43.4% vs 44.7% (NS); OHSS: 1.4% vs 4% (p=0.04); CCR: 8% vs 8% (p< 0.001); # Oocytes: €10745 (p=0.006) Drop-out rates: ↓ IHD-IVF. Gn requirement: ↓ with MD	

Continued

Table I Continued

Author Date Place, Population	Trial type and Method	Participants	Interventions: Mild (MD) versus Conventional-dose (CD) protocols	Cancellatn. criteria	Outcomes (per woman randomised/cycle) MD-IVF versus CD-IVF
Hohmann et al. (2003) The Netherlands n=169	Single centre, 3-arm RCT; Power: not done; Consent, ethical approval: yes	Age 20-38 yrs; BMI 19-29*; Regular cycle; No severe endometriosis/ uterine/ ovarian anomaly; <3 previous IVF; No previous poor response/ OHSS	MD: FSH 150 IU/d from D5, GnRH-ant (n=49). CD: Long GnRH-a DR; 150 IU/d FSH. (n=45). Dose adj: No. [3 rd arm had same 150 IU from D2, so excluded] Trigger: hCG 10000. LPS: P pessary. ET: 1-2 on D3/5.	<3 follicles	OPR/s cycle: 16% vs 18% (NS); Premature ovulation/ cycle: 2% vs 0%; # Grade I embryos: 61% vs 29% (p=0.008). Gn requirement: ↓ with MD
Karimzadeh et al. (2010), Iran, n=243	Single centre, 2-arm RCT	Age 18-35 yrs; BMI 18-30*; FSH < 10 IU/l; Regular cycle; 1 st IVF; Unexplained infertility	MD: CC 100 mg on D3-D7, FSH 75 IU/d from D5; ? GnRH-ant + hMG 75 IU/d (n=100) CD: Long GnRH-a DR; FSH 150-225 IU/d (n=100). Dose adj: No. Trigger: hCG. LPS: P intramuscular. ET D2/3	<3 follicles	OPR/s ET: 32% vs 26% (NS); OHSS: 0% vs 6% (p=0.02); CPR: 37% vs 31% (NS); CCR: 4% vs 0% (NS); # Oocytes/ embryo: ↓ with MD. # Top-quality embryo: 80% vs 75.5% (NS); Gn requirement: ↓ with MD
Lin et al. (2006), Taiwan, n= 70	Single centre, 2-arm RCT; Power: Difference in Gn dose; Consent, ethical approval: ?	Age 20-38 yrs; BMI 18.5-24.9*; FSH < 10 IU/l; Male factor; 1 st ICSI; No PCO or endometriosis	MD: CC 100 mg/d on D3-7, hMG 150-300 IU on D4, 6, 8 then daily; GnRH-ant (n= 60) CD: Long GnRH DR, hMG 150-300 IU/d (n=60). Dose adj: yes. Trigger: hCG 10000. LPS: Micro P. ET on D2/3	<2 follicles	LBR: 35.7% vs 35.0% (NS); CPR: 47% vs 40% (NS); OHSS: 1.7% vs 5% (NS); CCR: None in either group; # Oocytes: mean 5.4 vs 9 (p < 0.05); Gn requirement: ↓ with MD
Lou et al. (2010), China, n=60	Single centre, 2-arm RCT; Power: no; Consent, ethical approval: yes	Age <35 yrs; BMI 18- 28*; FSH < 10 IU/l; Regular cycle; Tubal factor; 1 st IVF attempt	MD: hMG 150 IU/d from D3; No GnRH-ant (n=30). CD: Long GnRH-a DR; rFSH 150-300 IU/d. (n=30). Dose adj: Yes. Trigger: hCG. LPS: P. ET: on D3	<2 follicles	OPR: 26.7% vs 23.3% (NS); OHSS: 0% vs 6.7% (NS); CCR: None; #Oocytes: mean 7.8 vs 12.2 (P<0.05). Gn requirement: ↓ with LD. Cost: mean 0.56 vs 1.6776 yuan (p< 0.001)
Mukherjee et al. (2012), India, n= 94	Single centre, 2-arm RCT; Power: no; Consent, ethical approval: yes	Age 25-35 yrs; Normal go-radotropin; Unexplained infertility; No PCOS/ endometriosis	MD: Let 5 mg/d on D3-7; rFSH 75 IU/d from D5 CD: rFSH 150-225 IU/d; GnRH-ant on both. Dose adj: No. Trigger hCG 10000. LPS: Micro P. ET on D2	<2 follicles	CPR: 36% vs 33% (NS); OHSS: 0% vs 14% (p=0.01). #Oocytes: mean 4.6 vs 4.9 (NS) Grade I embryo: mean 3.5 vs 3.7 (NS). Gn requirement: ↓ with MD. Cost: 34% less with MD-IVF
Oudshoorn et al. (2017), The Netherlands, n= 511	Multi-centre, 2-arm RCT; Power: adequate for cum LBR; Consent & ethical approval: yes	Age <44 years, high responder (AFC > 15); regular period (PCOS excluded); 1 st IVF.	MD: FSH 100 IU/d; GnRH-a or GnRH-ant (255) Dose adj: No. Trigger: hCG. LPS: Micro P. ET: on Day 3 or 5	>20 follicles	LBR: 25.7% vs 25.2% (NS); Cum LBRs: 36% vs 39.1% (NS); OHSS: 4.7% vs 4.7% (p<0.001); CCR: 24.1% vs 12.4% (p=0.001). #Oocytes: mean 9 vs 11.9 (p< 0.001). Cost: €622 vs €4714 per woman (NS)
Out et al. (2004), United Kingdom, n= 257	Multi-centre, 2-arm RCT Power: for # Oocytes, Consent & ethical approval: yes	Age 18-39 yrs; BMI 18-29*, Regular cycle.	MD: rFSH 150 IU/d from D2 (n= 131) CD: rFSH 200 IU/d from D2; on both arms, GnRH-ant from D6 (n= 126). Dose adj: No. Trigger: hCG 10000 IU. LPS: Progesterone. ET: max 3 embryos	<3 follicles	LBR: 31% vs 25% (NS); CPR: 34% vs 27% (NS); #Oocytes: mean 9.35 vs 10.4 (NS); # Good quality embryo: 4.6 vs 4.5 (NS)
Tan et al. (2005), Canada, n= 192	Multi-centre, 2-arm RCT; Power: for # Oocytes, Consent & ethical approval: yes	Age 18-39 yrs; BMI 18-29*, Regular cycle.	MD: rFSH 100 IU/d (n= 95) CD: rFSH 200 IU/d. On both arms, Long GnRH-a DR (n= 97). Dose adj: No. Trigger: hCG 5000-10000 IU. LPS: Progesterone. ET: max 3 embryos	<3 follicles	OPR: 27% vs 24% (NS); CCR: 5 vs 8 (NS); #Oocytes: mean 10.9 vs 12.2 (NS)

Continued

Table I Continued

Author Date Place, Population	Trial type and Method	Participants	Interventions: Mild (MD) versus Conventional-dose (CD) protocols	Cancellatn. criteria	Outcomes (per woman randomised/ cycle) MD-IVF versus CD-IVF
Tunmon et al. (1992), Canada, n= 408	Single centre, 2-arm RCT. Power: adequate for PR. Consent & ethical approval: yes	Any couple who need IVF treatment, except severe male factor.	MD: CC: 100 mg/d D5-D9 + hMG 75 IU/d from D6. No GnRH-ant. (n=229). CD: Long GnRH-a downregulation+ hMG dose adjusted with body weight. (n=179)	<2 follicles, high LH >25 IU, EZ drop or >12000 primol/l	PR: 10.7% vs 3.2% (NS); CCR: 30.8% vs 10.1% (P<0.001); # Oocytes: mean 3.6 vs 5.9 (p<0.001)
			Poor responders		
Ashrafi et al. (2005), Iran, n= 154	Single centre, 3-arm RCT. Power: not done	Previous poor response: <3 follicles >16 mm; E ₂ trigger <500 pg/l	MD: CC 100 mg/d on D3-7, hMG 150/d; no GnRH-ant (n=34) CD: Arm 1: hMG 150/d; no GnRH-ant (n=45) Arm 2: Long GnRH-a DR; hMG 225/d (n=52). Dose adj: yes. Trigger: hCG 10000.	?	CCR: 45.4% vs 38.8% & 50.1% (NS); # oocyte: mean 1.53 vs 1.52/ 2.28 (NS) Gn requirement: ↓ with MD.
Bastu et al. (2016), Turkey, n= 95	Single centre, 3-arm RCT. Power: adequate for # oocytes; Consent, ethical approval: yes	Age 18-42 yrs; BMI: 19.3-28.9*; POR according to Bologna criteria	MD: Let 5 mg/d on D 2/3-6/7, hMG 75+rFSH 75 IU from D2/3 (n= 33) CD: hMG 150+rFSH 150 IU from D2/3 (n= 31) OR hMG 225+rFSH 225 IU from D2/3 (n= 31); in all GnRH-ant. Dose adj: ? Trigger: rhCG 250 mg. LPS; P gel 8%. ET on D3, SET if <35 yrs in 1st cycle.	No follicle >11 mm on D8	OPR: 15% vs 13% & 16% (NS); CCR: 24% vs 26% & 29% (NS); # of oocytes/ embryos: mean 3.4 vs 3.3 & 3.6 (NS); Gn requirement: ↓ with MD.
Goswami et al. (2004), India, n= 38	Single centre, 2-arm RCT. Power: not done; Consent, ethical approval: yes	Age 36-41 yrs. Previous POR 1-3 cycles with long DR; Exclusion: FSH ≥ 12 iu/l, Endometriosis, pelvic surgery	MD: Let 2.5 mg/d on D3-7, FSH 75 IU/d D3 & 8; no GnRH-ant (n= 13) CD: Long GnRH-a DR; rFSH 300 iu/d (n= 25) Trigger: hCG 10000 IU. Dose adj: yes. LPS: Micro P; ET on D2.	No follicular development	CPR: 23.0% vs 24.0% (NS); CCR: 7% vs 4% (NS); # of oocytes/ embryos: mean 1.6 vs 2.1 (NS); Gn requirement: ↓ with MD.
Huang et al. (2015), China, n= 105	Single centre, 2-arm RCT. Power: not done; Consent? ethical approval: yes	Bologna criteria; Exclusion: >1 failed IVF, adenomyosis, drug allergy	MD: Let (? dose) on D3-7, rFSH 150 IU/d on D4,6&8; GnRH-ant (n= 50) CD: long GnRH-a DR; rFSH 300 IU/d (n= 55). Dose adj: ? Trigger: hCG 10000 IU. LPS: Micro P. ET on D2.	?	CPR: 26% vs 25.5% (NS); # oocytes: mean 2.7 vs 4.0 (p= 0.01); # good-grade embryo: mean 1.3 (NS)
Kim et al. (2009), South Korea, n=90	Single centre, 2-arm RCT; Power: not done; Consent, ethical approval: yes	Previous cycle with <4 follicles over 1.5 mm and <4 oocytes, with high dose	MD: rFSH 150 IU/d; GnRH-ant (n= 45) CD: rFSH 225 from D3; GnRH-ant (n= 45). Dose adj: yes. Trigger: rhCG 250 mgm. LPS: P 8% gel. DET if aged <38 yrs, 3 if >38 yrs	?	LBR/ET: 13.5% vs 16.7% (NS); CCR: 17.8% vs 6.7% (NS); # oocytes: mean 1.5 vs 3.1 (p<0.001) # top-grade embryo: mean 0.8 vs 1.2 (p= 0.007).
Klinkert et al. (2005), The Netherlands, n= 52	Single centre, 2-arm RCT; Power: for # oocytes. Consent, ethical approval: yes	AFC<5; 1 st IVF cycle; Regular period	MD: Long GnRH-a DR; rFSH 150 IU on D4, D6 CD: GnRH-a long DR; rFSH 300 iu/d (n= 26). Dose adj: yes. Trigger: hCG 10000 IU; LPS: micro P/ hCG.	No follicle developed	OPR: 8% vs 4% (NS); CCR: 12% vs 4% (NS); # of oocytes: median 3 vs 3; # embryos: median 2 vs 2 (NS)
Liu et al. (2020), China, n= 191	Single centre, 2-arm RCT; Power: for CLBR; Consent, ethical approval: yes	Bologna criteria; Exclusion: Severe endometriosis, repeated failed cycles.	MD: Let 5 mg/d on D3-7, rFSH 150 IU on D4, D6 & D8 onward; GnRH-ant (n= 97) CD: Long GnRH-a DR; rFSH 300 IU/d (n= 94) Dose adj: No. Trigger: hCG rhHCG 250. LPS: Micro P. ET on D3/5	No follicle developed	Cum LBR: 30.0% vs 15.0% (NS) LBR: 19.6% vs 18.1% (NS) OPR: 23.2% vs 22.9% (NS) CCR: # oocytes: mean 2.8 vs 4.0 (p <0.05); # top-grade embryo: mean 1.2 vs 1.4 (NS); Gn requirement: ↓ with MD

Continued

Table 1 Continued

Author Date Place, Population	Trial type and Method	Participants	Interventions: Mild (MD) versus Conventional-dose (CD) protocols	Cancellatn. criteria	Outcomes (per woman randomised/ cycle) MD-IVF versus CD-IVF
Martinez et al. (2003), Spain, n= 90	Single centre, 4-arm RCT; Power: ? Consent & ethical approval: yes	Previous POR	MD: CC 100mg/d on D4-8; hMG 150 IU from D5. CD: Arm 1: hMG 150+rFSH 150 IU from D2/3. Arm 2: hMG 225+rFSH 225 IU from D2/3. GnRH-ant in all arms. Trigger: rHCG 250 µg. Dose adj: ?: LPS; P gel 8%. ET on D3. SET if <35 yrs 1st cy, otherwise DET.	<3 follicle after 10 days	OPR: 13% 9.1% (NS); CPR: 21.7% vs 18.2% (NS). CCR: 23.9% vs 4.5% (NS); Gn requirement: ↓ with MD.
Mohsen et al. (2013), Egypt, n= 60	Single centre, 2-arm RCT; Power: not done; Consent, ethical approval: yes	Age unselected; BMI<30*; ≥1 previous cycle with POR; No endometriosis, pelvic/ ovarian surgery, no systemic disease, no severe male factor	MD: Let 2.5 mg on BD D2-6; hp-HMG 150 IU/d from D7; GnRH-ant (n= 30). CD: GnRH-a from D2 till ovulation trigger; hp-HMG 300 IU/d from D3 (n= 30). Dose adj: yes. Trigger: hCG 10000 IU. Pre-treatment: E ₂ 2mg BD from mid-luteal. LPS; P pessary 400 mg/d.	<2 follicles low/ plateau E ₂ despite increased Gn dose	CPR: 13.3% vs 16.6% (NS); CPR: 20% vs 16.6% (NS); # of oocytes: mean 5.1 vs 5.1 (NS)
Pilehvari et al. (2016), Iran, n= 77	Single centre, 2-arm RCT; Power: not done; Consent, Ethical approval: yes	Bologna criteria; No systemic disease, No Tx within last 3 months	MD: CC 100mg/d on D2-6, hMG 150 IU/d from D5 (n= 42). CD: hMG 300 IU/d from D2 (n= 42); GnRH-ant in both. Dose adj: yes. Trigger: hCG 10000 IU when 2 follicles > 16 mm. LPS; P pessary 400/ d.	low/ no follicular growth	CPR/ ET: 4% vs 5.6% (NS); CPR: 28.6% vs 31.4% (NS); # oocytes: mean 2.2 vs 2.8 (NS); Good quality of embryos: 58.5% vs 71.8% (NS)
Ragni et al. (2012), Italy, n=291	Multi-centre, 2-arm RCT.	Age 18-42 yrs; FSH>12; Previous POR (≤3 eggs); Exclusion: ≥ 1 failed cycle, surgically retrieved sperm	MD: CC 150 mg/d on D3-7, no gonadotropin (n= 145). CD: Long GnRH-a DR; rFSH 450 IU/d from D3; Dose adj: yes. Trigger: rHCG 250 mg; LPS; Micro P	low/ no follicular growth	LBR: 3% vs 5% (NS); CPR: 5% vs 6% (NS). CCR: 1.4% vs 1.4% (NS); Cost: mean C: €2803 vs €5423 per cycle (p<0.005)
Revelli et al. (2014), Italy, n= 695	Single centre, 2-arm RCT; Power: adequate for # oocytes; Consent & Ethical approval: yes	Age <43 yrs; FSH 10-20 IU; AMH 0.14-1.0 ng/ml; AFC 4-10	MD: CC 100 mg on D2-6, hMG 150 IU/d from D5; GnRH-ant (n= 309). CD: Long GnRH-a DR; hMG 300-450 IU/d (n= 331). Dose adj: no. Trigger: hCG 10000 IU. LPS; P gel 8%.	<1 fol. 10 mm & E2 <50 pg/ml d7/8	OPR/ ET: 17.8% vs 16.8% (NS); CPR: 13.2% vs 15.3% (NS). CCR: 13% vs 22% (p<0.01); # oocytes/ embryo: mean 2.7 vs 4.8 (p= 0.01) Top-grade embryo: 57.6% vs 54.8% (NS)
van Tillborg et al. (2017), The Netherlands, n= 511	Multi-centre, 2-arm RCT; Power: adequate for cumulative LBR; Consent & ethical approval: yes	Age <44 yrs; AFC<11; Regular cycle; 1 st IVF cycle; Normal pelvic scan.	MD: rFSH 150 IU/d either in long GnRH-a DR or GnRH-ant protocol (n= 260). CD: rFSH 225 if AFC 8-10 & 450 IU/d if AFC <8, either in long GnRH-a DR or GnRH-ant (n= 250). Dose adj: no. Trigger: hCG 10000 IU. LPS; P pessary; ET on D3/5.	<2 follicles of 12 mm, <3 follicles <17 mm.	LBR: 15.8% vs 14.8% (NS); Cum LBR: 20% vs 17.6% (NS); CCR:
Youssef et al. (2017), Egypt, Iran, Syria n= 394	Multi-centre, 2-arm RCT; Power: adequate for OPR; Consent & ethical approval: yes	Age <43 yrs; BMI<23*; FSH ≥15; ANH <1.5 ng/ml; AFC ≤8; Exclusion: Age >43, uterine anomaly Tx; COC for MS-IVF. LPS; P daily IM	MD: rFSH 150 IU/d from D5 of last COC, GnRH-ant (n= 195). CD: Long GnRH-a DR, hMG 450 IU/d (n= 199). Dose adj: no. Trigger: hCG 10000 IU. Pre-Tx: COC for MS-IVF. LPS; P daily IM	<2 fol. of <15 mm after d7	OPR: 12.8% vs 13.6% (NS); CPR: 15.3% vs 15.3%. CCR: 26% vs 18% (NS); # Top-quality embryo: mean 0.3 vs 0.8 (NS)
Yu et al. (2018), China, n= 106	Single centre, 3-arm RCT; Power: adequate for CPR; Consent & Ethical approval: yes	Age <43 yrs; BMI <23*; FSH ≥15; ANH <1.5 ng/ml; AFC ≤8; Exclusion: Endometriosis, endocrine disorder, pelvic surgery	MD: Arm 1 Let 5mg/d on D3-7, hMG 75 IU/d from D4; Arm 2 hMG 75 IU/d from D3; GnRH-ant in both CD: GnRH-a DR on D3, hMG 225-300 IU/d 28 days after GnRH-a. Dose adj: yes. Trigger: hCG 10000 IU. LPS; P 60 mg IM+ Dydrogesterone	?	LBR/ cycle: 15.4% vs 20.4% & 13.3% (NS); CPR: 23.1% vs 29.6% & 30%. CCR: 32.7% vs 11.1% (p<0.05)

RCT, randomised controlled trials; #, number; yrs, years; BMI body mass index, *kg/m²; Prev., previous; d, day; D, (cycle) day; GnRH-ant, GnRH antagonist; GnRH-a, GnRH agonist; DR, downregulation; adj, adjustment; LPS, luteal phase support; ?, not stated; ET, embryo transfer; OPR, ongoing pregnancy rate; NS, not significant; OHSS, ovarian hyperstimulation syndrome; CCR, cycle cancellation rate; P, progesterone; PR, pregnancy rate; Gn, gonadotrophin; AMH, anti-Mullerian hormone; AFC, antral follicle count; LBR, live birth rate; Cum LBR, cumulative live birth rate; CPR, clinical pregnancy rate; CC, clomiphene citrate; E2, oestradiol; Let, letrozole; hCG, recombinant hCG; rFSH, recombinant FSH; SET, single-embryo transfer; Micro, micronized; POR, poor ovarian reserve; COC, combined oral contraceptive.

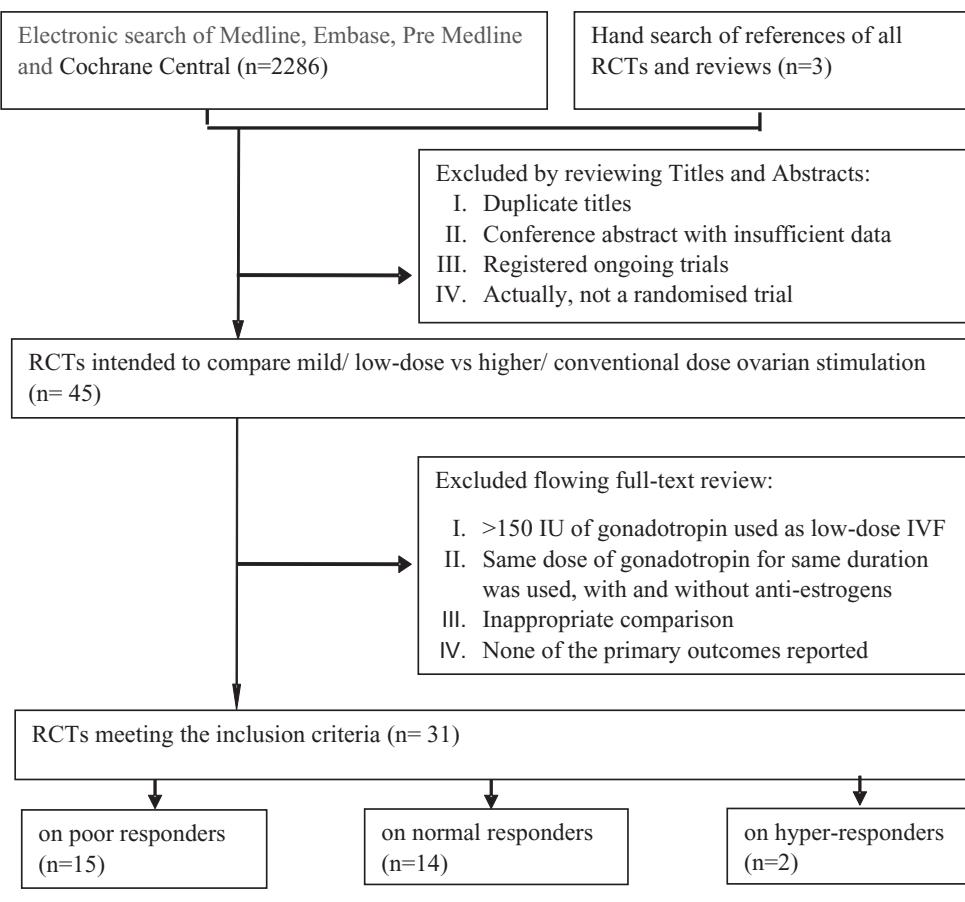


Figure 1. Flow-chart of the study selection process. RCT, randomised controlled trial.

versus high-dose gonadotrophin only protocols; CC+ mild-dose gonadotrophin versus CD-IVF protocols and Letrozole+ mild-dose gonadotrophin versus CD-IVF protocols.

Multi-arm studies

The methodology described by the *Cochrane Hand book for Systematic Review of Intervention* was followed in the meta-analysis of multi-arm studies (Higgins and Green, 2011). If MD-IVF was compared with two different CD-IVF protocols, both the events and populations (denominators) in MD-IVF were equally divided and incorporated under respective sub-groups. If MD-IVF or CD-IVF consisted of two different doses or types of gonadotrophin, they were combined into one taking the average of both events and populations. For continuous data in the above situations, the mean and SD of the common groups were kept the same, only the population was equally split into two subgroups.

Sensitivity analysis

We performed sensitivity analysis by repeating meta-analyses of all outcomes in the following ways: excluding and including small studies with RoB; excluding studies with permitted dose adjustment; gonadotrophin with and without oral compounds; applying a fixed as well as a

random effect model; and applying RR and peto odd-ratio (OR) as the method of determining effect size.

Results

The study selection process is demonstrated in the flow chart (Fig. 1). Three publications were found by hand searching (Out *et al.*, 2004; Tan *et al.*, 2005; Mukherjee *et al.*, 2012), the rest by electronic search. Forty-five shortlisted publications underwent full-test review for further assessment of eligibility criteria. Table II narrated the list of excluded studies with reasons. A large RCT applied single-embryo transfer policy in a 'minimal' group, with double-embryo transfer in a 'conventional' IVF group, but had both fresh and frozen-thawed transfer in both groups (Heijnen *et al.*, 2005)—this study was excluded for pregnancy outcomes per randomisation as these outcomes could have been affected by the differential embryo-transfer policy. However, cumulative pregnancy outcome, CCR and laboratory parameters would not have been affected hence this study was included in the meta-analyses for these outcomes. Finally, 31 RCTs were included: 15 RCTs in the poor, 14 RCTs in the normal and 2 RCTs in the hyper-responder group.

Table II The list of excluded studies.*

Studies	Reasons for exclusion
Poor responders	
(Siristatidis et al., 2016)	Started as RCT but ended with case-control trial
(Eftekhar et al., 2014)	CC+ hMG 225-300 IU/ day versus Letrozole+ hMG 225-300 IU/ day
(Ebrahimi et al., 2017)	Letrozole+ FSH 225 IU/ day versus Placebo+ FSH 225 IU/ day
(Fujimoto et al., 2014)	CC+ hMG ? dose versus hMG ? dose
(Jindal and Singh, 2013)	CC/ Letrozole+ gonadotropin, ? dose versus Gonadotropin ? dose
(Lee et al., 2011)	Letrozole+ FSH 225 IU/ day versus FSH 225 IU/ day
(Nabati et al., 2015)	Letrozole+ FSH 300 IU/ day versus FSH 300 IU/ day
(Ozcan Cenksoy et al., 2014)	CC+ FSH 450 IU/ day versus FSH 450 IU/ day
(Schimberni et al., 2016)	CC+ FSH 450 IU/ day versus FSH 450 IU/ day
(Selman and Rinaldi, 2016)	CC+ FSH 225 IU/ day versus CC+ FSH 225 IU/ day + corifollitropin alfa 150 IU
Normal/ hyper responders	
(Ghoshdastidar et al., 2010)	The denominators missing
(Grochowski et al., 1999)	Actually a non-randomised allocation
(Hoomans et al., 2002)	None of our primary outcomes was reported
(Jayaprakasan et al., 2010)	FSH 225 versus 300 IU/ day
(Kingsland et al., 1992)	CC+ hMG 150-300 IU/ day (depending on age) versus hMG 150-300 IU/ day
(Long et al., 1995)	CC+ hMG 150 IU/ day versus hMG 150 IU/ day (same dose and duration)
(Popovic-Todorovic, 2003)	FSH 150 versus 100-250 IU/ day
(Pruksananonda et al., 2004)	Full text could not be accessed
(Weigert et al., 2002)	Same dose was used, one with and the other group without oral compound
(Wiklund, 2001)	None of our primary outcomes was reported
Zhang et al. (2016)	Freeze-all embryo followed by single-embryo transfer for Mini-IVF, while fresh and frozen double-embryo transfer for conventional protocol

*The table explains on what basis some of the studies that were included in other related systematic reviews were considered not eligible for this review.

Characteristics of included studies

Table I summarised the studies included in this review and meta-analysis. All included papers were written in English except one (Martinez et al., 2003), which was written in Spanish and the translation was by Google Translator. Two were conference abstracts with sufficient data for meta-analysis (Huang et al., 2015; Elnashar et al., 2016).

Trial design

Seven included studies were multi-centre trials (Out et al., 2004; Tan et al., 2005; Heijnen et al., 2007; Ragni et al., 2012; Oudshoorn et al., 2017; van Tilborg et al., 2017; Youssef et al., 2018), the rest were from a single centre. Four trials conducted three-arm comparison (Harrison et al., 1994; Ashrafi et al., 2005; Bastu et al., 2016; Yu et al., 2018), one was a four-arm trial (Martinez et al., 2003), the rest were two-arm studies. Sample size calculation was done in six trials among poor responders: two for oocyte number (Revelli et al., 2014; Bastu et al., 2016), one for CPR (Yu et al., 2018), one for OPRs (Youssef et al., 2017), one for LBR (Ragni et al., 2012) and two for cumulative live birth (van Tilborg et al., 2017; Liu et al., 2020). Of the studies on unselected patients, four were powered for oocyte numbers (Out et al., 2004; Tan et al., 2005; Baart et al., 2007; Blockeel et al., 2011), two for pregnancy rate (PR) (Tummon et al., 1992; Dhont et al., 1995) and one for LBR (Heijnen et al., 2007). Both the RCTs on

hyper-responders were large: one had adequate power for number of oocytes ($n=412$) (Casano et al., 2012), the other for cumulative LBRs (Oudshoorn et al., 2017).

Participants

Recruitment in five RCTs was as per the Bologna consensus on poor ovarian response (POR) (Ragni et al., 2012; Huang et al., 2015; Bastu et al., 2016; Pilehvari et al., 2016; Liu et al., 2020); others were based on different combinations of age, FSH, AMH, AFC and previous poor response (Table I). Selection of patients in the non-PCOS hyper-responder group was on the sole criterion of AFC in both the RCTs (Casano et al., 2012; Oudshoorn et al., 2017). Unselected patients/ normal responders were recruited in absence of high or low ovarian reserve, mostly on the first cycle of IVF (detailed in Table I).

Interventions

Interventions in each individual trial were detailed in Table I. Comparison between low- and high-dose gonadotrophins only stimulation (without oral medication) was reported in six RCTs on poor responders (Ashrafi et al., 2005; Klinkert et al., 2005; Kim et al., 2009; van Tilborg et al., 2017; Youssef et al., 2017; Yu et al., 2018); seven RCTs on normal responders (Hohmann et al., 2003; Out et al., 2004; Tan et al., 2005; Baart et al., 2007; Heijnen et al., 2007; Lou and Huang, 2010; Blockeel et al., 2011) and both the RCTs on hyper-

responders. Ten trials in the patient with POR used oral compounds in the MD-IVF arm either alone (CC) (Ragni *et al.*, 2012) or in combination with low-dose gonadotrophins: CC was used in five and Letrozole in three trials (Table I). Among the normal responder group, six RCTs used CC+ gonadotrophin and two with Letrozole combination. Consistently, CC was used at 100 mg daily dose for 5 days, commencing on cycle Day 2–4, except the RCT by Ragini *et al.* where 150 mg daily dose was used. The dose for Letrozole was 5 mg daily, starting from Day 2–5, except in two trials: one used 2.5 mg daily (Goswami *et al.*, 2004) and the other 10 mg daily dose (Elnashar *et al.*, 2016). In all trials, the starting dose of gonadotrophin for MD-IVF was 150 IU daily, except in two RCTs on poor responder where a 75 IU dose was used (Goswami *et al.*, 2004; Yu *et al.*, 2018); in one trial for the normal (Tan *et al.*, 2005) and one for the high responders (Oudshoorn *et al.*, 2017)- 100 IU daily dose was used in both studies. However, the timing of commencement of gonadotrophin varied (Table I). Dose adjustment was allowed in 12 RCTs, fixed dose in 13 and not mentioned in remaining six trials (Table I). Pre-treatment was given in three RCTs (Dhont *et al.*, 1995; Mohsen and El Din, 2013; Youssef *et al.*, 2017). Cycle cancellation criteria varied between the studies (Table I).

Outcome measured

The definition of cumulative LBR differed among the studies: the RCT by Casano *et al.* (2012) and Liu *et al.* (2020) aggregated the outcome of fresh and all subsequent frozen-thawed transfer; while other three trials included all fresh and frozen cycles within a specified time-period of 12 months (Heijnen *et al.*, 2007) or 18 months (Oudshoorn *et al.*, 2017; van Tilborg *et al.*, 2017). Three studies reported pregnancy rates as positive beta-hCG (Dhont *et al.*, 1995; Hohmann *et al.*, 2003; Blockeel *et al.*, 2011) and two trials did not specify whether it was clinical pregnancy (Tummon *et al.*, 1992; Elnashar *et al.*, 2016) and therefore excluded from the meta-analysis on CPR. The criterion for cycle cancellation was not uniform (Table I). The clinical criteria for reporting of OHSS varied between the trials and were not clear in some studies. Three RCTs estimated total and mean per-patient cost of all fresh and frozen cycles together (Heijnen *et al.*, 2007; Oudshoorn

et al., 2017; van Tilborg *et al.*, 2017); one trial reported total and per-patient cost of only fresh cycle (Ragni *et al.*, 2012) and the remaining two reported the medication cost of stimulated cycles (Lou and Huang, 2010; Mukherjee *et al.*, 2012).

Risk of bias of the included studies

A summary of RoB was graphically presented in Fig. 2.

Selection bias

All RCTs were found to be 'low-risk' for random sequence generation, except five trials where the risk was unclear (Dhont *et al.*, 1995; Ashrafi *et al.*, 2005; Mukherjee *et al.*, 2012; Elnashar *et al.*, 2016; Pilehvari *et al.*, 2016). Allocation concealment was deemed to have low risk in all but seven RCTs where the risk was unclear (Tummon *et al.*, 1992; Dhont *et al.*, 1995; Martinez *et al.*, 2003; Lou and Huang, 2010; Pilehvari *et al.*, 2016; Yu *et al.*, 2018; Liu *et al.*, 2020). Performance and detection bias: All RCTs were of 'low-risk' for performance bias, as the blinding of both patients and assessors was neither possible nor required for these objective outcome measures. Attrition bias: The outcome data were not complete in one trial (high risk) (Huang *et al.*, 2015), and not clear in the three other studies (Ashrafi *et al.*, 2005; Mohsen and El Din, 2013; Elnashar *et al.*, 2016), the rest were of 'low risk'. Reporting bias: All RCTs had 'low risk' for reporting bias. Other bias: Baseline characteristics of both sides were not clear in eight RCTs (Ashrafi *et al.*, 2005; Mukherjee *et al.*, 2012; Ragni *et al.*, 2012; Huang *et al.*, 2015; Bastu *et al.*, 2016; Elnashar *et al.*, 2016).

Primary outcomes

Livebirth rates

Poor responders. Five RCTs compared LBRs ($n = 1248$), two of them compared mild and conventional-dose gonadotrophin only stimulation (Kim *et al.*, 2009; van Tilborg *et al.*, 2017), one CC and high-dose antagonist protocol (Ragni *et al.*, 2012) and two with letrozole combination, of which the study by Yu *et al.*, also had a 3rd arm with low-dose gonadotrophin only protocol (Yu *et al.*, 2018; Liu *et al.*, 2020).

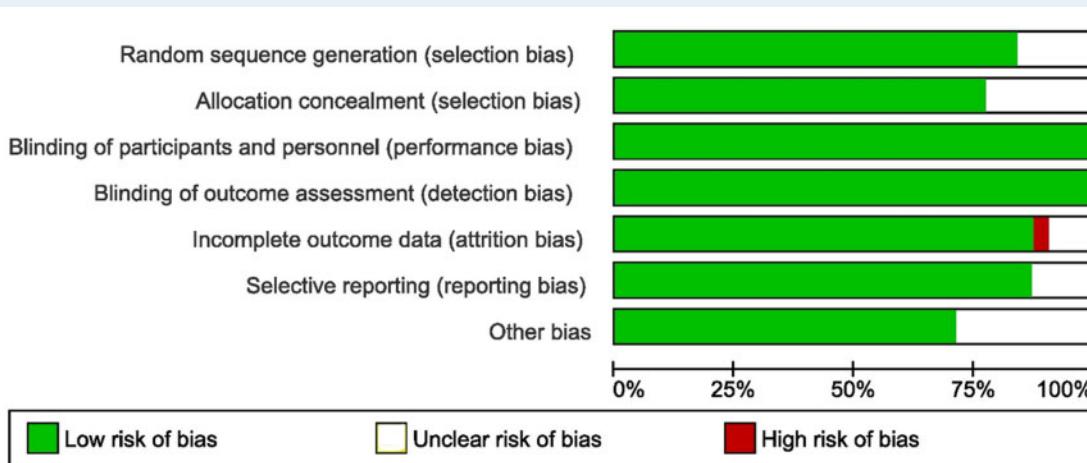


Figure 2. Risk of bias graph from the included studies.

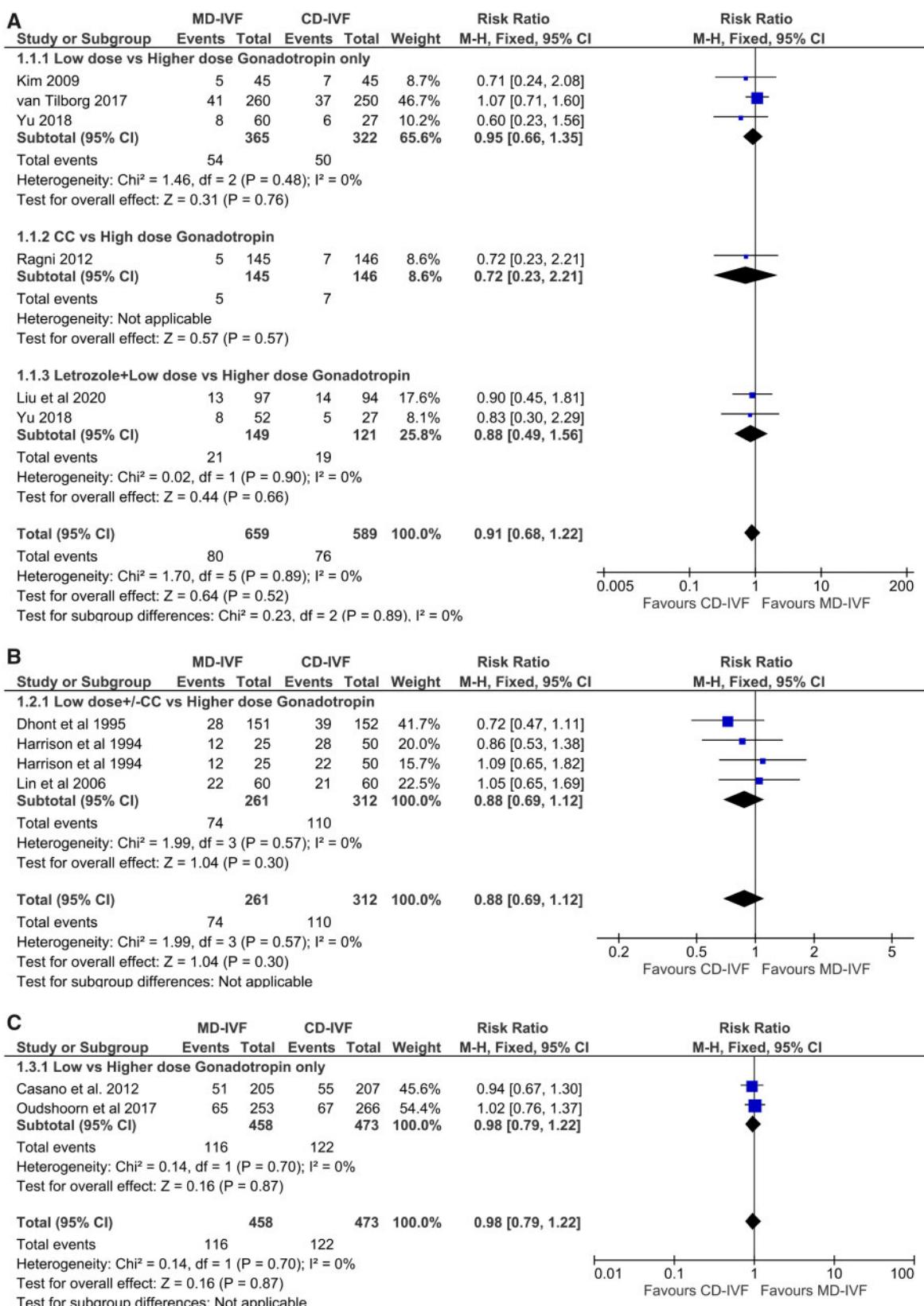


Figure 3. Forest plot of mild versus conventional-dose IVF: live birth rate per randomisation. A for poor responders, **B** for normal responders, **C** for hyper-responder. MD-IVF, mild-dose IVF; CD-IVF, conventional-dose IVF.

Table III Summary of evidence.

	Poor responders	Normal responders	Hyper-responders
Livebirth rates	No difference $\oplus\oplus\oplus\ominus$ RR 0.91 [0.68, 1.22] RCT= 5, n= 1248 ✓ I^2 0% ✓ Narrow CI ✓ 2 large RCTs low RoB ✓ No RCT contradicted ✗ 1 study with unclear RoB ✗ Clinical heterogeneity	No difference $\oplus\oplus\oplus\ominus$ RR 0.88 [CI 0.69, 1.12] RCT= 3, n= 573, ✓ I^2 0% ✓ Narrow CI ✓ ↓ Clinical heterogeneity ✓ No RCT contradicted ✗ Studies with unclear RoB	No difference $\oplus\oplus\oplus\ominus$ RR 0.98 [CI 0.79, 1.22] RCT= 2, n= 931 ✓ I^2 0% ✓ Narrow CI ✓ Only 1 unclear RoB ✓ No RCT contradicted ✗ ↑ Clinical heterogeneity
OHSS rates	–	↓ with MD-IVF $\oplus\oplus\oplus\ominus$ RR 0.26 [CI 0.14, 0.49] RCT= 9, n= 1925 ✓ I^2 0% ✓ Narrow CI ✓ Large effect size ✗ Unclear RoB ✗ Clinical heterogeneity	↓ with MD-IVF $\oplus\oplus\oplus\ominus$ RR 0.47 [CI 0.31, 0.72] RCT= 2, n= 931 ✓ I^2 0% ✓ Narrow CI ✓ Large effect size ✓ Low RoB (1 unclear) ✗ Clinical heterogeneity
Cycle cancellation rates	No difference $\oplus\ominus\ominus\ominus$ RR 1.33 [CI 0.96, 1.85] RCT= 15, n= 3459 ✗ I^2 64% ✗ Wide CI ✗ Most RCTs with RoB ✗ Clinical heterogeneity	↑ with MD-IVF $\oplus\ominus\ominus\ominus$ RR 2.08 [CI 1.38, 3.14]* RCT= 12, n= 2654 ✗ I^2 48% ✗ Wide CI ✗ Small RCTs, unclear RoB ✗ Clinical heterogeneity	No difference $\oplus\oplus\oplus\ominus$ RR 1.31 [CI 0.98, 1.77] RCT= 2, n= 1348 ✓ I^2 0% ✓ 2 large RCTs low RoB ✗ Moderately wide CI Clinical heterogeneity
Ongoing pregnancy rates	No difference $\oplus\oplus\oplus\ominus$ RR 1.02 [CI 0.81, 1.25] RCT= 7, n= 2006 ✓ I^2 0% ✓ Narrow CI ✓ 3 large RCTs low RoB ✓ No RCT contradicted ✗ Clinical heterogeneity	No difference $\oplus\oplus\oplus\ominus$ RR 1.10 [CI 0.88, 1.38] RCT= 7, n= 1026 ✓ I^2 0% ✓ ↓ Clinical heterogeneity ✓ No RCT contradicted ✗ Small studies with unclear RoB	No difference $\oplus\oplus\ominus\ominus$ RR 0.86 [CI 0.61, 1.23] RCT= 1, n= 521 ✓ Large RCT ✓ Low RoB ✗ Based on just 1 RCT with two different protocols
Number of oocytes retrieved (mean)	↓ with MD-IVF $\oplus\oplus\ominus\ominus$ SMD -0.43 [CI -0.58, -0.28] RCT= 14, n= 2773, ✓ Large effect size, narrow CI ✗ I^2 67% ✗ RCTs with RoB ✗ Clinical heterogeneity	↓ with MD-IVF $\oplus\ominus\ominus\ominus$ SMD -1.34 [CI -1.94, -0.75] RCT= 13, n= 3499, ✗ I^2 98% ✗ Wide CI ✗ RCTs with unclear RoB, ✗ Clinical heterogeneity	No difference $\oplus\oplus\ominus\ominus$ SMD -0.31 [CI -0.74, 0.13] RCT= 2, n= 931 ✓ Low RoB (1 unclear) ✗ I^2 91% ✗ Wide CI ✗ Clinical heterogeneity
Number of embryos created (Mean)	↓ with MD-IVF $\oplus\oplus\ominus\ominus$ SMD -0.39 [CI -0.59, -0.20] RCT= 9, n= 1559, ✓ Narrow CI ✓ 2 large RCTs with low RoB ✗ I^2 59% ✗ 1 RCT with high RoB ✗ Clinical heterogeneity	No difference $\oplus\ominus\ominus\ominus$ SMD -0.30 [-0.58, 0.08] RCT= 7, n= 1884, ✗ I^2 79% ✗ Small studies with wide CI ✗ Multiple unclear RoB ✗ Clinical heterogeneity	Data not available
Number of high-grade embryos (Mean)	No difference $\oplus\oplus\ominus\ominus$ MD -0.12 [-0.30, 0.05] RCT= 4, n= 723 ✓ I^2 0%	No difference $\oplus\oplus\ominus\ominus$ MD -0.18 [-0.49, 0.13] RCT= 6, n= 551, ✓ I^2 0%	Data not available

Table III Continued

	Poor responders	Normal responders	Hyper-responders
Proportion of high-grade embryos	<ul style="list-style-type: none"> ✓ 2 large RCTs with low RoB ✗ 1 small RCT with high RoB ✗ Clinical heterogeneity <p>No difference $\oplus\oplus\ominus$ Meta-analysis not possible ✓ All 3 RCTs including 1 large one with low RoB reported no difference</p>	<ul style="list-style-type: none"> ✗ Only 3 small RCTs (wide CI) with unclear RoB ✗ Clinical heterogeneity <p>No difference $\oplus\oplus\ominus$ RR 1.07 [0.93, 1.23] RCT = 3, n = 656 χ^2 0%</p>	<p>No difference $\oplus\oplus\ominus$ 46.7% vs 42.1% [$p > 0.05$] RCT = 1, n = 412 ✓ Only 1 RCT but large with low RoB</p>
Gonadotropin dose (mean)	<p>↓ with MS-IVF $\oplus\oplus\ominus$ SMD -3.17 [-3.80 -2.54] RCT = 13, n = 2314 ✓ Large effect size ✓ No RCT contradicted ✓ 3 large RCTs low RoB χ^2 96% ✗ RCTs with unclear/ high RoB ✗ Clinical heterogeneity</p>	<p>↓ with MS-IVF $\oplus\oplus\ominus$ SMD of -5.86 [CI -7.06, -4.66] RCT = 11, n = 2583 ✓ Large effect size ✓ No RCT contradicted χ^2 99% ✗ RCTs with unclear/ high RoB ✗ Clinical heterogeneity</p>	<p>↓ with MS-IVF $\oplus\oplus\ominus$ SMD -394.00 [-481.20 -306.80] RCT = 1, n = 412 ✓ Only 1 RCT but large with low RoB</p>

$\oplus\oplus\ominus$, moderate quality of evidence; $\oplus\oplus\ominus$, low quality of evidence; $\oplus\ominus\ominus$, very low quality of evidence; RR, relative risk; MD, mean difference; SMD, standardised mean difference; RoB, risk of bias.

There was no evidence of a difference in LBRs: RR 0.91 (CI 0.68, 1.22) (Fig. 3A). There was no statistical heterogeneity (χ^2 0%) and four RCTs were of low RoB (Kim et al., 2009; Ragni et al., 2012; van Tilborg et al., 2017; Liu et al., 2020). The finding remained unchanged in sensitivity analysis, when the smaller RCTs with possible RoB were excluded or whether trials with dose adjustments were included or excluded. The inference was the same, whether gonadotrophin only protocol or CC/Letrozole protocols were used. Due to the presence of significant clinical heterogeneity, the quality of evidence (QoE) was moderate (Table III).

Normal responders. Three included studies reported LBRs (n = 573), all compared CC+ gonadotrophin (Harrison et al., 1994; Dhont et al., 1995; Lin et al., 2006) and long downregulation protocol. There was no difference in LBRs: RR 0.88 (CI 0.69, 1.12) (Fig. 3B). There was no statistical heterogeneity (χ^2 0%) and very little clinical heterogeneity between the trials. The finding did not alter in the sensitivity analysis. However, the large RCT had multiple areas of unclear RoB (Dhont et al., 1995); the other two were small trials (Harrison et al., 1994; Lin et al., 2006); hence the QoE was moderate (Table III). The evidence with gonadotrophin only protocols for this outcome among normal responders was lacking.

Hyper-responders. Two large RCTs looked for livebirth, both were powered for their primary outcomes; both studies applied gonadotrophin only stimulation protocols (Casano et al., 2012; Oudshoorn et al., 2017). The meta-analysis found LBRs did not differ between the groups: RR 0.98 (CI 0.79, 1.22). There was no statistical heterogeneity (χ^2 0%) and the QoE was moderate on GRADE analysis owing to clinical heterogeneity (Table III).

Incidence of OHSS

One RCT on poor responders reported OHSS rates (van Tilborg et al., 2017). The incidence was not significantly different between

doses (1.8% with 150 IU dose vs. 1.2% with 225–450 IU dose, $P = 0.45$).

Normal responders. Nine RCTs (n = 1925) estimated OHSS rates: four small trials (Tan et al., 2005; Baart et al., 2007; Lou and Huang, 2010; Blockeel et al., 2011) and a large one (Heijnen et al., 2007) with gonadotrophin-only regimens showed lower incidence of OHSS (RR 0.43 (CI 0.21, 0.90)) in the MD-IVF group. Meta-analysis of four RCTs with oral compounds (Dhont et al., 1995; Lin et al., 2006; Karimzadeh et al., 2010; Mukherjee et al., 2012), as well as all eight studies together, also found the risk of OHSS to be significantly lower with MD-IVF (RR 0.26 (CI 0.14, 0.49)) (Fig. 4A). Overall, the effect-size was large; there was no statistical heterogeneity (χ^2 0%) and the CI was narrow. Multiple studies had one or more areas of unclear RoB; in addition, clinical heterogeneity, including varied criteria for reporting OHSS, made this evidence of a moderate quality (Table III).

Hyper-responders. Both RCTs on hyper-responders were with a gonadotrophin-only regimen (Casano et al., 2012; Oudshoorn et al., 2017). Meta-analysis of the pooled data found a significantly lower incidence of any grade of OHSS with MD-IVF, with a RR of 0.47 (0.31, 0.72) (Fig. 4B). There was no statistical heterogeneity (χ^2 0%) and no RoB. The QoE was moderate due to methodological diversity (Table III).

Cycle cancellation rate

Poor responders. All 15 RCTs investigated CCRs (n = 3459). There was no difference in the risk of cycle cancellation between both arms, with an RR of 1.33 (CI 0.96, 1.85). The CCR was found to be higher when the trials with dose adjustments were excluded (RR 1.73 (CI 1.02, 2.93)) (Fig. 5A). The presence of significant statistical (χ^2 63%) as well as clinical heterogeneity, wide CI and unclear RoB in most trials led to a very low QoE for this outcome (Table III).

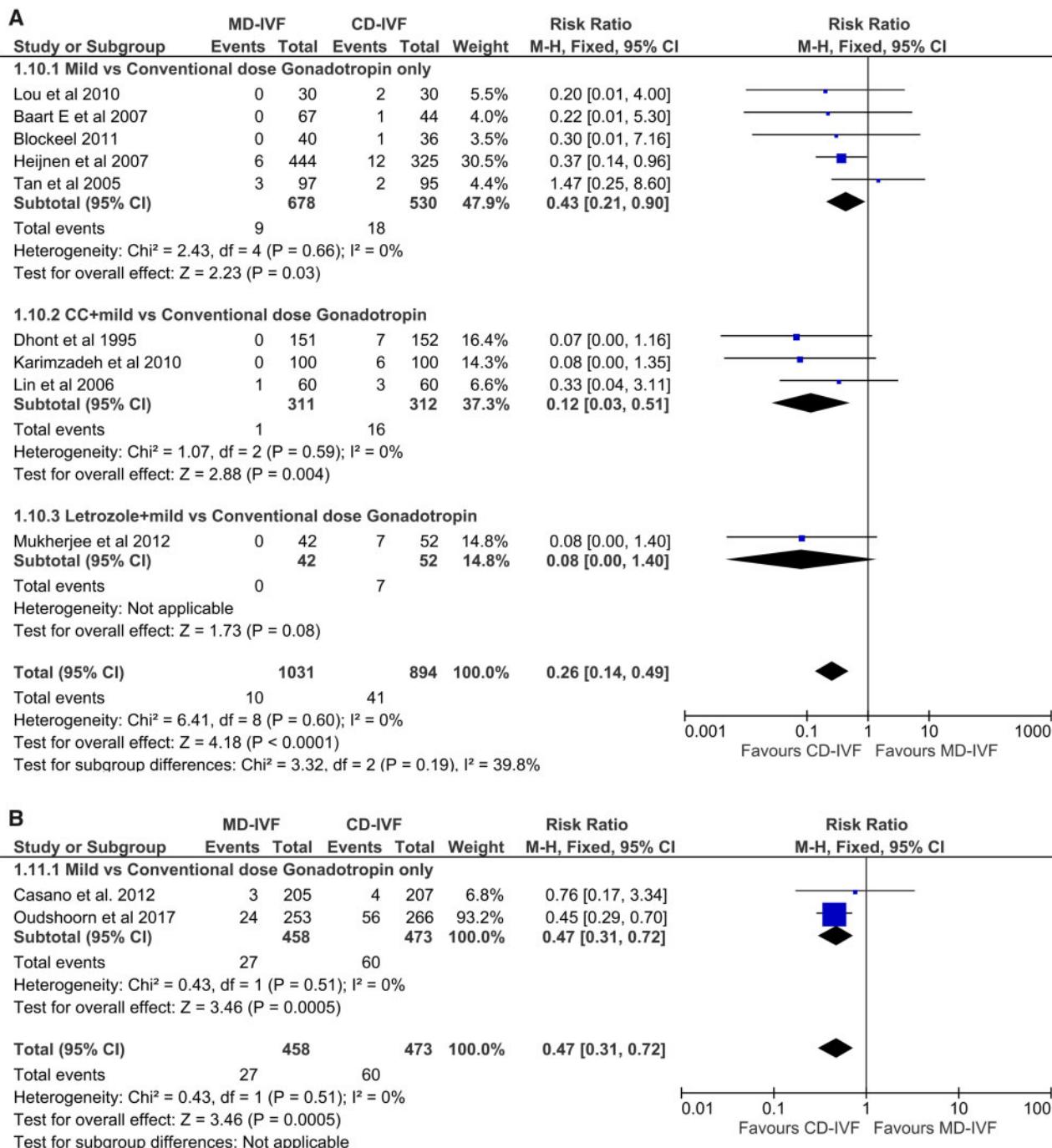


Figure 4. Forest plot of mild versus conventional-dose IVF: incidence of ovarian hyperstimulation syndrome per started cycle.

Normal responders. Seven RCTs (Hohmann et al., 2003; Out et al., 2004; Tan et al., 2005; Baart et al., 2007; Heijnen et al., 2007; Lou and Huang, 2010; Blockeel et al., 2011) with gonadotrophin only regimen ($n = 1430$) found no difference in CCRs in the meta-analysis (RR 1.60 (CI 0.96, 2.67)), while pooled data from five RCTs comprising a CC+ gonadotrophin regimen (Tummon et al., 1992; Harrison et al., 1994; Dhont et al., 1995; Lin et al., 2006; Karimzadeh et al., 2010) ($n = 1224$) showed a higher risk of cycle cancellation with MD-IVF (RR

2.87 (CI 1.46, 5.64)) (Fig. 5B). However, when three trials that did not use GnRH-agonist or antagonist for LH suppression (Tummon et al., 1992; Harrison et al., 1994; Dhont et al., 1995) were taken out of the meta-analysis, the CCR became comparable with CD-IVF. Overall, the CCR was higher with MD-IVF: RR 2.08 (CI 1.38, 3.14). The I^2 indicating statistical heterogeneity was 48%. The QoE was very low, due to clinical heterogeneity, a wide CI and multiple small studies with unclear RoB (Table III).

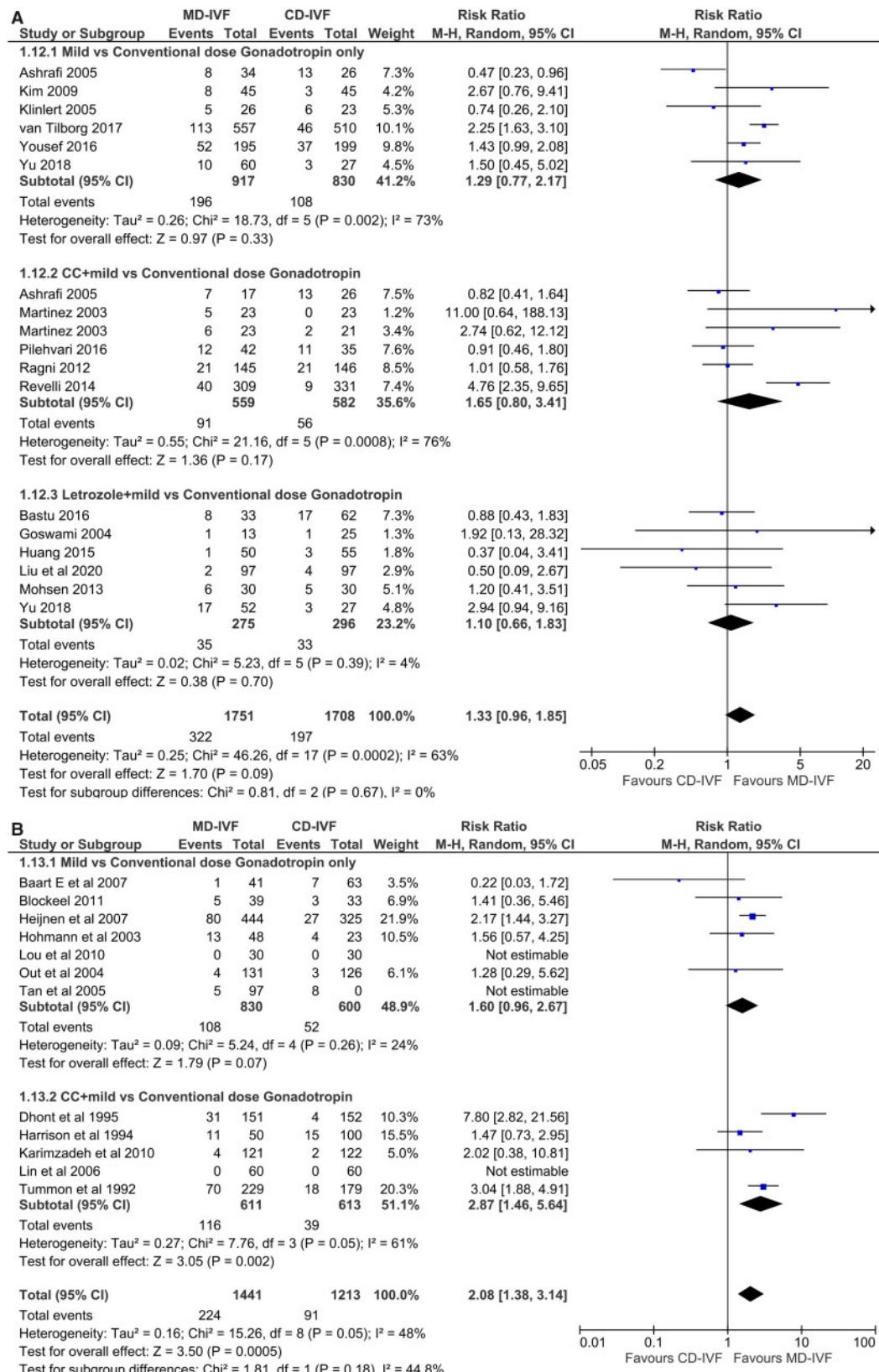
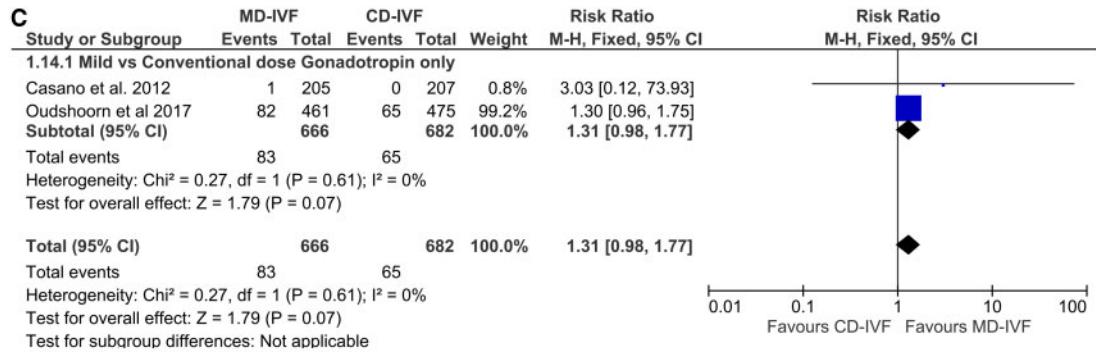
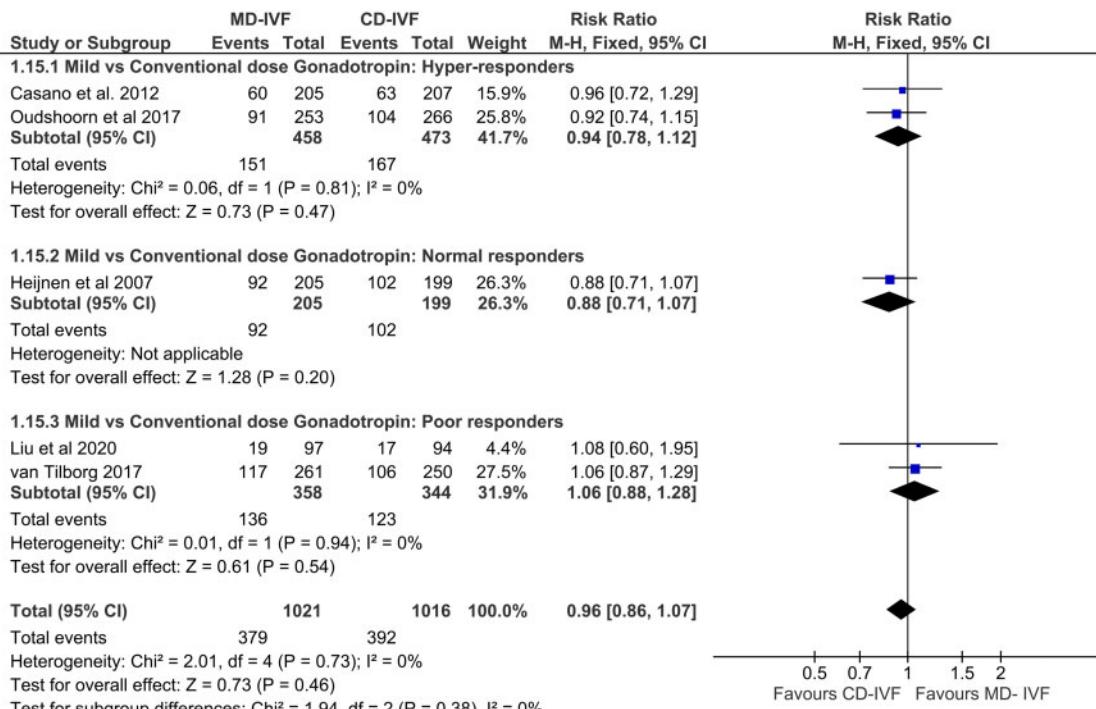


Figure 5. Forest plot of mild versus conventional-dose IVF: cycle cancellation rate per started cycle.

**Figure 5** Continued.**Figure 6.** Forest plot of mild versus conventional-dose IVF: cumulative live birth rate per randomisation.

Hyper-responders. CCRs were no different between MD-IVF and CD-IVF in both the RCTs with a gonadotrophin only agonist/antagonist protocol; RR 1.31 (CI 0.98, 1.77). Heterogeneity was absent ($I^2 = 0\%$), and studies were large with low RoB. This evidence, which was based on only two RCTs with diverse protocols, was of moderate quality.

Secondary outcomes

Cumulative LBR

This outcome was investigated in five RCTs in total ($n=2037$): two with hyper-responders (Casano et al., 2012; Oudshoorn et al., 2017),

one with normal (Heijnen et al., 2007) and two with poor responders (van Tilborg et al., 2017; Liu et al., 2020). All studies were large; all except one (Liu et al., 2020) were based on gonadotrophin only protocols. None of the individual RCTs found a difference between MD-IVF and CD-IVF protocols, and so the meta-analysis was of the pooled data (RR 0.96 (CI 0.86, 1.07)) (Fig. 6). All studies had a low RoB, one had an unclear RoB, I^2 was 0% and the conclusion remained unchanged when the only study that allowed dose adjustment (Casano et al., 2012) was excluded from the meta-analysis. However, the trials were conducted in different population types and the description of 'cumulative' livebirth was different (as described above).

Ongoing pregnancy rate

Poor responders. Seven RCTs ($n=2006$) reported OPRs: three compared low-dose with high-dose gonadotrophin protocols (Klinkert et al., 2005; van Tilborg et al., 2017; Youssef et al., 2017), two trials with CC (Martinez et al., 2003; Revelli et al., 2014) and the other two with Letrozole incorporated protocols (Bastu et al., 2016; Liu et al., 2020). Meta-analysis of pooled data found no difference in OPRs: RR of 1.01 (CI 0.81, 1.25). There was no statistical heterogeneity between the studies (I^2 0%), and all four large trials were of low RoB (Revelli et al., 2014; van Tilborg et al., 2017; Youssef et al., 2018; Liu et al., 2020). However, due to two small RCTs having an area of 'unclear RoB' and clinical heterogeneity among the study protocols, the overall QoE was moderate (Table III). If smaller studies with 'unclear RoB' or studies with dose adjustments were excluded, the inference of the meta-analysis remained the same. The inclusion of large RCTs having low RoB strengthened the QoE in the subgroup comparing low-dose with high-dose gonadotrophin only.

Normal responders. Seven RCTs ($n=1026$) in this category estimated OPR, six with gonadotrophin only stimulation (Hohmann et al., 2003; Out et al., 2004; Tan et al., 2005; Baart et al., 2007; Lou and Huang, 2010; Blockeel et al., 2011) and one was with CC combination (Karimzadeh et al., 2010). The RR of pooled data 1.10 (CI 0.88, 1.38) was not significant. Both statistical and clinical heterogeneity were low, and CI was narrow; however, this finding is based on predominantly gonadotrophin only protocols and presence of unclear RoB in multiple studies led to moderate QoE (Table III).

Hyper-responders. Only one RCT in this population reported OPR (Oudshoom et al., 2017). This large RCT with low RoB found no difference on OPR with a RR 0.86 (CI 0.61, 1.23).

Clinical pregnancy rate

Poor responders. Twelve RCTs ($n=2211$) on poor responders reported CPR. There was no significant difference in CPRs with an RR of 0.96 (CI 0.79, 1.16). The CI was narrow, statistical heterogeneity was absent (I^2 0%) and the finding of the meta-analysis remained the same in sensitivity analysis. However, diversity in the clinical protocol resulted in a moderate QoE.

Normal responders. Three RCTs on gonadotrophin only (Out et al., 2004; Tan et al., 2005; Lou and Huang, 2010) and four on oral compound+ gonadotrophin (Harrison et al., 1994; Lin et al., 2006; Karimzadeh et al., 2010; Mukherjee et al., 2012) analysed the CPR. Meta-analysis showed no difference in the CPR between mild and conventional-dose arms (RR 1.10 (CI 0.92, 1.31)). Three RCTs reported PR, defined as positive pregnancy test (urine or serum β -hCG): of them one study reported a significantly lower PR per cycle with MD-IVF (Dhont et al., 1995) while the other two trials found no difference (Hohmann et al., 2003; Blockeel et al., 2011). Two studies did not specify whether it was positive test or clinical pregnancy (Tummon et al., 1992; Elnashar et al., 2016), and these five studies were excluded from the meta-analysis on CPR. There was no statistical heterogeneity (I^2 0%) and CI was narrow; however, the studies were of small sample size with multiple 'unclear RoB' and diverse treatment protocols. Consequently, the QoE was low.

Hyper-responders. Meta-analysis of two large RCTs with low RoB in this group found no difference in CPR (RR 0.91 (CI 0.82, 1.01)). Differences in the methodology made the QoE moderate.

Total number of oocytes retrieved

Poor responders. All RCTs compared the number of oocytes retrieved; meta-analysis of 14 trials ($n=2773$) that reported the mean number of oocytes found a significantly lower number of oocytes recovered in MD-IVF group, with an SMD of -0.43 (CI -0.58 , -0.28). The other study (Klinkert et al., 2005) that expressed the figures in the median found no difference in the oocyte number. The effect size was large; but most of the studies had area(s) of 'unclear RoB' and one had an area of 'high RoB'; there was significant statistical (I^2 67%) and clinical heterogeneity, therefore, the QoE was low (Table III).

Normal responders. Meta-analysis from 13 RCTs ($n=3499$) revealed fewer oocytes with low-dose stimulation (SMD -1.34 (CI -1.94 , -0.75)), whether it was with a gonadotrophin only protocol (seven trials) or with CC (five trials). The only trial with Letrozole in the MD-IVF arm did not find any difference in the mean oocyte numbers (Mukherjee et al., 2012). The QoE, however, was very low in the presence of high statistical (I^2 98%) and clinical heterogeneity, multiple RoB and wide CI (Table III).

Hyper-responders. Pooled data from two large RCTs with low RoB found no difference in the mean oocyte number (SMD -0.31 (-0.74 , 0.13)). The QoE was low due to significant statistical (I^2 91%) and clinical heterogeneity (Table III).

Embryos created—total

No data on this outcome for the hyper-responders were available.

Poor responders. Ten RCTs on poor responders compared total number of embryos created; a meta-analysis of nine of them ($n=1559$) (Goswami et al., 2004; Kim et al., 2009; Mohsen and El Din, 2013; Huang et al., 2015; Bastu et al., 2016; van Tilborg et al., 2017; Youssef et al., 2018; Yu et al., 2018; Liu et al., 2020) found a lower mean of total embryos with MD-IVF than CD-IVF (SMD -0.39 (CI -0.59 , -0.20)), while the other trial expressed in the median (range) found no difference (Heijnen et al., 2005). Two large RCTs with low RoB used a gonadotrophin only regimen and found fewer embryos from MD-IVF (SMD -0.25 (CI -0.38 , -0.12); the finding was the same with CC/Letrozole regimens (SMD -0.36 (CI -0.62 , -0.10)), which were used in smaller studies with multiple areas of 'unclear RoB' and an area of high RoB. Overall, significant statistical (I^2 59%) and clinical heterogeneity resulted in low QoE (Table III). Exclusion of small studies with RoB did not change the inference.

Normal responders. Seven RCTs ($n=1884$) compared the mean of total embryos. Three trials were on gonadotrophin only protocols (Tan et al., 2005; Baart et al., 2007; Heijnen et al., 2007), the rest were a CC+ gonadotrophin regimen (Tummon et al., 1992; Harrison et al., 1994; Lin et al., 2006; Karimzadeh et al., 2010). The mean of total embryos created was lower in the MD-IVF group (SMD -0.30 (-0.58 , -0.08)). The difference was not significant in trials with CC+ gonadotrophin. Although the nature of the studies was more homogeneous, a high level of statistical heterogeneity (I^2 79%), predominantly small studies with wide CI and trials with multiple unclear RoB made this evidence of very low quality (Table III).

Embryos created—high grade

Poor responders. Seven RCTs compared 'top/high-grade' embryos between MD-IVF and CD-IVF: four of them compared the mean number (Kim et al., 2009; Huang et al., 2015; Youssef et al., 2018; Liu et al., 2020)—meta-analysis of these trials ($n=723$) showed no difference

(SMD -0.12 (CI -0.30 , 0.05)); three studies compared the proportion (%) of good-quality embryos (Revelli *et al.*, 2014; Pilehvari *et al.*, 2016; Yu *et al.*, 2018). A meta-analysis was not possible due to unavailability of denominators. However, all three studies found the proportion of good-quality embryos to be no different between the two approaches. A large RCT ($n=640$) with low RoB reported the proportion of embryos scoring >8 points to be 57.6% with MD-IVF and 54.8% with CD-IVF, the difference was not statistically significant (Revelli *et al.*, 2014). Overall, clinical heterogeneity was significant, plus wide CI, and three studies had multiple areas of unclear bias (one had a high RoB (Huang *et al.*, 2015)); hence the QoE was low (Table III).

Normal responders. High-grade embryos were compared in six RCTs, three of them ($n=551$) reported as mean (Harrison *et al.*, 1994; Out *et al.*, 2004; Mukherjee *et al.*, 2012), and three (total population 656) as a proportion (Baart *et al.*, 2007; Karimzadeh *et al.*, 2010; Elnashar *et al.*, 2016). All studies found no difference in mean or percentage of high-grade embryos. Meta-analysis of the mean number showed an MD of -0.18 (-0.49 , 0.13) and the proportion of high-grade embryos showed an RR of 1.07 (0.93, 1.23). Although statistical heterogeneity was absent ($I^2 0\%$), this evidence is based on mostly clinically heterogenous small trials with multiple unclear RoB and was therefore of low quality (Table III).

Hyper-responders. Only one large RCT reported the proportion of high-grade embryos to be 46.7% versus 42.1% ($P>0.05$) in MD-IVF and HD-IVF groups, respectively (Casano *et al.*, 2012).

Total gonadotrophin dose used

Poor responders. All included RCTs, except the one that did not use any gonadotrophin in the MD-IVF protocol (Ragni *et al.*, 2012), compared total amount of gonadotrophin used between the groups. There was a high level of clinical and statistical heterogeneity ($I^2 96\%$) among the studies and many RCTs had area(s) of 'unclear bias'. However, all individual trials found less gonadotrophin requirement in the MD-IVF programme, with a large effect-size (SMD -3.17 , CI -3.80 , -2.54) in the meta-analysis of 13 trials that measured the mean of stimulation dose (Table III). The two RCTs reporting the median dose also found the same result (Klinkert *et al.*, 2005; Liu *et al.*, 2020).

Normal responders. Eleven trials ($n=2583$) compared gonadotrophin use among normal responders, four of them used gonadotrophin only protocols (Out *et al.*, 2004; Tan *et al.*, 2005; Heijnen *et al.*, 2007; Lou and Huang, 2010); five studies used CC+ gonadotrophin as MD-IVF (Tummon *et al.*, 1992; Harrison *et al.*, 1994; Dhont *et al.*, 1995; Lin *et al.*, 2006; Karimzadeh *et al.*, 2010); the remaining two trials were based on Letrozole (Mukherjee *et al.*, 2012; Elnashar *et al.*, 2016). All individual studies reported lower gonadotrophin use with MD-IVF. Meta-analysis found an SMD of -5.86 (CI -7.06 , -4.66) with a large effect size. However, significant statistical ($I^2 99\%$) and clinical heterogeneity, along with studies with unclear RoB, led to a low QoE (Table III).

Hyper-responders. One large RCT with low RoB that compared gonadotrophin dose found a lower total dose used in the MD-IVF group: MD -394.00 (CI -481.20 , -306.80) (Casano *et al.*, 2012).

Treatment cost

Two included RCTs on poor responders (Ragni *et al.*, 2012; van Tilborg *et al.*, 2017), three among normal (Heijnen *et al.*, 2007; Lou and Huang, 2010; Mukherjee *et al.*, 2012) and one in the hyper-responder patient-category (Oudshoorn *et al.*, 2017) performed a

cost-analysis. Both trials on the poor responder were large with low RoB. One study found MD-IVF was associated with a per-cycle cost-saving of € 2620 with no use of gonadotrophin in the MD-IVF arm and a 450 IU daily dose in the CD-IVF arm (Ragni *et al.*, 2012). The other RCT reported a reduced cumulative treatment cost with the lower gonadotrophin dose regimen by €1099 (van Tilborg *et al.*, 2017). Cost with standard deviation was not reported in both the studies; therefore, a meta-analysis could not be performed. A larger RCT among the normal responders found a cumulative cost difference of €-2412.00 (Heijnen *et al.*, 2007). The other study by Lou reported the treatment cost of 1056 ± 111 and $16\ 776 \pm 3921$ yuan ($€136 \pm 14.3$ versus $€2160.04 \pm 505$) ($P<0.001$) for mild and conventional treatment, respectively (Lou and Huang, 2010). Converting yuan to euro at the current conversion rate, a meta-analysis of these two trials on normal responders also found MD-IVF to be less expensive (MD -2028.21 (CI -2208.00 , -1848.41)). The RCT by Mukherjee *et al.* reported 34% less average cost with letrozole-based protocol. The only RCT on the hyper-responders, however, did not find any approach cheaper than the other (Oudshoorn *et al.*, 2017). Overall, the study protocols including health-economic models were different between the trials, hence the QoE was low.

Discussion

Findings of the review

Meta-analyses of pregnancy outcome data found no difference in pregnancy outcomes: LBR, cumulative LBR, OPR and CPR between mild and conventional stimulation in poor, normal or high responders of IVF. The evidence came from a pooled population, which was adequately powered for these outcomes. The evidence was of moderate quality.

The incidence of OHSS was significantly lower with MD-IVF both in normal and high responders at a moderate QoE. Of note is that none of the included studies used GnRH agonist for ovulation trigger to prevent OHSS.

Overall, the risk of cycle cancellation was comparable among poor responders and high responders but was increased with MD-IVF in normal responders (very low QoE). Noteworthy is that multiple trials with normal responders were conducted before the introduction of GnRH antagonist for prevention of premature ovulation (Table I), and no difference in the CCRs was observed if the studies with no agonist or antagonist were excluded from the meta-analysis. On the other hand, if studies that allowed dose adjustments were taken out of the meta-analysis, the CCRs turned out to be higher with MD-IVF in poor responders only. Although MD-IVF was associated with a fewer oocytes retrieved or fewer embryos created, the chance of obtaining high-grade embryos was found to be no different in poor, normal as well as hyper-responders. MD-IVF appeared to reduce the use of gonadotrophins as well as treatment cost for poor and normal responders. The findings of our meta-analyses remained unchanged on sensitivity analysis or in subgroup analysis separating the gonadotrophin-only regimen from that with oral medications, except in the aforementioned situations.

Strength of this review

Our systematic review with meta-analysis is the first to include only RCTs that used less than or equal to 150 IU of daily gonadotrophin \pm

Table IV Existing systematic reviews on related topics.

Systematic reviews (chronological order)	Intervention and comparator	LBR	OPR	CPR	OHSS	CCR	Comments
Normal responders/ Unselected population							
Sterrenburg et al. (2011)	Gn only low- vs high-dose protocols		↔	↔	↔ / ↑ ^a	Low- vs high-dose with GnRH-a/ Antagonist	
Gibreel et al. (2012)	CC+ Gn vs C-IVF	↔ *	↔ **	↓ *	↑ **	↔ CCR with antagonist use	
Figueiredo et al. (2013)	CC+ Gn vs C-IVF	↔ **	↔ **	↓ *		7 RCTs	
Matsaseng et al. (2013)	Gn only/ CC+ Gn vs C-IVF	↓ [#]	↔	↓	↑	5 RCTs, 4 with Gn only	
Normal & poor responders							
Bechtejew et al. (2017)	CC/Let+ Gn vs C-IVF	↔ **	↔ ***	↓ **	?	Imprecise evidence with Let, more with CC	
Fan et al. (2017)	CC± Gn vs C-IVF	↔ [#]	↔	↔	↑↔ ^b	Normal 3, Poor 3 RCTs	
Kamath et al. (2017)	CC/Let± Gn vs C-IVF	↔ *	↔ */**	↓ *	↑ *	Same findings- normal or poor responders	
Poor responders only							
Song et al. (2016)	CC+ Gn ±. vs C-IVF	↔ [#]	↔		↔		
ASRM (2018)	Gn± CC/Let vs C-IVF	↔	↔			Same findings whether with or without CC/Let	
Youssef et al. (2018)	Lower- vs Higher-dose IVF	↔ [#]	↔	↔	↔ / ↑ ^c	Same findings whether with or without CC/Let, except for CCR	
Datta et al. (2020)	Gn± CC/Let vs C-IVF	↔ **	↔ */**	↔ **	↔ *	15 RCTs on poor, 14 on normal and 2 hyper-responders. Compared cumulative LBR and high-grade embryos	

↔, similar; ↑, high with MD-IVF; ↓, low with MD-IVF; *Low quality of evidence, **Moderate quality of evidence, ***High quality of evidence, ^ahigh with 100 IU dose versus 200 IU dose, no difference between 150 IU versus 225 IU dose, [#]Only 1 RCT, ^bHigh in normal responders, no difference in poor responders, ^cNo difference with Gn only protocol, high with oral agent incorporated protocol.

oral compounds in the mild-IVF group in treating the poor, normal as well as hyper-responders. The ASRM Practice Committee Guideline considered low-dose (≤ 150 IU daily, as did ours) as 'mild ovarian stimulation' for women with POR, however, this review backing the guideline was without a meta-analysis (Practice Committee of the American Society for Reproductive Medicine. Electronic address: ASRM@asrm.org, 2018). To our knowledge, our review is the only one to perform a meta-analysis on the number or proportion of high-grade embryos in the two approaches. We put due emphasis on the evidence related to the risk of OHSS as well as cycle cancellation, considering its health implications, treatment burden and emotional impact. Finally, this up-to-date review gave adequate statistical power to determine the difference in the live birth data.

Limitations

Significant clinical heterogeneity, including inconsistency in defining poor, normal and high responders, variations in the study protocols, dose adjustments in many studies and difference in the cycle cancellation criteria contributed to the evidence being of moderate quality in all pregnancy outcomes; however, this does increase the generalisability of the findings. A sensitivity analysis was carried out with or without the trials that allowed dose adjustment and excluding small studies with RoB. We have

included studies that compared a mild and higher stimulation dose in randomised studies irrespective of the pituitary suppression protocols; there is no evidence that pregnancy outcome differs whether agonist or antagonist protocols were used (Oudshoorn et al., 2017); the inference of our review remains unchanged when sensitivity analysis was carried out based on protocol. Two RCTs that used both GnRH agonist and antagonist protocols reported no difference in their findings when cycles with one of the protocols were excluded. Inclusion of a study that also used gamete/zygote intra-fallopian transfer (GIFT/ZIFT) might be questionable (Dhont et al., 1995). Given that GIFT/ZIFT was undertaken in an equal and small proportion of cases in both the arms, and that a separate sub-analysis excluding the GIFT/ZIFT technique did not alter the outcomes in this study, we felt the inclusion was justified. Different definitions used for cumulative LBR between the RCTs have affected the QoE for this outcome.

Comparison with other reviews

Existing systematic reviews on related topics have been narrated in Table IV. The majority have compared 'mild' versus 'conventional', or protocols with oral compound versus higher stimulation without oral compounds in either poor responders or the population in general with a subgroup analysis on poor responders. There is a narrative review on the hyper-responders (Gat et al., 2015). Some reviews

compared a 'standard' (150 IU) dose with 'individualised' lower or higher doses based on ovarian reserve (Lensen *et al.*, 2018). Only five previous systematic reviews have assessed the QoE as per the GRADE system (Gibreel *et al.*, 2012; Figueiredo *et al.*, 2013; Bechtejew *et al.*, 2017; Kamath *et al.*, 2017; Datta *et al.*, 2020). None of the reviews compared high-grade embryos between mild and higher dose, except our own recent meta-analysis on poor responders only (Datta *et al.*, 2020). One large multi-centre RCT that was excluded from our meta-analysis found no difference in top-quality blastocysts with incremental FSH dose in both low and high AMH groups (Arce *et al.*, 2014). It should be noted that pregnancy outcomes in all existing reviews showed no difference between lower and high gonadotrophin doses, whether gonadotrophin was combined with anti-oestrogens or not; nevertheless, our previous review on the poor responders (Datta *et al.*, 2020), the review by ASRM (Practice Committee of the American Society for Reproductive Medicine. Electronic address: ASRM@asrm.org, 2018) and the current review encompassing all patient types indicate that pregnancy outcomes are not compromised even when a stimulation dose of 150 IU or less was used as MD-IVF. Most of the reviews found lower risk of OHSS with a low-dose regimen; however, the evidence was contradictory on the risk of cycle cancellation (Table IV). The conclusion of our review was similar to other reviews with regard to total gonadotrophin dose. No difference in the number of retrieved oocytes was found in three reviews (Song *et al.*, 2016; Bechtejew *et al.*, 2017; Fan *et al.*, 2017), with a contradictory result in others (Kamath *et al.*, 2017; Datta *et al.*, 2020). Previous reviews identified lack of live birth data, predominance of small studies with RoB and heterogeneity between the trials as the limiting factors in reaching a firm conclusion. Our current systematic review, as well as our other review on the poor responders (Datta *et al.*, 2020), by adding few large recently published RCTs, has consolidated the evidence from a large pooled population.

Implications for clinical practice

The results of this review suggest that a mild ovarian stimulation for IVF can be considered for poor, normal and hyper-responders without compromising pregnancy outcomes but reducing treatment burden and cost. Our review supports the recent ASRM recommendation that mild ovarian stimulation should be considered for IVF treatment in poor responders. Our conclusion also is in the line with a recent Cochrane review that found increasing or decreasing the stimulation dose according to ovarian reserve did not improve the pregnancy outcome over a fixed dose of 150 IU a day, while a dose less than 150 IU significantly reduced the incidence of OHSS (Lensen *et al.*, 2018); many researchers have demonstrated that increasing stimulation dose according to follicular reserve may yield more oocytes but that did not translate into an improvement in the pregnancy outcomes (Arce *et al.*, 2014; Leijdekkers *et al.*, 2019). Overall, this updated review adds more data in favour of the mild approach, which could make IVF more patient-friendly, affordable and thereby more accessible worldwide.

Implications for research

Despite publications of large trials in recent years, the evidence on the effectiveness of mild IVF did not reach high quality, mainly due to the methodological differences between the studies and lack of agreement

on the study protocols. Future trials comparing mild IVF (≤ 150 IU dose) with or without oral compounds need to aim to compare the cumulative live birth outcomes from an adequately powered sample size and from the perspective of women with poor, normal or high ovarian response. Cumulative outcome from a single oocyte retrieval procedure with fresh and subsequent frozen transfer cycles would be preferred to a time-specific cumulative outcome. An agreement on defining participants' characteristics (e.g. Bologna criteria for poor responders) and defining mild stimulation protocols will minimise heterogeneity and make the conclusion more generalisable. More data on patient's acceptability are needed for a proper evaluation of mild IVF. As the numbers needed for adequate sample size will be large, an international collaboration must be explored to answer this very important question and provide a high QoE.

Acknowledgements

We thank the National Guideline Alliance (NGA) of Royal College of Obstetricians and Gynaecologists to conduct the electronic search for the studies.

Authors' roles

A.K.D. conceptualised, set methodology, reviewed included studies including the quality of evidence, analysed data and prepared the first draft. A.M. reviewed and edited the study principles and methodology and revised the manuscript. N.F. contributed in setting the selection criteria, reviewed included studies including the quality of evidence. S.C. reviewed study principles, resolved conflict between the study-reviewers and finalised the manuscript. G.N. reviewed and finalised the study principles, terminology and approved the manuscript.

Funding

No funding was required to conduct this systematic review and meta-analysis.

Conflict of interest

A.M. has been paid honoraria by Cooks, Ferring and Merck for speaking at educational meetings. The other authors have no conflicts of interest.

References

Arce JC, Andersen AN, Fernandez-Sanchez M, Visnova H, Bosch E, Garcia-Velasco JA, Barri P, de Sutter P, Klein BM, Fauser BC. Ovarian response to recombinant human follicle-stimulating hormone: a randomized, antimullerian hormone-stratified, dose-response trial in women undergoing in vitro fertilization/intracytoplasmic sperm injection. *Fertil Steril* 2014; **102**: 1633–1640.e5.

Ashrafi M, Ashtiani SK, Zafarani F, Samani RO, Eshrati B. Evaluation of ovulation induction protocols for poor responders undergoing assisted reproduction techniques. *Saudi Med J* 2005; **26**: 593–596.

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

Bastu E, Buyru F, Ozsurmeli M, Demiral I, Dogan M, Yeh J. A randomized, single-blind, prospective trial comparing three different gonadotropin doses with or without addition of letrozole during ovulation stimulation in patients with poor ovarian response. *Eur J Obstet Gynecol Reprod Biol* 2016; **203**: 30–34.

Bechtejew TN, Nadai MN, Nastri CO, Martins WP. Clomiphene citrate and letrozole to reduce follicle-stimulating hormone consumption during ovarian stimulation: systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 2017; **50**: 315–323.

Blockeel C, Sterrenburg MD, Broekmans FJ, Eijkemans MJ, Smitz J, Devroey P, Fauser BC. Follicular phase endocrine characteristics during ovarian stimulation and GnRH antagonist cotreatment for IVF: RCT comparing recFSH initiated on cycle day 2 or 5. *J Clin Endocrinol Metab* 2011; **96**: 1122–1128.

Casano S, Guidetti D, Patriarca A, Pittatore G, Gennarelli G, Revelli A. MILD ovarian stimulation with GnRH-antagonist vs. long protocol with low dose FSH for non-PCO high responders undergoing IVF: A prospective, randomized study including thawing cycles. *J Assist Reprod Genet* 2012; **29**: 1343–1351.

Datta AK, Maheshwari A, Felix N, Campbell S, Nargund G. Mild versus conventional ovarian stimulation for IVF in poor responders: a systematic review and meta-analysis. *Reprod Biomed Online* 2020; **41**: 225–238.

de Klerk C, Macklon NS, Heijnen EM, Eijkemans MJ, Fauser BC, Passchier J, Hunfeld JA. The psychological impact of IVF failure after two or more cycles of IVF with a mild versus standard treatment strategy. *Hum Reprod* 2007; **22**: 2554–2558.

Dhont M, Onghena A, Coetsier T, De Sutter P. Prospective randomized study of clomiphene citrate and gonadotrophins versus goserelin and gonadotrophins for follicular stimulation in assisted reproduction. *Hum Reprod* 1995; **10**: 791–796.

Drakopoulos P, Blockeel C, Stoop D, Camus M, de Vos M, Tournaye H, Polyzos NP. Conventional ovarian stimulation and single embryo transfer for IVF/ICSI. How many oocytes do we need to maximize cumulative live birth rates after utilization of all fresh and frozen embryos? *Hum Reprod* 2016; **31**: 370–376.

Ebrahimi M, Akbari-Asbagh F, Ghalandar-Attar M. Letrozole+ GnRH antagonist stimulation protocol in poor ovarian responders undergoing intracytoplasmic sperm injection cycles: an RCT. *IJRM* 2017; **15**: 101–108.

Eftekhari M, Mohammadian F, Davar RS. Comparison of pregnancy outcome after letrozole versus clomiphene treatment for mild ovarian stimulation protocol in poor responders. *Iranian Journal Ofreproductive Medicine* 2014; **12**: 725.

Elnashar I, Farghaly TA, Abdalbadie AS, Badran E, Abdelaleem AA, Ismail AM, Elsenosy E. Low cost ovarian stimulation protocols associated with lower pregnancy rate in normal responders in comparison to long protocol. *Fertil Steril* 2016; **106**(Suppl 3): e194–e195.

Fan Y, Zhang X, Hao Z, Ding H, Chen Q, Tian L. Effectiveness of mild ovarian stimulation versus GnRH agonist protocol in women undergoing assisted reproductive technology: a meta-analysis. *Gynecol Endocrinol* 2017; **33**: 746–756.

Ferraretti AP, La Marca A, Fauser BC, Tarlatzis B, Nargund G, Gianaroli L; ESHRE working group on Poor Ovarian Response Definition. ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. *Hum Reprod* 2011; **26**: 1616–1624.

Figueiredo JB, Nastri CO, Vieira AD, and Martins WP. Clomiphene combined with gonadotropins and GnRH antagonist versus conventional controlled ovarian hyperstimulation without clomiphene in women undergoing assisted reproductive techniques: systematic review and meta-analysis. *Arch Gynecol Obstet* 2013; **287**: 779–790.

Fujimoto A, Harada M, Hirata T, Osuga Y, Fujii T. Efficacy of clomiphene citrate supplementation to conventional GnRH antagonist protocols in poor responders undergoing assisted reproductive technology- a prospective randomized trial. *Fertility and Sterility* 2014; **102**: e65.10.1016/j.fertnstert.2014.07.224.

Gat I, Shlush E, Quach K, Librach CL. The continuum of high ovarian response: a rational approach to the management of high responder patient subgroups. *Syst Biol Reprod Med* 2015; **61**: 336–344.

Ghoshdastidar S, Maity S and Ghoshdastidar B. Improved ICSI outcome in poor responders using a novel stimulation regime with micro-dose flare followed by GnRH antagonist in mid follicular phase. *Human Reproduction* 2010; **25**(Suppl 1): i316.

Gibreel A, Maheshwari A, Bhattacharya S. Clomiphene citrate in combination with gonadotropins for controlled ovarian stimulation in women undergoing in vitro fertilization. *Cochrane Database Syst Rev* 2012; **11**: CD008528.

Goswami SK, Das T, Chattopadhyay R, Sawhney V, Kumar J, Chaudhury K, Chakravarty BN, Kabir SN. A randomized single-blind controlled trial of letrozole as a low-cost IVF protocol in women with poor ovarian response: a preliminary report. *Hum Reprod* 2004; **19**: 2031–2035.

Grochowski D, Wołczyński S, Kuczyński W, Domitrz J, Szamatowicz J, Szamatowicz M. Good results of milder form of ovarian stimulation in an in vitro fertilization/intracytoplasmic sperm injection program. *Gynecological Endocrinology* 1999; **13**: 297–304.

Harrison RF, Kondaveeti U, Barry-Kinsella C, Gordon A, Drudy L, Cottell E, Hennelly B, Frankish A, and Unwin A. Should gonadotropin-releasing hormone down-regulation therapy be routine in in vitro fertilization? *Fertil Steril* 1994; **62**: 568–573.

Heijnen EM, Eijkemans MJ, De Klerk C, Beckers NG, Klinkert ER, Broekmans FJ, Passchier J, Te Velde ER, Habbema JD, Macklon NS et al. Mild stimulation and single embryo transfer: towards safer and simpler IVF. The 21st annual meeting of the European Society of Human Reproduction and Embryology. *Human Reprod* 2020; **20**(Suppl 1): i141.

Heijnen EM, Eijkemans MJ, De Klerk C, Polinder S, Beckers NG, Klinkert ER, Broekmans FJ, Passchier J, Te Velde ER, Macklon NS et al. A mild treatment strategy for in-vitro fertilisation: a randomised non-inferiority trial. [Reprint in Ned Tijdschr Geneeskd. 2008 Apr 5;152(14):809–16; PMID: 18491824]. *Lancet* 2007; **369**: 743–749.

Higgins JPT, Green, S. (eds) *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 edn. The Cochrane Collaboration, 2011; London.

Hohmann FP, Macklon NS, Fauser BC. A randomized comparison of two ovarian stimulation protocols with gonadotropin-releasing hormone (GnRH) antagonist cotreatment for in vitro fertilization commencing recombinant follicle-stimulating hormone on cycle day 2 or 5 with the standard long GnRH agonist protocol. *J Clin Endocrinol Metab* 2003; **88**:166–173.

Hogaard A, Ingerslev HJ, Dinesen J. Friendly IVF: patient opinions. *Hum Reprod* 2001; **16**:1391–1396.

Hoomans EH, Mulder BB, and Asian Purgeon Study G. A group-comparative, randomized, double-blind comparison of the efficacy and efficiency of two fixed daily dose regimens (100- and 200-IU) of recombinant follicle stimulating hormone (rFSH, Puregon) in Asian women undergoing ovarian stimulation for IVF/ICSI. *J Assist Reprod Genet* 2002; **19**:470–476.

Huang R, Wang B, Yang X, Li TT, Liang XY. The comparison of mild stimulation vs. Controlled ovarian hyperstimulation protocol in poor ovarian responders: a prospective randomized study. *Hum Reprod* 2015; **1**:i49–i50.

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

Jindal A, Singh R. A prospective randomised controlled study comparing a low-cost antagonist protocol using oral ovulation inducing agents in IVF-ICSI cycles with a standard agonist long protocol. *Fertility and Sterility* 2013; **100**:S273 10.1016/j.fertnstert.2013.07.1137.

Kamath MS, Maheshwari A, Bhattacharya S, Lor KY, Gibreal A. Oral medications including clomiphene citrate or aromatase inhibitors with gonadotropins for controlled ovarian stimulation in women undergoing in vitro fertilisation. *Cochrane Database Syst Rev* 2017; **11**:CD008528.

Karimzadeh MA, Ahmadi S, Oskouian H, Rahmani E. Comparison of mild stimulation and conventional stimulation in ART outcome. *Arch Gynecol Obstet* 2010; **281**:741–746.

Kim CH, Kim SR, Cheon YP, Kim SH, Chae HD, Kang BM. Minimal stimulation using gonadotropin-releasing hormone (GnRH) antagonist and recombinant human follicle-stimulating hormone versus GnRH antagonist multiple-dose protocol in low responders undergoing in vitro fertilization/intracytoplasmic sperm injection. *Fertil Steril* 2009; **92**:2082–2084.

Kingsland C, Tan S-L, Bickerton N, Mason B, Campbell S. The routine use of gonadotropin-releasing hormone agonists for all patients undergoing in vitro fertilization. Is there any medical advantage? A prospective randomized study. *Fertility and Sterility* 1992; **57**:804–809.

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

Lee V, Chan C, Ng E, Yeung W, Ho PC. Will letrozole improve the ovarian response and live-birth rate in women with poor ovarian reserve who are undergoing in-vitro fertilisation treatment? A randomised controlled trial. *Human Fertility* 2011; **14**:70.

Leijdekkers JA, Torrance HL, Schouten NE, van Tilborg TC, Oudshoorn SC, Mol BWJ, Eijkemans MJC, Broekmans FJM. Individualized ovarian stimulation in IVF/ICSI treatment: it is time to stop using high FSH doses in predicted low responders. *Hum Reprod* 2019; **35**:1954–1963.

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). *Cochrane Database Syst Rev* 2018; **2**:CD012693.

Lin YH, Hwang JL, Seow KM, Huang LW, Hsieh BC, Tzeng CR. Comparison of outcome of clomiphene citrate/human menopausal gonadotropin/cetrorelix protocol and buserelin long protocol—a randomized study. *Gynecol Endocrinol* 2006; **22**:297–302.

Liu X, Li T, Wang B, Xiao X, Liang X, Huang R. Mild stimulation protocol vs conventional controlled ovarian stimulation protocol in poor ovarian response patients: a prospective randomized controlled trial. *Arch Gynecol Obstet* 2020; **301**:1331–1339.

Long CA, Sopelak VM, Lincoln SR, Cowan BD. Luteal phase consequences of low-dose gonadotropin-releasing hormone agonist therapy in nonluteal-supported in vitro fertilization cycles**Supported in part by the Vicksburg Hospital Medical Foundation, Vicksburg, Mississippi. *Fertility and Sterility* 1995; **64**: 573–576.

Lou HY, Huang XY. Modified natural cycle for in vitro fertilization and embryo transfer in normal ovarian responders. *J Int Med Res* 2010; **38**:2070–2076.

Martinez F, Coroleu B, Marques L, Parera N, Buxaderas R, Tur R, PN Barri. Comparison of "short protocol" versus "antagonists" with or without clomiphene citrate for stimulation in IVF of patients with "low response" [Spanish]. *Rev Iberoam Fertil Reprod Hum* 2003; **20**:355–360.

Matsaseng T, Kruger T, Steyn W. Mild ovarian stimulation for in vitro fertilization: are we ready to change? A meta-analysis. *Gynecol Obstet Invest* 2013; **76**:233–240.

Mohsen IA, El Din RE. Minimal stimulation protocol using letrozole versus microdose flare up GnRH agonist protocol in women with poor ovarian response undergoing ICSI. *Gynecol Endocrinol* 2013; **29**:105–108.

Mukherjee S, Sharma S, Chakravarty BN. Letrozole in a low-cost in vitro fertilization protocol in intracytoplasmic sperm injection cycles for male factor infertility: a randomized controlled trial. *J Hum Reprod Sci* 2012; **5**:170–174.

Nabati A, Peivandi S, Khalilian A, Mirzaeirad S, Hashemi S A. Comparison of GnRH Agonist Microdose Flare Up and GnRH Antagonist/Letrozole in Treatment of Poor Responder Patients in Intra Cytoplasmic Sperm Injection: Randomized Clinical Trial. *GJHS* 2015; **8**:166 10.5539/gjhs.v8n4p166.

Nargund G, Datta AK, Fauser B. Mild stimulation for in vitro fertilization. *Fertil Steril* 2017; **108**:558–567.

Nargund G, Fauser BC, Macklon NS, Ombelet W, Nygren K, Frydman R, Rotterdam ISMAAR Consensus Group on Terminology for Ovarian Stimulation for IVF. The ISMAAR proposal on terminology for ovarian stimulation for IVF. *Hum Reprod* 2007; **22**:2801–2804.

Nargund G, Fauser BCJM. Mild ovarian stimulation for IVF is the smartest way forward. *Reprod Biomed Online* 2020 (article in press).

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.

Out HJ, Rutherford A, Fleming R, Tay CC, Trew G, Ledger W, Cahill D. A randomized, double-blind, multicentre clinical trial comparing starting doses of 150 and 200 IU of recombinant FSH in women treated with the GnRH antagonist ganirelix for assisted reproduction. *Hum Reprod* 2004; **19**:90–95.

Ozcan Cenksoy P, Ficicioglu C, Kizilkale O, Suhha Bostanci M, Bakacak M, Yesiladali M, Kaspar C. The comparison of effect of microdose GnRH-a flare-up, GnRH antagonist/aromatase inhibitor letrozole and GnRH antagonist/clomiphene citrate protocols on IVF outcomes in poor responder patients. *Gynecological Endocrinology* 2014; **30**:485–489.

Pilehvari S, Shahrokh Tehraninejad E, Hosseiniashidi B, Keikhah F, Haghollahi F, Aziminekoo E. Comparison pregnancy outcomes between minimal stimulation protocol and conventional GnRH antagonist protocols in poor ovarian responders. *J Family Reprod Health* 2016; **10**:35–42.

Popovic-Todorovic B. 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. *Human Reproduction* 2003; **18**: 2275–2282.

Practice Committee of the American Society for Reproductive Medicine. Electronic address: ASRM@asrm.org. Comparison of pregnancy rates for poor responders using IVF with mild ovarian stimulation versus conventional IVF: a guideline. *Fertil Steril* 2018; **109**:993–999.

Pruksananonda K, Suwajanakorn S, Sereepapong W. Comparison of two different fixed doses of follitropin-beta in controlled ovarian hyperstimulation: A prospective randomized, double blind clinical trial. *J Med Assoc Thai* 2004; **87**:1151.

Ragni G, Levi-Setti PE, Fadini R, Brigante C, Scarduelli C, Alagna F, Arfuso V, Mignini-Renzini M, Candiani M, Paffoni A. Clomiphene citrate versus high doses of gonadotropins for in vitro fertilisation in women with compromised ovarian reserve: a randomised controlled non-inferiority trial. *Reprod Biol Endocrinol* 2012; **10**:114.

Revelli A, Chiadò A, Dalmasso P, Stabile V, Evangelista F, Basso G, and Benedetto C. "Mild" vs. "long" protocol for controlled ovarian hyperstimulation in patients with expected poor ovarian responsiveness undergoing in vitro fertilization (IVF): a large prospective randomized trial. *J Assist Reprod Genet* 2014; **31**:809–815.

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.

Schimberni M, Ciardo F, Schimberni M, Giallonardo A, De Pratti V, Sbracia M. Short gonadotropin-releasing hormone agonist versus flexible antagonist versus clomiphene citrate regimens in poor responders undergoing in vitro fertilization: a randomized controlled trial. *Eur Rev Med Pharmacol Sci* 2016; **20**:4354.

Selman H, Rinaldi L. Effectiveness of corifollitropin alfa used for ovarian stimulation of poor responder patients. *JWH* 2016; **609**–615.

Siristatidis C, Salamalekis G, Dafopoulos K, Basios G, Vogiatzi P, Papantoniou N. Mild versus conventional ovarian stimulation for poor responders undergoing IVF/ICSI: A prospective randomized study. *Human Reproduction* 2016; **31** (Supplement 1): i438–i439.

Song D, Shi Y, Zhong Y, Meng Q, Hou S, Li H. Efficiency of mild ovarian stimulation with clomiphene on poor ovarian responders during IVF/ICSI procedures: a meta-analysis. *Eur J Obstet Gynecol Reprod Biol* 2016; **204**:36–43.

Sterrenburg MD, Veltman-Verhulst SM, Eijkemans MJ, Hughes EG, Macklon NS, Broekmans FJ, Fauser BC. Clinical outcomes in relation to the daily dose of recombinant follicle-stimulating hormone for ovarian stimulation in in vitro fertilization in presumed normal responders younger than 39 years: a meta-analysis. *Hum Reprod Update* 2011; **17**:184–196.

Tan SL, Child TJ, Cheung AP, Fluker MR, Yuzpe A, Casper R, Leung P, Cadesky K, Davis VJ. A randomized, double-blind, multicenter study comparing a starting dose of 100 IU or 200 IU of recombinant follicle stimulating hormone (Puregon) in women undergoing controlled ovarian hyperstimulation for IVF treatment. *J Assist Reprod Genet* 2005; **22**:81–88.

Tummon IS, Daniel SA, Kaplan BR, Nisker JA, Yuzpe AA. Randomized, prospective comparison of luteal leuprolide acetate and gonadotropins versus clomiphene citrate and gonadotropins in 408 first cycles of in vitro fertilization. *Fertil Steril* 1992; **58**: 563–568.

van Tilborg TC, Torrance HL, Oudshoorn SC, Eijkemans MJC, Koks CAM, Verhoeve HR, Nap AW, Scheffer GJ, Manger AP, Schoot BC et al. Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 1: the predicted poor responder. *Hum Reprod* 2017; **32**:2496–2505.

Weigert M, Krischker U, Pöhl M, Poschalko G, Kindermann C, Feichtinger W. Comparison of stimulation with clomiphene citrate in combination with recombinant follicle-stimulating hormone and recombinant luteinizing hormone to stimulation with a gonadotropin-releasing hormone agonist protocol: a prospective, randomized study. *Fertility and Sterility* 2002; **78**:34–39.

Wiklund M. A prospective, randomized comparison of two starting doses of recombinant FSH in combination with cetrorelix in women undergoing ovarian stimulation for IVF/ICSI. *Human Reproduction* 2001; **16**:1676–1681.

Youssef MA, van Wely M, Al-Inany H, Madani T, Jahangiri N, Khodabakhshi S, Alhalabi M, Akhondi M, Ansaripour S, Tokhmechy R et al. A mild ovarian stimulation strategy in women with poor ovarian reserve undergoing IVF: a multicenter randomized non-inferiority trial. *Hum Reprod* 2017; **32**:112–118.

Youssef MA, van Wely M, Mochtar M, Fouda UM, Eldaly A, El Abidin EZ, Elhalwagy A, Mageed Abdallah AA, Zaki SS, Abdel Ghafar MS et al. Low dosing of gonadotropins in in vitro fertilization cycles for women with poor ovarian reserve: systematic review and meta-analysis. *Fertil Steril* 2018; **109**:289–301.

Yu R, Jin H, Huang X, Lin J, Wang P. Comparison of modified agonist, mild-stimulation and antagonist protocols for in vitro fertilization in patients with diminished ovarian reserve. *J Int Med Res* 2018; **46**:2327–2337.

Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, Sullivan E, van der Poel S, International Committee for Monitoring Assisted Reproductive Technology,

World Health Organization. The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary on ART terminology, 2009. *Hum Reprod* 2009;24:2683–2687.

Zhang JJ, Merhi Z, Yang M, Bodri D, Chavez-Badiola A, Repping S and Van Wely M. Minimal stimulation IVF vs conventional IVF: A randomized controlled trial. *American Journal of Obstetrics and Gynecology* 2016;214:96.e91–96.e98.