

REVIEW

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



BIOGRAPHY

Geeta Nargund is Lead Consultant in Reproductive Medicine at St George's Hospital, London and Medical Director of CREATE Fertility and abc ivf. She is President of the International Society for Mild Approaches in Assisted Reproduction (ISMAAR). Her research interests are natural and mild IVF and advanced ultrasound in reproductive medicine.

Adrija Kumar Datta^{1,*}, Abha Maheshwari², Nirmal Felix¹,
Stuart Campbell³, Geeta Nargund⁴

KEY MESSAGE

Our up-to-date review and meta-analysis, which gives adequate power to live birth outcomes, found genuinely low-dose gonadotrophin stimulation with or without oral agents likely to be as effective as high-dose conventional IVF for poor responders, with added apparent advantage of reduced treatment cost.

ABSTRACT

Mild ovarian stimulation is a treatment option for poor responders in IVF treatment. Our updated review evaluated mild IVF solely from randomized controlled trials (RCTs) that used genuine low-dose gonadotrophin (≤ 150 IU daily) alone or in combination with oral medications, comparing it with conventional-dose (>150 IU/ daily) IVF for poor responders. Electronic searches on *MEDLINE*, *Embase*, *The Cochrane Central Register of Controlled Trials* and *PreMEDLINE*, and hand searches from 2002 up to 31 January 2019, identified 14 RCTs, which were compiled with the above inclusion criteria. The risk of bias (ROB) and quality of evidence (QOE) were assessed as per Cochrane Collaboration. Meta-analyses found no difference in live birth rate (four RCTs, $n = 1057$, RR 0.91, CI 0.66 to 1.25) (moderate QOE), ongoing pregnancy rate (six RCTs, $n = 1782$, RR 1.01, CI 0.86 to 1.20) (moderate–high QOE) and cycle cancellation rates (14 RCTs, $n = 2746$, RR 1.38, CI 0.99 to 1.92) (low QOE). Fewer oocytes and embryos were obtained from mild IVF; however, the number and proportion of high-grade embryos were similar. Mild IVF resulted in reduced gonadotrophin use and cost. The inference remained unchanged when smaller studies with ROB were excluded, or whether gonadotrophin alone or combination with oral medication was used. The evidence of equal efficacy from a pooled population, which was adequately powered for live birth, supported a mild IVF strategy for poor responders in preference to more expensive conventional IVF. Although clinical heterogeneity remained a limiting factor, it increased the generalizability of the findings.

¹ CREATE Fertility, 6270 Bishops Court, Birmingham Business Park, Solihull Parkway, Birmingham B37 7YB, UK

² Aberdeen Fertility Centre, NHS Grampian, Foresterhill, Aberdeen, AB25 2ZL and University of Aberdeen, UK

³ CREATE Fertility, 150 Cheapside, London EC2V 6ET, UK and St George's University of London, UK

⁴ CREATE Fertility, 150 Cheapside, London EC2V 6ET, UK and St George's University Hospitals NHS Trust, London, UK

© 2020 Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.

*Corresponding author. E-mail address: adrija@createfertility.co.uk (A K Datta). <https://doi.org/10.1016/j.rbmo.2020.03.005>

1472-6483/© 2020 Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.

Declaration: Abha Maheshwari has been paid an honorarium by Cook Medical. The authors report no financial or commercial conflicts of interest.

KEYWORDS

Conventional stimulation
IVF
Meta-analysis
Mild ovarian stimulation
Poor responders
Systematic review

INTRODUCTION

A high proportion of women do not respond well to ovarian stimulation within an IVF programme. They are labelled as poor responders. To maximize the number of retrieved oocytes, women who are known to be, or predicted to be, a poor responder often receive high dose of gonadotrophins. This increases cost as well as treatment burden. An alternative approach using a low dose of gonadotrophins called 'mild stimulation IVF' has been suggested for this group of patients.

Mild ovarian stimulation for IVF (MS-IVF) is defined as 'a protocol in which the ovaries are stimulated with gonadotrophins, other pharmacological compounds, or both, with an intention of limiting the number of oocytes after stimulation for IVF' according to the *International Glossary on Infertility and Fertility Care* by The International Committee for Monitoring Assisted Reproductive Technology (Zegers-Hochschild et al., 2009). It has benefits of better tolerance of the stimulation process, less treatment-related stress (Hojgaard et al., 2001) and lower cost (Heijnen et al., 2005; Ragni et al., 2012). Several retrospective and prospective studies, including randomized controlled trials (RCTs) evaluated the clinical effectiveness of MS-IVF against conventional IVF (C-IVF) on normal and poor responders; most found no significant difference in pregnancy outcomes (Nargund et al., 2017).

Recently, however, there has been a tendency toward using conventional stimulation and a 'freeze all strategy', irrespective of whether the patients are high, normal or low responders. Nevertheless, the latest evidence from a meta-analysis of RCTs (Roque et al., 2019) clearly suggests no advantage of a conventional stimulation and a freeze-all strategy in poor responder patients. A few systematic reviews have compared conventional high stimulation and lower dose stimulation with or without oral agents on poor responders; however, no consensus was reached among the reviews as to what constitutes a 'mild stimulation' protocol and what gonadotrophin dose can be regarded as 'mild'; as a result, studies comparing high-dose gonadotrophin (>150 IU/day) protocols with even higher

gonadotrophin dose have been included in the reviews (Fan et al., 2017; Youssef et al., 2018).

We, therefore, conducted an up-to-date systematic review and meta-analysis collecting data from parallel-group RCTs to answer the question whether mild ovarian stimulation, as defined by 150 IU/day of gonadotrophin or below, is clinically a beneficial and cost-effective regimen for poor responders compared with widely used conventional high-dose stimulation. In the absence of global agreement, we have taken a daily dose of 150 IU of gonadotrophins or less, as defined by the Practice Committee of The American Society of Reproductive Medicine (*Practice Committee of the American Society for Reproductive Medicine, 2018*) to be considered as a MS-IVF.

MATERIALS AND METHODS

The present review and meta-analysis was conducted in accordance with The *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins, 2011), and the findings were presented according to the PRISMA guideline. The review protocol has been registered with PROSPERO 2018 CRD42018104879 (https://www.crd.york.ac.uk/PROSPERO/export_details_pdf.php)

Criteria for including studies in this review

No restriction was placed on language. Studies from 2002 (introduction of 'mild ovarian stimulation' concept in IVF) up to 31 January 2019 were included. Abstract 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

Randomized controlled trial with parallel group comparison.

Participants

Participants included infertile couples (any cause) who had undergone IVF and intracytoplasmic sperm injection, with the female partner known to be a poor responder, based on previous response to ovarian stimulation, or an anticipated poor responder. One or both of the following were taken into account: ovarian reserve, based on basal FSH, anti-

Müllerian hormone, antral follicle count and age in any combination; and or the participant fulfilled the Bologna criteria (Ferraretti et al., 2011)

Intervention

Mild stimulation studies included in the present systematic review used low-dose (≤ 150 IU daily) gonadotrophin alone or in combination with oral agents, e.g. clomiphene citrate, aromatase inhibitor or oral agents only.

Conventional stimulation studies used conventional high-dose (>150 IU daily starting dose) ovarian stimulation for poor responders, either within a 'long' or 'short' down-regulation or gonadotrophin releasing hormone antagonist (GnRH-ant) protocol, with or without agonist 'flare' protocols.

Comparison

Mild stimulation was compared with conventional stimulation IVF as defined above.

Studies were excluded if starting gonadotrophin dose was above 150 IU per day with or without oral agents (clomiphene citrate or aromatase inhibitor) in the MS-IVF group or the same dose of gonadotrophin was used in one arm with oral agents (as MS-IVF) and the other without oral agents as C-IVF.

Outcomes

Primary outcomes included live birth rate (LBR), ongoing pregnancy rate (OPR) per woman randomized and cycle cancellation rates per started cycle. Secondary outcomes included clinical pregnancy rates (CPR), total dose of gonadotrophin used, number of oocytes, number of embryos, number of high-grade embryos and cost comparison.

Search method

A high-sensitivity electronic search was conducted in MEDLINE, Embase, The Cochrane Central Register of Controlled Trials and PreMEDLINE between 2002 and 31 January 2019. Databases were searched using relevant medical subject headings, free-text terms and study type filters, where appropriate, without language restrictions. The reference list of all reviews or individual RCTs were also hand-searched to find additional RCTs.

Search term

Search terms were as follows: (IVF, ICSI, ovarian stimulation) AND ((mild

IVF stimulation, oral agents, aromatase inhibitors, clomiphene, letrozole, anastrozole) OR ((gonadotropin, FSH, follitropin, human menopausal gonadotrophin, menotrophin) AND (dose, low dose))) AND randomized controlled trials. We did not add poor responders in the initial search, as various terms were likely to be used to describe this subgroup of population.

Data collection and analysis

First, an electronic search was conducted using the search terms and databases described above. Full text of all shortlisted studies (RCTs) were reviewed by two reviewers (AKD and NF) independently; conflict, if any, was resolved by the third reviewers (SC and GN). 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, 2011) was consulted to prepare the data-extraction form, obtain the features of included studies, and assess risk of bias (ROB) and outcome data. Review Manager 5 (version 5.3) software was used to construct the ROB graph and Forest plots for all meta-analyses 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 methods

Data on trial methods included type of trials (two-arm/three-arm/four-arm), year and location of the trial (single or multi-centre), notes on power calculation, method of randomization, method of allocation concealment, exclusion of participants after randomization, proportion of, and reasons for losses at follow up, and reports of ethical approval and consent.

Participants

Information on participants included age, ovarian reserve of the women e.g. basal FSH, anti-Müllerian hormone, antral follicle count, previous IVF and intracytoplasmic sperm injection cycles, whether in accordance with Bologna criteria (Ferraretti *et al.*, 2011) and exclusion criteria of individual trials.

Intervention

Data on interventions included type of intervention and control with regards to type of medications, dose, protocol of ovarian stimulation, pre-treatment or co-intervention, if any, ovulation trigger type and dose, cancellation criteria and luteal phase regimen were noted. Additional information, including the number and stage of embryo transfer, were also looked for.

Outcomes

In addition to recording the reported outcomes, how the primary outcomes were defined and the timing of outcome measurement, e.g. per woman randomized or started cycle, were recorded. An 'intention to treat' analysis, e.g. per woman randomized or started cycle, was undertaken to incorporate the effect of cycle cancellation on the pregnancy outcomes, particularly in treating the poor responders (Orvieto *et al.*, 2017).

Assessment of risk of bias

The risk of bias (ROB) was assessed under the following headings: sequence generation, allocation concealment, blinding of participants and assessors and other source of bias.

Sequence generation was considered at low ROB if, for example a computer random number generator, a 'random number table' or coin tossing was used and considered unclear ROB if insufficient information was given. Central allocation or use of sealed envelopes was considered as at low ROB. An open random allocation was deemed at high risk and unclear ROB if the information about the process of allocation was insufficient. Blinding of patients and clinicians was neither possible nor applicable for this particular type of intervention and outcomes, e.g. clinical pregnancy rates. Studies with absence of blinding were considered as low ROB, as it was unlikely to influence outcomes. For selective outcome reporting, studies were considered at low ROB if they were free of selective reporting, e.g. all pre-specified outcomes had been reported with study protocol being available, at high ROB if not all pre-specified primary outcomes were reported and at unclear ROB if insufficient information was obtained. A difference in the baseline characteristics between the study and control groups constituted other sources of bias.

Treatment effect

For dichotomous data, relative risk and for continuous data, mean differences between treatment groups were calculated with 95% confidence intervals. In the presence of heterogeneous data, standardized mean difference (SMD) was used. An increase in the relative risk of an outcome, which may be beneficial, e.g. OPR, or detrimental, e.g. cycle cancellation rate (CCR), has been graphically displayed to the right of the centre line in the meta-analyses and a decrease in the odds was displayed to the left of the centre line.

Assessment of heterogeneity

The clinical and methodological characteristics of all included studies were examined (TABLE 1) and subgroup meta-analysis conducted with similar study protocols. Statistical heterogeneity was assessed by chi-squared test. A low *P*-value or a large chi-squared statistic relative to its degree of freedom was an indication of potentially heterogeneity. The I^2 statistic assessed the effect of the heterogeneity on the meta-analysis; an I^2 of above 50% indicated significant heterogeneity, in which case a 'random effect model' was applied. In all other situations, a 'fixed-effect model' was used.

Reporting bias

Although a funnel plot is not the ideal method of examining publication bias when the number of trials is few, a CPR without subgrouping of studies was plotted. An attempt was also made to minimize publication or reporting bias by not limiting our search by language or time or types of publication. Duplications arising from conference abstract and subsequent full-text paper were carefully excluded.

Subgroup analysis

Subgroup analysis was conducted with different types of mild and conventional stimulation protocols: low-dose GnRH-ant versus conventional-dose gonadotrophin GnRH-ant or GnRH agonist (GnRH-a) protocol; low-dose versus conventional-dose gonadotrophin: both GnRH-a down-regulation or antagonist protocol; clomiphene citrate plus low-dose versus conventional-dose gonadotrophin GnRH-a protocol; clomiphene citrate plus low-dose versus conventional-dose gonadotrophin GnRH-ant protocol; clomiphene citrate only versus conventional-dose GnRH-ant protocol; letrozole plus low-dose

TABLE 1 CHARACTERISTICS OF INCLUDED RANDOMIZED CONTROLLED TRIALS

Author, date (place) (sample size)	Trial type and method	Participants	Interventions	Cancellation criteria	Outcome
<i>Ashrafi et al., 2005</i> (Iran) (n = 131)	Single centre, three-arm RCT. Power? Consent, ethical approval: yes	Previous poor response: <3 follicles >16 mm; oestradiol at trigger <500 pg/l	Study arm 1: HMG 150 IU/day; Study arm 2: CC 100 mg/day day 3–7 + HMG 150 IU/day; no GnRH-ant in either arm. Control: GnRH-agonist long down-regulation nasal 800 µg or subcutaneous 500 µg/day; HMG 225 IU/day starting at oestradiol <50 pg/ml, LH <5 IU/l, progesterone < 2 pg/ml; trigger HCG 10000 IU; dose adjustment: yes; LPS: progesterone in oil IM.	?	Primary: CCR Secondary: number of oocytes, gonadotrophin dose
<i>Bastu et al., 2016</i> (Turkey) (n = 95)	Single centre, three-arm RCT. Power: adequate for number of oocytes. Consent, ethical approval: yes	Age 18–42 years, BMI 19.3–28.9, POR according to Bologna criteria	Study: letrozole 5 mg/day day 2–3 to 6–7 + HMG 75 + FSH 75 IU/day from day 2–3. Control arm 1: HMG 150 + FSH 150 IU/day from day 2/3. Control arm 2: HMG 225 + FSH 225 IU/day from day 2/3; in all GnRH-ant from day 6; trigger: recombinant HCG 250 IU; dose adjustment? LPS: progesterone gel 8%; embryo transfer day 3; SET if <35 years in first cycle, otherwise DET.	No follicle >11 mm on day 8	Primary: number of oocytes and embryos Secondary: OPR, CPR, gonadotrophin dose, CCR
<i>Goswami et al., 2004</i> (India) (n = 38)	Single centre, two-arm RCT. Power? Consent, ethical approval: yes	Age 36–41 years, previous POR 1–3 cycles with long down-regulation. Exclusion: FSH ≥12 iu/l, endometriosis and pelvic surgery	Study: letrozole 2.5 mg/day from day 3–7 + recombinant FSH 75 IU/day day 3 and 8; no GnRH-ant. Control: GnRH-a long down-regulation. Recombinant FSH 300 IU/day at down-regulation; trigger: HCG 10000 IU; dose adjustment up to 300 IU/day in control; LPS: micronized progesterone; day-2 embryo transfer.	No follicular development	Primary: CPR Secondary: number of oocytes and embryos, gonadotrophin dose, CCR
<i>Huang et al. 2015</i> (China) (n = 105)	Single centre, two-arm RCT. Power? Consent? ethical approval: yes	Inclusion: Bologna criteria. Exclusion: >1 failed IVF, adenomyosis, drug allergy	Study: letrozole (dose?) from day 3–7 + recombinant FSH 150 IU/day on day 4, 6 and 8. No GnRH-ant. Control: long down-regulation, recombinant FSH 300 IU/day at down-regulation; trigger: HCG 10000 IU; dose adjustment? LPS: micronized progesterone; day-2 embryo transfer.	–	Primary: CPR Secondary: oocytes and embryos, number of good-grade embryos, gonadotrophin dose, CCR
<i>Kim et al., 2009</i> (South Korea) (n = 90)	Single centre, two-arm RCT. Power? Consent, ethical approval: yes	Inclusion: previous cycle with <4 follicles over 15 mm and <4 oocytes, with high gonadotrophin dose	Study: recombinant FSH 150 IU/day along with GnRH-ant, when leading follicle reached 13–14 mm; trigger: recombinant HCG 250 µg. Control: GnRH-ant with recombinant FSH 225 IU/day from day 3; dose adjustment: yes; LPS: progesterone 8% gel.	?	Primary: LBR Secondary: CPR, gonadotrophin use, number of oocytes/embryo, number of top-grade embryos, donadotrophin dose, CCR
<i>Klinkert et al., 2005</i> (the Netherlands) (n = 52)	Single centre, two-arm RCT. Power: number of oocytes. Consent, ethical approval: yes.	Inclusion: AFC <5. First IVF cycle. Regular period	Study: GnRH-a long down-regulation, recombinant FSH 150 IU/day at down-regulation. Control: GnRH-a long down-regulation, rFSH 300 IU/day at down-regulation; trigger HCG 10000 IU; dose adjustment up to 300 IU/day; LPS: micronized progesterone or HCG. DET if <38 years, 3 embryos if >38 years.	No follicular development	Primary: number of oocytes. Secondary: OPR, CPR, number of embryos, gonadotrophin dose, CCR
<i>Martinez et al., 2003</i> (Spain) (n = 90)	Single centre, four-arm RCT. Power? Consent and ethical approval: yes	Previous POR	Study arm 1: CC 100 mg/day day 4–8 + HMG 150 IU/day from day 5. Study arm 2: CC 100 mg/day day 4–8 + recombinant FSH 150 IU/day from day 5. Control arm 1: HMG 150 + FSH 150 IU/day from day 2/3; GnRH-ant. Control arm 2: HMG 150 IU/day + FSH 150 IU/day plus GnRH-a short flare from day 2/3; trigger: recombinant HCG 250 ug; dose adjustment? LPS: progesterone gel 8%. Embryo transfer day 3, SET if <35 years in first cycle or DET.	<3 follicles after 10 days	Primary: CPR, Secondary: OPR, number of oocytes, gonadotrophin dose, CCR
<i>Mohsen and El Din, 2013</i> (Egypt) (n = 60)	Single centre, two-arm RCT. Power? Consent, ethical approval: yes	Age unselected, BMI<30, ≥1 previous cycle with POR. No endometriosis, pelvic or ovarian surgery, no systemic disease, no severe male factor	Study: Letrozole 2.5 mg bd from day 2–6 + highly purified HMG 150 IU/day from day 7; GnRH-ant at leading follicle 14 mm. Control: GnRH-a from day 2 until ovulation trigger; highly purified HMG 300 IU/day from day 3; trigger: HCG 10000 IU; pre-treatment: oestradiol 2 mg bd from mid-luteal; dose adjustment: yes; LPS: progesterone pessary 400 mg/day.	<2 follicles low/ plateau oestradiol despite increased gonadotrophin dose	Primary: CPR, Secondary: gonadotrophin dose, number of oocytes, number of embryos, gonadotrophin dose, CCR

(continued on next page)

Table 1 – (continued)

<i>Pilehvari et al., 2016</i> (Iran) (n = 77)	Single centre, two-arm RCT. Power? Consent, Ethical approval: yes	Inclusion according to Bologna criteria. Exclusion: systemic disease, IVF treatment within last 3 months	Study: CC 100 mg/day from day 2–6, HMG 150 IU/day from day 5. Control: HMG 300 IU/day from day 2; GnRH-ant at leading follicle 13–14 mm on both arms; trigger: HCG 10000 IU; dose adjustment: yes; LPS: progesterone pessary 400 mg bd.	?	Primary: CPR Secondary: number of oocytes, % high-grade of embryos, gonadotrophin dose, CCR
<i>Ragni et al., 2012</i> (Italy) (n = 291)	Multicentre, two-arm RCT. Power: under-powered for LBR. Consent and ethical approval: yes	Age 18–42 years, FSH>12 IU/l, previous POR (≤3 eggs). Exclusion: ≥1 failed cycle, surgically retrieved spermatozoa	Study: CC 150 mg/day from day 3–7; no gonadotrophin. Control: GnRH-a from day 1/2; recombinant FSH 450 IU/day from day 3; trigger: recombinant HCG 250 ug; dose adjustment: yes; LPS: micronized progesterone.	Low/no follicular growth	Primary: LBR Secondary: CPR, number of oocytes, CCR, cost
<i>Revelli et al., 2014</i> (Italy) (n = 695)	Single centre, two-arm RCT. Power: adequate for number of oocytes. Consent and ethical approval: yes.	Age <43 years, FSH 10–20 IU/l, AMH 0.14–1.0 ng/ml, AFC 4–10.	Study: CC 100 mg from day 2–6; HMG 150 IU/day from day 5; GnRH-ant from day 8. Control: GnRH-a long down-regulation, HMG 300–450 IU/day; trigger: HCG 10000 IU; dose adjustment: no; LPS: progesterone gel 8%.	<1 follicle 10 mm and oestradiol <50 pg/ml day 7/8	Primary: number of oocytes Secondary: OPR, CPR, % high-grade embryos, gonadotrophin dose, CCR
<i>van Tilborg et al., 2017</i> (the Netherlands) (n = 511)	Multicentre, two-arm RCT. Power: adequate for cumulative LBR. Consent and ethical approval: yes	Age <44 yrs, AFC<11, Regular cycle, 1 st IVF cycle, Normal pelvic scan.	Study: recombinant FSH 150 IU/day. Control: recombinant FSH 225 IU/day if AFC 8–10 and 450 IU/day if AFC <8, both arms with either in GnRH-a long down-regulation or GnRH-ant; trigger: HCG 10000 IU; dose adjustment: no. LPS: progesterone pessary; embryo transfer on day 3/5.	<2 follicles of 12 mm, <3 follicles <17 mm.	Primary: cumulative OPR leading to live birth. Secondary: number of oocytes, number of embryos, gonadotrophin dose, CCR, cost
<i>Youssef et al. 2017</i> (Egypt, Iran, Syria) (n = 394)	Multicentre, two-arm RCT. Power: adequate for OPR. Consent and ethical approval: yes	Age 35–43 years, FSH >10 IU/l, AFC <5, previous POR (≤5 eggs). Exclusion: age >43 years, congenital uterine anomaly	Study: recombinant FSH 150 IU/l, from day 5 of last COC, GnRH-ant from day 6 of stimulation. Control: GnRH-a long down-regulation; HMG 450 IU/day from down-regulation; trigger HCG 10000 IU; dose adjustment: no; pre-treatment: COC for MS-IVF. LPS: progesterone pessary or progesterone daily IM.	<2 follicles of <15 mm after day 7	Primary: OPR. Secondary: CPR, pregnancy, number of oocytes and embryos, number of high-grade embryos, gonadotrophin dose, CCR
<i>Yu et al., 2018</i> (China) (n = 166)	Single centre, three-arm RCT. Power: adequate for CPR. Consent and Ethical approval: yes	Age <43 years, BMI <23, FSH ≥15, AMH <1.5 ng/ml, AFC ≤8. Exclusion: endometriosis, endocrine disorder and pelvic surgery	Control arm: GnRH-a long down-regulation (decapetyl 3.7 mg IM single dose on day 3), HMG 225–300 IU/day 28 days after commencement of GnRH-a. Study arm 1: Letrozole 5 mg/day from day 3–7, HMG 75 IU/day from day 4. Study arm 2: HMG 75 IU/day from day 3; GnRH-ant from leading follicle 14 mm in both arms; trigger: HCG 10000 IU; dose adjustment: yes; LPS: progesterone 60 mg IM + dydrogesterone.	?	Primary: CPR Secondary: LBR, number of oocytes per embryo, proportion of high-quality embryos, gonadotrophin dose, CCR

AFC, antral follicle count; AMH, anti-Müllerian hormone; bd, twice daily; BMI, body mass index; CC, clomiphene citrate; CCR, cycle cancellation rate; COC, combined oral contraceptive; CPR, clinical pregnancy rate; DET, double embryo transfer; HMG, human menopausal gonadotrophin; LPS, luteal phase support; ?, not stated; DET, double embryo transfer; GnRH, gonadotrophin releasing hormone; GnRH-ant; GnRH-antagonist; IM, intramuscular; MS-IVF, study group; OPR, ongoing pregnancy rate; POR, poor ovarian response; RCT, randomized controlled trial; SET, single embryo transfer.

gonadotrophin versus conventional dose GnRH-a protocol; and letrozole plus low-dose gonadotrophin versus conventional dose GnRH-ant protocol.

Multi-arm studies

The *Cochrane Hand book for Systematic Review of Intervention* was followed in the meta-analysis of multi-arm studies (*Higgins, 2011*). If one protocol in the study group (MS-IVF) was compared with two different protocols in the comparator group (C-IVF), both the events and populations (denominators) of the single-arm group (MS-IVF) were equally divided and incorporated under respective sub-groups. The same principle applied if there were two different protocols in the study group (MS-IVF) and one in the control (C-IVF). On the other hand,

if experimental or control intervention consisted of two different doses or types of gonadotrophin in two separate arms, both the events and population were combined into one and placed under the same subgroup in the meta-analysis. For continuous data in the above situations, the mean and SD of the common groups were kept the same, only population was equally split into two subgroups.

Sensitivity analysis

A sensitivity analysis was conducted by repeating meta-analyses in following ways: excluding and including small studies with ROB; applying fixed as well as random effect model; and applying relative risk and peto odds ratio as the method of determining effect size. Apart from that, all meta-analyses were redone separately

with the RCTs that compared low-dose and high-dose gonadotrophin without oral ovarian stimulant to note any effect on the outcomes.

Results

The study selection process is presented in **FIGURE 1**. The search strategy (initial broad-coverage electronic search of database, including conference abstract and trial registry) identified 2286 citations. A review of all titles and abstracts enabled us to exclude the duplicate publications, conference abstracts with insufficient information, registered ongoing trials and publication that were not about the poor responders. Following this process, 22 shortlisted RCTs qualified for full-text review. Full-text review identified one trial that

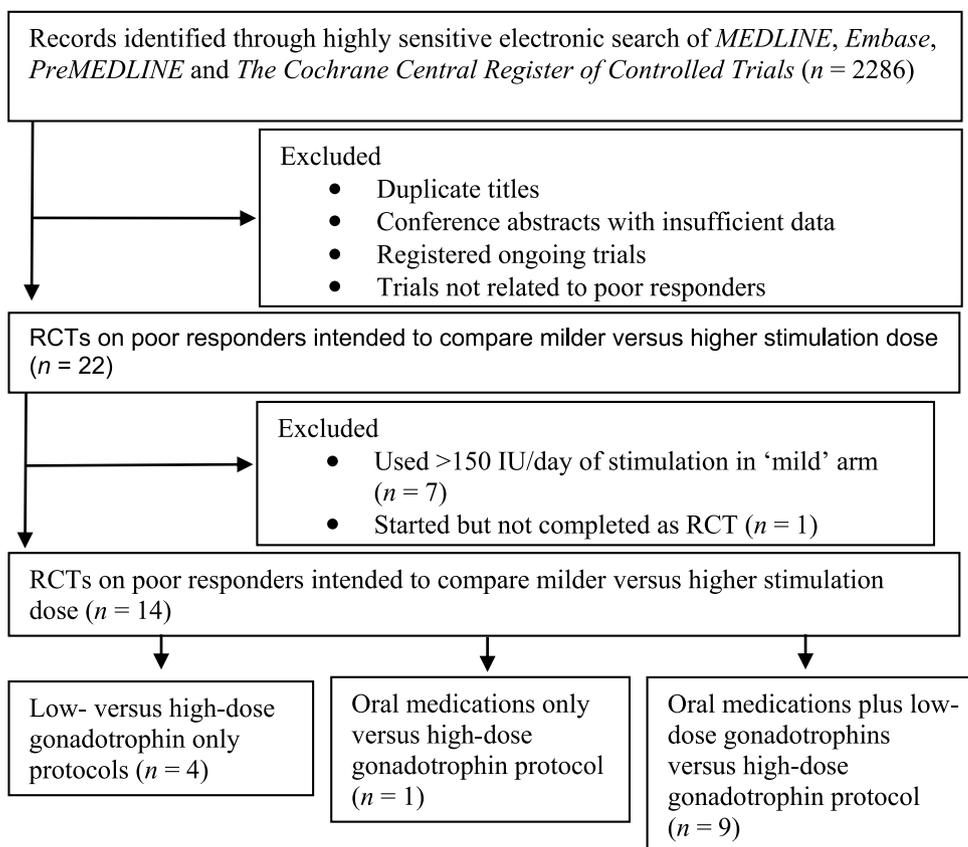


FIGURE 1 Study selection process. RCT, randomized controlled trial.

started as a RCT but was converted to a case-control study because of a problem in randomization and was therefore excluded (Siristatidis *et al.*, 2016); seven trials were excluded as >150 IU of gonadotrophin was combined with oral agents (clomiphene citrate or letrozole) in the mild or lower dose arm (Lee *et al.*, 2011; Ozcan Cenksoy *et al.*, 2014; Nabati *et al.*, 2015; Schimberni *et al.*, 2016; Selman and Rinaldi, 2016; Ebrahimi *et al.*, 2017; Davar *et al.*, 2018). Therefore, 14 RCTs were included for meta-analysis. Four trials exclusively compared low- and high-dose gonadotrophin protocols (Klinkert *et al.*, 2005; Kim *et al.*, 2009; van Tilborg *et al.*, 2017; Youssef *et al.*, 2017), whereas, two RCTs with a three-arm analysis had a low-dose gonadotrophin only arm as MS-IVF and the other MS-IVF arm combined with clomiphene citrate (Ashrafi *et al.*, 2005) or letrozole (Yu *et al.*, 2018). Among studies that compared oral agent-based mild stimulation and C-IVF protocols, three RCTs had clomiphene citrate low-dose gonadotrophin as MS-IVF (Martinez *et al.*, 2003; Revelli *et al.*, 2014; Pilehvari *et al.*, 2016) and four others used letrozole plus low-dose gonadotrophin in the mild-IVF arm

(Goswami *et al.*, 2004; Mohsen and El Din, 2013; Huang *et al.*, 2015; Bastu *et al.*, 2016). The study by Martinez *et al.* (2003) used a combination of clomiphene citrate and FSH (150 IU/day) in one arm and clomiphene citrate and human menopausal gonadotrophin (150 IU/day) in another arm as MS-IVF (the data from both the arms added together and the average included in the meta-analysis). One RCT used only clomiphene citrate, without gonadotrophin in the MS-IVF arm and compared it with a high gonadotrophin dose antagonist protocol (Ragni *et al.*, 2012).

Characteristics of included studies

The studies included in this review are presented in TABLE 1.

Trial design

Three included studies were multicentre trials (Ragni *et al.*, 2012; van Tilborg *et al.*, 2017; Youssef *et al.*, 2017); the rest were from single centres. Three trials conducted three-arm comparison (Ashrafi *et al.*, 2005; Bastu *et al.*, 2016; Yu *et al.*, 2018), one was a four-arm trial (Martinez *et al.*, 2003), and the remaining were two-arm studies. Power calculation was conducted in seven RCTs:

three of them were adequately powered for number of oocytes, which was their primary outcome (Klinkert *et al.*, 2005; Revelli *et al.*, 2014; Bastu *et al.*, 2016), one trial was powered for CPR (Yu *et al.*, 2018), one ongoing pregnancy rates (Youssef *et al.*, 2017), one for cumulative live birth (van Tilborg *et al.*, 2017) and the remaining one was underpowered for LBR (Ragni *et al.*, 2012).

Participants

Anticipated or proven poor responders of varied definitions were included in all trials; in three RCTs (Huang *et al.*, 2015; Bastu *et al.*, 2016; Pilehvari *et al.*, 2016), recruitment was as per the Bologna consensus on poor ovarian response. Age range of the participants, where available, is presented in TABLE 1.

Interventions

Interventions in each individual trial are detailed in TABLE 1. Briefly, six RCTs compared mild (150 IU) and high-dose (>150 IU) gonadotrophins only (without oral medication); three of them were on GnRH-ant protocols (Kim *et al.*, 2009; Youssef *et al.*, 2017; Yu *et al.*, 2018), two were on either agonist or both agonist and antagonist protocols (Klinkert *et al.*,

2005; van Tilborg *et al.*, 2017) and the study by Ashrafi *et al.* (2005) did not use any antagonist or agonist for LH suppression in the MS-IVF arms. Ten RCTs used oral agents in the MS-IVF arm either alone (clomiphene citrate) (Ragni *et al.*, 2012) or in combination with low-dose gonadotrophins; clomiphene citrate was used in four (Martinez *et al.*, 2003; Ashrafi *et al.*, 2005; Revelli *et al.*, 2014; Pilehvari *et al.*, 2016) and letrozole in five other trials (Goswami *et al.*, 2004; Mohsen and El Din, 2013; Huang *et al.*, 2015; Bastu *et al.*, 2016; Yu *et al.*, 2018). Consistently, clomiphene citrate was used at 100 mg daily dose for 5 days, commencing on cycle day 2 or 3 (from day 4 in one study); only Ragni *et al.* (2012) administered clomiphene citrate without gonadotrophin at 150 mg daily dose. The dose for letrozole was 5 mg daily, starting from day 2 or 3 for 5 days, except in the RCT by Goswami *et al.* (2004), in which 2.5 mg daily dose was used, and in the RCT by Huang *et al.* (2015), which did not mention the dose. In all trials, the starting dose of gonadotrophin for MS-IVF was 150 IU daily, except in two in which 75 IU daily dose was used (Goswami *et al.*, 2004; Yu *et al.*, 2018). The timing of commencement of gonadotrophin varied. The dose was fixed in three RCTs, dose adjustment was allowed in eight and not mentioned in three studies (TABLE 1). Pre-treatment was given in two RCTs: one with contraceptive pill (Youssef *et al.*, 2017) and the other with oral oestrogens (Mohsen and El Din, 2013). Cycle cancellation criteria varied between the studies (TABLE 1).

Outcomes measured

Live birth rates were compared in four RCTs: three of them, including the

third arm of the trial by Yu *et al.* (2018) compared low-dose and high-dose gonadotrophin (without oral agent) (Kim *et al.*, 2009; van Tilborg *et al.*, 2017; Yu *et al.*, 2018); the second arm of the study by Yu *et al.* (2018) used letrozole and the study by Ragni *et al.* (2012) used a clomiphene citrate only regimen and compared LBR with those of GnRH-a and GnRH-ant cycles, respectively. One adequately powered RCT compared cumulative live birth (van Tilborg *et al.*, 2017). Five included RCTs investigated OPR: two were adequately powered for this outcome (van Tilborg *et al.*, 2017; Youssef *et al.*, 2017); and another larger RCT was powered for number of oocytes (Revelli *et al.*, 2014). All included RCTs reported cancellation rates: whether the cycles were abandoned owing to poor ovarian response or premature ovulation was not often clear.

Among the secondary outcomes, CPRs were reported in all except one RCT by Ashrafi *et al.* (2005). All RCTs compared total dose of gonadotrophin, except Ragni *et al.* (2012) who used only clomiphene citrate as MS-IVF; another RCT expressed it in median (range) (Klinkert *et al.*, 2005). Mean or median number of oocytes retrieved was compared in all included RCTs. Nine trials reported the number of total embryos created: eight expressed the numbers in mean (SD) and one in median (range) (Klinkert *et al.*, 2005). Five RCTs compared the quantity of top-quality embryos (Kim *et al.*, 2009; Revelli *et al.*, 2014; Huang *et al.*, 2015; Youssef *et al.*, 2017; Yu *et al.*, 2018). The number of embryos transferred was not a true representation of the number of embryos; therefore, this outcome was not assessed in a meta-analysis. Two studies conducted

an economic evaluation (Ragni *et al.*, 2012; van Tilborg *et al.*, 2017).

Risk of bias in the included studies

Risk of bias is presented in FIGURE 2.

Selection bias

All RCTs were found to be ‘low-risk’ for random sequence generation, except two trials in which the risk was unclear (Ashrafi *et al.*, 2005; Pilehvari *et al.*, 2016). Allocation concealment was deemed to have low risk in all but three RCTs in which the risk was unclear (Martinez *et al.*, 2003; Pilehvari *et al.*, 2016; Yu *et al.*, 2018).

Performance and detection bias

All RCTs were considered to be of ‘low-risk’ for performance bias, as the blinding of both patients and assessors was neither possible nor required for these outcome measures. For attrition bias, the outcome data were not complete in one trial (high-risk) (Huang *et al.*, 2015), and not clear in the two other studies (Ashrafi *et al.*, 2005; Mohsen and El Din, 2013), the remaining were of ‘low-risk’. All RCTs had ‘low-risk’ for reporting bias. For other bias, baseline character of study groups were not clear in three RCTs (Ashrafi *et al.*, 2005; Huang *et al.*, 2015; Bastu *et al.*, 2016).

Primary outcomes

Live birth rates

Four RCTs (n = 1057) compared LBR: two studies compared low- and high-dose gonadotrophins (Kim *et al.*, 2009; van Tilborg *et al.*, 2017), one compared only clomiphene citrate with high-stimulation antagonist protocol (Ragni *et al.*, 2012) and the other three-arm study compared both (Yu *et al.*, 2018). One RCT achieved

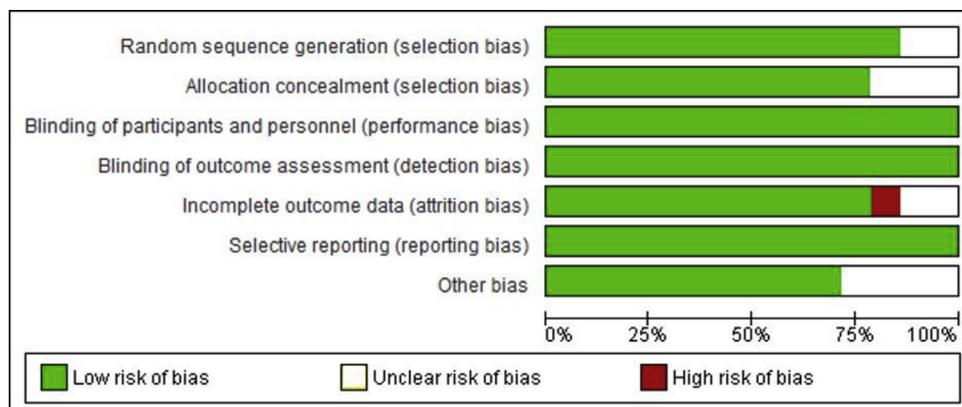


FIGURE 2 Risk of bias graph.

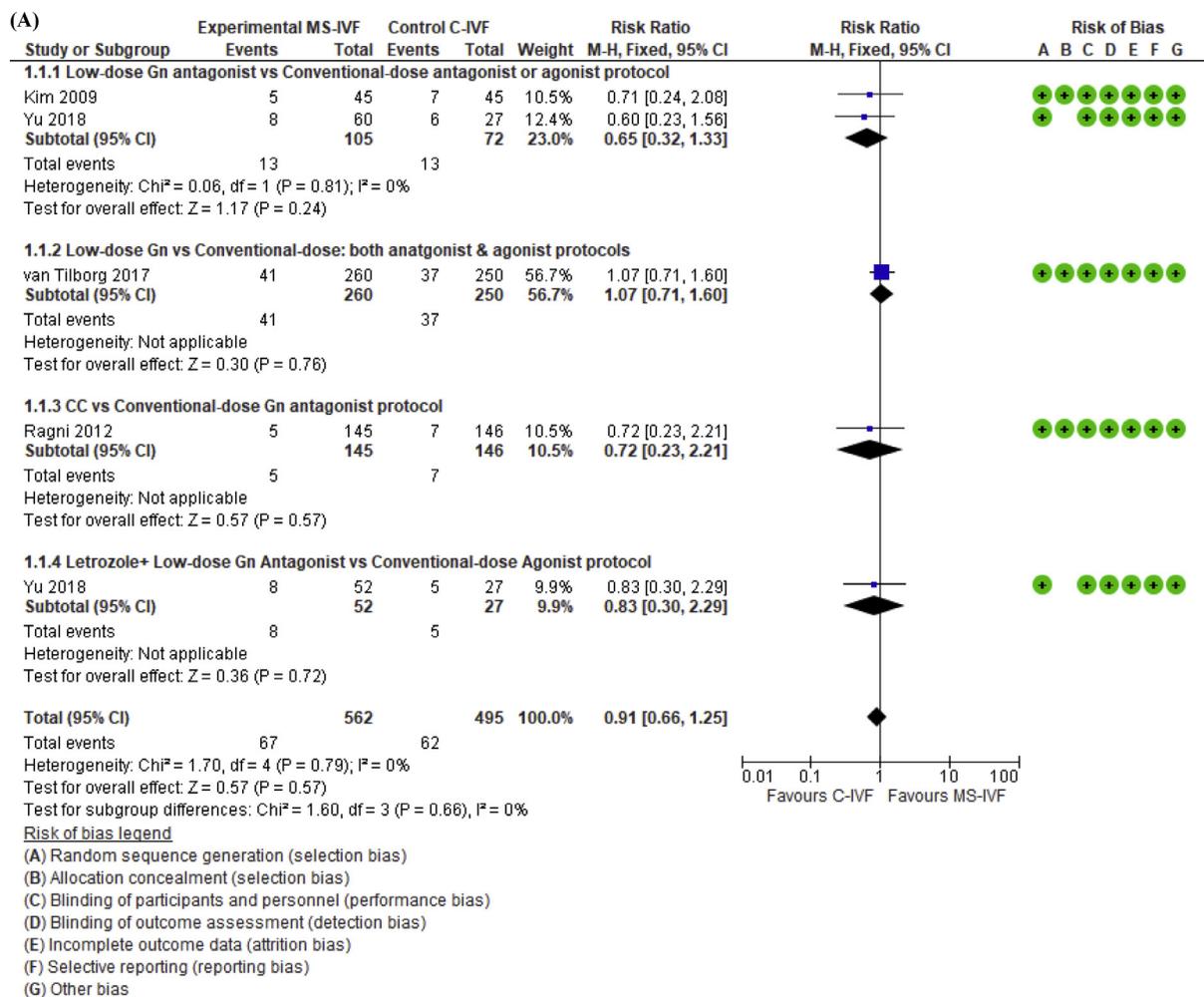


FIGURE 3A Live birth rates per patient randomized: mild versus conventional stimulation IVF. CC, clomiphene citrate; C-IVF, comparator group; Gn, gonadotrophin; MS-IVF, mild ovarian stimulation for IVF.

predefined sample size for cumulative live birth (*van Tilborg et al., 2017*), another stopped recruiting before the number of participants to give adequate power to live birth was reached (*Ragni et al., 2012*). The proportions of LBR

were higher with C-IVF in most of the individual studies, but the differences were not statistically significant; meta-analysis of the pooled data also found no evidence of a difference in LBRs (RR 0.91, CI 0.66 to 1.25) (**FIGURE 3A**). There was

no statistical heterogeneity, three trials were of low ROB (*Kim et al., 2009; Ragni et al., 2012; van Tilborg et al., 2017*) and the total pooled population was 1057, with fairly narrow confidence intervals (**TABLE 2** and **FIGURE 3A**). The finding

TABLE 2 EVIDENCE

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Number of Participants per cycle (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Conventional IVF	Mild stimulation IVF			
Live birth per women randomized	125 per 1000	114 per 1000 82 to 156	RR 0.91 (CI 0.66 to 1.25)	1057 (four RCTs)	⊕⊕⊕⊕ ^a Moderate
Ongoing pregnancy per women randomized	212 per 1000	214 per 1000 182 to 254	RR 1.01 (0.86 to 1.20)	1782 (six RCTs)	⊕⊕⊕⊕ ^b Moderate ^e
Cycle cancellation per started cycle	120 per 1000	166 per 1000 119 to 230	RR 1.38 (0.99 to 1.92)	2746 (14 RCTs)	⊕⊕⊕⊕ ^{c,d} Low

^a One step down owing to one study with an area of unclear risk of bias plus clinical heterogeneity (inference remains unchanged if studies with unclear risk of bias are excluded).

^b One step down owing to studies with an area of unclear risk of bias and clinical heterogeneity.

^c One step down owing to significant statistical heterogeneity plus clinical heterogeneity.

^d Another step down owing to multiple studies with unclear risk of bias and a study with high risk of bias.

^e Moderate to high quality of evidence in the subgroup comparing low versus high-dose gonadotrophin (without oral agent).

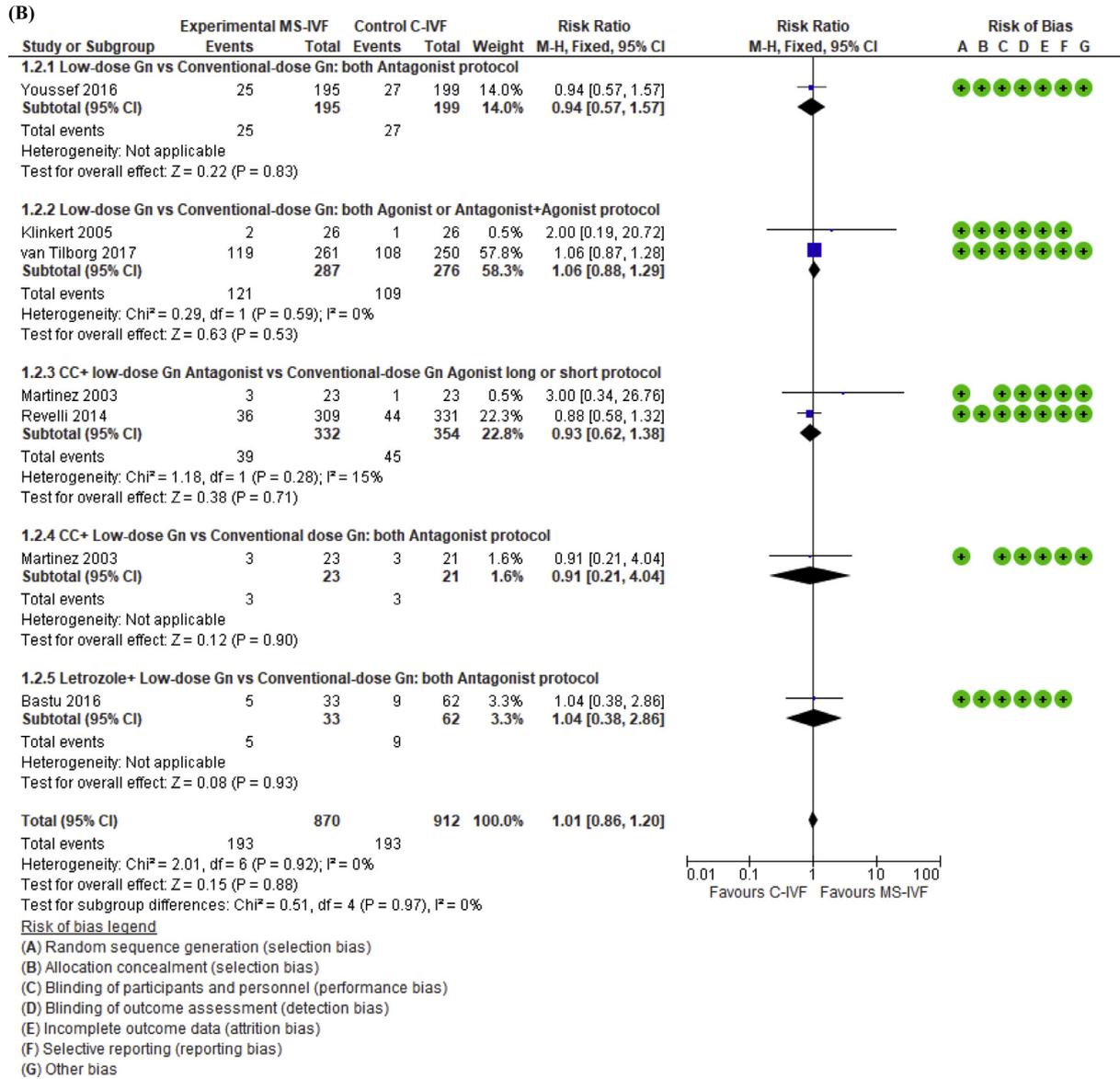


FIGURE 3B Ongoing pregnancy rates per patient randomized: mild versus conventional IVF; CC, clomiphene citrate; C-IVF, comparator group; Gn, gonadotrophin; MS-IVF, mild ovarian stimulation for IVF.

remained unchanged when the smaller RCT or the RCT with ROB were excluded (Kim et al., 2009; Yu et al., 2018). The quality of evidence was moderate owing to the presence of significant clinical heterogeneity.

Ongoing pregnancy rates

Six RCTs (n = 1782) 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 reported clomiphene citrate and gonadotrophin with agonist ‘long’ (Revelli et al., 2014) or ‘short’ protocol (Martinez et al., 2003) and one RCT compared letrozole plus gonadotrophin with high-dose antagonist protocol (Bastu et al., 2016). All individual

RCTs and meta-analyses of pooled data found no difference in OPRs (RR 1.01, CI 0.86 to 1.20) (TABLE 2 and FIGURE 3B). No statistical heterogeneity was found between the studies or subgroups, the confidence interval was narrow, with two large trials adequately powered for OPR (Youssef et al., 2017) and LBR (van Tilborg et al., 2017); these two RCTs and another large trial (Revelli et al., 2014) were of low ROB (FIGURE 3B). The overall quality of evidence was moderate, however, owing to two other small RCTs with an area of ‘unclear ROB’ (Martinez et al., 2003; Bastu et al., 2016) in the subgroups with oral agents and the presence of clinical heterogeneity among the study protocols. If small studies with unclear ROB (Martinez et al., 2003;

Klinkert et al., 2005; Bastu et al., 2016) were excluded, the inference of the meta-analysis remained unchanged. The quality of evidence seemed to be high in the subgroup comparing low-dose with high-dose gonadotrophin protocols (three RCTs, n = 957), and the confidence interval remained narrow (RR 1.04 CI 0.87 to 1.25); both large RCTs were adequately powered for OPR and were of low ROB (van Tilborg et al., 2017; Youssef et al., 2017).

Cycle cancellation rates

All 14 RCTs (n = 2746) investigated cancellation rates. The risk of cycle cancellation was comparable with a RR of 1.38 (CI 0.99 to 1.92) (TABLE 2 and FIGURE 3C). Although total population

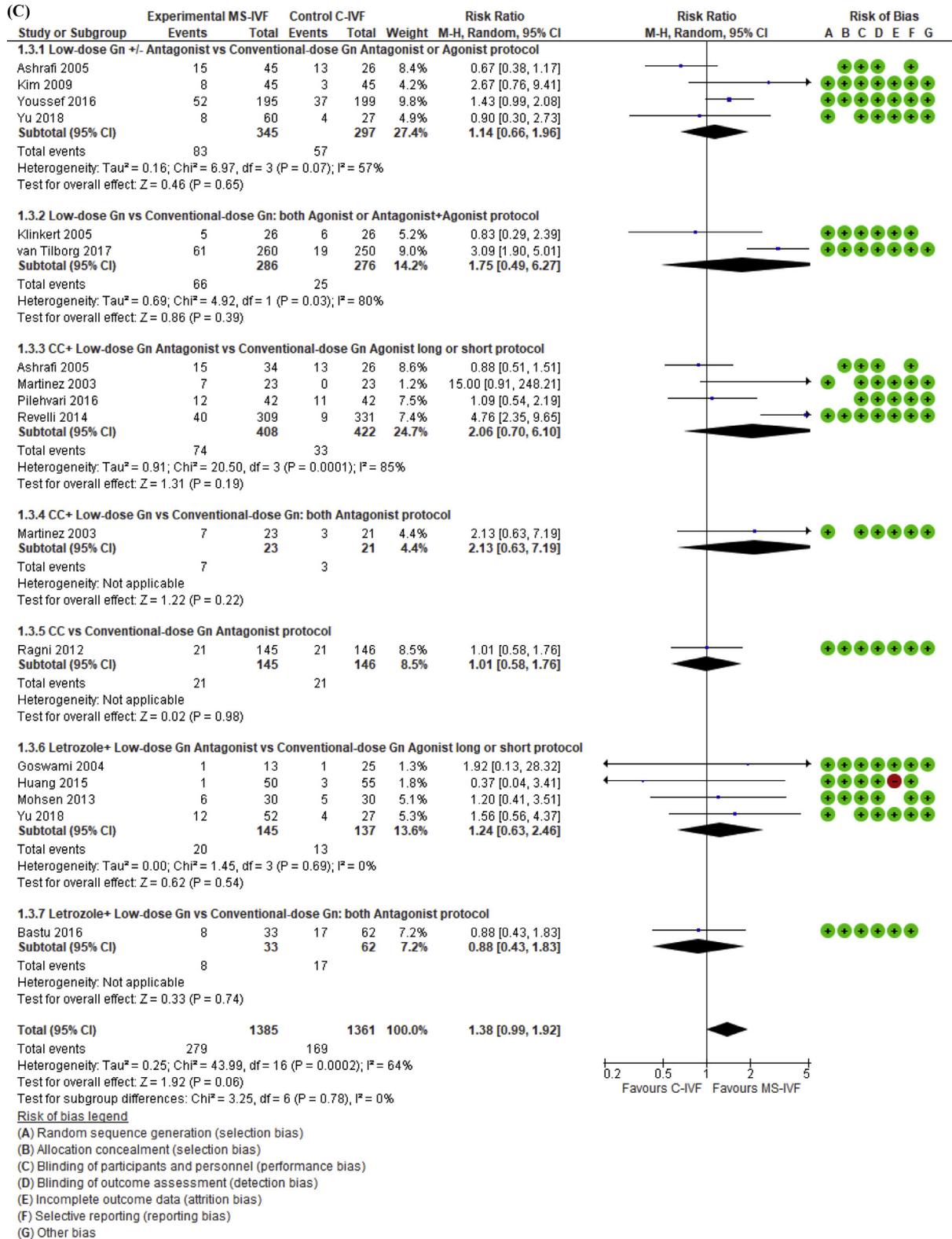


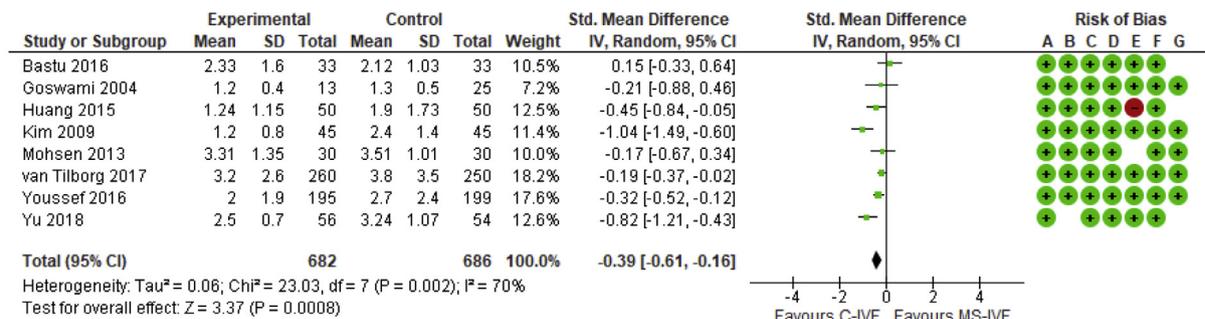
FIGURE 3C Cycle cancellation rates per cycle started: mild versus conventional IVF. CC, clomiphene citrate; C-IVF, comparator group; Gn, gonadotrophin; MS-IVF, mild ovarian stimulation for IVF.

was large, significant statistical as well as clinical heterogeneity was present, in addition to unclear ROB in one or more areas in most trials plus an area

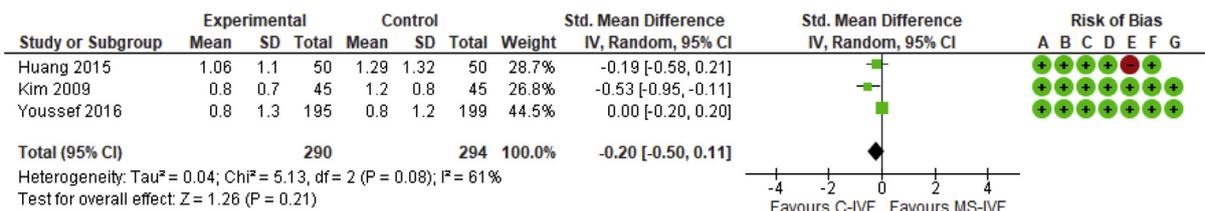
of high ROB in one trial; confidence intervals were moderately wide (FIGURE 3C). Moreover, no consistent cycle cancellation criteria among the

trials (some study even without a stated cancellation policy) led to a very low to low quality of evidence for this outcome.

(D) (i). Number of embryos created (mean). C-IVF, comparator group; MS-IVF, mild ovarian stimulation for IVF.



(ii): Number of high-grade embryos created. C-IVF, comparator group; MS-IVF, mild ovarian stimulation for IVF.



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

FIGURE 3D (i) Number of embryos created (mean). C-IVF, comparator group; MS-IVF, mild ovarian stimulation for IVF. (ii) Number of high-grade embryos created. C-IVF, comparator group; MS-IVF, mild ovarian stimulation for IVF.

Secondary outcomes

Clinical pregnancy rates

Twelve RCTs reported CPR as outcome. No significant difference was found in CPR between MS-IVF and C-IVF (RR 0.95, 95% CI 0.78 to 1.16). Although the confidence interval was narrow, with a total study population of 2097 and no statistical heterogeneity was observed, multiple studies had ‘unclear ROB’ and one study had an area of high ROB. Clinical heterogeneity remained. Consequently, the quality of evidence seemed to be moderate.

Total gonadotrophin dose used

All included RCTs, except the study by Ragni et al. (2012), which did not use any gonadotrophin in the MS-IVF protocol, compared total amount of gonadotrophin used between the groups. One study expressed the dose in median (range) (Klinkert et al., 2005); all others were expressed in mean, hence included in the meta-analysis. A high level of clinical and statistical heterogeneity (96%) was observed among the studies, and many RCTs had area(s) of ‘unclear bias’. All

individual trials, however, found lower gonadotrophin use in MS-IVF programme with a large effect-size (SMD 3.17, CI -3.80 to -2.54), which increased the reliability of this finding.

Total number of oocytes retrieved

All RCTs compared the difference in the number of oocytes retrieved. One of them expressed the data in median (range) and found no difference (mild: 3 [1–9] versus high-dose: 3 [1–6]; P = 0.79) (Klinkert et al., 2005). Meta-analysis of the remaining 13 RCTs that measured mean number demonstrated significantly lower number of oocytes recovered in the MS-VF group, with a SMD of -0.44 (CI -0.60 to -0.28). Most RCTs had area(s) of ‘unclear ROB’ and one had an area of ‘high ROB’. Statistical heterogeneity (71%) was significant and clinical heterogeneity observed. These factors contributed to the evidence being low quality.

Embryos created: total and high-grade

Nine RCTs compared total number of embryos created (studies that

mentioned the mean number of embryo transferred was not a reflection of total embryo created, and was therefore excluded from analysis) (Martinez et al., 2003; Ragni et al., 2012; Revelli et al., 2014; Pilehvari et al., 2016). A meta-analysis of eight trials that expressed the number of embryos as mean (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., 2017; Yu et al., 2018) found fewer mean of total embryos with MS-IVF compared with C-IVF (SMD -0.39, CI -0.61 to -0.16); the other trial expressed the number in median (range) (median 2 [0–6] in both mild and high dose arm; P = 0.86) (Klinkert et al., 2005), and therefore, remained outside the meta-analysis. Two adequately powered RCTs that contributed to a narrow confidence interval had no ROB (van Tilborg et al., 2017; Youssef et al., 2017); however, the evidence was of low quality owing to multiple studies with ‘unclear’ and one with high ROB, along with significant statistical and clinical heterogeneity (FIGURE 3D i). Exclusion of

small studies with ROB did not change the finding.

Six RCTs examined 'top- or high-grade' embryos between MS-IVF and C-IVF: three compared the mean number (Kim *et al.*, 2009; Huang *et al.*, 2015; Youssef *et al.*, 2017); one of them was large with low ROB (Youssef *et al.*, 2017). Meta-analysis of these three trials showed no difference (SMD -0.50, CI -0.34 to 0.11) (FIGURE 3D II); Three other studies comparing the proportion (%) of good-quality embryos (Revelli *et al.*, 2014; Pilehvari *et al.*, 2016; Yu *et al.*, 2018) individually did not find a difference between the two groups. A meta-analysis of the pooled data from these trials was not possible owing to unavailability of denominators. A large RCT ($n = 640$) with low ROB reported the proportion of embryo scoring greater than 8 points to be 57.6% with MS-IVF and 54.8% with C-IVF; the difference was not significant (Revelli *et al.*, 2014). Overall, clinical and statistical heterogeneity was significant; three studies had multiple areas of unclear bias, one had a high ROB (Huang *et al.*, 2015), hence the quality of evidence was low.

Treatment cost

Two of the included RCTs compared the cost of treatment: both were large trials with low ROB. One found MS-IVF was associated with a per-cycle cost-saving of €2620 (Ragni *et al.*, 2012), with no use of gonadotrophin in the MS-IVF arm and 450 IU of daily dose in the C-IVF arm. The other large RCT reported a reduced cumulative treatment cost with the lower gonadotrophin dose regimen by €1099 (van Tilborg *et al.*, 2017). Cost in each arm was not mentioned in this study; therefore, a meta-analysis could not be conducted. The study protocols were different in these two trials, resulting in significant clinical heterogeneity.

Discussion

Meta-analyses of pregnancy outcome data found no difference in live birth or ongoing pregnancy rates between mild stimulation and conventional higher dose stimulation used to treat the poor responders. The only RCT that investigated cumulative LBR did not find any difference (van Tilborg *et al.*, 2017) in this parameter. The risk of cycle cancellation and the chance of obtaining high-grade embryos were also comparable. Limited data showed that

MS-IVF was associated with reduced use of gonadotrophins as well as treatment cost. The findings of our meta-analyses remained unchanged on sensitivity analysis or in sub-group analysis separating gonadotrophin only regimen from that with oral medications.

The level of evidence on GRADE scoring ranged from low to high quality: moderate to high quality for OPR; and low quality for the risk of cycle cancellation, number of oocytes or embryos obtained. Two large RCTs with low ROB found no difference in the quantity or proportion of high-grade embryos (Revelli *et al.*, 2014; Youssef *et al.*, 2017). Albeit limited data, this could explain why the pregnancy outcome after mild ovarian stimulation remained as effective as that of high-dose stimulation, despite fewer oocytes retrieved and fewer embryos created, i.e. oocyte and embryo quality was not affected.

To the best of our knowledge, the present systematic review and meta-analysis is the first to include RCTs that used genuine mild stimulation regimen, i.e. a low-dose of gonadotrophin with or without oral medication) and compared with conventional high-dose protocols for poor responders. Among the systematic reviews related to this topic, only the American Society for Reproductive Medicine Practice Committee Guideline considered mild stimulation like ours to be 150 IU or less of gonadotrophin daily as 'mild ovarian stimulation'; however, this was a review without a meta-analysis (Practice Committee of the American Society for Reproductive Medicine. *Electronic address*, 2018). The present review is the only one to compare the number of high-grade embryos between the two approaches. We have taken cycle cancellation risk as one of our primary foci of attention, considering its clinical importance, especially treatment burden and emotional impact. Finally, the pooled live birth data of this up-to-date review are derived from a total population of 1057 women; this gives adequate statistical power to determine the difference in LBR. A sample size calculation aiming to give 80% power within 5% type 1 error for 5% difference in the LBRs (expecting between 5% and 10% LBRs among the poor responders) suggests 475 participants in each arm (StatsDirect version 2.8.0, 27th October 2013).

Lack of consensus on the definition of poor responders of IVF has always been an obstacle in uniform recruitment of participants (Polyzos and Devroey, 2011); many of the included RCTs in this review were conducted before the introduction of Bologna criteria. Other significant clinical heterogeneity includes variations in the study protocols and dose adjustments in many studies; however, the variations have increased generalizability of the findings. Difference in the cycle cancellation criteria significantly reduces the reliability of this important outcome. The denominator being per cycle started, the effect of cycle cancellation has been taken into account while analysing the pregnancy outcome; however, the difference in the cycle cancellation policies adds heterogeneity to the pooled evidence on this outcome. Economic evaluation is based on only few available studies. We have compared low versus high stimulation dose irrespective of whether GnRH-agonist or antagonist was used, in the absence of any evidence of difference between the two on pregnancy outcome in conventional-IVF programme protocols (Lambalk *et al.*, 2017); inference of our review remains unchanged when sensitivity analysis was performed based on protocols.

Four systematic reviews on poor responders have previously compared oral compounds plus gonadotrophin (any dose) and conventional IVF: three of them presented the outcomes of poor responders as a separate subgroup analysis (Bechtejew *et al.*, 2017; Fan *et al.*, 2017; Kamath *et al.*, 2017); the other review dealt with only poor responders (Song *et al.*, 2016). The pregnancy outcomes of our review are similar to all other published reviews, and also in agreement with the only systematic review with meta-analysis that compared low and high gonadotrophin-only regimens (Youssef *et al.*, 2018) in low responders. Cycle cancellation outcome was investigated in four reviews; no difference in CCR was found in three (Song *et al.*, 2016; Bechtejew *et al.*, 2017; Fan *et al.*, 2017); however, they did not judge the quality of experience. In one review (Kamath *et al.*, 2017), CCR was significantly higher with oral clomiphene citrate or letrozole (low quality of evidence). The conclusion of our review is the same as other reviews in terms of the effect of gonadotrophin dose. No difference in the number of retrieved

oocytes was found in some reviews (Song *et al.*, 2016; Bechtejew *et al.*, 2017; Fan *et al.*, 2017), whereas contradictory findings between the subgroups were found.

Our review supports the recent American Society for Reproductive Medicine recommendation (*Practice Committee of the American Society for Reproductive Medicine. Electronic address*, 2018) that mild ovarian stimulation should be considered for IVF treatment in poor responders. With a moderate to high quality of evidence on comparable pregnancy outcomes from an adequately powered aggregated population, along with accumulating evidence of cost-saving, we may more confidently consider mild ovarian stimulation for poor responder women in preference to conventional high-stimulation regimen, which has been reported to cause more patient discomfort (Hojgaard *et al.*, 2001; de Klerk *et al.*, 2007) and incur incremental treatment cost. No difference in the incidence of cycle cancellation also seems reassuring.

Although it is hard to completely eliminate clinical heterogeneity related to the study protocol, adherence to common consensus, e.g. Bologna criteria, for selecting poor responders, fixed-dose regimen in either arm and a universally agreed cycle cancellation criterion need to be considered in future randomized trials. More evidence on cumulative live birth, comparison of patient's tolerance (dropout rates) and treatment cost would aid in complete evaluation of mild stimulation IVF for poor responders. Research is also needed on patients' preferences for use of mild versus conventional stimulation, given the risks and benefits.

In conclusion, our review indicates that mild ovarian stimulation for IVF could be as effective as conventional high-dose stimulation for poor responders with a moderate to high quality evidence on pregnancy outcomes. Limited RCT evidence on cumulative live birth also infers the same. Despite fewer oocytes and embryos obtained after mild stimulation, no difference in the incidence of cycle cancellation or in the quantity of high-grade embryos resulted in comparable pregnancy outcome. Clinical effectiveness being similar, mild

stimulation seems to be associated with reduced use of gonadotrophins as well as cost (limited data). The present systematic review gives an insight into how future RCTs could be designed to obtain the best possible evidence on mild stimulation IVF.

ACKNOWLEDGEMENT

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

REFERENCES

- Ashrafi, M., Ashtiani, S.K., Zafarani, F., Samani, R.O., Eshtrati, B. **Evaluation of ovulation induction protocols for poor responders undergoing assisted reproduction techniques.** Saudi Medical Journal 2005; 26: 593–596
- 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.** European Journal of Obstetrics Gynecology and Reproductive Biology 2016; 203: 30–34
- Bechtejew, T.N., Nadai, M.N., Nastri, C.O., Martins, W.P. **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
- Davar, R., Neghab, N., Naghshineh, E. **Pregnancy outcome in delayed start antagonist versus microdose flare GnRH agonist protocol in poor responders undergoing IVF/ICSI: An RCT.** Int. J. Reprod. Biomed. (Yazd) 2018; 16: 255–260
- de Klerk, C., Macklon, N.S., Heijnen, E.M., Eijkemans, M.J., Fauser, B.C., Passchier, J., Hunfeld, J.A. **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
- Ebrahimi, M., Akbari-Asbagh, F., Ghalandarpoor, S.M. **Letrozole+ GnRH antagonist stimulation protocol in poor ovarian responders undergoing intracytoplasmic sperm injection cycles: An RCT.** International Journal of Reproductive BioMedicine 2017; 15: 101–108
- 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, A.P., La Marca, A., Fauser, B.C., Tarlatzis, B., Nargund, G., Gianaroli, L., Definition, E.w.g.o.P.O.R. **ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria.** Hum. Reprod. 2011; 26: 1616–1624
- Goswami, S.K., Das, T., Chattopadhyay, R., Sawhney, V., Kumar, J., Chaudhury, K., Chakravarty, B.N., Kabir, S.N. **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
- Heijnen, E.M., Eijkemans, M.J., De Klerk, C., Beckers, N.G., Klinkert, E.R., Broekmans, F.J., Passchier, J., Te Velde, E.R., Habbema, J.D., Macklon, N.S., et al., 2005. Mild stimulation and single embryo transfer: towards safer and simpler IVF. The 21st annual meeting of the european society of human reproduction and embryology:i141.
- Higgins, J.P.T., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 ed. The Cochrane Collaboration.
- Hojgaard, A., Ingerslev, H.J., Dinesen, J. **Friendly IVF: patient opinions.** Hum. Reprod. 2001; 16: 1391–1396

- Huang, R., Wang, B., Yang, X., Li, T.T., Liang, X.Y. **The comparison of mild stimulation vs. Controlled ovarian hyperstimulation protocol in poor ovarian responders: A prospective randomized study.** *Human Reproduction* 2015; 1: i49–i50
- Kamath, M.S., Maheshwari, A., Bhattacharya, S., Lor, K.Y., Gibreel, 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
- Kim, C.H., Kim, S.R., Cheon, Y.P., Kim, S.H., Chae, H.D., Kang, B.M. **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.** *Fertility and Sterility* 2009; 92: 2082–2084
- Klinkert, E.R., Broekmans, F.J., Looman, C.W., Habbema, J.D., te Velde, E.R. **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
- Lambalk, C.B., Banga, F.R., Huirne, J.A., Toftager, M., Pinborg, A., Homburg, R., van der Veen, F., van Wely, M. **GnRH antagonist versus long agonist protocols in IVF: a systematic review and meta-analysis accounting for patient type.** *Hum. Reprod. Update* 2017; 23: 560–579
- Lee, V.C.Y., Chan, C.C.W., Ng, E.H.Y., Yeung, W.S.B., Ho, P.C. **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
- Martinez, F., Coroleu, B., Marques, L., Parera, N., Buxaderas, R., Tur, R., Barri, P.N. **Comparison of short protocol versus antagonists with or without clomiphene citrate for stimulation in IVF of patients with low response.** [Spanish]. *Revista Iberoamericana de Fertilidad y Reproduccion Humana* 2003; 20: 355–360
- Mohsen, I.A., El Din, R.E. **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
- 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.** *Global journal of health science* 2015; 8: 166–171
- Nargund, G., Datta, A.K., Fauser, B. **Mild stimulation for in vitro fertilization.** *Fertil. Steril.* 2017; 108: 558–567
- Orvieto, R., Vanni, V.S., Gleicher, N. **The myths surrounding mild stimulation in vitro fertilization (IVF).** *Reprod. Biol. Endocrinol.* 2017; 15: 48
- Ozcan Cenksoy, P., Ficioglu, 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 Tehrani-nejad, E., Hosseini-rashidi, B., Keikhah, F., Haghollahi, F., Azimineko, E. **Comparison pregnancy outcomes between minimal stimulation protocol and conventional GnRH antagonist protocols in poor ovarian responders.** *Journal of Family and Reproductive Health* 2016; 10: 35–42
- Practice Committee of the American Society for Reproductive Medicine. Electronic address, A.a.o. **Comparison of pregnancy rates for poor responders using IVF with mild ovarian stimulation versus conventional IVF: a guideline.** *Fertil. Steril.* 2018; 109: 993–999
- Ragni, G., Levi-Setti, P.E., 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.** *Reproductive biology and endocrinology: RB&E* 2012; 10: 114
- Revelli, A., Chiado, A., Dalmaso, P., Stabile, V., Evangelista, F., Basso, G., 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, S.C., 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: 4354–4361
- Selman, H., Rinaldi, L. **Effectiveness of corifollitropin alfa used for ovarian stimulation of poor responder patients.** *International journal of women's health* 2016; 8: 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: 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
- van Tilborg, T.C., Torrance, H.L., Oudshoorn, S.C., Eijkemans, M.J.C., Koks, C.A.M., Verhoeve, H.R., Nap, A.W., Scheffer, G.J., Manger, A.P., Schoot, B.C. **Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 1: the predicted poor responder.** *Human reproduction (oxford, england)* 2017; 32: 2496–2505
- Youssef, M.A., van Wely, M., Al-Inany, H., Madani, T., Jahangiri, N., Khodabakhshi, S., Alhalabi, M., Akhondi, M., Ansari-pour, S., Tokhmechy, R., Zarandi, L., Rizk, A., El-Mohamedy, M., Shaeer, E., Khattab, M., Mochtar, M.H., van der Veen, F. **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, M.A., van Wely, M., Mochtar, M., Fouada, U.M., Eldaly, A., El Abidin, E.Z., Elhalwagy, A., Mageed Abdallah, A.A., Zaki, S.S., Abdel Ghafar, M.S., Mohesen, M.N., van der Veen, F. **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, G.D., de Mouzon, J., Ishihara, O., Mansour, R., Nygren, K., Sullivan, E., van der Poel, S., International Committee for Monitoring Assisted Reproductive, T., World Health, O. **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

Received 13 November 2019; received in revised form 4 February 2020; accepted 3 March 2020.