









# Ovarian stimulation strategies for intrauterine insemination in couples with unexplained infertility: a systematic review and individual participant data meta-analysis

J.A. Wessel <sup>1</sup>, N.A. Danhof<sup>1</sup>, R. van Eekelen <sup>1</sup>, M.P. Diamond <sup>2</sup>,  
R.S. Legro <sup>3</sup>, K. Peeraer <sup>4</sup>, T.M. D'Hooghe <sup>5,6,7</sup>, M. Erdem<sup>8</sup>,  
T. Dankert<sup>9</sup>, B.J. Cohlen <sup>10</sup>, C. Thyagaraju<sup>11</sup>, B.W.J. Mol <sup>12,13</sup>,  
M. Showell <sup>14</sup>, M. van Wely <sup>1</sup>, M.H. Mochtar<sup>1</sup>, and R. Wang <sup>12,\*</sup>

<sup>1</sup>Amsterdam UMC location University of Amsterdam, Centre for Reproductive Medicine, Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands <sup>2</sup>Department of Obstetrics and Gynecology, Augusta University, Augusta, GA 30912, USA <sup>3</sup>Department of Obstetrics and Gynecology, Penn State College of Medicine, Hershey, PA 17033, USA <sup>4</sup>UZ Leuven, Leuven University Fertility Center, Leuven 3000, Belgium <sup>5</sup>Merck Healthcare KGaA, Darmstadt 64293, Germany <sup>6</sup>Department of Development and Regeneration, Group Biomedical Sciences, KU Leuven/University of Leuven, Leuven 3000, Belgium <sup>7</sup>Department of Obstetrics and Gynecology, Yale University, New Haven, CT 06520, USA <sup>8</sup>Faculty of Medicine, Department of Obstetrics & Gynecology, Gazi University, Ankara 06560, Turkey <sup>9</sup>Department of Obstetrics and Gynecology, Rijnstate Hospital Arnhem, 06560 Ankara, The Netherlands <sup>10</sup>Department of Obstetrics and Gynaecology, Isala Fertility Center, 8025 AB Zwolle, The Netherlands <sup>11</sup>Department of OBG, Jawaharlal Institute of Postgraduate Medical education and Research (JIPMER), Pondicherry 605006, India <sup>12</sup>Department of Obstetrics and Gynaecology, Monash University, Clayton, VIC 3168, Australia <sup>13</sup>Aberdeen Centre for Women's Health Research, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen AB24 3FX, UK <sup>14</sup>Department of Obstetrics and Gynaecology, University of Auckland, Auckland 1142, New Zealand

\*Correspondence address. Department of Obstetrics and Gynaecology, Monash University, Level 5, Monash Medical Centre, 246 Clayton Road, Clayton, VIC 3168, Australia. E-mail: [r.wang@monash.edu](mailto:r.wang@monash.edu)  <https://orcid.org/0000-0002-6622-8134>

Submitted on October 19, 2021; resubmitted on March 23, 2022; editorial decision on April 12, 2022

## TABLE OF CONTENTS

- Introduction
- Methods
  - Registration and literature search
  - Eligibility criteria
  - Study selection and data collection
  - Outcomes
  - Risk of bias and overall certainty of evidence assessment
  - Statistical analysis
- Results
  - Study selection
  - Study characteristics
  - Risk of bias of individual RCTs
  - Primary outcomes
  - Secondary outcomes per woman

- Secondary outcomes per cycle
- Exploratory treatment–covariate interactions
- Additional analysis
- Discussion
  - Summary of evidence
  - Strengths and limitations
  - Interpretations
- Conclusion

**BACKGROUND:** Intrauterine insemination with ovarian stimulation (IUI-OS) is a first-line treatment for unexplained infertility. Gonadotrophins, letrozole and clomiphene citrate (CC) are commonly used agents during IUI-OS and have been compared in multiple aggregate data meta-analyses, with substantial heterogeneity and no analysis on time-to-event outcomes. Individual participant data meta-analysis (IPD-MA) is considered the gold standard for evidence synthesis as it can offset inadequate reporting of individual studies by obtaining the IPD, and allows analyses on treatment–covariate interactions to identify couples who benefit most from a particular treatment.

**OBJECTIVE AND RATIONALE:** We performed this IPD-MA to compare the effectiveness and safety of ovarian stimulation with gonadotrophins, letrozole and CC and to explore treatment–covariate interactions for important baseline characteristics in couples undergoing IUI.

**SEARCH METHODS:** We searched electronic databases including MEDLINE, EMBASE, CENTRAL, CINAHL, and PsycINFO from their inception to 28 June 2021. We included randomized controlled trials (RCTs) comparing IUI-OS with gonadotrophins, letrozole and CC among couples with unexplained infertility. We contacted the authors of eligible RCTs to share the IPD and established the IUI IPD-MA Collaboration. The primary effectiveness outcome was live birth and the primary safety outcome was multiple pregnancy. Secondary outcomes were other reproductive outcomes, including time to conception leading to live birth. We performed a one-stage random effects IPD-MA.

**OUTCOMES:** Seven of 22 (31.8%) eligible RCTs provided IPD of 2495 couples (62.4% of the 3997 couples participating in 22 RCTs), of which 2411 had unexplained infertility and were included in this IPD-MA. Six RCTs ( $n = 1511$ ) compared gonadotrophins with CC, and one ( $n = 900$ ) compared gonadotrophins, letrozole and CC. Moderate-certainty evidence showed that gonadotrophins increased the live birth rate compared to CC (6 RCTs, 2058 women, RR 1.30, 95% CI 1.12–1.51,  $I^2 = 26%$ ). Low-certainty evidence showed that gonadotrophins may also increase the multiple pregnancy rate compared to CC (6 RCTs, 2058 women, RR 2.17, 95% CI 1.33–3.54,  $I^2 = 69%$ ). Heterogeneity on multiple pregnancy could be explained by differences in gonadotrophin starting dose and choice of cancellation criteria. Post-hoc sensitivity analysis on RCTs with a low starting dose of gonadotrophins ( $\leq 75$  IU) confirmed increased live birth rates compared to CC (5 RCTs, 1457 women, RR 1.26, 95% CI 1.05–1.51), but analysis on only RCTs with stricter cancellation criteria showed inconclusive evidence on live birth (4 RCTs, 1238 women, RR 1.15, 95% CI 0.94–1.41). For multiple pregnancy, both sensitivity analyses showed inconclusive findings between gonadotrophins and CC (RR 0.94, 95% CI 0.45–1.96; RR 0.81, 95% CI 0.32–2.03, respectively). Moderate certainty evidence showed that gonadotrophins reduced the time to conception leading to a live birth when compared to CC (6 RCTs, 2058 women, HR 1.37, 95% CI 1.15–1.63,  $I^2 = 22%$ ). No strong evidence on the treatment–covariate (female age, BMI or primary versus secondary infertility) interactions was found.

**WIDER IMPLICATIONS:** In couples with unexplained infertility undergoing IUI-OS, gonadotrophins increased the chance of a live birth and reduced the time to conception compared to CC, at the cost of a higher multiple pregnancy rate, when not differentiating strategies on cancellation criteria or the starting dose. The treatment effects did not seem to differ in women of different age, BMI or primary versus secondary infertility. In a modern practice where a lower starting dose and stricter cancellation criteria are in place, effectiveness and safety of different agents seem both acceptable, and therefore intervention availability, cost and patients' preferences should factor in the clinical decision-making. As the evidence for comparisons to letrozole is based on one RCT providing IPD, further RCTs comparing letrozole and other interventions for unexplained infertility are needed.

**Key words:** unexplained infertility / intrauterine insemination / individual participant data / meta-analysis / ovarian stimulation / gonadotrophins / letrozole / clomiphene citrate

## Introduction

Intrauterine insemination with ovarian stimulation (IUI-OS) is a first-line treatment for couples with unexplained infertility (Practice Committee of the American Society for Reproductive Medicine, 2020). It aims to increase the pregnancy rates by increasing the number of dominant follicles per cycle, which is achieved by increasing the serum levels of FSH (van Rumste et al., 2008). Agents which increase FSH serum levels include exogenous gonadotrophins,

letrozole or clomiphene citrate (CC). Gonadotrophins have a direct effect on follicle growth as they contain FSH and may also contain recombinant LH- or HCG-driven LH activity. Letrozole is a third-generation aromatase inhibitor that interferes with the oestrogenic feedback at the pituitary by blocking oestrogen biosynthesis thus stimulating the production of serum FSH (Mitwally and Casper, 2001). CC is a selective oestrogen modulator and competes with oestrogen for binding to the hypothalamic oestrogen receptors, thus stimulating the production of serum FSH (Mitwally and Casper,

2001). While letrozole and CC are orally taken for 5 days, the gonadotrophins are injected subcutaneously.

Multiple systematic reviews have compared these ovarian stimulation agents with each other in women with unexplained infertility undergoing IUI. IUI with gonadotrophins increases live birth and/or ongoing pregnancy rates but also increased multiple pregnancy rates compared to other oral agents (Danhof *et al.*, 2020b; Zolton *et al.*, 2020). Nevertheless, IUI with adherence to strict cancellation criteria, i.e. withholding insemination if more than three dominant follicles develop, led to an acceptable multiple pregnancy rate without compromising the effectiveness (Danhof *et al.*, 2020b). However, a substantial unexplained heterogeneity across the primary trials comparing gonadotrophins to letrozole and CC was observed, and time-to-event outcomes were not reported in these meta-analyses (Eskew *et al.*, 2019; Danhof *et al.*, 2020b; Zolton *et al.*, 2020).

The population of couples with unexplained infertility is heterogeneous and the prognostic variables such as female age and duration of infertility affect pregnancy chances and safety issues independent of ovarian stimulation, such that on an individual level certain treatments may be more effective and/or safe than others (Steures *et al.*, 2004). Given the heterogeneous inclusion criteria of the primary trials, it is impossible to analyse interaction variables of couples with unexplained infertility in an aggregate data meta-analysis. To evaluate whether certain groups of couples benefit more from one treatment than from another, individual participant data meta-analysis (IPD-MA) of RCTs is optimal and therefore considered the as the gold standard for evidence synthesis (Riley *et al.*, 2010). In addition, IPD-MA also allows us to study time to conception leading to live birth which was impossible in aggregate data meta-analysis.

We therefore performed this IPD-MA to compare the effectiveness and safety of ovarian stimulation with gonadotrophins, letrozole and CC and to explore treatment-covariate interactions for important baseline characteristics in couples undergoing IUI-OS.

## Methods

### Registration and literature search

We conducted this IPD-MA according to a registered protocol (PROSPERO CRD42017053966) and reported it according to the Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data (PRISMA-IPD) statement (Stewart *et al.*, 2015).

We performed the search update on the 28th of June, 2021 based on an existing search strategy (Danhof *et al.*, 2020b). In brief, we searched the following electronic databases including MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, PsycINFO, Cochrane Gynaecology and Fertility Group trial register and Clinical Trial Registration Databases (clinicaltrials.gov and International Clinical Trials Registry Platform (ICTRP)). The detailed search strategy is presented in [Supplementary Table S1](#).

### Eligibility criteria

We included randomized controlled trials (RCTs) comparing IUI-OS with gonadotrophins, letrozole or CC among couples with unexplained infertility. We excluded dose comparing studies of the same

drug. If a trial also includes other factors of infertility, for instance ovulatory disorders, the trial was included but participants with other factors of infertility were excluded. We did not apply language restrictions.

### Study selection and data collection

Two authors (J.W. and M.v.W.) independently examined the studies for compliance with the inclusion criteria and selected eligible studies. Disagreements were resolved by discussion with a third author (R.W.).

We contacted the corresponding authors of all eligible studies to join the IUI IPD-MA collaboration and share their IPD and established the IUI IPD-MA Collaboration. We tried to obtain study protocols where possible. When we did not receive responses, we sent at least two more reminders. All authors sharing the IPD were asked to provide clarifications when information in the publications or datasets were unclear or inconsistent. We evaluated internal data consistency by checking duplicated and missing values as well as possible data errors and contacted the trial investigators for further clarification when needed.

### Outcomes

The primary effectiveness outcome was cumulative live birth per woman randomized and the primary safety outcome was multiple pregnancy. Secondary outcomes included ongoing pregnancy, clinical pregnancy, miscarriage, time to conception leading to live birth, cancellation and the total number of follicles > 14 mm at time of ovulation triggering. The unit of analysis was per couple randomized for all outcomes except for cancellation and total number of follicles > 14 mm at the time of ovulation triggering, in which the unit of analysis was per cycle. The definition of miscarriage was harmonized across different trials in this IPD-MA according to The International Glossary on Infertility and Fertility Care 2017 (Zegers-Hochschild *et al.*, 2017).

### Risk of bias and overall certainty of evidence assessment

Two authors (J.W. and R.W.) independently assessed the risk of bias of the included studies using the domain-based evaluation tool described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins *et al.*, 2011). Disagreements were resolved by discussion with a third author (M.v.W.). We assessed the following domains as low, unclear or high risk of bias: random sequence generation, allocation concealment, blinding of participant and personnel, blinding of outcome assessors, incomplete outcome data, selective reporting and other bias.

The overall certainty of evidence across RCTs were assessed when at least two studies were included by using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, including the risk of bias, consistency of effect, imprecision, indirectness and publication bias.

### Statistical analysis

We performed the analysis based on an intention-to-treat principle. We conducted a one-stage IPD-MA including random effects for trial in each pairwise comparison with studies contributing to IPD.

We also provided forest plots to visualize the results per trial and used the  $I^2$  statistic to quantify heterogeneity. Note that in these forest plots, the summary estimate was the one-stage estimate. For dichotomous outcomes, we estimated risk ratios (RR) using a generalized mixed model with a binomial distribution and a log link with random intercepts for study. For continuous outcomes, we estimated mean differences using a linear mixed model with random intercepts for study and cycle number. For time to conception leading to live birth, we used the number of IUI cycles as a time unit and calculated a pooled hazard ratio (HR) in Cox proportional hazards regression models for discrete time with a random effect (frailty with a normal distribution) for study (Fisher, 2015). Only conception that led to live birth were included.

Next, we explored treatment–covariate interaction of the following covariates on live birth: female age, type of infertility (primary/secondary) and body mass index (BMI). These treatment–covariate interactions were conducted using a two-stage approach, and were thus based solely on within-study information as recommended to avoid ecological bias (Fisher et al., 2017; Riley et al., 2020). As we limited the treatment–covariate interaction analysis to covariates that were available in at least 85% of participants, we did not perform analyses on other prespecified covariates duration of infertility, total motile sperm count, Hunault score, smoking status, ethnicity and antral follicle count due to missing data.

We then conducted pre-specified sensitivity analysis on studies with overall low risk of bias and studies with low risk of bias at allocation concealment to test the robustness of the findings. In addition, we also performed two post-hoc sensitivity analyses on studies with low starting dose of gonadotrophins ( $\leq 75$  IU) and on studies with stricter cancellation criteria ( $\leq 3$  dominant follicles), respectively. All sensitivity analyses were limited to the primary outcomes live birth and multiple pregnancy.

Finally, we intended to use funnel plots to explore the possibility of small study effects if at least 10 studies were present per comparison. To examine IPD availability bias, we also presented meta-analyses of trials without IPD.

Missing outcome data were not imputed. Data on the covariates female age, primary/secondary infertility and BMI were missing for 9% of participants and imputed using single imputation.

Data were prepared in Stata 16.1 and Microsoft Excel. IPD-MA was performed in R version 3.6.0 using the *rms*, *survival*, *foreign*, *mice*, *lme4*, *meta* and *miceadds* R packages and additional analysis was performed in Stata 16.1 using *admetan* package.

## Results

### Study selection

In total, we identified 338 studies, of which 313 studies were excluded after screening titles and abstract (Fig. 1). After screening full text, 22 studies were eligible. IPD was not sought from four studies due to insufficient contact information ( $n = 4$ , 224 couples) (Kamel, 1995; Sammour et al., 2001; Fatemi et al., 2003; Galal, 2015). The authors of the remaining 18 studies were contacted, among which IPD of 11 studies (1278 couples) were not available, due to either no response ( $n = 9$ , 1009 couples) (Balasch et al., 1994; Nakajima et al., 1999; El

Helw and El Sadek, 2002; Al-Fozan et al., 2004; Ozmen et al., 2005; Gregoriou et al., 2008; Fouda and Sayed, 2011; Ibrahim et al., 2012; Goldman et al., 2014) or data loss ( $n = 2$ , 269 couples) (Baysoy et al., 2006; Berker et al., 2011). These studies are listed in Supplementary Table SII. IPD of seven trials were provided by the trial authors (Ecochard et al., 2000; Dankert et al., 2007; Diamond et al., 2015; Erdem et al., 2015; Peeraer et al., 2015; Danhof et al., 2018; Naidu et al., 2020).

### Study characteristics

Characteristics of trials with and without IPD are presented in Table I and Supplementary Table SIII. No major issues were identified when checking the consistency of IPD.

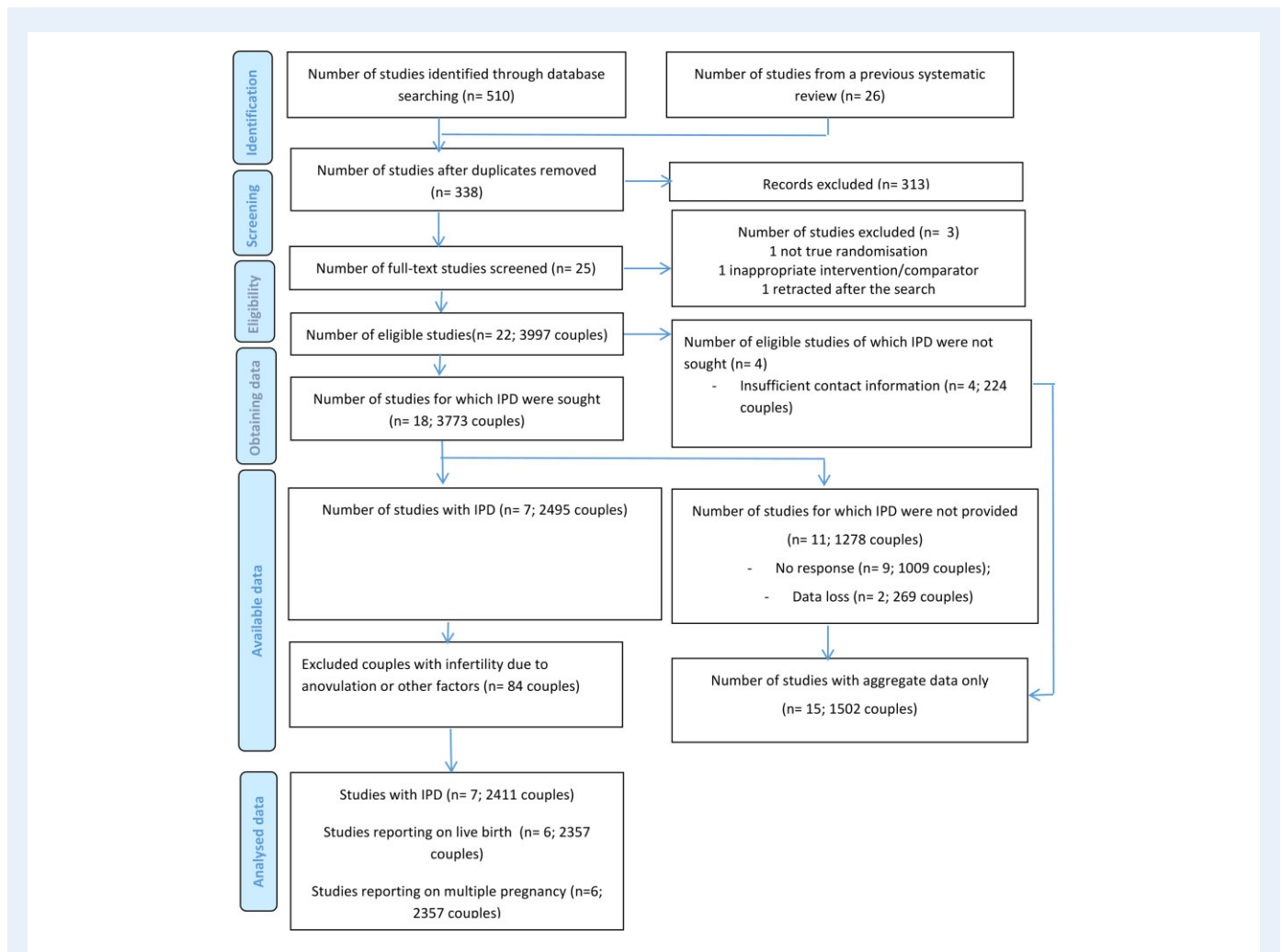
Of the included RCTs that provided IPD, five were multicentre studies (Ecochard et al., 2000; Dankert et al., 2007; Diamond et al., 2015; Peeraer et al., 2015; Danhof et al., 2018) and two were single-centre studies (Erdem et al., 2015; Naidu et al., 2020). All these RCTs were published in English between 2000 and 2020, including a conference abstract-only publication (Naidu et al., 2020). Five of the seven RCTs only included couples with unexplained infertility. The other two RCTs (Ecochard et al., 2000; Peeraer et al., 2015) also included women with other factors of infertility (e.g. ovulatory dysfunction and mixed factors) that were excluded from this IPD-MA ( $n = 84$ ).

In six RCTs involving 1511 women undergoing IUI-OS, gonadotrophins were compared to CC (Ecochard et al., 2000; Dankert et al., 2007; Erdem et al., 2015; Peeraer et al., 2015; Danhof et al., 2018; Naidu et al., 2020). One study, investigating 900 women, compared all three medications: gonadotrophins, letrozole and CC (Diamond et al., 2015). All RCTs had an IUI protocol with cancellation criteria, with five of them being more strict than the other two (maximum of three dominant follicles) (Ecochard et al., 2000; Dankert et al., 2007; Peeraer et al., 2015; Danhof et al., 2018; Naidu et al., 2020). One study performed selective ultrasound-guided follicular aspiration or cancelled the cycle (Peeraer et al., 2015).

The seven RCTs provided IPD on 2411 women who received 5678 IUI cycles. There were 1054 women allocated to gonadotrophins, 299 to letrozole and 1058 to CC. Overall characteristics and outcomes of all couples included in this IPD-MA are presented in Table II and Supplementary Table SIV.

### Risk of bias of individual RCTs

The details of the risk of bias assessment of the individual RCTs are presented in Fig. 2. One study (Erdem et al., 2015) was scored high risk of bias at allocation concealment because allocation was not concealed. One study (Diamond et al., 2015) was a double-blinded study for CC and letrozole and was scored low risk of performance bias for this comparison. All RCTs involving gonadotrophins were open label and therefore were scored at unclear risk of performance bias. Given that all reproductive outcomes of interest were objective outcomes, it is unlikely that the non-blinded design will affect these outcome measurements and therefore we scored all the included RCTs at low risk for detection bias. Attrition bias was scored at low risk for all trials. For selective reporting, we scored two studies to have unclear risk (Ecochard et al., 2000; Naidu et al., 2020) because study protocol or registration was not available for



**Figure 1. PRISMA IPD flow diagram.** IPD, individual participant data.

assessments. Other risk of bias was scored unclear for two studies due to lack of information on important baseline variables (Dankert *et al.*, 2007; Naidu *et al.*, 2020).

## Primary outcomes

### Live births

Six trials provided data on live birth. The results per study and the pooled one-stage estimated RRs are shown in Fig. 3a–c and Table III. Gonadotrophins increased the chance of a live birth compared to CC (RR 1.30, 95% CI 1.12–1.51,  $I^2 = 26\%$ ) and letrozole (RR 1.72, 95% CI 1.29–2.29). This implies that if the live birth rate following IUI with CC is assumed to be 22%, the live birth rate following IUI with gonadotrophins would be between 25% and 33% (NNT = 15 (9–38)). If the live birth rate following IUI with letrozole is assumed to be 19%, the live birth rate following IUI with gonadotrophins would be between 25% and 44% (NNT = 7 (4–17)). There was insufficient evidence of a difference between letrozole and CC on live birth (RR 0.80, 95% CI 0.59–1.10). If the live birth rate following IUI with CC is assumed to be 23%, the live birth rate following IUI with letrozole

would be between 13% and 25% (NNT = –22). We did not conduct a network meta-analysis as only one trial included letrozole (Diamond *et al.*, 2015).

Sensitivity analysis on five RCTs with low risk of bias at allocation concealment, thereby excluding Erdem *et al.*, were consistent with the main findings (RR 1.23, 95% CI 1.05–1.45,  $I^2 = 0\%$ ) (Table V). Sensitivity analysis on RCTs with overall low risk of bias was not performed due to the open-label design on the use of gonadotrophins in all RCTs.

Post-hoc sensitivity analyses on RCTs with low starting dose of gonadotrophins ( $\leq 75$  IU) showed similar results (RR 1.26, 95% CI 1.05–1.51) (Table V). This implies that if the live birth rate following IUI with CC is assumed to be 22%, the live birth rate following IUI with gonadotrophins would be between 23% and 33%. For stricter cancellation criteria, we are uncertain whether gonadotrophins lead to higher live birth rates than CC (RR 1.15, 95% CI 0.94–1.41) (Table V). This implies that if the live birth rate following IUI with CC is assumed to be 22%, the live birth rate following IUI with gonadotrophins would be between 21% and 31%.

**Table 1** Outline and design per trial.

Study	Year	Country	Number of participants	Intervention*			Cancellation criteria	Outcomes
				Clomiphene citrate	Letrozole	Gonadotrophins		
<b>Danhof et al.</b>	2018	The Netherlands	738	100 mg		75 IU	Max 3 follicles of $\geq$ 14 mm or max 5 follicles of $\geq$ 12 mm	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception, cancellation, number of follicles
<b>Dankert et al.</b>	2007	The Netherlands	138	100 mg		75 IU	Max 3 follicles of $\geq$ 14 mm	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception
<b>Diamond et al.</b>	2015	USA	900	100 mg	5 mg	150 IU	Max 4 follicles (mean diameter $>$ 18 mm) or max serum E2 levels of 3000 pg/ml	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception, cancellation, number of follicles
<b>Ecochard et al.</b>	2000	France	54	50 or 100 mg		150 IU	Max 3 follicles of $\geq$ 15 mm and/or max serum E2 levels of 1200 pg/ml	Clinical pregnancy, miscarriage
<b>Erdem et al.</b>	2015	Turkey	219	100 mg		75 IU	Max 4 follicles of $\geq$ 14 mm and/or max serum E2 levels of 1500 pg/ml	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception, cancellation, number of follicles
<b>Naidu et al.</b>	2020	India	112	100 mg		75 IU	Max 3 follicles of $\geq$ 18 mm <sup>‡</sup>	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception, cancellation, number of follicles
<b>Peeraer et al.</b>	2015	Belgium	250	50 mg		37.5 or 75 IU	Max 3 follicles of $\geq$ 14 mm	Live birth, multiple pregnancy, clinical pregnancy, ongoing pregnancy, miscarriage, time to conception, cancellation, number of follicles

\*Start dosing, during the cycles dosage can be adjusted. <sup>‡</sup>No participants had more than three dominant ( $\geq$ 14 mm) follicles in the trial.

**Table II Overall characteristics and outcomes.**

Characteristic or outcome	Number of studies	Number of women	Gonadotrophins mean (25th–75th percentile) or N (%)	Clomiphene citrate mean (25th–75th percentile) or N (%)	Letrozole mean (25th–75th percentile) or N (%)
<b>Female age (in years)</b>	6	2273	31.9 (29.0–35.0)	32.0 (29.0–35.0)	32.2 (29.0–35.0)
<b>Body mass index, BMI</b>	5	2161	24.9 (21.2–27.0)	24.7 (21.1–26.6)	27.3 (22.3–30.9)
<b>Primary infertility (%)</b>	6	2273	714 (72%)	695 (70%)	180 (60%)
<b>Number of cycles</b>	6	2381	2.5 (1.0–4.0)	2.6 (1.0–4.0)	3.2 (2.0–4.0)
<b>Live birth (%)</b>	6	2357	287 (28%)	222 (22%)	56 (19%)
<b>Multiple pregnancy (%)</b>	6	2357	47 (5%)	22 (2%)	9 (3%)
- Twin			35 (74%)	21 (95%)	9 (100%)
- Triplet			12 (26%)	1 (5%)	0 (0%)
<b>Ongoing pregnancy (%)</b>	6	2357	299 (29%)	233 (23%)	60 (20%)
<b>Clinical pregnancy (%)</b>	7	2411	326 (31%)	268 (25%)	67 (22%)
<b>Miscarriage (%)</b>	7	2411	71 (7%)	54 (5%)	9 (3%)
<b>Cancellations, per cycle (%)</b>	5	2163	231 (9%)	204 (8%)	35 (4%)
<b>Number of follicles &gt;14 mm, per cycle</b>	5	2163	2.2 (1–3)	2.1 (1–3)	2.0 (1–2)

### Multiple pregnancy

Six trials provided data on multiple pregnancy. The results per study and the pooled one-stage estimated RRs are shown in Fig. 3d–f and Table III. Gonadotrophins increased the risk of a multiple pregnancy compared to both CC (RR 2.17, 95% CI 1.33–3.54,  $I^2 = 69%$ ) and letrozole (RR 3.75, 95% CI 1.83–7.69), whereas there was insufficient evidence of a difference between letrozole and CC (RR 1.13, 95% CI 0.44–2.89). There were 12 triplet pregnancies in the gonadotrophins group, one in the CC group (RR 12.06, 95% CI 1.57–92.46,  $I^2 = 0$ ) and none in the letrozole group. Out of the 12 women in the gonadotrophins group who had triplets, 10 used a high dose (9 used 150 IU and 1 used 225 IU) and 2 used a low dose (75 IU) of gonadotrophins. Post-hoc sensitivity analyses on RCTs with low starting dose of gonadotrophins ( $\leq 75$  IU) and on RCTs with stricter cancellation criteria ( $\leq 3$  dominant follicles) showed insufficient evidence of a difference in multiple pregnancy (RR 0.94, 95% CI 0.45–1.96; and RR 0.81, 95% CI 0.32–2.03, respectively) (Table V).

### Secondary outcomes per woman

#### Ongoing pregnancy and clinical pregnancy

Six trials provided data on ongoing pregnancy while all seven trials provided data on clinical pregnancy (Table III). The results were consistent with those of the primary outcome; gonadotrophins increased the change of an ongoing pregnancy and clinical pregnancy compared to both CC and letrozole whereas there was insufficient evidence of a difference between letrozole and CC.

#### Miscarriage

Seven trials provided data on miscarriages (Table III). There was insufficient evidence of a difference between gonadotrophins and CC (RR 1.32, 95% CI 0.94–1.86,  $I^2 = 0%$ ) or between letrozole and CC (RR

0.65, 95% CI 0.28–1.47), whereas gonadotrophins increased the chance of a miscarriage compared to letrozole (RR 2.87, 95% CI 1.37–6.02).

#### Time to conception leading to live birth

Six trials provided data on time to conception leading to live birth (Table III) (Dankert *et al.*, 2007; Diamond *et al.*, 2015; Erdem *et al.*, 2015; Peeraer *et al.*, 2015; Danhof *et al.*, 2018; Naidu *et al.*, 2020). Gonadotrophins reduced the time to conception leading to a live birth compared to both CC (HR 1.37, 95% CI 1.15–1.63,  $I^2 = 22%$ ) and letrozole (HR 2.04, 95% CI 1.47–2.83). Letrozole appeared to increase the time to conception leading to a live birth compared to CC (HR 0.77, 95% CI 0.54–1.09).

#### Ovarian hyperstimulation syndrome

Three trials reported on ovarian hyperstimulation syndrome, with only one case in the gonadotrophins group. Therefore, a meta-analysis was not performed.

### Secondary outcomes per cycle

#### Cancellations

Five trials provided cycle-level data on cancellations of insemination (5958 cycles) for reasons of cancellation (Supplementary Table SV) (Diamond *et al.*, 2015; Erdem *et al.*, 2015; Peeraer *et al.*, 2015; Danhof *et al.*, 2018; Naidu *et al.*, 2020). The use of gonadotrophins resulted in a higher risk for cancellation than letrozole (RR 1.84, 95% CI 1.17–2.87), while there was insufficient evidence of a difference in cancellation of insemination between gonadotrophins and CC (RR 1.15, 95% CI 0.96–1.37,  $I^2 = 77%$ ), or between letrozole and CC (RR 1.14, 95% CI 0.69–1.88).

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Danhof 2018	+	+	?	+	+	+	+
Dankert 2007	+	+	?	+	+	+	?
Diamond 2015	+	+	?	+	+	+	+
Ecochard 2000	+	+	?	+	+	?	+
Erdem 2015	+	+	?	+	+	+	+
Naidu 2020	+	+	?	+	+	?	?
Peeraer 2015	+	+	?	+	+	+	+

■ Low risk of bias   
■ Unclear risk of bias   
■ High risk of bias

**Figure 2. Risk of bias summary of the included randomized controlled trials (RCTs) according to the bias assessment tool of the Cochrane Collaboration.** Performance bias for Diamond et al. (2015) was considered as low risk for letrozole versus clomiphene citrate (CC) and as unclear risk for the other two comparisons.

### Number of follicles

Five trials provided cycle-level data on the number of follicles of >14 mm (5958 cycles) (Diamond et al., 2015; Erdem et al., 2015; Peeraer et al., 2015; Danhof et al., 2018; Naidu et al., 2020). Gonadotrophins led to a larger mean number of follicles compared to CC (MD 0.16, 95% CI 0.07–0.26,  $I^2 = 90%$ ) and inconclusive findings compared to letrozole (MD 1.15, 95% CI 0.95–1.34), whereas letrozole resulted in a smaller mean number of follicles compared to CC (MD –0.59, 95% CI –0.75 to –0.43).

## Exploratory treatment–covariate interactions

### Female age

Five trials provided data on female age and live birth and were used for the treatment–covariate interaction analyses on age (Table IV) (Diamond et al., 2015; Erdem et al., 2015; Peeraer et al., 2015; Danhof et al., 2018; Naidu et al., 2020). When comparing gonadotrophins to CC, the estimated interaction RR per year of female age was

0.94 (95% CI 0.85–1.05,  $I^2 = 61%$ ). Comparing letrozole to CC, the estimated interaction RR was 1.03 (95% CI 0.96–1.10). Comparing gonadotrophins to letrozole, the estimated interaction RR was 1.02 (95% CI 0.95–1.09). Insufficient evidence on the treatment–covariate interaction of female age was found.

### BMI

Four trials provided data on BMI and live birth and were used for the treatment–covariate interaction analyses on BMI (Table IV) (Diamond et al., 2015; Erdem et al., 2015; Peeraer et al., 2015; Danhof et al., 2018). When comparing gonadotrophins to CC, the estimated interaction RR per unit of BMI was 1.03 (95% CI 0.95–1.11,  $I^2 = 40%$ ). Comparing letrozole to CC, the estimated interaction RR was 1.02 (95% CI 0.98–1.08). Comparing gonadotrophins to letrozole, the estimated interaction RR was 0.99 (95% CI 0.94–1.03). Insufficient evidence on the treatment–covariate interaction of BMI was found.

### Primary versus secondary infertility

Five trials provided data on live birth and type of infertility but due to small number of events, RR was not estimable in Naidu et al. (2020) and due to primary infertility as part of the inclusion criteria, the interaction was not estimable in Erdem et al. (2015); therefore three trials remained (Diamond et al., 2015; Peeraer et al., 2015; Danhof et al., 2018). When comparing gonadotrophins to CC, the estimated interaction RR for primary versus secondary infertility was 0.86 (95% CI 0.58–1.26,  $I^2 = 0%$ ). Comparing letrozole to CC, the estimated interaction RR was 1.33 (95% CI 0.70–2.54). Comparing gonadotrophins to letrozole, the estimated interaction RR was 0.60 (95% CI 0.33–1.09). Insufficient evidence on the treatment–covariate interaction of primary versus secondary infertility was found.

## Additional analysis

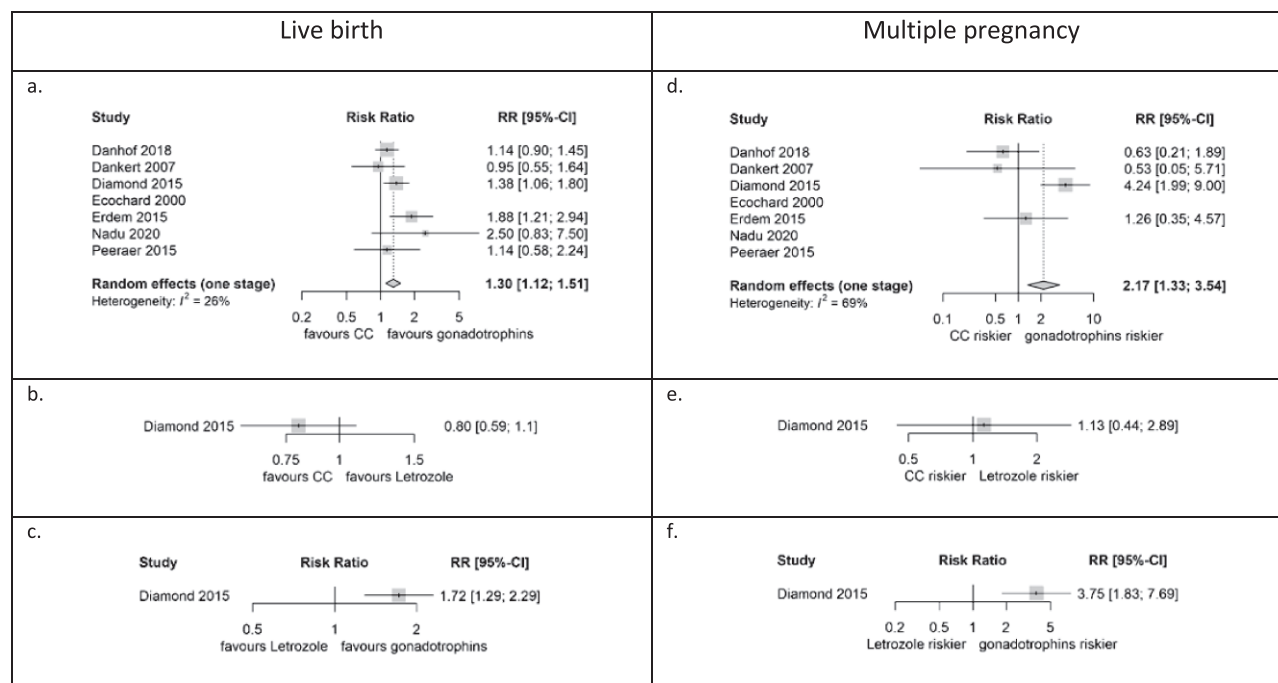
We did not present funnel plots as fewer than 10 trials were included. Meta-analyses of trials without IPD showed overlapping confidence intervals with the IPD-MAs in most comparisons, except for letrozole versus CC (Supplementary Table SVI). None of the trials comparing letrozole versus CC that did not contribute to IPD reported live birth. In addition, meta-analysis of trials not contributing to IPD showed higher clinical pregnancy rates in the letrozole group compared to CC (RR 1.66, 95% CI 1.24–2.24), which was inconsistent with the result from the trials contributing to IPD.

## Discussion

### Summary of evidence

In this IPD-MA, we evaluated the effectiveness and safety of ovarian stimulation with gonadotrophins, letrozole or CC in couples with unexplained infertility undergoing IUI-OS. Moderate-quality evidence showed that gonadotrophins increased the chance of a live birth compared to both CC and letrozole, while low-quality evidence due to substantial heterogeneity, suggested it may also increase the chance of a multiple pregnancy, especially for triplet pregnancy. Gonadotrophins





**Figure 3. Forest plots for live birth and multiple pregnancy.** (a–c) Live birth: (a) comparing gonadotrophins and CC; (b) comparing letrozole and CC; (c) comparing gonadotrophins and letrozole. (d–f) Multiple pregnancy: (d) comparing gonadotrophins and CC. (e) Comparing letrozole and CC. (f) Comparing gonadotrophins and letrozole. In each forest plot, the study level estimate was based on individual patient data (IPD) of each individual study and the summary estimate was based on a one-stage IPD meta-analysis (IPD-MA). In (a), Ecochard (2000) did not report live birth and therefore was not included in the IPD-MA. In (d), Ecochard (2000) did not report multiple pregnancy and therefore was not included in the IPD-MA. Note that study level estimates were not shown for two other studies (Peeraer, 2015; Naidu, 2020), due to the presence of 0 events in one group (Peeraer, 2015) or 0 events in both groups (Naidu, 2020), but the one-stage IPD-MA for multiple pregnancy included these two studies. CC, clomiphene citrate; RR, relative risk.

reduced the time to conception leading to live birth when compared to both CC and letrozole. Gonadotrophins gave a significantly higher number of dominant follicles compared to letrozole but also a significantly higher risk of cancellation. Compared to CC, the number of follicles was significantly higher when gonadotrophins were used but this does not necessarily lead to a higher cancellation rate, as the heterogeneity for the number of follicles was very high. We did not find treatment–covariate interactions on live birth for the pre-specified covariates: age, BMI or primary versus secondary infertility. The heterogeneity between trials comparing gonadotrophins and CC on multiple pregnancy could be explained by the choice of different starting doses of gonadotrophins and cancellation criteria in different trials. In the gonadotrophin group, there were 12 triplet pregnancies, and 10 of these women used a high dose of gonadotrophins (9 used 150 IU and 1 used 225 IU). When limiting the analysis to studies with a low starting dose of gonadotrophins ( $\leq 75$  IU), ovarian stimulation with gonadotrophins still significantly increased the probability of live birth. When limiting the analysis to studies with strict cancellation criteria, the difference was no longer significant, such that we are uncertain whether ovarian stimulation with gonadotrophins lead to higher live birth rates compared to CC. Both sensitivity analyses showed comparable risks of a multiple pregnancy between gonadotrophins and CC.

### Strengths and limitations

The strengths of the IPD-MA are the harmonization of the eligibility criteria (by excluding women with anovulatory infertility) and outcome definitions, the analyses of time to conception and treatment–covariate interactions.

A potential limitation of this IPD-MA is that we were not able to access the IPD of all eligible studies. IPD was available for just 32% (7/22) of the included trials, however the included studies accounted for 62% (2495/3997) of all participants. Previous empirical evidence in our research field has demonstrated that results of the RCTs without IPD have lower quality and more methodological issues compared to RCTs who shared IPD and therefore may result in different findings (Wicherts et al., 2011; Bordewijk et al., 2020). The willingness to share these data may indicate a good quality trial. For example, all 15 studies that did not share IPD lacked adequate trial registration (Supplementary Table SIII). Only two had trial registration, both of which were registered after start of the recruitment process. It is worth noting that one eligible trial with a large sample size ( $n=412$ ) comparing letrozole and CC was retracted due to concerns on its validity and this trial was not included in our IPD-MA (Badawy et al., 2009). As data checking is a mandatory process during IPD-MA, evidence from this IPD-MA should be considered

**Table III** Meta-analyses and GRADE assessments of all outcomes.

Comparison	Outcome	Number of RCTs	Number of participants	Risk ratio or hazard ratio	95% CI	I <sup>2</sup>	Overall certainty of evidence (GRADE)
<b>Gn vs CC</b>	Live birth	6	2058	1.30	1.12–1.51	26%	Moderate <sup>a</sup>
	Multiple pregnancy	6	2058	2.17	1.33–3.54	69%	Low <sup>a,b</sup>
	Ongoing pregnancy	6	2058	1.29	1.11–1.49	18%	Moderate <sup>a</sup>
	Clinical pregnancy	7	2112	1.22	1.07–1.40	0%	Moderate <sup>a</sup>
	Miscarriage	7	2112	1.32	0.94–1.86	0%	Low <sup>a,c</sup>
Time to conception <sup>†</sup>	6	2058	1.37	1.15–1.63	22%	Moderate <sup>a</sup>	
<b>Letrozole vs CC</b>	Live birth	1	599	0.80	0.59–1.10	n/a	n/a
	Multiple pregnancy	1	599	1.13	0.44–2.89	n/a	n/a
	Ongoing pregnancy	1	599	0.79	0.59–1.07	n/a	n/a
	Clinical pregnancy	1	599	0.79	0.60–1.04	n/a	n/a
	Miscarriage	1	599	0.65	0.28–1.47	n/a	n/a
	Time to conception <sup>†</sup>	1	599	0.77	0.54–1.09	n/a	n/a
<b>Gn vs Letrozole</b>	Live birth	1	600	1.72	1.29–2.29	n/a	n/a
	Multiple pregnancy	1	600	3.75	1.83–7.69	n/a	n/a
	Ongoing pregnancy	1	600	1.66	1.25–2.18	n/a	n/a
	Clinical pregnancy	1	600	1.59	1.22–2.06	n/a	n/a
	Miscarriage	1	600	2.87	1.37–6.02	n/a	n/a
	Time to conception <sup>†</sup>	1	600	2.04	1.47–2.83	n/a	n/a

<sup>†</sup>Hazard ratio for time to conception leading to live birth

<sup>a</sup>Downgraded by one level due to concerns on risk of bias

<sup>b</sup>Downgraded by one level due to inconsistency

<sup>c</sup>Downgraded by one level due to imprecision

Gn, gonadotrophins; CC, clomiphene citrate; GRADE, Grading of Recommendations, Assessment, Development and Evaluations; RCT, randomized controlled trial.

**Table IV** Meta-analyses of treatment-covariate interactions on live birth.

Comparison	Baseline covariate	Number of RCTs	Number of participants	Risk ratio (RR)	95% CI	I <sup>2</sup>
<b>Gn vs CC</b>	Age	5	1920	0.94	0.85–1.05	61%
	BMI	4	1808	1.03	0.95–1.11	40%
	Type of infertility (primary vs secondary)	3	1589	0.86	0.58–1.26	0%
<b>Letrozole vs CC</b>	Age	1	599	1.03	0.96–1.10	
	BMI	1	599	1.02	0.98–1.08	
	Type of infertility (primary vs secondary)	1	599	1.33	0.70–2.54	
<b>Gn vs Letrozole</b>	Age	1	600	1.02	0.95–1.09	
	BMI	1	600	0.99	0.94–1.03	
	Type of infertility (primary vs secondary)	1	600	0.60	0.33–1.09	

Gn, gonadotrophins; CC, clomiphene citrate; RCT, randomized controlled trial.

the gold standard to inform clinical practice. Another limitation is that some baseline variables including total sperm motile count, smoking status, ethnicity and Hunault score were not available in the databases of multiple trials and therefore these treatment-covariate interactions were not explored.

### Interpretations

Clinical decision-making should be based on a joint assessment of safety, effectiveness, availability, cost of the interventions as well as couples' preferences. Multiple pregnancy is an important measure for safety. The overall effectiveness on gonadotrophins in multiple

**Table V** Sensitivity analyses on live birth and multiple pregnancy comparing gonadotrophins versus clomiphene citrate.

Sensitivity analysis	Outcome	Number of RCTs	Number of participants	Risk Ratio (RR)	95% CI	I <sup>2</sup>
<b>RCTs with low risk of bias at allocation concealment*</b>	Live birth	5	1839	1.23	1.05–1.45	0%
	Multiple pregnancy	5	1839	2.37	1.39–4.04	78%
<b>RCTs with low starting dose of gonadotrophins (≤75IU)**</b>	Live birth	5	1457	1.26	1.05–1.51	37%
	Multiple pregnancy	5	1457	0.94	0.45–1.96	0%
<b>RCTs with stricter cancellation criteria (≤ 3 dominant follicles)***</b>	Live birth	4	1238	1.15	0.94–1.41	0%
	Multiple pregnancy	4	1238	0.81	0.32–2.03	0%

\*Erdem (2015) was excluded;

\*\*Diamond (2015) was excluded;

\*\*\*Both Erdem (2015) and Diamond (2015) were excluded.

RCT, randomized controlled trial.

pregnancy was dominated by a single study (Diamond *et al.*, 2015) with the highest number of events in the gonadotrophins group (34/301), in which a less strict cancellation criteria and a conventional starting dose (150IU) were used. Although the post-hoc sensitivity analyses should be interpreted with caution, they indicate that a lower starting dose of gonadotrophins and stricter cancellation criteria may result in a comparable low multiple pregnancy rate between gonadotrophins to CC, by reducing the chance of multifollicular growth and/or cancellation of the cycle (van Rumste *et al.*, 2008). In view of ongoing concerns on high multiple pregnancies resulting from non-IVF medically assisted reproduction procedures globally, a low starting dose of gonadotrophins with strict cancellation criteria should be considered (Danhof *et al.*, 2018; Bergh *et al.*, 2020), but it is not clear to which extent this recommendation has been applied in daily clinical practice worldwide. The findings of the overall IPD-MA on effectiveness including all studies favours gonadotrophins, but the advantage of gonadotrophins over CC on effectiveness (live birth) becomes smaller when factoring in a low starting dose or even disappears when applying strict cancellation criteria. In a modern practice where a lower starting dose and stricter cancellation criteria are in place, the effectiveness and safety of the two different agents seem both acceptable, and therefore intervention availability, cost and patients' preferences are valued as more important in clinical decision-making.

Letrozole is still an off-label agent for unexplained infertility in many countries, resulting in its unavailability, although it has been widely used in clinical practice with no evidence of an increased risk of congenital foetal malformation (Pundir *et al.*, 2021). Comparisons including letrozole were underpowered due to letrozole only being used in a single trial in this IPD-MA. The other 11 eligible trials involving letrozole did not contribute IPD, including eight studies which compared letrozole to CC and three studies which compared gonadotrophins to letrozole. RCTs contributing IPD and RCTs not contributing IPD showed inconsistent results on clinical pregnancy (Supplementary Table SVI) and none of the RCTs that did not contribute IPD reported on live birth. Therefore, evidence on its effectiveness in women with unexplained infertility is urgently needed and new RCTs comparing letrozole with other ovarian stimulation agents should be performed.

Recent cost-effectiveness analyses on ovarian stimulation agents in IUI showed that in settings where a live birth is valued at €3000 or less, between €3000 and €55 000 and above €55 000, CC, letrozole and gonadotrophins were the most cost-effective option in terms of net benefit, respectively (van Eekelen *et al.*, 2021). While recommending CC and letrozole, the authors also highlighted the high uncertainty surrounding such findings and call for more research on the relative effectiveness in this area (van Eekelen *et al.*, 2021). This could be due to the overall small differences between these agents in modern practice where a lower starting dose and stricter cancellation criteria are in place (Danhof *et al.*, 2020a). Given the cost variations across countries, future cost-effectiveness studies in different settings would be helpful to provide further health economic evidence to inform clinical decision-making. This is especially important for clinical decision-making in low or middle resource settings, where limited economic evidence is available. Finally, although preferences would depend on the health system in different countries, oral agents are still likely to be preferred in many settings when the effectiveness and safety are acceptable, given their convenience in use (Practice Committee of the American Society for Reproductive Medicine, 2020).

## Conclusion

In couples with unexplained infertility undergoing IUI-OS, gonadotrophins increased the chance of a live birth and reduced the time to conception compared to CC, at the cost of a higher multiple pregnancy rate, when not differentiating strategies on the cancellation criteria or the starting dose. The treatment effects did not seem to differ in women with different age, BMI or primary versus secondary infertility. In a modern practice where a lower starting dose and stricter cancellation criteria are in place, effectiveness and safety of different agents seem both acceptable, and therefore intervention availability, cost and patients' preferences should factor in the clinical decision-making. As the evidence for comparisons to letrozole is based on one RCT providing IPD, further RCTs comparing letrozole and other interventions for unexplained infertility are needed.

## Supplementary data

Supplementary data are available at *Human Reproduction Update* online.

## Data availability

The trial investigators (or their institutions) who shared the individual participant data (IPD) for this IPD meta-analysis retain ownership of their trial IPD. All requests for access to the IPD should be made directly to the trial investigators. The data dictionary and code for this IPD meta-analysis are available on request.

## Acknowledgements

We would like to thank all investigators that shared their data with special thanks to Dr R. Ecochard from the Hospices Civils de Lyon. We would like to acknowledge all the investigators and participants of the primary trials. The investigators of individual trials are listed in [Supplementary Table SVII](#). We would like to acknowledge the assistance of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the Reproductive Medicine Network (RMN) and the Protocol Subcommittee, in making the database of the AMIGOS trial available. The RMN Steering Committee was not involved in analysis, data interpretation and manuscript preparation. The contents of this article represent the views of the authors and do not represent the views of the NICHD RMN.

## Authors' roles

J.W. drafted the manuscript. N.D., B.W.M., M.v.W., M.M. and R.W. designed the study. M.S. developed the search strategies. N.D., R.v.E. and R.W. cleaned the data. R.v.E. and R.W. analysed the data. M.D., R.L., K.P., T.D., M.E., T.D., B.C., C.T. and M.M. provided data from the included studies. All authors interpreted the pooled data, critically revised the manuscript and approved the final version.

## Funding

This study was supported by the NHMRC Centre for Research Excellence in Women's Health in Reproductive Life (APP 1171592) through a fellowship grant (to RW) and an NHMRC Emerging Leadership Investigator grant (2009767).

## Conflict of interest

B.W.M. was supported by a NHMRC Investigator grant (GNT1176437). B.W.M. reports consultancy for ObsEva and has received research funding from Guerbet, Ferring and Merck. T.D. is vice president and head of Global Medical Affairs Fertility, Research and Development, Merck KGaA, Darmstadt, Germany. He is also a Guest professor in Reproductive Medicine and Biology at the Department of Development and Regeneration, Group Biomedical Sciences, KU

Leuven (University of Leuven), Belgium and an adjunct professor at the Department of Obstetrics and Gynecology in the University of Yale, New Haven, USA. T.D. also reports consulting fees from Forendo. The other authors do not have any conflict of interest to declare.

## References

- Al-Fozan H, Al-Khadouri M, Tan SL, Tulandi T. A randomized trial of letrozole versus clomiphene citrate in women undergoing superovulation. *Fertil Steril* 2004;**82**:1561–1563.
- Badawy A, Elnashar A, Totongy M. RETRACTED: clomiphene citrate or aromatase inhibitors for superovulation in women with unexplained infertility undergoing intrauterine insemination: a prospective randomized trial. *Fertil Steril* 2009;**92**:1355–1359.
- Balasz J, Balleca JL, Pimentel C, Creus M, Fabregues F, Vanrell JA. Late low-dose pure follicle stimulating hormone for ovarian stimulation in intra-uterine insemination cycles. *Hum Reprod* 1994;**9**:1863–1866.
- Baysoy A, Serdaroglu H, Jamal H, Karatekeli E, Ozornek H, Attar E. Letrozole versus human menopausal gonadotrophin in women undergoing intrauterine insemination. *Reprod Biomed Online* 2006;**13**:208–212.
- Bergh C, Kamath MS, Wang R, Lensen S. Strategies to reduce multiple pregnancies during medically assisted reproduction. *Fertil Steril* 2020;**114**:673–679.
- Berker B, Kahraman K, Taskin S, Sukur YE, Sonmezer M, Atabekoglu CS. Recombinant FSH versus clomiphene citrate for ovarian stimulation in couples with unexplained infertility and male subfertility undergoing intrauterine insemination: a randomized trial. *Arch Gynecol Obstet* 2011;**284**:1561–1566.
- Bordewijk EM, Wang R, van Wely M, Costello MF, Norman RJ, Teede H, Gurrin LC, Mol BW, Li W. To share or not to share data: how valid are trials evaluating first-line ovulation induction for polycystic ovary syndrome? *Hum Reprod Update* 2020;**26**:929–941.
- Danhof NA, Van Wely M, Repping S, Koks C, Verhoeve HR, De Bruin JP, Verberg MFG, Van Hooff MHA, Cohlen BJ, Van Heteren CF et al.; SUPER study group. Follicle stimulating hormone versus clomiphene citrate in intrauterine insemination for unexplained subfertility: a randomized controlled trial. *Hum Reprod* 2018;**33**:1866–1874.
- Danhof NA, van Wely M, Repping S, van der Ham DP, Klijn N, Janssen I, Rijn-van Weert JM, Twisk M, Traas MAF, Pelinck MJ et al; Amsterdam SUPER Study Group. Gonadotrophins or clomiphene citrate in couples with unexplained infertility undergoing intrauterine insemination: a cost-effectiveness analysis. *Reprod Biomed Online* 2020a;**40**:99–104.
- Danhof NA, Wang R, van Wely M, van der Veen F, Mol BWJ, Mochtar MH. IUI for unexplained infertility—a network meta-analysis. *Hum Reprod Update* 2020b;**26**:1–15.
- Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H, van Dop PA. A randomized clinical trial of clomiphene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. *Hum Reprod* 2007;**22**:792–797.
- Diamond MP, Legro RS, Coutifaris C, Alvero R, Robinson RD, Casson P, Christman GM, Ager J, Huang H, Hansen KR et al.

- Letrozole, gonadotropin, or clomiphene for unexplained infertility. *N Engl J Med* 2015;**373**:1230–1240.
- Ecochard R, Mathieu C, Royere D, Blache G, Rabilloud M, Czyba JC. A randomized prospective study comparing pregnancy rates after clomiphene citrate and human menopausal gonadotropin before intrauterine insemination. *Fertil Steril* 2000;**73**:90–93.
- El Helw B, El Sadek M. Single dose letrozole versus CC for superovulation prior to intrauterine insemination in a prospective RCT. *Hum Reprod* 2002;**17**:73.
- Erdem M, Abay S, Erdem A, Firat Mutlu M, Nas E, Mutlu I, Oktem M. Recombinant FSH increases live birth rates as compared to clomiphene citrate in intrauterine insemination cycles in couples with subfertility: a prospective randomized study. *Eur J Obstet Gynecol Reprod Biol* 2015;**189**:33–37.
- Eskew AM, Bedrick BS, Hardi A, Stoll CRT, Colditz GA, Tuuli MG, Jungheim ES. Letrozole compared with clomiphene citrate for unexplained infertility: a systematic review and meta-analysis. *Obstet Gynecol* 2019;**133**:437–444.
- Fatemi HM, Kolibianakis E, Tournaye H, Camus M, Van Steirteghem AC, Devroey P. Clomiphene citrate versus letrozole for ovarian stimulation: a pilot study. *Reprod Biomed Online* 2003;**7**:543–546.
- Fisher DJ. Two-stage individual participant data meta-analysis and generalized forest plots. *Stata J* 2015;**15**:369–396.
- Fisher DJ, Carpenter JR, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;**356**:j573.
- Fouda UM, Sayed AM. Extended letrozole regimen versus clomiphene citrate for superovulation in patients with unexplained infertility undergoing intrauterine insemination: a randomized controlled trial. *Reprod Biol Endocrinol* 2011;**9**:84.
- Galal AF. Letrozole step-up protocol: the effect of a novel superovulation induction protocol to enhance pregnancy rate in a couple with unexplained infertility undergoing intrauterine insemination. *Fertil Steril* 2015;**104**:e328.
- Goldman MB, Thornton KL, Ryley D, Alper MM, Fung JL, Hornstein MD, Reindollar RH. A randomized clinical trial to determine optimal infertility treatment in older couples: the Forty and Over Treatment Trial (FORT-T). *Fertil Steril* 2014;**101**:1574–1581.e1-2.
- Gregoriou O, Vlahos NF, Konidaris S, Papadias K, Botsis D, Creatsas GK. Randomized controlled trial comparing superovulation with letrozole versus recombinant follicle-stimulating hormone combined with intrauterine insemination for couples with unexplained infertility who had failed clomiphene citrate stimulation and intrauterine insemination. *Fertil Steril* 2008;**90**:678–683.
- Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
- Ibrahim MI, Moustafa RA, Abdel-Azeem AA. Letrozole versus clomiphene citrate for superovulation in Egyptian women with unexplained infertility: a randomized controlled trial. *Arch Gynecol Obstet* 2012;**286**:1581–1587.
- Kamel A. Effect of induction protocols on pregnancy rate in artificial insemination by husband. 1995, Abstract of the 11th Annual Meeting of ESHRE.
- Mitwally MF, Casper RF. Use of an aromatase inhibitor for induction of ovulation in patients with an inadequate response to clomiphene citrate. *Fertil Steril* 2001;**75**:305–309.
- Naidu A, Thyagaraju C, Chathurvedula L. Comparing low dose human menopausal gonadotrophins with clomiphene citrate for ovarian stimulation with intrauterine insemination in unexplained infertility and mild male factor infertility: a randomized clinical trial. *J Obstet Gynaecol Res* 2020;**46**:61.
- Nakajima A, Smith L, Wong B, Scot J, Cumming D, Tataryn I, Sagle M, McAra D, Nordstrom L. A randomized trial of clomiphene citrate plus intrauterine insemination versus recombinant follicle stimulation hormone plus intratruerine insemination for the treatment of unexplained infertility. *Fertil Steril* 1999;**72**:S208.
- Ozmen B, El Sadek M, Matar H, Fouad S, El Nomrosy K, Abbas H. Aromatase inhibitor versus clomiphene citrate in IUI cycles. *Abstracts of 21st Annual Meeting of ESHRE*, 2005.
- Peeraer K, Debrock S, De Loecker P, Tomassetti C, Laenen A, Welkenhuysen M, Meeuwis L, Pelckmans S, Mol BW, Spiessens C et al. Low-dose human menopausal gonadotrophin versus clomiphene citrate in subfertile couples treated with intrauterine insemination: a randomized controlled trial. *Hum Reprod* 2015;**30**:1079–1088.
- Practice Committee of the American Society for Reproductive Medicine. Evidence-based treatments for couples with unexplained infertility: a guideline. *Fertil Steril* 2020;**113**:305–322.
- Pundir J, Achilli C, Bhide P, Sabatini L, Legro RS, Rombauts L, Teede H, Coomarasamy A, Zamora J, Thangaratinam S. Risk of foetal harm with letrozole use in fertility treatment: a systematic review and meta-analysis. *Hum Reprod Update* 2021;**27**:474–485.
- Riley RD, Debray TPA, Fisher D, Hattle M, Marlin N, Hoogland J, Gueyffier F, Staessen JA, Wang J, Moons KGM et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. *Stat Med* 2020;**39**:2115–2137.
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;**340**:c221.
- Sammour A, Biljan MM, Tan SL, Tulandi T. Prospective randomized trial comparing the effects of letrozole (LE) and clomiphene citrate (CC) on follicular development, endometrial thickness and pregnancy rate in patients undergoing super-ovulation prior to intrauterine insemination (IUI). *Fertil Steril* 2001;**76**:S110.
- Steures P, van der Steeg JW, Mol BW, Eijkemans MJ, van der Veen F, Habbema JD, Hompes PG, Bossuyt PM, Verhoeve HR, van Kasteren YM et al.; CECERM (Collaborative Effort in Clinical Evaluation in Reproductive Medicine). Prediction of an ongoing pregnancy after intrauterine insemination. *Fertil Steril* 2004;**82**:45–51.
- Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF; PRISMA-IPD Development Group. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015;**313**:1657–1665.
- van Eekelen R, Wang R, Danhof NA, Mol F, Mochtar M, Mol BW, van Wely M. Cost-effectiveness of ovarian stimulation agents for IUI in couples with unexplained subfertility. *Hum Reprod* 2021;**36**:1288–1295.

- van Rumste MM, Custers IM, van der Veen F, van Wely M, Evers JL, Mol BW. The influence of the number of follicles on pregnancy rates in intrauterine insemination with ovarian stimulation: a meta-analysis. *Hum Reprod Update* 2008;**14**:563–570.
- Wicherts JM, Bakker M, Molenaar D. Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One* 2011;**6**: e26828.
- Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, Rienzi L, Sunde A, Schmidt L, Cooke ID et al. The international glossary on infertility and fertility care, 2017. *Hum Reprod* 2017;**32**:1786–1801.
- Zolton JR, Lindner PG, Terry N, DeCherney AH, Hill MJ. Gonadotropins versus oral ovarian stimulation agents for unexplained infertility: a systematic review and meta-analysis. *Fertil Steril* 2020;**113**:417–425.e1.